

# SMPTE STANDARD

## Interoperable Master Format — Application #2E



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 — xyYCC

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**

<b>Characteristic</b>	<b>Constraint</b>							
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		C O L O R. 4	C O L O R. 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	Monoscopic	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'Bc'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<sub>B</sub>C<sub>R</sub> 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<sub>B</sub>C<sub>R</sub> sampling, the Y' component shall be sampled at each pixel, but the C<sub>R</sub> and C<sub>B</sub> 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**

<b>System</b>	<b>Description</b>
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 xYCC709 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'<sub>B</sub> and C'<sub>R</sub> are referred to as signals E'<sub>R</sub>, E'<sub>G</sub>, E'<sub>B</sub>, E'<sub>Y</sub>, E'<sub>CB</sub> and E'<sub>CR</sub> respectively, i.e. they correspond to gamma pre-corrected signals.



## 5.4 Quantization

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

**Table 5 – Quantization Systems**

<b>System</b>	<b>Component Triplet</b>	<b>Quantization equations (<i>n</i> is the pixel bit depth)</b>	<b>Notes</b>
QE.1	R'G'B'	$D'_R = \text{INT}((219 \cdot R' + 16) \cdot 2^{n-8})$	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.
		$D'_G = \text{INT}((219 \cdot G' + 16) \cdot 2^{n-8})$	
		$D'_B = \text{INT}((219 \cdot B' + 16) \cdot 2^{n-8})$	
	Y'C <sub>B</sub> C <sub>R</sub>	$D'_Y = \text{INT}((219 \cdot Y' + 16) \cdot 2^{n-8})$	
		$D'_{CR} = \text{INT}((224 \cdot C'_R + 128) \cdot 2^{n-8})$	
		$D'_{CB} = \text{INT}((224 \cdot C'_B + 128) \cdot 2^{n-8})$	
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**

<i>Component Index</i>	<i>R'G'B'</i>	<i>Y'C'B'C'R</i>
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  $Y'C'_B C'_R$  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**

<i>Generic Picture Essence Descriptor Item</i>	<i>Constraints</i>
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.

DisplayHeight	
DisplayWidth	
DisplayXOffset	See Annex H for illustrative examples.
DisplayYOffset	
ActiveHeight	Note: Unless explicitly set, the Active Area Rectangle is by default equal to the Display Rectangle — see SMPTE ST 2067-2.
ActiveWidth	
ActiveXOffset	
ActiveYOffset	
DisplayF2Offset	Shall be present.
AspectRatio	See Section 6.2.1.2.
Active Format Descriptor	Shall be ignored.
Video Line Map	Shall be ignored.
Alpha Transparency	Shall not be present.
Transfer Characteristic	Shall be present. See Section 6.2.2
Image Alignment Offset	Shall not be present.
Image Start Offset	Shall not be present.
Image End Offset	Shall not be present.
FieldDominance	Shall be present if interlaced structure is used and shall not be present if progressive structure is used.
Picture Essence Coding	Shall be present. See Section 6.2.5.
Coding Equations	Shall be present if Y'C <sub>B</sub> C <sub>R</sub> sampling is used. See Section 6.2.3. Shall be ignored if R'G'B' sampling is used.
Color Primaries	Shall be present. See Section 6.2.4.
Alternative Center Cuts	See Section 6.2.6.

### 6.2.1.1 Stored Width and Stored Height

The values of the Stored Width and Stored Height items shall be set according to the image frame structure, as specified in Table 9.

**Table 9 – Stored Width and Stored Height values**

<i>Frame Structure</i>	<i>Progressive</i>	<i>Interlaced</i>
Stored Width	Image Frame Width	Image Frame Width
Stored Height	Image Frame Height	Image Frame Height / 2

### 6.2.1.2 Aspect Ratio

The value of the Aspect Ratio item shall be the ratio of width to height of the rectangular area into which the pixels within the Display Rectangle are intended to be displayed.

Note: For interlaced image structure, the *rectangular area into which the pixels within the Display Rectangle are intended to be displayed* is the rectangular area resulting from the interlacing of the pixel elements within the Display Rectangle of each of the two fields of the image frame.

Example: Illustrative examples are presented in Annex H.

### 6.2.1.3 Frame Layout

The value of the Frame Layout item shall be equal to:

- 00h (FULL\_FRAME) if the image structure is progressive.
- 01h (SEPARATE\_FIELDS) if the image structure is interlaced.

### 6.2.2 Transfer Characteristic

The value of the Transfer Characteristic item shall be equal to:

- 06.0E.2B.34.04.01.01.01.04.01.01.01.01.02.00.00 if COLOR.1, COLOR.2 or COLOR.3 systems are used.
- 06.0E.2B.34.04.01.01.0D.04.01.01.01.01.01.08.00.00 if COLOR.4 system is used.
- 06.0E.2B.34.04.01.01.0E.04.01.01.01.01.01.09.00.00 (see Annex A) if COLOR.5 system is used.
- 06.0E.2B.34.04.01.01.0D.04.01.01.01.01.01.0A.00.00 (see Annex C) if COLOR.6 or COLOR.7 systems are used.

### 6.2.3 Coding Equations

The value of the Coding Equations item shall be equal to:

- 06.0E.2B.34.04.01.01.01.04.01.01.01.01.02.01.00.00 if COLOR.1 or COLOR.2 systems are used.

- 06.0E.2B.34.04.01.01.01.04.01.01.01.02.02.00.00 if COLOR.3 or COLOR.4 systems are used.
- 06.0E.2B.34.04.01.01.0D.04.01.01.01.02.06.00.00 (see Annex D) if COLOR.5 or COLOR.7 systems are used.

#### 6.2.4 Color Primaries

The value of the Color Primaries item shall be equal to:

- 06.0E.2B.34.04.01.01.06.04.01.01.01.03.02.00.00 if the COLOR.1 system is used.
- 06.0E.2B.34.04.01.01.06.04.01.01.01.03.01.00.00 if the COLOR.2 system is used.
- 06.0E.2B.34.04.01.01.06.04.01.01.01.03.03.00.00 if the COLOR.3 or COLOR.4 systems are used.
- 06.0E.2B.34.04.01.01.0D.04.01.01.01.03.04.00.00 if the COLOR.5 or COLOR.7 systems are used.
- 06.0E.2B.34.04.01.01.0D.04.01.01.01.03.06.00.00 if the COLOR.6 system is used.

#### 6.2.5 Picture Essence Coding

The value of the Picture Essence Coding item shall reflect the JPEG 2000 profile and operating level used to encode the image essence.

Note: See Section 5.5.2 for list of JPEG 2000 profiles and operating levels.

#### 6.2.6 Alternative Center Cuts

The Alternative Center Cuts item is specified in SMPTE ST 2067-2.

An empty Alternative Center Cuts item shall be ignored. Any Alternative Center Cut value not specified in SMPTE ST 2067-2 may be safely ignored.

#### 6.2.7 Mastering Display Color Volume Metadata

If COLOR.3, COLOR.5, COLOR.6 or COLOR.7 system is used then:

- either all or none of the items specified in Annex B shall be present;
- the items specified in Annex B should characterize the mastering display; and
- if the items specified in Annex B are absent, no ST 2086 metadata values are assumed.

If neither COLOR.5, COLOR.6, nor COLOR.7 system is used, then no item specified in Annex B shall be present.

Note: Annex B.6 contains selected examples values for Mastering Display Color Volume Metadata. Other values, not specified in these examples, are also permitted.

### 6.3 RGBA Picture Essence Descriptor

#### 6.3.1 General

The RGBA Picture Essence Descriptor items shall be constrained as specified in Table 10.

**Table 10 – RGBA Essence Descriptor items**

<i>RGBA Picture Essence Descriptor Item</i>	<i>Constraints</i>
Component Max Ref	Shall be present. See Section 6.3.2.
Component Min Ref	Shall be present. See Section 6.3.2.
Alpha Max Ref	Shall not be present.
Alpha Min Ref	Shall not be present.
ScanningDirection	Shall be present and shall be equal to 00h.
PixelLayout	Shall be ignored.
Palette	Shall not be present.
PaletteLayout	Shall not be present.

### 6.3.2 Component Max Ref and Component Min Ref

The values of the Component Max Ref and Component Min Ref items shall be set according to the pixel bit depth and quantization system used, as specified in Table 11 below.

**Table 11 – Component Max Ref and Component Min Ref values**

<i>System</i>	<i>QE.1</i>				<i>QE.2</i>			
Pixel Bit Depth	8	10	12	16	8	10	12	16
Component Min Ref	16	64	256	4096	0	0	0	0
Component Max Ref	235	940	3760	60160	255	1023	4095	65535

## 6.4 CDCI Picture Essence Descriptor

### 6.4.1 General

The CDCI Picture Essence Descriptor items shall be constrained as specified in Table 12.

**Table 12 – CDCI Essence Descriptor items**

<b><i>CDCI Picture Essence Descriptor Item</i></b>	<b><i>Constraints</i></b>
Component Depth	Shall be present and shall be equal to the Pixel Bit Depth used (see Section 5.2.2.7).
Horizontal Subsampling	See Section 6.4.2.
Vertical Subsampling	Shall be 01h.
Color Siting	Shall be present and shall be 00h.
ReversedByteOrder	Shall not be present.
PaddingBits	Shall not be present.
Alpha Sample Depth	Shall not be present.
Black Ref Level	Shall be present. See Section 6.4.3.
White Ref level	Shall be present. See Section 6.4.3.
Color Range	Shall be present. See Section 6.4.3.

#### 6.4.2 Horizontal Subsampling

The value of Horizontal Subsampling item shall be equal to:

- 01h if 4:4:4 sampling is used per Section 5.2.2.8.
- 02h if 4:2:2 Y'C<sub>B</sub>C<sub>R</sub> sampling is used per Section 5.2.2.8.

#### 6.4.3 Black Ref Level, White Ref Level and Color Range Values

If COLOR.1, COLOR.2, COLOR.3, COLOR.4, COLOR.5 or COLOR.7 is used, the values of the Black Ref Level, White Ref level and Color Range items shall be set according to Table 13.

**Table 13 – Black Ref Level, White Ref level and Color Range values for COLOR.1, COLOR.2, COLOR.3, COLOR.4, COLOR.5 and COLOR.7**

Colorimetry	COLOR.1 COLOR.2		COLOR.3		COLOR.4		COLOR.5 COLOR.7		
	8	10	12	16	8	10	10	12	16
Pixel Bit Depth	8	10	12	16	8	10	10	12	16
Black Ref Level	16	64	256	4096	16	64	64	256	4096
White Ref level	235	940	3760	60160	235	940	940	3760	60160
Color Range	225	897	3585	57345	254	1013	897	3585	57345



Note 1: The White Ref level item applies only to the Y' component, and the Color Range item to the C'<sub>B</sub> and C'<sub>R</sub> components.

Note 2: In the case of COLOR.7, "White Ref" is occasionally referred to as "Nominal Peak".

## 6.5 JPEG 2000 Picture Sub Descriptor

### 6.5.1 General

The Top-Level File Package of the Image Track File shall reference a JPEG 2000 Picture Sub Descriptor SMPTE ST 422 as constrained by Table 14.

**Table 14 – JPEG 2000 Picture Subdescriptor items**

<i>JPEG 2000 Picture Subdescriptor Item</i>	<i>Constraints</i>
Coding Style	Shall be present.
J2CLayout	Shall be present. See Section 6.5.2.

### 6.5.2 J2CLayout

The value of the J2CLayout item shall be equal to:

- { 'R', x, 'G', x, 'B', x, 0, 0, 0, 0, 0, 0, 0, 0, 0 } if R'G'B' sampling is used, where x is the pixel bit depth; or
- { 'Y', x, 'U', x, 'V', x, 0, 0, 0, 0, 0, 0, 0, 0, 0 } if Y'C'<sub>B</sub>C'<sub>R</sub> sampling is used where x is the pixel bit depth, respectively.

## 7 Composition

### 7.1 ApplicationIdentification

The ApplicationIdentification element (see SMPTE ST 2067-2) shall include the value listed in Table 15.

**Table 15 – Application Identification**

<a href="http://www.smpte-ra.org/ns/2067-21/2020">http://www.smpte-ra.org/ns/2067-21/2020</a>
---

### 7.2 Homogeneous Image Essence

Within a given composition, the following shall remain constant:

- all image essence characteristics specified in Section 5.
- the JPEG 2000 profile and operating level combination (see Section 5.5.2).

### 7.3 Virtual Tracks

#### 7.3.1 Main Image Virtual Track

All Image Track Files referenced by Resource elements of type StereolImageTrackFileResourceType and type TrackFileResourceType shall conform to Section 6.

### 7.4 Segment Duration

If the average number of audio samples per Composition Edit Unit is not an integer, the duration of each Segment shall be an integer multiple of  $5/\text{Composition Edit Rate}$ .

### 7.5 MaxCLL and MaxFALL

If MainImage conforms to COLOR.6 or COLOR.7, the ExtensionProperties element of the Composition Playlist instance shall include

- zero or one instance of the MaxCLL element specified in Table 16; and
- zero or one instance of the MaxFALL element specified in Table 16.

**Table 16 – MaxCLL and MaxFALL elements**

```
<xs:element name="MaxCLL" type="xs:unsignedShort"/>
<xs:element name="MaxFALL" type="xs:unsignedShort"/>
```

The MaxCLL and MaxFALL values may be calculated as specified in Annex P.1 and P.2 of CTA 861-G, respectively. This calculation shall use the MainImage Virtual Track image essence contained:

- within the area specified by the Active Area Rectangle; and
- only from the First Frame of Composition to the Last Frame of Composition, if specified.

If the MaxCLL (MaxFALL) value is unknown, then the MaxCLL (MaxFALL) element shall be (i) absent or (ii) set to 0.

The MaxCLL and MaxFALL values shall be represented in units of  $1 \text{ cd/m}^2$ .

## Annex A ITU-R BT.2020 Transfer Characteristic Label (Normative)

Table A.1 – ITU-R BT.2020 Transfer Characteristic Label

<b>Byte No.</b>	<b>Description</b>	<b>Value (hex)</b>	<b>Meaning</b>
1-7		(see Transfer Characteristic node)	
8	Version Number	0Eh	Registry Version at the point of registration of this label
9-13		(see Transfer Characteristic node)	
14	ITU-R BT.2020 Transfer Characteristic	09h	Identifies ITU-R BT.2020 transfer characteristic
15-16		00h	

## Annex B Mastering Display Color Volume Metadata (Normative)

### B.1 General

Table B.1 specifies optional items for the Generic Picture Essence Descriptor (see SMPTE ST 377-1) based on the metadata parameters specified in SMPTE ST 2086.

**Table B.1 – Color Volume Metadata**

<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>
Mastering Display Primaries	MasteringDisplayPrimaries	ThreeColorPrimaries	12	dyn	urn:smpte:ul:060e2b34.0101010e.04200401.01010000	Opt	Display Primaries metadata as specified in ST 2086	n/a
Mastering Display White Point Chromaticity	MasteringDisplayWhitePointChromaticity	ColorPrimary	4	dyn	urn:smpte:ul:060e2b34.0101010e.04200401.01020000	Opt	Chromaticity of White Point metadata as specified in ST 2086	n/a
Mastering Display Maximum Luminance	MasteringDisplayMaximumLuminance	UInt32	4	dyn	urn:smpte:ul:060e2b34.0101010e.04200401.01030000	Opt	Maximum Display Mastering Luminance metadata as specified in ST 2086	n/a
Mastering Display Minimum Luminance	MasteringDisplayMinimumLuminance	UInt32	4	dyn	urn:smpte:ul:060e2b34.0101010e.04200401.01040000	Opt	Minimum Display Mastering Luminance metadata as specified in ST 2086	n/a

Note: The quantization of the metadata parameters match that specified in Section D.2.28 of Recommendation ITU-T H.265 (11/2019).

### B.2 Mastering Display Primaries

If present, the Mastering Display Primaries item shall be equal to the Display Primaries metadata specified in SMPTE ST 2086.

The ColorPrimary type shall consist of two UInt16 elements, in order, the normalized x and y chromaticity coordinates of the color primary in units of 0.00002.

EXAMPLE: The color primary characterized by the (x, y) chromaticity coordinates of (0.6800, 0.3200) is represented by a ColorPrimary value of {34000, 16000}.

Note 1: The elements of a ColorPrimary instance that conforms to the precision specified in ST 2086 (four decimal places) are multiples of 5 ( $0.0001 = 0.00002 \cdot 5$ ).

The ThreeColorPrimaries type shall be a fixed-size sequence of 3 instances of the ColorPrimary type, for a total of 12 bytes. The ColorPrimary instances should be ordered as follows: (i) instance with the largest x chromaticity

coordinate, (ii) instance with the largest y chromaticity coordinate, and (iii) instance with neither the largest y nor the largest x chromaticity coordinate.

Note 2: The recommended ordering of the ColorPrimary instances corresponds to RGB ordering in many common cases.

### **B.3 Mastering Display White Point Chromaticity**

If present, the Mastering Display White Point Chromaticity item shall be equal to the Chromaticity of White Point metadata specified in SMPTE ST 2086.

The ColorPrimary type is specified in B.2.

### **B.4 Mastering Display Maximum Luminance**

If present, the Mastering Display Maximum Luminance item shall be equal to the Maximum Display Mastering Luminance metadata specified in SMPTE ST 2086.

The value Mastering Display Maximum Luminance item shall be expressed in units of  $0.0001 \text{ cd/m}^2$ .

### **B.5 Mastering Display Minimum Luminance**

If present, the Mastering Display Minimum Luminance item shall be equal to the Minimum Display Mastering Luminance metadata specified in SMPTE ST 2086.

The value Mastering Display Minimum Luminance item shall be expressed in units of  $0.0001 \text{ cd/m}^2$ .

### **B.6 Examples** (Informative)

Table B.2 lists example values for the items specified in Table B.1.

**Table B.2 – Example values for the Mastering Display Color Volume Metadata  
(Informative)**

<i>Item</i>	<i>Example Values</i>	
Mastering Display White Point Chromaticity	Illuminant D65 specified in SMPTE RP 177	{ 15635, 16450 }
	Color primaries specified in Recommendation ITU-R BT.709	{ 32000, 16500 } { 15000, 30000 } { 7500, 3000 }
Mastering Display Primaries	Color primaries specified in Recommendation ITU-R BT.2020	{ 35400, 14600 } { 8500, 39850 } { 6550, 2300 }
	Color primaries of the P3D65 system specified in SMPTE ST 2113	{ 34000, 16000 } { 13250, 34500 } { 7500, 3000 }
Mastering Display Maximum Luminance	Reference white in Recommendation ITU-R BT.2035	1000000
	Mastering display with a maximum luminance of 4000 cd/m <sup>2</sup>	40000000
Mastering Display Minimum Luminance	Reference black in Recommendation ITU-R BT.2035	100
	Mastering display with a minimum luminance of 0.005 cd/m <sup>2</sup>	50

## Annex C SMPTE ST 2084 Transfer Characteristic Label (Normative)

Table C.1 – SMPTE ST 2084 Transfer Characteristic Label

<b>Byte No.</b>	<b>Description</b>	<b>Value (hex)</b>	<b>Meaning</b>
1-7		(see Transfer Characteristic node)	
8	Version Number	0Dh	Registry Version at the point of registration of this label
9-13		(see Transfer Characteristic node)	
14	SMPTE ST 2084 Transfer Characteristic	0Ah	Identifies the transfer characteristics as specified in SMPTE ST 2084
15-16		00h	

## Annex D ITU-R BT.2020 Non-Constant Luminance Coding Equations Label (Normative)

Table D.1 – ITU-R BT.2020 Non-Constant Luminance Coding Equations Label

<b>Byte No.</b>	<b>Description</b>	<b>Value (hex)</b>	<b>Meaning</b>
1-7		(see Coding Equations node)	
8	Version Number	0Dh	Registry Version at the point of registration of this label
9-13		(see Coding Equations node)	
14	ITU-R BT.2020 Non-Constant Luminance Coding Equations	06h	Identifies ITU-R BT.2020 coding equations for non-constant luminance
15-16		00h	



## Annex E Consolidated Schema (Informative)

This specification is accompanied by the following element, which is an XML schema document as specified in the XML Schema Part 1: Structures.

st2067-21a-2020.xsd

This element collects the XML schema definitions defined in this specification. It is informative and, in case of conflict, this specification takes precedence.

## Annex F Recommended selection of JPEG 2000 Profiles (Informative)

This annex provides guidance for selecting the appropriate JPEG 2000 profiles based on the characteristics of the image essence.

If image essence conforms to the following parameters, a broadcast application profile (as defined in ISO/IEC 15444-1) might be used. Refer to Table F.1. If any of these are exceeded, an IMF profile is the appropriate choice. Refer to Table F.2 and Table F.3.

- 1) Broadcast application profiles support:
  - a) a maximum frame dimension of 3840 x 2160 pixels;
  - b) a maximum 12-bit color depth;
  - c) a maximum compressed bit rate of 800 Mbit/s for non-reversible profile;
  - d) a maximum components sampling rate of 520 MSamples/s.

Note that in this standard there are many combinations of image characteristics where a profile from either the broadcast application profile sets or IMF profile sets might be applicable.

To determine maximum components sampling rate:

$$M = (W \times H \times F \times A) / 1,000,000$$

Where:

M is MegaSamples per second (MSamples/s)

W is Frame Width, in pixels

H is Image Frame Height, in pixels

F is Frame Rate, per second

A is average number of color components per pixel (2 for 4:2:2, 3 for 4:4:4)

Example:  $(3840 \times 2160 \times 24000/1001 \times 2) / 1,000,000 = 397.7330688$  MSamples/s.

In this example, a user might select either BCP Level 5 or 4k IMF Single Tile Lossy Profile Mainlevel 5 for non-reversible workflows, or either BCP Level 7 or 4k IMF Single/Multi-tile Reversible Profile Mainlevel 5 (Sublevel 0) for reversible workflows.

**Table F.1: Broadcast Contribution Profile Levels used in IMF App2E**

<b>Levels</b>	<b>Max. Sampling Rate (MSamples/s)</b>	<b>Max. compressed Bit Rate (Mbits/s)</b>
Broadcast Single Tile Profile Level 1	65	200
Broadcast Single Tile Profile Level 2	130	200
Broadcast Single Tile Profile Level 3	195	200
Broadcast Single Tile Profile Level 4	260	400
Broadcast Single Tile Profile Level 5	520	800
Broadcast Multi-Tile Reversible Profile Level 7	520	Unspecified *

\*Reversible compression

**Table F.2: IMF Profile Mainlevels used by IMF App2E**

<b>Mainlevel</b>	<b>Max. Sampling Rate (MSamples/s)</b>	<b>Available Sublevels for lossy profiles</b>
Mainlevel 1	65	Sublevel 1
Mainlevel 2	130	Sublevel 1
Mainlevel 3	195	Sublevel 1
Mainlevel 4	260	Sublevels 1-2
Mainlevel 5	520	Sublevels 1-3
Mainlevel 6	1200	Sublevels 1-4
Mainlevel 7	2400	Sublevels 1-5
Mainlevel 8	4800	Sublevels 1-6

**Table F.3: IMF Profile Sublevels used by IMF App2E**

<b><i>Sublevels</i></b>	<b><i>Max. compressed Bit Rate (Mbit/s)</i></b>
Sublevel 1	200
Sublevel 2	400
Sublevel 3	800
Sublevel 4	1600
Sublevel 5	3200
Sublevel 6	6400
Sublevel 0	Unspecified *

\*only used for reversible profiles in IMF App2E

Users are encouraged to select the smallest level, for broadcast application profiles, or Mainlevel and Sublevel, for IMF profiles, that meets the requirements for image frame dimensions, color depth, components sampling rate and, for non-reversible compression, compressed bit rate. This allows implementations to optimally allocate resources when processing the image essence.

It is recommended that the sampling rate and maximum bit rate used in determining the profiles be increased by 5% to allow for speed-up of the image, e.g. 24fps to 25fps, without having to update profile metadata in both the codestream and Track File.

## Annex G Additional JPEG 2000 Picture Essence Compression Labels (Normative)

**Table G.1 – Additional JPEG 2000 Picture Essence Compression Labels**

<b>Byte No.</b>	<b>Description</b>	<b>Value (hex)</b>	<b>Meaning</b>
1-7		See SMPTE ST 422	
8	Version Number	0Dh	Registry Version at the point of registration of this label
9-15		See SMPTE ST 422	
16	Broadcast Contribution Single Tile Profile Level 1	11h	Broadcast Contribution Single Tile Profile Level 1 as specified in ISO/IEC 15444-1
	Broadcast Contribution Single Tile Profile Level 2	12h	Broadcast Contribution Single Tile Profile Level 2 as specified in ISO/IEC 15444-1
	Broadcast Contribution Single Tile Profile Level 3	13h	Broadcast Contribution Single Tile Profile Level 3 as specified in ISO/IEC 15444-1
	Broadcast Contribution Single Tile Profile Level 4	14h	Broadcast Contribution Single Tile Profile Level 4 as specified in ISO/IEC 15444-1
	Broadcast Contribution Single Tile Profile Level 5	15h	Broadcast Contribution Single Tile Profile Level 5 as specified in ISO/IEC 15444-1
	Broadcast Contribution Multi-tile Reversible Profile Level 6	16h	Broadcast Contribution Multi-tile Reversible Profile Level 6 as specified in ISO/IEC 15444-1
	Broadcast Contribution Multi-tile Reversible Profile Level 7	17h	Broadcast Contribution Multi-tile Reversible Profile Level 7 as specified in ISO/IEC 15444-1

## Annex H Image Frame And Active Area Rectangle Examples (Informative)

This specification allows the user to choose which rectangular subset of an input image to store and which portion of the resulting stored image to identify as the Active Area Rectangle.

Figure H.1 and Table H.1 depict selected examples using HD progressive image frames. As shown in (a) and (b), two different users or the same user in different circumstances can elect to identify different portions of the same image as the Active Area Rectangle. In (c), only the Active Area Rectangle is stored. As illustrated by (b), the extent of the Display Rectangle is not specified and left to users: Table H.1 lists two valid sets of Display Rectangle and Active Area Rectangle values for (b).

The process by which the dimensions of the stored image and active area are set depends on individual workflows and can include a combination of manual and automated processing during and after ingest.

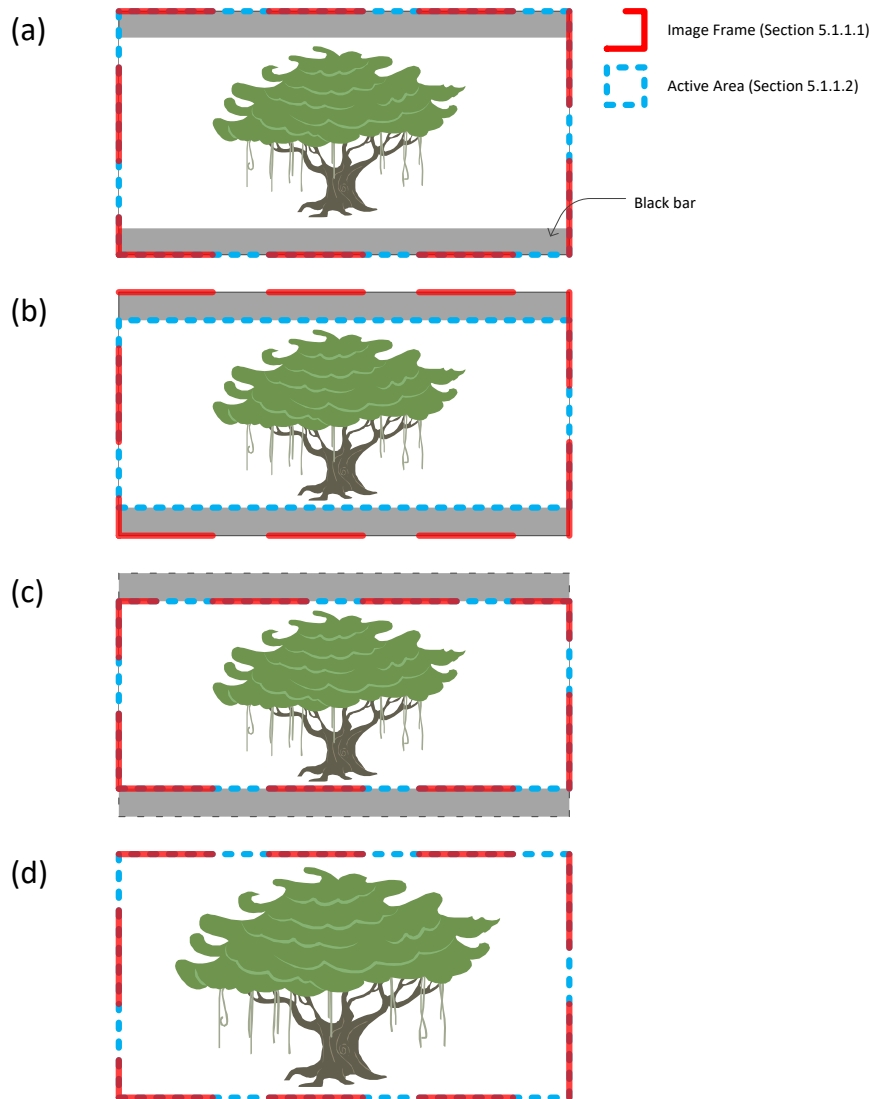


Figure H.1 – High-Definition Progressive Frame Active Area Rectangle Examples

**Table H.1 – Selected Property Values for the Examples of Figure H.1**

	Figure H.1a	Figure H.1b	Figure H.1b	Figure H.1c	Figure H.1d
Stored Height	1080	1080	1080	800	1080
Stored Width	1920	1920	1920	1920	1920
Sampled Height	1080	1080	1080	800	1080
Sampled Width	1920	1920	1920	1920	1920
Sampled X Offset	0	0	0	0	0
Sampled Y Offset	0	0	0	0	0
Display Height	1080	800	1080	800	1080
Display Width	1920	1920	1920	1920	1920
Display X Offset	0	0	0	0	0
Display Y Offset	0	140	0	0	0
Aspect Ratio	16/9	12/5	16/9	12/5	16/9
Active Height	1080	800	800	800	1080
Active Width	1920	1920	1920	1920	1920
Active X Offset	0	0	0	0	0
Active Y Offset	0	0	140	0	0

Figure H.2 and Table H.2 depict three means of storing the same source SD interlaced anamorphic image frame. The source image frame contains a 720x576 container as well as vertical blanking interval (VBI) information. The container is intended to be displayed at a 16:9 aspect ratio and contains a 2.40 Active Area Rectangle. In (a), the entire source image frame is ingested, including VBI information, which is not included in the Display Rectangle, as specified by SMPTE ST 377-1. In contrast to (a), the VBI information is not stored in (b). In (c), the Aspect Ratio property value is computed to preserve the exact pixel aspect ratio of the source image.

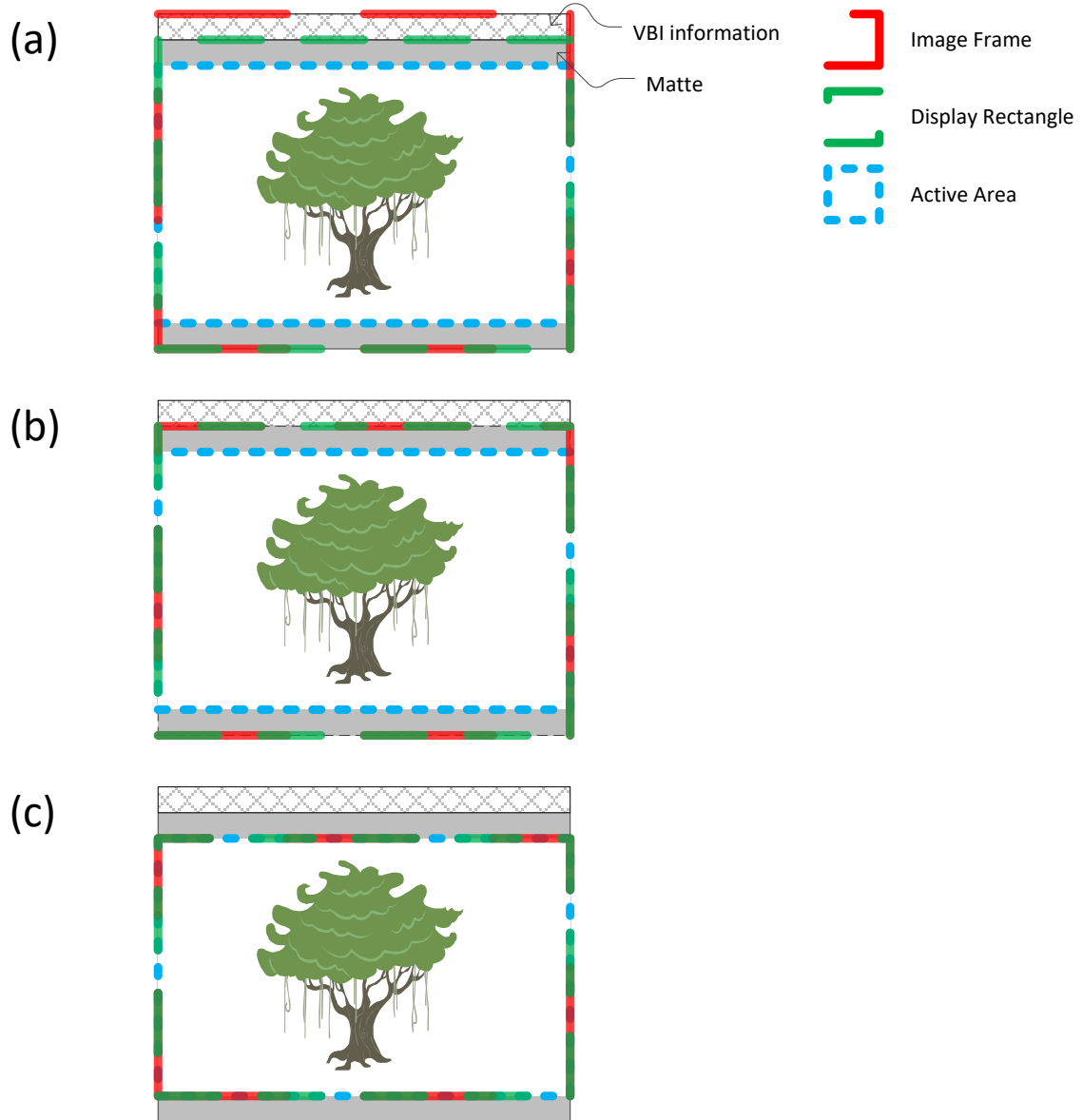


Figure H.2 – Standard Definition Interlaced Frame Examples



**Table H.2 – Selected Property Values for the Examples of Figure H.2**

	Figure H.2a	Figure H.2b	Figure H.2c
Stored Height	304	288	213
Stored Width	720	720	720
Sampled Height	304	288	213
Sampled Width	720	720	720
Sampled X Offset	0	0	0
Sampled Y Offset	0	0	0
Display Height	288	288	213
Display Width	720	720	720
Display X Offset	0	0	0
Display Y Offset	16	0	0
Aspect Ratio	16/9	16/9	512/213
Active Height	213	213	213
Active Width	720	720	720
Active X Offset	0	0	0
Active Y Offset	37	37	0

## **Bibliography (Informative)**

SMPTE RP 177:2003, Derivation of Basic Television Color Equations

Recommendation ITU-R BT.2035, A Reference Viewing Environment for Evaluation of HDTV Program Material or Completed Programmes

Recommendation ITU-T H.265 (11/2019), High Efficiency Video Coding