SMPTE ST 2036-1:2009
Revision of SMPTE 2036-1-2007

SMPTE STANDARD

Ultra High Definition Television — Image Parameter Values for Program Production

Table of Contents

Approved
December 18, 2009

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE ST 2036-1 was prepared by Technology Committee 10E.

Intellectual Property

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

1 Scope

1.1 This standard defines a family of progressive image sample structures for the representation of stationary or moving two-dimensional images sampled temporally at a constant frame rate and having an image format (sample structure) of 3840 x 2160 or 7680 x 4320 which has a hierarchical relationship with 1920 x 1080 and an aspect ratio of 16:9, called ultra high definition television (UHDTV), as defined in Table 1. This standard specifies:

- R'G'B' color encoding and digital representation
- $Y'C'_BC'_R$ color encoding and digital representation.

1.2 This standard specifies multiple samples, line and frame rate formats for ten-bit and twelve-bit systems. It is not necessary for an implementation to support all formats to be compliant with this standard. However, an implementation must state which of the Systems are supported.

Table 1 – Image sample structures and frame rates of UHDTV systems

System category	System nomenclature	Luma or R' G' B' samples per line	Lines per frame	Frame rate (Hz)
UHDTV1	3840 x 2160/23.98/P	3840	2160	24/1.001
	3840 x 2160/24/P	3840	2160	24
	3840 x 2160/25/P	3840	2160	25
	3840 x 2160/29.97/P	3840	2160	30/1.001
	3840 x 2160/30/P	3840	2160	30
	3840 x 2160/50/P	3840	2160	50
	3840 x 2160/59.94/P	3840	2160	60/1.001
	3840 x 2160/60/P	3840	2160	60
UHDTV2	7680 x 4320/23.98/P	7680	4320	24/1.001
	7680 x 4320/24/P	7680	4320	24
	7680 x 4320/25/P	7680	4320	25
	7680 x 4320/29.97/P	7680	4320	30/1.001
	7680 x 4320/30/P	7680	4320	30
	7680 x 4320/50/P	7680	4320	50
	7680 x 4320/59.94/P	7680	4320	60/1.001
	7680 x 4320/60/P	7680	4320	60

NOTE – Throughout this standard, references to signals represented by a single primed letter (e.g., R', G' and B') refer to signals to which the transfer characteristics in § 6.3 have been applied. Such signals are commonly described as being gamma corrected.

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 an Engineering Document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

3 Normative References

The following standard contains provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the edition indicated was 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 standard indicated below.

CIE S 014-2/E:2006. *Colorimetry Part 2: CIE standard illuminants. [ISO 11664-2:2007]*

SMPTE RP 177-1993, Derivation of Basic Television Color Equations

4 General

4.1 The specification of a system claiming compliance with this standard shall state:

- which of the systems of Table 1 are implemented;
- which of the signal representations are implemented (R’G’B’ or Y’C’$_B$C’$_R$);
- whether the digital representation employs uniformly quantized (linear), PCM, ten bits or twelve bits.

4.2 A twelve-bit codeword, when converted to a smaller number of sampling bits, should either be rounded or truncated. A ten-bit codeword, when converted to a twelve-bit codeword, shall have two padding bits added.

5 Image Structure

5.1 The sampling lattice shall be orthogonal. The pixel aspect ratio is 1:1 (square pixels). The aspect ratio of the image shall be 16:9.

5.2 UHDTV1 and UHDTV2 systems shall convey 2160 and 4320 picture lines per frame respectively at the interface.

5.3 Each line shall be represented by a number of samples, uniformly spaced.

5.4 The data numbering in a line shall be from left to right and the lines shall be numbered from top to bottom. The data numbering shall begin with 0. The top left sampling data shall be expressed with (0, 0) and the bottom right sampling data shall be (3839, 2159) for UHDTV1 and (7679, 4319) for UHDTV2.

6 System Colorimetry

6.1 To be compliant with this standard equipment shall be designed in accordance with the colorimetric analysis and opto-electronic transfer function defined in this section.

6.2 Picture information shall be linearly represented by red, green and blue tristimulus values (RGB), lying in the range 0 (reference black) to 1 (reference white), whose colorimetric attributes are based upon reference primaries with the following chromaticity coordinates (Table 2), whose reference white conforms to CIE D65 as defined by CIE S 014-2/E:2006.

Table 2 – Chromaticity coordinates of reference primaries and reference white

	CIE x	CIE y
Red primary	0.640	0.330
Green primary	0.300	0.600
Blue primary	0.150	0.060
Reference white	0.3127	0.3290

6.3 From the red, green and blue tristimulus values, three nonlinear primary components R', G' and B' shall be computed according to the opto-electronic transfer function as follows, where L denotes a tristimulus value and *V'* denotes a nonlinear primary signal:

$$V' = \begin{cases} 4.5L, & 0 \le L < \beta \\ \alpha L^{0.45} - (\alpha - 1), & \beta \le L \le 1 \end{cases}$$

For a 10-bit system $\alpha = 1.099$ and $\beta = 0.018$, and for a 12-bit system $\alpha = 1.0993$ and $\beta = 0.0181$ shall be used. (See Annex B.)

6.4 Signal levels shall be completely contained in the range specified between reference black and reference white specified in § 7.5, except for overshoots and undershoots due to processing.

6.5 The Y' component shall be computed as a weighted sum of nonlinear R'G'B' primary components, using coefficients calculated from the reference primaries according to the method given in SMPTE RP 177:

$$Y' = 0.2126R' + 0.7152G' + 0.0722B'$$

6.6 Color-difference component signals P'_B and P'_R having the same excursion as the Y' component shall be computed as follows:

$$P'_B = \frac{0.5}{1. - 0.0722}(B' - Y')$$

$$P'_R = \frac{0.5}{1. - 0.2126}(R' - Y')$$

P'_B and P'_R may be coded as C'_B and C'_R components for digital transmission.

7 Digital Picture Representation

7.1 Digital representation shall employ either R'G'B' or Y'C'$_B$C'$_R$ components as defined in § 6, uniformly sampled.

7.2 The digital signals described here are assumed to have been appropriately filtered to reduce or prevent aliasing upon sampling.

7.3 R'G'B' signals and Y' signals shall have the same nominal bandwidth. C'$_B$C'$_R$ shall have the same bandwidth as the associated Y' signal in the case of a 4:4:4 system, bandwidth nominally half that of the associated Y' signal in the case of a 4:2:2 system, and bandwidth nominally quarter that of the associated Y' signal in the case of a 4:2:0 system.

7.4 R'G'B' signals and the Y' signal of the Y'C'$_B$C'$_R$ interface shall be sampled orthogonally, line- and picture-repetitive. R'G'B' samples shall be cosited with each other.

7.5 Digital R'G'B' and Y' components shall be computed as follows:

$$L'_D = INT\left[219DL' + 16D\right]; \quad D = 2^{n-8}$$

where L' is the component value in abstract terms from 0 to 1.0, n takes the value 10 or 12 corresponding to the number of bits to be represented, and L_D' is the resulting digital code. The function INT[x] gives the value of 0 for fractional parts in the range of 0 to under 0.5, and +1 for fractional parts in the range from 0.5 to under 1.

NOTE – This scaling places the code words 040_h ($64_{(10)}$) and $3AC_h$ ($940_{(10)}$) in a 10-bit representation or the code words 100_h ($256_{(12)}$) and $EB0_h$ ($3760_{(12)}$) in a 12-bit representation for the signal level of 0% and 100% of R'G'B' and Y' as shown in Table 3.

NOTE – Throughout this section, subscript of "h" indicates the hexadecimal notation and subscript in parentheses with the integer decimal notation indicates the bit depth for video signal levels.

7.6 Digital C'$_B$ and C'$_R$ components of the Y'C'$_B$C'$_R$ set shall be computed as follows:

$$C'_D = INT\left[224DC' + 128D\right]; \quad D = 2^{n-8}$$

where C' is the component value in abstract terms from -0.5 to +0.5 and C'$_D$ is the resulting digital code.

NOTE – This scaling places the code words 040_h ($64_{(10)}$) , 200_h ($512_{(10)}$) and $3C0_h$ ($960_{(10)}$) in a 10-bit representation or the code words 100_h ($256_{(12)}$), 800_h ($2048_{(12)}$) and $F00_h$ ($3840_{(12)}$) in a 12-bit representation for the signal level of -50%, 0% and +50% of C'$_B$ and C'$_R$ as shown in Table 3.

7.7 C'$_B$ and C'$_R$ signals shall each have the same number of horizontal samples as the Y' component in the case of a 4:4:4 system.

C'$_B$ and C'$_R$ signals shall be horizontally subsampled by a factor of two with respect to the Y' component in the case of a 4:2:2 system. C'$_B$ and C'$_R$ samples shall be co-sited with even-numbered samples of Y'.

C'$_B$ and C'$_R$ shall be horizontally and vertically subsampled by a factor of two with respect to the Y' component in the case of a 4:2:0 system. C'$_B$ and C'$_R$ samples shall be co-sited with even-numbered samples and lines of Y'.

7.8 Code values having the eight most-significant bits all zero or all one, that is 10-bit codes 000_h ($0_{(10)}$) through 003_h ($3_{(10)}$) and $3FC_h$ ($1020_{(10)}$) through $3FF_h$ ($1023_{(10)}$) in the case of a ten-bit system, and 12-bit codes 000_h ($0_{(12)}$) through $00F_h$ ($15_{(12)}$) and $FF0_h$ ($4080_{(12)}$) through FFF_h ($4095_{(12)}$) in the case of a twelve-bit system as shown in Table 3, are employed for synchronizing purposes and shall be prohibited from video or ancillary data/signals.

7.9 A system having a ten-bit interface shall address the conversion of twelve-bit video data to ten-bits with an appropriate process that minimizes video artifacts such as quantization noise. When converting ten-bit data to twelve-bit data the two least significant bits of the twelve-bit word shall be set to 0.

7.10 For Y', R', G' and B' signals, undershoot and overshoot in video process may be accommodated by the use of code words 004_h ($4_{(10)}$) through $03F_h$ ($63_{(10)}$) and code words $3AD_h$ ($941_{(10)}$) through $3FB_h$ ($1019_{(10)}$) in a ten-bit system, code words 010_h ($16_{(12)}$) through $0FF_h$ ($255_{(12)}$) and code words $EB1_h$ ($3761_{(12)}$) through FEF_h ($4079_{(12)}$) in a twelve-bit system as shown in Table 3.

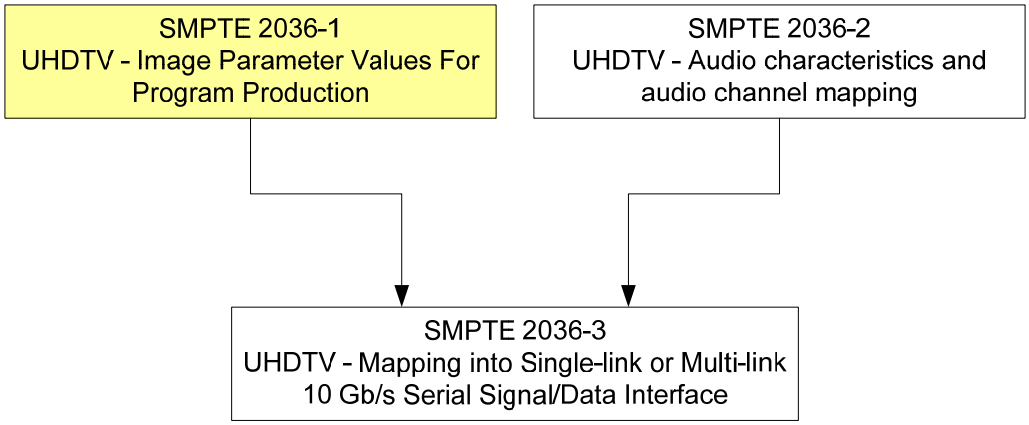
For C'_B and C'_R signals, undershoot and overshoot in video process may be accommodated by the use of code words 004_h ($4_{(10)}$) through $03F_h$ ($63_{(10)}$) and code words $3C1_h$ ($961_{(10)}$) through $3FB_h$ ($1019_{(10)}$) in ten-bit system, code words 010_h ($16_{(12)}$) through $0FF_h$ ($255_{(12)}$) and code words $F01_h$ ($3841_{(12)}$) through FEF_h ($4079_{(12)}$) in twelve-bit system as shown in Table 3.

Table 3 – Digital representation

Items		10-bit system	12-bit system	Note
R', G', B', Y'	Signal level of 100%	$3AC_h$ ($940_{(10)}$)	$EB0_h$ ($3760_{(12)}$)	Section 7.5
	Signal level of 0%	040_h ($64_{(10)}$)	100_h ($256_{(12)}$)	
	Headroom for Overshoot	$3AD_h$-$3FB_h$ ($941_{(10)}$-$1019_{(10)}$)	$EB1_h$-FEF_h ($3761_{(12)}$-$4079_{(12)}$)	Section 7.10
	Footroom for Undershoot	004_h-$03F_h$ ($4_{(10)}$-$63_{(10)}$)	010_h-FF_h ($16_{(12)}$-$255_{(12)}$)	
C'_B, C'_R	Signal level of 50%	$3C0_h$ ($960_{(10)}$)	$F00_h$ ($3840_{(12)}$)	Section 7.6
	Signal level of 0%	200_h ($512_{(10)}$)	800_h ($2048_{(12)}$)	
	Signal level of – 50 %	040_h ($64_{(10)}$)	100_h ($256_{(12)}$)	
	Headroom for Overshoot	$3C1_h$-$3FB_h$ ($961_{(10)}$-$1019_{(10)}$)	$F01_h$-FEF_h ($3841_{(12)}$-$4079_{(12)}$)	Section 7.10
	Footroom for Undershoot	004_h-$03F_h$ ($4_{(10)}$-$63_{(10)}$)	010_h-$0FF_h$ ($16_{(12)}$-$255_{(12)}$)	
Prohibited codes	Upper range	$3FC_h$-$3FF_h$ ($1020_{(10)}$-$1023_{(10)}$)	$FF0_h$-FFF_h ($4080_{(12)}$-$4095_{(12)}$)	Section 7.8
	Lower range	000_h-003_h ($0_{(10)}$-$3_{(10)}$)	000_h-$00F_h$ ($0_{(12)}$-$15_{(12)}$)	

Annex A (Informative)
Roadmap of UHDTV Standards

This figure shows the roadmap of the UHDTV standards for the program production and program exchange. This document covers UHDTV image parameters, including image format, colorimetry, digital representation etc.



Annex B (Informative)
Opto-Electronic Transfer Characteristics for Required Bit Accuracy

The opto-electronic transfer function is defined in § 6.3 as follows:

$$V' = \begin{cases} 4.5L, & 0 \le L < \beta \\ \alpha L^{0.45} - (\alpha - 1), & \beta \le L \le 1 \end{cases} \quad (1)$$

where L denotes a tristimulus value and *V'* denotes a nonlinear primary signal.

α and β are the solution to the following simultaneous equations:

$$\begin{cases} 4.5\beta = \alpha\beta^{0.45} - \alpha + 1 \\ 4.5 = 0.45\alpha\beta^{-0.55} \end{cases} \quad (2)$$

The simultaneous equations provide the required condition to have the two equations in (1) connected smoothly, and yield α=1.09929682680944… and β=0.018053968510807. Table B.1 shows the effective digits of α and β to get a sufficient accuracy in accordance with bit depth of the system.

Table B.1 – Effective digits of α and β

Bit depth	α	β
10-bit	1.099	0.018
11-bit	1.099	0.018
12-bit	1.0993	0.0181
18-bit	1.0993	0.0181
22-bit	1.099297	0.018054
25-bit	1.0992968	0.0180540

Annex C (Informative)
Bibliography

SMPTE 274M-2008, Television — 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

ITU-R BT.1201-1, Extremely High Resolution Imagery

ITU-R BT.709-5, Parameter Values for the HDTV Standards for Production and International Programme Exchange

ITU-R BT.1361, Worldwide Unified Colorimetry and Related Characteristics of Future Television and Imaging Systems

ITU-R BT.1769, Parameter Values for an Expanded Hierarchy of LSDI Image Formats for Production and International Programme Exchange

ITU-R BT.1846 (10/08), Notation for video systems