ETSI GS CCM 001 V1.1.1 (2017-02)

Compound Content Management Specification

Disclaimer

The present document has been produced and approved by intelligent Compound Content Management (CCM) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference

DGS/CCM-001

Keywords

CCM system, HDR, HDR composer, HDR Metadata Embedder, UHDTV, WCG

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Group Specification (GS) has been produced by ETSI Industry Specification Group (ISG) on intelligent Compound Content Management (CCM).

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

The present document defines and specifies the additional functionality required in consumer devices to enable the accurate recreation of both today's television signal, i.e. SDR signal, and the HDR/WCG signal from content created using the relevant production techniques standardized in SMPTE.

The requirement not to compromise the quality of the legacy signal using today's displays but yet to provide the highest quality HDR/WCG signal for a new generation of televisions is onerous and requires analysis to derive an optimal solution.

However, given the requirement to achieve backwards compatibility with existing receivers, some additional technology will be required in the next generation of HDR/WCG receivers to accurately recreate the HDR/WCG signal.

The full HDR system as shown in the figure 1 describes an end-to-end technology suite that enables the creation and distribution of content mastered with a high dynamic range and wide colour gamut. Display management may be used to more accurately match the capability of a given television to the HDR/WCG signal source. This creates seamless video experience for the viewer. The display management process may be guided by metadata generated in the HDR encoder.

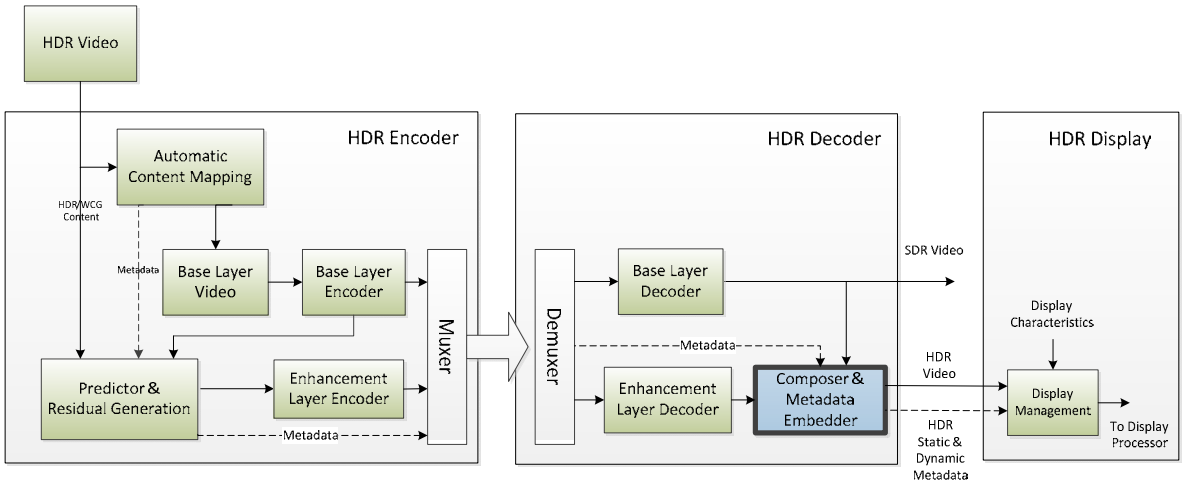

Figure 1: Overview of a generic HDR system and the additional HDR Compound Content Management functionality

1 Scope

The present document specifies the additional functionality required in Consumer devices for the reconstruction of an HDR/WCG signal using the "intelligent Compound Content Management" processing defined. Additionally it defines a method whereby this HDR/WCG signal and its associated metadata is transported over existing commonly used consumer baseband interfaces.

This process is one element of an end-to-end HDR video system. The definition of the overall end-to-end HDR video system is out of scope of the present document.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] SMPTE ST 2084:2014: "Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays".

[2] SMPTE ST 2086:2014: "Mastering Display Color Volume Metadata Supporting High Luminance And Wide Color Gamut Images".

[3] SMPTE ST 2094-1:2016: "Mapping VC-3 Coding Units into the MXF Generic Container".

[4] SMPTE ST 2094-10:2016: "Dynamic Metadata for Color Volume Transform - Application #1".

[5] Recommendation ITU-R BT.1886-0: "Reference electro-optical transfer function for flat panel displays used in HDTV studio production".

[6] Recommendation ITU-R BT.2100-0: "Image parameter values for high dynamic range television for use in production and international programme exchange".

[7] Recommendation ITU-R BT.709-6: "Parameter values for the HDTV standards for production and international programme exchange".

[8] Recommendation ITU-R BT.2020-2: "Parameter values for ultra-high definition television systems for production and international programme exchange".

[9] Recommendation ITU-T H.265 | ISO/IEC 23008-2: "Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 2: High efficiency video coding".

[10] ISO/IEC 13818-1: "Information technology - Generic coding of moving pictures and associated audio information - Part 1: Systems".

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] SMPTE 319M-2000: "For Television - Transporting MPEG-2 Recoding Information through 4:2:2 Component Digital Interfaces".

[i.2] Recommendation ITU-R BT.2246-2: "The present state of ultra-high definition television".

3 Definitions and abbreviations

3.1 Definitions

For the purpose of the present document, the following terms and definitions apply:

base layer: carries the video signal of the compound content which is used for the prediction of the HDR/WCG signal

block: MxN (M-column by N-row) array of samples

Compound Content Management (CCM): process deriving a picture with a high dynamic range

NOTE: The high dynamic range picture is derived from a BL, metadata and potentially an EL. This method can be part of the encoding / decoding process or not.

composer: component deriving a picture with a high dynamic range

NOTE: The composer is controlled by metadata and combines two baseband pictures derived from decodable compressed streams.

display management: process adapting the signal to the dynamic range and the colour gamut of the target display

dynamic range: ratio of the maximum light intensity to the minimum light intensity [i.2]

NOTE: In digital cameras the dynamic range is normally measured in terms of stops, which describe the total light range by power of 2.

enhancement layer: carries the additional information required to accurately recreate the HDR/WCG signal

high dynamic range: typically a dynamic range of more than 10 stops is referred to as high dynamic range

inverse mapping: colour mapping of the BL signal to form a predictor for the EL signal

inverse quantization: non-linearly scales the EL signal from a code word with a bit depth equal to EL_bit_depth to a code word with a higher bit depth

residual: EL signal which is the difference of the inverse mapped BL signal and the high dynamic range output signal

standard dynamic range: typically a dynamic range of up to 10 stops is referred to as standard dynamic range

3.2 Symbols

3.2.1 Arithmetic operators

For the purposes of the present document, the following arithmetic operators apply:

+	Addition
-	Subtraction (as a two-argument operator) or negation (as a unary prefix operator)
*	Multiplication, including matrix multiplication
÷	Used to denote division in mathematical equations where no truncation or rounding is intended.
/	Integer division with truncation of the result toward zero. For example, 7/4 and - 7/- 4 are truncated to 1 and - 7/4 and 7/- 4 are truncated to - 1.

3.2.2 Bit-wise operators

For the purposes of the present document, the following bit-wise operators apply:

x >> y	Arithmetic right shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values y. Bits shifted into the most significant bits (MSBs) as a result of the right shift have a value equal to the MSB of x prior to the shift operation.
x << y	Arithmetic left shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values y. Bits shifted into the least significant bits as a result of the left shift have a value equal to 0.

3.2.3 Relational operators

For the purposes of the present document, the following relational operators apply:

>	Greater than
>=	Greater than or equal to
<	Less than
<=	Less than or equal to
==	Equal to
!=	Not equal to

3.2.4 Assignment operators

For the purposes of the present document, the following assignment operators apply:

=	Assignment operator
++	Increment, i.e. *x* is equivalent to *x* = *x* + 1; when used in an array index, evaluates to the value of the variable prior to the increment operation.
--	Decrement, i.e. *x*- - is equivalent to *x* = *x* - 1; when used in an array index, evaluates to the value of the variable prior to the decrement operation.
+=	Increment by amount specified, i.e. x += 4 is equivalent to x = x + 4, and x += (-4) is equivalent to x = x + (-4).
-=	Decrement by amount specified, i.e. x -= 4 is equivalent to x = x - 4, and x -= (-4) is equivalent to x = x - (-4).

3.2.5 Mathematical functions

For the purposes of the present document, the following mathematical functions apply:

$$\text{Abs}(x) = \begin{cases} x & ; \quad x \geq 0 \\ -x & ; \quad x < 0 \end{cases}$$

$$\text{Clip3}(x, y, z) = \begin{cases} x & ; \quad z < x \\ y & ; \quad z > y \\ z & ; \quad otherwise \end{cases}$$

Floor(x) the largest integer less than or equal to x.

$$\text{Max}(x,y) = \begin{cases} x & ; \quad x \geq y \\ y & ; \quad x < y \end{cases}$$

Round(x)=Sign(x)*Floor(Abs(x)+0,5)

$$\text{Sign}(x) = \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x = 0 \\ -1 & ; \quad x < 0 \end{cases}$$

$$\text{Short}(x) = \begin{cases} x & ; \quad x < 2^{15} \\ x - 2^{16} & ; \quad x \geq 2^{15} \end{cases}$$, where x is a 16 bit integer value.

3.2.6 Order of operation precedence

When order of precedence in an expression is not indicated explicitly by use of parentheses, operations are evaluated sequentially from left to right.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

BL	Base Layer
CCM	Compound Content Management
CM	Composing Metadata
CRC	Cyclic Redundancy Check
DM	Display Management
DPB	Decoded Picture Buffer
EL	Enhancement Layer
EOS	End Of stream
EOTF	Electro-Optical Transfer Function
HDR	High Dynamic Range
ITU-R	International Telecommunications Union - Radiocommunications standardization sector
ITU-T	International Telecommunications Union - Telecommunications standardization sector
LSB	Least Significant Bit
MMR	Multivariate Multiple Regression
NLQ	Non-Linear Quantization
PPS	Picture Parameter Set
PQ	Perceptual Quantizer

NOTE: As defined in SMPTE ST 2084 [1].

SDR	Standard Dynamic Range
SHVC	Scalable High Efficiency Video Codec
SMPTE	Society of Motion Pictures and Television Engineers
ST	STandard
UHD	Ultra High Definition
UHDTV	Ultra High Definition Television
WCG	Wide Colour Gamut

4 CCM system model

The CCM system consists of two main components, the Composer and the HDR Display Management Metadata Embedder - see figure 2.

The composer reconstructs the HDR signal from the base layer (BL) image, the associated enhancement layer (EL) image and the related metadata information. First the EL signal is inverse quantized by applying inverse quantization coefficients delivered in the composing metadata (CM). Based on the BL signal and the related mapping coefficients of the CM, the predicted HDR signal is calculated. Finally the inverse quantized residual is added to the predicted HDR signal in order to recreate the HDR signal. If the characteristics of the BL input signal are different from the ones defined in SMPTE ST 2084 [1], a conversion of the recreated HDR signal according to SMPTE ST 2084 is performed to achieve the HDR output signal. The composer is defined in clause 5.

NOTE 1: Implementers can use different resolutions for the EL or the BL input pictures, thus requiring resampling of the input data before HDR composition in order to adjust either image resolution accordingly. Resampling is not a normative part of the present document and is assumed to be performed prior to the HDR reconstruction step defined in the present document. A reference implementation for the spatial resampling is provided in annex B.

Additional HDR, named HDR display management (DM) metadata, may be present at the input of the CCM system. The HDR DM metadata combines SMPTE ST 2086 static metadata [2] and SMPTE ST 2094-1/10 dynamic metadata [3] and [4], and is used to maintain the artistic intent when the content is mapped to the display capabilities. This HDR DM metadata should be passed-through to the output of the composer or the metadata embedder inserts the HDR DM metadata into the HDR picture for transport over baseband interfaces. The HDR DM metadata embedder is defined in clause 6.

NOTE 2: The composer metadata and functionality is independent of the type of SMPTE ST 2094 metadata carried.

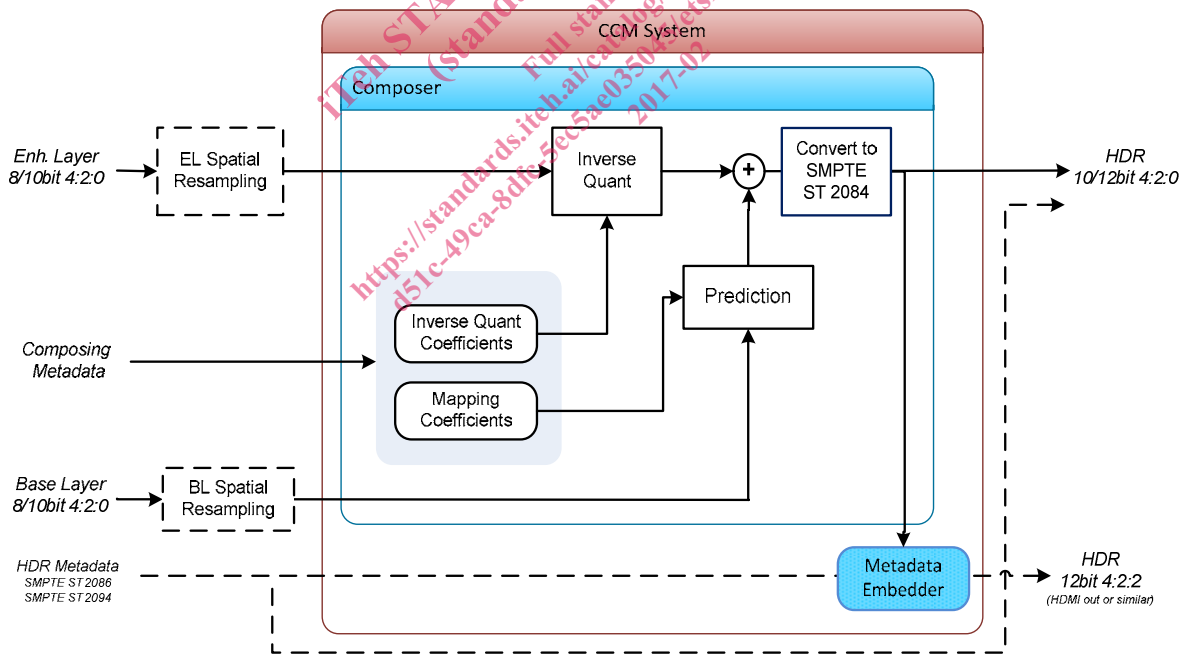

Figure 2: HDR Compound Content system high-level block diagram

5 Composer definition

5.1 Introduction

This clause defines the guidelines for combining the BL with the EL information and the composer metadata to recreate the HDR picture.

Clause 5.2 defines the input and output picture formats of the composer.

Clause 5.3 defines the CM, which is needed for the recreation of the HDR output picture.

Clause 5.4 defines the decoding process, which is performed to recreate the HDR output picture. Clause 5.4.2 applies to the BL mapping process. Clause 5.4.3 defines the EL processing and the reconstruction of the HDR signal.

Clause 5.5 defines the conversion of the recreated HDR picture to the transfer characteristic as defined in SMPTE ST 2084 [1].

5.2 Input and output picture format

5.2.1 General

The composer defines two input pictures, named BL and EL picture, and one output picture, named HDR picture.

Input and ouput pictures consist of three sample arrays of one of the following colour representation methods:

- Luma and two Chroma (YCbCr).
- Intensity and two Chroma (ICtCp) as defined in Recommendation ITU-R BT.2100-0 [6].

The dimension of a luma array is specified as (PicWidth)x(PicHeight) and the dimension of a chroma array is specified as (PicWidthC)x(PicHeightC). PicWidth specifies the width of each luma array in units of luma samples. PicHeight specifies the height of each luma array in units of luma samples. PicWidthC specifies the width of each chroma array in units of chroma samples. PicHeightC specifies the height of each chroma array in units of chroma samples.

NOTE: For convenience of notation and terminology in the present document, the variables and terms associated with these arrays are referred to as luma (or Y) and chroma, where the two chroma arrays are referred to as Cb and Cr; regardless of the actual colour representation method in use.

5.2.2 Input picture format

BL picture shall obey the following constraints:

a) Transfer characteristics of the BL picture shall be as defined in SMPTE ST 2084 [1] or as defined in Recommendation ITU-R BT.1886-0 [5].

b) Colour primaries shall be as defined in Recommendation ITU-R BT.2100-0 [6], Recommendation ITU-R BT.709-6 [7] or Recommendation ITU-R BT.2020-2 [8].

BL picture and EL picture shall obey the following constraints:

a) BL picture and EL picture shall have the same vertical and horizontal resolution (see also note 1 in clause 4).

b) For BL and EL picture, each of the two chroma arrays shall have half the height and half the width of the luma array.

c) BL and EL shall have the same framerate.

d) Samples in each array of BL picture and EL picture shall have the same bit depth.

e) The luma sample phase of the BL and EL shall be identical.