UHD World Association
世界超高清视频产业联盟

HDR Video Technology Part 2-1

Application Guide to System Integration

（Version NO. 1.0）

Release Time
2022-02-10

UHD World Association (UWA)
T/UWA 005.2-1-2022

Contents

High Dynamic Range Video Technology

Part 2-1: Application Guide to System Integration

1 Scope

This document stipulates HDR System Integration specified in T/UWA 005.1-2022, including production, transmission, reception and decoding.

2 Normative References

The following documents are indispensable to the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 14496-12:2020 Information technology — Coding of audio-visual objects, part 12: ISO base media file format.

ITU-T Rec. T.35:2000 SERIES T: Terminals for Telematic Services Procedure for the allocation of ITU-T defined codes for non-standard facilities.

T/UWA 005.1-2022 High Dynamic Range Video Technology – Part 1: Metadata and Tone Mapping.

3 Terms and Definitions

The terms and definitions below apply to this document.

3.1 HDR Vivid

HDR Vivid refers to an HDR technical standard provided in T/UWA 005.1-2022. It is the generic term for derivative technologies of the kind.

3.2 HDR Vivid Playback Device

A device that decodes video, processes metadata and images in accordance with T/UWA 005.1-2022, and outputs the result through a digital video interface.

3.3 HDR Vivid Display Device

A device that processes metadata and images in accordance with T/UWA 005.1-2022 and displays the images.

4 Abbreviations

For the purpose of this document, the abbreviations below apply.

HDR	High Dynamic Range
EOTF	Electro-Optical Transfer Function
OETF	Opto-Electrical Transfer Function
PQ	Perceptual Quantizer
HLG	Hybrid Log-Gamma
DASH	Dynamic Adaptive Streaming over HTTP
HLS	HTTP Live Streaming

5 Overview

HDR Vivid video streams contain dynamic metadata streams that comply with T/UWA 005.1-2022. The decoder needs to parse the dynamic metadata information of every frame in the video streams. The device needs to use dynamic metadata information, and process it according to T/UWA 005.1-2022. The process is shown in the Figure below.

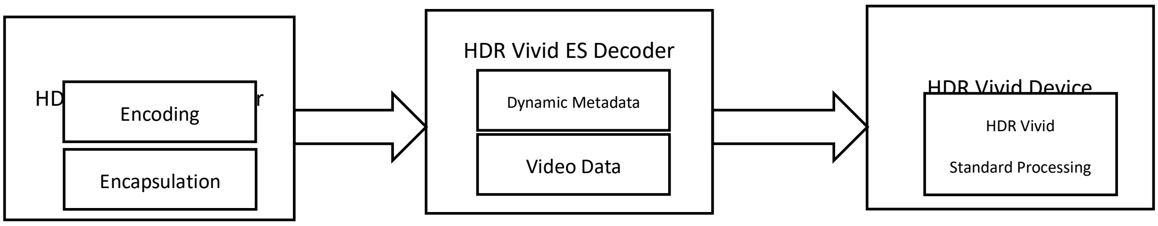

Figure 1 HDR Vivid Video Stream Processing Flow Diagram

The HDR Vivid ES Generator is responsible for generating HDR Vivid elementary streams (ES). Its Encoding Module generates ES via encoder, and its encapsulation module encapsulates ES into the formats needed, such as MP4, HLS, and DASH.

6 HDR Vivid ES Encapsulation

6.1 General Requirement

During HDR Vivid ES encapsulation, each frame must contain and only contains the dynamic metadata of that frame.

6.2 HEVC/H.265 and VVC/H.266 ES Encapsulation

The encapsulation of the dynamic metadata of every HDR Vivid frame shall comply with ITU-T Rec. T.35 (02/2000) and the following requirements:

1) Dynamic metadata is encapsulated in user_data_registered_itu_t_t35(payloadSize);

2) itu_t_t35_country_code in user_data_registered_itu_t_t35(payloadSize) is 0x26. itu_t_t35_country_code is the country code, and 0x26 indicates China.

3) terminal_provide_code in itu_t_t35_payload() is 0x0004, and terminal_provide_oriented_code is 0x0005. terminal_provide_code is the organization code, and 0x0004 indicates HDR Vivid. terminal_provide_oriented_code is the application code, defined by the application organization itself. 0x0005 indicates HDR Vivid Version 1.0. See Table 6.4 for the version evolution plan.

See Table 1 for the ES Encapsulation Syntax.

Table 1 ES Encapsulation Syntax

Pseudocode	**Descriptor**
user_data_registered_itu_t_t35(payloadSize) {	
itu_t_t35_country_code	0x26
if(itu_t_t35_country_code != 0xFF){	
i = 1	
}	
else {	
itu_t_t35_country_code_extension_byte	b(8)
i = 2	
}	
do {	
itu_t_t35_payload_byte	b(8)
i++	
} while(i < payloadSize)	
}	

Data in **itu_t_t35_payload_byte** is represented as below.

Table 2 T35 Payload Data Structure

itu_t_t35_payload (){	**Descriptor**
terminal_provide_code	0x0004
terminal_provide_oriented_code	0x0005 (v1.0)
dynamic_metadata ()	
}	

6.3 AVS2 and AVS3 ES Encapsulation

The HDR Vivid dynamic metadata of each frame in the HDR Vivid AVS2 ES and HDR Vivid AVS3 ES is encapsulated in hdr_dynamic_metadata_extension() of extension_data(i). In hdr_dynamic_metadata_extension(), extension_id is 0x5 (4 bits), and hdr_dynamic_metadata_type is 0x5 (4 bits). See the table below for more detail.

Table 3 HDR Image Extension Definition 1

Definition of HDR Image Metadata Extension	Descriptor
hdr_dynamic_metadata_extension( ) {	
extension_id	0x5
hdr_dynamic_metadata_type	0x5
while (next_bits(24) ! '0000 0000 0000 0000 0000 0001') {	
extension_data_byte	u(8)
}	
next_start_code()	
}	

The syntax of **extension_data_byte()** in Table 3 is shown in the table below.

Table 4 Data Table of extension_data_byte

Definition of Dynamic Metadata	Descriptor
extension_data_byte ( ) {	
itu_t_t35_country_code	0x26
terminal_provide_code	0x0004
terminal_provide_oriented_code	0x0005 (v1.0)
system_start_code	u(8)
if(system_start_code==0x01){	
num_windows=1	
for(w = 0; w < num_windows; w++) {	
minimum_maxrgb_pq[w]	u(12)
marker_bit	f(1)
average_maxrgb_pq[w]	u(12)
marker_bit	f(1)
variance_maxrgb_pq[w]	u(12)
marker_bit	f(1)
maximum_maxrgb_pq[w]	u(12)
marker_bit	f(1)
}	
for(w = 0; w < num_windows; w++) {	
tone_mapping_enable_mode_flag[w]	u(1)
if(tone_mapping_enable_mode_flag [w]==1){	
tone_mapping_param_enable_num [w]	u(1)
tone_mapping_param_enable_num [w]++	
for(i=0; i< tone_mapping_param_enable_num [w]; i++){	
targeted_system_display_maximum_luminance_ pq[i][w]	u(12)
base _enable_flag[i][w]	u(1)
marker_bit	f(1)

Table 4 Data Table of extension_data_byte (continued)

Definition of Dynamic Metadata	Descriptor
if(base _enable_flag[i][w]){	
base_param_m_p[i][w]	u(14)
base_param_m_m[i][w]	u(6)
marker_bit	f(1)
base_param_m_a[i][w]	u(10)
base_param_m_b[i][w]	u(10)
marker_bit	f(1)
base_param_m_n[i][w]	u(6)
base_param_K1[i][w]	u(2)
base_param_K2[i][w]	u(2)
base_param_K3[i][w]	u(4)
base_param_Delta_enable_mode[i][w]	u(3)
marker_bit	f(1)
base_param_enable_Delta[i][w]	u(7)
}	
3Spline _enable_flag[i][w]	u(1)
if(3Spline_enable_flag[i][w]){	
3Spline_enable_num[i][w]	u(1)
3Spline_enable_num[i][w]++;	
for(j = 0; j < 3Spline_enable_num[i][w]; j ++) {	
3Spline_TH_enable_mode[j] [i][w]	u(2)
if((3Spline_TH_mode[j][i] [w]==0)|| (3Spline_TH_mode[j][i] [w]==2)){	
3Spline_TH_enable_MB [j][i][w]	f(8)
}	
marker_bit	f(1)
3Spline_TH_enable[j][i][w]	f(12)
marker_bit	f(1)
3Spline_TH_enable_Delta1 [j][i][w]	f(10)
3Spline_TH_enable_Delta2 [j][i][w]	f(10)
marker_bit	f(1)
3Spline_enable_Strength[j][i][w]	f(8)
}	
}	
}	
}	
}	
color_saturation_mapping_enable_flag[w]	u(1)
if(color_saturation_mapping_enable_flag[w]) {	
color_saturation_enable_num[w]	u(3)

for(i = 0; i< color_saturation_enable_num [w]; i++) {	

Table 4 Data Table of extension_data_byte (continued)

Definition of Dynamic Metadata	Descriptor
color_saturation_enable_gain[i][w]	u(8)
marker_bit	f(1)
}	
}	
}	
}	
stuffing_bit	'1'
while(!byte_aligned())	
stuffing_bit	'0'
}	

terminal_provide_code and terminal_provide_oriented_code in **extension_data_byte()** correspond respectively to terminal_provide_code and terminal_provide_oriented_code in itu_t_t35_payload() described in 6.2. terminal_provide_code is 0x0004 (16 bits), and terminal_provide_oriented_code (16 bits) is the current version number.

marker_bit in **extension_data_byte()** is to avoid 21 or more consecutive zeros starting from any byte-aligned location. See T/UWA 005.1-2022 for other syntax elements of **extension_data_byte()**.

6.4 Backward Compatibility

There are 4 versions of HDR Vivid. Each version generates an independent set of dynamic metadata, encapsulated in one ES, meaning each ES can carry dynamic metadata of multiple, not necessarily continuous versions. For instance, one ES can carry the dynamic metadata of both versions 1 and 2; or it can carry the metadata of versions 1, 3 and 4. terminal_provide_oriented_code in itu_t_t35_payload() indicates version number. It is recommended that terminal devices extract the dynamic metadata of the highest version number for post-processing. See the table below for the mapping between version numbers and code words.

Table 5 Mapping Table of Version Numbers and Code Words

HDR Vivid Version Numbers	Code Words of terminal_provide_oriented_code in itu_t_t35_payload()
1.0	0x0005
2.0	0x0006
3.0	0x0007
4.0	0x0008

Note: If the terminal device cannot identify a version of HDR Vivid, then neglect that version. If it cannot identify any of the versions, then skip HDR Vivid processing.

7 HDR Vivid MP4 File Encapsulation

In HDR Vivid MP4 files, the original MP4 structure (in accordance with ISO/IEC 14496-12(2015-12-15)) remains unchanged except for two changes:

1) A new CUVV Configuration Box describing HDR Vivid video ES format is added in the extension of VisualSampleEntry Box of Sample Description Box(stsd) of the video track;
2) The data field "compressorname" in VisualSampleEntry Box is changed into "HDR Vivid video". (Optional, as applications identify whether a video is HDR Vivid video via CUVV Box.)
See the added CUVV Box descriptions below.

Table 6 CUVV Box Descriptions

class CUVVConfigurationBox extends Box('cuvv')	Descriptor
{	
unsigned int (16) cuva_version_map;	0x0009 (Version number information)
unsigned int (16) terminal_provide_code;	0x0004
unsigned int (16) terminal_provide_oriented_code;	0x0008 (Highest version number)
const unsigned int (32)[4] reserved = 0;	
}	

The four-character code of the newly-added box is CUVV. The internal structure specifies the version number information of HDR Vivid video. The 16-bit cuva_version_map can represent up to 16 version numbers, with each bit representing one. The highest bit indicates the highest version number, and the lowest bit the lowest version number. For instance, if cuva_version_map is 0x0009, the ES contains versions 4 and 1 (binary number 1001 being 9 in decimal).
The other two descriptors correspond to terminal_provide_code and terminal_provide_oriented_code described in the itu_t_t35_payload() structure in 6.2. terminal_provide_code (16 bits) is 0x0004, and terminal_provide_oriented_code (16 bits) is the value representing the highest version contained in the current ES. For instance, if cuva_version_map is 0x0009, the highest version in the current ES is 4; and according to Table 5 in section 6.4, terminal_provide_oriented_code for version 4 is 0x0008.
To add cuvv in a MP4 file, find moov->trak(video)->mdia->minf->stbl->stsd, and add a CUVV Box in VisualSampleEntry Box.
In accordance with ISO/IEC 14496-12(2015-12-15), VisualSampleEntry is described below.

Table 7 VisualSampleEntry Description

class VisualSampleEntry(codingname) extends SampleEntry (codingname){	Descriptor
unsigned int(16) pre_defined = 0;	
const unsigned int(16) reserved = 0;	
unsigned int(32)[3] pre_defined = 0;	
unsigned int(16) width;	
unsigned int(16) height;	
template unsigned int(32) horizresolution = 0x00480000; // 72 dpi	
template unsigned int(32) vertresolution = 0x00480000; // 72 dpi	
const unsigned int(32) reserved = 0;	

template unsigned int(16) frame_count = 1;	
string[32] compressorname;	
template unsigned int(16) depth = 0x0018;	
int(16) pre_defined = -1;	
// other boxes from derived specifications	
CleanApertureBox clap; // optional	
PixelAspectRatioBox pasp; // optional	
}	

The newly-added CUVV Box is placed in "other boxes from derived specifications", and added after other extension Boxes.

For instance, in a H.265 MP4 file, VisualSampleEntry is "hvc1" or "hev1". hvcC Box describes video coded ES. In a HDR Vivid MP4 file, a new Box is added after hvcC Box – CUVV Box, and the description of "Compression name" is changed to "CUVA HDR Video". The rest of the structure remains unchanged.

A sample HDR Vivid MP4 file:

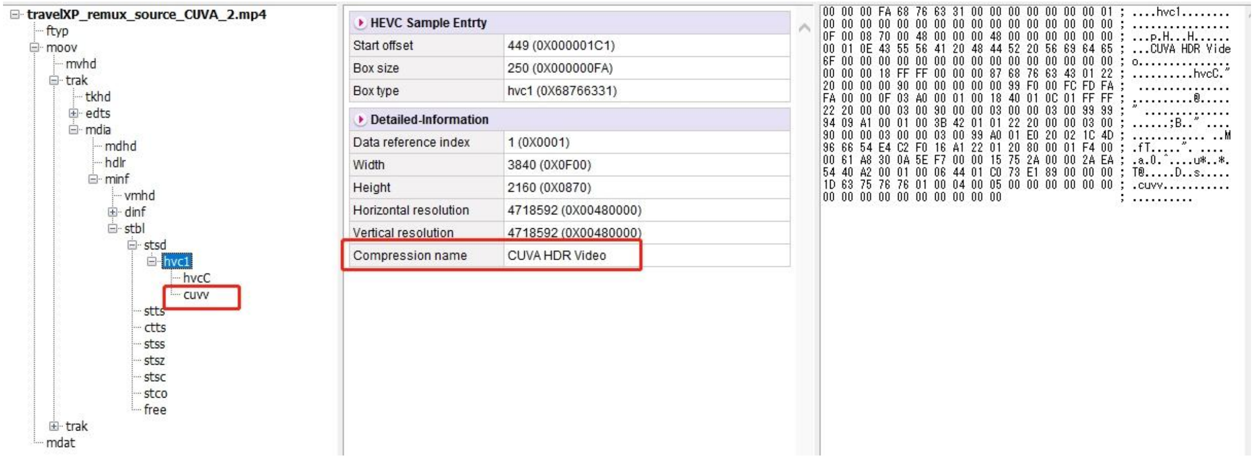

Figure 2 HDR Vivid MP4 File Description

8 HDR Vivid HLS Stream Encapsulation

8.1 General Requirements

The playlist file of HDR Vivid HLS stream is an m3u8 file with added HDR Vivid video information. Other than that, it is compliant with the HLS stream standard *HTTP Live Streaming 2nd Edition*.

8.2 HLS Stream Encapsulation

EXT-X-STREAM-INF and EXT-I-STREAM-INF tags are used to specify video or audio attributes in HLS streams. In additional to the m3u8 description, HDR Vivid HLS encapsulation has put in an EXT-X STREAM-INF or EXT-I-STREAM-INF tag for each video stream. The CODECS attribute uses HDR Vivid descriptor in the format of [CUVAHDR_video].[CUVA_Version_map] to replace the original description. CUVAHDR_video specifies the HDR Vivid elementary stream, and its value is "cuvv" for both MP4 and TS encapsulations. CUVA_Version_map specifies HDR Vivid version number, and each bit represents one version. The highest bit indicates the highest version number, and the lowest bit the lowest version number. For example, "1101" means that the ES contains versions 4, 3, and 1. Other audio and video attributes remain unchanged. Among them, the VIDEO-RANGE attribute is in line with what *HTTP Live Streaming 2nd Edition draft-pantos-hls-rfc8216bis-02* specifies.

Examples of HLS M3u8 description:

/*New description of two audio/video streams with different bit rates*/

#EXT-X-STREAM-INF:BANDWIDTH=5120000,VIDEO-RANGE=PQ,CODECS="cuvv.1101,mp4a.40.29", FRAME-RATE=50.000,RESOLUTION=3840x2160,AUDIO="aac"

low/video.m3u8

#EXT-X-STREAM-INF:BANDWIDTH=7680000,VIDEO-RANGE=PQ,CODECS="cuvv.1101,mp4a.40.29", FRAME-RATE=50.000,RESOLUTION=3840x2160,AUDIO="aac"

high/video.m3u8

/*Original description of two audio/video streams with different bit rates*/

#EXT-X-STREAM-INF:BANDWIDTH=5120000,VIDEO-RANGE=PQ, CODECS="hvc1.2.4.L153.b0, mp4a.40.29",FRAME-RATE=50.000,RESOLUTION=3840x2160,AUDIO="aac"

low/video.m3u8

#EXT-X-STREAM-INF:BANDWIDTH=7680000,VIDEO-RANGE=PQ,CODECS="hvc1.2.4.L153.b0, mp4a.40.29",FRAME-RATE=50.000,RESOLUTION=3840x2160,AUDIO="aac"

high/video.m3u8

9 HDR Vivid DASH Encapsulation

9.1 General Requirements

The media presentation description (MPD) file of HDR Vivid DASH (Dynamic Adaptive Streaming over HTTP) is in line with the requirements in ISO/IEC FDIS 23009, the international standard for MPEG-DASH, except for additional Representation information about HDR Vivid videos.

9.2 DASH Encapsulation

In a DASH MPD file, the Representation field specifies attributes of the video or audio, including the frame rate, resolution, bit rate, encoder, and bandwidth.

Apart from the original Representation description, the HDR Vivid DASH encapsulation adds another Representation for each video stream. The CODECS attributes here are described with HDR Vivid descriptors instead of the original ones, in the format of [CUVAHDR_video].[CUVA_Version]. The HDR Vivid_video tag specifies the HRD Vivid elementary stream, and its value is "cuvv" for both MP4 and TS encapsulations of the DASH streams. [CUVA _Version_map] specifies the HDR Vivid version numbers, and each bit represents one version. The highest bit indicates the highest version number, and the lowest bit the lowest version number. For example, "1101" means that the ES contains versions 4, 3, and 1. DASH audios and videos have their own descriptions respectively, and therefore the CODECS here only describes video information and the new Representation is only added to each of the video tracks.

Examples of MPD file description in HDR Vivid DASH elementary streams:

<Representation id="bbb_30fps_1024x576_2500k" codecs="cuvv.110" bandwidth="3134488"/> //added

<Representation id="bbb_30fps_1280x720_4000k" codecs="cuvv.110" bandwidth="4952892" /> //added

<Representation id="bbb_30fps_1024x576_2500k" codecs=" hvc1.2.4.L153.b0" bandwidth="3134488" />

<Representation id="bbb_30fps_1280x720_4000k" codecs=" hvc1.2.4.L153.b0" bandwidth="4952892" />

10 HDR Vivid DVB Transport Stream (TS) Encapsulation

10.1 General Requirements

The syntax of HDR Vivid DVB TS streams is in line with *ETSI EN 300 468 V1.16.1 (2019-08)* except for an additional description of component_descriptor regarding the HDR Vivid elementary streams. To be specific, the Program Mapping Table (PMT) is modified.
Fault tolerance strategies are applied to HDR programs with missing metadata.

10.2 TS Encapsulation

Based on the syntax of the original component_descriptor, the following fields are added to specify the HDR Vivid dynamic metadata in use. Except for the two additional descriptors, the rest remains the same.

(1) registration_descriptor () is added to the video description in the PMT. Its built-in field is defined as follows:

Table 8 Registration Descriptor

registration_descriptor ()	descriptor
{	
unsigned int (8) descriptor_tag;	**0x05**
unsigned int (8) descriptor_length;	**0x04**
unsigned int (32) Format_identifier;	**'cuvv'**
}	

descriptor_tag is 0x05, specifying it as a registration descriptor.

descriptor_length is 0x04, specifying the length of the descriptor "cuvv".

The value of Format_identifier ("cuvv") must be 0x63757676 which is used to instantly decide if the TS is an HDR Vivid ES.

(2) A user-defined descriptor is added to the video description in the PMT to indicate information such as the HDV Vivid version. This descriptor is defined as follows:

Table 9 User-Defined Descriptor

CUVV_video_stream_descriptor()	Descriptor
{	
unsigned int (8) descriptor_tag;	0xF3
unsigned int (8) descriptor_length;	0x0A
unsigned int (32) cuvv_tag ;	'cuvv'
unsigned int (16) cuva_version_map;	0x0005 (version numbers)
unsigned int (16) terminal_provide_code;	0x0004

unsigned int (16) terminal_provide_oriented_code;	0x0007 (the highest version number)
}	

descriptor_tag is 0xF3 defined by the user.

descriptor_length is 10.

The value of cuvv_tag is 0x63757676 ("cuvv"). It is used to further specify that the content of this descriptor is CUVV information when there is a conflict due to the user-defined descriptor_tag 0xF3.

CUVA_Version_map specifies the HDR Vivid version numbers. It has 16 bits and each bit represents one version. The highest bit indicates the highest version number, and the lowest bit the lowest version number. For example, when CUVA_Version_map is 0x0005 (5 in decimal is 0101 in binary), it means that there are versions 3 and 1 in the ES.

The last two descriptors are corresponding to "terminal_provide_code" and "terminal_provide_oriented_code" in the "itu_t_t35_payload ()" statement mentioned in Section 6.2 .Terminal_provide_code is 0x0004 (16 bits). Terminal_provide_oriented_code (16 bits) is the value specifying the highest version number of the bit stream. For example, when CUVA_Version_map is 0x0005, it means the highest version of the ES is version 3. From Table 5 in Section 6.4 we can see that terminal_provide_oriented_code corresponding to version 4 is 0x0007.

10.3 Strategy for TS packet loss

If the HDR Vivid DVB streams in TS encapsulation cannot synchronize or there are over three lost packets, the synchronization should be restarted.

11 Explanation of Terminal Application

11.1 Composition of the Terminal System

Functional Modules

The HDR Vivid terminal system has two functions, video receiving and decoding, and video display. The process is shown in the following figure.

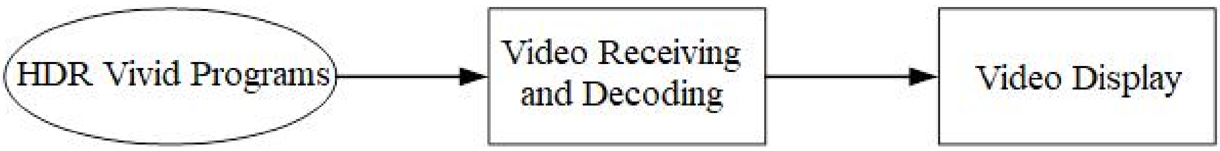

Figure 3 Functions of HDR Vivid Terminal System

Video receiving and decoding is done in playback devices which have network ports, coaxial connectors, USB ports, and video decoding abilities, mainly set-top-boxes, media players, and display devices with decoding abilities. HDR Vivid playback devices refers to those that receive, decode, and output HDR Vivid video content and metadata.

Display devices are used to present the content of videos, mainly televisions, projectors, monitors, laptops, mobile phones and other devices that can present images. Display devices can be put into the following categories based on their ability to support HDR Vivid.

Table 10 Display Device Categories

Category	HDR Vivid Display Device	HDR Display Device	SDR Display Device
Description	In line with HDR Vivid specification, can receive and process the HDR Vivid metadata, and display the video signals	Can only receive, decode, and display PQ-HDR or HLG_HDR video signals and static metadata defined under BT.2100	Do not support HDR signals defined by BT.2100, can only receive, decode, and display SDR video signals under BT.2020, BT.709, and BT.601

11.1.1 Application Methods

There are two application methods for the terminal system:

(1) All-in-one method

The display device receives HDR Vivid signals directly from the video distribution network, and then decodes, plays, and displays video on its own.

(2) Playback device method

The display device is connected via the digital visual interface to a playback device which receives HDR Vivid signals from the video distribution network and decodes the signals. Then the display device receives the decoded video signals from the playback device and displays the video content.

A playback device only needs to support the playback device method. A display device should support either one of the methods.

11.2 Decoding and Presentation of HDR Vivid Signals in the Device

11.2.1 Basic Requirements

A playback device should provide at least one adaptation mode, under which it performs only the processing required under this standard and other necessary image processing, to ensure that the output image and information meet the requirement of this standard.

A display device should provide at least one adaptation mode, under which it performs image processing based on the requirements of the authentication mode defined in this standard, ensuring the presentation of HDR Vivid videos is in line with this standard.

11.2.2 Adaptation Modes under the Playback Device Method

Under the playback device method where the playback device and the display device work together for the decoding and presentation of HDR Vivid videos, there are two adaptation modes.

(1) Receiving end adaptation mode

Under this mode, the playback device sends HDR Vivid signals and the dynamic metadata to the display device. Then, the display device performs color signal dynamic range conversion and color correction based on the requirements specified in T/UWA 005.1-2022. Here, when the display device receives video images with varying information inputs such as different transfer character curves, color gamuts, or color space conversion matrices, the display should remain stable with no visible flickering or black screen.

(2) Monitor adaptation mode

Under this mode, the playback device performs color signal dynamic range conversion and color correction based on the requirements specified in T/UWA 005.1-2022 with the dynamic range information in the Extended Display Identification Data (EDID) provided by the display device. Here, when the display device receives image signals sent by the playback device, it does not perform color dynamic range conversion or color correction. It performs signal conversion based on the transfer character curve and other necessary processing, and then present the video.

An HDR Vivid playback device should fully support both adaptation modes mentioned above. If an HDR Vivid display device has a digital visual interface which supports the HDR Vivid format, it needs to support at least one of the above modes. The HDR Vivid playback device should prioritize the receiving end adaptation mode when connecting with the HDR Vivid display device, and only use the monitor adaptation mode when the HDR Vivid display device only supports this mode.

11.2.3 Adaptation for Interworking with PQ HDR Signals

When the HDR Vivid playback device receives HDR Vivid signals in PQ HDR format, it needs to convert the signal format based on the type of display device it connects with. Details are in the following table:

Table 11 PQ Signal Format Conversion for Display Device

Type of Display Device	HDR Vivid Display Device		HDR Display Device	SDR Display Device
	Mastering Display Adaptation Mode	Receiving End Adaptation Mode		
Output Signal from Playback Device	Processed PQ HDR signals and HDR Vivid VS-IF information frames	PQ HDR signals and VS-EMDS information frames of HDR Vivid dynamic metadata	(1) PQ HDR signals and static metadata (if the video signals contain static metadata); (2) Or processed PQ HDR signals	SDR signals after conversion (the most preferable output color gamut is that defined in BT2020, followed by BT709 and lastly BT601)

11.2.4 Adaptation for Interworking with HLG HDR Signals

When the HDR Vivid playback device receives HDR Vivid signals in HLG HDR format, it needs to convert the signal format based on the type of display device it connects with. Details are in the following table:

Table 12 HLG Signal Format Conversion for Display Device

Type of Display Device	HDR Vivid Display Device		HDR Display Device	SDR Display Device
	Mastering Display Adaptation Mode	Receiving End Adaptation Mode		
Output Signal from Playback Device	Processed PQ HDR signals and HDR Vivid VS-IF	HLG HDR signals and VS-EMDS information	(1) HLG HDR ([2]PQ HDR) signals and static metadata	SDR signals after conversion (the most preferable

	information frames	frames of HDR Vivid dynamic metadata ([1] if the video signals contain HDR Vivid dynamic metadata)	(if the video signals contain static metadata) (2) Or processed PQ HDR signals	color gamut is that defined in BT2020, followed by BT709 and lastly BT601)

[1] HDR Vivid does not mandatorily require HLG HDR signals to carry dynamic metadata.

[2] Some HDR display devices support only the PQ HDR format and not the HLG HDR format. In this case, the playback device need to convert the HLG HDR signals to PQ HDR signals before sending them to the display device.

11.3 Dynamic Metadata Transfer over HDMI

Overview

When an HDR Vivid playback device and an HDR Vivid display device are connected via HDMI to transfer image data, dynamic metadata must be transferred as is required in this chapter.

HDR Vivid signal over HDMI has its optical-electric transfer feature defined in *transfer_character* field of VS-IF or VS-EMDS with BT. 2020 as its color gamut. The playback device does not have to transmit HDR static metadata frames specified under CTA-861.3.

When an HDR Vivid playback device is connected to an HDR Vivid display device, the playback device would query EDID of the display device for VS-VDB data defined in 11.3.1 to know how well the display device supports HDR Vivid. If VS-VDB indicates that HDR Vivid receiving end adaptation mode is supported, the playback device would encapsulate the HDR Vivid dynamic metadata in information frames as per 11.3.3, and send them along with HDR Vivid video image to the display device. If VS-VDB indicates that HDR Vivid monitor mode is supported, the playback device would refer to VS-VDB for maximum luminance and minimum luminance of the display device, and adjust video content accordingly as per Chapter 9 and Appendix B.3 of T/UWA 005.1-2022 before sending to the display device. The playback device would also send to the display device VS-IF information as per 11.3.2.

11.3.1 Define VS-VDB

HDR Vivid display device uses the VS-VDB data block in HDMI EDID to indicate its ability to receive HDR Vivid dynamic metadata.

11.3.1.1 Syntax

Below is the syntax of the HDR Vivid VS-VDB data block:

Table 13 Syntax of VS-VDB data block

Byte\Bit	7	6	5	4	3	2	1	0
0	Tag Code= (0x07)			Length=(14)				
1	Extended Tag Code=0x01 (VSVDB)							
2	IEEE OUI/CID (0x03）							
3	IEEE OUI/CID (0x75)							
4	IEEE OUI/CID (0x04）							
5	system_start_code[7…0]							
6	version_code[3…0]				0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)
7	display_maximum_luminance [31…0]							
8								
9								
10								
11	display_minimum_luminance [15…0]							
12								
13	monitor_mode_support	rx_mode_support	0 (Reserved)					
14	0 (Reserved)							

11.3.1.2 Semantics

system_start_code

8-bit unsigned integer. It indicates the HDR Vivid version supported by a device. If the device supports the current version, the value is 1.

If the adopted version is higher than what is represented in system_start_code of dynamic metadata data in the stream, dynamic metadata will be packed in HDR Vivid VS-EMDS package structure. Or, in current transmitting device, display adaptation is performed according to HDR Vivid metadata criteria before the metadata is sent to the receiving device.

version_code

4-bit unsigned integer. It indicates the HDR Vivid version number of a device. If it is current version, the value is 1. The transmitting device refers to the version information of the receiving device and then sends the dynamic metadata in the corresponding format.

display_maximum_luminance

32-bit unsigned integer. It indicates maximum luminance of a display device under D65. The unit of measurement is 0.0001 nit (cd/m^2).Code word 0x00000001 represents 0.0001 nit. The value of MaxDisplay is display_maximum_luminance/10000. If the display device supports monitor mode, value of display_maximum_luminance in the display device must be set properly for playback device to refer to when it does imaging processing. If not, the value of display_maximum_luminance is set as 0.

display_minimum_luminance

16-bit unsigned integer. It indicates the minimum luminance of a display device under D65. The unit of measurement is 0.0001 nit (cd/m^2). Code word 0x0001 represents 0.0001 nit. The value of MinDisplay is display_minmum_luminance/1000. If the display device supports monitor mode, the value of display_minimum_luminance in the display device must be set properly for playback device to refer to when it does imaging processing. If not, the value of display_minimum_luminance is set as 0.

monitor_mode_support

1 bit unsigned integer. It indicates if the display device supports monitor mode. If monitor_mode_support=1, the display device supports monitor mode. If monitor_mode_support=0, the display device does not support the monitor mode.

rx_mode_support

1 bit unsigned integer. It indicates whether a display device supports the receiving end adaptation mode. If rx_mode_support=1, the display device supports receiving end adaptation mode; if rx_mode_support=0, the device does not support the receiving end adaptation mode.

11.3.2 Define VS-IF

11.3.2.1 Syntax

In HDR Vivid monitor mode, static metadata is transferred from playback device to display device using a VS-IF information frame of HDMI. Below is the syntax of VS-IF information frames:

Table 14 Syntax of VS-IF information frames

Byte\Bit	7	6	5	4	3	2	1	0
HB0	VSIF Type Code =0x81							
HB1	VSIF Version=0x01							
HB2	0	0	0	0	0	Payload Length=27		
PB00	checksum							
PB01	IEEE OUI/CID (0x03)							
PB02	IEEE OUI/CID (0x75)							
PB03	IEEE OUI/CID (0x04)							
PB04	system_start_code [7…0]							
PB05	version_code [3…0]				Monitor_mode_enable =0x01	transfer_ character=0	0 (Rsvd)	0 (Rsvd)
PB06~PB27	0 (Reserved)							

11.3.2.2 Syntax

monitor_mode_enable

1 bit unsigned integer. In this guideline, monitor_mode_enable has a fixed value of 1. It indicates that current HDR Vivid signals are connected and processed in monitor mode.

transfer_character

1 bit unsigned integer. It represents transfer character of graphs. In this guideline, transfer_character has a fixed value of 0, meaning that under monitor mode, the graph transfer character is a ST2084 EOTF curve.

11.3.3 Define VS-EMDS

In HDR Vivid receiving end adaptation mode, dynamic metadata are transferred from playback device to display device through VS-EMDS information frames of HDMI.

11.3.3.1 Syntax

Below is the syntax of HDR Vivid VS-EMDS:

Table 15 Syntax of VS-EMDS

<table>
<tr><th>Byte\ Bit</th><th>7</th><th>6</th><th>5</th><th>4</th><th>3</th><th>2</th><th>1</th><th>0</th></tr>
<tr><td>HB0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr>
<tr><td>HB1</td><td>1 (First)</td><td>0 (Last)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td></tr>
<tr><td>HB2</td><td></td><td colspan="7">0 (Sequence_Index)</td></tr>
<tr><td>PB0</td><td>1 (New)</td><td>0 (End)</td><td colspan="2">1 (DS_Type)</td><td>0 (AFR)</td><td>1 (VFR)</td><td>1 (Sync)</td><td>0 (Rsvd)</td></tr>
<tr><td>PB01</td><td colspan="8">0 (Rsvd)</td></tr>
<tr><td>PB02</td><td colspan="8">0 (Organization_ID)</td></tr>
<tr><td>PB03</td><td colspan="8">0x00 (Data_Set_Tag_MSB)</td></tr>
<tr><td>PB04</td><td colspan="8">0x02 (Data_Set_Tag_LSB)</td></tr>
<tr><td>PB05</td><td colspan="8">0x00 (Data_Set_Length_MSB)</td></tr>
<tr><td>PB06</td><td colspan="8">0x38 (Data_Set_Length_LSB)</td></tr>
<tr><td>PB07</td><td colspan="8">IEEE OUI/CID (03)</td></tr>
<tr><td>PB08</td><td colspan="8">IEEE OUI/CID (75)</td></tr>
<tr><td>PB09</td><td colspan="8">IEEE OUI/CID (04)</td></tr>
<tr><td>PB10</td><td colspan="8">system_start_code [7…0]</td></tr>
<tr><td>PB11</td><td colspan="4">version_code [3…0]</td><td colspan="4">minimum_maxrgb_pq [11…8]</td></tr>
<tr><td>PB12</td><td colspan="8">minimum_maxrgb_pq [7…0]</td></tr>
<tr><td>PB13</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td colspan="4">average_maxrgb_pq [11…8]</td></tr>
<tr><td>PB14</td><td colspan="8">average_maxrgb_pq [7…0]</td></tr>
<tr><td>PB15</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td colspan="4">variance_maxrgb_pq [11…8]</td></tr>
<tr><td>PB16</td><td colspan="8">variance_maxrgb_pq [7…0]</td></tr>
<tr><td>PB17</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td colspan="4">maximum_maxrgb_pq [11…8]</td></tr>
<tr><td>PB18</td><td colspan="8">maximum_maxrgb_pq [7…0]</td></tr>
<tr><td>PB19</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td colspan="4">targeted_system_display_maximum_luminance_ pq [11…8]</td></tr>
<tr><td>PB20</td><td colspan="8">targeted_system_display_maximum_luminance_ pq [7…0]</td></tr>
<tr><td>PB21</td><td>Transfer _Character</td><td>base_ enable flag[0]</td><td colspan="6">base_param_m_p [13…8]</td></tr>
<tr><td>PB22</td><td colspan="8">base_param_m_p [7…0]</td></tr>
<tr><td>PB23</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td colspan="6">base_param_m_m [5…0]</td></tr>
<tr><td>PB24</td><td>0(Rsvd)</td><td>0(Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td>0 (Rsvd)</td><td colspan="2">base_param_m_a[9,8]</td></tr>
</table>

Byte\ Bit	7	6	5	4	3	2	1	0
PB25	base_param_m_a[7…0]							
PB26	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	base_param_m_b[9…8]	
PB27	base_param_m_b [7…0]							
HB0	0	1	1	1	1	1	1	1
HB1	0 (First)	0 (Last)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)
HB2		1(Sequence_Index)						
PB0	0 (Rsvd)	0 (Rsvd)	base_param_m_n [5…0]					
PB1	0 (Rsvd)	0 (Rsvd)	base_param_K1 [1..0]		base_param_K2[1..0]		base_param_K3[1..0]	
PB2	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0(Rsvd)	base_param_Delta_enable_mode[2..0]		
PB3	0 (Rsvd)	base_param_enable_Delta[6…0]						
PB4	0(Rsvd)	0(Rsvd)	0(Rsvd)	3Spline_enable_num		3Spline _enable_flag[0]	3Spline_TH_enable_mode0[1, 0]	
PB5	3Spline_TH_enable_MB0[7…0]							
PB6	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	3Spline_TH_enable0[11…8]			
PB7	3Spline_TH_enable0[7…0]							
PB8	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	3Spline_TH_enable_0Delta1[9,8]	
PB9	3Spline_TH_enable_0Delta1[7…0]							
PB10	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	3Spline_TH_enable_0Delta2[9,8]	
PB11	3Spline_TH_enable_0Delta2[7…0]							
PB12	3Spline_enable_Strength0[7…0]							
PB13	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	3Spline_TH_enable_mode1[1, 0]	
PB14	3Spline_TH_enable_MB1[7…0]							
PB15	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	3Spline_TH_enable1[11…8]			
PB16	3Spline_TH_enable1[7…0]							
PB17	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	3Spline_TH_enable_1Delta1[9,8]	
PB18	3Spline_TH_enable_1Delta1[7…0]							
PB19	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	0 (Rsvd)	3Spline_TH_enable_1Delta2[9,8]	
PB20	3Spline_TH_enable_1Delta2[7…0]							
PB21	3Spline_enable_Strength1[7…0]							
PB22	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	color_saturation_enable_num[2…0]		

Byte\Bit	7	6	5	4	3	2	1	0
PB23	color_saturation_enable_gain0 [7…0]							
PB24	color_saturation_enable_gain1 [7…0]							
PB25	color_saturation_enable_gain2 [7…0]							
PB26	color_saturation_enable_gain3 [7…0]							
PB27	color_saturation_enable_gain4 [7…0]							
HB0	0	1	1	1	1	1	1	1
HB1	0(First)	1(Last)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)	0(Rsvd)
HB2		2(Sequence_Index)						
PB0	color_saturation_enable_gain5 [7…0]							
PB1	color_saturation_enable_gain6 [7…0]							
PB2	color_saturation_enable_gain7 [7…0]							
PB3	graphic_ source_display_value [7…0]							
PB4	0 (Reserved)							
PB5	max_display_mastering_luminance,MSB							
PB6	max_display_mastering_luminance,LSB							
PB7~PB27	0 (Reserved)							

11.3.3.2 Syntax

version_code
4 bit unsigned integer. It indicates the system version. If it is current version, the value is 1.
transfer_character
1 bit unsigned integer. It indicates the graph transfer character. If the value of transfer_character is 0, the transfer character is a ST2084 EOTF curve. If transfer_character is 1, the transfer character is a Hybrid Log-Gamma(HLG) OETF curve.
graphic_source_display_value_pq
8 bit unsigned integer. It indicates the display luminance code values of PQ nonlinear normalized content of the image that is transferred from a playback device to a display device. Graph content's color gamut and transfer character are consistent with those of the video. The transfer character is determined by transfer_character field of the image. If graphic_source_value_pq is 0, there is no graph transfer. Graph's fixed curve source luminance parameter graphic_source_value_pq takes graphic_source_display_value_pq/255 as its value. See 11.3.4 graph processing for the display device adaptation under receiving end adaptation mode.
max_display_mastering_luminance
16-bit unsigned integer. It is generated by combining max_display_mastering_luminance, LSB and max_display_mastering_luminance, MSB. If the code value is 0x0001, the max_display_mastering_luminance value is 1 nit, and if the code value is 0xFFF, the max_display_mastering_luminance value is 65535 nits.
For the rest parameters, see grammar related to dynamic metadata in T/UWA 005.1-2022.

11.3.4 Graph processing for display device adaptation in receiving end adaptation mode

a) When the receiving end receives VS-EMDS packages, it extracts from them HDR Vivid metadata and graphic fixed curve range parameters.

b) See Chapter 9 of T/UWA 005.1-2022 to generate medium tone mapping curve P, including base curve parameter $P_{tone\_mapping}$, which includes m_p, m_m, m_n, m_a, m_b, K1, K2, and K3, one dimensional spline curve parameter $P_{1spline}$, which includes MB[0][0] and TH3[0], and cubic spline $P_{3spline}$, which includes TH1[3Spline_num], TH2[3Spline_num], TH3[3Spline_num], MA[2][3Spline_num], MB[2][3Spline_num], MC[2][3Spline_num], and MD[2][3Spline_num].

c) If graphic_source_value_pq = 0, set the final tone mapping curve S as P;

d) If graphic_source_value_pq > 0, graph's target display luminance is

$$graphic\_target\_value\_pq = \max(\min(graphic\_source\_value\_pq, MaxDisplayPQ \times 0.9), 0.5081)$$

Set the one dimensional spline curve $S_{1spline}$ including MB[0][0] and TH3[0] of S as follows:

MB[0][0] = graphic_target_value_pq /graphic_source_value_pq

TH3[0] = graphic_source_value_pq

base_offset = 0

Set the base curve parameter $S_{tone\_mapping}$, including s_m_p, s_m_m, s_m_n, s_m_a, s_m_b, s_K1, s_K2 and s_K3 of S as base curve $P_{tone\text{-}mapping}$, including m_p, m_m, m_n, m_a, m_b, K1, K2 and K3.

See 9.3 of T/UWA 005.1-2022 and generate cubic spline parameter $S_{3spline}$ following steps below.

1) Set TH1[1] = TH3[0], and calculate TH2[1] = TH1[1] + B, TH3[1] = TH2[1] + C, where B has a default value of 0.15, and C takes B/2 as its value; if TH3[1] > MaxSource, then TH3[1] = MaxSource. TH2[1] = TH1[1] + 2* (TH3[1] - TH1[1])/3. Generate cubic spline parameter $S_{3spline}$ as per 9.3.3.1 or 9.3.3.2 of T/UWA 005.1-2022, and check if this cubic spline curve is monotonically increasing (see 11.3.4.1). If yes, go to step 2; otherwise, go to step 3.

2) If TH1[3Spline_num] is greater than TH3[1], pass this cubic spline's parameter $P_{3spline}$, including TH1[3Spline_num], TH2[3Spline_num], TH3[3Spline_num], MA[2][3Spline_num], MB[2][3Spline_num], MC[2][3Spline_num] and MD[2][3Spline_num] to cubic spline parameter $S_{3spline}$; otherwise, values of TH1[2], TH2[2], TH3[2], and the latter cubic spline parameter are all set as 0, and the value of 3spline_num of metadata is set as 1. Terminate the cubic spline parameter generation process.

3) Set TH3[1] = MaxSource, TH2[1] = TH1[1] + 2*(TH3[1] - TH1[1])/3, and generate cubic spline parameter $S_{3spline}$, and check if the curve is monotonically increasing (see 11.3.4.1). If yes, go to step 4. Otherwise, go to step 5.

4) Find minimum TH3 in [TH1[1]+B+C, MaxSource] using binary search, and generate cubic spline parameter $S_{3spline}$ as per 9.3.3.1 or 9.3.3.2 of T/UWA 005.1-2022, and make it monotonically increasing (see 11.3.4.1). Search 10 times at most, and then terminate the curve parameter generation process.

5) TH3[1] = MaxSource, TH2[1] = TH3[1], MA[0][1] = graphic_target_value_pq, MB[0][1] = (VA2-MA[0][1])/(TH2[1]-TH1[1]), MC[0][1], MD[0][1], MA[1][1], MB[1][1], MC[1][1], and MD[1][1] all have a value of 0. Terminate the cubic spline parameter generation process.

11.3.4.1 Check if the Cubic Spline is Monotonically Increasing

In a cubic spline range, there are two cubic splines, and the equations are as follows:

$$F(L) = MD[0][n] \times (L - TH1[n])^3 + MC[0][n] \times (L - TH1[n])^2 + MB[0][n] \times (L - TH1[n])^1 + MA[0][n]$$

where L is an independent variable of range [TH1[n], TH2[n]];

$$F(L) = MD[1][n] \times (L - TH2[n])^3 + MC[1][n] \times (L - TH2[n])^2 + MB[1][n] \times (L - TH2[n])^1 + MA[1][n]$$

where L is an independent variable of range [TH2[n], TH3[n]], and 0 < n <= 3Spline_num.

Both cubic splines have to be monotonically increasing. Check with the following conditions.

a) Conditions to be met for the first cubic spline curve to increasing monotonically are:

If $\mathrm{TH1[n]} \leq (\mathrm{TH1[n]} - \frac{MC[0][n]}{3 \times MD[0][n]}) \leq TH2[n]$, then

$$MB[0][n] - \frac{MC[0][n]^2}{3 \times MD[0][n]} \geq 0$$

.

If $\mathrm{TH1[n]} > \left(\mathrm{TH1[n]} - \frac{MC[0][n]}{3 \times MD[0][n]}\right) \quad \left(\mathrm{TH1[n]} - \frac{MC[0][n]}{3 \times MD[0][n]}\right) > TH2[n]$, then

$$\begin{cases} MB[0][n] \geq 0 \\ 3 \times MD[0][n] \times (DTH2)^2 + 2 \times MC[0][n] \times (DTH2)^1 + MB[0][n] \geq 0 \end{cases}$$

where $DTH2 = (TH2[n] - TH1[n])$

b) Conditions to be met for the second cubic spline curve to increase monotonically are:

If $\mathrm{TH2[n]} \leq (\mathrm{TH2[n]} - \frac{MC[1][n]}{3 \times MD[1][n]}) \leq TH3[n]$, then

$$MB[1][n] - \frac{MC[1][n]^2}{3 \times MD[1][n]} \geq 0$$

.

If $\mathrm{TH2[n]} > \left(\mathrm{TH2[n]} - \frac{MC[1][n]}{3 \times MD[1][n]}\right) \quad \left(\mathrm{TH2[n]} - \frac{MC[1][n]}{3 \times MD[1][n]}\right) > TH3[n]$, then:

$$\begin{cases} MB[1][n] \geq 0 \\ 3 \times MD[1][n] \times (DTH3)^2 + 2 \times MC[1][n] \times (DTH3)^1 + MB[1][n] \geq 0 \end{cases}$$

Where $DTH3 = (TH3[n] - TH2[n])$.

12 Application Guide for Post-production

12.1 Input and Output Suggestions

The main tasks of post-production are: extracting and adjusting dynamic metadata and encapsulating files.

The basic requirement for post-production input is BT.2020 color gamut. It could be an RGB signal in a linear gamut, an RGB signal in a PQ gamut, or a YUV signal in a PQ gamut. It is recommended that the bit width of the linear gamut be at least 16 bits and that of the PQ gamut be at least 12 bits. YUV444 sampling format or YUV422 sampling format is recommended.

The output must be the 10-bit video data of YUV420, YUV422, or YUV444 format in the PQ gamut and dynamic metadata. If the output content is to be distributed directly, it is recommended that the ES file in the specified encoding format be output and dynamic metadata be embedded in the SEI header. If the output content is to be archived, it is recommended that an MXF file be output. Video content is stored in single frames in a high-bitrate format, such as XAVC. Dynamic metadata is embedded in the SEI data of each frame in a XAVC file.

12.2 Image Quality Monitoring

Post-production needs to output two channels for display. The first channel is output to a professional display monitor, as a high-quality reference. The second channel is processed using HDR Vivid dynamic metadata (tone mapping) and output to a consumer display device, such as a 500-nit HDR Vivid TV, as a consumer device reference.

12.3 Dynamic Metadata Generation Mode

Dynamic metadata can be generated in two modes: Auto Mode and Director Mode.

12.3.1 Auto Mode

In Auto Mode, the HDR Vivid dynamic metadata is extracted from each frame of the video, processed in real time using HDR Vivid, and output to the display device for playback. Auto Mode needs to integrate the automatic dynamic metadata extraction algorithm and post-processing algorithm specified in T/UWA 005.1-2022.

The Auto Mode UI should support the video read-in and playback, dynamic metadata extraction, HDR Vivid post-production, and display output of processed video content.

12.3.2 Director Mode

12.3.2.1 Overview

Director Mode is for colorists to adjust the dynamic metadata of each scene or frame manually, to achieve their desired effect. The basic requirements are to provide the HDR Vivid tone mapping curve adjustment controls, display the current tone mapping curve during adjustment, and support real-time post-processing based on the current dynamic metadata or curve and output to display.

Director Mode needs to integrate the dynamic metadata extraction algorithm, the post-processing algorithm and the Director Mode curve adjustment algorithm.

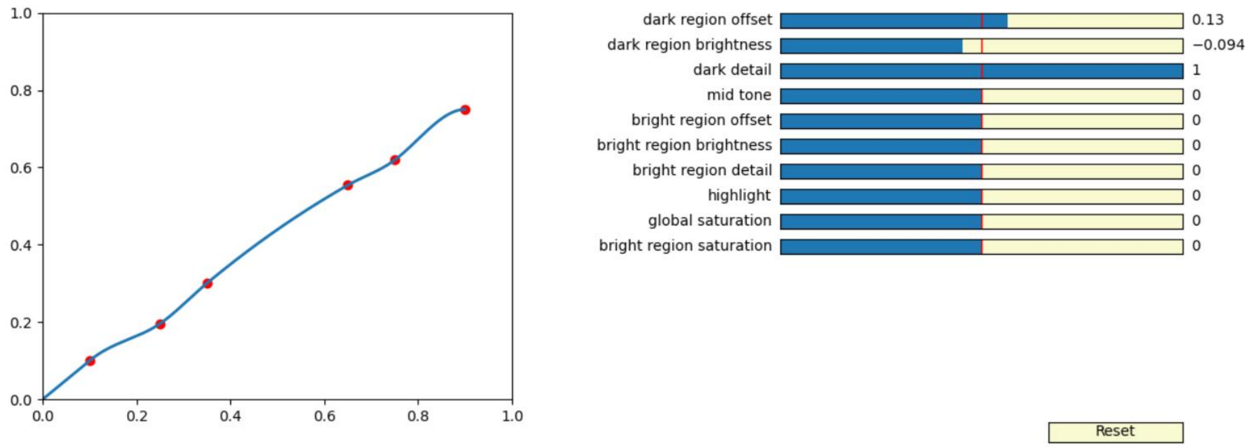

Figure 4 UI reference in HDR Vivid Director Mode

12.3.2.2 Director Mode Composition

The specifications of Director Mode consist of two parts. First, one must run Auto Mode before running Director Mode. The dynamic metadata extracted and the corresponding tone mapping curve in Auto Mode shall be used as a reference in Director Mode. If the colorist does not make any adjustments, the Director Mode output shall be equivalent to Auto Mode output. Alternatively, if the colorist restores the default settings, the dynamic metadata and tone mapping curves shall be what they were in Auto Mode. After the adjustment, the dynamic metadata adjusted by the colorist replaces the dynamic metadata in Auto Mode and is written into the output file. Second, the number of controls, the adjustment range and the adjustment function shall comply with HDR Vivid specifications. HDR Vivid recommends 10 sliding controls, each with an adjustment range between -1.0 and 1.0, and a default value of 0, same as in Auto Mode. When the values of all controls are 0, the dynamic metadata and tone mapping curves in Director Mode shall be the same as those in Auto Mode.

12.3.2.3 Curve Anchor Point

The HDR Vivid curve is determined by six anchor points. It is recommended that the six anchor points be shown in the UI. The controls in Director Mode can directly adjust the locations of these six anchor points, and the HDR Vivid curve is calculated and the corresponding dynamic metadata is generated according to the adjustment made on the controls. The x values of the six anchor points in ascending order are TH1, TH2, TH3, TH1_HIGH, TH2_HIGH, and TH3_HIGH, as defined in the HDR Vivid dynamic metadata. The y coordinates are the y values of the HDR Vivid curve on these six locations, and are expressed as TH1_Y, TH2_Y, TH3_Y, TH1_HIGH_Y, TH2_HIGH_Y, and TH3_HIGH_Y. Auto Mode is the basis for Director Mode. The anchor values in Auto Mode are TH1_ref, TH2_ref, TH3_ref, TH1_HIGH_ref, TH2_HIGH_ref, TH3_HIGH_ref, TH1_Y_ref, TH2_Y_ref, TH3_Y_ref, TH1_HIGH_Y_ref, TH2_HIGH_Y_ref, TH3_HIGH_Y_ref.

The curve adjustment algorithm is responsible for mapping the control values to the locations of the anchor points, and calculating the curve parameters and dynamic metadata. The locations of the six anchors are interdependent, and the curve adjustment algorithm strives to coordinate them to generate curves that conform to the HDR Vivid standards. However, in some extreme cases, if by adjusting the six anchor points, no curve that meets the HDR Vivid standards can be generated, the algorithm will return fitting failed. In this case, the controls, curve, and dynamic metadata shall return to their last normal states.

12.3.2.4 Controls

The controls functions and the corresponding anchor locations and dynamic metadata are described in Table 16.

Table 16 Control Description

No.	Control	Description
1	Dark region offset	The value of TH3 is adjusted, which corresponds to TH3 defined in the dynamic metadata. When the value of the control is 0, the value of TH3 is equal to the value of TH3_ref in Auto Mode. When the value of the control is -1, TH3 takes the minimum value of the adjustment range. When the value of the control is +1, TH3 takes the maximum value of the adjustment range. [TH2_ref, 0.45] is the recommended adjustment range. The region whose value is smaller than TH3 is defined as dark region. The three anchor points TH1, TH2, and TH3 determine the cubic spline of the dark region. Adjusting the value of TH3 changes the coverage of the cubic spline of the dark region. In addition, the main curve is determined by two anchor points: TH3 and TH1_HIGH. Therefore, when the value of TH3 is changed, parameters of the main curve shall be re-calculated, including m_p, m_a, m_b, m_m, m_n, K1, K2, and K3 defined in the dynamic metadata.
2	Dark region brightness	The value of TH1_Y is adjusted, which corresponds to the value of MB defined in the dynamic metadata, that is, the slope of the linear spline at the beginning of the curve, which affects the brightness of the darkest part of the image. When the value of the control is 0, TH1_Y takes the value of MB_ref in Auto Mode. When the value of the control is -1, TH1_Y takes the minimum value of the adjustment range. When the value of the control is +1, TH1_Y takes the maximum value of the adjustment range. [0, 1] is the recommended adjustment range.
3	Dark detail	The value of TH2_Y is adjusted, which corresponds to the strength defined in the dynamic metadata. It affects the shape of the cubic spline in the dark region and thus affects the detail performance of the dark region. When the value of the control is 0, TH2_Y takes the value strength_ref in Auto Mode. When the value of the control is -1, TH2_Y takes the minimum value of the adjustment range. When the value of the control is +1, TH2_Y takes the maximum value of the adjustment range. [-0.5, 0.5] is the recommended adjustment range.
4	Mid tone	The value of TH3_Y is adjusted. TH3_Y is defined as the mid tone. It does not directly correspond to any dynamic meta-parameter, but affects m_p, m_a, m_b, m_m, m_n, K1, K2, and K3 defined in the dynamic metadata. When the value of the control is 0, TH3_Y takes the value TH3_Y_ref in Auto Mode. When the value of the control is -1, TH3_Y takes the minimum value of the adjustment range. When the value of the control is +1, TH3_Y takes the maximum value of the adjustment range. [TH2_Y_ref, TH1_HIGH_Y_ref] is the recommended adjustment range. When the value of TH3_Y is changed, parameters of the main curve shall be re-calculated, including m_p, m_a, m_b, m_m, m_n, K1, K2, and K3 defined in the dynamic metadata. If a HDR Vivid-compliant curve cannot be obtained after TH3_Y is adjusted to a certain location, the coordination mechanism is triggered – that is, a new TH1_HIGH_Y value is searched in the same direction of TH3_Y adjustment until a HDR Vivid-compliant curve can be calculated. If the adjustment fails, the current adjusted location is invalid and the

No.	Control	Description
		system returns to the last valid status and control values.
5	Bright region offset	The value of TH1_HIGH is adjusted, which corresponds to the TH1_HIGH defined in the dynamic metadata. When the value of the control is 0, the value of TH1_HIGH is equal to the value of TH1_HIGH_ref in Auto Mode. When the value of the control is -1, TH1_HIGH takes the minimum value of the adjustment range. When the value of the control is +1, TH1_HIGH takes the maximum value of the adjustment range. [0.48, TH2_HIGH_ref] is the recommended adjustment range. The region whose value is greater than TH1_HIGH is defined as the bright region. The three anchor points TH1_HIGH, TH2_HIGH, and TH3_HIGH determine the cubic spline of the bright region. Adjusting the value of TH1_HIGH changes the coverage of the cubic splines in the bright region. In addition, the main curve is determined by two anchor points: TH3 and TH1_HIGH. Therefore, when the value of TH1_HIGH is changed, parameters of the main curve shall be re-calculated, including m_p, m_a, m_b, m_m, m_n, K1, K2, and K3 defined in the dynamic metadata.
6	Bright region brightness	The value of TH1_HIGH_Y is adjusted. It does not directly correspond to any dynamic meta-parameter, but affects m_p, m_a, m_b, m_m, m_n, K1, K2, and K3 defined in the dynamic metadata. When the value of the control is 0, TH1_HIGH_Y takes the value TH1_HIGH_Y_ref in Auto Mode. When the value of the control is -1, TH1_HIGH_Y takes the minimum value of the adjustment range. When the value of the control is +1, TH1_HIGH_Y takes the maximum value of the adjustment range. [TH3_Y_ref, TH2_HIGH_Y_ref] is the recommended adjustment range. When the value of TH1_HIGH_Y is changed, parameters of the main curve shall be re-calculated, including m_p, m_a, m_b, m_m, m_n, K1, K2, and K3 defined in the dynamic metadata. If a compliant HDR Vivid curve cannot be obtained after TH1_HIGH_Y is adjusted to a certain location, the coordination mechanism is triggered, that is, a new TH3_Y value is searched in the same direction of TH1_HIGH_Y adjustment until a HDR Vivid-compliant curve can be calculated. If the adjustment fails, the current adjusted location is invalid and the system returns to the last valid status and control values.
7	Bright region detail	The value of TH2_HIGH_Y is adjusted, which corresponds to strength_HIGH defined in dynamic metadata. It affects the shape of the cubic spline of the bright region and thus affects the detail performance of the bright region. When the value of the control is 0, strength_HIGH takes the value of strength_HIGH_ref in Auto Mode. When the value of the control is -1, strength_HIGH takes the minimum value of the adjustment range. When the value of the control is +1, strength_HIGH takes the maximum value of the adjustment range. [-0.5, 0.5] is the recommended adjustment range.
8	Highlight	The value of TH3_HIGH is adjusted, which corresponds to the TH3_HIGH defined in the dynamic metadata. When the value of the control is 0, TH3_HIGH takes the value of TH3_HIGH_ref in Auto Mode. When the value of the control is -1, TH3_HIGH takes the minimum value of the adjustment range. When the value of the control is +1, TH3_HIGH takes the maximum

No.	Control	Description
		value of the adjustment range. [TH2_HIGH, max_source] is the recommended adjustment range, with max_source indicating the PQ value of the maximum luminance of the current frame. In Director Mode, the value of TH3_HIGH_Y must be equal to the maximum luminance of the reference display device. Therefore, the region whose value is greater than TH3_HIGH is the overexposed region. This control is used to adjust the range of overexposed region to control the effect of highlight.
9	Global saturation	The saturation of the global luminance range is adjusted, which corresponds to color_saturation_gain[0] defined in the dynamic meta parameter. When the value of the control is 0, it takes the value of color_saturation_gain[0] in Auto Mode. When the value of the control is -1, it takes the minimum value 0. When the value of the control is +1, it takes the maximum value 255.
10	Bright region saturation	The saturation for the bright region is adjusted, which corresponds to color_saturation_gain[1] defined in the dynamic meta parameter. When the value of the control is 0, it takes the value of color_saturation_gain[1] in Auto Mode. When the value of the control is -1, it takes the minimum value 0. When the value of the control is +1, it takes the maximum value 255.

13 Content Protection

HDR Vivid video streams can be encrypted and decrypted by regular digital copyright protection systems. Because the associated metadata cannot be restored or reversely derived from video streams if encrypted, it is recommended not to apply additional content protection to HDR metadata, which increases system complexity.