



UHD World Association

世界超高清视频产业联盟



# Coding of UHD video and audio broadcasting system for “Bai Cheng Qian Ping” : System

(Version NO. 1.0)

Release Time

2023-08-22

UHD World Association (UWA)

T/UWA 012.2-2023



## Contents

<b>1 Scope .....</b>	<b>1</b>
<b>2 Normative References .....</b>	<b>1</b>
<b>3 Terms and Definitions .....</b>	<b>1</b>
<b>4 Abbreviations .....</b>	<b>1</b>
<b>5 Requirements for multiplex transmission of coded audio and video .....</b>	<b>2</b>
<b>6 IP transmission requirements for transport streams .....</b>	<b>7</b>
<b>APPENDIX A Independent audio transmission (Informative appendix) .....</b>	<b>8</b>



# Coding of UHD video and audio broadcasting system for “Bai Cheng Qian Ping” : system

## 1 Scope

This document specifies the requirements for multiplex transmission of video basic stream and audio basic stream encoded in the UHD video and audio transmission system for "Bai Cheng Qian Ping" in the transport stream defined in GB/T 17975.1-2010, and also specifies the transmission requirements of the transport stream in the IP network.

This document is applicable to the multiplexing and transmission of video and audio encoding stream in the UHD video and audio transmission system for "Bai Cheng Qian Ping".

## 2 Normative References

The contents in the following documents, through normative references in the text, constitute indispensable provisions of this document. Among them, the dated reference documents are only applicable to the version corresponding to that date; For undated references, the latest version (including all amendments) is applicable to this document.

ISO/IEC 13818-1:2019 Information technology Generic coding of moving pictures and associated audio information Part 1: Systems

ETSI TS 102 034 Transport of MPEG-2 TS Based DVB Services over IP Based Networks

GB/T 17191.2-1997 Information Technology - Coding of Moving Pictures and Associated Sound on Digital Storage Media with a Data Transmission Rate of 1.5Mbit/s - Part 2: Video

GB/T 17975.1-2010 Information Technology - Generic Coding of Moving Images and Associated Audio Information - Part 1: Systems

GB/T 17975.2-2000 Information Technology - Generic Coding of Moving Pictures and Associated Audio Signals - Part 2: Video

GB/T 20090.2-2006 Information Technology Advanced Audio and Video Coding Part 2: Video

GB/T 33475.2-2016 Information Technology Efficient Multimedia Coding Part 2: Video

T/AI 109.2 Information technology - Intelligent media coding - Part 2: Video

T/UWA 009.1 3D Sound Technical Specification Part 1: Coding, Distribution and Presentation

## 3 Terms and Definitions

The following terms and definitions are applicable to this document.

### 3.1 library picture

Images in the library stream specified in T/AI 109.2 can be referenced by images in other bitstreams.

### 3.2 library stream

Bit stream containing knowledge image specified in T/AI 109.2

### 3.3 Main bit stream

Bit stream specified in T/AI 109.2 that can be decoded by referring to the library picture provided by information other than this bit stream.

### 3.4 Bai Cheng Qian Ping

'Bai Cheng Qian Ping' means 'a hundred cities and a thousand large screens', which is a public promotion project, refers to the 8K UHD HDR image and 3D audio played on over a thousand large screens at commercial streets in more than hundred major cities.

## 4 Abbreviations

The following abbreviations are applicable to this document.

<b>AAC</b>	Advanced Audio Coding
<b>AVS2</b>	Information Technology High Efficiency Audio Video Coding Part 2: Video
<b>AVS3</b>	Information Technology Intelligent Media Coding Part 2: Video
<b>HLS</b>	HTTP Live Streaming
<b>RTMP</b>	Real Time Messaging Protocol
<b>RTP</b>	Real time Transport Protocol

**Bslbf** bit string, i.e. binary bit string, with the left bit first. (Bit string, left bit first, where "left" is the order in which bit strings are written in the specification.)

**Uimsbf** Unsigned integer, most significant bit first

**UDP** User Datagram Protocol

**UTC** Coordinated Universal Time

**WebRTC** network real-time communication

## 5 Requirements for multiplex transmission of coded audio and video

### 5.1 General requirements

T/AI 109.2 benchmark 10 position, 10.0.60 level or GB/T 33475.2-2016 benchmark 10 position, 8.0.60 level shall be adopted for video coding.

The video basic stream and audio basic stream of UHD video and audio coding should be transmitted in the transport stream defined in GB/T 17975.1-2010.

### 5.2 Stream\_Id requirements

On the basis of "Table 2-18 stream\_id assignment" in GB/T 17975.1-2010, the stream code with stream\_id 1110 xxxx is specified as "GB/T 17975.2 or GB/T 20090.2 or GB/T 17191.2 or GB/T 33475.2, and the video stream number is xxxx". The stream code with stream\_id 1111 1101 is specified as "T/AI 109.2 video stream"; the stream code with stream\_id 1101 1101 is specified as "T/UWA 009.1-2022 audio stream".

Audio stream\_id refers to ISO/IEC 13818-1:2019.

### 5.3 Stream\_Type requirements

On the basis of "Table 2-29 Stream Type Designation" in GB/T 17975.1-2010, the stream\_type value which is supplemented in the stream of GB/T 33475.2 video is 0xD2, the stream\_type value of T/AI109.2 video is 0xD4, the stream\_type value of the T/UWA 009.1 audio stream is 0xD5.

Audio stream\_Type refers to ISO/IEC 13818-1:2019.

### 5.4 AVS3 video descriptor requirements

#### 5.4.1 AVS3 video descriptor syntax

AVS3 video description is added on the basis of Section 2.6 of GB/T 17975.1-2010. AVS3 video descriptor syntax should meet the requirements of Table 1. On the basis of "Table 50 Program and Program Element Descriptors" in GB/T 17975.1-2010, add the tag value 62, which should be the tag described by AVS3 video.

Table 1 AVS3 video descriptor syntax

Grammar	Number of digits	Mnemonic
AVS3_Video_Descriptor(){		
Descriptor_Tag	8	Uimsbf
Descriptor_Length	8	Uimsbf
Profile_Id	8	Uimsbf
Level_Id	8	Uimsbf
Multiple_Frame_Rate_Flag	1	Bslbf
Frame_Rate_Code	4	Uimsbf
Sample_Precision	3	Uimsbf
Chroma_Format	2	Uimsbf
Temporary_Id_Flag	1	Bslbf
Td_Mode_Flag	1	Bslbf
Library_Stream_Flag	1	Uimsbf
Reserved	3	Bslbf

Grammar	Number of digits	Mnemonic
Colour_Primitives	8	Uimsbf
Transfer_Characteristics	8	Uimsbf
Matrix_Coefficients	8	Uimsbf
If (! Library_stream_flag){		
Num_Ref_Library_Stream	7	Uimsbf
Id_Type_Flag	1	Bslbf
For (i=0; i<num_ref_library_stream; i++){		
If (id_type_flag)		
Ref_Library_Stream_PEID [i]	13	Uimsbf
Else{		
Ref_Library_Stream_Id [i]	8	Uimsbf
Reserved	5	Bslbf
}		
Reserved	3	Bslbf
}		
}		
}		

#### 5.4.2 Semantics of AVS3 video description field

##### **profile\_id**

This field is 8 bits. Indicates the profile of the video bitstream. This field is the same as the profile\_id in T/AI109.2 video bitstream.

##### **level\_id**

This field is 8 bits. Indicates the level of the video bitstream. This field is the same as the level\_id in T/AI109.2 video bit stream.

##### **multiple\_frame\_rate\_flag**

This field is 1 bit. Setting '1' means that there may be multiple frame rates in the video stream, and setting '0' means that there is only a single frame rate.

##### **frame\_rate\_code**

This field is 4-bit, which is the same as the frame\_rate\_code in T/AI109.2 video bit stream. When the multiple\_frame\_rate\_flag is set to '1', a specific frame rate means that some other frame rates are allowed in the video stream. The frame rate code shall meet the requirements of Table 2.

Table 2 **Frame rate code**

Encoding rate	Simultaneously allowed rate
23.976	
24.0	23.976
25.0	
29.97	23.976
30.0	23.976 24.0 29.97
50.0	25.0
59.94	23.976 29.97

60.0	23.976 24.0 29.97 30.0 59.94
100.0	50.0
119.88	59.94
120.0	59.94 60.0 119.88

**sample\_precision**

This field is 3 bits. Specify the accuracy of luminance and chrominance samples. This field is the same as the `sample_precision` in T/AI109.2 video bit stream.

**chroma\_format**

This field is 2 bits. Specifies the format of the chromaticity component. This field is the same as the `chroma_format` in T/AI109.2 video bitstream.

**temporal\_id\_flag**

This field is 1 bit. Indicates whether the video stream is allowed to use the time layer identifier. This field is the same as the `temporal_id_enable_flag` in the T/AI109.2 video stream.

**td\_mode\_flag**

This field is 1 bit. Indicates that the video stream is a monocular video stream or a multi view video stream. This field is the same as `td_mode_flag` in T/AI109.2 video stream.

**colour\_primaries**

This field is 8 bits. Describes the chromaticity coordinates of the three primary colors of the source image in the video stream. This field is the same as the `colour_primaries` in T/AI109.2 video stream.

**transfer\_characteristics**

This field is 8 bits. Explain the photoelectric conversion characteristics of the source image in the video stream. This field is the same as the `transfer_characteristics` in the T/AI109.2 video stream.

**matrix\_coefficients**

This field is 8 bits. Explain the conversion matrix used when converting red, green and blue primary colors to brightness and chroma signals. This field is the same as the `matrix_coefficients` in T/AI109.2 video stream.

**library\_stream\_flag**

This field is 1 bit. Indicates whether the basic stream corresponding to the descriptor in the program is a library bit stream. A value of '1' indicates that the basic stream corresponding to the descriptor is a library stream; A value of '0' indicates that the basic stream corresponding to this descriptor is the main bit stream.

**num\_ref\_library\_stream**

This field is 7 bits. When the basic stream corresponding to the descriptor is a main bit stream, the number of library bitstreams that the main bit stream depends on is defined.

**id\_type\_flag**

This field is 1 bit. Indicates the index type of the library bitstream on which the main bit stream depends when the basic stream corresponding to the descriptor is a main bit stream. The value of this field is '1', which means that the index of the dependent library bit stream uses the PEID of the transport stream packet where the library bit stream is located; The value of this field is '0', which means that the index of the dependent library bit stream uses the stream of the PES packet where the library bit stream is located. When the descriptor is included in the program stream, the field value should be '0'.

**ref\_library\_stream\_PEID[i]**

This field is 13 bits. It is defined that when the basic stream corresponding to this descriptor is a main bit stream, the field is the PEIT\_indicator of the i-th library bit stream that this master bit stream depends on in the transport stream packet.

**ref\_library\_stream\_id[i]**

This field is 8 bits. It is defined that if the basic stream corresponding to this descriptor is a main bitstream, the field is the stream\_id of the i-th library bitstream that this main bitstream depends on in the PES packet header.



## 5.5 AVS2 video descriptor requirements

### 5.5.1 AVS2 video descriptor syntax

AVS2 video descriptor is supplemented on the basis of Section 2.6 of GB/T 17975.1-2010. AVS2 video descriptor syntax shall meet the requirements of Table 3. On the basis of "Table 50 Program and Program Element Descriptors" in GB/T 17975.1-2010, the tag value 64 shall be supplemented. The tag value 64 shall be the tag described by AVS2 video.

Table 3 AVS2 video descriptor syntax

Grammar	digit	Mnemonic
AVS2_Video_Descriptor(){		
Descriptor_Tag	8	Uimsbf
Descriptor_Length	8	Uimsbf
Profile_Id	8	Uimsbf
Level_Id	8	Uimsbf
Extension_Layer_Number	8	Uimsbf
For (i=1; i<=extension_layer_number; i++){		
Layer_Profile_Id [i]	8	Uimsbf
Layer_Level_Id [i]	8	Uimsbf
Layer_Type [i]	8	Uimsbf
Dependent_Layer_Number [i]	8	Uimsbf
For (j=0; j<dependent_layer_number [i]; j++){		
Dependent_Layer_Id [i] [j]	8	Uimsbf
}		
}		
Multiple_Frame_Rate_Flag	1	Bslbf
Frame_Rate_Code	4	Uimsbf
AVS_Still_Present	1	Bslbf
Chroma_Format	2	Uimsbf
Sample_Precision	3	Uimsbf
Reserved	5	Bslbf
Colour_Primitives	8	Uimsbf
Transfer_Characteristics	8	Uimsbf
Matrix_Coefficients	8	Uimsbf
}		

### 5.5.2 Semantics of AVS2 video description fields

#### profile\_id

8-bit field. Indicates the profile of the bitstream. This field is the same as profile\_id in GB/T 33475.2 video stream.

#### level\_id

8-bit field. Indicates the level of the bitstream. This field is the same as the level\_id in the GB/T 33475.2 video stream.

#### extension\_layer\_number

8-bit field. Indicates the number of expansion layers of the bitstream. This field is the same as the extension\_layer\_number in GB/T 33475.2 video stream.

#### layer\_profile\_id

8-bit field. Indicates the profile of the bitstream layer. This field is the same as the layer\_profile\_id in the GB/T 33475.2 video stream.

#### layer\_level\_id

8-bit field. Indicates the level of the bitstream layer. This field is the same as the layer\_level\_id in the GB/T 33475.2 video stream.

#### layer\_type

8-bit field. This field is the same as level\_id in the GB/T 33475.2 video stream.

**dependent\_layer\_number**

8-bit field. Indicates the number of other layers the current layer depends on. This field is the same as the dependent\_layer\_number in the GB/T 33475.2 video stream.

**dependent\_layer\_id**

8-bit field. Indicates the index of the layer on which the current layer depends on. This field is the same as the dependent\_level\_id in the GB/T 33475.2 video stream.

**multiple\_frame\_rate\_flag**

1-bit field. When '1' is set, it indicates that there may be multiple frame rates in the video stream. When '0' is set, it indicates that there is only a single frame rate.

**frame\_rate\_code**

4-bit field. The code field definition is the same as frame\_rate\_code in GB/T 33475.2 video stream. The difference is that a specific frame rate means that some other frame rates are allowed in the video stream when multiple\_frame\_rate\_flag is set to '1'. The frame rate code shall meet the requirements of Table 4.

Table 4 **Frame rate code**

Encoding rate	Simultaneously allowed rate
23.976	
24.0	23.976
25.0	
29.97	23.976
30.0	23.976 24.0 29.97
50.0	25.0
59.94	23.976 29.97
60.0	23.976 24.0 29.97 30.0 59.94

**AVS\_still\_present**

1-bit field. When '1' is set, it means that the video stream only contains static image data; When set to '0', it can contain moving or static image data.

**chroma\_format**

2-bit field. Specifies the format of the chromaticity component. This field is the same as the encoding method of chroma\_format in GB/T 33475.2 video stream.

**sample\_precision**

3-bit field. Specify the accuracy of luminance and chrominance samples. This field is the same as the encoding method of sample\_precision in the GB/T 33475.2 video stream.

**color\_primaries**

8-bit field. Represents the chromaticity coordinates of the three primary colors of the source image in the video stream. This field is the same as the colour\_primaries in the GB/T 33475.2 video stream.

**transfer\_characteristics**

8-bit field. Represents the photoelectric transfer characteristics of the source image in the video stream. This field is the same as the transfer\_characteristics in the GB/T 33475.2 video stream.

**matrix\_coefficients**

8-bit field. Represents the conversion matrix used for converting red, green and blue primary colors into luminance and chrominance signals. This field is the same as the matrix\_coefficients in GB/T 33475.2 video stream.

**5.6 Extended PES Grouping**

Expand the PES packet, and use the private data segment of the PES packet to transmit the encoded time information TimeStamp.

**5.6.1 TimeStamp syntax**

Based on GB/T 17975.1-2010, PES\_Extension\_Flag set to 1, PES\_Private\_Data\_Flag set to 1, PES\_Private\_ The syntax structure of data transmission should meet the requirements of Table 5. The application of this grammatical structure refers to Appendix A.

Table 5 TimeStamp Syntax Structure

Grammar	digit	Mnemonic
TimeStamp(){		
Syncword	12	Uimsbf
Version	2	Uimsbf
Utc_Time_Valid	1	Uimsbf
Reserved	1	Uimsbf
Reserved	64	Uimsbf
Utc_Time	48	Uimsbf
}		

### 5.6.2 Semantics of TimeStamp fields

#### Syncword

Syntax structure synchronization, should be 0xFEE.

#### Version

Indicates the version of the syntax structure, and the value is 1.

#### Reserved

All 1 are reserved in this syntax structure.

#### UTC time effective flag bit utc\_Time\_Valid

Whether to set the flag bit of utc time, accounting for 1 bit, 1 represents utc\_Time has an actual value, 0 represents utc\_Invalid time.

#### World Standard Time utc\_Time

The time when the frame was generated is the number of milliseconds counted from 0:00:00sec on January 1, 1970.

## 6 IP transmission requirements for transport streams

### 6.1 The transport stream is transmitted as the load of UDP

It shall comply with the provisions of Section 7.1.2 of ETSI TS 102 034.

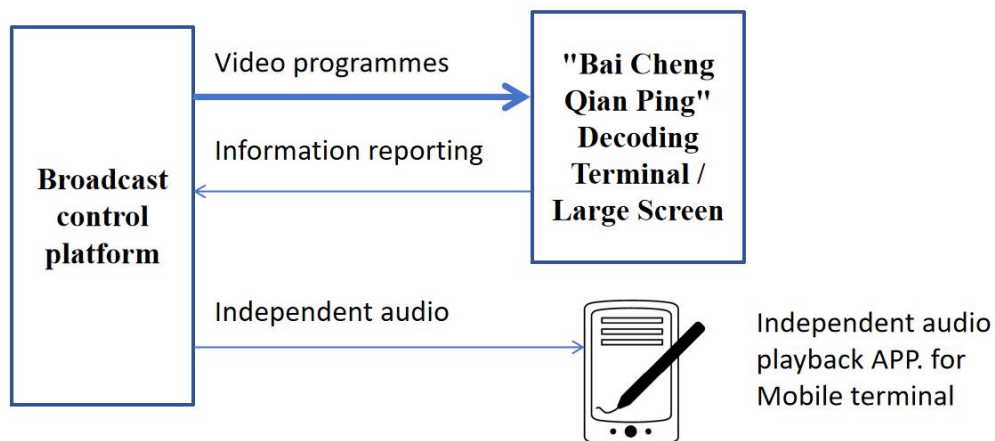
### 6.2 The transport stream is transmitted as the load of RTP

It shall comply with the provisions of Section 7.1.1 of ETSI TS 102 034.

## APPENDIX A Independent audio transmission (Informative appendix)

Some large screens are not suitable for playing sound, but the audience needs to hear synchronous audio when watching programs on large screen. To meet this demand, it is necessary to transmit audio corresponding to the program through independent channel, so that the audience can listen to the audio of the program synchronously with the portable mobile device when watching programs on large screen. See Figure 1 for details. Due to different positions of large screens, different delays of different transmission networks, and different decoding buffer strategies of different types of terminals, the playback delay of decoded video is different; The time stamp needs to be added to the program code stream. See Section 5.6 for the specific syntax structure. The decoding terminal needs to report the decoding time and the corresponding time stamp; The mobile device application that users carry with them needs to obtain the location of the device so that the service platform can find the corresponding large screen and provide appropriate audio streams.

FIGURE 1 Schematic diagram of independent audio transmission for "Bai Cheng Qian Ping"



### Independent audio internet transmission

The audio encoding adopts Audio Vivid. In order to reduce delay and ensure data integrity, it is not recommended to perform secondary transcoding during transmission..

In order to control the independent audio player terminal to play synchronously with the video on the corresponding large screen, it is recommended to embed the UTC timestamp corresponding to the time when the frame is generated in the audio transport stream. The specific timestamp embedding method is same as the video transport stream.

### Independent audio synchronization requirements

During independent audio transmission, the delay between audio and corresponding video on large screen shall be controlled within (- 120,120) ms.

The mobile terminal APP playing independent audio should have the function of synchronous fine tuning.