ENGLISH TRANSLATION

MULTIMEDIA CODING SPECIFICATION FOR DIGITAL BROADCASTING (SECOND GENERATION)

ARIB STANDARD

ARIB STD-B62 Version 1.7 (Fascicle 1)

Version 1.0	July 31,	2014
Version 1.1	March 17,	2015
Version 1.2	July 3,	2015
Version 1.3	December 3,	2015
Version 1.4	March 25,	2016
Version 1.5	July 6,	2016
Version 1.6	September 29,	2016
Version 1.7	July 27,	2017

Association of Radio Industries and Businesses

General Notes to the English Translation of ARIB Standards and Technical Reports

1. Notes on Copyright

- The copyright of this document is ascribed to the Association of Radio Industries and Businesses (ARIB).
- All rights reserved. No part of this document may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, without the prior written permission of ARIB.

2. Notes on English Translation

- ARIB Standards and Technical Reports are usually written in Japanese. This document is a translation into English of the original document for the purpose of convenience of users. If there are any discrepancies in the content, expressions, etc. between the original document and this translated document, the original document shall prevail.
- ARIB Standards and Technical Reports, in the original language, are made publicly available through web posting. The original document of this translation may have been further revised and therefore users are encouraged to check the latest version at an appropriate page under the following URL:
http://www.arib.or.jp/english/index.html.

Foreword

The Association of Radio Industries and Businesses (ARIB) investigates and summarizes the basic technical requirements for various radio systems in the form of "ARIB Standards". These standards are developed with the participation of and through discussions amongst radio equipment manufacturers, telecommunication operators, broadcasting equipment manufacturers, broadcasters and users.

ARIB Standards include "government technical regulations" (mandatory standard) that are set for the purpose of encouraging effective use of frequency and preventing interference with other spectrum users, and "private technical standards" (voluntary standards) that are defined in order to ensure compatibility and adequate quality of radio equipment and broadcasting equipment as well as to offer greater convenience to radio equipment manufacturers, telecommunication operators, broadcasting equipment manufacturers, broadcasters and users.

This ARIB Standard is developed for "MULTIMEDIA CODING SPECIFICATION FOR DIGITAL BROADCASTING (SECOND GENERATION)". In order to ensure fairness and transparency in the defining stage, the standard was set by consensus at the ARIB Standard Assembly with the participation of both domestic and foreign interested parties from radio equipment manufacturers, telecommunication operators, broadcasting equipment manufacturers, broadcasters and users.

ARIB sincerely hopes that this ARIB Standard will be widely used by radio equipment manufacturers, telecommunication operators, broadcasting equipment manufacturers, broadcasters and users.

NOTE:

Although this ARIB Standard contains no specific reference to any Essential Industrial Property Rights relating thereto, the holders of such Essential Industrial Property Rights state to the effect that the rights listed in the Attachment 1 and 2, which are the Industrial Property Rights relating to this standard, are held by the parties also listed therein, and that to the users of this standard, in the case of Attachment 1, such holders shall not assert any rights and shall unconditionally grant a license to practice such Industrial Property Rights contained therein, and in the case of Attachment 2, the holders shall grant, under reasonable terms and conditions, a non-exclusive and non-discriminatory license to practice the Industrial Property Rights contained therein. However, this does not apply to anyone who uses this ARIB Standard and also owns and lays claim to any other Essential Industrial Property Rights of which is covered in whole or part in the contents of the provisions of this ARIB Standard.

Attachment 1 (Selection of Option 1)
(N/A)

Attachment 2 (Selection of Option 2)

PATENT HOLDER	NAME OF PATENT	REGISTRATION NO./ APPLICATION NO.	REMARKS
Sony Corporation	Submitted comprehensive confirmation of patents for ARIB STD-B62 Ver1.0 *1		
Sharp Corporation	Submitted comprehensive confirmation of patents for ARIB STD-B62 Ver1.0*1		

Note) *1: Valid for ARIB STD-B62 Ver1.0 (accepted on July 24, 2014)

TOTAL CONTENTS

VOLUME 1

Data Coding Scheme

Part 1　Reference Model for Multimedia Coding Scheme

Contents

<Blank Page>

Chapter 1: General Terms

1.1 Purpose

This part of the standard defines a reference model for the second generation multimedia coding scheme for advanced broadband satellite digital broadcasting specified as the Japanese standard specifications.

1.2 Scope

This part of the standard defines a reference model for multimedia coding scheme for advanced broadband satellite digital broadcasting.

1.3 References

1.3.1 Normative references

The latest version is used if a specific version is not specified for any of the following documents and recommendations that are listed as normative references below. A specific version is used if it is specified for a normative reference; however, users of this standard are encouraged to investigate the possibility of applying the most recent version of the normative references.

[1] IPTVFJ STD-0010 "Integrated Broadcast-Broadband System Specification"

1.3.2 Informative references

[1] ARIB STD-B10 "Service Information for Digital Broadcasting System"

[2] ARIB STD-B24 "Data Coding and Transmission Specification for Digital Broadcasting"

[3] ARIB STD-B32 "Video Coding, Audio Coding and Multiplexing Specifications for Digital Broadcasting"

[4] ARIB STD-B60 "MMT-based Media Transport Scheme in Digital Broadcasting Systems"

1.4 Terms, definitions and abbreviations

1.4.1 Definitions

This standard defines following terms.

Term	Definition
Logical coordinate	A logical coordinate system for a receiver's decoder model in a presentation process. Each plane has its own logical coordinate.
Display coordinate	A coordinate system for displaying on the display.
Reference model	A model to be referred to as the standard for a system, protocol, receiver, and presentation process etc., in the specifications for the multimedia coding scheme and transmission systems.
α blending	A technique of mixing images whose transparency is controlled by the α value.
Colorimetry	Specifications for color reproduction.
Plane	Display screen to display monomedia and multimedia contents.

Monomedia	Independent media element, such as video, still picture, figure, audio and character, that can be presented using only its own data without referring to other media.
Closed caption	Textual information service of TV broadcasting to provide information related to video content. Closed captions are sometimes super imposed over video content.
Super impose	Subtitling service that is not related to the main video, audio or data, e.g., news flash, program remarks, warning, time signal, etc.
Real time presentation element	Monomedia that has time constraint in decoding and presenting, e.g., video, audio, closed captions, and super impose.

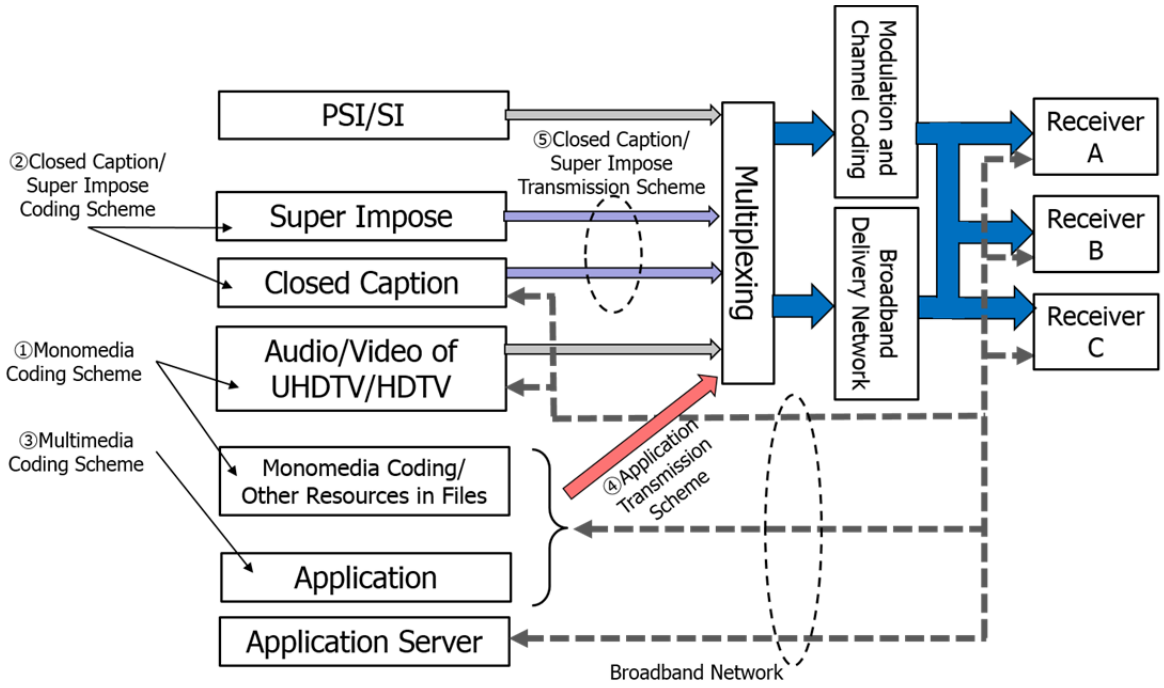
1.4.2 Abbreviations

AAC	Advanced Audio Coding
AIT	Application Information Table
ALS	Audio Lossless Coding
HEVC	High Efficiency Video Coding
HTTP	Hyper Text Transfer Protocol
LCT	Layout Configuration Table
MMT	MPEG Media Transport
MPEG-DASH	MPEG Dynamic Adaptive Streaming over HTTP
NTP	Network Time Protocol
RTP	Real-time Transport Protocol
TMCC	Transmission and Multiplexing Configuration and Control
TTML	Timed Text Markup Language
TTS	Time stamped TS
UHDTV	Ultra High Definition TV

Chapter 2: System

It is required to define some interfaces from sending through receiving data in broadcast chain in order to provide multimedia broadcasting services. Specifications for receivers are also necessary to enable viewers to receive transmitted data and enjoy the service as the sender intended. Based on the perspective described above, this chapter defines the reference model for the entire system that is related to multimedia broadcasting via the advanced broadband satellite digital broadcasting.

This standard specifies the following two types of data transmission protocols, MPEG-2 TS and MMT. Figure 2-1 shows the configuration of the system for the multimedia broadcasting services via the advanced broadband satellite digital broadcasting.



Note) The "broadband delivery network" shown in the figure above is used when distributing multiplexed signals using the MMT. The "broadband network" indicated by the dashed lines represents the acquisition of monomedia content, such as video, audio and still pictures, or multimedia application upon a request from a receiver, or a processing request to a server.

Fig. 2-1: System configuration

Each configuration element of the system shown in Fig. 2-1 is specified as follows:

①Monomedia coding scheme

This scheme defines coding schemes for video, audio, character, figure, etc., that are used for the multimedia services.

②Closed caption/super impose coding scheme

This scheme defines coding schemes related to closed caption and super impose.

③Multimedia coding scheme

This scheme defines coding schemes to create a multimedia application. This scheme is defined in Volume 2 of this standard by taking into account the compatibility with the Hybridcast technical specifications defined by the IPTV Forum Japan.

④Application transmission scheme

This format defines a transmission scheme to transmit a multimedia application and other control signals.

⑤Closed caption/super impose transmission scheme

This scheme defines a transmission methods for closed caption and super impose.

The configuration elements described above are defined in the standard written in Table 2-1.

Table 2-1: Standard that defines the configuration elements of this system

Configuration element	MPEG-2 TS system	MMT system
Monomedia coding scheme	STD-B62 Volume 1 Part 2	
Closed caption/ super impose coding scheme	STD-B62 Volume 1 Part 3	
Multimedia coding scheme	STD-B62 Volume 2	
Application transmission scheme	STD-B24 Volume 3	STD-B60 Chapter 10
Closed caption/super impose transmission scheme	STD-B24 Volume 3	STD-B60 Chapter 9

Chapter 3: Protocol

In the multimedia broadcasting system defined in this standard, video, audio and all other data provided are multiplexed as a service on the broadcast stream by MPEG-2 TS defined in STD-B32, STD-B24 and STD-B10 or by MMT defined in STD-B60. A protocol stack used for each system is shown in Fig. 3-1 and Fig. 3-2. The protocol stack for broadband network that is common to both MPEG-2 TS and MMT systems is shown in Fig. 3-3.

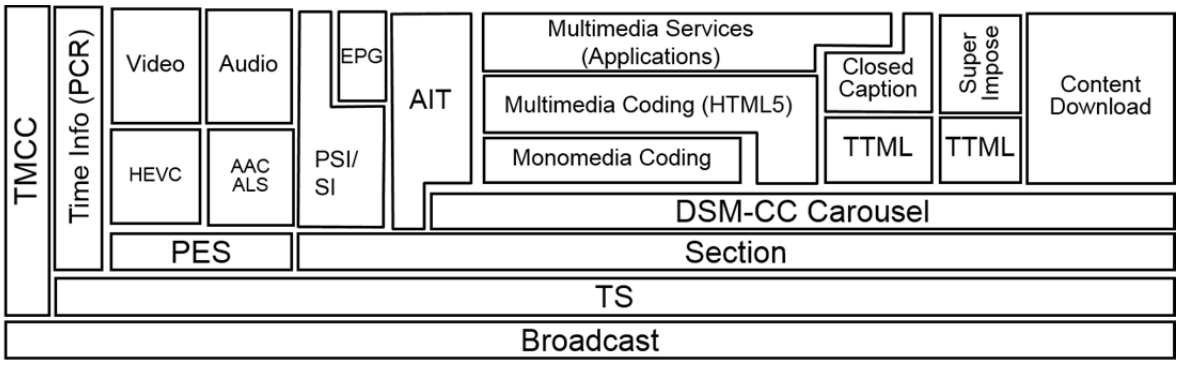


Fig. 3-1: Protocol stack for the MPEG-2 TS system

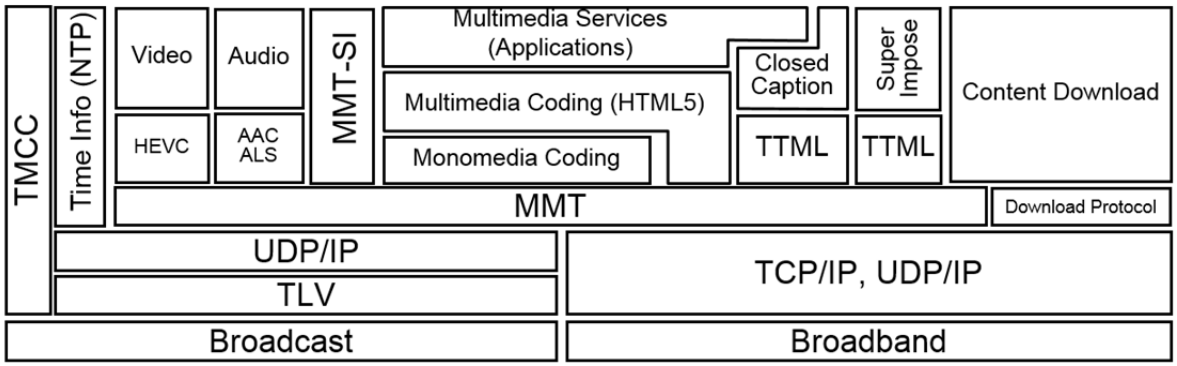


Fig. 3-2: Protocol stack for the MMT system

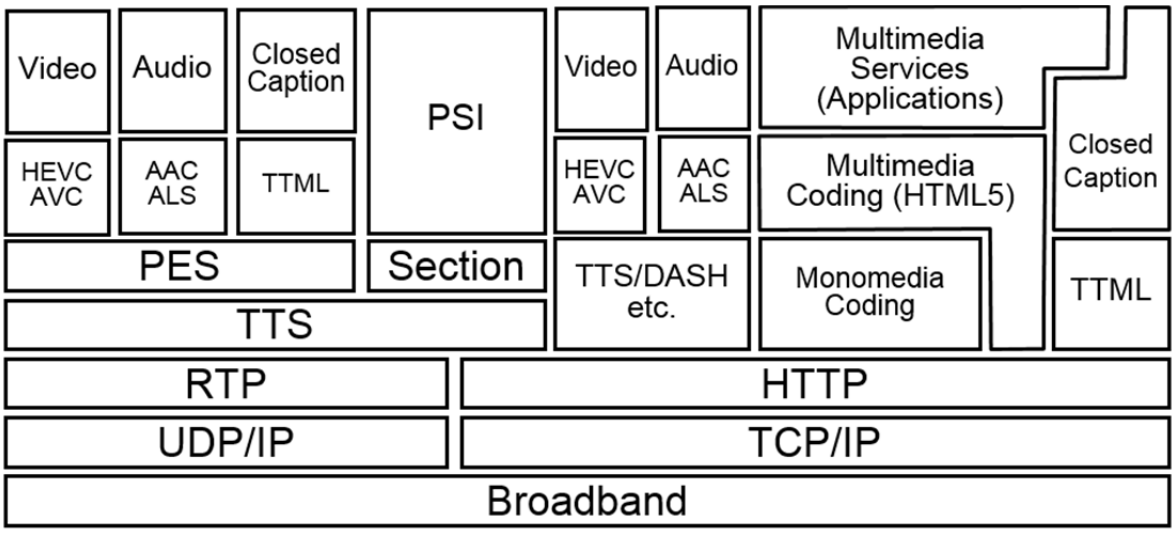


Fig. 3-3: Protocol stack for broadband network that is common to the MPEG-2 TS and MMT systems

The protocols used for broadband network are not limited to the protocols shown in the figures. Actual receiver implementation may use additional protocols.

Chapter 4: Application Model for Multimedia Application

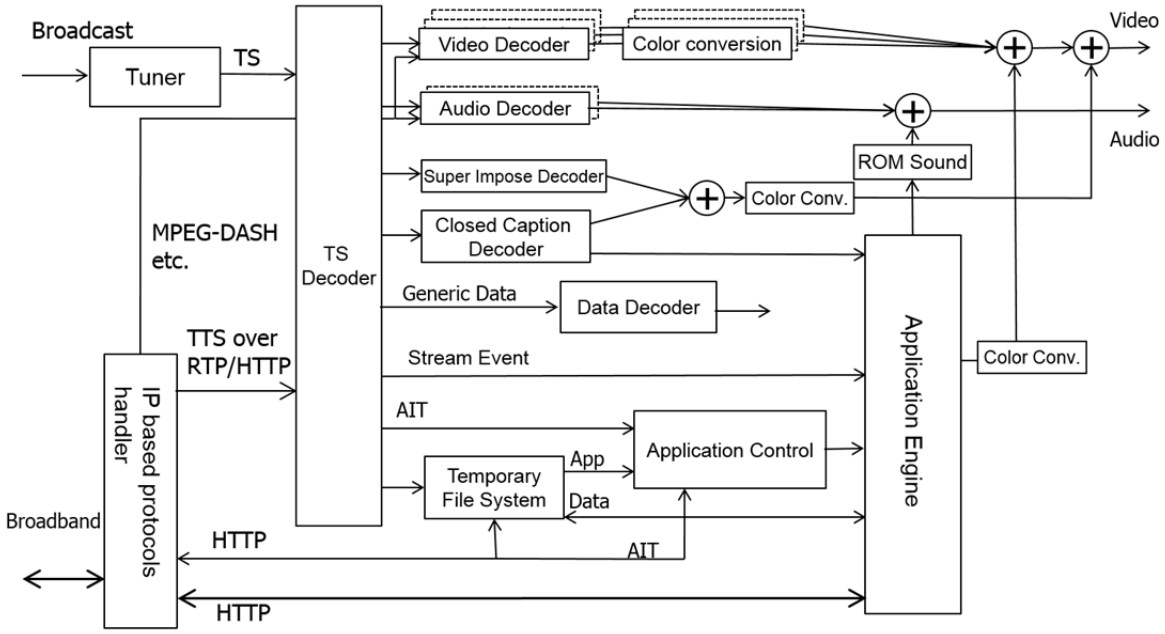
The application models that are defined in IPTVFJ STD-0010 "Integrated Broadcast-Broadband System Specification" developed by the IPTV Forum Japan is used.

Chapter 5: Receiver

This chapter defines the basic functions of receivers so as to receive/provide multimedia services regardless of the implementation of the receiver. The receivers, which can receive/provide multimedia services, must have functions to receive/provide/store/communicate with the multimedia broadcasting service in addition to the functions to view normal TV programs. These functions enable viewers to enjoy multimedia services. In addition, the receivers that can receive/provide multimedia services based on this standard must be designed to offer maximum compatibility with the receiver functions that satisfy the Hybridcast technical specification developed by the IPTV Forum Japan.

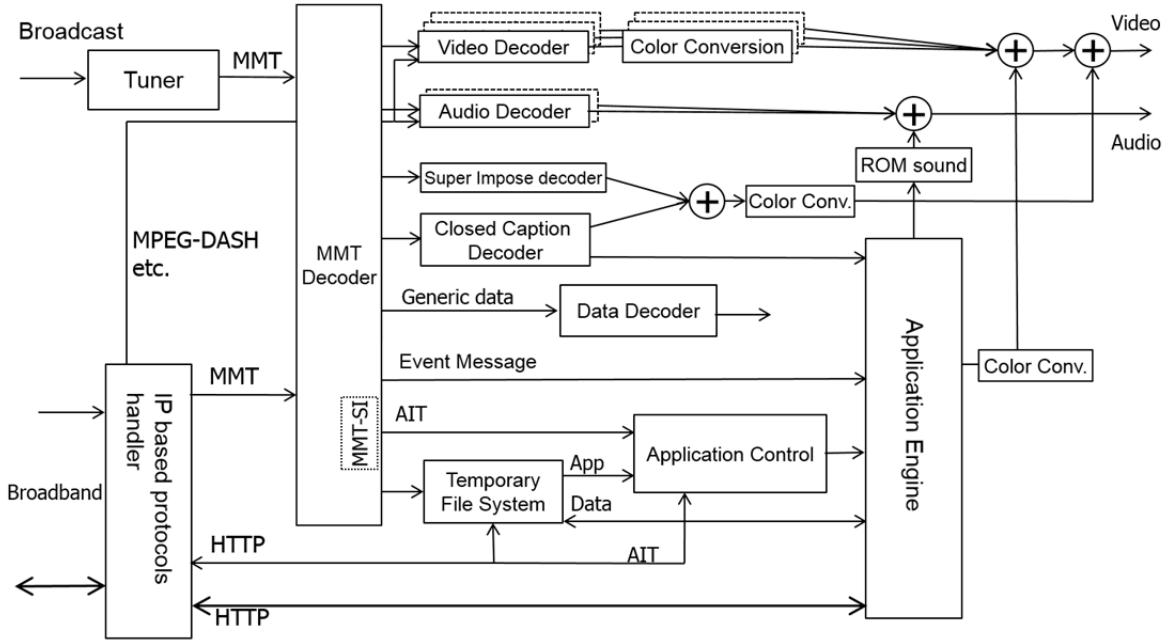
5.1 Functions for Broadcast and Broadband

The receivers that can receive/provide multimedia services based on this standard need to have the broadcasting receive function and broadband access function based on at least one type of protocol stack defined in Chapter 3. Figure 5-1 shows the example of a receiver structure that supports the MPEG-2 TS and Fig. 5-2 shows the example of a receiver structure that supports the MMT. The protocol names along side of specific paths in the figures are supposed to be major protocols for those paths.



DR conversion: Dynamic Range conversion

Fig. 5-1: Example structure of a receiver that supports the MPEG-2 TS



DR conversion: Dynamic Range conversion

Fig. 5-2: Example structure of a receiver that supports the MMT

The broadcast signals sent out from a broadcasting station are input to the TS decoder or the MMT decoder. As for the receivers that support the MMT, the MMT decoder processes broadcast signals or a part of them that are delivered over a broadband network. The TS decoder or MMT decoder hands the real time presentation elements data to the video, audio, closed caption and super impose decoders in accordance with control signals included in the received signals.

The multimedia application and files of elements of the multimedia application are decoded as files by the TS decoder or the MMT decoder. Such files are stored in a cache (temporary file system). A browser acquires the files from a cache (temporary file system) or a server via a broadband network to execute applications, in accordance with the application control information that is included in the broadcast signals or obtained broadband network using HTTP.

Generic data can be used for a player that presents data other than video, audio and closed captions or can be used for streaming to an application.

5.2 Presentation Function

In order to execute and display multimedia services on a receiver as intended, it is necessary to define the display and presentation functions that must be implemented as minimum requirements for receivers. Fig. 5-3 shows the logical plane structure of a receiver to satisfy requirements for presentation..

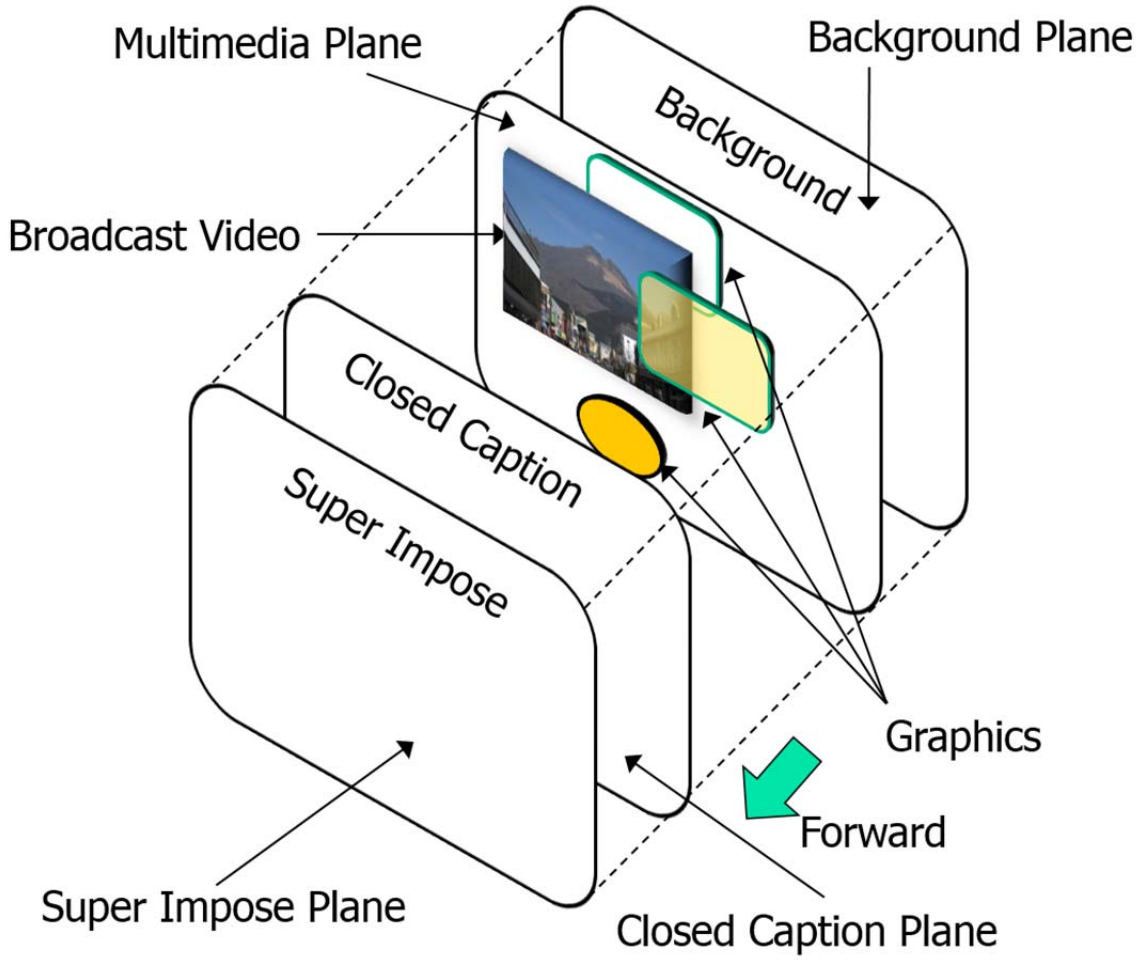


Fig. 5-3: Logical plane structure of a receiver

Each plane displays the following.

(1) Super impose plane

This plane displays super impose. This plane is in the most foreground of other display planes. The output from the super impose decoder shown in Fig. 5-1 and Fig. 5-2 is rendered on this plane.

(2) Closed caption plane

This plane displays closed captions. This place is located behind the super impose plane and in front of the multimedia plane. The output from the closed caption decoder shown in Fig. 5-1 and Fig. 5-2 is rendered on this plane.

(3) Multimedia plane

This plane displays composite images using a single monomedia, such as video, or the multimedia application. Multimedia application, which is coded by the multimedia coding scheme defined in volume 2 of this standard, controls when superimposing figures and video on this plane. This plane is located behind the closed caption plane and in front of the background plane. Use of multiple planes is allowed in accordance with the number of the video decorders shown in Fig. 5-1 or Fig. 5-2. However, the output of the browser shown in Fig. 5-1 or Fig. 5-2

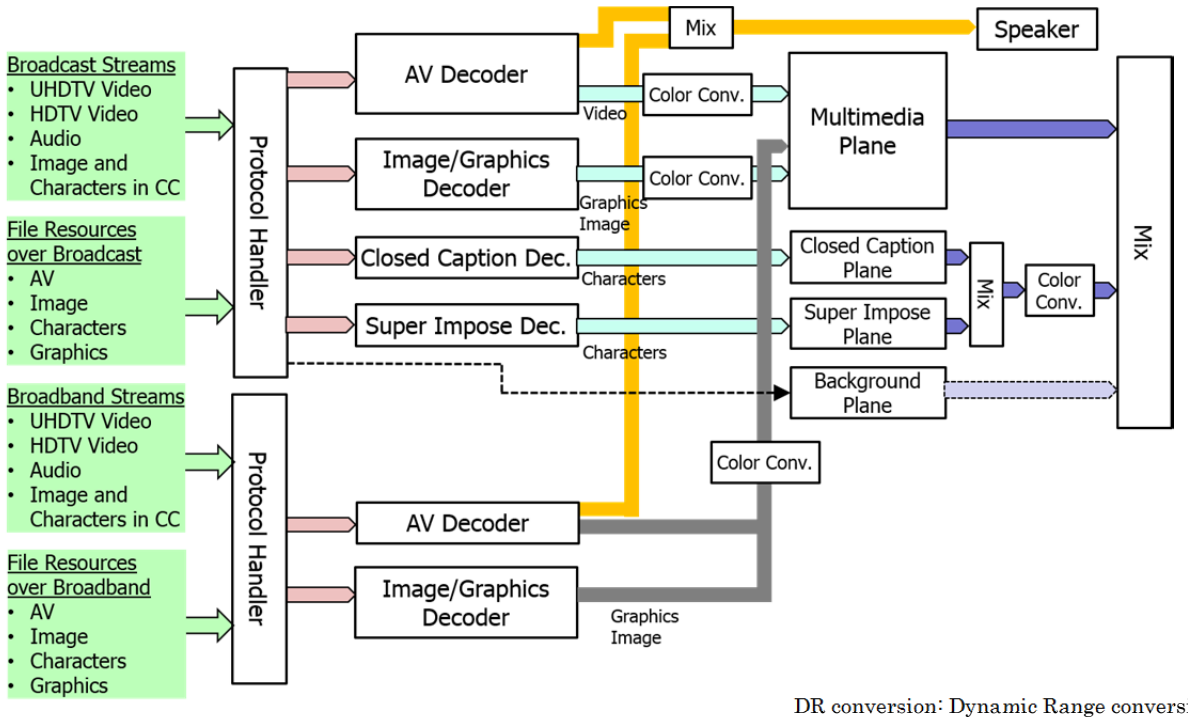
that executes the multimedia application is rendered only on the most foreground multimedia plane.

(4) Background plane

This plane displays the background color. This plane is located rearmost. When all planes located in front of this plane are transparent, then the entire surface of this plane is painted in a specified color and displayed. Or if there are blank areas on the planes located in front of this plane, the areas that match the blank areas are painted in a specified color and displayed.

5.3 Decoding Process and Display

Figure 5-4 shows decoding and presentation processes in a receiver by presenting the relationship with data flow and the output.



DR conversion: Dynamic Range conversion

Fig. 5-4: Receiver's decoder model based on the data flow

(1) Transmission data decoding process

The data transmitted to a receiver is decoded through a transport process. In this process, the data transmitted via stream (UHDTV or HDTV video, audio, closed captions, etc.) and file data (still picture, character, figure, etc.) are separated from the received signals and handed to the relevant decoder to decode them. The transmitted data is decoded based on the protocol stack for the data transmitted over broadcast channel (when a receiver supports the MMT system, the data transmitted over a broadband network is included) or the data transmitted via a broadband network (mainly pull-type data, including UHTV or HDTV video and audio transmitted in stream, whose acquisition is initiated by a receiver).

(2) Monomedia decoding process

Monomedia data, such as video, audio and still pictures, are delivered to a respective monomedia decoders through the transport process. The monomedia decoder converts the color-gamut or dynamic range when outputting the data to the multimedia plane if necessary and adjusts to match the color space or dynamic range of each monomedia.

(3) Presentation process

The decoded monomedia data that is outputted to each plane is composited with control, such as size and position. Before composition, it is necessary to adjust the data or dynamic range to match the color space for a respective presentation element. To align with this behavior, the data output to the multimedia plane is adjusted when outputting to the plane. On the other hand, since the same color space is used for closed captions and super impose, the color-gamut or dynamic range conversion is performed if necessary when composing the data displayed on the closed caption plane and super impose plane, and then composing them with multimedia plane. After that, the color space or dynamic range for closed captions and super impose and that for the multimedia plane are adjusted.

When a receiver supports the MMT, the background plane is painted in a color specified in the Layout Configuration Table (LCT) in the received MMT signal.

Chapter 6: Presentation Process

This chapter defines models related to the presentation process.

6.1 Coordinates and Composition

6.1.1 Configuration of each plane

The super impose plane, closed caption plane, multimedia plane and background plane that are logical planes to which decoded data are outputted from the monomedia decoder have the logical plane configuration shown in Table 6-1.

Table 6-1: Configuration of each plane

Plane	Specified range
Super impose plane	7680 x 4320　RGB with an 8-bit in each*1 256 levels of α blending
Closed caption plane	7680 x 4320　RGB with an 8-bit in each*1 256 levels of α blending
Multimedia plane	7680 x 4320　RGB with an 8/10-bit in each*2 The graphics elements in the plane supports 256 levels of α blending
Background plane	7680 x 4320　RGB with an 8-bit in each*1

*1 A color space represented using RGB with an 8 out of 10bits in each in the Rec. ITU-R BT.2020. LSB and the second LSB are neglected.

*2 In case of use of the decoders which output using color space define in Rec. ITU-R BT.2020, the output signal will be used directly. Otherwise, the same color space is represented using RGB with an 8-bit (LSB and the second LSB are neglected).

The coordinates of each plane represent the logical coordinates in each plane. The coordinates of the upper left corner are (0, 0) and the coordinates of the lower right corner are (7679, 4319). If a resolution is specified in the monomedia signals, data is rendered in a plane by converting the resolution into the coordinates above. For instance, if 3840-by-2160 resolution is specified, the coordinates (3839, 2159) that correspond to the bottom right corner are converted to (7679, 4319).

6.1.2 Plane and Layout Configuration

When a receiver supports MMT, each plane performs the following operations on the Layout Configuration Table included in the received MMT signal.

(1) Multimedia plane

If an area for multimedia application or video asset is specified by the LCT, size and position of the objects belonging to corresponding asset in this plane is changed in accordance with the LCT. In this case, the logical coordinates in the plane do not change. If a video asset is specified as the asset for a specified area while executing the multimedia application, the execution of the multimedia application and the presentation of the multimedia application on the applicable plane continue. If the multimedia application asset is specified, the asset, including the other assets referenced from the multimedia application, is displayed in accordance with a specified area. Figure 6-1 shows this example. As (a) in the figure below, if a layout setting is not specified, the physical coordinates match the logical coordinates in the multimedia plane. However, as (b) in the figure below, if an area other than the 1/4 of the right area and 1/4 of the bottom area is specified and if the multimedia application asset is assigned to the area, the coordinates of the lower right corner in the plane (7679, 4319) are automatically converted to the physical coordinates (5759, 3239) to display the multimedia application by reducing the height and width to 3/4 size while keeping the same coordinates for the multimedia application..

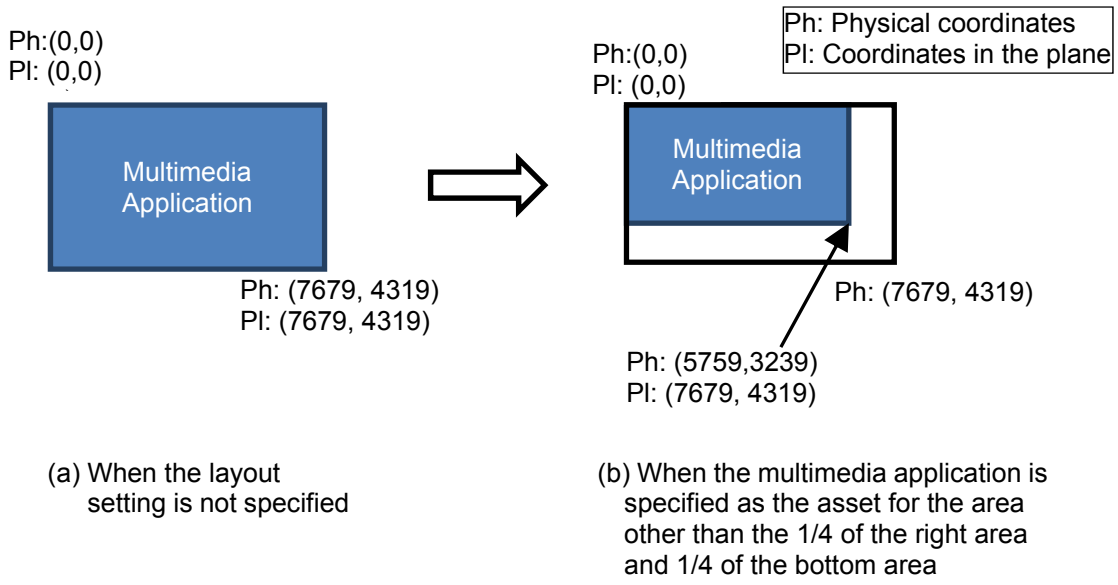


Fig. 6-1: Operation example of multimedia plane on the LCT in the MMT signal

(2) Super impose and closed caption planes

If the areas for the super impose and closed caption planes are specified in the LCT, the sizes and positions of these planes are maintained, but their presentation is clipped in the specified area. Figure 6-2 shows this example. As (a) in the figure below, if a layout setting is not specified, then the three-lines closed caption is displayed at the bottom of the screen. In contrast, as shown in (b) in the figure below, the closed caption area is specified to the lower 1/4 area by LCT and if the closed caption asset is assigned there, the presentation is not made in the upper 3/4 area. Rendering in the area by the logical coordinates of the closed caption plane, ranging from (0,0) to (7679, 3239) will not be presented. As a result, the last line of the closed caption only is displayed.

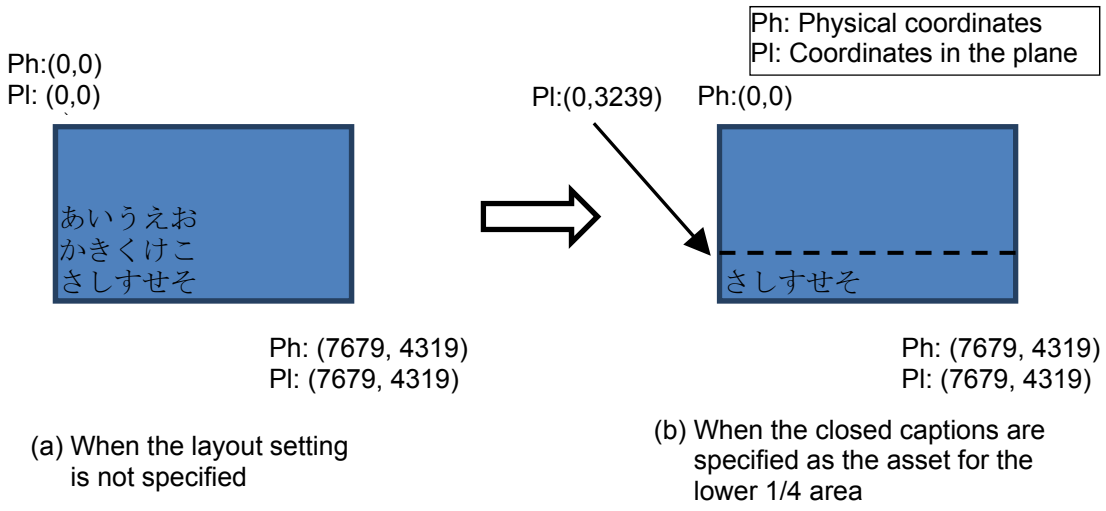


Fig. 6-2: Operation example of closed caption plane on the LCT in the MMT signal

6.2 Colorimetry

6.2.1 Harmonizatoin of color space

UHDTV video uses color spaces defined in the Rec. ITU-R BT.2020. The multimedia application may present a HDTV video, which uses a color space defined in the Rec. ITU-R BT.709. The multimedia application may also use graphics at the same time. Therefore, color-gamut must be harmonize to composite each presentation element.

6.2.2 Conversion from a color space defined in the Rec. ITU-R BT.709 to a color space defined in the Rec. ITU-R BT.2020

Figure 6-3 shows the process of converting a color-gamut defined in the Rec. ITU-R BT.709 to a color-gamut defined in the Rec. ITU-R BT.2020. In the process, the DQ, M1, M3 and Q levels represent the conversion between $Y'C'_BC'_R$ and $R'G'B'$ and they are used according to the implementation of a receiver.

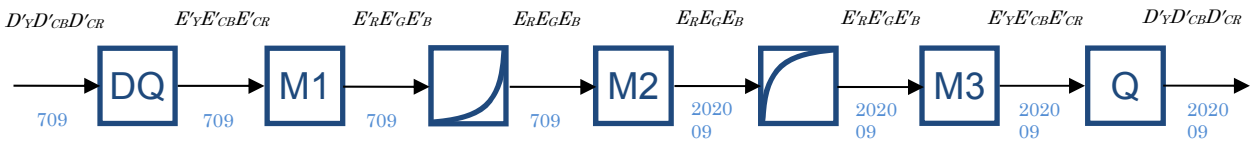


Fig. 6-3: Conversion process from a color space defined in the Rec. ITU-R BT.709 to a color space defined in the Rec. ITU-R BT.2020

Figure 6-4 shows the conversion equation for each stage.

DQ Inverse quantization (Nbit) from $D'_Y D'_{CB} D'_{CR}$ (Recommendation BT.709) to $E'_Y E'_{CB} E'_{CR}$ (Recommendation BT.709)

$$E'_Y = \left(D'_Y / 2^{N-8} - 16\right) / 219$$

$$E'_{CB} = \left(D'_{CB} / 2^{N-8} - 128\right) / 224$$

$$E'_{CR} = \left(D'_{CR} / 2^{N-8} - 128\right) / 224$$

M1 Conversion from $E'_Y E'_{CB} E'_{CR}$ (Recommendation BT.709) to $E'_R E'_G E'_B$ (Recommendation BT.709)

$$\begin{bmatrix} E'_R \\ E'_G \\ E'_B \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dfrac{1.5748}{0.7152} \\ 1 & -\dfrac{0.0722 \times 1.8556}{0.7152} & -\dfrac{0.2126 \times 1.5748}{0.7152} \\ 1 & 1.8556 & 0 \end{bmatrix} \begin{bmatrix} E'_Y \\ E'_{CB} \\ E'_{CR} \end{bmatrix}$$

EOTF from $E'_R E'_G E'_B$ (Recommendation BT.709) to $E_R E_G E_B$ (Recommendation BT.709)

$$E = (E')^2 \quad , \quad 0 \le E' \le 1$$

M2 Conversion from $E_R E_G E_B$ (Recommendation BT.709) to $E_R E_G E_B$ (Recommendation BT.2020)

$$\begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{2020} = \begin{bmatrix} 0.6370 & 0.1446 & 0.1689 \\ 0.2627 & 0.6780 & 0.0593 \\ 0 & 0.0281 & 1.0610 \end{bmatrix}^{-1} \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{709}$$

$$= \begin{bmatrix} 1.7167 & -0.3557 & -0.2534 \\ -0.6667 & 1.6165 & 0.0158 \\ 0.0176 & -0.0428 & 0.9421 \end{bmatrix} \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{709}$$

$$= \begin{bmatrix} 0.6274 & 0.3293 & 0.0433 \\ 0.0691 & 0.9195 & 0.0114 \\ 0.0164 & 0.0880 & 0.8956 \end{bmatrix} \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{709}$$

OETF (gamma pre-correction) from $E_R E_G E_B$ (Recommendation BT.2020) to $E'_R E'_G E'_B$ (Recommendation BT.2020)

$$E' = E^{1/2} \quad , \quad 0 \le E \le 1$$

M3 Conversion from $E'_R E'_G E'_B$ (Recommendation BT.2020) to $E'_Y E'_{CB} E'_{CR}$ (Recommendation BT.2020)

$$\begin{bmatrix} E'_Y \\ E'_{CB} \\ E'_{CR} \end{bmatrix} = \begin{bmatrix} 0.2627 & 0.6780 & 0.0593 \\ -\dfrac{0.2627}{1.8814} & -\dfrac{0.6780}{1.8814} & 0.5000 \\ 0.5000 & -\dfrac{0.6780}{1.4746} & -\dfrac{0.0593}{1.4746} \end{bmatrix} \begin{bmatrix} E'_R \\ E'_G \\ E'_B \end{bmatrix}$$

Q Quantization (Nbit) from $E'_Y E'_{CB} E'_{CR}$ (Recommendation BT.2020) to $D'_Y D'_{CB} D'_{CR}$ (Recommendation BT.2020)

$$D'_Y = \mathrm{INT}\left[\left(219 \times E'_Y + 16\right) \times 2^{N-8}\right]$$
$$D'_{CB} = \mathrm{INT}\left[\left(224 \times E'_{CB} + 128\right) \times 2^{N-8}\right]$$
$$D'_{CR} = \mathrm{INT}\left[\left(224 \times E'_{CR} + 128\right) \times 2^{N-8}\right]$$

Fig. 6-4: Conversion from a color space defined in the Rec. ITU-R BT.709 to a color space defined in the Rec. ITU-R BT.2020

6.3 Composition between Planes

As defined in Section 6.1.1, the composition between planes, except for the background plane, supports 256 levels of α blending. However, a receiver does not need to perform α blending to a video (motion picture) that is allocated to the multimedia plane.

6.4 Dynamic range conversion

As defined in Section 5.3, it is necessary to properly convert dynamic range in the processing of composition and presentation between monomedia. Conversion method of dynamic range is specified in operation.

<Blank Page>

Part 2　Monomedia Coding

Contents

Chapter 1: General Terms

1.1 Purpose

This part of the standard defines the monomedia coding for the second generation multimedia coding scheme for digital broadcasting that is specified as the Japanese standard specifications.

1.2 Scope

This part of the standard defines to the monomedia coding among the multimedia coding scheme used for the advanced broadband satellite digital broadcasting.

1.3 References

1.3.1 Normative References

The latest version is used if a specific version is not specified for any of the following documents and recommendations that are listed as normative references below. A specific version is used if it is specified for a normative reference; however, users of this standard are encouraged to investigate the possibility of applying the most recent version of the normative references.

- [1] ARIB STD-B32 "Video Coding, Audio Coding and Multiplexing Specifications for Digital Broadcasting" Part 1, Part 2
- [2] ISO/IEC 11172-3:1993 “Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s -- Part 3: Audio”
- [3] ISO/IEC 10646:2014 “Information technology -- Universal Coded Character Set (UCS)”
- [4] JIS X0201:1997 "7-bit and 8-bit Coded Character Sets for Information Interchange"
- [5] JIS X0213:2000 "7-bit and 8-bit Double Byte Coded Extended KANJI Sets for Information Interchange"
- [6] JIS X0213:2000/Amendment 1:2004 "7-bit and 8-bit Double Byte Coded Extended KANJI Sets for Information Interchange (Addendum)" [1][1]
- [7] “Scalable Vector Graphics (SVG) 1.1 (Second Edition)”, W3C Recommendation, 16 August 2011
- [8] “WOFF File Format 1.0”, W3C Recommendation, 13 December 2012
- [9] ISO/IEC 10918-1:1994 | Recommendation ITU-T T.81, “Information technology -- Digital compression and coding of continuous-tone still images: Requirements and guidelines”
- [10] Portable Network Graphics (PNG) specification (Second Edition), W3C Recommendation, 10 November 2003
- [11] “GRAPHICS INTERCHANGE FORMAT Version 89a”, CompuServe Inc.
- [12] MNG (Multiple-image Network Graphics) Format Version 1.0, W3C Memo, 31 January 2001
- [13] DAVIC 1.4.1 Specification Part 9, “Information Representation”, Annex B, “Coding of Linear Audio”
- [14] Unicode Standard 9.0.0[2]

[1] [5] to which [6] is applied is collectively called "JIS X0213:2004".

[2] It is expected that similar contents are provided in ISO/IEC 10646 Version 5.

1.4 Terms, definitions and abbreviations

1.4.1 Definitions

This standard defines following terms.

Term	Definition
Monomedia	Independent media element, such as video, still picture, figure, audio and character, that can be presented using only its own data without referring to other media.
Chunk	General name of the structure that represents a section of PNG coded or MNG coded data.
Synthesized sound	Expression media for music playback using electronic sound, etc. The music data to be replayed are saved in a receiver or given by another device.

1.4.2 Abbreviations

AAC	Advanced Audio Coding
AIFF	Audio Interchanger File Format
ALS	Audio Lossless Coding
AVC	Advanced Video Coding
GIF	Graphics Interchange Format
HEVC	High Efficiency Video Coding
ISO	International Organization for Standardization
ITU	International Telecommunication Union
JIS	Japan Industrial Standard
JPEG	Joint Photographic Experts Group
MNG	Multiple image Network Graphics
MPEG	Motion Picture Experts Group
PCM	Pulse Code Modulation
PNG	Portable Network Graphics
W3C	World Wide Web Consortium
UCS	Universal multiple-octet coded Character Set
WOFF	Web Open Font Format

Chapter 2: Video Coding

2.1 MPEG-2 VIDEO

The scheme defined in Section 5.1.1, Chapter 1 of STD-B32 is used for MPEG-2 video coding. When a low resolution is used, a scheme defined in Section 5.2.1, Chapter 1 of STD-B32 is used.

2.2 H.264|MPEG-4 AVC

A scheme defined in Section 5.1.2, Chapter 1 of STD-B32 is used for H.264|MPEG-4 AVC video coding. When a low resolution is used, a scheme defined in Section 5.2.2, Chapter 1 of STD-B32 is used.

2.3 H.265|HEVC

A scheme defined in Section 5.1.3, Chapter 1 of STD-B32 is used for H.265|HEVC video coding. When a low resolution is used, a scheme defined in Section 5.2.3, Chapter 1 of STD-B32 is used.

Chapter 3: Still Picture and Figure Coding

3.1 JPEG

A scheme specified in ISO/IEC10918-1:1994 is used for JPEG coding..

3.2 PNG

The W3C Recommendation, "Portable Network Graphics (PNG) Specification (Second Edition) , 10 November 2003", is used for PNG coding.

3.3 MNG

The specification based on MNG Format Version 1.0 (released January 31, 2001) is used for the MNG (Multiple-image Network Graphics) file format including animation and graphics.

3.3.1 Constraints on MNG

MNG should be operated according to the following rules.

- Multiple PNG images are included in a MNG file. Only a method to display the PNG images in order can be used.
- Only the object whose Object_id is 0 can be used.
- The following three frame rewriting rules are applied.
  (1) Use the previous frame, Framing_mode (Framing_mode = 0).
  (2) Simply overwrite a PNG image one at a time in every single frame cycle (Framing_mode = 1).
  (3) Display a PNG image one at a time in every single frame cycle after erasing the background with transparent color (Framing_mode = 3).
- Only the following two rules are applied to the animation repetition process.
  (1) Continue to display the last displayed PNG image (default).
  (2) Repeat displaying all images starting from the picture in the first file for a specified number of times (Termination_action = 3).

3.3.2 Available chunks

This section defines available chunks. The constraints are also provided when the value of each field is restricted.

3.3.2.1 MHDR

One MHDR always exists in the head. The field is fixed to 28-byte.

Field	Number of bytes	Meaning	Restriction
Frame_width	4	Frame width	
Frame_height	4	Frame height	
Ticks_per_second	4	Unit time for the time between frame	Other than 0
Nominal_layer_count	4	Number of layers	Fixed to 0
Nominal_frame_count	4	Number of frames	Fixed to 0
Nominal_play_time	4	Play time	Fixed to 0
Simplicity_profile	4	File profile information	Fixed to 0

3.3.2.2 MEND

One MEND always exists at the end. There is no field.

3.3.2.3 PLTE Global palette

The PLTE Global palette can be omitted. Only one PLTE Global palette can exist immediately before the first IHDR chunk. For the index color PNG image with an empty PLTE chunk that is included in MNG, see the PLTE chunk. When the PLTE Global palette is omitted, all index color PNG images that are included in MNG have individual PLTE chunks.

3.3.2.4 tRNS Global transparency

The tRNS Global transparency can be omitted. Only one tRNS Global transparency can exist immediately before the first IHDR chunk only when the Global PLTE chunk exists. For the index color PNG image, which refers to the Global PLTE chunk, included in MNG, see the tRNS chunk as the alpha value of each color.

3.3.2.5 IHDR, PNG chunks, IEND

IHDR, PNG chunks, IEND must be the same as the PNG image defined in Section 3.2.

3.3.2.6 TERM

TERM can be omitted. If it exists, only one TERM can exist immediately after the MHDR chunk. The field is fixed to 10-byte. If the TERM chunk is omitted, continue to display the last displayed image after showing the images of all files.

Field	Number of bytes	Meaning	Restriction
Termination_action	1	Specifying repetition process	Fixed to 3
Action_after_iterations	1	Action after the repetition process	Fixed to 0
Delay	4	Delay time before starting repetition	Fixed to 0
Iteration_max	4	Number of repetitions	

3.3.2.7 FRAM

Multiple FRAMs can exist. The field is fixed to 1 byte or 10-byte.

Field	Number of bytes	Meaning	Restriction
Framing_mode	1	Specifying frame rewrite mode	Restricted to 0, 1, or 3

The following fields can be omitted.

Subframe_name , Separator	1	Frame name	Fixed to 0
Change_interframe_ Delay	1	Time change flag between frames	Fixed to 2
Change_timeout_ and_termination	1	Timeout value change flag	Fixed to 0
Change_layer_clipping _boundaries	1	Clip value change flag	Fixed to 0
Change_sync_id_list	1	Sync id change flag	Fixed to 0
Interframe_delay	4	Time between frames	

3.3.2.8 DEFI

Multiple DEFIs can exist. This sets the display position of the subsequent PNG image. The field is fixed to 12-byte.

Field	Number of bytes	Meaning	Restriction
Object_id	2	Object ID	Fixed to 0
Do_not_show	1	Object hidden flag	Fixed to 0
Concrete_flag	1	Object attribute flag	Fixed to 0
X_location	4	X coordinate of an object	
Y_location	4	Y coordinate of an object	

3.4 GIF

For GIF (Graphics Interchange Format) file format, the file should be in accordance with the GRAPHICS INTERCHANGE FORMAT Version 89a defined by the U.S. Compuserve Inc.

3.5 SVG

When graphics are coded in SVG (Scalable Vector Graphics), the W3C Recommendation "Scalable Vector Graphics (SVG) 1.1 (Second Edition) 16 August 2011", should be applied.

Chapter 4: Audio Coding

4.1 MPEG-2 AAC

For the audio coding scheme using MPEG-2 AAC, the LC profile defined in Chapter 5 of STD-B32 Volume 2 is used.

4.2 MPEG-4 AAC

For the audio coding scheme using MPEG-4 AAC, the scheme defined in Chapter 6 of STD-B32 Volume 2 is used.

4.3 PEG-4 ALS

For an audio coding scheme using MPEG-4 ALS, the scheme defined in Chapter 7 of STD-B32 Volume 2 is used.

4.4 PCM (AIFF-C)

For the audio coding file format using PCM, AIFF-C (Audio Interchange File Format) defined in DAVIC 1.4 Specification Part 9 Annex B is used. The file is coded based on the conditions shown in Table 4-1.

Table 4-1: Constraint conditions on PCM coding parameter

Sampling frequency of television sound	Condition of PCM coding	
	Sampling frequency	Number of bits
48kHz	48kHz, 24kHz, 12kHz	8-bit or 16-bit

4.5 MP3

MP3 file format should be coded using MPEG-1 Audio Layer-3 defined in ISO/IEC11172-3 and MPEG-2BC defined in ISO/IEC13818-3 as audio coding scheme. The file is coded based on the conditions shown in Table 4-2.

Table 4-2: Constraint conditions on the coding parameters for MP3 file

Coding scheme	Condition of MP3 coding
	Sampling frequency
MPEG-1 Audio Layer-3	48kHz
MPEG-2BC	24kHz

Chapter 5: Character Coding

5.1 Character Coding Scheme

ISO/IEC 10646:2014 is used.

5.2 Repertoire of graphic character

Repertoire of graphic character employs the union of the following coded character sets is used as the subset of repertoire defined in ISO/IEC 10646:2014.

- The following value defined in ISO/IEC 10646:2014 Annex A
  - BASIC LATIN
  - LATIN-1 SUPPLEMENT
- Coded Kanji set defined in JIS X0213:2004
- Coded character set composed of UCS character referred by character name described in JIS X0213:2004 Annex 5, Table 1
- Coded character set composed of UCS character referred by character name described in JIS X0213:2004 Annex 5, Table 2
- Additional characters and symbols shown in Table 5-2 and Table 5-3

When using the variants of the above subset, the "Variation selector" and "Variation sequence" defined in Section 16.6 of ISO/IEC 10646:2014 are used. The details of its use are specified in an operational guideline.

5.3 Repertoire of control character

The control characters provided in Table 5-1 are used.

Table 5-1: Control coding

Code value	Name	Control function
U+0009	CHARACTER TABULATION	Tab
U+000A	LINE FEED (LF)	Line feed
U+000D	CARRIAGE RETURN (CR)	Carriage return

5.4 Encoding

Characters are encoded in UTF-8 defined in ISO/IEC 10646:2014 9.1. The signature, "<EF BB BF>" that represents UTF-8 is not used.

5.5 Encoding of Gaiji Characters

The schemes defined in the following documents are used for encoding of Gaiji characters (characters not included in repertoire defined in 5.2).

- Scalable Vector Graphics (SVG) 1.1 (Second Edition), W3C Recommendation, 16 August 2011, Chapter20 “Fonts”
- WOFF File Format 1.0, W3C Recommendation, 13 December 2012

Table 5-2: Additional symbols and characters

UCS code	UCS Character Name	Character's Meaning (Informative)	Example Glyph	Row-Cell in STD-B24
U+20158	CJK Unified Ideograph 20158		亭	85-02
U+20BB7	CJK Unified Ideograph 20BB7		吉	85-15
U+56E4	CJK Unified Ideograph 56E4		囤	85-23
U+5880	CJK Unified Ideograph 5880		墀	85-27
U+FA6B	CJK Compatibility Ideograph FA6B		恵	85-39
U+9FC4	CJK Unified Ideograph 9FC4		梁	85-47
U+6911	CJK Unified Ideograph 6911		椑	85-48
U+233CC	CJK Unified Ideograph 233CC		杞	85-53
U+233FE	CJK Unified Ideograph 233FE		栈	85-54
U+6DCA	CJK Unified Ideograph 6DCA		沼	85-61
U+6DF8	CJK Unified Ideograph 6DF8		清	85-62
U+FA6C	CJK Compatibility Ideograph FA6C		熙	85-67
U+242EE	CJK Unified Ideograph 242EE		熙	85-68
U+7421	CJK Unified Ideograph 7421		琡	85-77
U+3EDA	CJK Unified Ideograph 3EDA		珊	85-84

U+7575	CJK Unified Ideograph 7575		畵	85-85
U+9FC6	CJK Unified Ideograph 9FC6		祓	85-93
U+4103	CJK Unified Ideograph 4103		袂	85-94
U+9FC5	CJK Unified Ideograph 9FC5		褦	86-01
U+79DA	CJK Unified Ideograph 79DA		秚	86-02
U+7FA1	CJK Unified Ideograph 7FA1		羙	86-08
U+813A	CJK Unified Ideograph 813A		脺	86-10
U+FA6D	CJK Compatibility Ideograph FA6D		舘	86-11
U+8FF6	CJK Unified Ideograph 8FF6		迶	86-26
U+91B2	CJK Unified Ideograph 91B2		醲	86-30
U+9592	CJK Unified Ideograph 9592		閒	86-35
U+9AD9	CJK Unified Ideograph 9AD9		髙	86-39
U+26CC	CROSSING LANES	accident		90-01
U+26CD	DIABLED CAR	disabled car		90-02
U+2757	HEAVY EXCLAMATION MARK SYMBOL	obstacles on the road		90-03
U+26CF	PICK	under construction		90-04
U+26D0	CAR SLIDING	icy road		90-05

U+26D1	HELMET WITH WHITE CROSS	maintenance		90-06
U+26D2	CIRCLED CROSSING LANES	road closed		90-08
U+26D5	ALTERNATE ONE-WAY LEFT WAY TRAFFIC	alternate one-way traffic		90-09
U+26D3	CHAINS	tire chains required		90-10
U+26D4	NO ENTRY	no thoroughfare		90-11
U+1F17F	NEGATIVE SQUARED LATIN CAPITAL LETTER P	parking space (empty, full)		90-16
U+1F18A	CROSSED NEGATIVE SQUARED LATIN CAPITAL LETTER P	parking space (closed)		90-17
U+26D6	BLACK TWO-WAY LEFT WAY TRAFFIC	two-way traffic 1		90-20
U+26D7	WHITE TWO-WAY LEFT WAY TRAFFIC	two-way traffic 2		90-21
U'+26D8	BLACK LEFT LANE MERGE	lane merge 1		90-22
U+26D9	WHITE LEFT LANE MERGE	lane merge 2		90-23
U+26DA	DRIVE SLOW SIGN	drive slow 1		90-24
U+26DB	HEAVY WHITE DOWN-POINTING TRIANGLE	drive slow 2		90-25
U+26DC	LEFT CLOSED ENTRY	closed entry 1		90-26
U+26DD	SQUARED SALTIRE	closed entry 2		90-27

U'+26DE	FALLING DIAGNOL IN WHITE CIRCLE IN BLACK SQUARE	closed to large cars 1		90-28
U+26DF	BLACK TRUCK	closed to large cars 2		90-29
U+26E0	RESTRICTED LEFT ENTRY-1	restricted entry 1		90-30
U+26E1	RESTRICTED LEFT ENTRY-2	restricted entry 2		90-31
U+2B55	HEAVY LARGE CIRCLE	basic symbol for speed limit		90-32
U+3248	CIRCLED NUMBER TEN ON BLACK SQUARE	10km/h		90-33
U+3249	CIRCLED NUMBER TWENTY ON BLACK SQUARE	20km/h		90-34
U+324A	CIRCLED NUMBER THIRTY ON BLACK SQUARE	30km/h		90-35
U+324B	CIRCLED NUMBER FORTY ON BLACK SQUARE	40km/h		90-36
U+324C	CIRCLED NUMBER FIFTY ON BLACK SQUARE	50km/h		90-37
U+324D	CIRCLED NUMBER SIXTY ON BLACK SQUARE	60km/h		90-38
U+324E	CIRCLED NUMBER SEVENTY ON BLACK SQUARE	70km/h		90-39
U+324F	CIRCLED NUMBER EIGHTY ON BLACK SQUARE	80km/h		90-40

U+2491	NUMBER TEN FULL STOP	time of day (10:00)	10.	90-45
U+2492	NUMBER ELEVEN FULL STOP	time of day (11:00)	11.	90-46
U+2493	NUMBER TWELVE FULL STOP	time of day (12:00)	12.	90-47
U+1F14A	SQUARED HV	HDTV	HV	90-48
U+1F14C	SQUARED SD	SDTV	SD	90-49
U+1F13F	SQUARED LATIN CAPITAL LETTER P	progressive broadcasting	P	90-50
U+1F146	SQUARED LATIN CAPITAL LETTER W	wide-format (16:9) broadcasting service	W	90-51
U+1F14B	SQUARED MV	multi-view television	MV	90-52
U+1F210	SQUARED CJK UNIFIED IDEOGRAPH-624B	sign language interpretation	手	90-53
U+1F211	SQUARED CJK UNIFIED IDEOGRAPH-5B57	closed-captioned broadcasting	字	90-54
U+1F212	SQUARED CJK UNIFIED IDEOGRAPH-53CC	two-way broadcasting service	双	90-55
U+1F213	SQUARED KATAKANA DE	data broadcasting service	デ	90-56
U+1F142	SQUARED LATIN CAPITAL LETTER S	stereo broadcasting service	S	90-57
U+1F214	SQUARED CJK UNIFIED IDEOGRAPH-4E8C	bilingual broadcasting service	二	90-58
U+1F215	SQUARED CJK UNIFIED IDEOGRAPH-591A	sound-multiplex broadcasting service	多	90-59
U+1F216	SQUARED CJK UNIFIED IDEOGRAPH-89E3	commentary broadcasting	解	90-60
U+1F14D	SQUARED SS	surrounding stereo broadcasting service	SS	90-61

U+1F131	SQUARED LATIN CAPITAL LETTER B	B-mode stereo compression		90-62
U+1F13D	SQUARED LATIN CAPITAL LETTER N	news		90-63
U+2B1B	BLACK LARGE SQUARE	background, rectangle		90-64
U+2B24	BLACK LARGE CIRCLE	background, circle		90-65
U+1F217	SQUARED CJK UNIFIED IDEOGRAPH-5929	weather forecast		90-66
U+1F218	SQUARED CJK UNIFIED IDEOGRAPH-4EA4	traffic information		90-67
U+1F219	SQUARED CJK UNIFIED IDEOGRAPH-6620	drama film		90-68
U+1F21A	SQUARED CJK UNIFIED IDEOGRAPH-7121	free broadcasting service		90-69
U+1F21B	SQUARED CJK UNIFIED IDEOGRAPH-6599	pay broadcasting service		90-70
U+26BF	SQUARED KEY	parental lock		90-71
U+1F21C	SQUARED CJK UNIFIED IDEOGRAPH-524D	the first part		90-72
U+1F21D	SQUARED CJK UNIFIED IDEOGRAPH-5F8C	the latter part		90-73
U+1F21E	SQUARED CJK UNIFIED IDEOGRAPH-518D	rebroadcast		90-74
U+1F21F	SQUARED CJK UNIFIED IDEOGRAPH-65B0	new series of programs		90-75
U+1F220	SQUARED CJK UNIFIED IDEOGRAPH-521D	first released program		90-76
U+1F221	SQUARED CJK UNIFIED IDEOGRAPH-7D42	the last episode		90-77
U+1F222	SQUARED CJK UNIFIED IDEOGRAPH-751F	live broadcast		90-78

U+1F223	SQUARED CJK UNIFIED IDEOGRAPH-8CA9	mail-order		90-79
U+1F224	SQUARED CJK UNIFIED IDEOGRAPH-58F0	voice actors		90-80
U+1F225	SQUARED CJK UNIFIED IDEOGRAPH-5439	dubbed version		90-81
U+1F14E	SQUARED PPV	pay-per-view		90-82
U+3299	CIRCLED IDEOGRAPH SECRET	confidential		90-83
U+1F200	SQUARE HIRAGANA HOKA	and others		90-84
U+26E3	HEAVY CIRCLE WITH STROKE AND TWO DOTS ABOVE	public office, governmental agency		91-01
U+2B56	HEAVY OVAL WITH OVAL INSIDE	prefectural office		91-02
U+2B57	HEAVY CIRCLE WITH CIRCLE INSIDE	municipal office		91-03
U+2B58	HEAVY CIRCLE	town office, village office		91-04
U+2B59	HEAVY CIRCLED SALTIRE	police office		91-05
U+2613	SALTIRE	police satellite office		91-06
U+328B	CIRCLED IDEOGRAPH FIRE	fire station		91-07
U+26E8	BLACK CROSS ON SHIELD	hospital, clinic		91-09
U+3246	CIRCLED IDEOGRAPH SCHOOL	school		91-10
U+3245	CIRCLED IDEOGRAPH KINDERGARTEN	kindergarten		91-11

U+26E9	SHINTO SHRINE	shrine		91-12
U+0FD6	LEFT FACING SVASTI SIGN	temple		91-13
U+26EA	CHURCH	church		91-14
U+26EB	CASTLE	remains of a castle		91-15
U+26EC	HISTORIC SITE	historic site, place of scenic beauty		91-16
U+26ED	GEAR WITHOUT HUB	factory		91-18
U+26EE	GEAR WITH HANDLES	power plant, power substation		91-19
U+26EF	MAP SYMBOL FOR LIGHTHOUSE	lighthouse		91-20
U+2693	ANCHOR	harbor		91-21
U+2708	AIRPLANE	airport		91-22
U+26F0	MOUNTAIN	mountain		91-23
U+26F1	UMBRELLA ON GROUND	bathing beach		91-24
U+26F2	FOUNTAIN	park		91-25
U+26F3	FLAG IN HOLE	golf course		91-26
U+26F4	FERRY	ferryboat terminal		91-27
U+26F5	SAILBOAT	marina, yacht harbor		91-28

U+1F157	NEGATIVE CIRCLED LATIN CAPITAL LETTER H	hotel		91-29
U+24B9	CIRCLED CAPITAL LETTER D	department store		91-30
U+24C8	CIRCLED CAPITAL LETTER S	station		91-31
U+26F6	SQUARE FOUR CORNERS	intersection		91-32
U+1F15F	NEGATIVE CIRCLED LATIN CAPITAL LETTER P	parking space		91-33
U+1F18B	NEGATIVE SQUARED IC	interchange (part of the highway system)		91-34
U+1F18D	NEGATIVE SQUARED SA	service area (part of the highway system)		91-35
U+1F18C	NEGATIVE SQUARED PA	parking area (part of the highway system)		91-36
U+1F179	NEGATIVE SQUARED LATIN CAPITAL LETTER J	junction (part of the highway system)		91-37
U+26F7	SKIER	skiing field		91-38
U+26F8	ICE SKATE	ice skating field		91-39
U+26F9	PERSON WITH BALL	track and field, gymnasium		91-40
U+26FA	TENT	camping site		91-41
U+1F17B	NEGATIVE SQUARED LATIN CAPITAL LETTER L	leisure center		91-42
U+26FB	JAPANESE BANK SYMBOL	bank		91-44

U+26FC	HEADSTONE GRAVEYARD SYMBOL	graveyard, memorial park, cemetery		91-45
U+26FD	FUEL PUMP	gas station		91-46
U+26FE	CUP ON BLACK SQUARE	drive-in restaurant		91-47
U+1F17C	NEGATIVE SQUARED LATIN CAPITAL LETTER M	museum, cultural center		91-48
U+26FF	WHITE FLAG WITH HJORIZONTAL MIDDLE BLACK STRIPE	Self-Defense Forces site		91-49
U+27A1	BLACK RIGHTWARDS ARROW	to be continued		92-01
U+2B05	LEFTWARDS BLACK ARROW			92-02
U+2B06	UPWARDS BLACK ARROW			92-03
U+2B07	DOWNWARDS BLACK ARROW			92-04
U+2B2F	WHITE VERTICAL ELLIPSE			92-05
U+2B2E	BLACK VERTICAL ELLIPSE			92-06
U+33A5	SQUARE M CUBED	cubic meter		92-12
U+33A0	SQUARE CM SQUARED	square centimeter		92-14
U+33A4	SQUARE CM CUBED	cubic centimeter		92-15
U+1F100	DIGIT ZERO FULL STOP	time of day (0:00)		92-16
U+2488	DIGIT ONE FULL STOP	time of day (1:00)		92-17

U+2489	DIGIT TWO FULL STOP	time of day (2:00)	2.	92-18
U+248A	DIGIT THREE FULL STOP	time of day (3:00)	3.	92-19
U+248B	DIGIT FOUR FULL STOP	time of day (4:00)	4.	92-20
U+248C	DIGIT FIVE FULL STOP	time of day (5:00)	5.	92-21
U+248D	DIGIT SIX FULL STOP	time of day (6:00)	6.	92-22
U+248E	DIGIT SEVEN FULL STOP	time of day (7:00)	7.	92-23
U+248F	DIGIT EIGHT FULL STOP	time of day (8:00)	8.	92-24
U+2490	DIGIT NINE FULL STOP	time of day (9:00)	9.	92-25
U+1F101	DIGIT ZERO COMMA		0,	92-32
U+1F102	DIGIT ONE COMMA		1,	92-33
U+1F103	DIGIT TWO COMMA		2,	92-34
U+1F104	DIGIT THREE COMMA		3,	92-35
U+1F105	DIGIT FOUR COMMA		4,	92-36
U+1F106	DIGIT FIVE COMMA		5,	92-37
U+1F107	DIGIT SIX COMMA		6,	92-38
U+1F108	DIGIT SEVEN COMMA		7,	92-39
U+1F109	DIGIT EIGHT COMMA		8,	92-40

U+1F10A	DIGIT NINE COMMA		9,	92-41
U+3233	PARENTHESIZED IDEOGRAPH SOCIETY	corporation aggregate	㈳	92-42
U+3236	PARENTHESIZED IDEOGRAPH FINANCIAL	incorporated foundation	㈶	92-43
U+3244	CIRCLED IDEOGRAPH QUESTION		㉾	92-47
U+27D0	WHITE DIAMOND WITH CENTERED DOT		⟐	92-52
U+1F12D	CIRCLED CD	circled "Compact Disc"	ⓒⓓ	92-55
U+1F12C	CIRCLED LATIN CAPITAL LETTER R	disc record	®	92-86
U+1F12B	CIRCLED ITALIC LATIN CAPITAL LETTER C	single disc record, compact disc	©	92-87
U+3247	CIRCLED IDEOGRAPH KOTO	koto (Japanese harp)	㋏	92-88
U+1F190	SQUARED DJ	disc jockey	DJ	92-89
U+1F226	SQUARED CJK UNIFIED IDEOGRAPH-6F14	performed by	演	92-90
U+213B	FACSIMILLE SIGN	facsimile	Fax	92-91
U+322A	PARENTHESIZED IDEOGRAPH MOON	Monday	(月)	93-01
U+322B	PARENTHESIZED IDEOGRAPH FIRE	Tuesday	(火)	93-02
U+322C	PARENTHESIZED IDEOGRAPH WATER	Wednesday	(水)	93-03
U+322D	PARENTHESIZED IDEOGRAPH WOOD	Thursday	(木)	93-04
U+322E	PARENTHESIZED IDEOGRAPH METAL	Friday	(金)	93-05

U+322F	PARENTHESIZED IDEOGRAPH EARTH	Saturday	㈯	93-06
U+3230	PARENTHESIZED IDEOGRAPH SUN	Sunday	㈰	93-07
U+3237	PARENTHESIZED IDEOGRAPH CONGRATULATION	public holiday	㈷	93-08
U+3036	CIRCLED POSTAL MARK	post office	〒	93-15
U+26BE	BASEBALL		⚾	93-16
U+1F240	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-672C	home run	🉀	93-17
U+1F241	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-4E09	three-base hit	🉁	93-18
U+1F242	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-4E8C	two-base hit	🉂	93-19
U+1F243	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-5B89	single hit	🉃	93-20
U+1F244	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-70B9	score	🉄	93-21

U+1F245	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-6253	at bats	〔打〕	93-22
U+1F246	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-76D7	stolen base	〔盗〕	93-23
U+1F247	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-52DD	win	〔勝〕	93-24
U+1F248	TORTOISE SHELL BRACKETED CJK UNIFIED IDEOGRAPH-6557	loss	〔敗〕	93-25
U+1F12A	TORTOISE SHELL BRACKETED LATIN CAPITAL LETTER S	save	〔S〕	93-26
U+1F227	SQUARED CJK UNIFIED IDEOGRAPH-6295	pitcher	投	93-27
U+1F228	SQUARED CJK UNIFIED IDEOGRAPH-6355	catcher	捕	93-28
U+1F229	SQUARED CJK UNIFIED IDEOGRAPH-4E00	first baseman	一	93-29
U+1F214	SQUARED CJK UNIFIED IDEOGRAPH-4E8C	second baseman	二	93-30
U+1F22A	SQUARED CJK UNIFIED IDEOGRAPH-4E09	third baseman	三	93-31
U+1F22B	SQUARED CJK UNIFIED IDEOGRAPH-904A	shortstop	遊	93-32
U+1F22C	SQUARED CJK UNIFIED IDEOGRAPH-5DE6	left fielder	左	93-33

U+1F22D	SQUARED CJK UNIFIED IDEOGRAPH-4E2D	center fielder	🈭	93-34
U+1F22E	SQUARED CJK UNIFIED IDEOGRAPH-53F3	right fielder	🈮	93-35
U+1F22F	SQUARED CJK UNIFIED IDEOGRAPH-6307	designated hitter	🈯	93-36
U+1F230	SQUARED CJK UNIFIED IDEOGRAPH-8D70	pinch runner	🈰	93-37
U+1F231	SQUARED CJK UNIFIED IDEOGRAPH-6253	pinch hitter	🈱	93-38
U+3390	SQUARE HZ	hertz	㎐	93-41
U+33CA	SQUARE HA	hectare	㎗	93-42
U+33A2	SQUARE KM SQUARED	square kilometer	㎢	93-44
U+3371	SQUARE HPA	hectopascal	㎩	93-45
U+2189	VULGAR FRACTION ZERO THIRDS		↉	93-49
U+2156	VULGAR FRACTION TWO FIFTHS		⅖	93-55
U+2157	VULGAR FRACTION THREE FIFTHS		⅗	93-56
U+2158	VULGAR FRACTION FOUR FIFTHS		⅘	93-57
U+2159	VULGAR FRACTION ONE SIXTH		⅙	93-58
U+215A	VULGAR FRACTION FIVE SIXTHS		⅚	93-59
U+2150	VULGAR FRACTION ONE SEVENTH		⅐	93-60
U+215B	VULGAR FRACTION ONE EIGHTS		⅛	93-61

U+2151	VULGAR FRACTION ONE NINETH			93-62
U+2152	VULGAR FRACTION ONE TENTH			93-63
U+26C4	SNOWMAN WITHOUT SNOW	snow		93-67
U+26C9	TURNED WHITE SHOGI PIECE	white on shogi (Japanese chess)		93-70
U+26CA	TURNED BLACK SHOGI PIECE	black on shogi (Japanese chess)		93-71
U+26CB	WHITE DIAMOND IN SQUARE			93-76
U+2A00	N-ARY CIRCLED DOT OPERATOR			93-77
U+26C5	SUN BEHIND CLOUD	cloudy or fair		93-80
U+2614	UMBRELLA WITH RAIN DROPS	shower		93-81
U+26C6	RAIN	rain		93-82
U+26C7	BLACK SNOWMAN	heavy snow		93-84
U+26A1	HIGH VOLTAGE SIGN	thunder		93-85
U+26C8	THUNDER CLOUD AND RAIN	thunderstorm		93-86
U+269E	THREE LINES CONVERGING RIGHT			93-88
U+269F	THREE LINES CONVERGING LEFT			93-89
U+2474	PARENTHESIZED DIGIT ONE			94-17
U+2475	PARENTHESIZED DIGIT TWO			94-18

U+2476	PARENTHESIZED DIGIT THREE		⑶	94-19
U+2477	PARENTHESIZED DIGIT FOUR		⑷	94-20
U+2478	PARENTHESIZED DIGIT FIVE		⑸	94-21
U+2479	PARENTHESIZED DIGIT SIX		⑹	94-22
U+247A	PARENTHESIZED DIGIT SEVEN		⑺	94-23
U+247B	PARENTHESIZED DIGIT EIGHT		⑻	94-24
U+247C	PARENTHESIZED DIGIT NINE		⑼	94-25
U+247D	PARENTHESIZED DIGIT TEN		⑽	94-26
U+247E	PARENTHESIZED DIGIT ELEVEN		⑾	94-27
U+247F	PARENTHESIZED DIGIT TWELVE		⑿	94-28
U+1F110	PARENTHESIZED LATIN CAPITAL LETTER A		🄐	94-33
U+1F111	PARENTHESIZED LATIN CAPITAL LETTER B		🄑	94-34
U+1F112	PARENTHESIZED LATIN CAPITAL LETTER C		🄒	94-35
U+1F113	PARENTHESIZED LATIN CAPITAL LETTER D		🄓	94-36
U+1F114	PARENTHESIZED LATIN CAPITAL LETTER E		🄔	94-37
U+1F115	PARENTHESIZED LATIN CAPITAL LETTER F		🄕	94-38
U+1F116	PARENTHESIZED LATIN CAPITAL LETTER G		🄖	94-39

U+1F117	PARENTHESIZED LATIN CAPITAL LETTER H		(H)	94-40
U+1F118	PARENTHESIZED LATIN CAPITAL LETTER I		(I)	94-41
U+1F119	PARENTHESIZED LATIN CAPITAL LETTER J		(J)	94-42
U+1F11A	PARENTHESIZED LATIN CAPITAL LETTER K		(K)	94-43
U+1F11B	PARENTHESIZED LATIN CAPITAL LETTER L		(L)	94-44
U+1F11C	PARENTHESIZED LATIN CAPITAL LETTER M		(M)	94-45
U+1F11D	PARENTHESIZED LATIN CAPITAL LETTER N		(N)	94-46
U+1F11E	PARENTHESIZED LATIN CAPITAL LETTER O		(O)	94-47
U+1F11F	PARENTHESIZED LATIN CAPITAL LETTER P		(P)	94-48
U+1F120	PARENTHESIZED LATIN CAPITAL LETTER Q		(Q)	94-49
U+1F121	PARENTHESIZED LATIN CAPITAL LETTER R		(R)	94-50
U+1F122	PARENTHESIZED LATIN CAPITAL LETTER S		(S)	94-51
U+1F123	PARENTHESIZED LATIN CAPITAL LETTER T		(T)	94-52
U+1F124	PARENTHESIZED LATIN CAPITAL LETTER U		(U)	94-53
U+1F125	PARENTHESIZED LATIN CAPITAL LETTER V		(V)	94-54
U+1F126	PARENTHESIZED LATIN CAPITAL LETTER W		(W)	94-55
U+1F127	PARENTHESIZED LATIN CAPITAL LETTER X		(X)	94-56

U+1F128	PARENTHESIZED LATIN CAPITAL LETTER Y		(Y)	94-57
U+1F129	PARENTHESIZED LATIN CAPITAL LETTER Z		(Z)	94-58

Table 5-3 Additional symbols used for EPG

UCS code	UCS Character Name	Character's Meaning (Informative)	Example Glyph
U+1F19B	SQUARED THREE D	3DTV	3D
U+1F19C	SQUARED SECOND SCREEN	second screen	2nd Scr
U+1F19D	SQUARED TWO K	2K broadcasting	2K
U+1F19E	SQUARED FOUR K	4K broadcasting	4K
U+1F19F	SQUARED EIGHT K	8K broadcasting	8K
U+1F1A0	SQUARED FIVE POINT ONE	5.1 channel surround sound	5.1
U+1F1A1	SQUARED SEVEN POINT ONE	7.1 channel surround sound	7.1
U+1F1A2	SQUARED TWENTY-TWO POINT TWO	22.2channel surround sound	22.2
U+1F1A3	SQUARED SIXTY P	60Hz progressive scan broadcasting	60P
U+1F1A4	SQUARED ONE HUNDRED TWENTY P	120Hz progressive scan broadcasting	120P
U+1F1A5	SQUARED LATIN SMALL LETTER D	data broadcasting	d
U+1F1A6	SQUARED HC	Hybridcast	HC

U+1F1A7	SQUARED HDR	High Dynamic Range	
U+1F1A8	SQUARED HI-RES	High-Resolution Audio	
U+1F1A9	SQUARED LOSSLESS	Lossless Audio	
U+1F1AA	SQUARED SHV	SHV broadcasting	
U+1F1AB	SQUARED UHD	UHDTV	
U+1F1AC	SQUARED VOD	VOD	
U+1F23B	SQUARED CJK UNIFIED IDEOGRAPH-914D	Simulcast on the internet	

Description 1 (Vacant number)

Description 2 Coded character set composing graphic character repertoire

Graphic character repertoire used in character encoding scheme defined in Chapter 5 is a union of coded character sets listed in Section 5.2 (each set is called sub repertoire in this description for convenience) as defined in the section. This description describes the details of each sub repertoire for the purpose of assistance to interpret what kind of set the repertoire of this graphic character is.

1 BASIC LATIN

This sub repertoire is composed of 95 characters represented by code value U+20~U+7E. Note that this is a set different from graphic character set used for Lain character defined in JIS X0201:1997.

Table D2-1 Difference between BASIC LATIN and JIS X0201:1997 graphic character set used for Latin character

	ISO/IEC 10646:2014 BASIC LATIN	JIS X0201:1997 graphic character set used for Latin character
Character represented by code value U+20(2/0)	SPACE	(Not included in a set)
Character represented by code value U+5C (5/12)	REVERSE SOLIDUS	YEN SIGN
Character represented by code value U+7E(7/14)	TILDE	OVERLINE

2 LATIN-1 SUPPLEMENT

This sub repertoire is composed of 96 characters represented by code value U+A0~U+FF. Ninety-five (95) characters excluding U+B5(MICRO SIGN) are also contained in the “coded Kanji set defined in JIS X0213:2004”, sub repertoire mentioned later.

3 Coded Kanji set defined in JIS X0213:2004

This sub repertoire is composed of 11233 characters defined as coded Kanji set in JIS X0213:2004, and is a coded character set in which each character’s coded value is defined in the values shown in “0221” column from Table 1 to Table 24 in JIS X0213:2004 Annex 4, “UCS” column in JIS X0213:2004 Annex 6, or “CJK” column in JIS X0208:1997 Annex 6. Note that the coded kanji set defined in JISX0213:2004 includes that defined in JIS X0208:1997 and is called Kanji set which contains non-kanji in both standard.

Application of JIS X0213:2004 Annex 5 Table 2 is not indicated intentionally. This sub repertoire is a set which does not apply the Table. That is to say, this sub repertoire contains U+21~U+7E (94 characters excluding SPACE from BASIC LATIN), U+A5 (YEN SIGN) and U+203E (OVERLINE). On the other hand each character of full-width coded values, U+FF01~U+FF5E, U+FFE5 and U+FFE3 is contained in the sub repertoire “coded character set composed of UCS characters referring to character names described in JIS X0213:2004 Annex 5 Table 2” not in this sub repertoire.

As shown in Table D2-2, this sub repertoire contains characters whose coded values are UCS code value sequence. Note that characters represented by each coded character composing these sequences can be only used for the sequence concerned unless it is contained by itself in any sub repertoire. Specifically U+309A(COMBINING KATAKANA-HIRAGANA

SEMIVOICEDSOUND MARK) are applicable to this and only used in the sequence shown in Table D2-2.

Table D2-2 Character whose coded value is the sequence of multiple UCS code value out of characters contained in coded Kanji set defined in JIS X0213:2004

JIS X0213 Plane, Row, Cell	JIS X0213:2004 Character name	Code value
1-4-87	[HIRAGANA LETTER BIDAKUON NGA]	<304B,309A>
1-4-88	[HIRAGANA LETTER BIDAKUON NGI]	<304D,309A>
1-4-89	[HIRAGANA LETTER BIDAKUON NGU]	<304F,309A>
1-4-90	[HIRAGANA LETTER BIDAKUON NGE]	<3051,309A>
1-4-91	[HIRAGANA LETTER BIDAKUON NGO]	<3053,309A>
1-5-87	[KATAKANA LETTER BIDAKUON NGA]	<30AB,309A>
1-5-88	[KATAKANA LETTER BIDAKUON NGI]	<30AD,309A>
1-5-89	[KATAKANA LETTER BIDAKUON NGU]	<30AF,309A>
1-5-90	[KATAKANA LETTER BIDAKUON NGE]	<30B1,309A>
1-5-91	[KATAKANA LETTER BIDAKUON NGO]	<30B3,309A>
1-5-92	[KATAKANA LETTER AINU CE]	<30BB,309A>
1-5-93	[KATAKANA LETTER AINU TU]	<30C4,309A>
1-5-94	[KATAKANA LETTER AINU TO]	<30C8,309A>
1-6-88	[KATAKANA LETTER AINU P]	<31F7,309A>
1-11-36	[LATIN SMALL LETTER AE WITH GRAVE]	<E6,300>
1-11-40	[LATIN SMALL LETTER OPEN O WITH GRAVE]	<254,300>
1-11-41	[LATIN SMALL LETTER OPEN O WITH ACUTE]	<254,301>
1-11-42	[LATIN SMALL LETTER TURNED V WITH GRAVE]	<28C,300>
1-11-43	[LATIN SMALL LETTER TURNED V WITH ACUTE]	<28C,301>
1-11-44	[LATIN SMALL LETTER SCHWA WITH GRAVE]	<259,300>
1-11-45	[LATIN SMALL LETTER SCHWA WITH ACUTE]	<259,301>
1-11-46	[LATIN SMALL LETTER HOOKED SHUWA WITH GRAVE]	<25A,300>
1-11-47	[LATIN SMALL LETTER HOOKED SHUWA WITH ACUTE]	<25A,301>
1-11-69	[RISING SYMBOL]	<2E9,2E5>
1-11-70	[FALLING SYMBOL]	<2E5,2E9>

4 Coded character set composed of UCS character referred by character name described in JIS X0213-2004 Annex 5, Table 1

This sub repertoire is a set of so-called half-width katakana and composed of 63 characters of coded value U+FF61~U+FF9F.

5 Coded character composed of UCS character referred by character name described in JIS X0213-2004 Annex 5, Table 2

This sub repertoire is a set of so-called full-width alphanumeric character and composed of 96 characters of coded value U+FF01~U+FF5E, U+FFE3 and U+FFE5.

6 Characters and symbols shown in Table 5-2

This sub repertoire aims to cover a set of additional symbols defined in ARIB STD-B24 Fascicle 1 Part 2, 7.1.1.3. Out of characters consisting the set the sub repertoire is a coded character set composed of 283 characters excluding 36 characters which are not covered in ISO/IEC 10646:2014 (92Ku26Ten~31Ten, 56Ten~85Ten) by making a set of 319 characters which is not covered by JIS X0213:2004 as a base.

7 Symbol shown in Table 5-3

This sub repertoire, which assumes to be mainly used for EPG, is a set of symbol defined additionally in Unicode Standard 9.0.0, and a coded character set corresponding to 19 characters proposed in ISO/IEC JTC 1/SC2/WG2 N4671.

<Blank Page>

Part 3　Coding of Closed Caption and Super impose

Contents

Chapter 1: General Terms

1.1 Purpose

This part of the standard defines the coding of closed caption and super impose for the second generation multimedia coding scheme for digital broadcasting that is specified as the Japanese standard specifications.

1.2 Scope

This standard applies to the coding of closed caption and super impose among the multimedia coding scheme for the advanced broadband satellite digital broadcasting.

1.3 References

1.3.1 Normative references

The latest version is used if a specific version is not specified for any of the following documents and recommendations that are listed as normative references. A specific version is used if it is specified for a normative reference; however, iusers of this standard are encouraged to investigate the possibility of applying the most recent version of the normative references.

- [1] Timed Text Markup Language 1 (TTML1) (Second Edition), W3C Recommendation 24 September 2013 (hereinafter referred to as "TTML1 standard")
- [2] SMPTE Timed Text Format (SMPTE-TT), SMPTE ST 2052-1:2013

1.3.2 Informative references

The standard related to this standard is as follows.

- [1] ARIB STD-B60 "MMT-based Media Transport Scheme in Digital Broadcasting Systems"

1.4 Terminology

1.4.1 Abbreviations

AIFF	Audio Interchange File Format
AIFF-C	Audio Interchange File Format - Compressed
MMT	MPEG Media Transport
MPU	Media Processing Unit
PCM	Pulse Code Modulation
PNG	Portable Network Graphics
SMPTE	Society of Motion Picture & Television Engineers
SMPTE-TT	SMPTE - Timed Text
SVG	Scalable Vector Graphics
TTML	Timed Text Markup Language
UCS	Universal multiple-octet coded Character Set
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
WOFF	Web Open Font Format

Chapter 2: Presentation Function of Closed Caption and Super impose

2.1 Types of Closed Caption and Super impose

In this standard, it is called closed caption for a service which is presented by characters, still pictures, or sound without using runtime environment in a receiver and is related to the video content. A service which not related to video content is called super impose. For viewpoint of specification of the transmission and coding scheme, they are not distinguished from each other, and are collectively called closed caption.

2.2 Overview of Presentation Function of Closed Caption

Table 2-1 shows the overview of the presentation function of closed caption in this standard.

Table 2-1: Presentation function of closed caption

Presentation function	Format	1920 x 1080, 3840 x 2160, 7680 x 4320
	Orientation	horizontal, vertical, mixture of horizontal and vertical
	Character repertoire	Characters defined in Vol.1, Part. 2 of this standard and additional downloadable characters
	Character encoding	UTF-8
	Font	Selectable
	Gaiji	Available by SVG 1.1 and WOFF File Format 1.0
	Font size	Selectable in pixels
	Font color	256 steps each for red, green, blue and alpha (transparency)
	Color designation unit	Each character
	Character attribute	flashing, underline, overline, strike-out, enclosure, shadow, bold, italic, italic bold
	Graphics	PNG, SVG
Presentation control	Timing control	start time, end time, duration
	Switching control	cut, pop-on, paint-on, roll-up, scroll, key-frame animation
Misc.	Audio clip	PCM (AIFF-C), MP3, MPEG2-AAC, MPEG4-AAC
	Built-in sound	PCM (AIFF-C), MP3, MPEG2-AAC, MPEG4-AAC

Table 2-2: Closed Caption display mode

Display mode		Display function
Live	Automatic presentation	Presented regardless of viewer's operation
	Automatic erasure	Erased regardless of viewer's operation
	Selectable	Control for presentation by the viewer's operation and receiver setting
Playback of recording	Automatic display	Recorded automatically when recording and always displayed when playing a video regardless of viewer's operation
	Automatic non-display	Non-displayed when playing a video
	Selectable display	Recorded automatically when recording and displayed (or non-displayed) when playing a video based on the viewer's operation.

Format is 1920 x 1080, 3840 x 2160 or 7680 x 4320.

Table 2-3: Display format and display area size

Display format	Size of display area
1920 x 1080	Width 1920 x Height 1080
3840 x 2160	Width 3840 x Height 2160
7680 x 4320	Width 7680 x Height 4320

Chapter 3: Description Language for Closed Caption and Super impose

This chapter defines the description language, ARIB-TTML, for closed caption and super impose, to which some functions are added for the advanced broadband satellite digital broadcasting, based on the W3C Recommendation "Timed Text Markup Language 1 (TTML1) (Second Edition)" and SMPTE ST 2052-1:2013 "Timed Text Format (SMPTE-TT)".

The documents described using the ARIB-TTML description language are called "ARIB-TTML document".

3.1 Character Coding

3.1.1 Character coding scheme

UTF-8 defined in Volume 1, Part 2 of this standard is used as the character coding scheme for the ARIB-TTML documents.

3.1.2 Repertoire

Repertoire including Gaiji defined in Volume 1, Part 2 of this standard are used for the ARIB-TTML document encoded in UTF-8. The following control codes are used.

Table 3-1: Control codes available in the ARIB-TTML document

Code location	Value
U+0009	Character tabulation (TAB)
U+000A	LINE FEED (LF)
U+000D	CARRIAGE RETURN (CR)

3.2 Namespace

The following namespaces and prefixes are used for the elements and attributes of the ARIB-TTML document. The prefix "arib-tt:" is used for the elements and attributes for extension defined by this standard.

Table 3-2: Namespace

Name	Prefix	Namespace
TT	tt:	http://www.w3.org/ns/ttml
TT Parameter	ttp:	http://www.w3.org/ns/ttml#parameter
TT Style	tts:	http://www.w3.org/ns/ttml#styling
TT Metadata	ttm:	http://www.w3.org/ns/ttml#metadata
SMPTE	smpte:	http://www.smpte-ra.org/schemas/2052-1/2013/smpte-tt
ARIB	arib-tt:	http://www.arib.or.jp/ns/arib-ttml/v1_0

3.3 Structure of the ARIB-TTML Document

An ARIB-TTML document conforms to the TTML1 standard [1] and the SMPTE-TT standard [2] and is written in a document structure using the elements and attributes extended by this standard. Figure 3-1 shows the upper structure of the ARIB-TTML document.

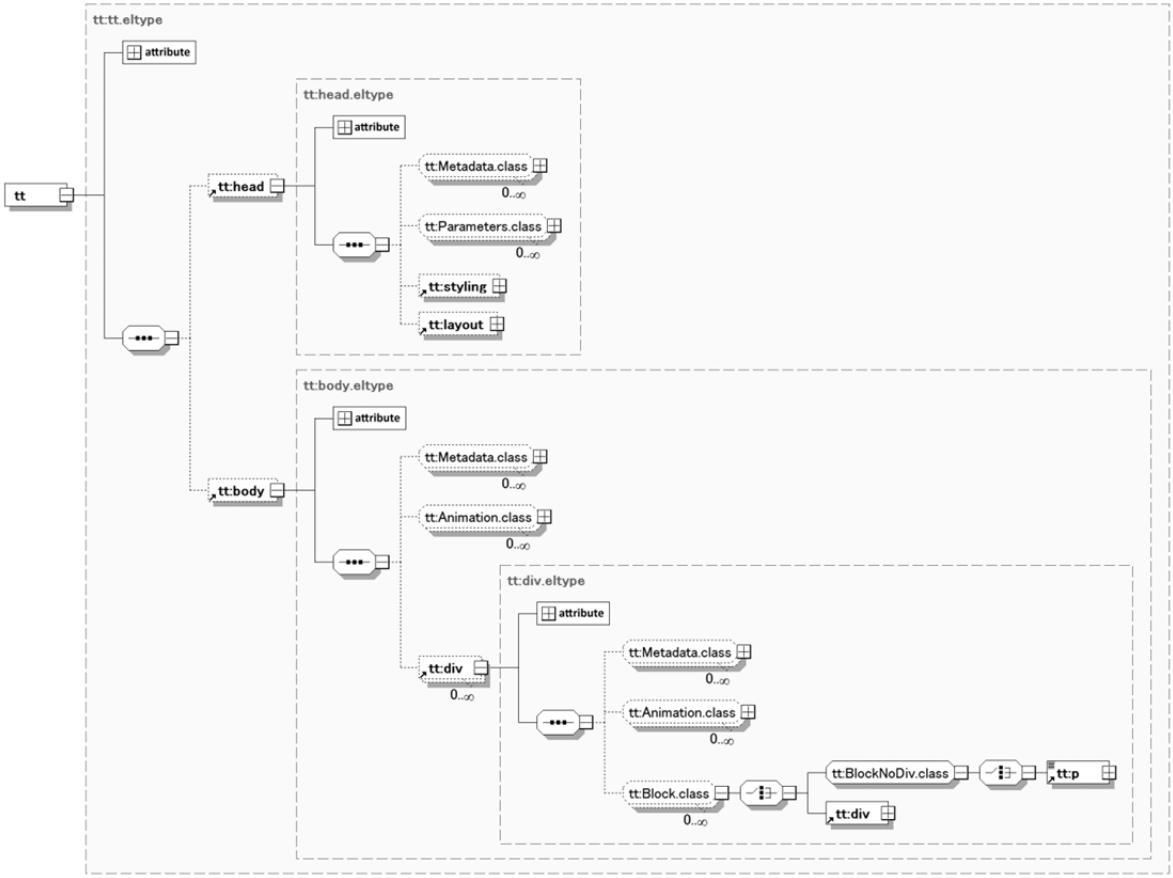


Fig. 3-1: Upper structure of the ARIB-TTML document

tt element is the root element, under which the head element and body element can be placed. The under the head element can be the styling element and the layout element that define layout, etc. of the entire document. div element and p element can be child elements of body element. div element and p element can contain the text and other data provided as closed caption and super impose. In div element and p element, the information for controlling time, such as begin attribute can be described to display closed caption text and other data in a specified time. Each element described above conforms to the TTML1 standard. Elements and attributes belonging to a level lower than the elements described above include extensions by SMPTE-TT standard and this standard. The following section defines the elements and attributes that are extended by this standard.

3.3.1 Profile of the ARIB-TTML

The ARIB-TTML profile identifiers is

- http://www.arib.or.jp/ns/profiles/arib-ttml-full/v1_0

The feature element description and extension element description that are specified in the ARIB-TTML profile definition, which conform to the TTML1 standard, are defined in the operational rules.

3.3.2 Extended elements

3.3.2.1 arib-tt:font-face element

arib-tt:font-face element is an element for specifying a non-built-in font that can be used for Gaiji in an ARIB-TTML document. Zero or more arib-tt:font-face element can be placed in a styling element of an ARIB-TTML document. For the specification related to non-built-in fonts, see Section 3.5.2. Table 3-3 shows syntax of arib-tt:font-face element and Fig. 3-2 shows its structure.

Table 3-3: Syntax of arib-tt:font-face element

```
<xsd:element name="font-face" type="arib:font-faceType"/>

<xsd:complexType name="font-faceType">
    <xsd:sequence>
        <xsd:element name="src" type="arib:font-face-srcType"
                maxOccurs="unbounded"/>
    </xsd:sequence>
    <xsd:attribute name="font-family" type="xsd:string" use="required"/>
    <xsd:attribute name="unicode-range" type="xsd:string"/>
    <xsd:attribute ref="xml:id"/>
</xsd:complexType>

<xsd:complexType name="font-face-srcType">
    <xsd:attribute name="url" type="xsd:anyURI" use="required"/>
    <xsd:attribute name="format" type="xsd:string"/>
    <xsd:attribute ref="xml:id"/>
</xsd:complexType>
```

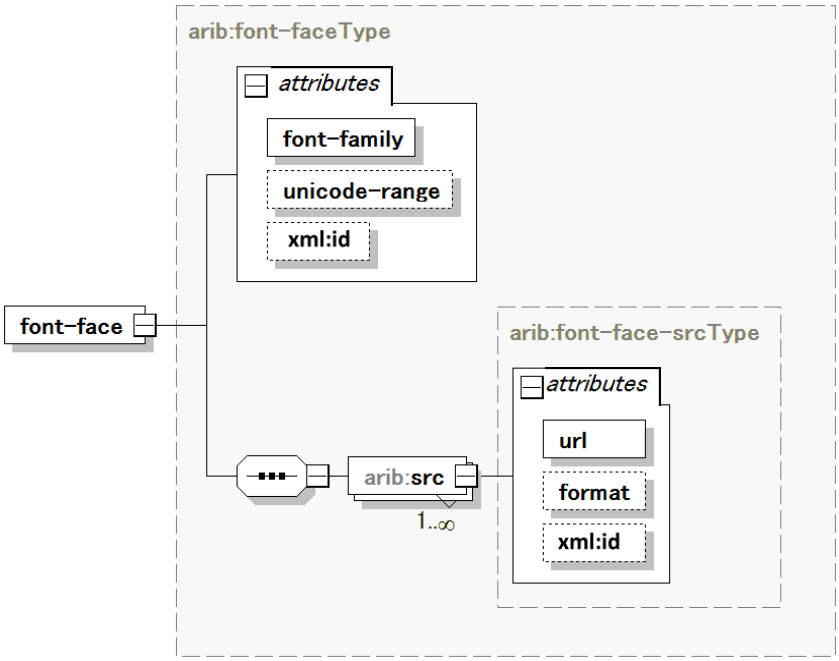


Fig. 3-2: Structure of arib-tt:font-face element

Semantics for arib-tt:font-face element:

font-family: This is a mandatory attribute for a non-built-in font to define a font family name to be called by tts:fontFamily attribute.

unicode-range: This attribute represents the range of character code values that is supported by a font. The following shows description examples.

- unicode-range = "U+A5": The example of applying a range only to a specific character
- unicode-range = "U+0-7F": The example of specifying one range
- unicode-range = "U+A5, U+0-7F, …": The example of specifying multiple ranges

xml:id: A unique identifier inthe ARIB-TTML document

arib-tt:src: This attribute is a parent element to specify url attribute and format attribute. One or more src elements must be defined.

url: This is a mandatory attribute that defines a URL of a font file location.

format: This attribute defines a font file format. Either "svg" or "woff" can be written.

xml:id: A unique identifier in the ARIB-TTML document

3.3.2.2 arib-tt:keyframes element

arib-tt:keyframes element is an element to define animation sequence. Every moment of a sequence of an animation, from the start to the end, is specified by arib-tt:keyframe element which is a child of this element. For instance, if the tts:origin attribute of multiple keyframe elements is changed, the target element moves according to value. The animation defined by this element is called by arib-tt:animation element using a name of the animation which is also defined by this element. Zero or more arib-tt:keyframes element can be placed in styling

element of an ARIB-TTML document. Table 3-4 shows syntax of arib-tt:keyframes element and Fig. 3-3 shows its structure.

Table 3-4: The syntax of the arib-tt:keyframes element

```
<xsd:element name="keyframes" type="arib:keyframesType"/>

<xsd:complexType name="keyframesType">
    <xsd:sequence>
        <xsd:element name="keyframe" type="arib:keyframeType" minOccurs="2"
                maxOccurs="unbounded"/>
    </xsd:sequence>
    <xsd:attribute name="animationName" type="xsd:string" use="required"/>
    <xsd:attribute ref="xml:id"/>
</xsd:complexType>

<xsd:complexType name="keyframeType">
    <xsd:attribute name="position" type="xsd:string" use="required"/>
    <xsd:attribute ref="tts:backgroundColor"/>
    <xsd:attribute ref="tts:color"/>
    <xsd:attribute ref="tts:fontSize"/>
    <xsd:attribute ref="tts:extent"/>
    <xsd:attribute ref="tts:opacity"/>
    <xsd:attribute ref="tts:origin"/>
    <xsd:attribute ref="xml:id"/>
</xsd:complexType>
```

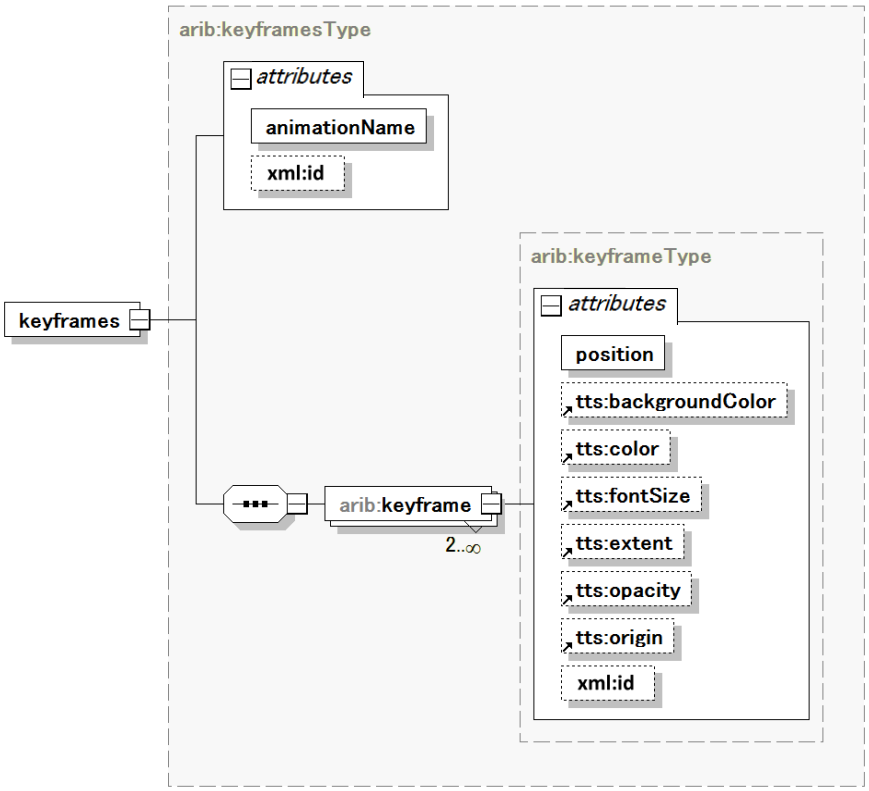


Fig. 3-3: Structure of arib-tt:keyframes element

Semantics of arib-tt:keyframes element:

animationName: an attribute to specify the name of an animation sequence defined by this element. An animation is referenced by specifying the animation name with the arib-tt:animation attribute in an ARIB-TTML document.

xml:id : A unique identifier in the ARIB-TTML document

keyframe: an element to define each step (keyframe) of an animation. The position attribute whose value is set to "0%" and the keyframe element whose value is set to "100%" must be defined. Multiple attribute can be specified with the keyframe element.

position: a percent value that represents a time when an applicable key frame occurs in an animation sequence.

tts:backgroundColor: Background color of an element in which the arib-tt:animation attribute is specified in a moment when an applicable keyframe is presented.

tts:color: Character color of the element in which the arib-tt:animation attribute is defined in a moment when an applicable keyframe is presented.

tts:extent: Width and height of an element in which the arib-tt:animation attribute is defined in a moment when an applicable keyframe is presented.

tts:fontSize: Font size of an element in which the arib-tt:animation attribute is defined in a moment when an applicable keyframe is presented.

tts:opacity: An alpha value for semi-transparent composition of an element in which the arib-tt:animation attribute is defined in a moment when an applicable keyframe is presented.

tts:origin: Position of an element in which the arib-tt:animation attribute is defined in a moment when an applicable keyframe is presented.

xml:id : A unique identifier in the ARIB-TTML document* The detailed behavior of each attribute is defined separately in the operational guidelines.

3.3.2.3 arib-tt:audio element

In ARIB-TTML, arib-tt:audio element can be used to play audio clip and the receiver's built-in sound. This element can be placed as a child element of the div element or p element in an ARIB-TTML document. Before playing audio, the parsing of an ARIB-TTML document is performed. If begin attribute is specified in the div element or p element, the start time specified in the attribute is the time to start playing the audio. If begin attribute is not specified, play of the audio starts at the time when presentation of the ARIB-TTML document starts. For the specification related to audio play, see Section 3.6. Table 3-5 shows syntax of the arib-tt:audio element and Fig. 3-4 shows its structure.

Table 3-5: Syntax of arib-tt:audio element

```
<xsd:element name="audio" type="arib:audioType"/>

<xsd:complexType name="audioType">
    <xsd:attribute name="src" type="xsd:anyURI" use="required"/>
    <xsd:attribute name="loop" type="xsd:boolean"/>
    <xsd:attribute ref="xml:id"/>
</xsd:complexType>
```

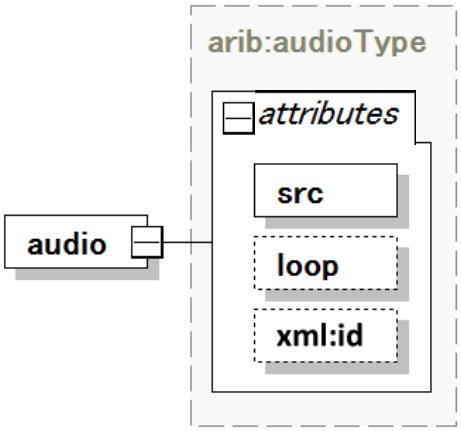


Fig. 3-4: Structure of arib-tt:audio element

Semantics of the arib-tt:audio element:

src: A mandatory attribute that defines a URL indicating the location of an audio file.

loop: A boolean value that specifies whether to repeat replaying a specified audio. If the loop attribute is omitted, the loop is specified as "false".

xml:id : A unique identifier in the ARIB-TTML document

3.3.3 Extended attribute

3.3.3.1 arib-tt:animation

arib-tt:animation attribute is used when an animation that is defined using the arib-tt:keyframes element is applied to the element in which this attribute is defined. This attribute can be specified to style element, body element, div element, p element and span element. Table 3-6 shows the syntax of the arib-tt:animation attribute.

Table 3-6: Syntax of arib-tt:animation attribute

```
<xsd:attribute name="animation" type="xsd:string"/>
```

Semantics of the arib-tt:animation attribute:

animation: A character string that specifies the following six values collectively in the following order by separating them with a space: <animation-name>, <animation-duration>, <animation-timing-function>, <animation-delay>, <animation-iteration-count> and

<animation-direction>.

(Example: arib-tt:animation = "myAnimation1 1000ms linear 0ms 1 normal")

・<animation-name>: A name of an animation sequence defined in an arib-tt:keyframes element.

・<animation-duration>: Duration time taken to play an animation once in milliseconds (ms).

・<animation-timing-function>: This specifies how an animation progress. The following value can be specified.

"ease": The progress follows three-dimensional bezier curve, cubic-bezier (0.25, 0.1, 0.25, 1.0) (this enables an animation to start and end in a smooth manner).

"linear": Animation changes at a constant speed from the first to the last.

"ease-in": The progress follows cubic-bezier (0.42, 0, 1.0, 1.0) (start slowly).

"ease-out": The progress follows cubic-bezier (0, 0, 0.58, 1.0) (end slowly).

"ease-in-out": The progress follows cubic-bezier (0.42, 0, 0.58, 1.0) (start and end slowly).

"step-start": At the time for the start keyframe, the end keyframe is animated, and held until the time for the end keyframe. This is equivalent to steps(1,start).

"step-end": By the time for the end keyframe, the start keyframe is animated. At the time for the end keyframe, the end keyframe is animated. This is equivalent to steps(1,end).

"steps(<number>,[start | end])" : This animates keyframes by a number that divides the domain of operation into equally sized intervals. An argument <number> specifies the number of intervals in the function. It must be a positive integer (greater than 0). The second argument, which is optional, is either the value *start* or *end*, and specifies the point at which the progress of animation occur within the interval. If the second argument is *start*, the animation progresses at the beginning of the interval. If it is *end*, the animation progresses at the end of the interval.

・<animation-delay>: A value how much time is required for delaying the start of an animation. 0 means no delay.

・<animation-iteration-count>: specifies value to specify the number of repeats of the animation.

・<animation-direction>: specifies value to specify whether to play an animation by alternatively reversing when repeating the animation. The following value can be specified.

"normal": Repeat an animation by replaying in the normal direction.

"alternate": Repeat an animation by replaying in the normal direction when the number of repetitions is odd and in the reverse direction when the number of repetitions is even.

3.3.3.2 arib-tt:border

The arib-tt:border attribute is an attribute used to specify the style, thickness and color of the enclosing border of four sides of a sequence of characters. To specify the character borders in left, right, up and down independently, arib-tt:border-top attribute, arib-tt:border-bottom attribute, arib-tt:border-left attribute, arib-tt:border-right should be specified separately. These attributes can be used in style element, body element, div element, p element and span element. Table 3-7 shows syntax of arib-tt:border attribute.

Table 3-7: Syntax of arib-tt:border attribute

```
<xsd:attribute name="border" type="xsd:string"/>
<xsd:attribute name="border-top" type="xsd:string"/>
<xsd:attribute name="border-bottom" type="xsd:string"/>
<xsd:attribute name="border-left" type="xsd:string"/>
<xsd:attribute name="border-right" type="xsd:string"/>
```

Semantics of arib-tt:border attribute:

border, border-top, border-bottom, border-left, border-right: A character string that specifies the following three values collectively in the following order by separating them with a space: < border-style>, <border-width> and <border-color>. “border” specifies the same borders which are surrounding character(s). “border-top”, “border-bottom”, “border-left” and “border-right” are used to specify corresponding border independently. (Example: arib-tt: border = "solid 3px white")

・< border-style >: A value to specify the border style. The following value can be specified.

"none": Border is not displayed and the thickness becomes 0. If another border is overlapping, the overlapped border is displayed.

"hidden": Border is not displayed and the thickness becomes 0. If another border is overlapping, the overlapped border is not displayed neither.

"solid": The border is displayed with a single line.

"double": The border is displayed with two lines.

"groove": The border is displayed as if it were curved into the background.

"ridge": The border is displayed as if it were rising out of the background.

"inset": The border is displayed as if it were embedded in the background by displaying the top and left sides in a darker shade and the lower and right sides in a lighter shade.

"outset": The border is displayed as if it were rising out of the background by displaying the top and left sides in a lighter shade and the lower and right sides in a darker shade.

"dashed": The border is displayed with a dashed line.

"dotted": The border is displayed with a dotted line.

・< border-width >: A value to specify thickness of the border in a pixel (px).

・< border-color >: A value to specify border color. The method to specify a color conforms to Section 3.5.5.

3.3.3.3 arib-tt:letter-spacing

arib-tt:letter-spacing attribute is to specify the space between letters. This attribute can be used for style element, body element, div element, p element and span element. Table 3-8 shows syntax of arib-tt:letter-spacing attribute.

Table 3-8: Syntax of arib-tt:letter-spacing attribute

```
<xsd:attribute name="letter-spacing" type="xsd:string"/>
```

Semantics of the arib-tt:letter-spacing attribute:

letter-spacing: A value to specify an additional space to the default space between letters in pixel (px).

3.3.3.4 arib-tt:marquee

arib-tt:marquee attribute is used to automatically scroll an element and to specify style, direction, speed, and the number of repetitions collectively. If an image is included in the element in which the arib-tt:marquee attribute is defined, the image is scrolled together. This element can be specified as the attribute of the style element, body element, div element, p element and span element. Table 3-9 shows syntax of arib-tt:marquee attribute.

Table 3-9: Syntax of arib-tt:marquee attribute

```
<xsd:attribute name="marquee" type="xsd:string"/>
```

Semantics of arib-tt:marquee attribute:

marquee:A character string that specifies the following four values collectively in the following order by separating them with a space: < marquee-style>, <marquee-direction>, <marquee-speed> and <marquee-play-count>.

(Example: arib-tt:marquee = "scroll forward normal 1")

・<marquee-style>: A value to specify how the scroll is moved. The following value can be specified.

"scroll": Scroll content of the element entirely, from beginning to end.

"slide": Scroll content of element from beginning to end, and then stop.

"alternate": Scroll content of element from beginning to end entirely, and then scroll backward to beginning.

・< marquee-direction>: A value to specify scroll direction. The following value can be specified.

"forward": When using horizontal writing, scroll the element "from left to right". When using vertical writing, scroll the element "from top to bottom".

"reverse": When using horizontal writing, scroll the element "from right to left". When using vertical writing, scroll the element "from bottom to top".

・<marquee-speed >: This specifies the speed of scroll movement. "slow", "normal" or

"fast" can be specified. The specific speed is defined in the operational guidelines.

・<marquee-play-count >: This specifies the number of times of scrolls. A numerical value or "infinite" can be specified.

3.3.3.5 arib-tt:ruby

The arib-tt:ruby attribute specifies that a specified element is ruby. And this attribute associates the specified element to an element that includes a character string, which is subject to ruby, in the same ARIB-TTML document. For instance, a span element contains character string for ruby and has arib-tt:ruby attribute with value of 'ruby_1'. This element, span1, is an element for ruby. Another span element, span2, contains character string which is subject to ruby and has xml:id attribute with value of 'ruby_1'. The span1 element is associated to as ruby for span2 by reference xml:id and arib-tt:ruby attributes in these elements with the same value. arib-tt:ruby attribute can be specified to div element, p element and span element. Table 3-10 shows syntax of the arib-tt:ruby attribute.

Table 3-10: Syntax of arib-tt:ruby attribute

```
<xsd:attribute name="ruby" type="xsd:string"/>
```

Semantics of arib-tt:ruby attribute:

ruby: A value of xml:id attribute of an element that is subject to ruby.

3.3.3.6 arib-tt:text-shadow

arib-tt:text-shadow attribute is used when a shadow is drawn around letters. This attribute collectively specifies horizontal position, vertical position, blur distance and the color of the shadow. This attribute can be specified to style element, body element, div element, p element and span element. Table 3-11 shows syntax of the arib-tt:text-shadow attribute.

Table 3-11: Syntax of arib-tt:text-shadow attribute

```
<xsd:attribute name="text-shadow" type="xsd:string"/>
```

Semantics of the arib-tt:text-shadow attribute:

text-shadow: This is a character string that specifies the following four values collectively in the following order by separating them with a space: < offset-x>, <offset-y>, < blur-radius> and <color>. (E.g.: arib-tt:text-shadow = "5px 5px 2px blue")

・<offset-x>: This specifies the position of the horizontal shadow in a pixel value (px). The shadow moves right when a positive value is specified. The shadow moves left when a negative value is specified.

・<offset-y>: This specifies the position of the vertical shadow in a pixel value (px). The shadow moves downward when a positive value is specified. The shadow moves upward when a negative value is specified.

・<blur-radius>: This specifies the radius of the shadow's blur in a pixel

value (px).

・<color>: This specifies the color of the shadow. For the method of specifying a color, follow the specification described in Section 3.5.5.

3.4 Referencing Resources from the ARIB-TTML document

3.4.1 Referencing resources transmitted over broadcast

For the case when the ARIB-TTML document and a resource that is referenced from the ARIB-TTML document are broadcasted in the same service, resource reference by URL shown in Table 3-12 should be used to specify the resource from the ARIB-TTML document.

Table 3-12: Resource reference used by an ARIB-TTML document for a resource in MMT broadcast signal

subt://<subsample_index>

Semantics of **<subsample_index>**:

This is the sub-sample number of a resource in the same closed caption data transmission unit. An ARIB-TTML document is always placed at the beginning (sub-sample number = 0). The sub-sample number, which is placed for the second sub-sample (sub-sample number 1) or later, is specified.

A resource that can be referenced using the method above is limited to that transmitted in the same closed caption data transmission unit as the referencing ARIB-TTML document. The resource is valid during after receiving the applicable transmission unit and before receiving the next transmission unit. Once the next transmission unit including the ARIB-TTML document has been received, a resource included in the previous transmission unit is removed from the cache of the receiver and the availability of referenced resource is not guaranteed. MPU is used as a closed caption data transmission unit in MMT.

3.4.2 Referencing receiver's built-in sound

Resource reference shown in Table 3-13 should be used to play the receiver's built-in sound by audio element in an ARIB-TTML document.

Table 3-13: Resource reference used by an ARIB-TTML document for receiver's built-in sound

romsound://**<sound_id>**

Semantics of **<sound_id>**:

The number of the sound source of the receiver's built-in sound (an integer value greater than or equal to 0) is specified.

3.5 Font

3.5.1 Font-family

If round gothic, round gothic bold or gothic is specified as the font-family attribute in an ARIB-TTML document, letters must be displayed in an appropriate corresponding font style. If other values for font-family are specified, font style selected by such values is up to receiver implementation. For Latin fonts, letters are shown in proportional font if sans-serif is specified as the generic-family, and letters are shown in fixed-width font if monospace is specified.

3.5.2 Specifying non-built-in font including Gaiji

If the arib-tt:font-face element defined in Section 3.3.2.1 is specified, the receiver uses a specified non-built-in font. In broadcasting signal, a font file simultaneously transmitted with an ARIB-TTML document is used, which is referenced in a way defined in Section 3.4.1. Available formats of font file are in accordance with volume, Part 2, Section 5.5 of this standard.

Table 3-14 shows an example ARIB-TTML document to display Gaiji by a non-built-in font. In this example, the non-built-in font file is transmitted in a unit referenced as the sub-sample number 1 in the same transmission unit with the ARIB-TTML document. The file is referenced with "subt: //1".

Table 3-14: Example ARIB-TTML document for Gaiji using non-built-in font

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns="http://www.w3.org/ns/ttml"
    xmlns:arib-tt="http://www.arib.or.jp/ns/arib-ttml/v1_0"
    xmlns:smpte="http://www.smpte-ra.org/schemas/2052-1/2013/smpte-tt"
    xml:lang="ja">
  <head>
    <styling>
      <arib-tt:font-face font-family="extChar1" unicode-range="U+E000">
        <arib-tt:src url="subt://1" format="svg"/>
      </arib-tt:font-face>
    </styling>
  </head>
  <body>
    <div>
      <p begin="00:00:03" end="00:00:06">display supplemental characters &#xe000;</p>
    </div>
  </body>
</tt>
```

3.5.3 Font-size

A font size is specified by font-size attribute in a pixel (px) unit. A single value or two values separated with a space can be specified. When a single value is specified, the value indicates the font height. When two values are specified, the former indicates the font width and the latter indicates the font height.

3.5.4 Font-weight

A font weight is specified by font-weight attribute. If normal or bold is specified, actual font weight is defined in the operational guidelines.

3.5.5 Color

A font color is specified by color attribute by 256 steps of red, green, blue and transparency. A font color can be specified for each character. A font color can be specified using a color name, #rrggbb (red, green and blue in 2-digit hexadecimals), or #rrggbbaa (red, green, blue, and transparency in 2-digit hexadecimals). The available color names are defined separately in the operational guidelines.

3.6 Audio playback

If arib-tt:audio element defined in Section 3.3.2.3 is specified, audio clip and the receiver's built-in sound are played while effective valid ARIB-TTML document is presented.

3.6.1 Audio clip file

In broadcasting, if an audio clip file and an ARIB-TTML document are transmitted simultaneously and the ARIB-TTML document refers the audio clip file using the resource reference defined in Section 3.4.1, a relevant audio clip file is played. Available coding schemes for audio clip are PCM (AIFF-C), MP3, MPEG-2 AAC and MPEG-4 AAC defined in Volume 1, Part 2 of this standard.

3.6.2 Receiver's built-in sound

If receiver's built-in sound is referenced using resource reference defined in Section 3.4.2, relevant built-in sound is played. The sound_id value, which is defined separately in the operational guidelines, is used to select sound.

3.7 Displaying images

Images should be able to be displayed using the smpte:backgroundImage attribute defined in the SMPTE-TT standard [2] as described below.

3.7.1 Images embedded in an ARIB-TTML document

Image data can be embedded in an ARIB-TTML document using the smpte:image element that is defined in the SMPTE-TT standard. By referencing an id attribute of the smpte:image element, smpte:backgroundImage attribute enables to render it. Coding scheme of such embedded images is PNG defined in Volume 1, Part 2 of this standard. Table 3-15 shows an example of an ARIB-TTML document in which image data is embedded. In this example, "Img1" is an id value of smpte-image element which contains image data and the data is referenced by an smpte:backgroundImage attribute using "#Img1".

Table 3-15: An example ARIB-TTML document for embedded image data in the document

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns="http://www.w3.org/ns/ttml"
      xmlns:arib-tt="http://www.arib.or.jp/ns/arib-ttml/v1_0"
      xmlns:smpte="http://www.smpte-ra.org/schemas/2052-1/2013/smpte-tt"
      xml:lang="ja">
    <head>
        <metadata>
            <smpte:image xml:id="Img1" imageType="PNG" encoding="Base64">
                [Image data encoded in Base64]
                ・・・・・
            </smpte:image>
        </metadata>
    </head>
    <body>
        <div>
            <div begin="00:00:05" end="00:00:10" smpte:backgroundImage="＃Img1">
                <p>image display</p>
            </div>
        </div>
    </body>
</tt>
```

3.7.2 Reference to an external image resource of ARIB-TTML document

In broadcasting using the MMT, if image files transmitted simultaneously with an ARIB-TTML document are referenced by the smpte:backgroundImage attribute in the ARIB-TTML document using the resource reference defined in Section 3.4.1, content of the image files is displayed. Available coding schemes for such image files are PNG and SVG defined in Volume 1, Part 2 of this standard. Table 3-16 shows an example ARIB-TTML document for displaying an image in an external resource. In this example, an image file is transmitted in the same transmission unit and allocated as sub-sample number 1. The file is referenced using "subt: //1".

Table 3-16: Example ARIB-TTML document for displaying an image in an external resource

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns="http://www.w3.org/ns/ttml"
      xmlns:arib-tt="http://www.arib.or.jp/ns/arib-ttml/v1_0"
      xmlns:smpte="http://www.smpte-ra.org/schemas/2052-1/2013/smpte-tt"
      xml:lang="ja">
    <head/>
    <body>
        <div>
            <div begin="00:00:05" end="00:00:10" smpte:backgroundImage="subt://1">
                <p>image display</p>
            </div>
        </div>
    </body>
</tt>
```

3.8 Clearing Screen

If an empty ARIB-TTML document (<tt></tt>) is received, it is considered as a clear screen command and therefore all closed captions that are being presented are cleared.

3.9 Default Value of the ARIB-TTML Document

The default values of attribute value, style and layout in an ARIB-TTML document are defined separately in the operational guidelines.

Chapter 4: Initialization

Table 4-1 shows the control information and initialization. Table 4-2 shows the initial status after initialization.

Table 4-1: Control information and initialization

Initialization operation item / Control information		Display screen	Audio play	Image	Non-built-in font
Updating the additional identification information of closed captions *1		X	X	X	X
Closed Caption data	Updating the ARIB-TTML document (live mode) *2	X *3	X *3	X *3	X *3
	Updating the ARIB-TTML document (any mode other than live mode) *4	X	X	X	X
	Clearing screen (empty element)	X	X	X	X

Note)*1: In MMT, this information is placed in the MH-data coding scheme descriptor. In MPEG2-TS, this information is placed in the data coding scheme descriptor or DII message.

*2: This mode is engaged when the OPM (operation mode) value is "00" in the additional dentification informationof closed caption and super imposetransmission in MH-data coding scheme descriptor.

*3: Initialization is not performed only when "indefinite" is specified in the end attribute or dur attribute of an element in the ARIB-TTML document. For more detail, see Description 1e.

*4: Even if "indefinite" is specified in the end attribute or the dur attribute of an element in an ARIB-TTML document, initialization operation is performed and presentation is not continued.

Table 4-2: Initial status

Item	Initial status
Display screen	The screen is cleared. The presentation time control is in a state that is passed the time to finish presentation.
Audio play	Not playing. No data is held for audio clip.
Image	No data is held.
Non-built-in font	Reference to a non-built-in font is released and no data is held.

Chapter 5: Transmission of Closed Caption and Super impose

5.1 Transmission of Closed Caption and Super impose Using MMT

The transmission scheme of the closed captions and super impose using MMT is defined in ARIB STD-B60, Chapter 9.

5.2 Transmission of Closed Caption and Super impose Using MPEG-2 TS

The transmission method of the closed caption and super impose using the MPEG-2 TS is defined in ARIB STD-B24, Part 3. Chapter 10.

Description 1: Example of continuous presentation across multiple TTML documents in live mode

TTML is a data format that describes closed caption data for one video. Since one TTML document comprises of a complete data set, there is no relationship between multiple TTML documents. As an exception, the specification for the fragmented streaming of TTML is described as the Non-Normative (optional) specification in Annex L to TTML1 Recommendation. However, this optional specification does not support the case in which a set of presented closed caption survive across multiple TTML documents. Therefore, this standard introduced a mechanism to continuously present closed caption across multiple TTMLs. This section explains the mechanism.

Figure D1-1 shows an example in which continuous display of closed captions is required. In the example, 'String 4' in TTML-2 is the closed caption that needs to be presented until 50.856 seconds. However, the TTML document needs to be separated in order to transmit 'String5' in TTML-3. This example is assumed that it is in live broadcast, and the transmission of 'String5' has not been determined when transmitting TTML-2. In this case, change of TTML documents from TTML-2 to TTML-3 takes switching time and thus presentation of the closed captions will be interrupted temporarily. This situation is inevitable because even if the end attribute of the p element for 'String4' in TTML-2 is not specified, the screen must be cleared for presenting TTML-3 unless a TTML processor is instructed not to clear.

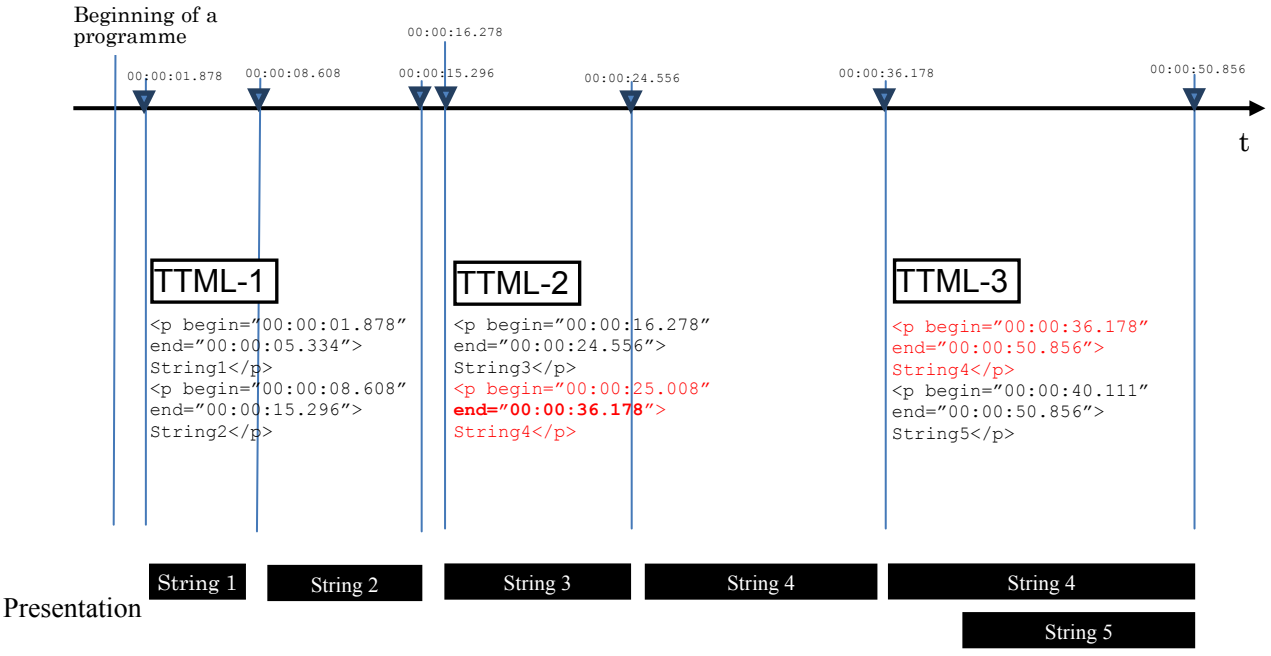


Fig. D1-1: A case which requires continuous presentation of the closed caption across multiple TTML documents

For this purpose, the constant, "indefinite" in the end attribute (or dur attribute) of the p element can be used to instruct the TTML processor to continue presenting content in the p element.

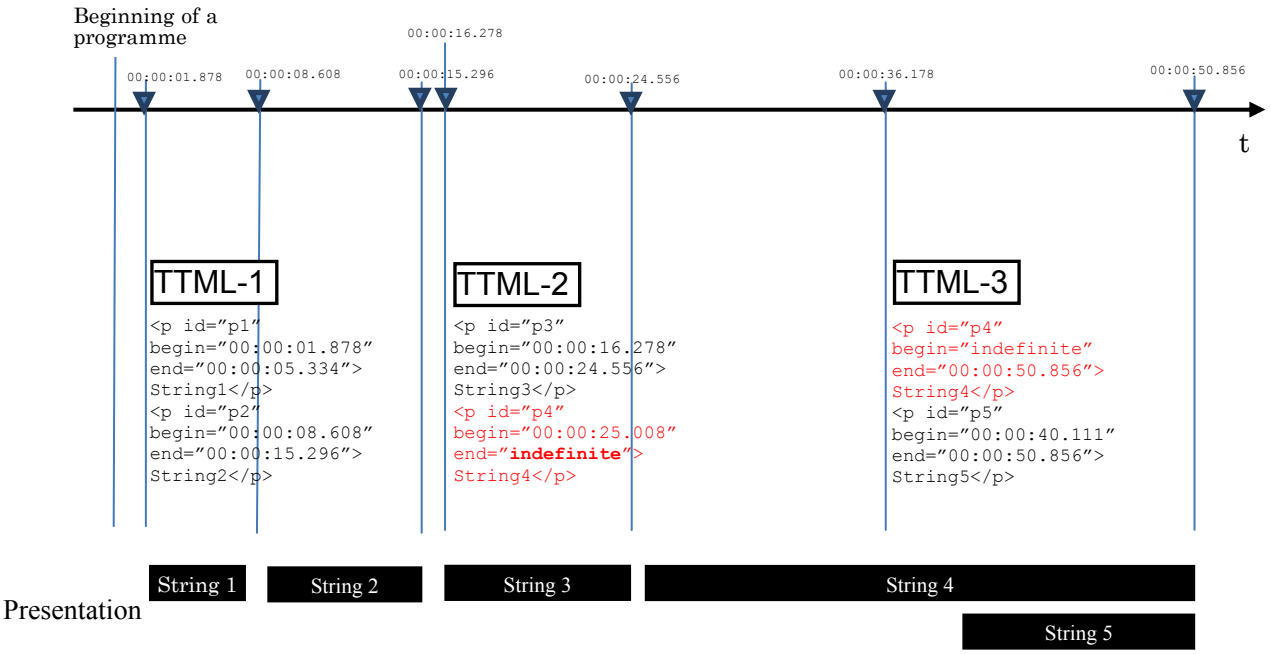


Fig. D1-2: A control of continuous presentation of the closed caption across multiple TTML documents

Figure D1-2 shows an example of specifying "indefinite" in the end attribute. As shown in the example, if there are closed captions that need to be presented continuously, "indefinite" can be specified as time description of the p element for 'String4' in TTML-2. In TTML-3 that is transmitted next, "indefinite" is also specified in the begin attribute of the element of the closed captions that are presented continuously. Relationship between these two elements in TTML-2 and TTML-3 is indicated by the same id attribute. Use of 'indefinite' in begin attribute and end attribute, and the use of the same value in id attribute enables to continuous presentation of 'String4'. A TTML processor parses a new TTML document, and replaces content of the node by existing content of the node which has the same id attribute. Please note that, as shown in Figure D1-2, the value of the applicable end attribute of the replaced p element in the DOM tree is overwritten with the value in a new TTML document, if the end attribute value of the p element is determined, as seen in the p element for 'String4' in TTML-3. Other DOM node values of the applicable p element are kept as those in TTML-2. It is preferred that other the DOM node values of an applicable p element are kept as those in TTML-2, regardless the sameness of property values of the element with the same id attribute, including style such as the presentation area, between TTML-2 and TTML-3. In addition, if "indefinite" is not set in the begin attribute of an element in TTML-3, initialization is performed soon after the parsing of TTML-3 is complete and the p element's begin attribute of the applicable id is determined. As a result, the closed captions are not presented continuously. If "indefinite" is not specified in any end attribute (or the dur attribute) of TTML-2, a receiver performs the initialization process when it starts receiving TTML-3.

<Blank Page>

MULTIMEDIA CODING SPECIFICATION FOR DIGITAL BROADCASTING (SECOND GENERATION) ARIB STANDARD

ARIB STD-B62 VERSION 1.6-E1 (Fascicle 1)
(July 27, 2017)

This Document is based on the ARIB standard of “MULTIMEDIA CODING SPECIFICATION FOR DIGITAL BROADCASTING (SECOND GENERATION)” in Japanese edition and translated into English inSeptember, 2017.

Published by

Association of Radio Industries and Businesses

11F, Nittochi Building
1-4-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-0013, Japan

TEL 81-3-5510-8590
FAX 81-3-3592-1103

Printed in Japan