ENGLISH TRANSLATION

VIDEO CODING, AUDIO CODING, AND MULTIPLEXING SPECIFICATIONS FOR DIGITAL BROADCASTING

ARIB STANDARD

ARIB STD-B32 Version 3.9

Established May 31, 2001	Version 1.0
Revised July 27, 2001	Version 1.1
Revised January 24, 2002	Version 1.2
Revised March 28, 2002	Version 1.3
Revised June 5, 2003	Version 1.4
Revised February 5, 2004	Version 1.5
Revised May 25, 2004	Version 1.6
Revised September 28, 2004	Version 1.7
Revised December 14, 2004	Version 1.8
Revised March 14, 2006	Version 1.9
Revised September 28, 2006	Version 2.0
Revised March 14, 2007	Version 2.1
Revised July 29, 2009	Version 2.2
Revised April 26, 2010	Version 2.3
Revised November 5, 2010	Version 2.4
Revised March 28, 2011	Version 2.5
Revised July 3, 2012	Version 2.6
Revised September 25, 2012	Version 2.7
Revised December 18, 2012	Version 2.8
Revised March 18, 2014	Version 2.9
Revised July 31, 2014	Version 3.0
Revised December 16, 2014	Version 3.1
Revised March 17, 2015	Version 3.2
Revised July 3, 2015	Version 3.3
Revised September 30, 2015	Version 3.4
Revised December 3, 2015	Version 3.5
Revised March 25, 2016	Version 3.6
Revised July 6, 2016	Version 3.7
Revised September 29, 2016	Version 3.8
Revised December 9, 2016	Version 3.9

Association of Radio Industries and Businesses

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

1. Notes on Copyright

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 "VIDEO CODING, AUDIO CODING, AND MULTIPLEXING SPECIFICATIONS FOR DIGITAL BROADCASTING". 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 Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Japan Broadcasting Corporation (NHK)	ﾃﾞｨｼﾞﾀﾙ情報伝送方式、ﾃﾞｨｼﾞﾀﾙ情報送信装置およびﾃﾞｨｼﾞﾀﾙ情報受信装置	特願平 05-65183 特開平 06-276169	Japan
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.1.*[15]		
NEC Corporation	画像信号の動き補償フレーム間予測符号化・復号化方法とその装置	特許 1890887	Japan
	画像の圧縮記録システム	特許 2036887	Japan, United States, United Kingdom, Germany, France, Netherlands, Canada
	適応変換符号化の方法及び装置	特許 2569842	Japan, United States, United Kingdom, Germany, France, Netherlands
	適応変換符号化の方法及び装置	特許 2778161	Japan, United States, United Kingdom, Germany, France, Netherlands
	適応変換符号化の方法及び装置	特許 2569849	Japan, United States, United Kingdom, Germany, France, Netherlands
	適応変換符号化復号化の方法及び装置	特許 2638208	Japan, United States, United Kingdom, Germany, France, Netherlands
	符号化方式及び復号方式	特許 2820096	Japan, Korea, Australia
	改良 DCT の順変換計算装置および逆変換計算装置	特許 3185214	Japan, United States, United Kingdom, Germany, France, Netherlands, Canada

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
NEC Corporation	適応変換符号化方式および適応変換復号方式	特許 3255022	Japan, United States, United Kingdom, Germany, France, Netherlands, Italy, Sweden, Canada, Australia, Korea
	変換符号化方法及び装置	特許 3444261	Japan
	適応変換符号化の方法及び装置	特許 2890522	
	適応変換符号化の方法及び装置	特許 2890523	
NEC Corporation & Matsushita Electric Industrial Co., LTD. *1 (Joint application)	オーディオ復号装置と復号方法およびプログラム	特許 3579047	Japan, United States, United Kingdom, Germany, France, Netherlands, Italy, Sweden, Finland, Canada, Korea, Taiwan, China, Brazil, Hong Kong, India, Hungary, Czech, Spain
	オーディオ復号化装置およびオーディオ復号化方法	特許 3646938	Japan, United States, United Kingdom, Germany, France, Netherlands, Italy, Sweden, Finland, Canada, Korea, Taiwan, China, Brazil, Hong Kong, India, Hungary, Czech, Spain
	オーディオ復号装置およびオーディオ復号方法	特許 3646939	Japan, United States, United Kingdom, Germany, France, Netherlands, Italy, Sweden, Finland, Canada, Korea, Taiwan, China, Brazil, Hong Kong, India, Hungary, Czech, Spain

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Matsushita Electric Industrial Co., LTD.	画像信号のﾌﾚｰﾑ間内挿符号化方法とその装置	特許 1,949,701	Japan, (MPEG Essential Patent)
	動き補償予測方法とそれを用いた画像信号符号化方法	特許 2,699,703	Japan, (MPEG Essential Patent)
	画像信号符号化装置と画像信号復号化装置及び画像信号符号化方法と画像信号復号化方法	特許 2,695,244	Japan, (MPEG Essential Patent)
	画像符号化方法及び画像符号化装置	特許 2,684,941	Japan, (MPEG Essential Patent)
Panasonic Corporation	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.0.*[14]		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.6.*[18]		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.7.*[19]		
Sony Corporation	音声信号圧縮方法及びメモリ書き込み方法	特許 1952835	Japan
	オーディオ信号処理方法	特許 3200886	Japan, United States, United Kingdom, Germany, France, Austria, Australia, Korea, Hong Kong
	オーディオ信号処理方法	特許 3141853	Japan, United States, United Kingdom, Germany, France, Austria, Australia, Korea, Hong Kong

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	信号符号化又は復号化装置、及び信号符号化又は復号化方法、並びに記録媒体	WO94/28633	Japan, United States, United Kingdom, Germany, France, Netherlands, Austria, Italy, Spain, Canada, Australia, Korea, China
Sony Corporation	信号符号化方法及び装置、信号復号化方法及び装置、並びに信号記録媒体	特開平 7-168593	Japan, United States, United Kingdom, Germany, France, Korea, Taiwan, China, Malaysia, Indonesia, India, Thailand, Mexico, Turkey
	符号化データ復号化方法及び符号化データ復号化装置	特許 2874745	Japan, Hong Kong, Korea, United States, Germany, France, United Kingdom
	映像信号符号化方法	特許 2877225	Japan, Hong Kong, Korea, United States, Germany, France, United Kingdom
	符号化データ編集方法及び符号化データ編集装置	特許 2969782	Japan, Hong Kong, Korea, United States, Germany, France, United Kingdom
	動画像データエンコード方法および装置、並びに動画像データデコード方法および装置	特許 2977104	Japan, United States
	動きベクトル伝送方法及びその装置並びに動きベクトル復号化方法及びその装置	特許 2712645	Japan, Australia, Canada, Korea, United States, Germany, France, United Kingdom

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	画像情報符号化装置及び方法、並びに画像情報復号装置及び方法*8	特開 2005-039743	Japan, Brazil, China, Germany, France, United Kingdom, Indonesia, India, Korea, Mexico, Russia, United States, Viet Nam
Sony Corporation	信号処理装置および方法、並びにプログラム*8	特許第 3800427	Japan, China, Germany, France, United Kingdom, Indonesia, India, Korea, Malaysia, Netherlands, Singapore, Thailand, Taiwan, United States
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.0.*6		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.1.*7		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver2.9.*13		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.0.*14		
Motorola Japan Ltd.	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.5.*1		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.6.*2		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.7.*3		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.8.*4		
Philips Japan Ltd.	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.5.*1		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.6.*2		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.7.*3		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.8.*4		

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Mitsubishi Electric Corporation	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.1.*7		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver1.9.*5		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver2.2.*8		
	画像符号化装置、画像符号化方法、画像復号装置及び画像復号方法*14	PCT/JP2014/ 003107	WO
Nippon Telegraph and Telephone Corporation	ディジタル信号処理方法､その処理器､そのプログラム、及びそのプログラムを格納した記録媒体*9	特許 3871672	Japan, United States, United Kingdom France, Germany, Italy, China
	浮動小数点形式ディジタル信号可逆符号化方法、及び復号化方法と、その各装置、その各プログラム*9	特許 4049791	Japan, United States, United Kingdom France, Germany, Italy, China
	浮動小数点形式ディジタル信号可逆符号化方法、及び復号化方法と、その各装置、その各プログラム*9	特許 4049792	Japan, United States, United Kingdom France, Germany, Italy, China
	浮動小数点信号可逆符号化方法、復号化方法、及びそれらの装置、プログラム及びその記録媒体*9	特許 4049793	Japan, United States, United Kingdom France, Germany, Italy, China
	多チャネル符号化方法、復号化方法、これらの装置、プログラムおよびその記録媒体*9	特許 3886482	Japan
	多チャネル信号符号化方法、多チャネル信号復号化方法、それらの方法を用いた装置、プログラム、および記録媒体*9	特許 4348322	Japan
	情報符号化方法、復号化方法、共通乗数推定方法、これらの方法を利用した装置、プログラム及び記録媒体*9	特許 4324200	Japan, United States, China
	情報圧縮符号化装置､その復号化装置､これらの方法、及びこれらのプログラムとその記録媒体*9	特許 4328358	Japan, United States, China
	信号の符号化装置、復号化装置、方法、プログラム、記録媒体、及び信号のコーデック方法*9	特許 4359312	Japan, United States, China

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	動画像の輝度変化補償方法、動画像符号化装置、動画像復号装置、動画像符号化もしくは復号プログラムを記録した記録媒体および動画像の符号化データを記録した記録媒体*14	特許第 2938412	Japan
Nippon Telegraph and Telephone Corporation	動画像符号化方法、動画像復号方法、画像符号化装置、画像復号装置、動画像符号化プログラム、動画像復号プログラムおよびそれらのプログラムの記録媒体*14	特許第 3866628	Japan
Nippon Telegraph and Telephone Corporation & The University of Tokyo (Joint application) *9	多チャネル信号符号化方法、その復号化方法、これらの装置、プログラム及びその記録媒体	特許 4461144 (特願 2006-531829)	Japan, United States, China
	長期予測符号化方法、長期予測復号化方法、これら装置、そのプログラム及び記録媒体	特許 4469374 (特願 2006-552928)	Japan, United States, China
Nippon Telegraph and Telephone Corporation & TODAI TLO, Ltd. (Joint application) *9	多チャネル信号符号化方法、その復号化方法、これらの装置、プログラム及びその記録媒体	特許 4374448	Japan, United States, China
QUALCOMM Incorporated	Adaptive filter*10	JP 3771275	US 6,724,944; US 7,242,815; DE;EP;FI;FR;GB; HK;JP;NL
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver2.3.*9		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver2.4.*11		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver2.5.*12		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.0.*14		

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Parameter Selection in Data Compression and Decompression*16	JP4819361	US7,593,582; US7,388,993; US6,975,773; CN;EP;HK;IN; KR;MX;TH;TW
	Pixel-by-pixel weighting for intra-frame coding*16	JP5372911	US8,238,428; US20120300835; CN;IN;KR;TW
QUALCOMM Incorporated	Mode uniformity signaling for intra-coding*16	JP5096561	US8,488,672; CN; EP; IN; KR; TW
	Adaptive coding of video block prediction mode*16	JP5254324	US8,428,133; US8,520,732; JP; AT; BE; BR; CA; CH; CN; DE; DK; EP; ES; FI; FR; GB; GR; HU; IE; IN; IT; KR; NL; NO; PL; PT; RO; RU; SE; TW
	Filtering video data using a plurality of filters*16	JP5650183	US20100008430; JP; BR; CA; CN; EP; HK; IN; KR; RU; SG; TW
	Non-zero rounding and prediction mode selection techniques in video encoding*16	JP2012-533225	US20110007802; JP; CN; EP; IN; TW
	Video coding using transforms bigger than 4x4 and 8x8*16	JP5259828	US8,483,285; AU; CA; CN; EP; ID; IN; KR; PH; RU; SG; TW; UA; VN; ZA
	Video coding with large macroblocks *16	JP5384652	US8,634,456; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TW; UA; VN; ZA
	Video coding with large macroblocks *16	JP5547199	US8,619,856; JP; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TW; UA; VN; ZA
	Video coding with large macroblocks *16	JP2012-504908	US8,503,527; US20130308701; JP; CN; EP; HK; IN; KR; TW

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Chrominance high precision motion filtering for motion interpolation*16	JP5646654	US20110200108; CN; EP; HK; IN; KR; TW
	Block type signalling in video coding *16	JP5642806	US20110206123; BR; CN; EP; IN; KR; TW
QUALCOMM Incorporated	Mixed tap filters*16	JP5607236	US20110249737; JP; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Adapting frequency transforms for intra blocks coding based on size and intra mode or based on edge detection*16	JP2013-531445	US20120008683; JP; AT; BE; CH; CN; DE; DK; EP; ES; FI; FR; GB; GR; HU; IE; IN; IT; KR; NL; NO; PL; PT; RO; SE
	Indicating intra-prediction mode selection for video coding*16	JP2013-539940	US20120082223; CN; EP; IN; KR
	Intra smoothing filter for video coding*16	JP5587508	US20120082224; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Entropy coding coefficients using a joint context model*16	JP2013-543317	US8,913,666; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Adaptive support for interpolating values of sub-pixels for video coding *16	JP2014-502800	US20120147967; JP; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Separately coding the position of a last significant coefficient of a video block in video coding*16	JP2014-504077	US20120140813; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Coding the position of a last significant coefficient within a video block based on a scanning order for the block in video coding*16	JP2013-542151	US20120140814; US20140341274; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
QUALCOMM Incorporated	Indicating intra-prediction mode selection for video coding using CABAC*16	JP2014-506067	US8,913,662; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Signaling quantization parameter changes for coded units in high effiency video coding (HEVC) *16	JP2014-506752	US20120189052; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Performing motion vector prediction for video coding*16	JP2014-509480	US20120195368; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Multi-metric filtering*16	JP2014-511613	US20120213291; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Quantized pulse code modulation in video coding*16	JP2014-511649	US20120224640; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
	Coding of transform coefficients for video coding*16	JP2014-509158	US20120230419; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Coding of transform coefficients for video coding*16	JP2014-511657	US20120230420; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Video coding techniques for coding dependent pictures after random access*16	JP2014-513456	US20120230433; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Hierarchy of motion prediction video blocks*16	JP2014-511618	US20120219064; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Coding of transform coefficients for video coding*16	JP2014-511656	US20120230418; US20140307777; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Bi-predictive merge mode based on uni-predictive neighbors in video coding*16	JP2014-514814	US20120243609; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
	Motion vector prediction in video coding*16	JP2014-514861	US20120269270; US20130272408; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Offset type and coefficients signaling method for sample adaptive offset*16	JP2014-516217	US20120287988; US20140241417; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Enhanced intra-prediction mode signaling for video coding using neighboring mode*16	JP2014-517630	US20120314766; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Memory efficient context modeling *16	JP2014-522603	US20120328003; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
QUALCOMM Incorporated	Coding of transform coefficients for video coding*16	JP2014-511655	US20120230417; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Unified merge mode and adaptive motion vector prediction mode candidates selection*16	JP2014-517656	US20120320969; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
	Derivation of the position in scan order of the last significant transform coefficient in video coding*16	JP2014-521249	US20130003834; BR; CA; CN; EP; IN; KR; RU
	Signaling syntax elements for transform coefficients for sub-sets of a leaf-level coding unit*16	JP2014-521256	US20130003821; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Video coding using adaptive motion vector resolution*16	JP2014-523714	US20130003849; US20140341297; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
	Unified merge mode and adaptive motion vector prediction mode candidates selection*16	JP2014-516989	US20120320968; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Signaling picture size in video coding*16	JP2014-521281	US20130016769; US20140341275; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Buffering prediction data in video coding*16	JP2014-525198	US20130022119; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Adaptive center band offset filter for video coding*16	JP2014-533048	US20130114674; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Motion vector determination for video coding*16	JP2014-526840	US20130070854; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Motion vector predictor candidate clipping removal for video coding*16	JP2014-531873	US20130083853; BR; CN; EP; IN; KR; TW
	Coding reference pictures for a reference picture set*16	JP2014-530570	US20130077687; AR; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Video coding with subsets of a reference picture set*16	JP2014-530571	US20130077679; AR; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Reference picture list construction for video coding*16	JP2014-530567	US20130077677; AE; AR; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Reference picture list construction for video coding*16	JP2014-530568	US20130077678; AR; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Reference picture list construction for video coding*16	JP2014-526858	US20130077685; AE; AR; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Decoded picture buffer management *16	JP2014-530569	US20130077680; AR; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Performing transform dependent de-blocking filtering*16	JP2014-531879	US20130094572; CN; EP; IN; KR; TW
	Parallelization friendly merge candidates for video coding*16	JP2014-517658	US20130077691; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
	Intra PCM (IPCM) and lossless coding mode video deblocking*16	JP2014-531169	US20130101025; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Determining boundary strength values for deblocking filtering for video coding*16	JP2014-534733	US20130101024; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Loop filtering around slice boundaries or tile boundaries in video coding*16	JP2014-533008	US20130101016; CN; EP; IN; KR

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Coefficient scanning in video coding *16	JP2014-525200	US20130051475; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Random access with advanced decoded picture buffer (DPB) management in video coding*16	JP2014-540043	US20130107953; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Unified design for picture partitioning schemes*16	JP2014-534737	US20130107952; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Loop filtering control over tile boundaries*16	JP2014-534738	US20130107973; BR; CN; EP; IN; KR
	Video coding with network abstraction layer units that include multiple encoded picture partitions *16	JP2014-540122	US20130114735; BR; CN; EP; IN; KR; TW
	Intra-mode video coding*16	JP2014-540129	US20130114707; AR; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Context state and probability initialization for context adaptive entropy coding*16	JP2014-540089	US20130114675; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Signaling quantization matrices for video coding*16	JP2014-541203	US20130114695; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Generating additional merge candidates*16	JP2014-541199	US20130114717; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Padding of segments in coded slice NAL units*16	JP2014-540073	US20130114736; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Progressive coding of position of last significant coefficient*16	JP2014-541158	US20130114738; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Context reduction for context adaptive binary arithmetic coding *16	JP2014-541069	US20130114671; US20140355681; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Number of contexts reduction for context adaptive binary arithmetic coding*16	JP2014-541070	US20130114672; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Number of context reduction for context adaptive binary arithmetic coding*16	JP2014-541071	US20130114673; US20140355669; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Border pixel padding for intra prediction in video coding*16	JP2014-520454	US20120314767; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
	Largest coding unit (LCU) or partition-based syntax for adaptive loop filter and sample adaptive offset in video coding*16	JP2014-543556	US20130136167; BR; CN; EP; IN; KR; TW

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Performing motion vector prediction for video coding*16	JP2014-549122	US20130163668; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Signaling of deblocking filter parameters in video coding*16	JP2014-553475	US20130188733; US20140369404; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Determining contexts for coding transform coefficient data in video coding*16	JP2014-552329	US20130182772; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Determining contexts for coding transform coefficient data in video coding*16	JP2014-552336	US20130182773; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Determining contexts for coding transform coefficient data in video coding*16	JP2014-552342	US20130182758; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Coding parameter sets and NAL unit headers for video coding*16	JP2014-552328	US20130182755; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Throughput improvement for CABAC coefficient level coding*16	JP2014-552197	US20130182757; AR; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; TW; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Indication of use of wavefront parallel processing in video coding *16	JP2014-553300	US20130182774; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Sub-streams for wavefront parallel processing in video coding*16	JP2014-553301	US20130182775; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Context optimization for last significant coefficient position coding*16	JP2014-541161	US20130114676; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Restriction of prediction units in B slices to uni-directional inter prediction*16	JP2014-556674	US20130202037; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Motion vector coding and bi-prediction in HEVC and its extensions*16	US20130243093*	JP; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Deriving context for last position coding for video coding*16	US20130251041*	JP; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Chroma slice-level QP offset and deblocking*16	US20130259141*	JP; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Coded block flag coding*16	US20130266074*	JP; AE; AR; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; VN; ZA
	Low-delay video buffering in video coding*16	US20130266075*	JP; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Low-delay video buffering in video coding*16	US20130266076*	JP; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Grouping bypass coded syntax elements in video coding*16	WO2013154939*	US20130272380; JP; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Wavefront parallel processing for video coding*16	WO2013154687*	US20130272370; JP; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MX; MY; PH; RU; SG; TH; UA; VN; ZA
	Bypass bins for reference index coding in video coding*16	WO2013154866*	US20130272377; JP; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Transform coefficient coding*16	WO2013158642*	US20130272423; JP; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
	Coding least signficant bits of picture order count values identifying long-term reference pictures*16	JP2014-544936	US20130142256; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Coding picture order count values identifying long-term reference frames*16	JP2014-544938	US20130142257; AE; AU; BR; CA; CN; EP; HK; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Video coding with enhanced support for stream adaptation and splicing *16	WO2013158415*	US20130279564; JP; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Quantization parameter (QP) coding in video coding*16	WO2013163526*	US20130287103; JP; AE; AR; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Parameter set updates in video coding*16	WO2013163563*	US20130294499; JP; BR; CN; EP; IN; KR; TW
	Full random access from clean random access pictures in video coding*16	WO2013163569*	US20130294500; JP; AR; CN; EP; IN; KR; TW
	Decoded picture buffer processing for random access point pictures in video sequences*16	WO2013158461*	US20130279599; JP; CN; EP; IN; KR
	Marking reference pictures in video sequences having broken link pictures*16	WO2013158462*	US20130279575; JP; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Signaling data for long term reference pictures for video coding *16	WO2013184305*	US20130329787; JP; AE; AU; BR; CA; CN; IL; IN; MX; MY; PH; SG; TH
	Grouping of bypass-coded bins for SAO syntax elements*16	WO2013188558*	US20130336382; AR; CN; TW
	High-level syntax extensions for high efficiency video coding*16	US20130243081*	JP; AE; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; UA; VN; ZA
	Signaling long-term reference pictures for video coding*16	WO2014004391*	US20140003538; AR; AU; CA; CN; EP; IL; IN; MY; SG; TW
	Streaming adaption based on clean random access (CRA) pictures*16	WO2014004150*	US20140003536; AU; CA; EP; IL; IN; MX; MY; SG; TW
	Tiles and wavefront parallel processing*16	WO2014005087*	US20140003531; JP; AR; CN; EP; IN; TW
	Random access and signaling of long-term reference pictures in video coding*16	WO2014004201*	US20140003537; AR; AU; BR; CA; EP; IL; IN; MX; MY; SG; TW

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Coefficient groups and coefficient coding for coefficient scans*16	WO2013158563*	US20130272378; JP; AE; AR; AU; BR; CA; CN; EP; ID; IL; IN; KR; MY; PH; RU; SG; TH; TW; UA; VN; ZA
QUALCOMM Incorporated	Video parameter set for HEVC and extensions*16	WO2014008286*	US20140003491; AU; CA; EP; IN; MY; SG; TW
	Video parameter set for HEVC and extensions*16	WO2014008287*	US20140003492; TW
	Video parameter set for HEVC and extensions*16	WO2014008290*	US20140003493; EP; IN; TW
	SEI messages including fixed-length coded video parameter set ID (VPS_ID) *16	WO2014011363*	US20140010277; IN; TW
	Coding random access pictures for video coding*16	WO2014011567*	US20140016697; IN; TW
	Coding SEI NAL units for video coding*16	WO2014011569*	US20140016707; AU; IN; SG; TW
	Coding timing information for video coding*16	WO2014011570*	US20140016708; AU; IN; MY; SG; TW
	Video coding with improved random access point picture behaviors*16	WO2014046850*	US20140079140; AR; TW
	Indication of interlaced video data for video coding*16	WO2014047202*	US20140079116; AR; TW
	Indication of frame-packed stereoscopic 3D video data for video coding*16	WO2014047204*	US20140078249; TW
	Indication and activation of parameter sets for video coding*16	WO2014046812*	US20140086317; AR; TW
	Indication and activation of parameter sets for video coding*16	WO2014046813*	US20140086337; AR; TW
	Hypothetical reference decoder parameters in video coding*16	WO2014047183*	US20140086336; AR; TW
	Bitstream conformance test in video coding*16	WO2014047178*	US20140086303; AR; TW
	Bitstream conformance test in video coding*16	WO2014047175*	US20140086331; AR; TW
	Access unit independent coded picture buffer removal times in video coding*16	WO2014047577*	US20140086332; AR; TW
	Coded picture buffer removal times signaled in picture and sub-picture timing supplemental enhancement information messages*16	WO2014047580*	US20140086341; AR; TW

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Sequence level flag for sub-picture level coded picture buffer parameters*16	WO2014047582*	US20140086342; AR; TW
	Expanded decoding unit definition *16	WO2014047583*	US20140086340; AR; TW
	Buffering period and recovery point supplemental enhancement information messages*16	WO2014047584*	US20140086343; AR; TW
QUALCOMM Incorporated	Coded picture buffer arrival and nominal removal times in video coding*16	WO2014047586*	US20140086344; AR; TW
	Long-term reference picture signaling in video coding*16	PCT/US2013/060416*	US20140086324; AR; TW
	Error resilient decoding unit association*16	WO2014051892*	US20140092993; TW
	Supplemental enhancement information message coding*16	WO2014051893*	US20140092994; AR; TW
	Signaling of regions of interest and gradual decoding refresh in video coding*16	WO2014051915*	US20140092963;
	Signaling layer identifiers for operation points in video coding*16	WO2014052013*	US20140092955; AR; TW
	Improved signaling of layer identifiers for operation points of a video coder*16	WO2014055536*	US20140092996; AR; TW
	Hypothetical reference decoder parameter syntax structure*16	WO2014058598*	US20140098895; AR; TW
	Identification of operation points applicable to nested SEI message in video coding*16	PCT/US2013/060925*	US20140098894;
	Sub-bitstream applicability to nested SEI messages in video coding*16	PCT/US2013/060940*	US20140098896; AR; TW
	Low-delay buffering model in video coding*16	WO2014099489*	US20140169448; TW
	Progressive refinement with temporal scalability support in video coding*16	WO2014105485*	US20140185670; TW
	Conditional signaling of picture order count timing information for video timing in video coding*16	WO2014107360*	US20140192901; AR; TW
	Signaling of clock tick derivation information for video timing in video coding*16	WO2014107362*	US20140192902; AR; TW
	Signaling of clock tick derivation information for video timing in video coding*16	WO2014107361*	US20140192903; AR; TW
	Video buffering operations for random access in video coding*16	WO2014107250*	US20140192882; TW

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
QUALCOMM Incorporated	Non-nested SEI messages in video coding*16	WO2014107396*	US20140192149; TW
	Gradual decoding refresh with temporal scalability support in video coding*16	WO2014107721*	US20140192896; TW
	Coding of transform coefficients for video coding*16	US20130058407	
	Determining quantization parameters for deblocking filtering for video coding*16	US20130101031	TW
JVC KENWOOD Holdings, Inc.	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver2.4.*11		
SHARP CORPORATION	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.0.*14		
Dolby Japan K. K.	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.5.*17		
Dolby International AB	Motion Vector Coding Method and Motion Vector Decoding Method※20	US 8,401,080	US
	Moving Picture Coding Method and Moving Picture Decoding Method※20	US 8,396,116	US
	Picture Coding Method, Picture Decoding Method, Picture Coding Apparatus, Picture Decoding Apparatus, and Program Thereof※20	US 8,385,409	US
	Image Sequence Compression Featuring Independently Coded Regions※20	JP 4777583	DE; EP; FR; JP; US
	Compressed Video Signal Including Independently Coded Regions※20	US 6,507,618	US
	Method of Coding and Decoding Images, Coding and Decoding Device and Computer Programs Corresponding thereto※20	PCT/FR2012/050380 JP 2013-557151	BR; CN; EP; HK; IN; JP; KR; RU; US
	Method of Coding and Decoding Images, Coding and Decoding Device and Computer Programs Corresponding thereto※20	PCT/FR2012/051391 JP 2014-516422	BR; CN; EP; HK; IN; JP; KR; RU; US

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Dolby International AB	Method of Coding and Decoding Images, Coding and Decoding Device and Computer Programs Corresponding thereto※20	PCT/FR2012/052552 JP 2014-539392	AL; AT; BE; BG; CH; CY; CZ; DE; DK; EE; EP; ES; FI; FR; GB; GR; HR; HU; IE; IS; IT; LI; LT; LU; LV; MC; MK; MT; NL; NO; PL; PT; RO; RS; SE; SI; SK; SM; TR; US; BR; HK; IN; JP; KR; CN; RU
	Method of Coding and Decoding Images, Coding and Decoding Device and Computer Programs Corresponding thereto※20	PCT/FR2012/052551 JP 2014-539391	AL; AT; BE; BG; CH; CY; CZ; DE; DK; EE; EP; ES; FI; FR; GB; GR; HR; HU; IE; IS; IT; LI; LT; LU; LV; MC; MK; MT; NL; NO; PL; PT; RO; RS; SE; SI; SK; SM; TR; US; BR; HK; IN; JP; KR; CN; RU
	Methods and Systems for Parallel Video Encoding and Decoding※20	PCT/JP2009/056778 JP 5529937 JP 5075988 JP 5786061 JP 5075988 JP 2015-147980 JP 2015-147981	CN; JP; RU; US; EP; JK; BR; IN
	Tracking a Reference Picture Based on an Designated Picture on an Electronic Device※20	PCT/JP2012/077021 JP 2014-516128	US; CN; EP; JP

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Dolby International AB	Moving Picture Decoder※20	JP 3664626 JP 3710464 JP 4462914 JP 4508627	JP
	Method and System for Selectively Breaking Prediction in Video Coding ※20	PCT/CA2011/001412	CN; EP; US
	Method and system for picture segmentation using columns※20	PCT/CA2011/001411	CN; EP; US
	Method and System for Dynamic Selection of Transform Size in a Video Decoder Based on Signal Content※20	US 7,894,530	CN; US; TW
	Method and Apparatus for Controlling Loop Filtering or Post Filtering in Block Based Motion Compensationed Video Coding※20	JP 3688248 JP 3714944 JP 4120989 JP 4565010 JP 4666411 JP 4666413 JP 4666414 JP 4666415 JP 4717136 JP 4717137 JP 4717138 JP 4723024 JP 4723025 JP 4723026 JP 4723027	DE; EP; FR; GB; JP; US

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Dolby International AB	Adaptive filtering Based Upon Boundary Strength※20	PCT/JP02/09306 JP 3688283 JP 3688288 JP 4372019 JP 4094019 JP 4372197 JP 4672065 JP 4672074 JP 4672077 JP 4672078 JP 4723022 JP 4723023 JP 5346908 JP 5222343 JP 5216070 JP 5216071	AT; BE; CA; CN; DE; EP; ES; FR; GB; HK; IE; IT; JP; KR; NL; PT; SE; TR; US
	Encoding Device and Decoding Device※20	JP 2001-348412 PCT/JP2002/011605 JP 3926726 JP 4308229 JP 5048697	CN; DE; FR; GB; ID; JP; KR; NL; US
	Embedded Block Coding with Optimized Truncation※20	US 6,778,709	US
	Source Coding Enhancement Using Spectral-Band Replication※20	PCT/IB1998/000893 JP 4220461 JP 3871347	AT; BE; BR; CH; CN; DE; DK; ES; FI; FR; GB; HK; IE; IT; JP; LI; NL; PT; RU; SE; US
	Efficient Spectral Envelope Coding Using Variable Time/Frequency Resolution and Time/Frequency Switching※20	PCT/SE2000/001887 JP 4035631 JP 4334526 JP 4628921	AT; BE; BR; CH; CN; DE; DK; ES; FI; FR; GB; HK; IE; IT; JP; LI; NL; PT; RU; SE
	Efficient Spectral Envelope Coding Using Variable Time/Frequency Resolution and Time/Frequency Switching※20	PCT/SE2000/000158	US

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Dolby International AB	Enhancing Perceptual Performance of SBR and Related HFR Coding Methods by Adaptive Noise-Floor Addition and Noise Limiting※20	PCT/SE2000/000159 JP 4377302 JP 4511443 JP 4519783 JP 4519784 JP 4852122 JP 4852123 JP 3603026	AT; BE; BR; CH; CN; DE; DK; ES; FI; FR; GB; GR; HK; IE; IT; JP; LI; LU; NL; PT; SE; US
	Spectral Translation/Folding in the Subband Domain※20	PCT/SE2001/001171 JP 4289815 JP 5090390	BR; CN; DE; FI; FR; GB; HK; JP; NL; RU; SE; US
	Enhancing Perceptual Performance of High Frequency Reconstruction Coding Methods by Adaptive Filtering※20	PCT/SE2001/002510 JP 3954495	AT; BE; CH; CN; DE; DK; ES; FI; FR; GB; HK; IE; IT; JP; KR; LI; NL; PT; SE; US
	Enhancing the Performance of Coding Systems that Use High Frequency Reconstruction Methods ※20	PCT/SE2001/002533 JP 3983668 JP 4991397 JP 2011-269144 JP 2014-002174	AT; BE; CH; CN; DE; DK; ES; FI; FR; GB; HK; IE; IT; JP; KR; LI; NL; PT; SE; TR; US
	Aliasing Reduction Using Complex-Exponential Modulated Filterbanks※20	PCT/SE2002/000626 JP 3977744	CN; DE; ES; FI; FR; GB; HK; IN; IT; JP; KR; NL; SE; TR
	Efficient and Scalable Parametric Stereo Coding for Low Bitrate Audio Coding Applications※20	PCT/SE2002/001372 JP 4447317 JP 4474347 JP 4700467 JP 4786987 JP 4878384 JP 5133397 JP 5186444 JP 5186543 JP 5427270	AT; BE; CH; CN; CZ; DE; DK; ES; FI; FR; GB; GR; HK; IE; IN; IT; JP; KR; LI; LU; NL; SE; TR; US

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Dolby International AB	Methods for Improving High Frequency Reconstruction※20	PCT/EP2002/013462 JP 3870193	AT; BE; CH; CN; DE; DK; ES; FI; FR; GB; HK; IE; IN; IT; JP; KR; LI; NL; PT; SE; US
	Method for Reduction of Aliasing Introduced by Spectral Envelope Adjustment in Real-Valued Filterbanks※20	PCT/EP2003/009485 JP 4328720 JP 5132627 JP 5326020 JP 5557467 JP 5577187	AT; AU; BE; CA; CH; CN; DE; DK; ES; FI; FR; GB; HK; IN; IT; JP; KR; LI; MX; NL; NO; SE; SG; TR; UA; US; VN; ZA
	Method for Reduction of Aliasing Introduced by Spectral Envelope Adjustment in Real-Valued Filterbanks※20	US 7,548,864 US 7,577,570 US 7,590,543 US 8,145,475 US 8,346,566 US 8,498,876 US 8,606,587	US
	Advanced Processing Based on a Complex-Exponential-Modulated Filterbank and Adaptive Time Signalling Methods※20	PCT/EP2004/004607 JP 4527716 JP 4602375	AT; CH; CN; DE; DK; ES; FI; FR; GB; HK; IN; IT; JP; KR; LI; NL; PL; SE; TR; US
	Audio Data Decoding Device and Audio Data Coding/Decoding System※20	JP 3765622 CN ZL97114604.7	CN
	Method for Reduced Bit-Depth Quantization※20	PCT/JP02/08146 JP3678365 JP 3862725 JP 4030558 JP 4067558 JP 4745325 JP 4745425 JP 4745433 JP 4745434 JP 4745435 JP 4745436	CA; US; AT; BE; BG; CH; CY; CZ; CN; DE; DK; EE; EP; ES; FI; FR; GB; GR; HK; IE; IT; JP; KR; LI; LU; MC; NL; PT; SE; SK; TR

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	Methods and Systems for Image Intra-Prediction Mode Estimation, Communication, and Organization ※20	PCT/JP03/06623 JP 3734492 JP 3734494 JP 4357427 JP 4357543 JP 4357590	CN; DE; EP; ES; FR; GB; HK; IT; JP; KR; NL; TW; US
	Video Encoder※20	PCT/JP2004/004374 JP 5025289 JP 5444047 JP 5536811	AT; BE; CN; DE; EP; ES; FI; FR; GB; HK; IE; IT; JP; NL; PL; PT; SE; US
Dolby Laboratories Licensing Corporation	Device and Method of Improving the Perceptual Luminance Nonlinearity-Based Image Data Exchange Across Different Display Capabilities※20	PCT/US2012/068212 JP 2016-032053	AU; BR; CA; CN; DE; EP; ES; FR; GB; HK; IN; IT; JP; KR; MX; MY; NL; RU; SG; TH; US; VN
	Enhanced Temporal and Resolution Layering in Advanced Television※20	PCT/US2001/11204	CA; CG; US
	High Precision Encoding and Decoding of Video Images※20	PCT/US2002/06078	AT; BE; CH; CN; DE; DK; EP; ES; FI; FR; GB; HK; IT; LI; NL; SE; SG; TR; US
	Interpolation of Video Compression Frames※20	PCT/US2002/22063 JP 4339680	AU; CA; CN; JP; MX; SG; US
	Method and System for Improving Compressed Image Chroma Information※20	PCT/US2002/22205 JP 5178389 JP 5506645 JP 5506901 JP 5506902 JP 5506903 JP 5506904 JP 5506905	AU; BN; CA; CN; DE; EP; ES; FI; FR; GB; HK; IN; IT; JP; MX; NL; SE; SG; SK; TR; US

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Dolby Laboratories Licensing Corporation	Interpolation of Video Compression Frames※20	PCT/US2003/20397	AU; CA; CN; EP; HK; IN; KR; MX; MY; SG; TW; US; MO; VE
	Quantization Control for Variable Bit Depth※20	US 8,548,047	US
	Compatible Stereoscopic Video Deliver※20	PCT/US2009/050809	CN; EP; US
	Methods and Devices for Sub-Sampling and Interleaving Multiple Images, EG Stereoscopic ※20	PCT/US2010/022445 JP 5406942	CN; EP; HK; JP; KR; US
	Directed Interpolation and Data Post-Processing※20	PCT/US2010/031762 JP 5562408	CN; EP; JP; US
	Methods and Systems for Reference Processing in Image and Video Codecs※20	PCT/US2011/020168 JP 5680674	CN; EP; JP; KR; US
	Image Processing Methods and Apparatus Using Localized Gamut Definitions※20	PCT/US2011/050484	CN; EP; KR; US
	Systems and Methods for Multi-Layered Frame-Compatible Video Delivery※20	PCT/US2011/044757 JP 5749340 JP 2016-081931 JP 2016-081932	CN; EP; HK; JP; US
	Inter-layer Reference Picture Processing for Coding Standard Scalability※20	PCT/US2013/061352 JP 2015-534595	AU; BR; ID; KR; MX; MY; PA; RU; SG; TH; UA; VN; CA; CN; EO; HK; IL; IN; JP; TW; US
	High Precision Up-sampling in Scalable Coding of High Bitdepth Video※20	PCT/US2013/073006 JP 2015-549434	TW; BR; HK; IN; KR; MY; RU; CN; EP; JP; US

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
Dolby Laboratories Licensing Corporation	Audio Data Decoding Device and Audio Data Coding/Decoding System[※20]	JP 3765622 US 6,240,388	US
	Method and Apparatus for Encoding and Decoding Multiple Audio Channels at Low Bit Rates Using Adaptive Selection of Encoding Method[※20]	US 5,890,125 PCT/US1998/008647 JP 4223679	JP; US
	Reconstruction of the Spectrum of an Audio Signal With Incomplete Spectrum Based on Frequency Translation[※20]	PCT/US2003/008895 JP 4345890	AU; BG; CA; CN; DE; EE; FR; GB; HK; ID; IE; IN; JP; KR; MY; SG; SI; SK; TR; US
	Processing Audio Signals with Adaptive Time or Frequency Resolution[※20]	PCT/US2002/005999 JP 4763965	JP; US

*1 : Valid for the revised parts of ARIB STD-B32 Ver1.5
*2 : Valid for the revised parts of ARIB STD-B32 Ver1.6
*3 : Valid for the revised parts of ARIB STD-B32 Ver1.7
*4 : Valid for the revised parts of ARIB STD-B32 Ver1.8
*5 : Valid for the revised parts of ARIB STD-B32 Ver1.9
*6 : Valid for ARIB STD-B32 Ver1.0
*7 : Valid for the revised parts of ARIB STD-B32 Ver1.1
*8 : Valid for the revised parts of ARIB STD-B32 Ver2.2
*9 : Valid for the revised parts of ARIB STD-B32 Ver2.3 (received on April 16, 2010)
*10 : Valid for the revised parts of ARIB STD-B32 Ver2.3 (received on October 22, 2010)
*11 : Valid for the revised parts of ARIB STD-B32 Ver2.4 (received on October 28, 2010)
*12 : Valid for the revised parts of ARIB STD-B32 Ver2.5 (received on March 18, 2011)
*13 : Valid for the revised parts of ARIB STD-B32 Ver2.9 (received on March 11, 2014)
*14 : Valid for the revised parts of ARIB STD-B32 Ver3.0 (received on July 24, 2014)
*15 : Valid for the revised parts of ARIB STD-B32 Ver3.1 (received on December 9, 2014)
*16 : Valid for the revised parts of ARIB STD-B32 Ver3.0 (received on January 26, 2015)
*17 : Valid for the revised parts of ARIB STD-B32 Ver3.5 (received on November 26, 2015)
*18 : Valid for the revised parts of ARIB STD-B32 Ver3.6 (received on March 18, 2016)
*19 : Valid for the revised parts of ARIB STD-B32 Ver3.7 (received on June 29, 2016)
*20 : Valid for the revised parts of ARIB STD-B32 Ver3.5 (received on June 29, 2016)

TOTAL CONTENTS

Part 1: Video Signal and Coding System

<Blank Page>

Part 1: Video Signal and Coding System

Contents

<Blank Page>

Chapter 1: General Terms

1.1 Objective

The purpose of this standard is to prescribe a video signal and video coding system for digital broadcasting.

1.2 Scope

This standard applies to digital broadcasting that comply with the "Standard transmission system for digital broadcasting among standard television broadcasting and the like" (Ordinance of the Ministry of Internal Affairs and Communications No. 87, 2011) and also applies to digital broadcasting that comply with the "Standard transmission system for general satellite broadcasting" (Ordinance of the Ministry of Internal Affairs and Communications No.94, 2011).

1.3 References

1.3.1 Normative References

The following documents are those which a part of the items provided in the documents is quoted in this standard.

(1) Ordinance of the Ministry of Internal Affairs and Communications No.87, 2011: "Standard transmission system for digital broadcasting among standard television broadcasting and the like" (Partial Amendment: Dec. 10, 2013, July 3, 2014, July 29, 2016. hereinafter referred to as "Ordinance.")

(2) Ordinance of the Ministry of Internal Affairs and Communications No.94, 2011: "Standard transmission system for general satellite broadcasting" (Partial Amendment: Dec. 10, 2013, July 3, 2014. hereinafter referred to as "Ordinance No.94."

(3) Notification of the Ministry of Internal Affairs and Communications No.234, 2014: "On defining the procedure of compression and procedure of transmission of video signal and audio signal" (hereinafter referred to as "Notification.")

(4) Rec. ITU-R BT.2100-0(2016): Image parameter values for high dynamic range television for use in production and international programme exchange

(5) Rec. ITU-T H.262 (2012) | ISO/IEC 13818-2:2013: Information technology -- Generic coding of moving pictures and associated audio information: Video (hereinafter referred to as "MPEG-2 Video Standard".)

(6) Rec. ITU-T H.264 (2014) | ISO/IEC 14496-10:2014: Advanced video coding for generic audiovisual services (hereinafter referred to as "MPEG-4 AVC Standard".)

(7) Rec. ITU-T H.265 (2013) | ISO/IEC 23008-2:2013: High efficiency video coding (hereinafter referred to as "HEVC Standard".)

(8) ARIB standard BTA S-001C Version 1.0 (2009): "Studio Standard of 1125/60 High Definition television system"

(9) ARIB standard STD-B56 Version 1.1 (2014): "Studio Standard of Ultra High Definition television system" (Hereinafter referred to as "UHDTV Studio Standard")

1.4 Terminology

1.4.1 Definitions

(1) Digital terrestrial television broadcasting:
Digital broadcasting and High Definition broadcasting among standard television broadcasting which are operated by key terrestrial broadcasting stations that are provided in Ordinance, Chapter 3.

(2) Multimedia broadcasting:
Television broadcasting and Multimedia broadcasting which are operated by key terrestrial broadcasting stations that are provided in Ordinance, Chapter 4.

(3) BS digital broadcasting:
Digital broadcasting among Standard television broadcasting, High-definition television broadcasting, very high frequency broadcasting and data broadcasting by wide band transmission system, which are operated by key satellite broadcasting stations (including key satellite broadcasting test station and practical test station for key satellite broadcasting) which use radio wave whose frequency is more than 11.7GHz and less than or equal to 12.2GHz that is provided in Ordinance, Chapter 5, Clause 2.

(4) Advanced BS digital broadcasting:
Digital broadcasting among Standard television broadcasting, High-definition television broadcasting, Ultra high definition television broadcasting, very high frequency broadcasting and data broadcasting by advanced wide band transmission system, which are operated by key satellite broadcasting stations (including key satellite broadcasting test station and practical test station for key satellite broadcasting) which use radio wave whose frequency is more than 11.7GHz and less than or equal to 12.2GHz, provided in Ordinance, Chapter 5, Clause 3.

(5) Narrow band CS digital broadcasting:
Standard television broadcasting, High-definition television broadcasting, very high frequency broadcasting and data broadcasting by narrow band transmission system, which are operated as a general satellite broadcasting by the satellite broadcasting stations which use radio wave whose frequency is more than 12.2GHz and less than or equal to 12.75GHz, provided in Ordinance No. 94, Chapter 3, Clause 1.

(6) Wide band CS digital broadcasting:
Standard television broadcasting, High-definition television broadcasting, very high frequency broadcasting and data broadcasting by wide band transmission system, which are operated by the satellite broadcasting stations which use radio wave whose frequency is more than 12.2GHz and less than or equal to 12.75GHz, provided in Ordinance, Chapter 6, Clause 3.

(7) Advanced narrow band CS digital broadcasting:
Standard television broadcasting, High-definition television broadcasting, Ultra high definition television broadcasting, very high frequency broadcasting and data broadcasting by advanced wide band transmission system as a general satellite broadcasting, which are operated by the satellite broadcasting stations which use radio wave whose frequency is more than 12.2GHz and less than or equal to 12.75GHz, provided in Ordinance No.94, Chapter 3, Clause 1.

(8) Advanced wide band CS digital broadcasting:
Standard television broadcasting, High-definition television broadcasting, Ultra high definition television broadcasting, very high frequency broadcasting and data broadcasting by

advanced wide band transmission system, which are operated by the key satellite broadcasting station which use radio wave whose frequency is more than 12.2GHz and less than or equal to 12.75GHz, provided in Ordinance, Chapter 6, Clause 5.

1.4.2 Abbreviations

AVC	Advanced Video Coding
CIE	Commission Internationale de l'Éclairage
DCT	Discrete Cosine Transform
DTS	Decoding Time-Stamp
EOTF	Electro-Optical Transfer Function
GOP	Group of Pictures
HDTV	High Definition Television
HDR-TV	High Dynamic Range Television
HEVC	High Efficiency Video Coding
HL	High Level
HLG	Hybrid Log-Gamma
H14L	High-1440 Level
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
ITU-T	International Telecommunication Union, Telecommunication Standardization Sector
ML	Main Level
MMTP	MPEG Media Transport Protocol
MP	Main Profile
MPEG	Moving Picture Experts Group
OETF	Opto-Electronic Transfer Function
PES	Packetized Elementary Stream
PQ	Perceptual Quantization
PTS	Presentation Time-Stamp
SDR-TV	Standard Dynamic Range Television
SDTV	Standard Definition Television
SEI	Supplemental Enhancement Information
UHDTV	Ultra High Definition Television
VUI	Video Usability Information

<Blank Page>

Chapter 2: Video Input Format

2.1 Video signal

Video signals shall be composed of a signal representing the luminance of the subject (hereinafter referred to as "luminance signal") and two other signals representing the hue and chroma of the subject (hereinafter referred to as "color difference signals").

(Ordinance, Article 2; Article 23, Paragraph 1; Article 31, Paragraph 1; Article 63, Paragraph 1; Article 81, No.2, Paragraph 1)

2.1.1 Video signal characteristics based on HDTV [1]

(1) Three primary color signals

Three primary color signals E'_R, E'_G, and E'_B shall represent voltage levels (voltage levels normalized by reference white level) resulting from gamma pre-correction (this means on the transmitting side to compensate by providing signal voltage levels E'_R, E'_G, and E'_B with inverse characteristics to the CRT so that the luminance of the CRT for red, green, and blue is correctly reproduced) of the red, green, and blue signal voltage levels developed when a pixel is scanned. E'_R, E'_G, and E'_B shall conform to CRT using red, green, and blue with "x" and "y" values in the following table as three primary colors in the CIE display system (referring to the quantitative display system of colors by means of plane coordinates established by the Commission Internationale d'Éclairage (CIE)).

	x	y
Red	0.640	0.330
Green	0.300	0.600
Blue	0.150	0.060

The reference white level shall be as shown below. The color difference signals shall be 0 when the subject is white.

	x	y
White	0.3127	0.3290

Gamma pre-correction shall be performed according to the following characteristics:

$$E' = \alpha L^{0.45} - (\alpha - 1) \qquad (\beta \le L)$$
$$E' = 4.50L \qquad (-\beta < L < \beta)$$
$$E' = -\alpha(-L)^{0.45} + (\alpha - 1) \qquad (L \le -\beta)$$

Here, E' is the output video signal of camera and L is the voltage proportional to input light of the camera, and both are normalized by reference white level mentioned above. Only in case of wide color gamut system, L＜0 and 1＜L are permitted. α and β shall be the solutions of the following simultaneous equations, and a fraction below the second decimal place may be rounded in the calculation.

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

When α and β are rounded off to the second decimal place, α=1.099, β=0.018, and so the characteristics of gamma pre-correction are the following.

[1] This provision is based on Ordinance, Article 23, Paragraph 1; Article 31, Paragraph 1; Article 63, Paragraph 1; Article 81, No.2, Paragraph 1: Appended Table 19: Appended Table 69.

$$E' = 1.099\,L^{0.45} - 0.099 \qquad (0.018 \le L)$$
$$E' = 4.500\,L \qquad (-0.018 < L < 0.018)$$
$$E' = -1.099\,(-L)^{0.45} + 0.099 \qquad (L \le -0.018)$$

(2) Luminance and color difference signals

Luminance signal E'_Y, color difference signals E'_{CB} and E'_{CR} are generated by the following equations by three primary color signals E'_R, E'_G and E'_B which are gamma pre-corrected.

$$E'_Y = 0.2126\ E'_R + 0.7152\ E'_G + 0.0722\ E'_B$$
$$E'_{CB} = (E'_B - E'_Y) / 1.8556$$
$$E'_{CR} = (E'_R - E'_Y) / 1.5748$$

The quantization characteristics of luminance signal and color difference signals are the following.

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

Note 1: m denotes quantized bit number of luminance signal and color difference signals.
2: Operator INT [A] represents the integer given by round off to the first decimal of real number A.

2.1.2 Video signal characteristics based on SDTV [2]

(1) Three primary color signals

Three primary color signals E'_R, E'_G, and E'_B shall represent voltage levels (voltage levels normalized by reference white level) resulting from gamma pre-correction (this means on the transmitting side to compensate by providing signal voltage levels E'_R, E'_G, and E'_B with the 1/γ th power of each value so that the luminance of the subject is correctly reproduced, as the luminance of CRT for red, green and blue is proportional to the γ th power of each voltage which are input to grid) of the red, green, and blue signal voltage levels developed when a pixel is scanned. E'_R, E'_G, and E'_B shall conform to CRT which use red, green, and blue with “x” and “y” values in the following table as three primary colors in the CIE display system (referring to the quantitative display system of colors by means of plane coordinates established by the Commission Internationale d’Éclairage (CIE)), and whose gamma value is 2.2.

	x	y
Red	0.67	0.33
Green	0.21	0.71
Blue	0.14	0.08

The color difference signals shall be 0 when the subject is white.

(2) Luminance and color difference signals

Luminance signal E'_Y, color difference signals E'_{CB} and E'_{CR} are generated by the following equations using three primary color signals E'_R, E'_G and E'_B which are gamma pre-corrected.

$$E'_Y = 0.299\ E'_R + 0.587\ E'_G + 0.114\ E'_B$$
$$E'_{CB} = -0.169\ E'_R - 0.331\ E'_G + 0.500\ E'_B$$
$$E'_{CR} = 0.500\ E'_R - 0.419\ E'_G - 0.081\ E'_B$$

[2] This provision is based on Ordinance, Article 73, Appended Table 55.

The quantization characteristics of luminance signal and color difference signals are as the following.

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

Note 1: m denotes quantized bit number of luminance signal and color difference signals.
2: Operator INT [A] represents the integer given by round off to the first decimal of real number A.

2.1.3 Video signal characteristics based on UDTV [3]

(1) Three primary color signals

Three primary color signals E'_R, E'_G, and E'_B shall represent voltage levels (voltage levels normalized by reference white level) resulting from gamma pre-correction (this means on the transmitting side to compensate by providing signal voltage levels E'_R, E'_G, and E'_B with inverse characteristics to the CRT so that the luminance of the CRT for red, green, and blue is correctly reproduced) of the red, green, and blue signal voltage levels developed when a pixel is scanned. E'_R, E'_G, and E'_B shall conform to CRT using red, green, and blue with "x" and "y" values in the following table as three primary colors in the CIE display system (referring to the quantitative display system of colors by means of plane coordinates established by the Commission Internationale d'Éclairage (CIE)).

	x	y
Red	0.708	0.292
Green	0.170	0.797
Blue	0.131	0.046

The reference white level shall be as shown below. The color difference signals shall be 0 when the subject is white.

	x	y
White	0.3127	0.3290

(2) Gamma pre-correction

In case of standard dynamic range (SDR-TV), gamma pre-correction shall be performed according to the following characteristics:

$$E' = \alpha L^{0.45} - (\alpha - 1) \qquad (\beta \le L \le 1)$$
$$E' = 4.50L \qquad (0 \le L < \beta)$$

Here, E' is the voltage proportional to the output video signal of camera and L is the voltage proportional to input light of the camera, and both are normalized by reference white level mentioned above. α and β are the solutions of the following simultaneous equations, and a fraction below the second decimal place may be rounded in the calculation.

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

When α and β are rounded off to the second decimal place, α=1.099 and β=0.018 and so the characteristics of gamma pre-correction are shown as the following.

$$E' = 1.099\ L^{0.45} - 0.099 \qquad (0.018 \le L \le 1)$$
$$E' = 4.500L \qquad (0 \le L < 0.018)$$

[3] This provision is based on Ordinance, Article 63, Paragraph 1; Article 81 No.2, Paragraph 1, Appended Table 69 No.2.

In case of high dynamic range (HDR-TV), gamma pre-correction (only for sampled values of luminance signal and color difference signals which are quantized by a binary number in ten figures) shall be according to each of the followings.

(i) HLG

$$E' = r\sqrt{L} \qquad (0 \le L \le 1)$$
$$E' = a \cdot \ln(L - b) + c \qquad (1 < L)$$

Here, E' is the voltage proportional to the output video signal of the camera, L is the voltage proportional to the input light of the camera, and each are normalized in the range from 0 to 1. a, b, and c are constants, and they are as the following.

a=0.17883277
b=0.28466892
c=0.55991073.

(ii) PQ

$$E' = \left(\frac{c_1 + c_2 L^{m_1}}{1 + c_3 L^{m_1}}\right)^{m_2} \qquad (0 \le L \le 1)$$

Here, E' is the voltage proportional to the output video signal of the camera. L is the voltage proportional to the input light of the camera, L=1 corresponds to luminance of the display 10,000 cd/m^2. m_1, m_2, c_1, c_2, and c_3 are constants, and they are as the followings.

$m_1 = 2610/4096 \times \frac{1}{4} = 0.1593017578125$

$m_2 = 2523/4096 \times 128 = 78.84375$

$c_1 = 3424/4096 = 0.8359375 = c_3 - c_2 + 1$

$c_2 = 2413/4096 \times 32 = 18.8515625$

$c_3 = 2392/4096 \times 32 = 18.6875$

(3) Luminance and color difference signals

Luminance signal E'_Y, color difference signals E'_{CB} and E'_{CR} are generated by gamma pre-corrected prime color signals E'_R, E'_G and E'_B as the followings;

$$E'_Y = 0.2627\ E'_R + 0.6780\ E'_G + 0.0593\ E'_B$$
$$E'_{CB} = (E'_B - E'_Y) / 1.8814$$
$$E'_{CR} = (E'_R - E'_Y) / 1.4746$$

The quantization characteristics of luminance and color difference signals are as the followings;

$$D'_Y = \mathrm{INT}\left[876\ E'_Y + 64\right]$$
$$D'_{CB} = \mathrm{INT}\left[896\ E'_{CB} + 512\right]$$
$$D'_{CR} = \mathrm{INT}\left[896\ E'_{CR} + 512\right]$$

Note 1: Operator INT [A] denotes the integer given by round off of real number A.

2.2 Sample value of signals

2.2.1 HDTV and SDTV

The sampled values for luminance and color-difference signals shall be quantized by 8-bit or 10-bit.

(Ordinance, Article 23, Paragraph 2; Article 31, Paragraph 2; Article 63, Paragraph 2; Article 81, No.2, Paragraph 2)

2.2.1 UHDTV

The sampled values for luminance and color-difference signals shall be quantized by 10-bit.

(Ordinance, Article 63, Paragraph 2; Article 81, No.2, Paragraph 2)

2.3 Scanning direction

Pictures shall be scanned at a constant rate from left to right and from top to bottom.

(Ordinance, Article 23 Paragraph 3; Article 63, Paragraph 3; Article 81, No.2, Paragraph 3)

2.4 Video signal parameters

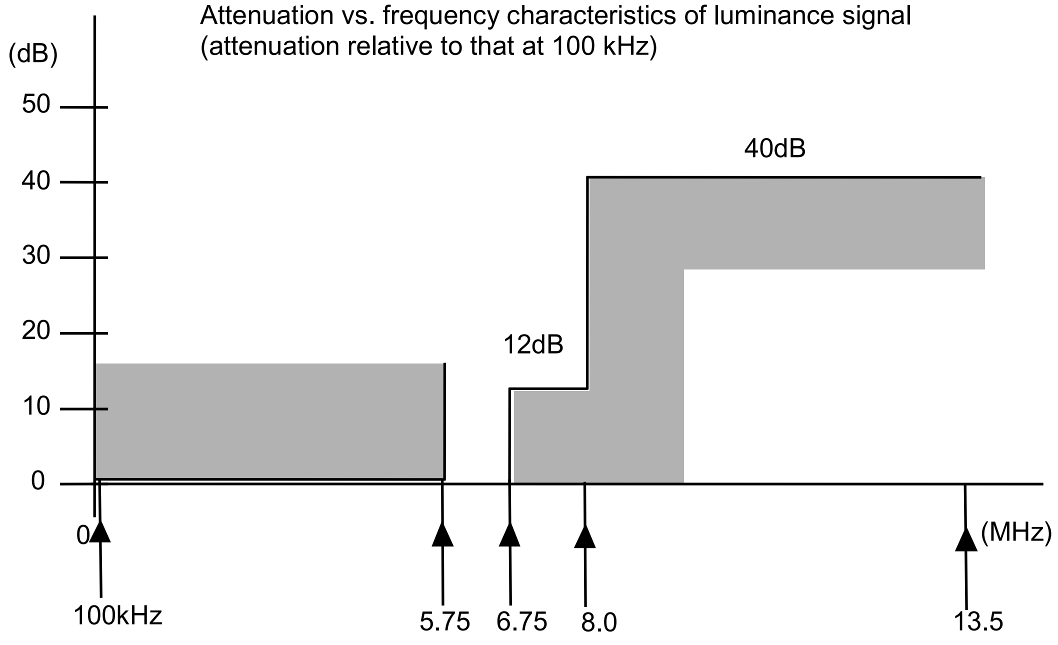
2.4.1 HDTV and SDTV

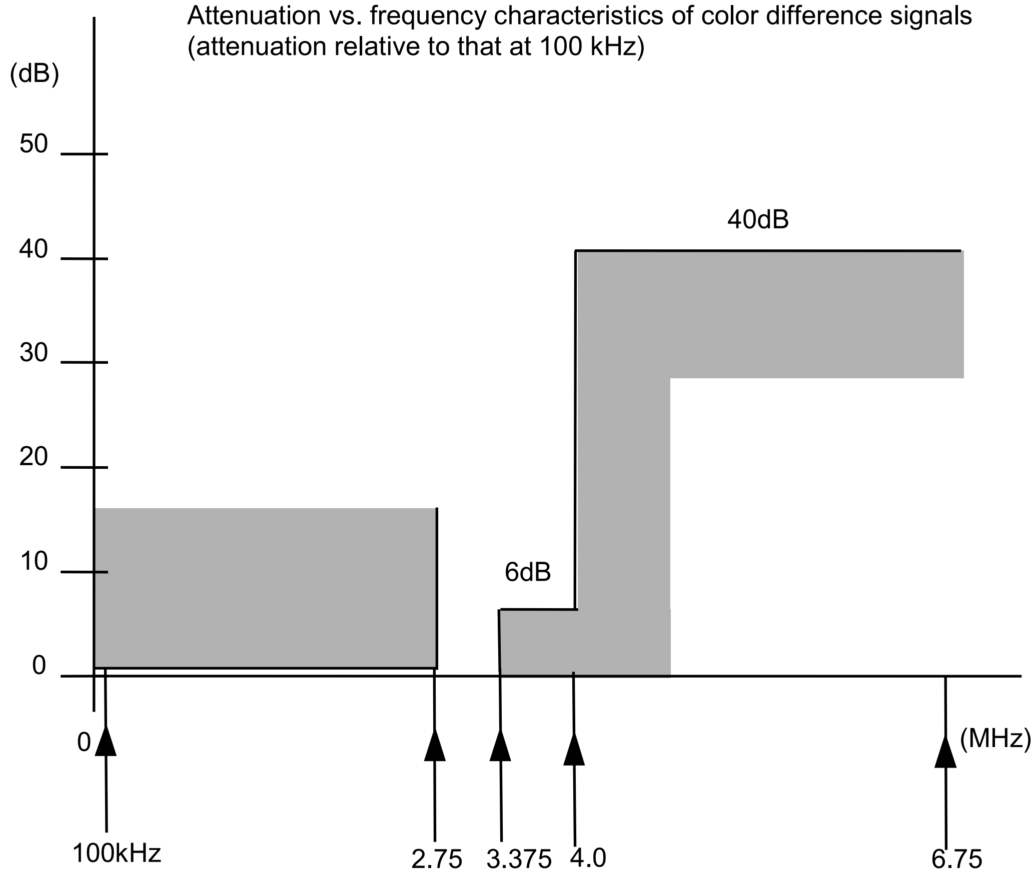
Number of lines, number of active lines, scanning system, frame frequency, field frequency, aspect ratio, horizontal scanning frequency, sampling frequencies (for luminance and color difference signals), number of samples per line (for luminance and color-difference signals), number of samples per active line (for luminance and color-difference signals), filter characteristics, and horizontal and vertical synchronizing signals for video shall be as the table shown below.

Video signal parameters

Video format		480/60/I	480/60/P	720/60/P	1080/60/I
Number of scanning lines		525	525	750	1125
Number of active scanning lines		483	483	720	1080
Scanning system		Interlaced	Progressive	Prgressive	Interlaced
Frame frequency		30/1.001 Hz	60/1.001 Hz	60/1.001 Hz	30/1.001 Hz
Field frequency		60/1.001 Hz			60/1.001 Hz
Aspect ratio		16:9 or 4:3	16:9	16:9	16:9
Horizontal frequency f_H		15.750/1.001 kHz	31.5/1.001 kHz	45.000/1.001 kHz	33.750/1.001 kHz
Sampling frequency	Luminance signal	13.5 MHz	27 MHz	74.25/1.001 MHz	74.25/1.001 MHz
	Color difference signals	6.75 MHz	13.5 MHz	37.125/1.001 MHz	37.125/1.001 MHz
Number of samples per line	Luminance signal	858	858	1650	2200
	Color difference signals	429	429	825	1100
Number of active samples per line	Luminance signal	720	720	1280	1920
	Color difference signals	360	360	640	960
Filter characteristics		Appendix 1	Appendix 2	Appendix 3	
Horizontal synchronizing signal		Appendix 4		Appendix 5	Appendix. 6
Vertical synchronizing signal		Appendix 7	Appendix 8	Appendix 9	Appendix 10

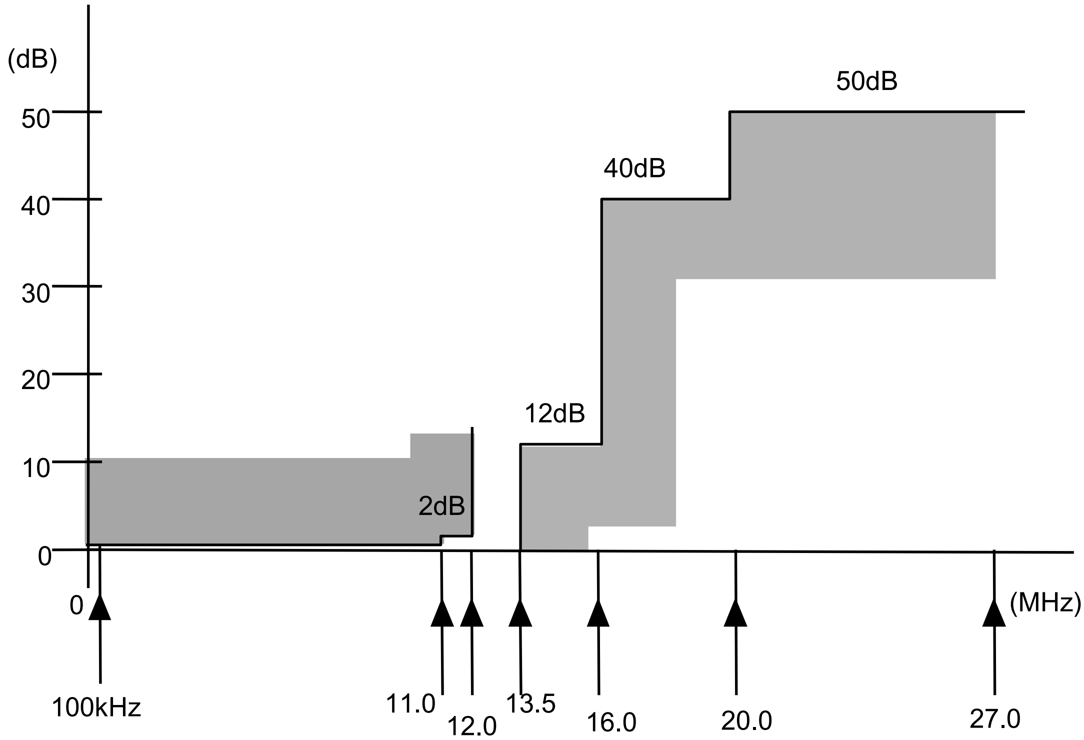
Appendix 1: Filter characteristics in case that number of lines is 525 and scanning system is interlaced



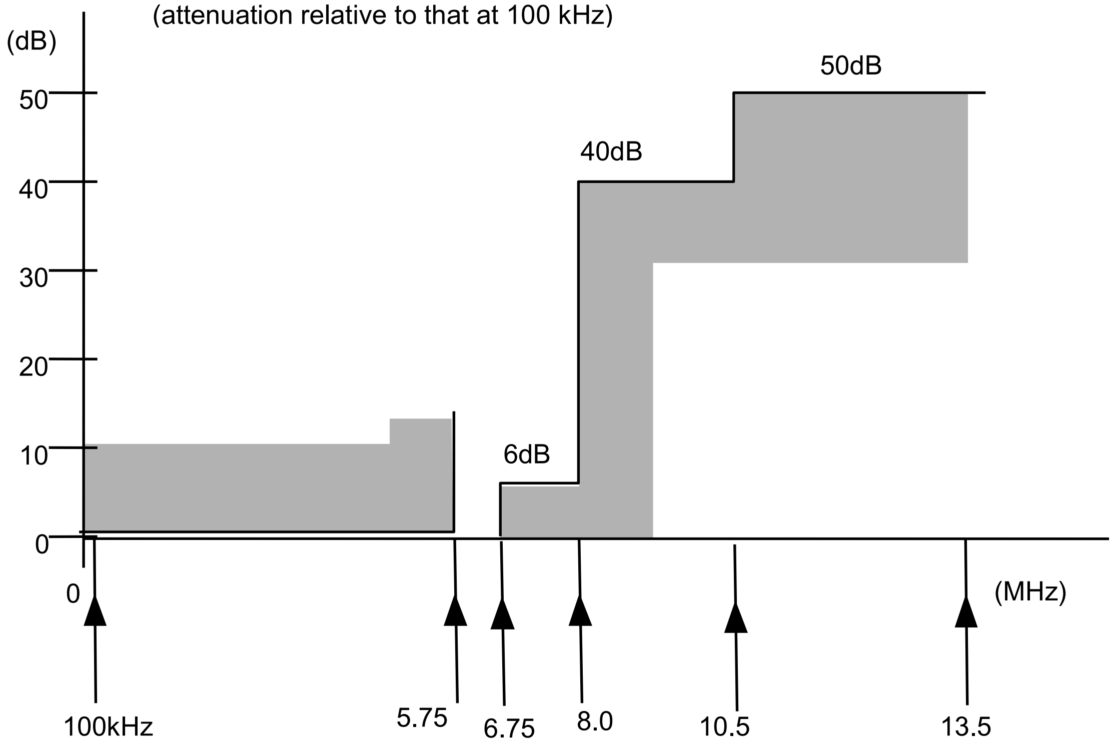


Appendix 2: Filter characteristics in case that number of lines is 525 and scanning system is progressive

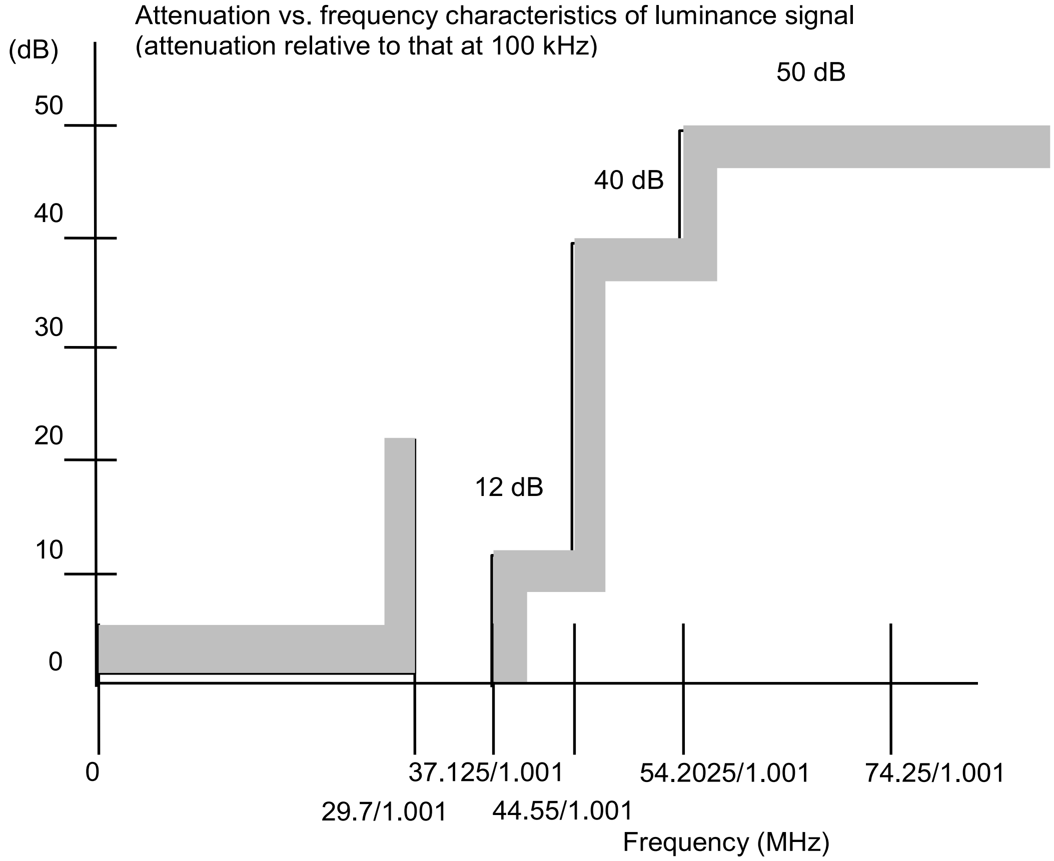
Attenuation vs. frequency characteristics of luminance signal
(attenuation relative to that at 100 kHz)

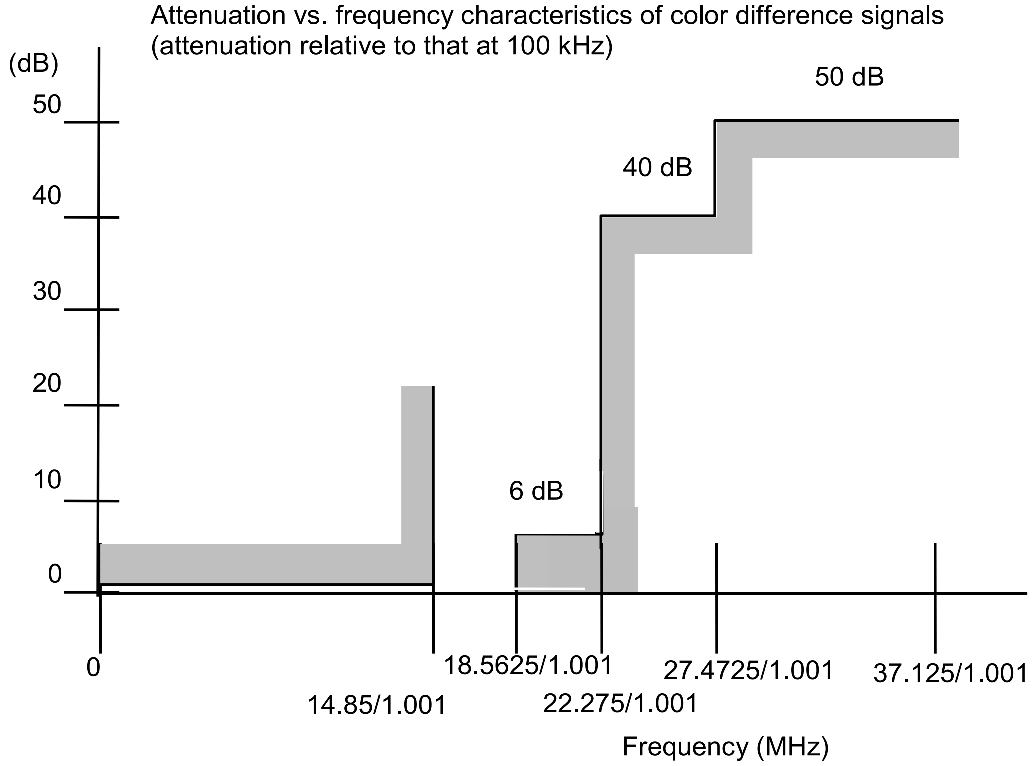


Attenuation vs. frequency characteristics of color difference signals
(attenuation relative to that at 100 kHz)

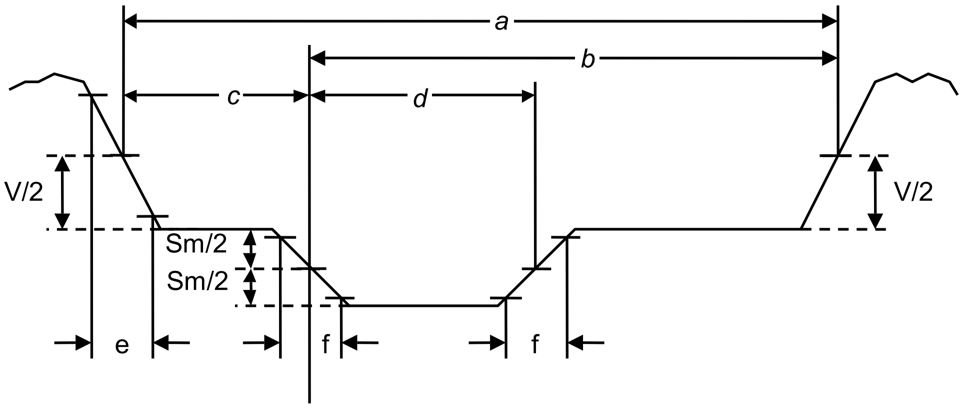


Appendix 3: Filter characteristics in case that number of lines is 750 and scanning system is progressive and in case that number of lines is 1125 and scanning system is interlaced





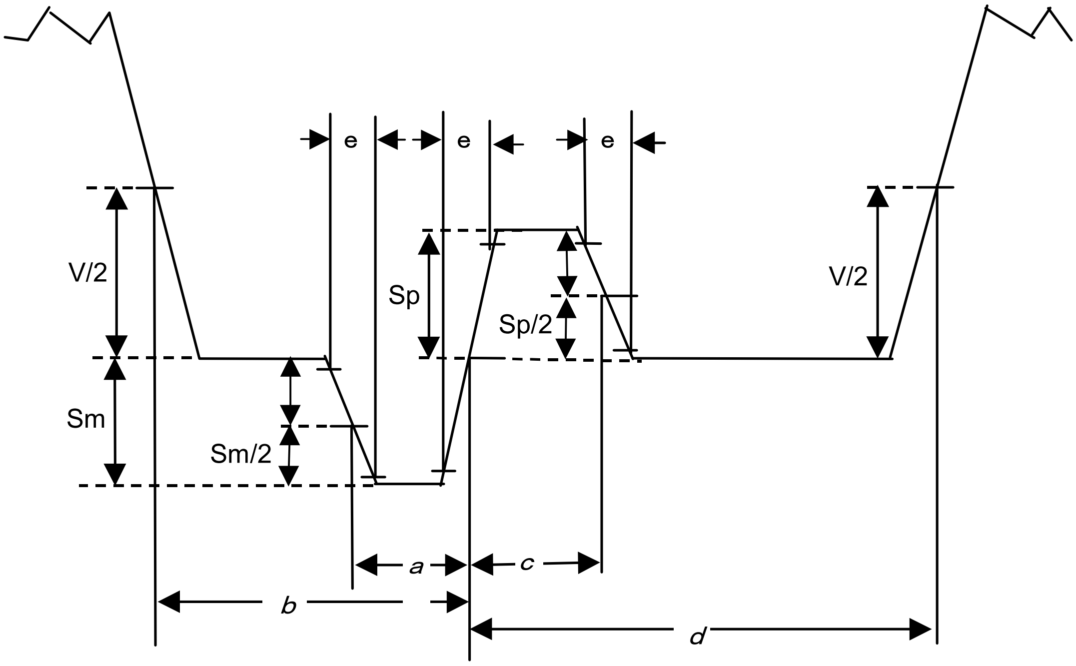
Appendix 4: Horizontal synchronizing signal in case that number of lines is 525 and scanning system is interlaced and progressive.



Timing and level specification of line synchronizing signal

Symbol	Parameter	Nominal value	
		525/59.94/2:1	525/59.94/1:1
H	Horizontal scanning period (μs)	1001/15.75	1001/31.5
a	Horizontal blanking interval (μs)	10.70	5.35
b	Start of active video (μs)	9.20	4.60
c	End of active video (μs)	1.50	0.75
d	Negative pulse width (μs)	4.70	2.35
e	Horizontal blanking fall time (10 - 90%) (μs)	0.14	0.07
f	Horizontal sync signal rise/fall time (10 - 90%) (μs)	0.14	0.07
Sm	Amplitude of negative pulse (mV)	300	
V	Amplitude of video signal (mV)	700	

Appendix 5: Horizontal synchronizing in case that number of lines is 750 and scanning system is progressive.



Level specification of horizontal synchronizing signal

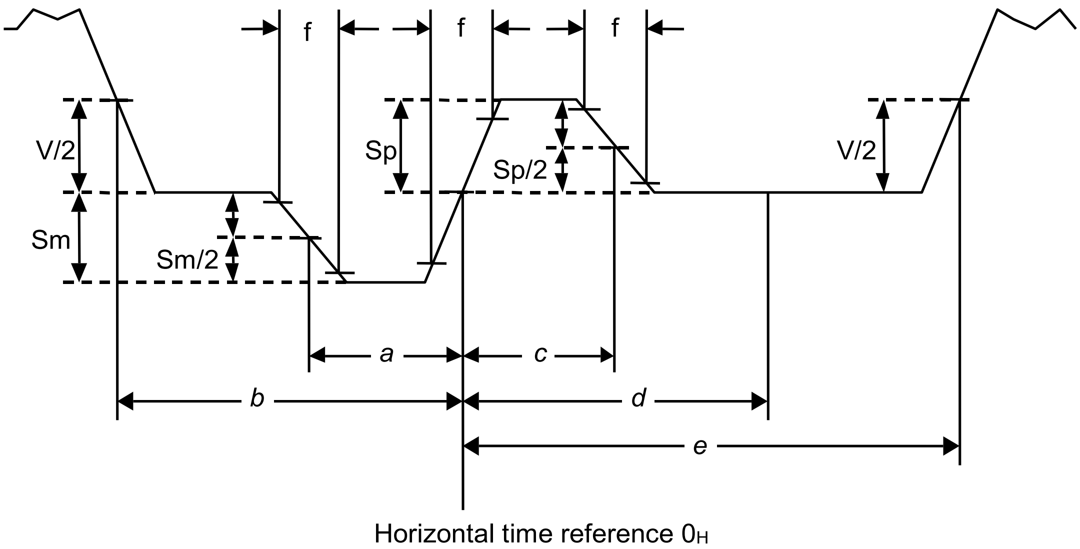
Symbol	Parameter	Nominal value
Sm	Amplitude of negative pulse (mV)	300
Sp	Amplitude of positive pulse (mV)	300
V	Amplitude of video signal (mV)	700

Timing specification of horizontal synchronizing signal

Symbol	Parameter	Nominal value
a	Negative horizontal sync width (T)	40
b	End of active video signal (T)	110
c	Positive horizontal sync width (T)	40
d	Start of active video signal (T)	260
e	Rise/fall time (T)	4

Note: " T " denotes the duration of a reference clock, and the reciprocal of the luminance sampling frequency.

Appendix 6: Horizontal synchronizing signal in case that number of lines is 1125 and scanning system is interlaced.



Level specification of horizontal synchronizing signal

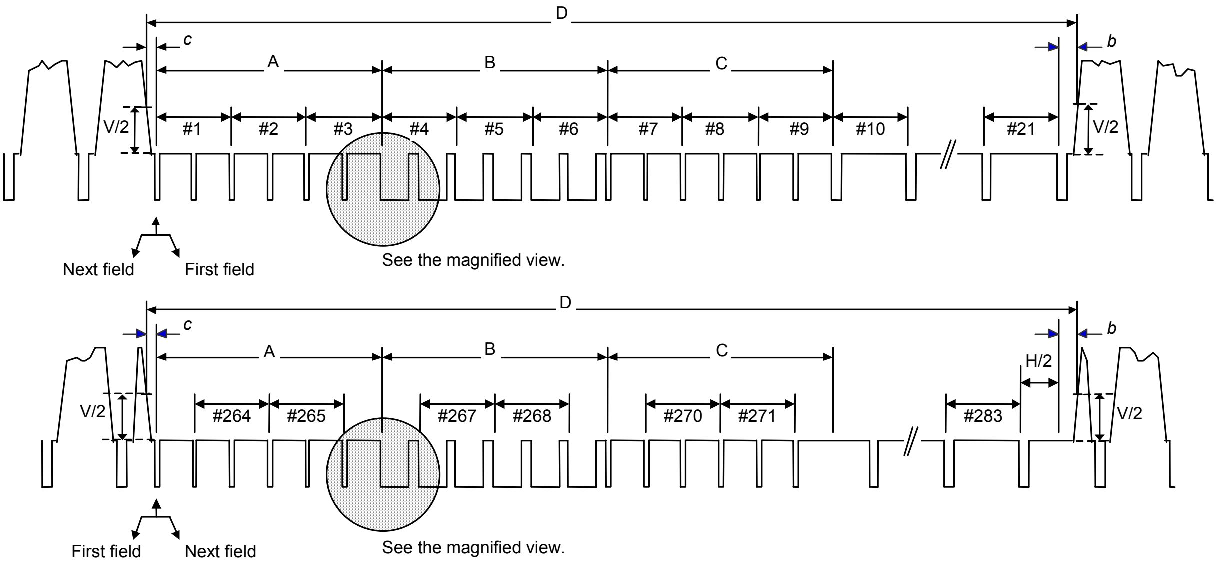
Symbol	Parameter	Nominal value
Sm	Amplitude of negative pulse (mV)	300
Sp	Amplitude of positive pulse (mV)	300
V	Amplitude of video signal (mV)	700

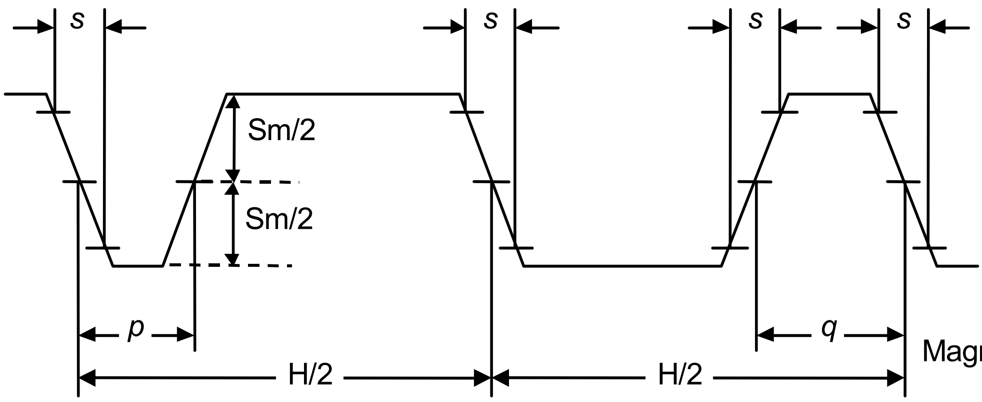
Timing specification of horizontal synchronizing signal

Symbol	Parameter	Nominal value
a	Negative horizontal sync width (T)	44
b	End of active video signal (T)	88
c	Positive horizontale sync width (T)	44
d	End of clamp (T)	132
e	Start of active video signal (T)	192
f	Rise/fall time (T)	4

Note: " T " denotes the duration of a reference clock,
and the reciprocal of the luminance sampling frequency.

Appendix 7: Vertical synchronizing signal in case that number of lines is 525 and scanning system is interlaced.





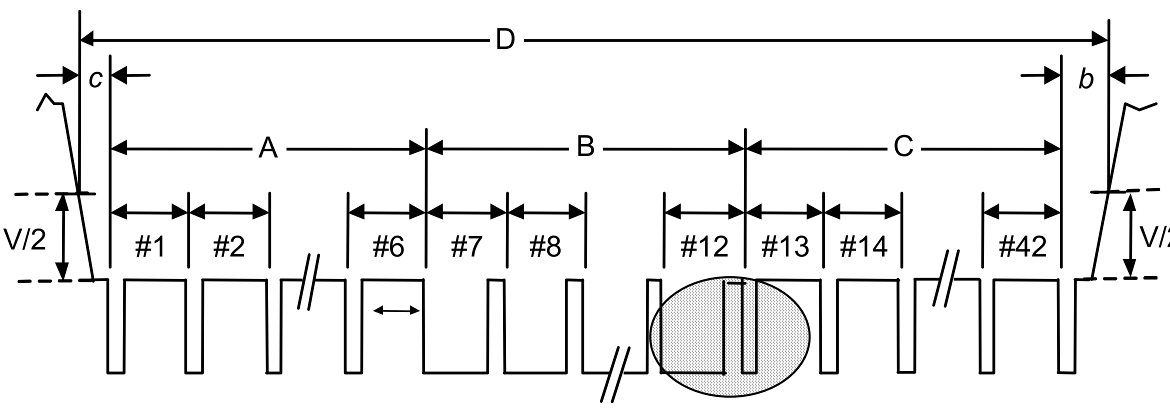
Magnified view

Timing specification of vertical synchronizing signal

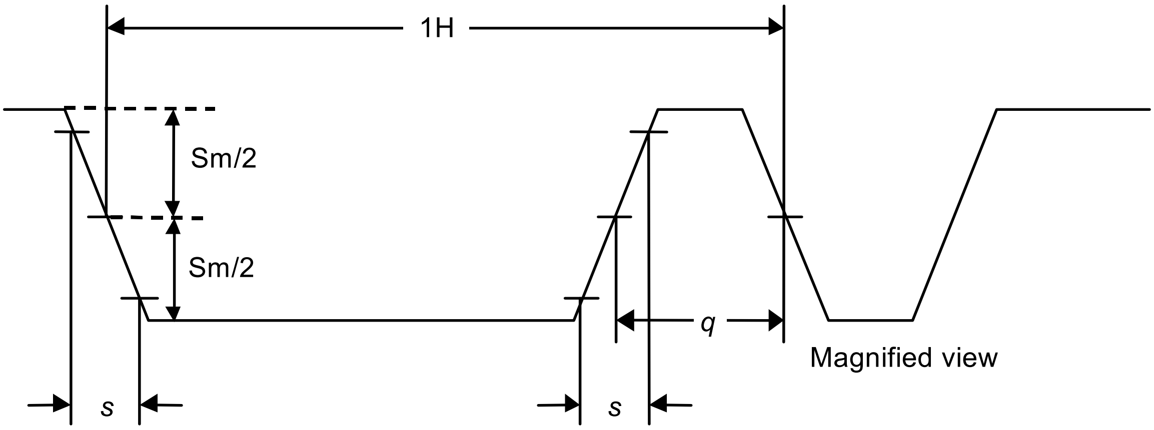
Symbol	Parameter	Nominal value
F	Vertical scanning interval (ms)	1001/30
D	Vertical blanking interval	21H+a
A	Equalizing pulse interval	3H
B	Vertical sync pulse interval	3H
C	Equalizing pulse interval	3H
s	Vertical sync pulse rise/fall time (10 - 90 %) (μs)	0.14
p	Equalizing pulse width (μs)	2.30
q	Vertical serration pulse width (μs)	4.70

Note: " H ", "*a*", "*b*", "*c*", "Sm", and " V " shall have the values shown in Appendix 5.

Appendix 8: Vertical synchronizing signal in case that number of lines is 525 and scanning system is progressive.



See the magnified view.

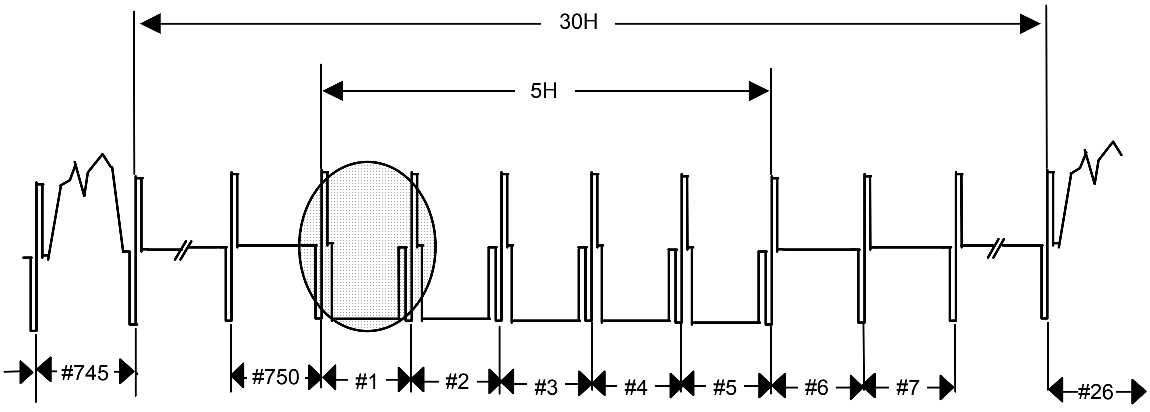


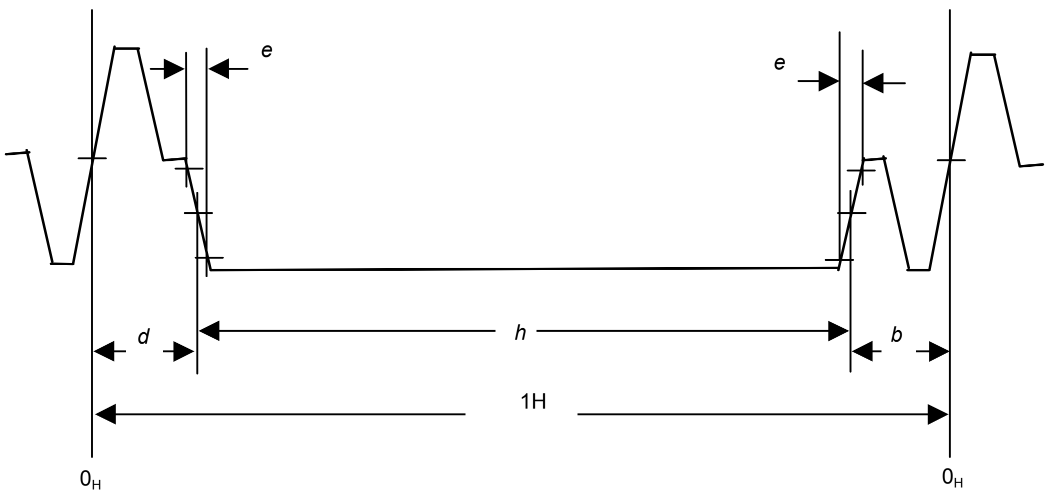
Timing specification of vertical synchronizing signal

Symbol	Parameter	Nominal value
F	Vertical scanning interval (ms)	1001/60
D	Vertical blanking interval	42H+a
A	From start of horizontal sync pulse immediately after start of vertical blanking interval to start of vertical sync pulse	6H
B	Vertical sync pulse interval	6H
C	From start of horizontal sync pulse immediately after end of vertical sync pulse to start of horizontal sync pulse immediately before end of vertical blanking interval	30H
s	Vertical sync pulse rise/fall time (10 - 90 %) (μs)	0.07
q	Vertical serratuion pulse width (μs)	2.35

Note: “ H ”, “*a*”, “*b*”, “*c*”, “Sm”, and “ V ” shall have the values shown in Appendix 5.

Appendix 9: Vertical synchronizing signal in case that number of lines is 750 and scanning system is progressive.



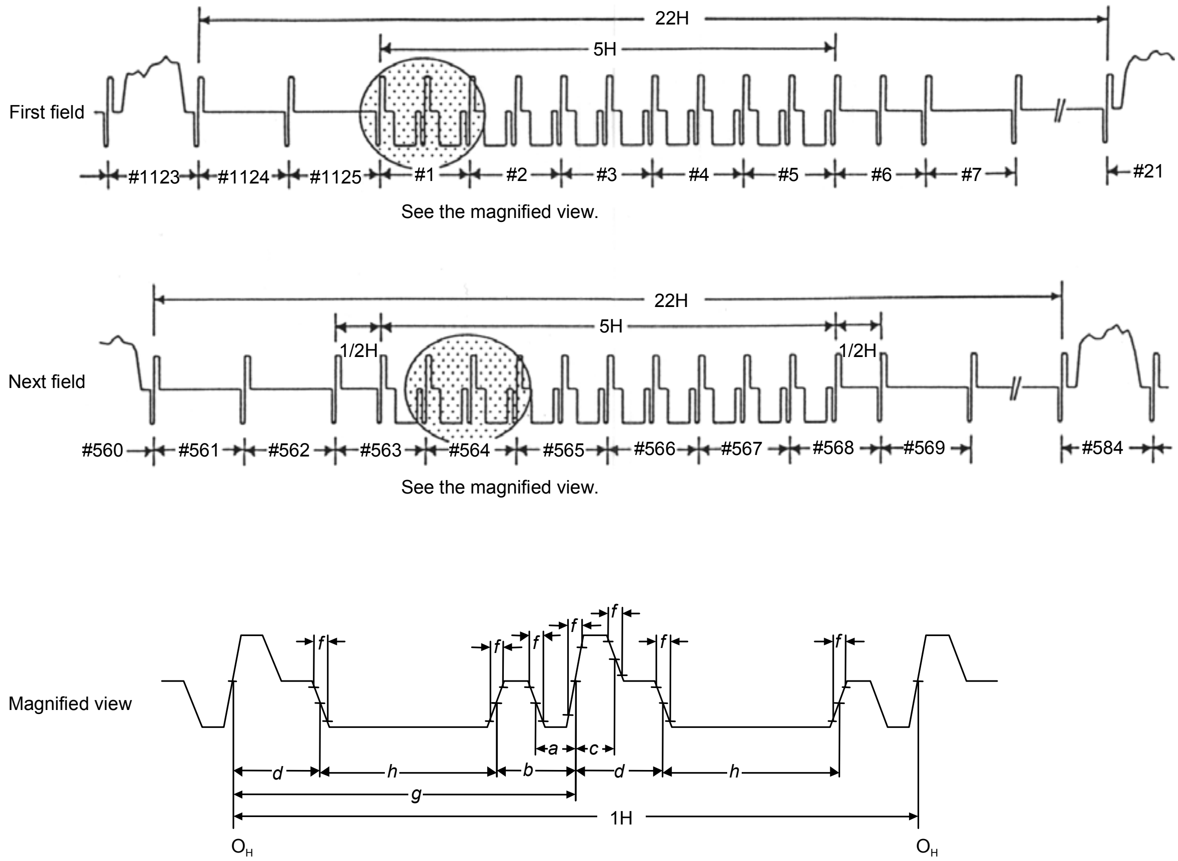


Provisions for vertical synchronizing signal and field

Symbol	Parameter	Nominal value
H	1 line interval (T)	1650
h	Vertical sync width (T)	1280
	Top line of picture	#26
	Bottom line of picture	#745
	Vertical blanking interval	30H
	Start of frame	#1

Note: " T " denotes the duration of a reference clock, and the reciprocal of the luminance sampling frequency.

Appendix 10: Field synchronizing signal in case that number of lines is 1125 and scanning system is interlaced.



Provisions for vertical synchronizing signal and field

Symbol	Parameter		Nominal value
H	1 line interval (T)		2200
g	1/2 line interval (T)		1100
h	Vertical sync pulse width (T)		880
	Top line of picture	First field	# 21
		Next field	#584
	Bottom line of picture	First field	#560
		Next field	#1123
	Vertical blanking interval	First field	22H
		Next field	23H
	Start of field	First field	# 1
		Next field	#564

Note: " T " denotes the duration of a reference clock, and the reciprocal of the luminance sampling frequency.

(Ordinance, Article 23, paragraph 4; Article 63, Paragraph 4, Appended Table 20)

2.4.2 UHDTV and progressive HDTV

Number of active lines, scanning system, frame frequency, aspect ratio of picture, and number of active samples per line (for luminance and color-difference signals) for video shall be as the table shown below.

		1080/60/P	2160/60/P	2160/120/P	4320/60/P	4320/120/P
Number of active lines		1080	2160	2160	4320	4320
Scanning system		Progressive	Progressive	Progressive	Progressive	Progressive
Frame frequency (Hz)		60/1.001, 60	60/1.001, 60	120/1.001, 120	60/1.001, 60	120/1.001, 120
Picture aspect ratio		16:9	16:9	16:9	16:9	16:9
Number of active samples per line	Luminance signal	1920	3840	3840	7680	7680
	Color difference signals	960	1920	1920	3840	3840

(Ordinance, Article 63, Paragraph 4; Article 81, No.2, Paragraph 4; Appended Table No.70; Appended Table No.71)

Chapter 3: Video Coding System

3.1 System based on MPEG-2 Video Standard

Video coding shall be achieved by a combination of the systems defined below, and video compression and transmission procedures shall comply with the notification separately issued by the Minister of Internal Affairs and Communications. (See Chapter 4, 4.1.)

(1) Motion compensated prediction coding (system in which the amount of information to be transmitted is reduced by detecting the motion vectors for previous and future frames or fields and sending two signals: (a) signal representing the difference between the original signal and motion compensated frame or field signal , and (b) motion vector information)

(2) Discrete cosine transform (system in which the amount of information to be transmitted is reduced by transforming the original picture from 8 × 8 pixels to spatial frequency components, and quantizing these frequency components in consideration of their visual characteristics.)

(3) Variable length coding (system in which the number of bits to be transmitted is reduced by representing codes that are statistically high and low in frequency of occurrence, respectively, using short and long bit strings)

(Ordinance, Article 2; Article 4, Paragraph 1)

3.2 System based on MPEG-4 AVC Standard

Video coding shall be achieved by a combination of the systems defined below, and video compression and transmission procedures shall comply with the notification separately issued by the Minister of Internal Affairs and Communications. (See Chapter 4, 4.2.)

(1) Intra-picture prediction coding (system in which the amount of information to be transmitted is reduced by coding the difference between the pixel to be coded and the neighbor pixel in the original signal)

(2) Motion compensated prediction coding (system in which the amount of information to be transmitted is reduced by detecting the motion vectors for previous and future frames or fields and sending two signals: (a) signal representing the difference between the original signal and motion compensated frame or field signal , and (b) motion vector information)

(3) Integer transform (system in which the amount of information to be transmitted is reduced by transforming the original picture from 4 × 4 pixels or 8 × 8 pixels to spatial frequency components by orthogonal transform with integer precision, and quantizing these frequency components in consideration of their visual characteristics.)

(4) Entropy coding (system in which the number of bits to be transmitted is reduced by representing codes as different bit sequences in accordance with the appearance probability of codes)

(Ordinance, Article 2; Article 24, No.5, Paragraph 1)

3.3 System based on HEVC Standard

Video coding shall be achieved by a combination of the systems defined below, and video compression and transmission procedures shall comply with the notification separately issued by the Minister of Internal Affairs and Communications. (See Chapter 4, 4.3.)

(1) Intra-picture prediction coding (system in which the amount of information to be transmitted is reduced by coding the difference between the pixel to be coded and the neighbor pixel in the original signal)

(2) Motion compensated prediction coding (system in which the amount of information to be transmitted is reduced by detecting the motion vectors for previous and future frames or fields and sending two signals: (a) signal representing the difference between the original signal and motion compensated frame or field signal , and (b) motion vector information)

(3) Integer transform (system in which the amount of information to be transmitted is reduced by transforming the original picture from 4 × 4 pixels or 8 × 8 pixels to spatial frequency components by orthogonal transform with integer precision, and quantizing these frequency components in consideration of their visual characteristics.)

(4) Entropy coding (system in which the number of bits to be transmitted is reduced by representing codes as different bit sequences in accordance with the appearance probability of codes)

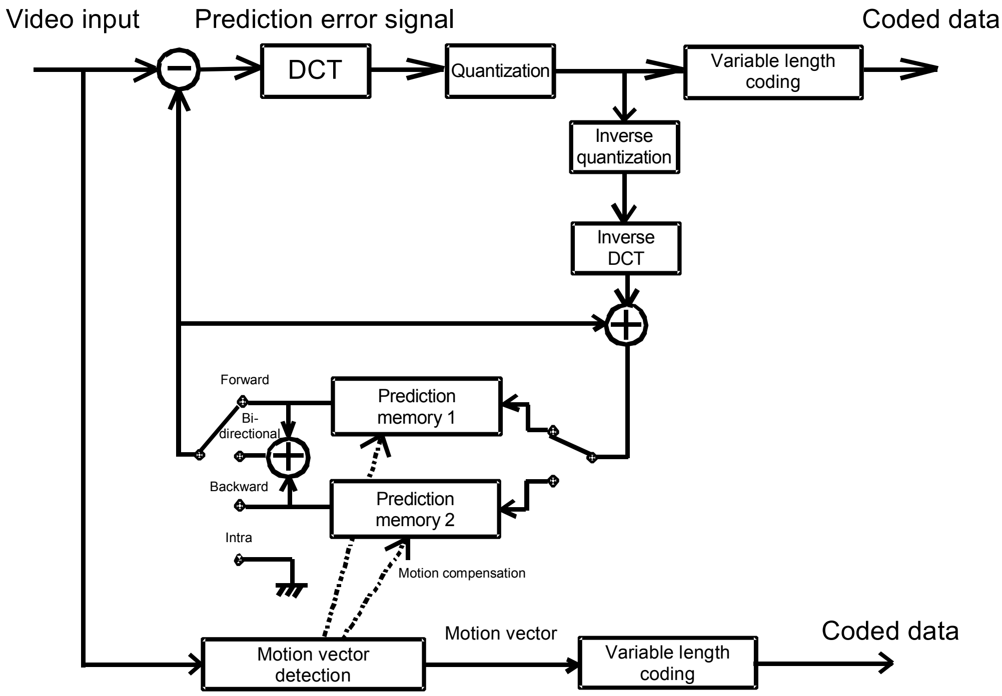
(5) Pixel adaptive offset filtering (system in which picture quality is improved by adding offset in accordance with the pixel value after de-blocking filtering)

(Ordinance, Article 2; Article 62, Paragraph 2)

Chapter 4: Video Compression Procedure, Transmission Procedure, and Signal Configuration after Coding

4.1 System based on MPEG-2 Standard

4.1.1 Compression and transmission procedures



Notes:

1. DCT represents a discrete cosine transform in which two-dimensional DCT coefficients F (u, v) for N × N pixels f (x, y) are defined as the followings when the horizontal and vertical directions of the picture are assumed to be the x and y axes, respectively:

$$F(u,v) = \frac{2C(u)C(v)}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left\{\frac{(2x+1)u\pi}{2N}\right\}\cos\left\{\frac{(2y+1)v\pi}{2N}\right\}$$

Provided that

$$C(u), C(v) = \begin{cases} 1/\sqrt{2} & \text{for } u, v = 0 \\ 1 & \text{for } u, v \neq 0 \end{cases}$$

2. Inverse DCT represents an inverse discrete cosine transform and is defined as the followings:

$$f(x,y) = \frac{2}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} C(u)C(v)F(u,v)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N}$$

3. In the figure shown above, "Forward" represents forward prediction coding in which motion compensation is based on past picture information. "Bi-directional" denotes bi-directional prediction coding in which motion compensation is based on future and past picture information. "Backward" refers to backward prediction coding in which motion compensation is based on future picture information. "Intra" represents

intra-coding in which no prediction is performed and in which only the current picture information is used.

4. Inverse quantization and variable length coding shall comply with ITU-T Rec. H.262. Note that the order of output data of a variable length coder shall be one of the following:

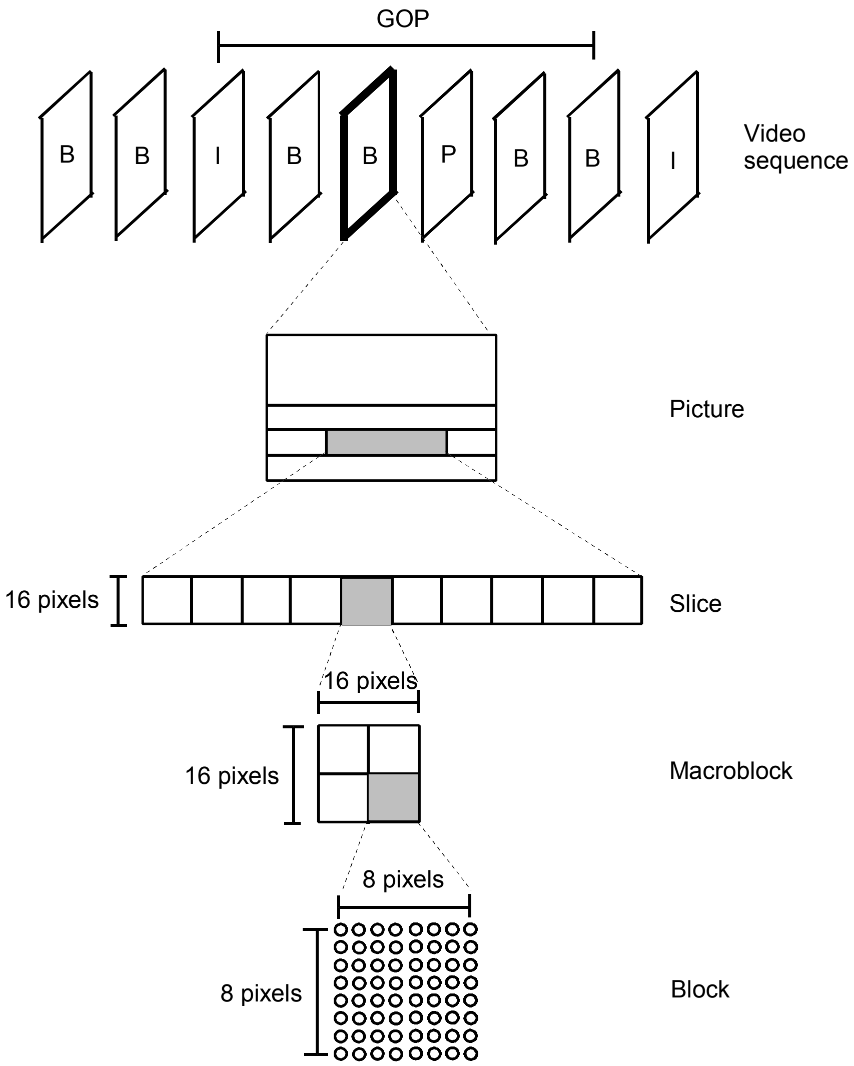
								u
	0	1	2	3	4	5	6	7
0	0	1	5	6	14	15	27	28
1	2	4	7	13	16	26	29	42
2	3	8	12	17	25	30	41	43
3	9	11	18	24	31	40	44	53
4	10	19	23	32	39	45	52	54
5	20	22	33	38	46	51	55	60
6	21	34	37	47	50	56	59	61
ν 7	35	36	48	49	57	58	62	63

								u
	0	1	2	3	4	5	6	7
0	0	4	6	20	22	36	38	52
1	1	5	7	21	23	37	39	53
2	2	8	19	24	34	40	50	54
3	3	9	18	25	35	41	51	55
4	10	17	26	30	42	46	56	60
5	11	16	27	31	43	47	57	61
6	12	15	28	32	44	48	58	62
ν 7	13	14	29	33	45	49	59	63

5. Motion vector detection shall be conducted for each macroblock.
6. Coded data shall be generated in compliance with the video bitstream syntax given in ITU-T Rec. H.262.

(Notification, Appended Table 1)

4.1.2 Signal Configuration



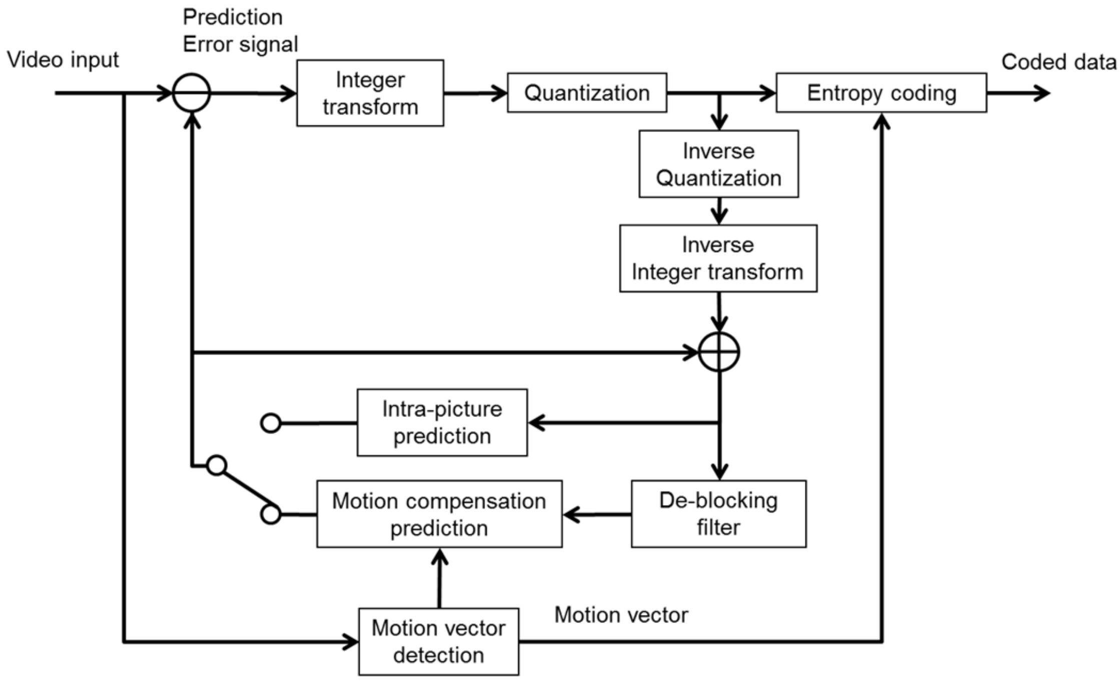
Notes:

1. Video sequence is the highest syntactic configuration for video coding and refers to a series of images that comprise a video signal.
2. GOP consists of I-pictures (pictures encoded using only current picture information), B-pictures (pictures encoded using current, past and future picture information) and P-pictures (pictures encoded using current and past picture information) and contains at least one I-picture.
3. A picture refers to a single image.
4. A slice consists of an arbitrary number of macroblocks in the same horizontal row.
5. A macroblock consists of a luminance signal of 16 × 16 pixels and two color difference signals of spatially corresponding to 8 × 8 or 16 × 8 pixels.

(Notification, Appended Table 2)

4.2 System based on MPEG-4 AVC Standard

4.2.1 Compression and transmission procedures



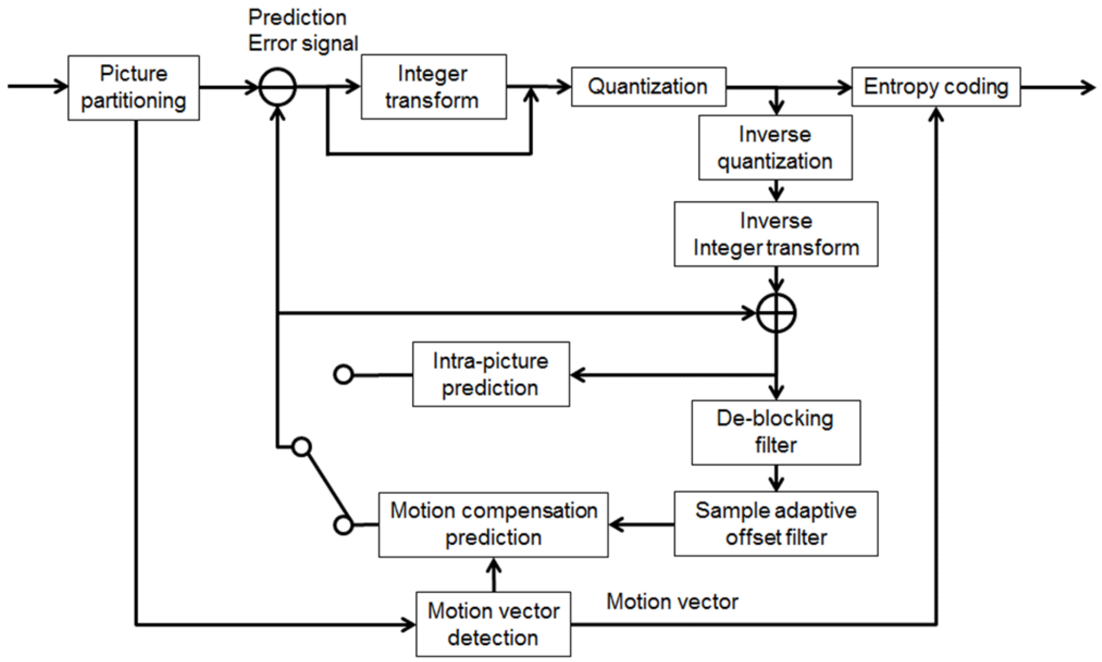
Notes：

1. Integer transform represents the procedure of orthogonal transform with the integer precision for block of 4×4 pixels and 8×8 pixels.
2. Intra-picture prediction represents the procedure of prediction by using neighbor block in which coding has been finished in present picture information, and motion compensation prediction represents the procedure of prediction with variable block size by using future and past picture information.
3. De-blocking filter represents the procedure to reduce noise which occurs in the boundary of block in accordance with coding.
4. Inverse quantization, inverse integer transform and entropy coding shall comply with ITU-T Rec. H.264.
5. Motion vector detection is conducted for each block of 4×4 pixels, 4×8 pixels, 8×4 pixels, 8×16 pixels, 16×8 pixels and 16×16 pixels.
6. Coded data is generated in compliance with video bitstream syntax given in ITU-T Rec. H.264.

(Notification, Appended Table 3)

4.3 System based on HEVC Standard

4.3.1 Compression and transmission procedures



Notes：

1. Picture partitioning represents the procedure of partitioning to square domain of 8×8 pixels, 16×16 pixels, 32×32 pixels and 64×64 pixels.
2. Integer transform represents the procedure of orthogonal transform with the integer precision for block of 4×4 pixels, 8×8 pixels, 16×16 pixels and 32×32 pixels.
3. Intra-picture prediction represents the procedure of prediction by using neighbor block in which coding has been finished in present picture information, and motion compensation prediction represents the procedure of prediction with variable block size by using future and past picture information.
4. De-blocking filter represents the procedure to reduce noise which occurs in the boundary of block in accordance with coding.
5. Pixel adaptive offset filter represents the procedure to reduce noise which occurs inside the block in accordance with coding.
6. Inverse quantization, inverse integer transform and entropy coding shall comply with ITU-T Rec. H.265.
7. Motion vector detection is conducted for each block of 4×8 pixels, 4×16 pixels, 8×4 pixels, 8×8 pixels, 8×16 pixels, 8×32 pixels, 12×16 pixels, 16×4 pixels, 16×8 pixels, 16×12 pixels, 16×16 pixels, 16×32 pixels, 16×64 pixels, 24×32 pixels, 32×8 pixels, 32×16 pixels, 32×24 pixels, 32×32 pixels, 32×64 pixels, 48×64 pixels, 64×16 pixels, 64×32 pixels, 64×48 pixels and 64×64pixels.
8. Coded data is generated in compliance with video bitstream syntax given in ITU-T Rec. H.265.

(Notification, Appended Table 4)

<Blank Page>

Chapter 5: Restrictions on Coding Parameters

5.1 Restrictions on video coding parameters for television services

5.1.1 System based on MPEG-2 Video Standard

Video coding shall conform to the Main Profile syntax defined in the MPEG-2 Video Standard. Additionally, “Set 1 of the coding parameter constraints” given in Table 5-1 shall be met if the display area is not specified by sequence_display_extension, while “Set 2 of the coding parameter constraints” given in Table 5-2 shall be satisfied if the display area is specified by sequence_display_extension. The Main Profile syntax values defined in the MPEG-2 Video Standard shall be used for parameters not listed in this standard as constraints.

Also note that Table 5-3 shows the meanings of code numbers assigned to MPEG-2 Video Standard coding parameters in Table 5-1 and Table 5-2, and desirable display formats on monitors with 4:3 and 16:9 aspect ratios for each parameters values, respectively in Fig. 5-1.

On the transmission side, vbv_delay shall always be set to 0xFFFF, and the operation is at a variable bit-rate. Video PES shall consist of video data of a single frame, and PTS (or DTS as necessary) shall always be transmitted in the PES Header. In the receiver, start control and output control of video and audio decoding shall be performed by PTS or DTS within each PES Header. Decoding control shall not be performed by vbv_delay.

Table 5-1: Set 1 of coding parameter restrictions (when the display area is not specified by sequence_display_extension)

<table>
<tr><th colspan="4">sequence_header restriction</th><th>sequence_ extension restriction</th><th colspan="3">sequence_delay_ extension restriction (Note 4)</th><th rowspan="2">Other parameters (Note 7)</th><th rowspan="2">Figure 5-1 (Note 8)</th></tr>
<tr><th>vertical_ size_value</th><th>horizontal_ size_value</th><th>aspect_ratio_information</th><th>frame_rate_ code</th><th>progressive_ sequence</th><th>color_ primaries</th><th>transfer_ characteristics</th><th>matrix_ coefficients</th></tr>
<tr><td>1080 (Note 1)</td><td>1920, 1440</td><td>3</td><td>4 (Note 3)</td><td>0</td><td rowspan="2">1 (Note 5)</td><td rowspan="2">1 (Note 5)</td><td rowspan="2">1 (Note 5)</td><td rowspan="2">Nominal value for MP@HL</td><td>①</td></tr>
<tr><td>720</td><td>1280</td><td>3</td><td>7 (Note 3)</td><td>1</td><td>①</td></tr>
<tr><td>480</td><td>720</td><td>3</td><td>7 (Note 3)</td><td>1</td><td rowspan="3">1
(Note 5, Note 6)</td><td rowspan="3">1
(Note 5, Note 6)</td><td rowspan="3">1
(Note 5, Note 6)</td><td>Nominal value for MP@H14L</td><td>①</td></tr>
<tr><td rowspan="2">480</td><td rowspan="2">720, 544, 480 (Note 2)</td><td>3</td><td rowspan="2">4 (Note 3)</td><td rowspan="2">0</td><td rowspan="2">Nominal value for MP@ML</td><td>①</td></tr>
<tr><td>2</td><td>①</td></tr>
</table>

Note 1: A total of 1088 lines is actually coded in the MPEG-2 Video Standard. That is, the encoder adds 8 lines of fictional video data (dummy data) below the active lines. Therefore, 1088 lines of video data are actually coded. The decoder discards the dummy data from these 1088 lines of video data, outputting only the upper 1080 active lines.

Note 2: To ensure media crossover and allow preparation for flexible future operations, 544 and 480 samples are also available as horizontal_size_value. However, due to the high-quality services required of digital broadcasting, 720 samples are desirable when possible. When 544 samples are used, the center position shall be aligned to that for 720 samples. Additionally, these 544 samples shall consist of 540 samples of actual video data and two samples of fictional video data (basically black) on each side of the actual video data.

Note 3: For film materials, the repeat_first_field, top_field_first and progressive_frame flags shall be controlled to allow encoding without changing frame_rate_code. (See Attachment 1, Chapter 5.)

Note 4: If sequence_display_extension is not transmitted, display_vertical_size and display_horizontal_size are assumed by the receiving side to be equal, respectively, to the vertical_size_value and horizontal_size_value specified by sequence_header. However, note that if horizontal_size_value is 544 samples, the receiving side displays the area of 540 samples while excluding two samples on each side, as when display_horizontal_size is transmitted as 540 samples.

Note 5: If sequence_display_extension is not transmitted, color_primaries, transfer_characteristics and matrix_coefficients are assumed by the receiving side to be equal to “1.”

Note 6: For narrow band CS digital broadcast, this shall depend on the provision of operation by broadcast company.

Note 7: Nominal values given in the MPEG-2 Video Standard are used for Main Profile levels. However, note that bit_rate_value shall be the maximum transferable rate or less for MP@HL and MP@H14L, while it shall be 15 Mbps or less for MP@ML. A variable bitrate is required, and vbv_delay shall always be set to 0xFFFF.

Note 8: See “Desirable display formats on 4:3 and 16:9 aspect ratio monitors” in Fig. 5-1.

Table 5-2: Set 2 of the coding parameter restrictions (when the display area is specified by sequence_display_extension)

Parameter value of sequence_header				Parameter value of sequence_extension	Parameter value of sequence_display_extension					Other parameters (Note 8) (Note 7)	Figure 5-1 (Note 10)
vertical_size_value	horizontal_size_value	aspect_ratio_information (Note 3)	frame_rate_code (Note 4)	progressive_sequence	display_vertical_size	display_horizontal_size (Note 5)	color_primaries	transfer_characteristics	matrix_coefficients		
1080 (Note 1)	1920, 1440	3	4	0	1080	1920. 1440	1 (Note 6)	1 (Note 6)	1 (Note 6)	Nominal value for MP@HL	①
	1920	2				1440					②
	1440					1080					
720	1280	3	7	1	720	1280					①
		2				960					②
480	720	3	7	1	480	720	1 (Note 6, Note 7)	1 (Note 6, Note 7)	1 (Note 6, Note 7)	Nominal value for MP@HL	①
		2				540					②
480	720, 544, 480 (Note 2)	3	4	0	480	720, 540, 480				Nominal value for MP@HL	①
	720	2				540					②
	720, 544, 480 (Note 2)	3				720, 540, 480					③
		2			360						④

Note 1: A total of 1088 lines is actually coded in the MPEG-2 Video Standard. That is, the encoder adds 8 lines of fictional video data (dummy data) below the active lines. Therefore, 1088 lines of video data are actually coded. The decoder discards the dummy data from these 1088 lines of video data, outputting only the upper 1080 active lines.

Note 2: To ensure media crossover and allow preparation for flexible future operations, 544 and 480 samples are also available as horizontal_size_value. Due to the high-quality services required of digital broadcasting, 720 samples shall be used when possible. When 544 samples are used, the center position shall be aligned to that for 720 samples. Additionally, these 544 samples shall consist of 540 samples of active video data and two samples of fictional video data (basically black) on each side of active video data.

Note 3: The MPEG-2 Video Standard stipulates that aspect_ratio_information represent the aspect ratio of the area specified by display_vertical_size and display_horizontal_size when sequence_display_extension is transmitted.

Note 4: For film materials, the repeat_first_field, top_field_first and progressive_frame flags shall be controlled to allow encoding without changing frame_rate_code. (See Chapter 5 in the Appendix.)

Note 5: If there are two or more numbers in a box under display_horizontal_size, this means that of those numbers only the same value as that of horizontal_size_value can be selected, except where horizontal_size_value is 544, in which case 540 can be selected.

Note 6: If color_primaries, transfer_characteristics or matrix_coefficients (sequence_display_extension parameters) is not transmitted, the value of the parameter that is not transmitted is assumed by the receiving side to be equal to "1."

Note 7: In case of narrow band CS digital broadcasting, operational rule shall be applied.

Note 8: The nominal values given in the MPEG-2 Video Standard are used for Main Profile levels. However, note that bit_rate_value shall be equal to or less than the maximum transferable rate for MP@HL and MP@H14L, and 15 Mbps or less for MP@ML. A variable bitrate shall be used, and vbv_delay shall always be set to 0xFFFF.

Note 9: Ideally, receiver functionality shall be examined before using frame_center_horizontal_offset (FCHO) and frame_center_vertical_offset (FCVO) (picture_display_extension parameters). If picture_display_extension is not transmitted, FCHO and FCVO are assumed by the receiving side to be "0."

Note 10: See "Desirable display formats on 4:3 and 16:9 aspect ratio monitors" in Fig. 5-1.

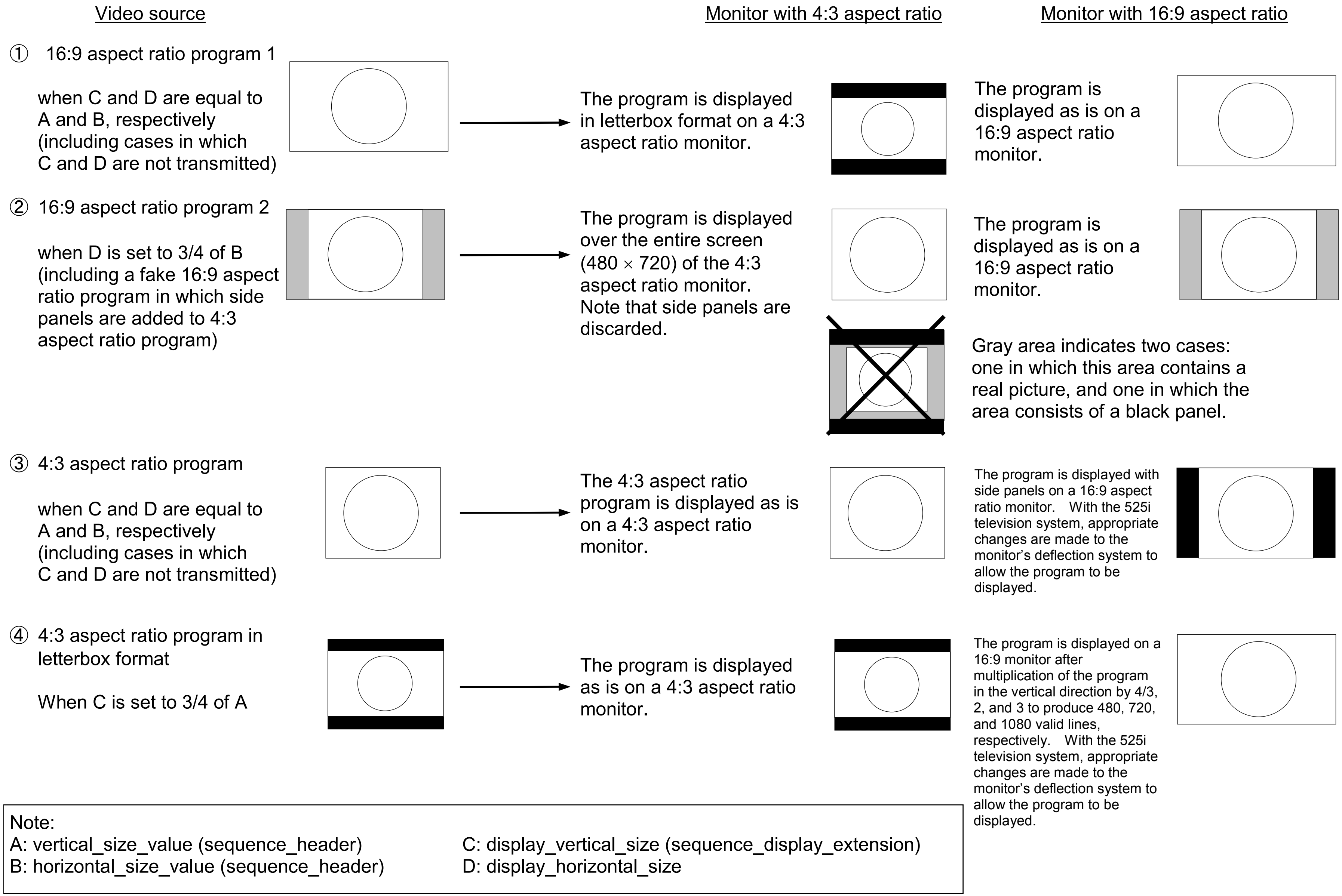


Fig. 5-1: Desirable display formats on 4:3 and 16:9 aspect ratio monitors

Table 5-3: Meanings of each code number of coding parameter in MPEG-2 Video Standard in Table 5-1 and Table 5-2

aspect_ratio_information	2 = 4:3 display	3 = 16:9 display
frame_rate_code	4 = 30/1.001 Hz	7 = 60/1.001 Hz
progressive_sequence	0 = interlaced	1 = progressive
color_primaries	1 = nominal value in Rec. ITU-R BT.709 (BT.1361)	
transfer_characteristics	1 = nominal value in Rec. ITU-R BT.709 (BT.1361)	
matrix_coefficients	1 = nominal value in Rec. ITU-R BT.709 (BT.1361)	

5.1.2 System based on MPEG4-AVC Standard

5.1.2.1 Profile and Level

Video coding system shall be based on High 4:2:2 Profile (including Main, High, High10 Profile). And according to video format, any of Level 3, 3.1, and 3.2 for 480/60/I, either Level 3.1 or 3.2 for 480/60/P, either Level 3.2 or 4 for 720/60/P, Level 4 for 1080/60/I, Level 4.2 for 1080/60/P and Level 5.2 for 2160/60/P shall be adopted.

5.1.2.2 Video Coding Format

4:2:0 or 4:2:2 of Y', C'_B, C'_R shall be used. Sampling point is the same as those provided in MPEG-2 Video standard. Three primary colors, gamma characteristics, matrix for luminance and color difference signals shall be based on the provision of Chapter 2, 2.1.1. Bit number of each sample is 8 bits or 10 bits.

The combination of video format to be coded, Profile and Level shall be as shown in Table 5-4.

Table 5-4: Combination of Video coding format and Profile, Level

Input video format	Chroma format	Bit precision (bit)	Number of horizontal pixels	Number of Vertical pixels	Frame rate (Hz)	Scanning system	Picture aspect ratio	Profile	Level
480/60/I	4:2:0	8	720	480	29.97	Interlaced	4 : 3	Main, High	3, 3.1, 3.2
	4:2:0	8	720	480	29.97	Interlaced	16 : 9	Main, High	3, 3.1, 3.2
480/60/P	4:2:0	8	720	480	59.94	Progressive	16 : 9	Main, High	3.1, 3.2
720/60/P	4:2:0	8	1280	720	59.94	Progressive	16 : 9	Main, High	3.2, 4
1080/60/I	4:2:0	8	1440	1080	29.97	Interlaced	16 : 9	Main, High	4
	4:2:0	8	1920	1080	29.97	Interlaced	16 : 9	Main, High	4
	4:2:0	10	1920	1080	29.97	Interlaced	16 : 9	High10	4
	4:2:2	8, 10	1920	1080	29.97	Interlaced	16 : 9	High4:2:2	4
1080/60/P	4:2:0	8	1920	1080	59.94	Progressive	16 : 9	High	4.2
	4:2:0	10	1920	1080	59.94	Progressive	16 : 9	High10	4.2
	4:2:2	8, 10	1920	1080	59.94	Progressive	16 : 9	High4:2:2	4.2
2160/60/P *1	4:2:0	8	3840	2160	59.94	Progressive	16 : 9	High	5.2
	4:2:0	10	3840	2160	59.94	Progressive	16 : 9	High10	5.2
	4:2:2	8, 10	3840	2160	59.94	Progressive	16 : 9	High4:2:2	5.2

*1 Practical operation is limited to the coverage provided in Ordinance and Notification. (See Annex A)

5.1.2.3 Syntax

(1) Sequence parameter set

Syntax element	Value	Remarks
profile_idc	77, 100, 110, or 122	77 : Main profile 100 : High profile 110 : High 10 profile 122 : High 4:2:2 profile (Note) See Table 5-4 for combination of video format
level_idc	30, 31, 32, 40, 42, or 52	30: level 3 31: level 3.1 32: level 3.2 40: level 4 42: level 4.2 52: level 5.2 (Note) See Table 5-4 for combination of video format
chroma_format_idc	1 or 2	1: 4:2:0 format 2: 4:2:2 format
bit_depth_luma_minus8	0 or 2	0: Luminance pixel has 8 bit precision. 2: Luminance pixel has 10 bit precision.
bit_depth_chroma_minus8	0 or 2	0: Chroma pixel has 8 bit precision. 2: Chroma pixel has 10 bit precision.
pic_width_in_mbs minus1	See Table 5-5	This represents macro block number-1 in horizontal direction.
pic_height_in_map_unit_minus1	See Table 5-5	This represents macro block number-1 in vertical direction.
frame_mbs_only_flag	0 or 1 See Table 5-5 and Table 5-6	1: only frame macro block 0: field macro block or MBAFF is permitted. mb_adaptive_frame_field_flag is set to 0 or 1. Set to 1 only for progressive video
frame_cropping_flag	See Table 5-6	0: display all decoded video 1: display a part of decoded and partitioned video
frame_crop_left_offset	See Table 5-6	This represents a half of the number of leftmost pixels which are not displayed in decoded video.
frame_crop_right_offset	See Table 5-6	This represents a half of the number of rightmost pixels which are not displayed in decoded video.
frame_crop_top_offset	See Table 5-6	This represents a half or a quarter of the number of topmost pixels which are not displayed in decoded video.
frame_crop_bottom_offset	See Table 5-6	This represents a half or a quarter of the number of bottommost pixels which are not displayed in decoded video.
vui_parameters_present_flag	1	1: decode VUI (Video Usability Information)

(2) VUI

Syntax element	Value	Remarks
aspect_ratio_info_present_flag	1	Information of aspect ratio is essential.
aspect_ratio_idc	See Table 5-6	This represents pixel aspect ratio.
sar_width	4	When resolution is 1440×1080, if aspect_ratio_idc=255, this syntax is essential.
sar_height	3	When resolution is 1440×1080, if aspect_ratio_idc=255, this syntax is essential.
video_full_range_flag	0	0: Based on Rec. ITU-R BT.709-5
colour_primaries	1	1: Based on Rec. ITU-R BT.709-5
transfer_characteristics	1 or 11	1: Based on Rec. ITU-R BT.709-5, Rec. ITU-R BT.1361 conventional color gamut system 11: Based on IEC61966-2-4 (wide color gamut system)
matrix_coefficients	1	1: Based on Rec. ITU-R BT.709-5
chroma_loc_info_present_flag	0	0: the same as the sample position of 4:2:0 color difference signals in MPEG-2 Video Standard
timing_info_present_flag	1	1: This represents the frame rate for fixed frame rate. This includes num_units_in_tick,time_scale, fixed_frame_rate_flag in syntax elements. Frame-rate = time_scale / num_units_in_tick / 2 Note: About calculating frame rate in detail, see the semantics of fixed_frame_rate_flag in Annex E of MPEG-4 AVC Standard.
num_units_in_tick	1001	1001 fixed
time_scale	60000 or 120000	When frame rate is 29.97Hz, this shall be set to 60000. And when it is 59.94Hz, this shall be set to 120000.

Table 5-5: Combination of parameters representing picture size (No.1)

Number of horizontal pixels	Number of vertical pixels	pic_width_in_mbs_minus1	pic_height_in_map_units_minus1	frame_mbs_only_flag	Frame rate (Hz)	Scanning
720	480	44	14	0	29.97	Interlaced
720	480	44	29	1	59.94	Progressive
1280	720	79	44	1	59.94	Progressive
1440	1080	89	33	0	29.97	Interlaced
1920	1080	119	33	0	29.97	Interlaced
1920	1080	119	67	1	59.94	Progressive
3840	2160	239	134	1	59.94	Progressive

Table 5-6: Combination of parameters representing picture size (No.2)

Picture aspect ratio	Number of horizontal pixels	Number of vertical pixels	aspect_ratio_idc	frame_mbs_only_flag	frame_cropping_flag	frame_crop_left_offset	frame_crop_right_offset	frame_crop_top_offset	frame_crop_bottom_offset
4:3	720	480	3	0	0	0	0	0	0
16:9	720	480	5	0	0	0	0	0	0
16:9	720	480	5	1	0	0	0	0	0
16:9	1280	720	1	1	0	0	0	0	0
16:9	1440	1080	255 or 14 (Note)	0	1	0	0	0	2
16:9	1920	1080	1	0	1	0	0	0	2
16:9	1920	1080	1	1	1	0	0	0	4
16:9	3840	2160	1	1	0	0	0	0	0

(Note) Operation by 255 is desirable because aspect_ratio_idc=14 is not provided in the early standard of MPEG-4 AVC Standard.

(3) Pan-scan rectangle SEI

When video signal is transmitted with different aspect ratio from original video source such as side panel or letter box, it is possible to avoid displaying black frame (picture frame) by setting parameters of pan-scan as the following. So according to Fig. 5-2 “Desirable display form for monitor with aspect ratio 4:3/16:9”, Pan-scan rectangle SEI is coded on head I-picture (IDR picture for closed GOP, and I-picture of non-IDR for open GOP) as necessity.

Also, when Pan-scan is operated (② or ④ in Fig. 5-2), Pan-scan rectangle SEI must be coded.

The values of each parameter in operation mentioned above are shown in Table 5-7 and Table 5-8.

Table 5-7: Parameters for Pan-scan operation

		Parameters of VUI	Parameters of Sequence parameter set			Parameters of Pan-scan rectangle SEI				
Picture width	Picture height	aspect_ratio_idc	pic_width_in_mbs_minus1	pic_height_in_map_units_minus1	frame_mbs_only_flag	pan_scan_rect_left_offset	pan_scan_rect_right_offset	pan_scan_rect_top_offset	pan_scan_rect_bottom_offset	Figure 5-2
720	480	5	44	29	1	0	0	0	0	①
720	480	5	44	29	1	1440	−1440	0	0	②
720	480	5	44	14	0	0	0	0	0	①
720	480	5	44	14	0	1440	−1440	0	0	②
720	480	3	44	14	0	0	0	0	0	③
720	480	3	44	14	0	0	0	960	−960	④
1280	720	1	79	44	1	0	0	0	0	①
1280	720	1	79	44	1	2560	−2560	0	0	②
1440	1080	255 (sar_width=4, sar_height=3)	89	33	0	0	0	0	0	①
1440	1080	255 (sar_width=4, sar_height=3)	89	33	0	2880	−2880	0	0	②
1920	1080	1	119	33	0	0	0	0	0	①
1920	1080	1	119	33	0	3840	−3840	0	0	②
1920	1080	1	119	67	1	0	0	0	0	①
1920	1080	1	119	67	1	3840	−3840	0	0	②
3840	2160	1	239	134	1	0	0	0	0	①
3840	2160	1	239	134	1	7680	−7680	0	0	②

Table 5-8: Another syntax elements of Pan-scan SEI

Syntax elements	Value	Remarks
Pan_scan_rect_id	0	Pan-scan information is not discriminated by ID.
Pan_scan_rect_cancel_flag	0	Pan-scan information is always transmitted.
Pan_scan_cnt_minus1	0	Pan-scan information is only one kind.
Pan_scan_rect_repetition_period	1	Pan-scan information is effective until next sequence, or immediately before next picture to which Pan-scan SEI is added.

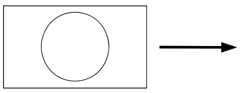
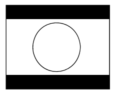
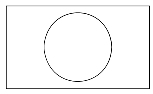
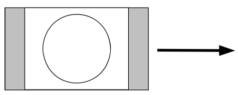
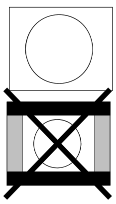
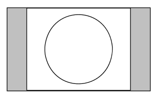
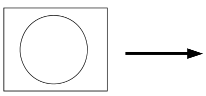
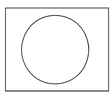
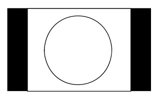
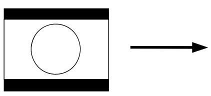
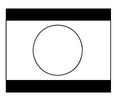
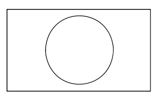
	Video source	Monitor with 4:3 aspect ratio	Monitor with 16:9 aspect ratio
① 16:9 program No.1		The program is displayed in letterbox format on a 4:3 monitor. 	The program is displayed as is on a 16:9 monitor. 
② 16:9 program No.2		The program is displayed over the entire screen (480×720)of the 4:3 monitor. Note that side panels are discarded. 	The program is displayed as is on a 16:9 monitor.  Gray area indicates two cases: one in which this area contains a real picture and one is in which the area consists of a black panel.
③ 4:3 program		The 4:3 program is displayed as is on a 4:3 monitor. 	The program is displayed with side panels on a 16:9 monitor. With 480/I system, appropriate change are made to the monitor's deflection system to allow the program to be displayed. 
④ 4:3 program in letterbox format		The program is displayed as is on a 4:3 monitor. 	The program is displayed on a 16:9 monitor after multiplication of the program in the vertical deflection by 4/3, 2, and 3 to produce 480, 720, and 1080 valid lines respectively. With the 480/I system, appropriate change are made to the monitor's deflection system to allow the program to be displayed. 

Fig. 5-2: Desirable display formats on 4:3 and 16:9 aspect ratio monitors

5.1.3 System based on HEVC Standard

5.1.3.1 Profile, Tier, and Level

Video coding system shall be based on Main 10 Profile (including Main Profile) and Main Tier provided in HEVC Standard. In case that coding pixel bit number is 10 bits, Main 10 Profile shall be used and in case that coding pixel bit number is 8 bits, Main Profile shall be used. According to video format, Level 4.1 for 1080/60/I, Level 4.1 for 1080/60/P, Level 5.1 for 2160/60/P, Level 5.2 for 2160/120/P, Level 6.1 for 4320/60/P, and Level 6.2 for 4320/120/P shall be adopted.

For 2160/120/P and 4320/120/P, temporal scalable coding which is always consisted of two sub-layers. Lower sub-layer (sub-bitstream) is equivalent to 60Hz (or 60/1.001Hz) video based on Level 5.1 (for 2160/120/P) or Level 6.1 (for 4320/120/P).

5.1.3.2 Video Coding Format

4:2:0 of Y', C'_B, C'_R shall be used. The sample position of color difference signals is the same as those in UHDTV studio standard. Three primary colors of SDR-TV, gamma characteristics, and matrix for luminance and color difference signals are based on the provision in Chapter 2, 2.1.1 or 2.1.3 for HDTV, and in Chapter 2, 2.1.3 for UHDTV. Bit number of each sample in SDR-TV is 8 bits or 10 bits for HDTV (although 10 bits in case of using video signal characteristics in Chapter 2, 2.1.3), and 10 bits for UHDTV. Bit number of each sample in HDR-TV is 10 bits.

The combination of video format to be coded, profile and level is shown in Table 5-9.

Table 5-9: Combination of video coding format, Profile and Level

Input video format	Chroma format	Bit precision (bit)	Number of horizontal pixels	Number of Vertical pixels	Frame rate (Hz)	Scanning	Display aspect ratio	SDR/HDR	Profile	Level
1080/60/I	4:2:0	8	1920	1080	30/1.001	Interlaced	16 : 9	SDR-TV	Main	4.1
	4:2:0	8	1920	1080	30	Interlaced	16 : 9		Main	4.1
	4:2:0	10	1920	1080	30/1.001	Interlaced	16 : 9	SDR-TV or HDR-TV	Main 10	4.1
	4:2:0	10	1920	1080	30	Interlaced	16 : 9		Main 10	4.1
1080/60/P	4:2:0	8	1920	1080	60/1.001	Interlaced	16 : 9	SDR-TV	Main	4.1
	4:2:0	8	1920	1080	60	Interlaced	16 : 9		Main	4.1
	4:2:0	10	1920	1080	60/1.001	Interlaced	16 : 9	SDR-TV or HDR-TV	Main 10	4.1
	4:2:0	10	1920	1080	60	Progressive	16 : 9		Main 10	4.1
2160/60/P	4:2:0	10	3840	2160	60/1.001	Progressive	16 : 9		Main 10	5.1
	4:2:0	10	3840	2160	60	Progressive	16 : 9		Main 10	5.1
2160/120/P	4:2:0	10	3840	2160	120/1.001	Progressive	16 : 9		Main 10	5.2
	4:2:0	10	3840	2160	120	Progressive	16 : 9		Main 10	5.2
4320/60/P	4:2:0	10	7680	4320	60/1.001	Progressive	16 : 9		Main 10	6.1
	4:2:0	10	7680	4320	60	Progressive	16 : 9		Main 10	6.1
4320/120/P	4:2:0	10	7680	4320	120/1.001	Progressive	16 : 9		Main 10	6.2
	4:2:0	10	7680	4320	120	Progressive	16 : 9		Main 10	6.2

5.1.3.3 Syntax

(1) NAL unit header

Syntax element	Value	Remark
nuh_layer_id	0	0 fixed in HEVC Standard

(2) Profile, Tier, Level

Syntax element	Value	Remarks
general_profile_space	0	0 fixed in HEVC Standard
general_tier_flag	0	Main tier
general_profile_idc	1, 2	1: Main Profile 2: Main10 Profile (Note) See Table 5-9 about the combination of video format
general_profile_compatibility_flag[j] (j = [0, 31])	0, 1	Description on compatible Profile In Main Profile, only when j=1 or 2 this value is 1 and otherwise it is 0. In Main 10 Profile, only when j=2 this value is 1 and otherwise it is 0.
general_progressive_source_flag	0, 1	0: in case of 1080/I 1: other than 1080/I
general_interlaced_source_flag	0, 1	0: other than 1080/I 1: in case of 1080/I
general_frame_only_constraint_flag	0, 1	0: in case of 1080/I 1: other than 1080/I
general_level_idc	123, 153, 156, 183, 186	123: Level 4.1 153: Level 5.1 156: Level 5.2 183: Level 6.1 186: Level 6.2 (Note) See Table 5-9 about the combination of video format

(3) Video parameter set

Syntax element	Value	Remarks
vps_max_layers_minus1	0	0 fixed in HEVC Standard
vps_timing_info_present_flag	0	Describe timing information by VUI, and omit description by VPS.

(4) Sequence parameter set

Syntax element	Value	Remarks
chroma_format_idc	1	1: "4:2:0"
pic_width_in_luma_samples pic_height_in_luma_samples	See (7)	MinCbSizeY(=8) times
conformance_window_flag conf_win_left_offset conf_win_right_offset conf_win_top_offset conf_win_bottom_offset	See (7)	
bit_depth_luma_minus8 bit_depth_chroma_minus8	0, 2	0: Main Profile (8-bit) 2: Main10 Profile (10-bit)
log2_min_luma_coding_block_size_minus3	0	MinCbSizeY = 8
log2_diff_max_min_luma_coding_block_size	2, 3	2: CtbSizeY = 32 3: CtbSizeY = 64
log2_min_transform_block_size_minus2	0	Log2MinTrafoSize = 2 (4x4)
log2_diff_max_min_transform_block_size	3	Log2MaxTrafoSize = 5 (32x32)
vui_parameters_present_flag	1	This surely describes VUI
vui_parameters( )	See (5)	

(5) VUI parameters

Syntax element	Value	Remarks
aspect_ratio_info_present_flag	1	This describes aspect_ratio_idc.
aspect_ratio_idc	1	1:1 ("square pixel")
video_signal_type_present_flag	1	This describes video information.
video_format	0	Component
video_full_range_flag	0	Based on the provision of luminance and color difference signals.
colour_description_present_flag	1	This describes color description information.
colour_primaries	1, 9	1: Rec. ITU-R BT.709, IEC 61966-2-4 (in case of HDTV(SDR-TV)) 9: Rec. ITU-R BT.2020 (in case of HDTV, UHDTV or HDR-TV)
transfer_characteristics	1, 11, 14, 16, 18	1: Rec. ITU-R BT.709 (in case of conventional color gamut of HDTV (SDR-TV)) 11: IEC 61966-2-4 (in case of wide color gamut of HDTV (SDR-TV)) 14: Rec. ITU-R BT.2020, 10-bit (in case of UHDTV (SDR-TV)) 16: Rec. ITU-R BT.2100 PQ (in case of HDR-TV) 18: Rec. ITU-R BT.2100 HLG (in case of HDR-TV)
matrix_coefficients	1, 9	1: Rec. ITU-R BT.709, IEC 61966-2-4 (in case of HDTV (SDR-TV)) 9: Rec. ITU-R BT.2020 Non-constant luminance (in case of HDTV, UHDTV or HDR-TV)
chroma_loc_info_present_flag	1	This describes 4:2:0 color difference signals

Syntax element	Value	Remarks
		position.
chroma_sample_loc_type_top_field chroma_sample_loc_type_bottom_field	0, 2	0: The middle of 2 luminance lines in vertical direction (in case of 1080/I). 2: This agrees with luminance line in vertical direction (other than 1080/I).
neutral_chroma_indication_flag	0	The value of color difference signal is the value coded in bitstream.
field_seq_flag	See (7)	
frame_field_info_present_flag	1	This describes pic_struct information.
vui_timing_info_present_flag	1	This describes timing information.
vui_num_units_in_tick vui_time_scale	See (8)	Either “case 1” or “case 2” is applied.
vui_poc_proportional_to_timing_flag	1	POC value is proportional to display time.
vui_num_ticks_poc_diff_one_minus1	See (9)	
vui_hrd_parameters_present_flag	1	HRD information is described on VUI.

(6) HRD parameters

Syntax element	Value	Remarks
nal_hrd_parameters_present_flag	1	This describes NAL HRD information.
sub_pic_hrd_params_present_flag	0	only HRD description of the unit of AU
fixed_pic_rate_general_flag[i] (i = [0, sps_max_sub_layers_minus1])	0, 1	0: only in case of different picture rate between CVS.
fixed_pic_rate_within_cvs_flag[i] (i = [0, sps_max_sub_layers_minus1])	0, 1	highest layer. (i = sps_max_sub_layers_minus1) shall always be 1 (picture rate is fixed in CVS).
elemental_duration_in_tc_minus1[i] (i = [0, sps_max_sub_layers_minus1])	See (9)	

(7) Parameters representing picture size

Video coding format	field_seq_flag	aspect_ratio_idc	general_progressive_source_flag	general_interlace_source_flag	pic_width_in_luma_samples	pic_height_in_luma_samples	conformance_window_flag	conf_win_left_offset	conf_win_right_offset	conf_win_top_offset	conf_win_bottom_offset
1080/I	0	1	0	1	1,920	1,088	1	0	0	0	4
	1	1	0	1	1,920	544	1	0	0	0	2
1080/P	0	1	1	0	1,920	1,080	0	0	0	0	0
	0	1	1	0	1,920	1,088	1	0	0	0	4
2160/P	0	1	1	0	3,840	2,160	0	0	0	0	0
4320/P	0	1	1	0	7,320	4,320	0	0	0	0	0

Note: Both the case pic_height_in_luma_samples is 1,080 (without cropping) and the case it is 1,088 (with cropping) are permitted in 1080/P.

(8) Time scale

Frame/field frequency and scanning	vui_time_scale		vui_num_units_in_tick	
	Case 1	Case 2	Case 1	Case 2
59.94/I	60,000	27,000,000	1,001	450,450
59.94/P	60,000		1,001	450,450
60.00/I	60,000		1,000	450,000
60.00/P	60,000		1,000	450,000
119.88/P	120,000		1,001	225,225
120.00/P	120,000		1,000	225,000

(9) Picture interval

Syntax element	Value	Remarks
vui_num_ticks_poc_diff_one_minus1	vui_num_units_in_tick – 1	
elemental_duration_in_tc_minus1[sps_max_sub_layers_minus1]	0	

5.1.4 Desirable encoding areas

With reference to video encoding, the areas shown in Table 5-10 shall be the desirable encoding areas for the respective video input formats. It is desirable that after decoding, active video lines of signals generated by the receiver shall match the lines shown in Table 5-10.

Table 5-10: Desirable encoding areas

Video input format	Number of active lines	Number of lines to be encoded	Desirable encoding area
4320/120/P	4320	4320	All active lines
4320/60/P	4320	4320	All active lines
2160/120/P	2160	2160	All active lines
2160/60/P	2160	2160	All active lines
1080/60/P	1080	1080	All active lines (line numbers: 42～1121)
1080/60/I	1080	1080	All active lines (line numbers: 21～560 and line numbers: 584～1123)
720/60/P	720	720	All active lines (line numbers: 26～745)
480/60/P	483	480	line numbers: 45～524
480/60/I	483	480	line numbers: 23～262 and line numbers: 286～525

5.2 Restrictions on video coding parameters for low definition video services[4]

5.2.1 System based on MPEG-2 Video Standard

Video coding shall conform to the Main Profile defined in the MPEG-2 Video Standard, and to code with the condition shown in Table 5-11.

Table 5-11: Restrictions on coding parameters

<table>
<tr><th colspan="4">Sequence Header restriction</th><th>Sequence extension restriction</th><th colspan="3">Sequence display extension restriction(Note4)</th><th rowspan="2">Other parameter (Note5)</th></tr>
<tr><th>vertical_ size_value</th><th>horizontal_ size_ value</th><th>aspect_ ratio_ information</th><th>frame_ rate_ code</th><th>progressive_ sequence</th><th>color_ primaries</th><th>transfer_ characteristics</th><th>matrix_ coefficients</th></tr>
<tr><td>480</td><td>352</td><td rowspan="3">2, 3</td><td rowspan="4">4 (Note2,3)</td><td>0</td><td rowspan="4">1</td><td rowspan="4">1</td><td rowspan="4">1</td><td>nominal value for MP@ML</td></tr>
<tr><td>240</td><td>352</td><td rowspan="2">1</td><td rowspan="2">nominal value for MP@LL</td></tr>
<tr><td>120 (Note1)</td><td>176</td></tr>
<tr><td>less than 480</td><td>less than 720</td><td>1</td><td>0, 1</td><td>nominal value for MP@ML</td></tr>
</table>

(Note 1) In practice, 128 lines are coded in MPEG-2 Video Standard.
(Note 2) In case that transmittable bitrate is really low, the encoding method that frame rate to be coded is made to be small by using skipped macroblock etc. is also effective.
(Note 3) In case of film material, the encoding method to control flag of repeat_first_field, top_field_first, progressive_frame without changing frame_rate_code shall be possible. (See Attachment 1, Chapter 5)
(Note 4) In case that sequence_display_extension is not transmitted, color_primaries, transfer_characteristics, and matrix_coefficients are processed in the receiver as each value is equal to 1.
(Note 5) For each level of Main profile, provided values in MPEG-2 Video standard are adopted. Here, bit_rate_value takes maximum capacity which is transmittable for MP@HL and MP@H14, and for MP@ML it takes below 15Mbit/s. It shall be operated with variable bit rate, and vbv_delay shall be always 0xFFFF.

Table 5-12: Meanings of each code number of coding parameter in MPEG-2 Video Standard in Table 5-11

aspect_ratio_information	1 = square pixel 2 = 4:3 display 3 = 16:9 display
frame_rate_code	4 = 30/1.001 Hz
progressive_sequence	0 = interlaced scanning 1 = progressive scanning
color_primaries	1 = nominal value in Rec. ITU-R BT.709(BT.1361)
transfer_characteristics	1 = nominal value in Rec. ITU-R BT.709(BT.1361)
matrix_coefficients	1 = nominal value in Rec. ITU-R BT.709(BT.1361)

[4] In this paragraph, "low definition video service" represents the video service in which less pixels than those of video coding format (horizontal and vertical coding pixels) provided in "5.1 Restrictions on video coding parameters for television services" are used for video coding format, and the video service in multimedia broadcasting.

5.2.2 System based on MPEG-4 AVC Standard

Video coding system shall be based on Baseline or Main Profile provided in MPEG-4 AVC Standard, and Level is one of 1, 1.1, 1.2, 1.3, 2, 2.1, 2.2, 3.

The restrictions on coding parameters are shown in Table 5-13. As for parameters which are not described here such as buffer size, they shall conform to the provision of MPEG-4 AVC Standard.

Table 5-13: Restrictions on coding parameters

Item		Restrictions
Video format		YC_BC_R 4:2:0
Input bit number		8 bit
Scanning system		progressive
Maximum frame rate		30 Hz
Maximum picture size		nominal value in MPEG-4 AVC Standard (Table 5-14)
Maximum bitrate		nominal value in MPEG-4 AVC Standard (Table 5-14)
Picture interval		Within 0.7 second in case that video ES is multiplexed to PES.
Color description	colour_primaries	based on Rec. ITU-R BT.1361(Rec. ITU-R BT.709)
	transfer_characteristics	based on Rec. ITU-R BT.1361 conventional color gamut system (Rec. ITU-R BT.709) or wide color gamut system (IEC61966-2-4)
	matrix_coefficients	based on Rec. ITU-R BT.1361(Rec. ITU-R BT.709)

Table 5-14: Picture size and maximum bitrate

Profile	Level	Maximum picture size (number of macro block) (nominal value in MPEG-4 AVC Standard)	horizontal pixels × vertical lines (luminance signal) and picture aspect ratio (horizontal : vertical)	Maximum bitrate (nominal value in MPEG-4 AVC Standard)
Baseline or Main	1	99	160×90 (16:9) 160×120 (4:3) 176×120 (4:3, 16:9) 176×144 (4:3)	64kbit/s
	1.1	396	320×180 (16:9) 320×240 (4:3) 352×240 (4:3, 16:9) 352×288 (4:3)	192kbit/s
	1.2	396		384kbit/s
	1.3	396		768kbit/s
	2	396		2Mbit/s
	2.1	792	352×480 (4:3, 16:9)	4Mbit/s
	2.2	1620	640×480 (4:3) 720×480 (4:3, 16:9)	4Mbit/s
	3	1620		10Mbit/s

5.2.3 System based on HEVC Standard

Video coding system shall be based on Main profile or Main 10 profile and Main Tier provided in HEVC Standard. Main 10 profile is used when bit number of coding pixel is 10 bits, and Main profile is used when bit number of coding pixel is 8 bits. Level is one of 2, 2.1, 3, 3.1, 4, 4.1 shall be used.

Restrictions on coding parameters are shown in Table 5-15. The parameters which are not described here such as buffer size shall conform to HEVC Standard.

Table 5-15: Restrictions on coding parameters

Items		Restrictions
Video format		$Y'C'_BC'_R$ 4:2:0
Input bit number		8 bit or 10 bit
scanning		progressive or interlaced
Display size		See Table 5-16
Frame rate		See Table 5-16
Maximum bit rate		Upper limit value of bit rate in NAL level provided in HEVC Standard (See Table 5-16)
Color description	colour_primaries	based on Rec. ITU-R BT.709
	transfer_characteristics	based on Rec. ITU-R BT.709 or IEC61966-2-4
	matrix_coefficients	based on Rec. ITU-R BT.709

Table 5-16: Picture size, frame rate and maximum bitrate

Profile	Level	horizontal pixels × vertical lines (luminance signal) and picture aspect ratio (horizontal : vertical)	scanning system	frame frequency (progressive) or field frequency (interlaced) [Hz]	maximum bit rate (nominal value in HEVC Standard) [Mbps]
Main or Main10	2	320×180 (16:9)	progressive	59.94, 60.00	1.65
	2.1	320×240 (4:3), 352×240 (4:3, 16:9)	progressive	59.94, 60.00	3.30
	3	640×480 (4:3), 720×480 (4:3, 16:9)	interlaced	59.94, 60.00	6.60
	3.1	640×480 (4:3), 720×480 (4:3, 16:9)	progressive	59.94, 60.00	11.0
	4	1280×720 (16:9)	progressive	59.94, 60.00	13.2
	4	1440×1080 (16:9)	interlaced	59.94, 60.00	13.2
	4.1	1440×1080 (16:9)	progressive	59.94, 60.00	22.0

The restrictions on syntax are as the followings. The restriction on syntax which is not described particularly is the same as the restriction described in 5.1.3.3.

(1) Profile, Tier, Level

Syntax element	Value	Remarks
general_progressive_source_flag	0, 1	0: interlaced 1: progressive
general_interlaced_source_flag	0, 1	0: progressive 1: interlaced
general_frame_only_constraint_flag	0, 1	0: interlaced 1: progressive
general_level_idc	60, 63, 90. 93. 120, 123	60: Level 2 63: Level 2.1 90: Level 3 93: Level 3.1 120: Level 4 123: Level 4.1 (Note) See Table 5-16 about the combination with video format

(2) Sequence parameter set

Syntax element	Value	Remarks
pic_width_in_luma_samples pic_height_in_luma_samples	See (4)	MinCbSizeY(=8) times
conformance_window_flag conf_win_left_offset conf_win_right_offset conf_win_top_offset conf_win_bottom_offset	See (4)	

(3) VUI Parameters

Syntax element	Value	Remarks
aspect_ratio_idc	See (6)	Sample aspect ratio
video_format	0	Component
colour_primaries	1	1: Rec. ITU-R BT.709, IEC 61966-2-4
transfer_characteristics	1, 11	1: Rec. ITU-R BT.709 (in case of HDTV conventional color gamut) 11: IEC 61966-2-4 (in case of HDTV wide color gamut)
matrix_coefficients	1	1: Rec. ITU-R BT.709, IEC 61966-2-4
chroma_loc_info_present_flag	1	4:2:0 color difference signals position is described
chroma_sample_loc_type_top_field chroma_sample_loc_type_bottom_field	0, 2	0: The middle of 2 luminance lines in vertical direction (in case of interlace). 2: This agrees with luminance line in vertical direction (in case of progressive)
field_seq_flag	See (4)	
vui_num_units_in_tick vui_time_scale	See (5)	Either “case 1” or “case 2” is applied.

(4) Parameters representing picture size

Input video format (“i ”represents interlace	field_seq_flag	general_progressive_source_flag	general_interlace_source_flag	pic_width_in_luma_samples	pic_height_in_luma_samples	conformance_window_flag	conf_win_left_offset	conf_win_right_offset	conf_win_top_offset	conf_win_bottom_offset
320×180	0	1	0	320	184	1	0	0	0	2
320×240	0	1	0	320	240	0	0	0	0	0
352×240	0	1	0	352	240	0	0	0	0	0
640×480i	1	0	1	640	240	0	0	0	0	0
640×480i (Note 1)	0	0	1	640	480	0	0	0	0	0
640×480	0	1	0	640	480	0	0	0	0	0
720×480i	1	0	1	720	240	0	0	0	0	0
720×480i (Note 1)	0	0	1	720	480	0	0	0	0	0
720×480	0	1	0	720	240	0	0	0	0	0
1280×720	0	1	0	1280	720	0	0	0	0	0
1440×1080i	1	0	1	1440	544	1	0	0	0	2
1440×1080i (Note1)	0	0	1	1440	1080	0	0	0	0	0
1440×1080	0	1	0	1440	1080	0	0	0	0	0

(Note 1) In case of coding as a unit of frame

(5) Time scale

Frame/Field frequency and scanning	vui_time_scale		vui_num_units_in_tick	
	Case 1	Case 2	Case 1	Case 2
59.94/I, 59.94/P	60,000	27,000,000	1,001	450,450
60.00/I, 60.00/P	60,000	27,000,000	1,000	450,000

(6) Sample aspect ratio

Picture size	Aspect ratio of display	aspect_ratio_idc	Aspect ratio of sample
320×180	16:9	1	1:1
320×240	4:3	1	1:1
352×240	4:3	3	10:11
352×240	16:9	5	40:33
640×480	4:3	1	1:1
720×480	4:3	3	10:11
720×480	16:9	5	40:33
1280×720	16:9	1	1:1
1440×1080	16:9	14	4:3

Annex A Technical system which is applied to digital broadcasting

Technical system which is applied to each standard system of digital broadcasting provided in Ordinance (Ordinance of MIC No.87, 2011 or Ordinance of MIC No.94, 2011) is shown in Table A-1.

Table A-1: Technical system which is applied to standard system (○: Applied)

		Digital broadcasting		Digital terrestrial broad casting	Multi-media broad casting	BS digital broad casting	Advanced BS digital broad casting (Note 2)	Narrow band CS digital broad casting (Note 3)	Advanced narrow band CS digital broad casting (Note 4)	Wide band CS digital broad casting	Advanced wide band CS digital broad casting (Note 2)
Video coding format	Video signal characteristics (See: Chapter 2, 2.1)	based on HDTV	conventional color gamut	○	○	○	○	○	○	○	○
			wide color gamut(xvYCC)		○		○		○		○
		based on SDTV						○			
		based on UHDTV					○		○		○
	Video format (See: Chapter 2, 2.4)	480/60/I		○	○ (Note 1)	○		○	○	○	
		480/60/P		○		○		○	○	○	
		720/60/P		○		○		○	○	○	
		1080/60/I		○		○	○	○	○	○	○
		1080/60/P					○		○		○
		2160/60/P					○		○		○
		2160/120/P					○		○		○
		4320/60/P					○				○
		4320/120/P					○				○
Video coding system (See: Chapter 3)		MPEG-2 Video Standard		○		○		○	○	○	
		MPEG-4 AVC Standard			○				○		
		HEVC Standard					○		○		○

(Note 1) Various kinds of video format with 160×90～720×480 (maximum frame frequency is 30Hz) are provided in Ordinance.

(Note 2) Video signal characteristics which is applied to video format in Advanced BS digital broadcasting and Advanced wide band CS digital broadcasting conforms to Table A-2.

(Note 3) Video signal characteristics which is applied to video format in Narrow band CS digital broadcasting conforms to Table A-3.

(Note 4) Video signal characteristics which is applied to video format in Advanced narrow band CS digital broadcasting conforms to Table A-4.

Table A-2: Video signal characteristics which is applied to video format for SDR-TV in Advanced BS digital broadcasting and Advanced wide band CS digital broadcasting (○: Applied)

		1080/60/I	1080/60/P	2160/60/P	2160/120/P	4320/60/P	4320/120/P
based on HDTV	conventional color gamut	○	○				
	wide color gamut(xvYCC)	○	○				
based on UHDTV				○	○	○	○

Table A-3: Video signal characteristics which is applied to video format in Narrow band CS digital broadcasting (○: Applied)

		480/60/I	480/60/P	720/60/P	1080/60/I
based on HDTV	conventional color gamut			○	○
based on SDTV		○	○		

Table A-4: Video signal characteristics and video coding system which are applied to video format for SDR-TV in Advanced narrow band CS digital broadcasting (○: Applied)

		480/60/I	480/60/P	720/60/P	1080/60/I	1080/60/P	2160/60/P	2160/120/P
based on HDTV	conventional color gamut	○	○	○	○	○		
	wide color gamut(xvYCC)				○ (Note 5)	○ (Note 5)		
based on UHDTV							○	○
MPEG-2 Video Standard		○	○	○	○	○		
MPEG-4 AVC Standard		○	○	○	○	○		
HEVC Standard					○	○	○	○

(Note 5) Only in case of using HEVC Standard

Attachment: Operational Guidelines

Attachment 1: Operational Guidelines for MPEG-2 Video Standard on television services

Chapter 1: General Terms

1.1 Objective

The purpose of these operational guidelines is to present technical recommendations of MPEG-2 Video Standard in the practical operation concerning to video signals and video coding systems in digital television services.

1.2 Scope

These operational guidelines apply to video signals using MPEG-2 Video Standard among the video signals in television services that comply with the “Standard transmission system for digital broadcasting among standard television broadcasting and the like” (Ordinance).

1.3 References

1.3.1 Normative References

(1) Rec. ITU-T H.262|ISO/IEC 13818-2:2000 Information technology--Generic coding of moving pictures and associated audio information: Video (hereinafter referred to as “MPEG-2 Video Standard”)

(2) Rec. ITU-T H.222.0|ISO/IEC 13818-1:2000: Information technology--Generic coding of moving pictures and associated audio information: Systems (hereinafter referred to as “MPEG-2 Systems Standard”)

1.4 Terminology

1.4.1 Abbreviations

CA	Conditional Access
CAT	Conditional Access Table
DTS	Decoding Time-Stamp
ECM	Entitlement Control Message
EMM	Entitlement Management Message
ES	Elementary Stream
GOP	Group of Pictures
HDTV	High Definition Television (Note 1)
NIT	Network Information Table
PAT	Program Association Table
PES	Packetized Elementary Stream
PID	Packet Identifier
PMT	Program Map Table
PSI	Program Specific Information
PTS	Presentation Time-Stamp

SDTV	Standard Definition Television (Note 2)
TMCC	Transmission & Multiplexing Configuration Control
TS	Transport Stream

Note 1: In this standard, the term represents "High Definition Television Broadcasting" provided in Ordinance.

Note 2: In this standard, the term represents "Standard Television Broadcasting" provided in Ordinance.

Chapter 2: Transmitting Sequence Header and Sequence End Code

2.1 Transmitting sequence header (sequence_header)

sequence_header and sequence_extension (and sequence_display_extension if necessary) shall be transmitted immediately before the GOP header. But if the GOP header is not transmitted, they shall be transmitted immediately before I-picture data at the beginning of GOP.

2.2 Transmitting sequence end code (sequence_end_code)

The timing which sequence_end_code is transmitted shall be immediately after a single frame of video data has been transmitted.

Note: When a sequence_end_code is received at the receiver side, it is recommended that the freeze-frame of the video data received immediately before the sequence_end_code be displayed until the following video data is correctly decoded and displayed. This permits continuous display of video data if the video data transmitted following the sequence_end_code is decoded and displayed without delay, and it does not necessarily mean that the freeze-frame of video data is displayed for affixed period of time.

<Blank Page>

Chapter 3: Channel-hopping time

The following operation is recommended to keep channel-hopping time below a given duration:

sequence_header shall be encoded at least every 500 ms, and the picture shall be updated in intra mode.

Note: The sequence_header that contains video format and other information, transmission frequency of intra mode picture, and delay at the buffer are among the video coding parameters related to channel-hopping time.

Fig. 3-1(a) shows a flowchart of various stages related to channel-hopping in BS digital broadcasting. Fig. 3-1(b) shows figures for terrestrial digital broadcasting. The channel-hopping time in terrestrial digital broadcasting is the same as for BS digital broadcasting, except for the front-end part. The channel-hopping time at the front-end part in terrestrial digital broadcasting is shown in Fig. 3-1(b).

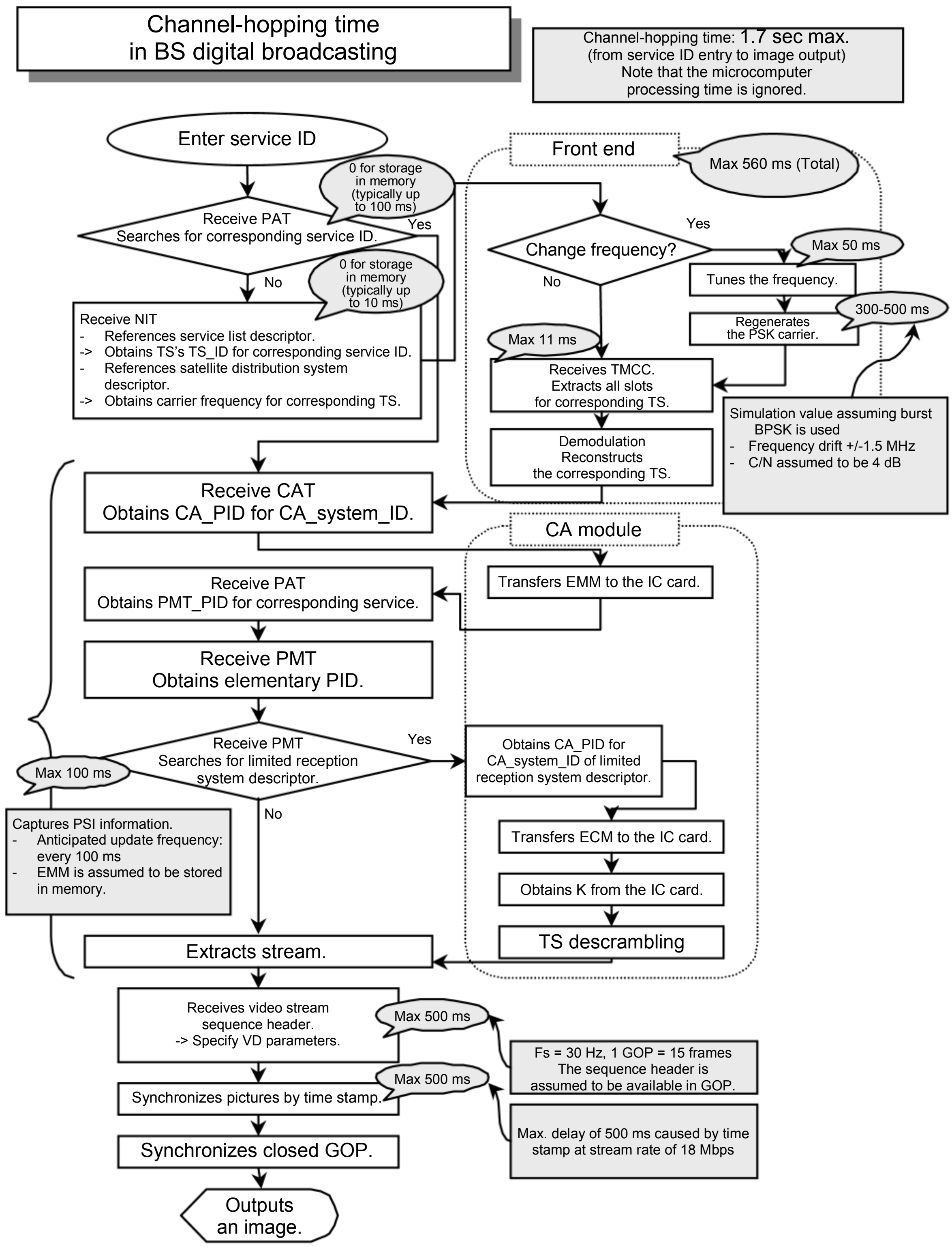
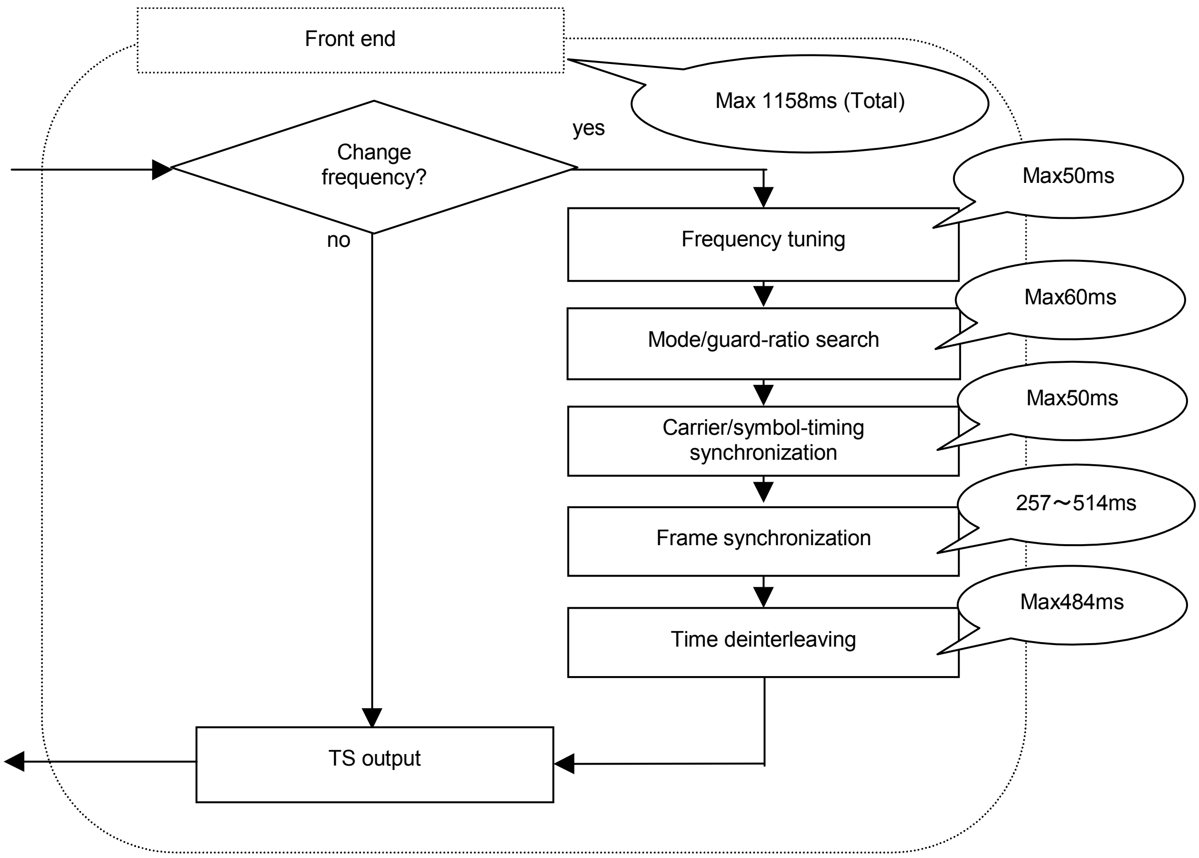


Fig. 3-1(a): Channel-hopping time in BS digital broadcasting



Note: The above channel-hopping time is in the case of mode3, guard-ratio 1/4, and time-interleaving I=4. (This combination is the case in which total delay time in the front-end becomes maximum.)

- Frequency tuning: the same as BS digital broadcasting
- Mode/guard-ratio search: only for combinations used in terrestrial digital broadcasting
- Carrier/symbol-timing synchronization: tens of symbols
- Frame synchronization: 1 to 2 frames (TMCC acquisition time)

Fig. 3-1(b) Channel-hopping time in terrestrial digital broadcasting (front end)
(Switching time except for front-end part is the same as for BS digital broadcasting)

<Blank Page>

Chapter 4: Seamless Switching

In order to realize seamless display by the receiver when switching between video formats, the following procedure is recommended for the transmitting side and the receiving side:

4.1 Changing the number of active samples

(1) Procedure on the transmitting side
- The sequence is terminated at the operation switching point by sequence_end_code. A new number of samples is specified by the next sequence_header.
- The head GOP of the new operation sequence sets the closed_gop flag at the GOP header.
- vbv_buffer_size remains unchanged before and after switching.
- The continuity of PTS and DTS is assured.

(2) Receiver operation
- The operating mode is specified by the parameter of number of pixels included in the received sequence_header. The new operation mode is specified according to information included in the received sequence_header even if sequence_end_code is not received.

4.2 Changing picture aspect ratio for 480/60/I system

(1) Procedure on the transmitting sid
- The sequence is terminated at the operation switching point by sequence_end_code. A new aspect ratio is specified by the next sequence_header.
- The head GOP of the new operation sequence sets the closed_gop flag in the GOP header.
- vbv_buffer_size remains unchanged after switching.
- The continuity of PTS and DTS is assured.

(2) Receiver operation
- The operating mode is specified by the parameter of aspect ratio included in the received sequence_header. The new operating mode is specified according to information included in the received sequence_header even if sequence_end_code is not received.

4.3 Changing bitrate

(1) Procedure on the transmitting sid e
- The variable bit rate mode is always used. (vbv_delay: 0xFFFF)
- sequence_end_code is not inserted at change point of the transfer bit rate.
- vbv_buffer_size remains unchanged after this change.
- The continuity of PTS and DTS is assured.

(2) Receiver operation
- The receiver shall be operated seamlessly by controlling the start of video and audio decoding and output according to PTS and DTS described in the PES Header.

Note: The bitrate for transmission is changed at the transmitting side based on the operational method mentioned above. In this case, it shall be controlled so that the decoder's buffer shall not fail. Among the total delay arising between coding and decoding, the interval during which data passes through the buffer is expressed as the "buffer capacity/bitrate." That is, when vbv_buffer_size remains constant, the passing time changes with a change of bit rate. As a result, when the passing time increases, the decoder's buffer transfers to the underflow state (in which case it takes more time for data to be received). And conversely, when the passing time decreases, the buffer transfers to the overflow state. The buffer will fail if this transition in buffer state exceeds the capacity.

4.4 Video format switching method

This section describes the operational method for transmitting and receiving sides to ensure seamless or near-normal display of pictures when switching between video stream formats (1080/I, 720/P, 480/P, 480/I, etc.) for a specific service ID.

To allow perfectly seamless switching, both transmitting and receiving sides shall be capable of seamless switching. However, it is possible to assume that either the transmitting or receiving side or neither is capable of seamless switching when broadcasting services begin. Even in this case, the procedure given in this section is recommended for video format switching, given the fact that both the transmitting and receiving sides can be changed independently to a perfectly seamless switching-capable system by displaying a freeze-frame or black frame screen – an approach that is less visually disruptive.

Sections 4.4.1 and 4.4.2 respectively describe the procedure for the transmitting side that permits perfectly seamless switching and the simple procedure. However, there are also other methods positioned between both of the above. It is possible to make the transmitting side gradually capable of perfectly seamless switching in parallel with upgrading the system on the transmitting side. As an example, this section discusses the switching of three SDTV programs to one HDTV program. However, switching from HDTV to SDTV or switching between different formats (480/I ↔ 480/P, 1080/I ↔ 720/P, etc.) can be handled the same way on both transmitting and receiving sides. When switching from any video format to another for a specific service ID, the video stream ES PID for the original format shall be changed after switching to another format.

When switching from three SDTV programs to one HDTV program or vice versa, broadcasting stations intending to provide seamless display shall transmit the same number of PMTs that specify the same service_id as SDTV during HDTV broadcasting, and shall specify as HDTV's ES_PID a unique value to distinguish it any PID of components broadcast when transmission of the new PMT starts. Moreover, both SDTV and HDTV PMTs shall contain the video decode control descriptor given in the ARB STD-B10. In this section, we temporarily specify the following values as service_id and ES_PID, by assuming that the above requirements are met:

SDTV 1 program: service_id = 01, ES_PID = 101 → HDTV program: service_id = 01, ES_PID = 104
SDTV 2 program: service_id = 02, ES_PID = 102 → HDTV program: service_id = 02, ES_PID = 104
SDTV 3 program: service_id = 03, ES_PID = 103 → HDTV program: service_id = 03, ES_PID = 104

4.4.1 Procedure for perfect seamless switching (method with sequence_end_code is transmitted)

(1) Procedure on the transmitting side

1. Assume that switching between SDTV and HDTV occurs at time T1. The SDTV's PMT shall contain video_decode_control_descriptor (sequence_end_code_flag: 1, video_encode_format: 0100 (480i), 0011 (480p)).
2. Three SDTV encoders and one HDTV encoder synchronize PCR and PTS (or DTS) to ensure seamless PCR at the time of switching.
3. Transmission of the HDTV program's PMT (ES_PID = 104) starts one second (standard time) before switching time T1. HDTV's PMT shall contain video_decode_control_descriptor (sequence_end_code_flag: 1, and video_encode_format: 0001 (1080i), 0010 (720p)). (Note 1)
4. Transmission of the SDTV stream terminates immediately before the switching time as the end of GOP, and sequence_end_code is added at the end. (Note 2)
5. At switching time, the muliplexer halts TS multiplexing for SDTV and starts TS multiplexing for HDTV. The HDTV sequence_header shall be transmitted as soon as possible after the switching to the HDTV stream is complete. The HDTV sequence_header shall begin with GOP. The first GOP shall be treated as a "closed GOP". Null data is multiplexed between the SDTV stream's sequence_end_code and HDTV stream's sequence_header_code. (Note 2)

Note 1:

Timing at which new PMT is to be transmitted

- For broadcasts of free programs only, the receiver can handle program switching as long as a new PMT is transmitted at least 0.5 second before switching time T1. Because the transmitting side is typically operated in units of exactly seconds, transmitting a new PMT one second before T1 shall be the standard. There are no problems with the receiver as long as transmission of a new PMT starts 0.5 to 2.0 seconds before the switching time.
- For broadcasts of pay per view programs, if there are a number of keys subject to program switching, transmission of a new ECM two seconds before switching time may in certain cases be too late, given the IC card response time. However, if a new PMT is transmitted more than two seconds before switching time, an individual selecting the station at that timing will be unable to see any picture at all for a lengthy duration. Therefore, a new PMT shall be transmitted sometime between 0.5 and 2.0 seconds before the switching time. CAS operation shall be ensured (for example) by unifying keys or using temporal non-scrambling so that no inconvenience arises, even when station selection is made at this timing.

Note 2:

Schedule control is performed in units of seconds at the broadcast stations. This control timing does not generally coincide with GOP end timing, due to GOP length or the frame/field frequency of 59.94 Hz. Therefore, stream end and start timings come slightly before or after the control timing. The gap between the end of SDTV stream and the start of HDTV stream shall be sufficiently narrow to prevent underflow at the decoder on the receiving side.

(2) Procedure on the receiver side

(a) A seamless switching-capable receiver

1. The receiver obtains the new version of PMT.
2. The Demux is set up so that it feeds the ES_PID stream data of both SDTV and HDTV to the AV decoder when the receiver (based on the contents of the PMT descriptor) finds that switching from SDTV to HDTV will occur and that sequence_end_code will be transmitted in a stream. However, note that SDTV and HDTV real data is not fed to the decoder at the same time, regardless of transmission timing. Instead, SDTV stream data is first stored in the buffer. HDTV stream data is stored in the buffer only when the storage of SDTV stream data is complete.
3. The video decoder displays a freeze-frame picture and mutes the audio upon it obtaining sequence_end_code.
4. The decoder performs the appropriate decoding through automatic tracking upon obtaining sequence_header of HDTV stream. When ready to output normal video and audio data, the decoder cancels video freeze-frame and audio muting. (To display pictures in an apparently seamless manner, the HDTV stream shall be received soon after the SDTV stream so that the buffer does not underflow. In this case, no freeze-frame picture is displayed. If the period between the end and start of the SDTV stream is not sufficiently short, and if the buffer underflows as a result, a freeze-frame picture is transmitted immediately before sequence_end_code is displayed.)
5. When the receiver finds that HDTV decoding has begun, the Demux only feeds HDTV's ES_PID to the AV decoder.

(b) A seamless switching-incapable receiver

1. The receiver obtains the new version of PMT.
2. Freeze-frame or black frame is displayed and audio muted if, based on the contents of the PMT descriptor (regardless of whether sequence_end_code is present), the receiver finds that switching from SDTV to HDTV will occur.
3. The video decoder halts SDTV decoding.
4. The Demux is set up to stop receiving streams with SDTV's ES_PID and feeds streams with HDTV's ES_PID to the decoder buffer.
5. Using the host CPU to monitor the sequence_header monitor register of the video decoder, the receiver awaits HDTV stream input.
6. When the decoder obtains HDTV stream sequence_header, it begins HDTV decoding. When ready to output normal video and audio data, the decoder cancels video freeze-frame and audio muting.

○ Precaution:
If a receiver is available that is not seamless switching-capable, but can display a freeze-frame picture upon reception of new PMT, it is preferable that a virtually flicker-free picture be transmitted even when the freeze-frame picture is displayed 0.5 second (delay at the buffer) or more before start of new PMT transmission.

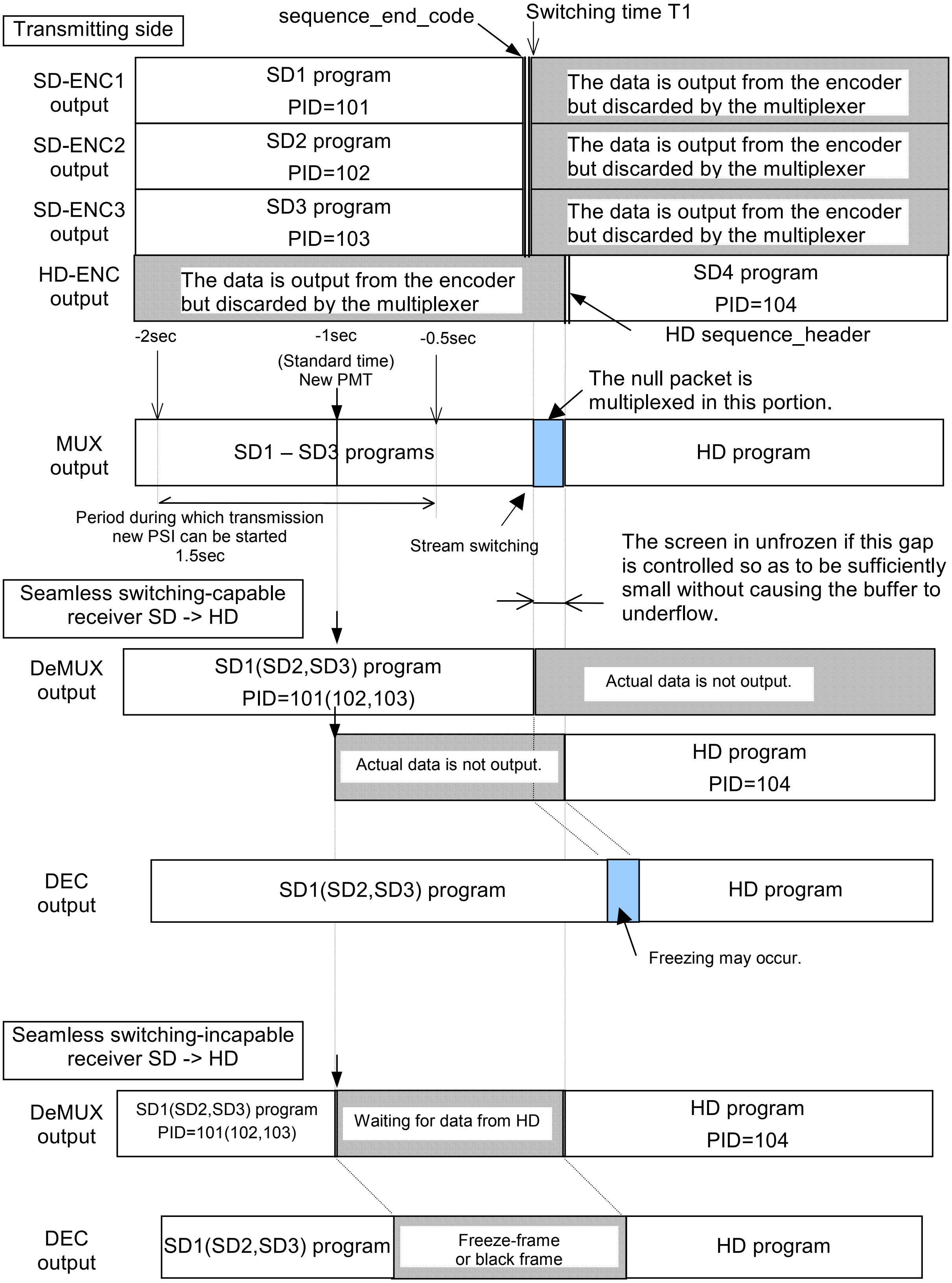


Fig. 4-1: Conceptual diagram of timing of transmitting and receiving sides that enables perfect seamless switching between SDTV and HDTV
(when sequence_end_code_flag of video_decode_control_descriptor is 1)

4.4.2 Simple procedure for switching between SDTV and HDTV (method by which sequence_end_code is not transmitted)

This section assumes that three SDTV encoders and one HDTV encoder are operating asynchronously, and that PCR is not seamless. Thus, the objective is to achieve synchronous encoder operations and seamless PCR.

(1) Procedure on the transmitting side

1. Assume that switching between SDTV and HDTV occurs one second before the start of an actual HDTV program and denote this moment as T1. SDTV's PMT shall contain video_decode_control_descriptor (sequence_end_code_flag: 0, and video_encode_format: 0100 (480/I), 0011 (480/P)).
2. The SDTV stream encoders begin transmitting still-frame pictures—pictures that may be displayed as black- or freeze-frames—0.5 second or more before the scheduled start of the HDTV program's PMT transmission relative to switching time T1. These encoders transmit mute as audio data.
3. The HDTV stream encoder starts transmitting still-picture and mute, respectively, as video and audio data one second or more before switching time T1.
4. Transmission of the HDTV program's PMT (ES_PID = 104) starts one to 0.2 second before switching time T1. HDTV's PMT shall contain video_decode_control_descriptor (sequence_end_code_flag: 0, and video_encode_format: 0001 (1080/I), 0010 (720/P)). (Note 1)
5. At switching time T1, the multiplexer halts TS multiplexing for SDTV and starts TS multiplexing for HDTV. Transmission of the SDTV stream shall be terminated as GOP end immediately before the switching time. (sequence_end_code may be added at the end.) The HDTV sequence_header shall be transmitted as soon as possible after switching to the HDTV stream is complete.
6. The transmission of still-picture and mute signals, respectively, as video and audio data continues until the HDTV program starts (one second after the switching time). The actual HDTV program starts one second after T1.

Note 1: See Note 1 in 4.4.1, "Procedure for perfect seamless switching."

(2) Receiver operation

If a seamless switching-capable receiver processes signals according to the method described in Section 4.4.1 (2) (a), the SDTV stream is suddenly terminated halfway through processing, resulting in a state similar to when a serious transmission error occurs. Depending on decoder performance, it is possible to assume that a screen with block error is displayed because the picture decoded before the error cannot be displayed as a freeze-frame picture. Therefore, it is recommended that seamless switching-capable receivers process signals as follows in the same manner as seamless switching-incapable receivers in cases in which sequency_end_code_flag is 0:

1. The receiver obtains the new version of PMT.
2. Based on the contents of the PMT descriptor, when the receiver finds that switching from SDTV to HDTV will occur, it displays a freeze-frame picture and mutes the audio.
3. The video decoder halts SDTV decoding.
4. The Demux is set up to stop receiving streams with SDTV's ES_PID and to start feeding streams with HDTV's ES_PID to the decoder buffer.
5. By using the host CPU to monitor the sequence_header monitor register of the video decoder, the receiver awaits input of the HDTV stream.

6. When the decoder obtains HDTV stream sequence_header, it begins HDTV decoding. When ready to output normal video and audio data, the decoder cancels video freeze-frame and audio muting.

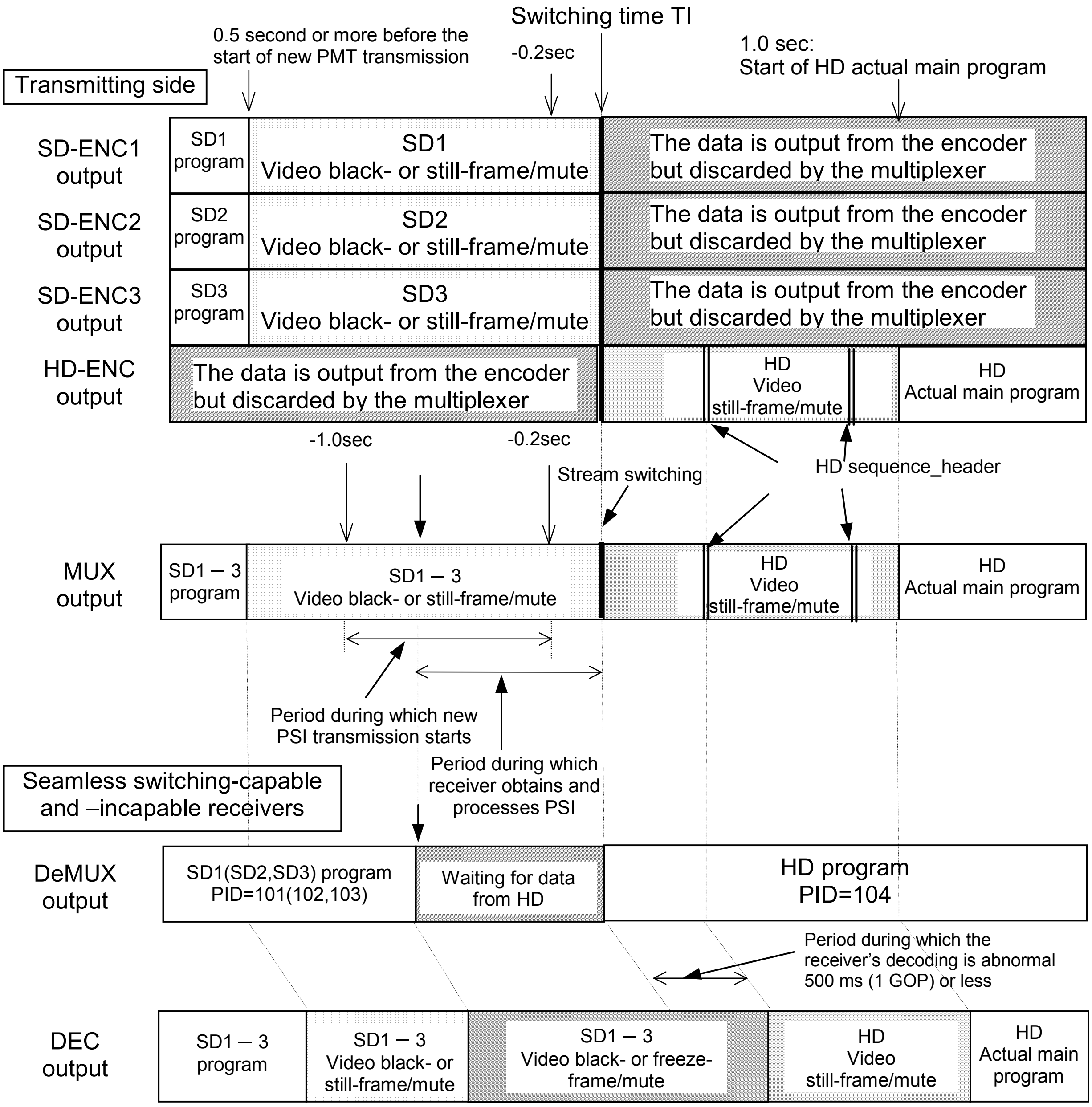


Fig. 4-2: Conceptual diagram of timing of transmitting and receiving sides in simplified procedure switching for SDTV and HDTV
(when sequence_end_code_flag of video_decode_control_descriptor is 0)

<Blank Page>

Chapter 5: An example of Encoding Film Materials

This chapter presents an example of encoding film materials by controlling the repeat_first_field, top_field_first, and progressive_frame flags of the picture layer. At this time, the same values are used for frame_rate_code and progressive_sequence of the sequence layer as for ordinary television pictures.

With interlaced scanning, when the encoder detects 2-3 pull-down, it sums two temporally equal fields, encodes both as a progressive frame, and sets the flag that indicates that the field corresponding to the third field of the 2-3 pull-down system is identical to the first field. No video data for that field is transmitted.

With progressive scanning, only 24 frames of video data are transmitted by setting the flag indicating that the first of 24 frames of film per second is displayed twice, the second three times, the third twice, the fourth three times, and so.

At this time, the decoder can reproduce the 2-3 sequence when the repeat_first_field and top_field_first flags are set or reset as shown below. (See “Fig. 5-1: Example of encoding film materials.”)

- Interlaced scanning

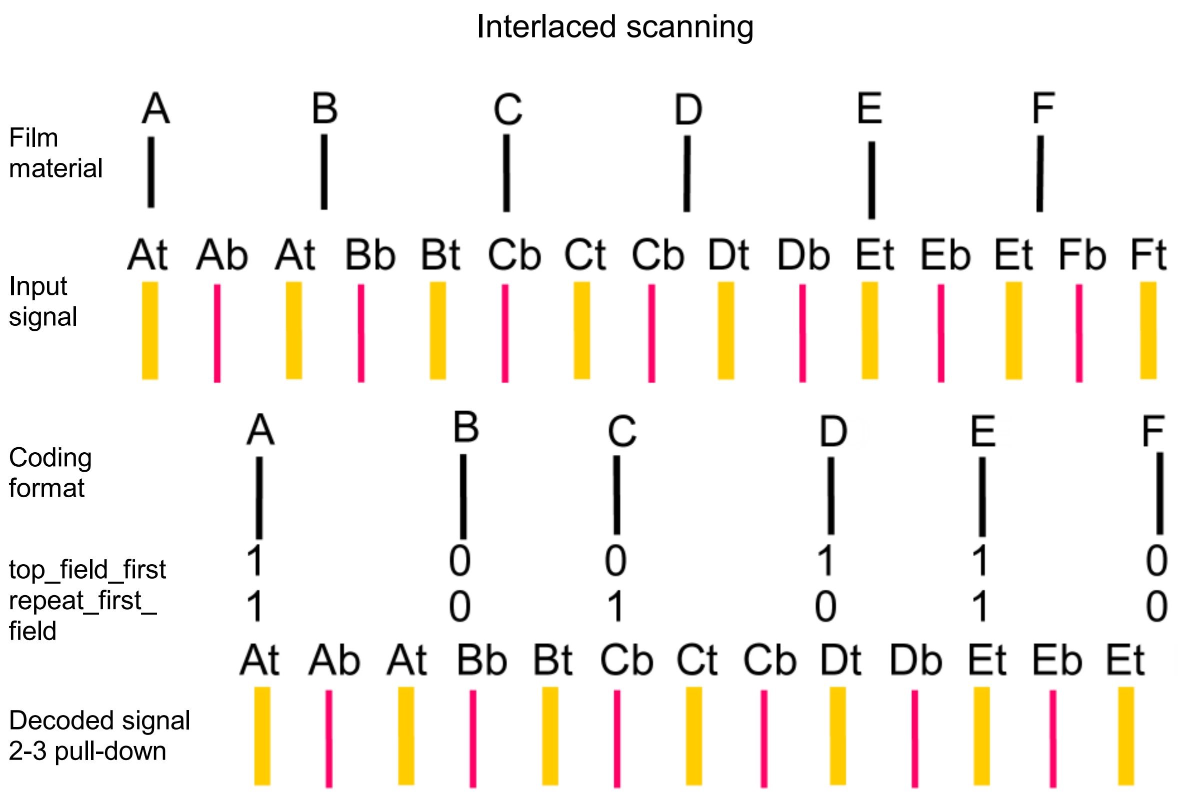
When repeat_first_field = 0, the decoded picture consists of two fields. Conversely, when repeat_first_field = 1, the decoded picture consists of three fields. Whether the top or bottom field is displayed first is specified by top_field_first.

repeat_first_field	top_field_first	Decoded picture (fields)
0	1	top / bottom
1	1	top / bottom / top
0	0	bottom / top
1	0	bottom / top / bottom

- Progressive scanning

The number of times each frame is to be displayed is specified by the combination of repeat_first_field and top_field_first.

repeat_first_field	top_field_first	Number of times each frame is to be displayed
0	0	1
1	0	2
1	1	3



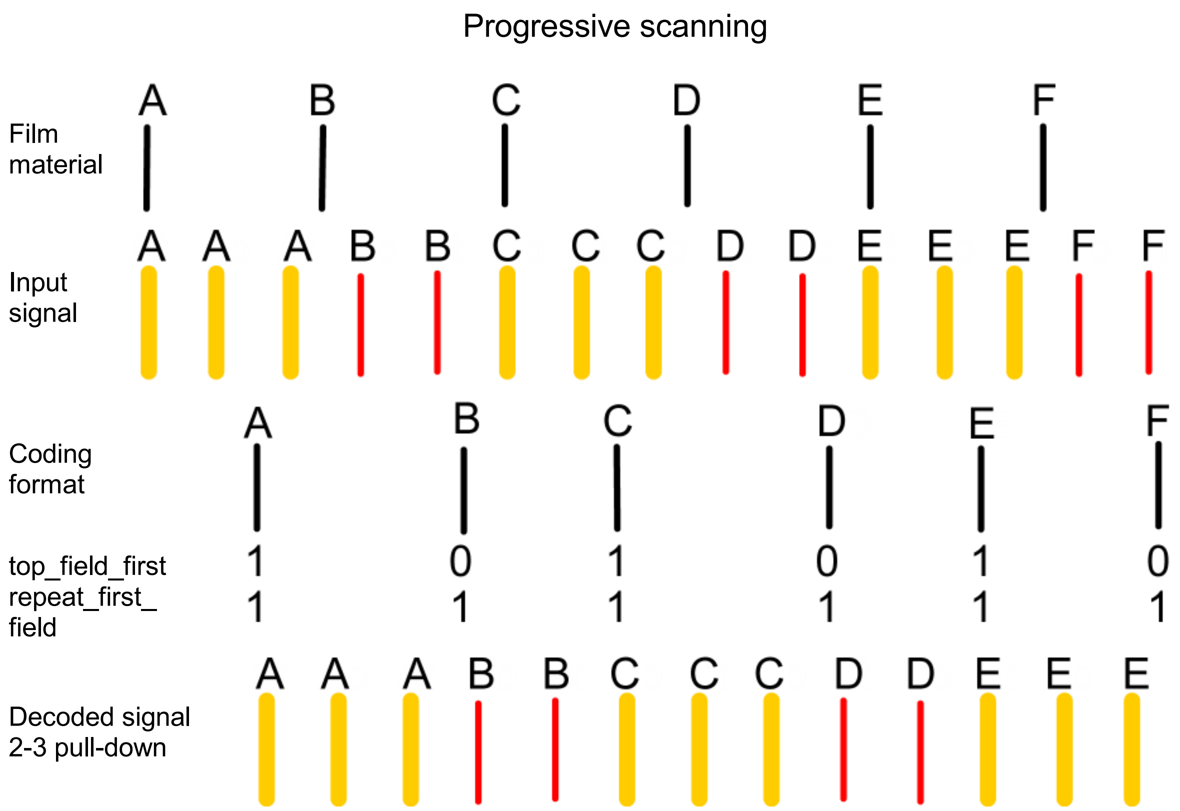


Fig. 5-1: Example of encoding film materials

Attachment 2: Operational Guidelines for MPEG-4 AVC Standard on television services

Chapter 1: General Terms

1.1 Objective

The purpose of these operational guidelines is to present technical recommendations of MPEG-4 AVC Standard in the practical operation concerning to video signals and video coding systems in digital television services.

1.2 Scope

These operational guidelines apply to video signals using MPEG-4 AVC Standard among video signals in television service that comply with the “Standard transmission system for digital broadcasting among standard television broadcasting and the like” (Ordinance).

1.3 References

1.3.1 Normative References

(1) Rec. ITU-T H.264|ISO/IEC 14496-10:2012: Advanced video coding for generic audiovisual services (hereinafter referred to as “MPEG-4 AVC Standard”)

(2) Rec. ITU-T H.222.0|ISO/IEC 13818-1:2006: Information technology -- Generic coding of moving pictures and associated audio information: Systems (hereinafter referred to as “MPEG-2 Systems Standard”)

1.4 Terminology

1.4.1 Abbreviations

CABAC	Context-based Adaptive Binary Arithmetic Coding
CAVLC	Context-based Adaptive Variable Length Coding
CBR	Constant Bit Rate
CPB	Coded Picture Buffer
DTS	Decoding Time-Stamp
GOP	Group of Pictures
HRD	Hypothetical Reference Decoder
IDR	Instantaneous Decoding Refresh
MBAFF	Macroblock-Adaptive Frame-Field Coding
NAL	Network Abstraction Layer
PES	Packetized Elementary Stream

POC	Picture Order Count
PPS	Picture Parameter Set
PTS	Presentation Time-Stamp
SEI	Supplemental Enhancement Information
SPS	Sequence Parameter Set
STD	System Target Decoder
VCL	Video Coding Layer
VLC	Variable Length Coding
VUI	Video Usability Information

Chapter 2: Summary of MPEG-4 AVC Standard

MPEG-4 AVC Standard has been standardized as Rec. ITU-T H.264 | ISO/IEC 14496-10 by the collaboration of VCEG in ITU-T and MPEG in ISO/IEC for next generation motion picture compression system with higher compression efficiency, comparing with MPEG-2 Video (ISO/IEC 13818-2) and MPEG-4 Visual (ISO/IEC 14496-2), Rec. ITU-T H.263.

In MPEG-4 AVC, the following technology has been introduced, aiming to improve picture quality comparing with conventional system. Adoption of integer precision DCT with 4×4 block size, smaller than 8×8 (8×8 integer precision DCT is available according to profile) and DCT transform of the differential value by intra-prediction for I-picture resulted in the reduction of the code amount for I-picture. On the other hand, for prediction to time direction, reference frame number in past and future direction is extended. Motion compensation block size which is from minimum 4×4 to maximum 16×16 is adaptively changed, so the precision of motion compensation is operated with 1/4 pixel unit. And making correlation with pictures maximum resulted in the reduction of the code amount.

Also, adoption of B-picture which can be referred, weighting prediction when the luminance of picture changes, introduction of quantized parameter which enables picture quality control easy, block noise reduction by using de-blocking filter (loop filter) when making reconstructed picture, these show various ideas on picture quality in detail.

Furthermore, CABAC (arithmetic code) with higher coding efficiency is added to CAVLC (Hufmann code) in Main profile and High profile.

In this way, it has been possible to make big reduction of generated code amount as a whole, and comparing with MPEG-2 Video around two times improvement of coding efficiency has been realized.

Here, also in MPEG-4 AVC Standard as in MPEG-2 Video Standard, in order to ensure inter-connectivity of encoder and decoder, profile and level restricts on coding data respectively. Profile restricts on coding tool and process range and so on, and level restricts on pixel size to be processed and frame frequency and so on.

<Blank Page>

Chapter 3: Restrictions on coding parameters

3.1 Profile and Level

Video coding system shall be based on High4:2:0 profile (including Main, High, High 10 profile) which is provided in MPEG-4 AVC Standard, and level shall correspond to video format; any of level 3, 3.1, and 3.2 for 480/60/I, either level 3.1 or 3.2 for 480/60/P, either level 3.2 or 4 for 720/60/P, level 4 for 1080/60/I, level 4.2 for 1080/60/P, and level 5.2[5] for 2160/60/P.

(Description)

As for video format 480/60/I, 480/60/P, 720/60/P and 4:2:0 of 1080/60/I, it shall be based on Main or High profile, and as for video format 10 bit and 4:2:2 of 1080/60/I, 1080/60/P, and 2160/60/P, it shall be based on any of High, High 10, High4:2:2 profile (See Table 3-1).

3.2 Video coding format

4:2:0 or 4:2:2 of Y, C_B, C_R shall be taken. The sampling position of color difference signals shall be the same as the position provided in MPEG-2 Video Standard. Three primary colors, gamma characteristics, matrix of luminance and color difference signals shall be based on the provision of this standard Part 1, 2.1.1. Bit number of each sample shall be 8 bits or 10 bits.

Combination of video format to be coded, profile and level shall be as shown in Table 3-1.

[5] As for adaptation of 2160/60/P, the preparation of environment such as the realization of LSI is the condition.

Table 3-1: Combination of Video format, Profile and Level

Input video format	Chroma format	Bit precision (bit)	Number of horizontal pixels	Number of vertical pixels	Frame rate (Hz)	Scanning system	Picture aspect ratio	Profile	Level
480/60/I	4:2:0	8	720	480	29.97	interlaced	4 : 3	Main, High	3, 3.1, 3.2
	4:2:0	8	720	480	29.97	interlaced	16 : 9	Main, High	3, 3.1, 3.2
480/60/P	4:2:0	8	720	480	59.94	progressive	16 : 9	Main, High	3.1, 3.2
720/60/P	4:2:0	8	1280	720	59.94	progressive	16 : 9	Main, High	3.2, 4
1080/60/I	4:2:0	8	1440	1080	29.97	interlaced	16 : 9	Main, High	4
	4:2:0	8	1920	1080	29.97	interlaced	16 : 9	Main, High	4
	4:2:0	10	1920	1080	29.97	interlaced	16 : 9	High10	4
	4:2:2	8, 10	1920	1080	29.97	interlaced	16 : 9	High4:2:2	4
1080/60/P	4:2:0	8	1920	1080	59.94	progressive	16 : 9	High	4.2
	4:2:0	10	1920	1080	59.94	progressive	16 : 9	High10	4.2
	4:2:2	8, 10	1920	1080	59.94	progressive	16 : 9	High4:2:2	4.2
2160/60/P	4:2:0	8	3840	2160	59.94	progressive	16 : 9	High	5.2
	4:2:0	10	3840	2160	59.94	progressive	16 : 9	High10	5.2
	4:2:2	8, 10	3840	2160	59.94	progressive	16 : 9	High4:2:2	5.2

(Description)

Though only luminance signal (without color difference signals) can be coded in High profile of MPEG-4 AVC Standard, it is not used for broadcasting. Also in MPEG-4 AVC Standard, though the sampling position of 4:2:0 color difference signals can be selected from plural patterns, phase change process of color difference signals becomes unnecessary and the processing becomes simple by selecting the same sampling position as that of MPEG-2 Video Standard, even if the mutual transform with MPEG-2 Video is necessary. As for colorimetry, conventional color gamut system or wide color gamut system can be selected according to the provision in this standard Part 1, 2.1.1. Furthermore, considering the objective of advancement of broadcasting, the video format in which the number of horizontal pixels is 1440 is not used for 10 bit and 4:2:2 of 1080/60/I.

3.3 Bitrate

Upper limit value of bitrate in NAL (Network Abstraction Layer) level is as the following.

Profile	Level	Upper limit value of bitrate
Main	3	12Mbit/s*
High	3	15Mbit/s*
Main	3.1	16.8Mbit/s*
High	3.1	20Mbit/s
Main, High	3.2, 4	20Mbit/s
High	4.2	30Mbit/s
High	5.2	80Mbit/s

* provided value in MPEG-4 AVC Standard

(Description)

Upper limit value of bitrate described here is to limit the result of multiplication of MaxBR and cpbBrNalFactor in MPEG-4 AVC Standard.

For video format with 8 bit sample, in MPEG-4 AVC Standard, almost good picture quality is expected by using bitrate mentioned above. As for video format with 10 bit sample and 4:2:2 sampling, it will be described after the result of picture quality inspection is obtained.

3.4 Frame rate

Frame rate shall be fixed in the sequence.

(Description)

Though frame interval can be variable in sequence in MPEG-4 AVC Standard, fixed frame rate shall be used in broadcasting.

3.5 Frame structure

Frame shall be the unit for coding, and 2 fields shall be included in 1 frame in case of interlace. Here, the value of picture timing SEI pic_struct shall correspond to 0～8 (including 3:2 pulldown、frame-doubling/tripling).

(Description)

Though a structure of one field (e.g. only top field) is permitted in MPEG-4 AVC Standard, in broadcasting, a frame must consist of 2 fields with top and bottom.

3.6 GOP structure

3.6.1 Arrangement of header information

GOP (Group of Pictures) starts from I-picture for decoding order, and only one sequence parameter set (equivalent to sequence header in MPEG-2 Video Standard) is arranged in the I-picture. The sequence parameter set must be essential for decoding this sequence. There are two kinds of I-picture; one is composed of only IDR slice, the other is composed of only I-slice.

Picture parameter set (equivalent to picture header in MPEG-2 Video Standard) must be arranged in the picture which refers to this picture parameter set.

(Description)

It is possible to send plural picture parameter sets together, or to make picture parameter set sent in advance to be the stream structure to which succeeding picture refers in MPEG-4 AVC Standard. But by operating to put picture parameter set on each picture as in MPEG-2 Video Standard, it becomes unnecessary to keep picture parameter set in decoder, so the processing is simplified.

3.6.2 Open/Closed GOP

Closed GOP and Open GOP are defined as the following.

- Closed GOP
  - GOP in which head I-picture in decoding order is IDR picture. When decoding starts from head of GOP, it is guaranteed all pictures in GOP can be decoded.

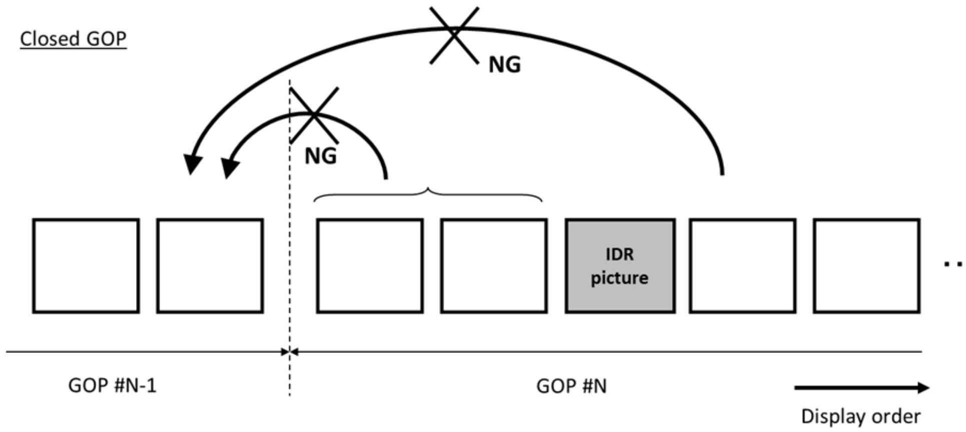


Fig. 3-1: Closed GOP

- Open GOP
  - GOP in which head I-picture in decoding order is non-IDR I-picture. When decoding starts from head of GOP, the picture whose display order is earlier than head I-picture may not be decoded normally.
  - Picture whose display order is later than head I-picture must be decoded normally.

In order to ensure decoding picture whose display order is later than head I-picture, open GOP shall satisfy the following provisions.

- Picture whose display order is earlier than head I-picture of GOP can refer to the picture in immediately before GOP.
- Picture whose display order is later than head I-picture of GOP cannot refer to the picture in immediately before GOP.

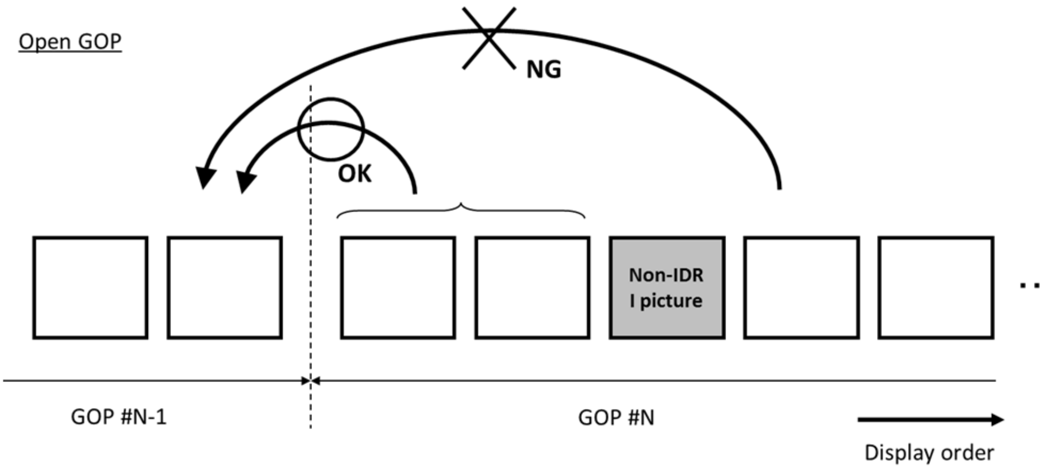


Fig. 3-2: Open GOP

3.6.3 Inter-picture prediction structure

Picture must be composed of only the same slice type.

In the picture of field structure, 1 frame must be composed of only I-field, only P-field, I-field and P-field, or only B-field. Hereinafter, “I-frame”, “P-frame”, “B-frame” are defined as pictures coded in frame structure, classifying pictures of field structure mentioned above.

I-picture and P-picture consist of only reference picture (nal_ref_idc is not 0), and decoding order must agree with display order. P-picture can be decoded by referring to only I-picture or P-picture in the same GOP (must not refer to other GOP or B-picture).

Here, there are two kinds of B-picture; B-picture to which can be referred by other picture (nal_ref_idc is not 0) and B-picture to which is not referred by other picture (nal_ref_idc is 0). And they are called as reference B-picture and non-reference B-picture respectively in this clause.

The decoding order of non-reference B-picture and reference B-picture shall be immediately after I-picture or P-picture whose display order is immediately after. Here, I-picture or P-picture shall be the picture in the same GOP as non-reference B-picture or reference B-picture.

Non-reference B-picture shall refer to only

(a) the frame pair or the field pair of I-picture or P-picture whose display order is immediately before or immediately after, or

(b) the frame pair or the field pair of reference B-picture whose display order is nearer than immediately before or immediately after I-picture or P-picture, and display order is immediately before or immediately after.

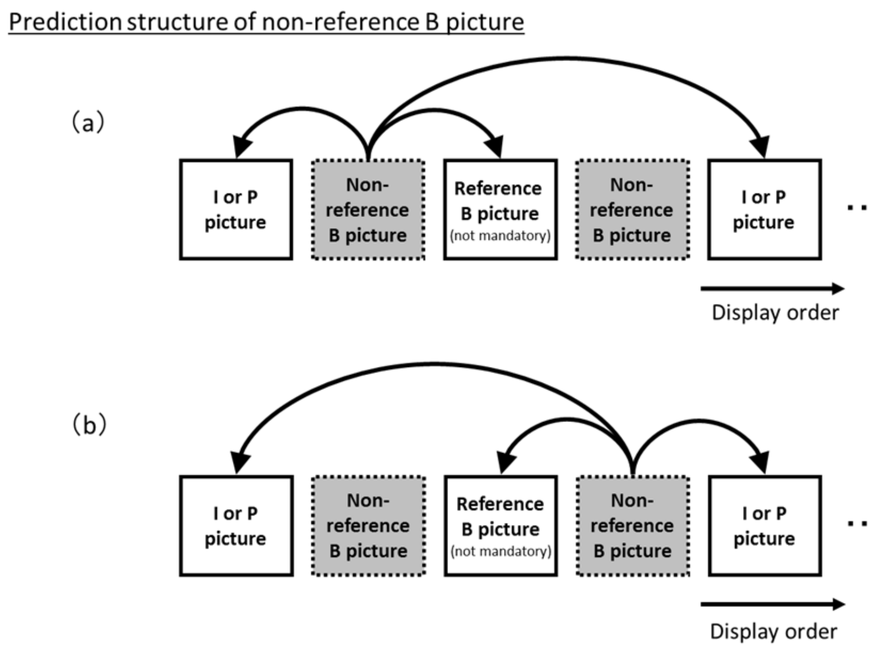


Fig. 3-3: Prediction structure of non-reference B picture

Reference B-picture shall refer to only

(a) the frame pair or the field pair of I-picture or P-picture whose display order is immediately before or immediately after, or

(b) the field of reference B-picture which composes the same frame.

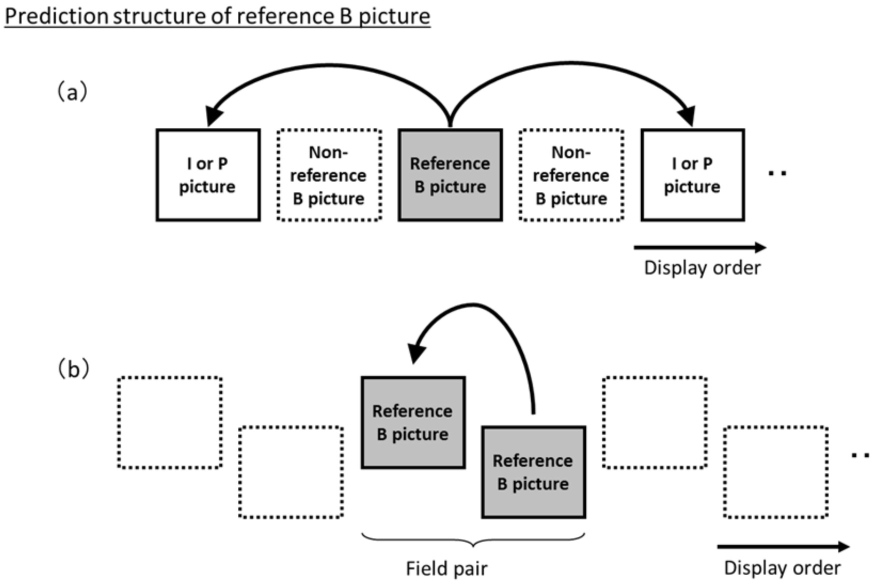


Fig. 3-4: Prediction structure of reference B-picture

Re-ordering between non-reference B-pictures must not be operated. (For non-reference B-pictures, decoding order must agree with display order.)

When decoded reference picture marking is operated in reference B-picture, in I-picture or P-picture whose decoding order is immediately after, the content of this decoded reference picture marking shall be transmitted again by decoded reference picture marking SEI.

Maximum number of the frames of successive B-picture (non-reference B-picture or reference B-picture) or field pair shall be 3.

The difference between the decoding time of I-picture whose decoding order is top in GOP and the display time of the picture whose display order is top in GOP shall be within 2 frame intervals. (When decoding starts from top of GOP, display can be started after waiting for maximum 2 frames interval.)

(Description)

In MPEG-4 AVC Standard, I-slice/P-slice/B-slice are permitted to be mingled in picture, and it is possible to realize a flexible prediction structure such that P-picture can use B-picture for prediction. But by limiting the rule to GOP structure similar to MPEG-2 Video Standard, it becomes easy to realize decoder. Moreover, it becomes possible to realize the function like fast-forward play in case of recording stream of broadcasting.

3.6.4 GOP length

GOP length shall be 500ms as a rule, and maximum 1s. (See Note)

(Note)

Considering an effect of picture quality improvement, an influence of channel-hopping and CPB buffer volume by making GOP length long, it is appropriate that GOP length shall be 500msec as a rule, and maximum 1s.

3.7 Data structure of bitstream

The order of NAL unit and SEI (Supplemental Enhancement Information) message which compose access unit shall be as shown in Figure 3-5 in the head of GOP and the access unit other than that. And NAL unit and SEI message which are not described must not be operated. The operations of each NAL unit and SEI message are as described in clause 3.7.1 and 3.7.2. However, Pan-scan SEI, decoded reference picture marking repetition SEI, filler data, end of sequence are not mandatory in access unit.

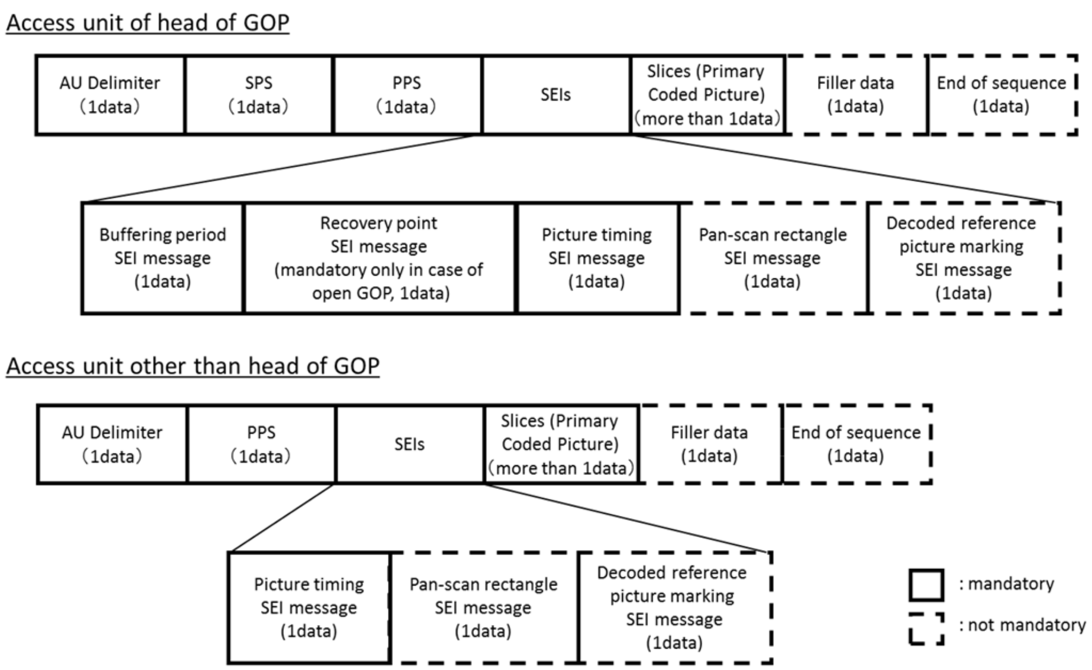


Fig. 3-5: Data structure of access unit

(Description)

Although decoding order of NAL unit which composes bitstream (ES) is much freely arranged in MPEG-4 AVC Standard, it is easier to make decoder by arranging data in order of processing.

It is provided that access unit delimiter is essential to transmit MPEG-4 AVC bitstream by MPEG-2 Systems TS (Rec. ITU-T H.222.0 | ISO/IEC 13818-1:2006/Amd3).

3.7.1 Access unit of the head of GOP

- Access unit delimiter (mandatory)
- Sequence parameter set (SPS) (mandatory)
  Always set vui_parameters_present_flag to 1, and send VUI (Video Usability Information).
  Always set nal_hrd_parameters_present_flag to 1 in VUI, and send HRD (Hypothetical Reference Decoder) information with NAL level.
- Picture parameter set (PPS) (mandatory)
- SEI
  - Buffering interval SEI (mandatory)
  - Recovery point SEI
    In order to avoid the disorder of picture at the head of Open GOP, recovery point SEI is surely added at the head of Open GOP. Also, it is operated as broken_link_flag = 0 in other case of transmission with splicing of bitstream.
  - Picture timing SEI (mandatory)
    Set pic_struct_present_flag to 1, and show the display order of each field by pic_struct of picture timing SEI.
  - Pan-scan SEI
    Details of operation are separately provided in clause 3.11 "Pan-scan rectangle SEI".
  - Decoded reference picture marking repetition SEI
- Slice data (decoded data of pixel value) (mandatory)
  Minimum unit of slice shall be 1 macro block sequence, or 1 macro block pair sequence.
  In GOP which starts from I-picture other than IDR, both in case of successive decoding from preceding GOP and in case of starting decoding from head I-picture of GOP by channel-hopping etc., slice data must be encoded so that decode and display of slice which do not refer to picture belonging to preceding GOP without being conscious of continuation from preceding GOP. (See Note 2)
- Filler data
- End of sequence

(Description)

As VUI includes important information about aspect ratio, matrix of luminance and color difference signals, frame rate, etc., it needs to include sequence parameter set equivalent to sequence header in MPEG-2 Video Standard.

In order to send HRD (Hypothetical Reference Decoder) information with NAL level, HRD parameter is included in VUI of sequence parameter set, and buffering interval SEI is sent at access unit of the head of GOP, and picture timing SEI is sent at every access unit.

Display order of each field is shown by pic_struct in picture timing SEI for the certain operation of display system.

3.7.2 Access unit except the head of GOP

- Access unit delimiter (mandatory)
- Picture parameter set (mandatory)
- SEI (Supplemental Enhancement Information)
  - Picture timing SEI (mandatory)
    Set pic_struct_present_flag to 1, and show the display order of each field by pic_struct of Picture timing SEI.
  - Pan-scan SEI
    Details of operation are separately provided in Clause 3.11 "Pan-scan rectangle SEI".
  - Decoded reference picture marking repetition SEI
- Slice data (decoded data of pixel value) (mandatory)
  Minimum unit of slice shall be 1 macro block sequence, or 1 macro block pair sequence.
  In GOP which starts from I-picture other than IDR, both in case of successive decoding from preceding GOP and in case of starting decoding from head I-picture of GOP by channel-hopping etc., slice data must be decoded so that decode and display of slice which do not refer to picture belonging to preceding GOP without being conscious of continuation from preceding GOP. (See Note)
- Filler data
- End of sequence

(Description)

Display order of each field is shown by pic_struct in picture timing SEI for the certain operation of display system.

(Note)

It should be guaranteed that later picture of display order from the head I-picture of OpenGOP can be decoded correctly. For example, when the head of OpenGOP is I/P-picture which has a field structure, if ref_pic_list_reordering() information etc. are not encoded, the translation of index number of the reference picture of I-field to P-field differs in the case that decoding succeeds from the preceding GOP and in the case that decoding starts from the head

I-picture by channel-hopping etc., so there is a possibility not to decode correctly.

By coding ref_pic_list_reordering() information etc. in P-field and reference B-field as necessity, index number of the reference picture can be translated uniquely, and correct decoding is possible.

3.8 Identifier which represents sequence end

- As for identifier which represents sequence end, end of sequence NAL unit shall be used. (End of stream NAL unit is not used.)
- When end of sequence NAL unit is transmitted, it shall be transmitted immediately before head access unit of closed GOP (GOP which starts from IDR picture).

On the receiver side, it is recommended that when end of sequence NAL unit is received, the picture of video data which is received immediately before end of sequence NAL unit is displayed in frozen until the video data transmitted after is correctly decoded and displayed. This means that if the video data transmitted after end of sequence NAL unit can be promptly decoded and displayed, data processing continues seamlessly, and does not mean that the picture always be displayed in frozen for affixed period of time.

3.9 Coding tool

3.9.1 Two-way motion compensation block size

MinLumaBiPredSize shall be 8×8. (even if it is below level 4)

(Description)

There is a coding tool which can be used only below level 4 (two-way motion compensation in block size with less than 8×8 pixels) in MPEG-4 AVC Standard. As level 4 is mandatory in order to implement HDTV, prohibiting coding tool which cannot work below level 4 makes it easy to realize decoder.

3.9.2 Number of reference frame

Maximum number of reference frame (num_ref_frames) shall be 4.

(Description)

The number of reference frame which can be used when HDTV is decoded at level 4 is 4. Therefore, setting maximum number of reference frame to 4 for other pixels or level makes it easy to realize decoder.

3.9.3 CPB size

CPB size shall use the values described in the following.

Profile	Level	CPB size
Main	3	less than or equal to 12Mbit*
High	3	less than or equal to 15Mbit *
Main	3.1	less than or equal to 16.8Mbit*
High	3.1	less than or equal to 20Mbit
Main, High	3.2, 4	less than or equal to 20Mbit
High	4.2	less than or equal to 30Mbit
High	5.2	less than or equal to 80Mbit

* provided value in MPEG-4 AVC Standard

(Description)

CPB size described here is that size which limits the result of multiplication of MaxCPB by cpbBrNalFactor in MPEG-4 AVC Standard.

CPB size shall be the size of necessary minimum considering memory capacitance and ease of AV synchronization. As maximum value of GOP length is 1s, the size corresponding to this for upper limit bitrate is appropriate. Further, as for video format of 10 bit sample and 4:2:2 sampling, it is not described because its bitrate is not provided in these guidelines.

3.9.4 Minimum compression ratio

MinCR (Minimum Compression Ratio) shall use the following value. Also when slice partitioning, each slice which is included in 1 access unit shall be satisfied with MinCR.

Level	MinCR
3	2
Other than 3 (3.1～5.2)	4

(Description)

As MinCR is limited to 4 in MPEG-4 AVC Standard for level 3.1 to 4 including level 4 which is widely used for HDTV, by corresponding the limitation even in higher level than level 4, it makes easy to realize decoder. Also when slice partitioning, by limiting MinCR to each slice, the amount of codes does not cluster to specific slice, and it becomes easy to realize decoder.

3.9.5 Combination of Profile, bit precision and chroma format

Profile, bit precision and chroma format which can be combined shall be as the following.

Profile	Bit precision, Chroma format
Main, High (Profile_idc =77, 100)	bit_depth_luma_minus8 = 0 bit_depth_chroma_minus8 = 0 chroma_format_idc = 1
High10 (Profile_idc = 110)	bit_depth_luma_minus8 = 2 (only) bit_depth_chroma_minus8 = 2 (only) chroma_format_idc = 1
High4:2:2 (Profile_idc = 122)	bit_depth_luma_minus8 = 0 or 2 bit_depth_chroma_minus8 = 0 or 2 chroma_format_idc = 2 (only)

(Description)

Profile of necessary minimum shall be used according to bit precision and chroma format. Also, 9 bit shall not be used for bit precision of sample.

3.9.6 Slice partitioning

As for level 5.2, the slice structure shall be that which the number of macro block in the slice is not more than MaxFS / 4, namely 1 picture is partitioned into more than 4 pictures.

(Description)

It is provided in MPEG-4 AVC Standard that High4:2:2 and High10 profile, etc. whose processing load is heavier than High profile shall apply to this restriction in case that PicSizeInMbs is bigger than 1620.

By using slice partitioning structure, CABAC data etc. can be processed in parallel for each slice. So, it becomes easy to realize a decoder and flexibility of design raises for receivers such as decoder with software as a base. In addition to this, improvement of error tolerance can be expected. On the other hand, motion compensation prediction and de-blocking filter can be processed over slices even in slice partitioning structure, but intra-prediction is limited to the processing in slice.

As for the application of this restriction to below level 4.2 in High profile, it is necessary to judge after a further evaluation on the effect to coding efficiency and picture quality by slice partitioning, required processing ability of decoder, trend of performance of devices for decoder and so on.

3.9.7 Entropy coding

CAVLC and CABAC which are two kinds of variable length coding system shall be able to change only over a suspension of broadcasting.

(Description)

Implementation for changing 2 variable length codings instantaneously in order not to suspend video becomes not necessary, and it becomes easy to realize decoder.

3.9.8 Picture Order Count

Type of POC (Picture Order Count) shall be 0.

(Description)

Type 1 is complicated and it is not so effective in reducing the amount of code for bitrate of broadcasting. Type 0 includes all types and it makes the value of syntax element necessary minimum, so only type 0 shall be used.

3.10 HRD conformance

HRD must satisfy Output Timing Conformance of type 2 (NAL level). Also, data which is input to CPB must be decoded within 1 second.

3.11 Syntax

3.11.1 NAL unit

Syntax element	Operation	Remarks
nal_ref_idc	Any of 0, 1, 2, 3	0: non reference picture 1,2,3: reference picture ・not 0: in case of I-picture or P-picture 0 or not 0: in case of B-picture ・The values of nal_ref_idc of 2 fields which compose 1 frame are the same.
Nal_unit_type	Any of 1, 5, 6, 7, 8, 9, 10, 12	1: slice other than IDR picture 5: slice of IDR picture 6: SEI 7: sequence parameter set 8: picture parameter set 9: access unit delimiter 10: end of sequence 12: filler data

3.11.2 Sequence parameter set

Syntax element	Operation	Remarks
profile_idc	Any of 77, 100, 110, 122	77: Main profile 100: High profile 110: High 10 profile 122: High 4:2:2 profile (Note) See Table 3-1 on combination with video format
level_idc	Any of 30, 31, 32, 40, 42, 52	30: level 3 31: level 3.1 32: level 3.2 40: level 4 42: level 4.2 52: level 5.2 (Note) See Table 3-1 on combination with video format
seq_parameter_set_id	0	0 fixed
chroma_format_idc	1 or 2	1: 4:2:0 format 2: 4:2:2 format
bit_depth_luma_minus8	0 or 2	0: luminance pixel value consists of 8 bits 2: luminance pixel value consists of 10 bits
bit_depth_chroma_minus8	0 or 2	0: chroma pixel value consists of 8 bits 2: chroma pixel value consists of 10 bits
qpprime_y_zerotransform_bypass_flag	0	0: reversible coding mode is not used
pic_order_cnt_type	0	0: mode which represents display order by difference from IDR immediately before
num_ref_frames	1～4	This represents maximum value of reference frame number.
gaps_in_frame_num_value_allowed_flag	0	0: Decoding movement is not provided when frame number is discontinuous.
pic_width_in_mbs minus1	See Table 3-2	This represents macro block number－1 in horizontal direction.
pic_height_in_map_unit_minus1	See Table 3-2	This represents macro block number－1 in vertical direction.
frame_mbs_only_flag	0 or 1 See Table 3-2 and Table 3-3	1: only frame macro block 0: field macro block or MBAFF is permitted. mb_adaptive_frame_field_flag is set to 0 or 1. Only in case of progressive scan video, set this to 1.
direct_8x8_inference_flag	1	1: Direct mode prediction coding with less than 8×8 block size is not used.
frame_cropping_flag	See Table 3-3	0: All decoded video is displayed. 1: A part of decoded video is picked out and displayed.
frame_crop_left_offset	See Table 3-3	This represents half value of the number of leftmost pixels which are not displayed in decoded video.
frame_crop_right_offset	See Table 3-3	This represents half value of the number of rightmost pixels which are not displayed in decoded video.

frame_crop_top_offset	See Table 3-3	This represents half or quarter value of the number of upper end pixels which are not displayed in decoded video.
frame_crop_bottom_offset	See Table 3-3	This represents half or quarter value of the number of lower end pixels which are not displayed in decoded video.
vui_parameters_present_flag	1	1: VUI (Video Usability Information) is coded.

3.11.3 Picture parameter set

Syntax element	Operation	Remarks
pic_parameter_set_id	0	0 fixed
entropy_coding_mode_flag	0 or 1	0: CAVLC 1: CABAC This value shall be constant in all picture parameter set which is included in sequence.
pic_order_present_flag	1	1 fixed
num_slice_groups_minus1	0	0: Slice group is not used.
num_ref_idx_l0_active_minus1	0～7	This represents maximum value of L0 reference picture number －1 in the setting value of num_ref_frames. P frame: 0～3 B frame: 0～1 P field: 0～7 B field: 0～3
num_ref_idx_l1_active_minus1	0～3	This represents maximum value of L1 reference picture number －1 in the setting value of num_ref_frames. B frame: 0～1 B field: 0～3
pic_init_qs_minus26	0	0 fixed: Because SP or SI slice is not used.
redundant_pic_cnt_present_flag	0	0: Redundant slice is not used.

3.11.4 Access unit delimiter

Syntax element	Operation	Remarks
primary_pic_type	Any of 0, 1, 2	I-picture: 0 P-picture: 1 B-picture: 2

3.11.5 Slice header

Syntax element	Operation	Remarks
slice_type	Any of 7, 5, 6	I-picture: 7 P-picture: 5

		B-picture: 6
num_ref_idx_l0_active_minus1	0～7	This represents maximum value of L0 reference picture number－1 in the setting range of num_ref_frames. P frame: 0～3 B frame: 0～1 P field: 0～7 B field: 0～3
num_ref_idx_l1_active_minus1	0～3	This represents maximum value of L1 reference picture number－1 in the setting range of num_ref_frames. B frame: 0～1 B field: 0～3

3.11.6 Decoded Reference Picture Marking Syntax

Syntax element	Operation	Remarks
no_output_of_prior_pics_flag	0	0: Decoded picture is displayed. (Picture which is not displayed is not decoded.)

3.11.7 VUI

Syntax element	Operation	Remarks
aspect_ratio_info_present_flag	1	Information of aspect ratio is essential.
aspect_ratio_idc	See Table 3-3	This represents pixel aspect ratio.
sar_width	4	When resolution is 1440×1080, if aspect_ratio_idc=255, this syntax shall be essential.
sar_height	3	When resolution is 1440×1080, if aspect_ratio_idc=255, this syntax shall be essential.
video_full_range_flag	0	0: based on Rec. ITU-R BT.709-5
colour_primaries	1	1: based on Rec. ITU-R BT.709-5
transfer_characteristics	Either 1 or 11	1: based on Rec. ITU-R BT.709-5, Rec. ITU-R BT.1361 conventional color gamut system 11: based on IEC61966-2-4 (wide color gamut system)
matrix_coefficients	1	1: based on Rec. ITU-R BT.709-5
chroma_loc_info_present_flag	0	0: the same as the sample position of 4:2:0 color difference signals in MPEG-2 Video Standard
timing_info_present_flag	1	1: num_units_in_tick, time_scale, fixed_frame_rate_flag which represent the frame rate in case of fixed frame rate are included in syntax elements. Frame-rate = time_scale /

		num_units_in_tick / 2 Note: About detail calculation method of frame rate, see the semantics of fixed_frame_rate_flag in Annex E of MPEG-4 AVC Standard.
num_units_in_tick	1001	1001 fixed
time_scale	Either 60000 or 120000	60000 is set in case that frame rate is 29.97Hz, and 120000 is set in case that frame rate is 59.94Hz.
fixed_frame_rate_flag	1	1: fixed frame rate
nal_hrd_parameters_present_flag	1	1: NAL HRD parameter which represents bitrate and buffer information is included in syntax elements.
vcl_hrd_parameters_present_flag	0	0: VCL HRD parameter which represents bit rate and buffer information is not included in syntax element.
low_delay_hrd_flag	0	0: Under flow of receiving buffer is not permitted.
pic_struct_present_flag	1	1 fixed

Table 3-2: Combination of parameters which represent picture size (No. 1)

Number of horizontal pixels	Number of vertical pixels	pic_width_in_mbs_minus1	pic_height_in_map_units_minus1	frame_mbs_only_flag	Frame rate (Hz)	Scanning system
720	480	44	14	0	29.97	interlaced
720	480	44	29	1	59.94	progressive
1280	720	79	44	1	59.94	progressive
1440	1080	89	33	0	29.97	interlaced
1920	1080	119	33	0	29.97	interlaced
1920	1080	119	67	1	59.94	progressive
3840	2160	239	134	1	59.94	progressive

Table 3-2: Combination of parameters which represents picture size (No. 2)

Picture aspect ratio	Number of horizontal pixels	Number of vertical pixels	aspect_ratio_idc	frame_mbs_only_flag	frame_cropping_flag	frame_crop_left_offset	frame_crop_right_offset	frame_crop_top_offset	frame_crop_bottom_offset
4:3	720	480	3	0	0	0	0	0	0
16:9	720	480	5	0	0	0	0	0	0
16:9	720	480	5	1	0	0	0	0	0
16:9	1280	720	1	1	0	0	0	0	0
16:9	1440	1080	255 or 14 (Note)	0	1	0	0	0	2
16:9	1920	1080	1	0	1	0	0	0	2
16:9	1920	1080	1	1	1	0	0	0	4
16:9	3840	2160	1	1	0	0	0	0	0

(Note) aspect_ratio_idc=14 is not provided in early standard of MPEG-4 AVC Standard, so operation using 255 is recommended.

3.11.8 Pan-scan rectangle SEI

When video data is transmitted with different picture aspect ratio such as side panel or letter box from original video source, it is possible to avoid displaying black frame (picture frame) by display angle of the receiver, by setting parameters of pan-scan shown as the following. So, according to Fig. 3-6: "Desirable display formats on 4:3 and 16:9 aspect ratio monitors," Pan-scan rectangle SEI is coded to head I-picture (IDR picture for closed GOP, and I-picture of non-IDR for open GOP) as the necessity.

Here, in case that pan-scan is operated (Fig. 3-6 ② or ④), Pan-scan rectangle SEI must be coded.

Each parameter in case of the above operation is shown in Table 3-4 and Table 3-5.

Table 3-4: Parameters for pan-scan operation

		VUI Parameter	Parameters of Sequence parameter set			Parameters of Pan-scan rectangle SEI				
Picture width	Picture height	aspect_ratio_idc	pic_width_in_mbs_minus1	pic_height_in_map_units_minus1	frame_mbs_only_flag	pan_scan_rect_left_offset	pan_scan_rect_right_offset	pan_scan_rect_top_offset	pan_scan_rect_bottom_offset	Fig. 3-6
720	480	5	44	29	1	0	0	0	0	①
720	480	5	44	29	1	1440	−1440	0	0	②
720	480	5	44	14	0	0	0	0	0	①
720	480	5	44	14	0	1440	−1440	0	0	②
720	480	3	44	14	0	0	0	0	0	③
720	480	3	44	14	0	0	0	960	−960	④
1280	720	1	79	44	1	0	0	0	0	①
1280	720	1	79	44	1	2560	−2560	0	0	②
1440	1080	255(sar_width=4,sar_height=3)	89	33	0	0	0	0	0	①
1440	1080	255(sar_width=4,sar_height=3)	89	33	0	2880	−2880	0	0	②
1920	1080	1	119	33	0	0	0	0	0	①
1920	1080	1	119	33	0	3840	−3840	0	0	②
1920	1080	1	119	67	1	0	0	0	0	①
1920	1080	1	119	67	1	3840	−3840	0	0	②
3840	2160	1	239	134	1	0	0	0	0	①
3840	2160	1	239	134	1	7680	−7680	0	0	②

Table 3-5: Another syntax elements of Pan-Scan SEI

Syntax elements	Operation	Remarks
pan_scan_rect_id	0	Pan-scan information is not discriminated by ID.
pan_scan_rect_cancel_flag	0	Pan-scan information is always transmitted.
pan_scan_cnt_minus1	0	Pan-scan information is only one kind.
pan_scan_rect_repetition_period	1	Pan-scan information is valid until next sequence, or immediately before next picture to which Pan-Scan SEI is added.

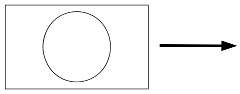
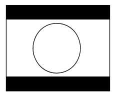
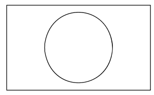
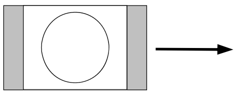
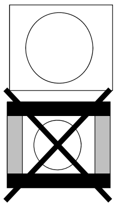
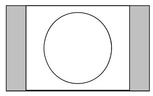
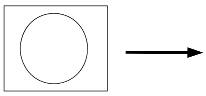
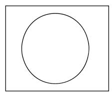
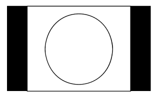
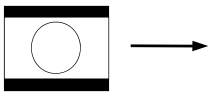
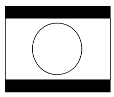
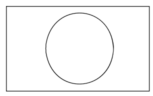
	Video source	Monitor with 4:3 aspect ratio	Monitor with 16:9 aspect ratio
① 16:9 program No.1		The program is displayed in letterbox format on a 4:3 monitor. 	The program is displayed as is on a 16:9 monitor. 
② 16:9 program No.2		The program is displayed over the entire screen (480 ×720) of the 4:3 monitor. Note that side panels are discarded. 	The program is displayed as is on a 16:9 monitor.  Gray area indicates two cases: one in which this area contains a real picture and one is in which the area consists of a black panel.
③ 4:3 program		The 4:3 program is displayed as is on a 4:3 monitor. 	The program is displayed with side panels on a 16:9 monitor. With 480/I system, appropriate change are made to the monitor's deflection system to allow the program to be displayed. 
④ 4:3 program in letterbox format		The program is displayed as is on a 4:3 monitor. 	The program is displayed on a 16:9 monitor after multiplication of the program in the vertical deflection by 4/3, 2, and 3 to produce 480, 720, and 1080 valid lines respectively. With the 480/I system, appropriate change are made to the monitor's deflection system to be displayed. 

Fig. 3-6: Desirable display formats on 4:3 and 16:9 aspect ratio monitors

<Blank Page>

Chapter 4: Seamless Switching

As for seamless switching in the same codec, operation shall be the same as this standard Part1, Attachment 1, Chapter 4 "Seamless Switching."

4.1 Changing the number of active samples

(1) Procedure on the transmitting side

- Sequence is stopped using end of sequence NAL unit at operation switching point, and new sample number is specified in sequence parameter set of next GOP.
- The top GOP of new working sequence shall be closed GOP.
- cpb_size_scale and cpb_size_value_minus1 in hrd_parameters() shall not be changed before and after the switching.
- The difference between decoding time of I-picture whose decoding order is head in GOP and display time of picture whose display order is head in GOP shall not be changed before and after the switching.
- Continuity of PTS and DTS is guaranteed.

(2) Receiver operation

Working mode is set by the parameter of pixel number which is included in received sequence parameter set. Even if end of sequence NAL unit is not received, new working mode is set according to the contents of received sequence parameter set.

4.2 Changing picture aspect ratio for 480/60/I system

(1) Procedure on the transmitting side

- Sequence is stopped using end of sequence NAL unit at operation switching point, and new aspect ratio is specified in sequence parameter set of next GOP.
- The top GOP of new working sequence shall be closed GOP.
- cpb_size_scale and cpb_size_value_minus1 in hrd_parameters() shall not be changed before and after the switching.
- The difference between decoding time of I-picture whose decoding order is head in GOP and display time of picture whose display order is head in GOP shall not be changed before and after the switching.
- Continuity of PTS and DTS is guaranteed.

(2) Receiver operation

Working mode will be set by the parameter of aspect ratio which is included in received sequence parameter set. Even if end of sequence unit is not received, new working mode will be set according to the contents of received sequence parameter set.

4.3 Changing bitrate

(1) Procedure on the transmitting side

- Operation is always in variable bit rate mode.
  - cbr_flag in hrd_parameters() is set to 0.
- End of Sequence NAL unit is not inserted at transmission bit rate changing point.
- cpb_size_scale and cpb_size_value_minus1 in hrd_parameters() shall not be changed before and after the switching.
- The difference between decoding time of I-picture whose decoding order is head in GOP and display time of picture whose display order is head in GOP shall not be changed before and after the switching.
- Continuity of PTS and DTS is guaranteed.

(2) Receiver operation

In the receiver, decoding start and output of video and audio signals shall be controlled in accordance with PTS and DTS which are described in PES Header, and the receiver shall work seamlessly by this control.

Chapter 5: Multiplex by MPEG-2 Systems Standard

5.1 PES packet

- PES packet must be composed of the access unit which always constructs one frame or one field pair (it must not include plural frames or field pair).
- PTS must be transmitted in PES Header. In the receiver, start of decoding and output shall be controlled in accordance with PTS and DTS of PES Header. And the setting value of PTS_DTS_flag is as the following.
  - 11b: in the case that PES packet includes I-picture, P-picture, or B-picture in which PTS and DTS are different
  - 10b: in the case that PES packet includes B-picture in which PTS and DTS are the same

5.2 STD delay

- STD delay shall be maximum 1 second in normal operation.

5.3 Descriptors

- The following descriptors defined in MPEG-2 Systems Standard shall not be operated.
  - AVC video descriptor
  - AVC timing and HRD descriptor

<Blank Page>

Attachment 3: Operational Guidelines for MPEG-4 AVC Standard on low definition video services

Chapter 1: General Terms

1.1 Objective

The purpose of these operational guidelines is to present technical recommendations of MPEG-4 AVC Standard in the practical operation concerning to video signals and video coding systems in low definition video services (See main clause, Chapter 5, 5.2).

1.2 Scope

These operational guidelines apply to video signals using MPEG-4 AVC Standard among the video signals in digital broadcasting that comply with the "Standard transmission system for digital broadcasting among standard television broadcasting and the like" (Ordinance).

1.3 References

1.3.1 Normative References

(1) Rec. ITU-T H.264|ISO/IEC 14496-10:2012: Advanced video coding for generic audiovisual services (hereinafter referred to as "MPEG-4 AVC Standard")

1.4 Terminology

1.4.1 Abbreviations

CIF	Common Intermediate Format
HHR	Half Horizontal Resolution
IDR	Instantaneous Decoding Refresh
QCIF	Quarter Common Intermediate Format
QSIF	Quarter Source Input Format
QVGA	Quarter VGA
SD	Standard Definition
SIF	Source Input Format
SQVGA	Sub Quarter VGA
VGA	Video Graphics Array
VUI	Video Usability Information

<Blank Page>

Chapter 2: Restrictions on coding parameters

2.1 Profile and Level

Video coding system shall be based on Baseline or Main profile provided in MPEG-4 AVC Standard, and level shall be any of 1, 1.1, 1.2, 1.3, 2, 2.1, 2.2, and 3.

Maximum picture size and frame rate (macro block number per unit time) are provided according to level, and it is desirable to decide level to be operated and video coding format, considering format of resource, display device of the receiver, the processing and so on. Also, usable tool is defined according to profile, and it is desirable to select profile considering requirement and services.

2.2 Video format and corresponding parameters

2.2.1 Assumed video format

Assumed video format and corresponding parameters are shown in Table 2-1. About 16:9 picture of SQVGA and QVGA, the pixel aspect shall be the same as those of 4:3 picture, and picture size shall be those which the number of vertical pixels is reduced.

Table 2-1: Assumed video format

Format	Picture size (H×V)	Picture aspect ratio (H : V)	seq_parameter_set_rbsp( )		vui_parameters( )	
			pic_width_in_mbs_minus1	pic_height_in_map_units_minus1	aspect_ratio_info_present_flag	aspect_ratio_info
SQVGA	160×120	4:3	9	7 ※	1	1
SQVGA	160×90	16:9	9	5 ※		1
525QSIF	176×120	4:3	10	7 ※		3
525QSIF	176×120	16:9	10	7 ※		5
QCIF	176×144	4:3	10	8		2
QVGA	320×240	4:3	19	14		1
QVGA	320×180	16:9	19	11 ※		1
525SIF	352×240	4:3	21	14		3
525SIF	352×240	16:9	21	14		5
CIF	352×288	4:3	21	17		2
525HHR	352×480	4:3	21	29		3
525HHR	352×480	16:9	21	29		5
VGA	640×480	4:3	39	29		1
525SD	720×480	4:3	44	29		3
525SD	720×480	16:9	44	29		5

※ When width or height of the picture is not divided by 16 without remain, the encoder adds fictional video data (dummy data) on the right side of active sample or below active line, and the data is practically processed by the number of samples or the number of lines as a multiple of 16. The decoder discards the dummy data, outputting as active samples or active lines.

2.2.2 Frame rate

Frame rate shall be integer times of 1000/1001, calculating as = time_scale/num_units_in_tick by using variables of VUI Parameters. But maximum frame rate [Hz] in each level shall be as shown in Table 2-2 for the video format in operation.

Table 2-2: Maximum frame rate in each level [Hz]

	1	1.1	1.2	1.3	2	2.1
SQVGA(4:3)	15000/1001	30000/1001	30000/1001	30000/1001	30000/1001	30000/1001
SQVGA(16:9)	24000/1001	30000/1001	30000/1001	30000/1001	30000/1001	30000/1001
525QSIF(4:3)	15000/1001	30000/1001	30000/1001	30000/1001	30000/1001	30000/1001
525QSIF(16:9)	15000/1001	30000/1001	30000/1001	30000/1001	30000/1001	30000/1001
QCIF	15000/1001	30000/1001	30000/1001	30000/1001	30000/1001	30000/1001
QVGA(4:3)	-	10000/1001	15000/1001	30000/1001	30000/1001	30000/1001
QVGA(16:9)	-	12000/1001	24000/1001	30000/1001	30000/1001	30000/1001
525SIF(4:3)	-	15000/2002	15000/1001	30000/1001	30000/1001	30000/1001
525SIF(16:9)	-	15000/2002	15000/1001	30000/1001	30000/1001	30000/1001
CIF	-	15000/2002	15000/1001	30000/1001	30000/1001	30000/1001
525HHR(4:3)	-	-	-	-	-	30000/1001
525HHR(16:9)	-	-	-	-	-	30000/1001
VGA	-	-	-	-	-	-
525SD(4:3)	-	-	-	-	-	-
525SD(16:9)	-	-	-	-	-	-

	2.2	3
SQVGA(4:3)	30000/1001	30000/1001
SQVGA(16:9)	30000/1001	30000/1001
525QSIF(4:3)	30000/1001	30000/1001
525QSIF(16:9)	30000/1001	30000/1001
QCIF	30000/1001	30000/1001
QVGA(4:3)	30000/1001	30000/1001
QVGA(16:9)	30000/1001	30000/1001
525SIF(4:3)	30000/1001	30000/1001
525SIF(16:9)	30000/1001	30000/1001
CIF	30000/1001	30000/1001
525HHR(4:3)	30000/1001	30000/1001
525HHR(16:9)	30000/1001	30000/1001
VGA	15000/1001	30000/1001
525SD(4:3)	15000/1001	30000/1001
525SD(16:9)	15000/1001	30000/1001

2.2.3 Color description

In case that video_signal_type_present_flag = 0 or colour_description_present_flag = 0 in VUI Parameters, all values of colour_primaries, transfer_characteristics,and matrix_coefficients will be 2 (Unspecified). But all values shall be interpreted as equivalent to 1 (Rec. ITU-R BT.709) in the decoder.

2.3 Operational guidelines considering channel-hopping time

① Insert I-picture of IDR type for maximum 5 seconds, with usually 2 seconds interval.

② In case that the parameters of Sequence Parameter Set are different, it is desirable to use different seq_parameter_set_id.

2.4 Desirable operational guidelines in Baseline profile

(1) Assumed service requirement

・bitrate: 64～768 kbit/s

・video format: SQVGA, 525QSIF, QCIF, QVGA, 525SIF, CIF

・frame rate: 5, 10, 12, 15, 24, 30Hz (each shall be 1000/1001 times frequency)

But frame skip shall not be limited.

・Picture aspect ratio: 4:3, 16:9

(2) Operation level

Any of level 1, 1.1, 1.2, and 1.3 is operated. Level shall be selected according to coding format.

(3) Major restriction on operation

・FMO (Flexible Macroblock Ordering), ASO (Arbitrary Slice Order), RS (Redundant Slices) are not operated. constraint_set0_flag=1 and constraint_set1_flag=1 for Sequence Parameter Set,.

2.5 Desirable operational guidelines in Main profile

(1) Assumed service requirement

・bit rate: upper limit 10 Mbit/s

・video format: SQVGA, 525QSIF, QCIF, QVGA, 525SIF, CIF, 525HHR, VGA, 525SD

・frame rate: 5, 10, 12, 15, 24, 30Hz (each shall be 1000/1001 times frequency)

But frame skip shall not be limited.

・Picture aspect ratio: 4:3, 16:9

(2) Operation level

Any of level 1, 1.1, 1.2, 1.3, 2, 2.1, 2.2, and 3 is operated. Level shall be selected according to coding format.

<Blank Page>

Attachment 4: Operational Guidelines for 3D video services by frame compatible system

Chapter 1: General Terms

1.1 Objective

The purpose of these operational guidelines is to present technical recommendations in practical operations concerning to video signals and video coding system for 3D video service by frame compatible system.

1.2 Scope

These operating guidelines apply to 3D video services by frame compatible system operated in digital broadcasting that comply with the "Standard transmission system for digital broadcasting among standard television broadcasting and the like" (Ordinance).

1.3 References

1.3.1 Normative References

(1) Rec. ITU-T H.262|ISO/IEC 13818-2:2012: (hereinafter referred to as "MPEG-2 Video Standard")

(2) Rec. ITU-T H.264|ISO/IEC 14496-10:2012: Advanced video coding for generic audiovisual services (hereinafter referred to as "MPEG-4 AVC Standard")

1.4 Terminology

1.4.1 Definitions

(1) Binocular disparity system

The system to display three dimensional image by using binocular disparity which is one of the factors of three dimensional sight of human being; three dimensional sight is given by the method that pictures of two cameras corresponding to left eye and right eye are given to viewer's left eye and right eye respectively.

(2) Binocular vision system

3D system that viewpoint (shooting direction) is one, and two images which are given to left eye and right eye respectively are used in binocular disparity system.

(3) Frame compatible system

The system that two images for left eye and right eye of 3D image by binocular vision system are transmitted by composing to one picture frame for each frame.

(4) Side-by-side system

The system that the horizontal pixel number of two images for left eye and right eye by frame compatible system is sub-sampled to 1/2, and sub-sampled images are composed to one picture frame by arranging at left and right of one frame.

Chapter 2: Frame compatible 3D video format

2.1 Structure of side-by-side 3D video format

Left image on video frame shall be for left eye, and right image shall be for right eye.

The boundary of two images composed to one picture shall be in the position at 1/2 pixels of the number of active pixels per 1 line.

Left and right images shall be down-sampled at the same phase in all lines of left and right images.

Fig. 2-1 shows the method of composing left and right images to one frame, and Fig. 2-2 shows the method of reconstructing left and right images.

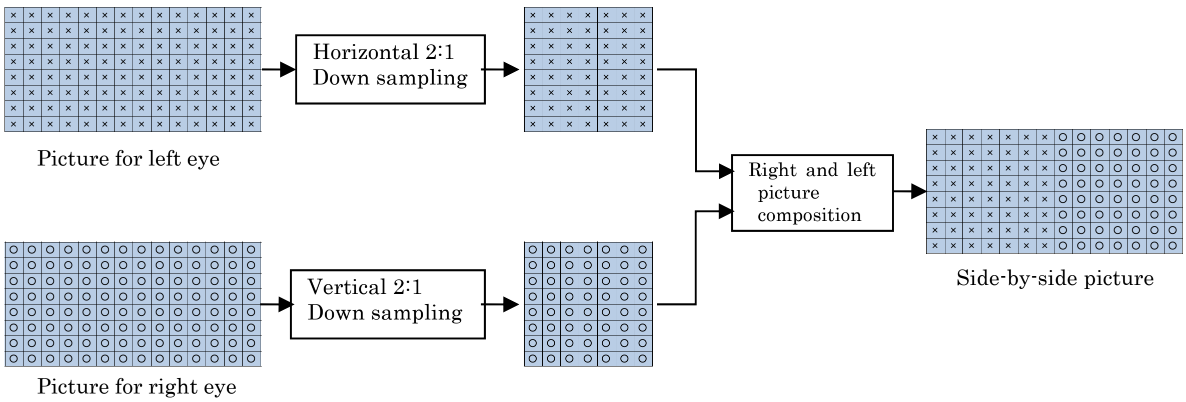


Fig. 2-1: Method of composing side-by-side 3D video

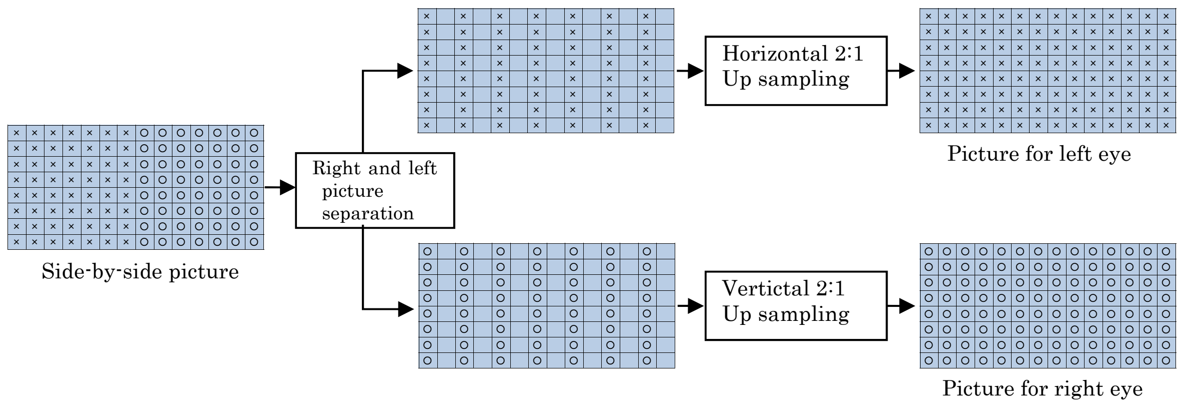


Fig. 2-2: Method of reconstructing side-by-side 3D video

2.1.1 1080/60/I

1080/60/I side-by-side 3D video format is shown in Table 2-1.

Table 2-1: 1080/60/I side-by-side 3D video format

Video parameters		Left eye image	Right eye image
Number of lines		1125	
Number of active lines		1080	
Scanning system		Interlaced	
Frame frequency		30/1.001Hz	
Field frequency		60/1.001Hz	
Picture aspect ratio		16:9	
Horizontal frequency		33.750/1.001 kHz	
Sampling frequency	Luminance signal	74.25/1.001 MHz	
	Color difference signals	37.125/1.001 MHz	
Number of samples per line	Luminance signal	1100	1100
	Color difference signals	550	550
Number of samples per active line	Luminance signal	960	960
	Color difference signals	480	480

Chapter 3: 3D identification information in MPEG-2 Video

3.1 Frame packing arrangement data

In order to identify 2D/3D and 3D video format, according to the provision in MPEG-2 Video Standard Annex L, in addition to the provision of this standard Part 1, Attachment 1, frame_packing_arrangement_data() is arranged in user_data of extensions_and_user_data(2) succeeding to picture_header() and picture_coding_extension(). When 3D video service is broadcasted, every video frame must include frame_packing_arrangement_data(). If there is no frame_packing_arrangement_data(), the service is judged to be 2D video.

Table 3-1: Arrangement of frame_packing_arrangement_data() in user data

Data structure	Bit number	Bit sequence notation
user_data( ) {		
user_data_start_code	32	bslbf
while(nextbits() != '0000 0000 0000 0000 0000 0001'){		
frame_packing_arrangement_data( )		
}		
next_start_code()		
}		

user_data_start_code: fixed value “0x000001B2”

As for frame_packing_arrangement_data(), only one is arranged in user_data().

Table 3-2: frame_packing_arrangement_data()

Data structure	Bit number	Bit sequence notation
frame_packing_arrangement_data( ){		
frame_packing_user_data_identifier	32	bslbf
remaining_data_length	8	uimsbf
reserved_bit	1	uimsbf
arrangement_type	7	bslbf
reserved_data	16	bslbf
for (i = 3; i < remaining_data_length; i++)		
additional_reserved_data_byte	8	bslbf
}		

frame_packing_user_data_identifier: This shall be fixed value “0x4a503344”

remaining_data_length: This represents byte length after main field. It shall be fixed value “3”.

arrangement_type (3D video format identification): This is 7 bit field, and represents the kind of 3D video format according to Table 3-3.

Table 3-3: 3D video format identification

arrangement_type	explanation
0000011	3D side-by-side (See Figure 2-1, Figure 2-2)
0000100	Reserved
0001000	2D video
Other than the above	Reserved

In case of changing 3D and 2D, it is recommended that the frame in which Picture data() is changed agree with the frame which arrangement_type is changed.

reserved_data: This shall be fixed value “0x04FF”

additional_reserved_data_byte: This is not used in main field.

Chapter 4: 3D identification information in MPEG-4 AVC

4.1 Frame Packing Arrangement SEI

In order to identify 2D/3D video, according to the provision in MPEG-4 AVC Standard Annex D, in addition to the provision of this standard Part 1, Attachment 2, 3.7, Frame Packing Arrangement SEI is operated in access unit. The order of NAL unit and SEI message which compose access unit shall be as shown in Fig. 4-1 in the head of GOP and other access unit, and NAL unit and SEI massage which are not described are not operated.

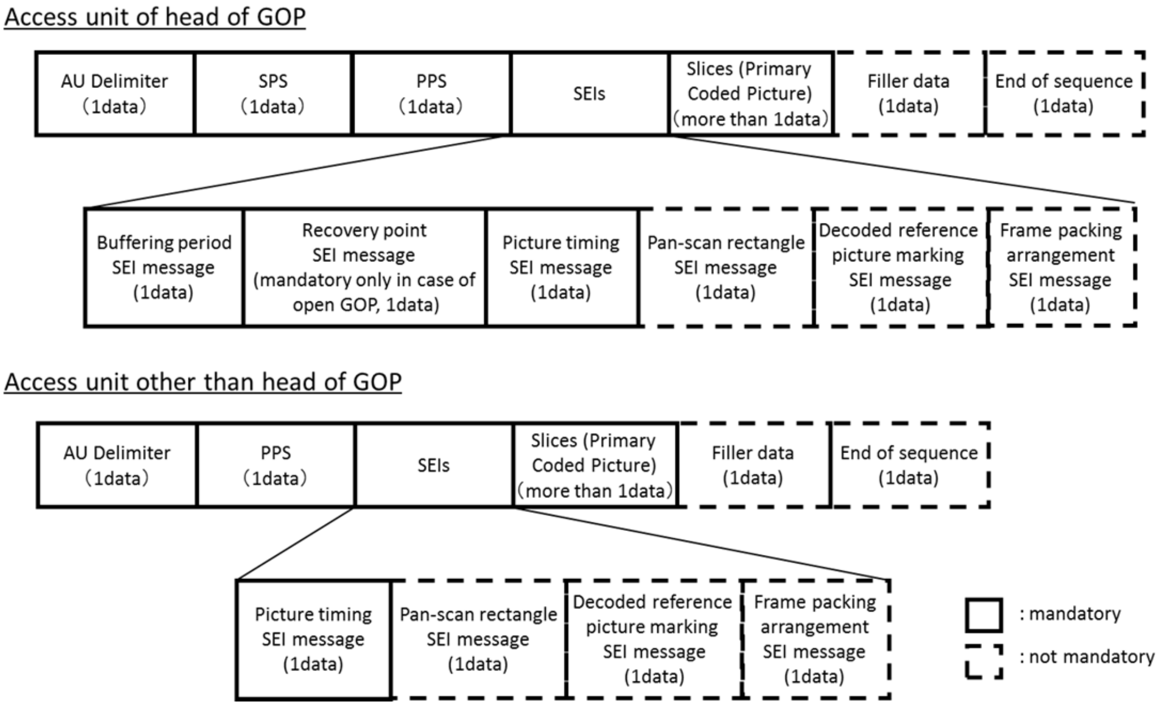


Fig. 4-1: Data structure of Access unit

Operation of Frame Packing Arrangement SEI is mandatory in access unit where frame compatible 3D video system is arranged. Operations of Frame Packing Arrangement SEI in each field are shown in Table 4-2.

Table 4-2: Operations of Frame Packing Arrangement SEI

Syntax element	Operation	Remarks
frame_packing_arrangement_id	0	0 fixed
frame_packing_arrangement_cancel_flag	0 or 1	0: Field that continues to the following exists. 1: Field that continues to the following does not exist. (note 1)
frame_packing_arrangement_type	3	3: 3D side-by-side system (See Fig. 2-1, Fig. 2-2)
quincunx_sampling_flag	0	0: quincunx sampling is not operated. (See Fig 2-1, Fig. 2-2)
content_interpretation_type	1	1: Left half of image is for left eye.
spatial_flipping_flag	0	0: Neither left nor right half of image is reversed.
frame0_flipped_flag	0	0 fixed
field_views_flag	0	0 fixed
current_frame_is_frame0_flag	0	0 fixed
frame0_self_contained_flag	0	0 : There is a possibility for prediction that the image for left eye refers to the image for right eye.
frame1_self_contained_flag	0	0 : There is a possibility for prediction that the image for right eye refers to the image for left eye.
frame0_grid_position_x	0	0 fixed (See Fig. 2-1, Fig. 2-2)
frame0_grid_position_y	0	0 fixed (See Fig. 2-1, Fig. 2-2)
frame1_grid_position_x	0	0 fixed (See Fig. 2-1, Fig. 2-2)
frame1_grid_position_y	0	0 fixed (See Fig. 2-1, Fig. 2-2)
frame_packing_arrangement_reserved_byte	0	0 fixed
frame_packing_arrangement_repetition_period	0	0: This SEI is valid in only concerned frame.
frame_packing_arrangement_extension_flag	0	0 fixed

(Note 1) frame_packing_arrangement_cancel_flag shall be "0" for 3D video, and "1" for 2D video. When changing between 3D and 2D, it is recommended that the frame where slice data is changed and the frame where frame_packing_arrangement_cancel_flag is changed will be made to be agreed with.

Chapter 5: Attention to operating 3D identification information

Pay attention not to occur frequent changes between 2D/3D in operating 3D identification signal for the following reasons.

(1) When this signal is used for control of changing video signal in the receiver, it may happen to occur that it takes several seconds to change.

(2) In case of 3D display by binocular vision system, according to changing between 2D/3D video, the viewer has to change viewing circumstances such as putting on or taking off glasses and so on.

<Blank Page>

Attachment 5: Operational Guidelines for HEVC Standard on television services

Chapter 1: General Terms

1.1 Objective

The purpose of these operational guidelines is to present technical recommendations of HEVC Standard in the practical operation concerning to video signals and video coding systems in digital television services.

1.2 Scope

These operational guideline apply to video signals using HEVC Standard among the video signals in television service that comply with the "Standard transmission system for digital broadcasting among standard television broadcasting and the like" (Ordinance).

1.3 References

1.3.1 Normative References

(1) Rec. ITU-T H.265 (2013) | ISO/IEC 23008-2:2013 - Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 2: High efficiency video coding (hereinafter referred to as "HEVC Standard")

(2) Rec. ITU-T H.222.0 (2012)| ISO/IEC 13818-1:2013 - Information technology -- Generic coding of moving pictures and associated audio information: Systems (including AMD1 to AMD4) (hereinafter referred to as "MPEG-2 System Standard")

(3) ISO/IEC 23008-1:2014 - Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 1: MPEG media transport (MMT) (hereinafter referred to as "MMT Standard")

(4) ARIB STD-B60: "Media transport system in digital broadcast"

1.4 Terminology

1.4.1 Abbreviations in HEVC Standard

AU	Access Unit
AUD	AU Delimiter
BLA	Broken Link Access
CABAC	Context-based Adaptive Binary Arithmetic Coding
CB	Coding Block

CBR	Constant Bit Rate
CRA	Clean Random Access
CPB	Coded Picture Buffer
CTB	Coding Tree Block
CTU	Coding Tree Unit
CU	Coding Unit
CVS	Coded Video Sequence
DCT	Discrete Cosine Transform
DPB	Decoded Picture Buffer
DST	Discrete Sine Transform
EOB	End of Bitstream
EOS	End of Sequence
HRD	Hypothetical Reference Decoder
IDR	Instantaneous Decoding Refresh
IRAP	Intra Random Access Point
LP	Leading Picture
NAL	Network Abstraction Layer
NALU	NAL Unit
PB	Prediction Block
POC	Picture Order Count
PPS	Picture Parameter Set
PU	Prediction Unit
RADL	Random Access Decodable Leading
RASL	Random Access Skipped Leading
RBSP	Raw Byte Sequence Payload
RPS	Reference Picture Set
SAO	Sample Adaptive Offset
SEI	Supplemental Enhancement Information
SOP	Structure of Pictures
SPS	Sequence Parameter Set
STSA	Step-wise Temporal Sub-layer Access
TB	Transform Block
TP	Trailing Picture
TSA	Temporal Sub-layer Access
TU	Transform Unit

VBR	Variable Bit Rate
VCL	Video Coding Layer
VPS	Video Parameter Set
VUI	Video Usability Information
WPP	Wavefront Parallel Processing

1.4.2 Abbreviations in MPEG-2 Systems Standard

CA	Conditional Access
CAT	Conditional Access Table
DTS	Decoding Time-Stamp
ECM	Entitlement Control Message
EMM	Entitlement Management Message
ES	Elementary Stream
GOP	Group of Pictures
NIT	Network Information Table
PAT	Program Association Table
PES	Packetized Elementary Stream
PID	Packet Identifier
PMT	Program Map Table
PSI	Program Specific Information
PTS	Presentation Time-Stamp
TS	Transport Stream
TMCC	Transmission & Multiplexing Configuration Control

1.4.3 Abbreviations in MMT Standard

AIT	Application Information Table
AMT	Address Map Table
AL-FEC	Application Layer Forward Error Correction
BIT	Broadcaster Information Table
CA	Conditional Access
CDT	Common Data Table
CRI	Clock Relation Information
CRID	Content Reference Identifier
DCI	Device Capability Information
ECM	Entitlement Control Message

EIT	Event Information Table
EMM	Entitlement Management Control
GFD	Generic File Delivery
HRBM	Hypothetical Receiver Buffer Model
LCT	Layout Configuration Table
LDT	Linked Description Table
MFU	Media Fragment Unit
MMT	MPEG Media Transport
MMTP	MMT Protocol
MPI	Media Presentation Information
MPT	MMT Package Table
MPU	Media Processing Unit
NIT	Network Information Table
PA	Package Access
PLT	Package List Table
SDT	Service Description Table
SDTT	Software Download Trigger
TLV	Type Length Value
URL	Uniform Resource Locator

Chapter 2: Summary of HEVC Standard

HEVC Standard is the standard that ITU-T SG16 and ISO/IEC JTC1/SC29/WG11 (MPEG) have developed in cooperation as next generation coding system for motion picture by which higher effective compression can be realized than the existing standard such as AVC standard (Rec. ITU-T H.264 | ISO/IEC 14496-10) and so on.

The following key technologies are introduced in HEVC Standard.

- variable block size (16×16 ~ 64×64 CTU, 8×8 ~ 64×64 CU)
- asymmetrical inter-PU partitioning (2N×nU, 2N×nD, nL×2N, nR×2N)
- orthogonal transform (4×4 ~ 32×32 DCT, 4×4 DST)
- intra-prediction (35 modes, reference pixel filter processing, prediction coding by prediction mode)
- inter-prediction (4×8/8×4 ~ 64×64 PU, 1/4 pixel precision motion vector, improved interpolation filter, prediction coding of motion vector)
- entropy coding (High throughput CABAC)
- in-loop filter (Deblocking filter, SAO)

It is said that a big reduction in amount of generated code becomes possible as a whole by these key technologies, and improvement of twice coding efficiency is realized comparing with AVC standard.

Also, in order to secure inter-operability between encoders and decoders, coding data is restricted for profile and level in HEVC Standard as in AVC Standard. As for profile, coding tool and processing range, etc. are restricted. And as for level, number of pixels to be handled, bitrate, frame frequency, etc. are restricted.

<Blank Page>

Chapter 3: Video coding format

3.1 Parameters

Parameters of video coding format shall be as shown in Table 3-1.

Table 3-1: Video coding format

Parameter		1080/60/I	1080/60/P	2160/60/P	2160/120/P	4320/60/P	4320/120/P
picture aspect ratio		16:9					
number of active samples per line		1,920		3,840		7,680	
number of active lines per frame		1,080		2,160		4,320	
coding sampling structure		Y′, C′B, C′R (non-constant luminance)					
		4:2:0					
pixel aspect ratio		1:1 (square pixel)					
frame frequency [Hz]		30/1.001, 30	60/1.001, 60	60/1.001, 60	120/1.001, 120	60/1.001, 60	120/1.001, 120
field frequency [Hz]		60/1.001, 60	—				
scanning		interlaced	progressive				
SDR -TV	bit number of pixel	8-bit, 10-bit (10-bit in case that colorimetry is based on ITU-R BT.2020)		10-bit			
	colorimetry, transfer function	Rec. ITU-R BT.709, IEC 61966-2-4(xvYCC), Rec. ITU-R BT.2020		Rec. ITU-R BT.2020			
HDR -TV	bit number of pixel	10-bit					
	colorimetry	Rec. ITU-R BT.2020					
	transfer function	Rec. ITU-R BT.2100 (HLG or PQ)					

(Description)

These parameters correspond to video signal based on ARIB STD-B56 and S-001C.

HDR-TV corresponds to the transfer function of HLG or PQ based on Rec. ITU-R BT.2100.

It is assumed that HDTV program is served by down convert from HEVC signal, making use of wide color gamut video in ARIB STD-B56. In case that colorimetry is based on Rec. ITU-R BT.2020, the number of bits for pixel shall be 10-bit.

4:4:4 and 4:2:2 for sampling structure are defined besides 4:2:0, and 12 bit for bit number of video sample is defined besides 10 bit in ARIB STD-B56. But considering sampling structure and bit number of video sample defined in HEVC Standard, video coding format is provided.

Chapter 4: Restrictions on video coding parameters

4.1 Profile and Level

Profile, Level and related condition for coding shall be as shown in Table 4-1.

Table 4-1: Profile and Level

Parameters	1080/60/I	1080/60/P	2160/60/P	2160/120/P	4320/60/P	4320/120/P
Video coding system	Rec. ITU-T H.265 | ISO/IEC 23008-2					
Profile	Main Note 1 Main10 Note 2		Main10			
Level	4.1 Note 3	4.1	5.1	5.2	6.1	6.2
Tier	Main Tier					
upper limit of bitrate [Mbps] Note 4	22	22	44	55	132	165
Temporal scalable coding Note 5	-	-	-	used	-	used
Chroma format	Y'C'BC'R 4:2:0					

Note 1: In case that bit number of coded pixel is 8 bit.

Note 2: In case that bit number of coded pixel is 10 bit.

Note 3: Considering the result of evaluation experiments, and in order to make upper limit of bitrate higher, level4.0 (upper limit of bitrate: 13.2Mbps) is not used but level4.1 shall be used in which upper limit of bitrate is 22Mbps.

Note 4: upper limit of bitrate in NAL level provided in HEVC Standard except for 2160/120/P and 4320/120/P

Note 5: scalable coding by which a part of 2160/120/P or 4320/120/P bitstream (sub-bitstream) can be decoded in the receiver which can deal with 2160/60/P or 4320/60/P.

(Description)

These are based on the provision of HEVC which is the latest international standard.

Extensibility of frame frequency is considered.

4.2 Syntax

The values shown in from Table 4-2 to Table 4-11 shall be used.

Table 4-2: NAL Unit Header

Syntax element	Value	Remarks
nuh_layer_id	0	0 fixed in HEVC Standard
nuh_temporal_id_plus1	See Table 4-11	

Table 4-3: Profile, Tier and Level

Syntax element	Value	Remarks
general_profile_space	0	0 fixed in HEVC Standard
general_tier_flag	0	Main tier
general_profile_idc	1, 2	1: Main Profile 2: Main10 Profile (Note) See Table 4-1 about combination of video format.
general_profile_compatibility_flag[j] (j = [0, 31])	0, 1	Description of compatible Profile In Main Profile, this is 1 only when j is 1 or 2, otherwise 0. In Main10 Profile, this is 1 only when j is 2, otherwise 0.
general_progressive_source_flag	0, 1	0: 1080/I 1: other than 1080/I
general_interlaced_source_flag	0, 1	0: other than 1080/I 1: 1080/I
general_frame_only_constraint_flag	0, 1	0: 1080/I 1: other than 1080/I
general_level_idc	123, 153, 156, 183, 186	123: Level 4.1 153: Level 5.1 156: Level 5.2 183: Level 6.1 186: Level 6.2 (Note) See Table 4-1 about combination of video format.

Table 4-4: Video Parameter Set, VPS

Syntax element	Value	Remarks
vps_max_layers_minus1	0	0 fixed in HEVC Standard
vps_max_sub_layers_minus1	See the right.	Identical value with sps_max_sub_layers_minus1
profile_tier_level()	See the right.	Identical value with profile_tier_level() of SPS
vps_sub_layer_ordering_info_present_flag	0	The values of vps_max_dec_pic_buffering_minus1, vps_max_num_reorder_pics, vps_max_latency_increase_plus1 in sub-layer are not described.
vps_max_layer_id	0	0 fixed in HEVC Standard
vps_num_layer_sets_minus1	0	0 fixed in HEVC Standard
vps_timing_info_present_flag	0	Timing information is described in VUI, and description is omitted in VPS.

Table 4-5: Sequence Parameter Set, SPS

Syntax element	Value	Remarks
chroma_format_idc	1	1: “4:2:0”
pic_width_in_luma_samples pic_height_in_luma_samples	See Table 4-8	MinCbSizeY(=8) times
conformance_window_flag conf_win_left_offset conf_win_right_offset conf_win_top_offset conf_win_bottom_offset	See Table 4-8	
bit_depth_luma_minus8 bit_depth_chroma_minus8	0, 2	0: Main Profile (8-bit) 2: Main10 Profile (10-bit)
log2_min_luma_coding_block_size_minus3	0	MinCbSizeY = 8
log2_diff_max_min_luma_coding_block_size	2, 3	2: CtbSizeY = 32 3: CtbSizeY = 64
log2_min_transform_block_size_minus2	0	Log2MinTrafoSize = 2 (4x4)
log2_diff_max_min_transform_block_size	3	Log2MaxTrafoSize = 5 (32x32)
vui_parameters_present_flag	1	VUI shall always be described.
vui_parameters( )	See Table 4-6	

Table 4-6: VUI parameters

Syntax element	Value	Remarks
aspect_ratio_info_present_flag	1	aspect_ratio_idc is described
aspect_ratio_idc	1	1:1 (“square pixel”)
video_signal_type_present_flag	1	Video information is described.
video_format	0	Component
video_full_range_flag	0	based on the provision of luminance signal and color difference signals
colour_description_present_flag	1	Color description information is described.
colour_primaries	1, 9	1: Rec. ITU-R BT.709, IEC 61966-2-4 (in case of HDTV (SDR-TV)) 9: Rec. ITU-R BT.2020 (in case of HDTV, UHDTV or HDR-TV)
transfer_characteristics	1, 11, 14, 16, 18	1: Rec. ITU-R BT.709 (in case of HDTV (SDR-TV) conventional color gamut) 11: IEC 61966-2-4 (in case of HDTV wide color gamut) 14: Rec. ITU-R BT.2020, 10-bit (in case of UHDTV (SDR-TV)) 16: Rec. ITU-R BT.2100 PQ (in case of HDR-TV) 18: Rec. ITU-R BT.2100 HLG (in case of HDR-TV)
matrix_coefficients	1, 9	1: Rec. ITU-R BT.709, IEC 61966-2-4 (in case of HDTV(SDR-TV)) 9: Rec. ITU-R BT.2020 Non constant luminance (in case of HDTV, UHDTV or HDR-TV)
chroma_loc_info_present_flag	1	4:2:0 color difference signals position is described
chroma_sample_loc_type_top_field chroma_sample_loc_type_bottom_field	0, 2	0: The middle of 2 luminance lines in vertical direction (in case of 1080/I) 2: This agrees with luminance line in vertical direction (other than 1080/I)
neutral_chroma_indication_flag	0	The values of color difference signals are coded values in the bitstream.
field_seq_flag	See Table 4-8	
frame_field_info_present_flag	1	pic_struct information is described.
vui_timing_info_present_flag	1	Timing information is described.
vui_num_units_in_tick vui_time_scale	See Table 4-9	Either “case 1” or “case 2” is applied.
vui_poc_proportional_to_timing_flag	1	
vui_num_ticks_poc_diff_one_minus1	See Table 4-10	
vui_hrd_parameters_present_flag	1	HRD information is described in VUI.

Table 4-7: HRD parameters

Syntax element	Value	Remarks
nal_hrd_parameters_present_flag	1	NAL HRD description is described.
sub_pic_hrd_params_present_flag	0	Only HRD description of AU unit
fixed_pic_rate_general_flag[i] (i = [0, sps_max_sub_layers_minus1])	0, 1	0 only in case that picture rate is different between CVS.
fixed_pic_rate_within_cvs_flag[i] (i = [0, sps_max_sub_layers_minus1])	0, 1	highest layer. (i = sps_max_sub_layers_minus1) shall always be 1 (picture rate is fixed in CVS).
elemental_duration_in_tc_minus1[i] (i = [0, sps_max_sub_layers_minus1])	See Table 4-10	

Table 4-8: Parameters which represent picture size

Video coding format	field_seq_flag	aspect_ratio_idc	general_progressive_source_flag	general_interlace_source_flag	pic_width_in_luma_samples	pic_height_in_luma_samples	conformance_window_flag	conf_win_left_offset	conf_win_right_offset	conf_win_top_offset	conf_win_bottom_offset
1080/I	0	1	0	1	1,920	1,088	1	0	0	0	4
	1	1	0	1	1,920	544	1	0	0	0	2
1080/P	0	1	1	0	1,920	1,080	0	0	0	0	0
	0	1	1	0	1,920	1,088	1	0	0	0	4
2160/P	0	1	1	0	3,840	2,160	0	0	0	0	0
4320/P	0	1	1	0	7,680	4,320	0	0	0	0	0

Note: 1080/P system permits both the case pic_height_in_luma_samples is 1,080 (without cropping) and the case it is 1,088 (with cropping).

Table 4-9: Time scale

	vui_time_scale		vui_num_units_in_tick	
	Case 1	Case 2	Case 1	Case 2
59.94/I	60,000	27,000,000	1,001	450,450
59.94/P	60,000		1,001	450,450
60.00/I	60,000		1,000	450,000
60.00/P	60,000		1,000	450,000
119.88/P	120,000		1,001	225,225
120.00/P	120,000		1,000	225,000

Table 4-10: Picture interval

Syntax element	Value	Remarks
vui_num_ticks_poc_diff_one_minus1	vui_num_units_in_tick – 1	
elemental_duration_in_tc_minus1[sps_max_sub_layers_minus1]	0	

Table 4-11: nuh_temporal_id_plus1

NAL unit	nuh_temporal_id_plus1	Explanations
AUD	[1, sps_max_sub_layers_minus1 + 1]	The value equivalent to Temporal ID of corresponding AU
VPS	1	HEVC Standard
SPS	1	HEVC Standard
PPS	[1, sps_max_sub_layers_minus1 + 1]	HEVC Standard
Picture timing SEI	[1, sps_max_sub_layers_minus1 + 1]	The value equivalent to Temporal ID of corresponding AU
SEI other than the above	1	
Slice segment	[1, sps_max_sub_layers_minus1 + 1]	The value equivalent to Temporal ID of AU
Filler	[1, sps_max_sub_layers_minus1 + 1]	The value of corresponding Slice segment
EOS	1	HEVC Standard

(Description)

Though VPS is mandatory in bitstream, the receiver may read VPS and discard it in HEVC Standard. In these operational guidelines, VPS shall be set to the same value as syntax for SPS.

POC increases by 1 every field picture (interval is 1,001/60,000 sec) in case of 59.94/I system.

On the other, POC increases 1 every frame picture in case of progressive system (59.94/P, interval is 1,001/60,000 sec).

The scope of general_progressive_source_flag and general_interlaced_source_flag shall be within CVS.

The value of elemental_duration_in_tc_minus1[0] can also deal with the case of 24Hz material. The number of repetition is considered accoding to the formula E-51 in HEVC Standard.

4.3 Bitstream structure

Bitstream is constructed as the structure shown in Fig. 4-1 and Table 4-12.

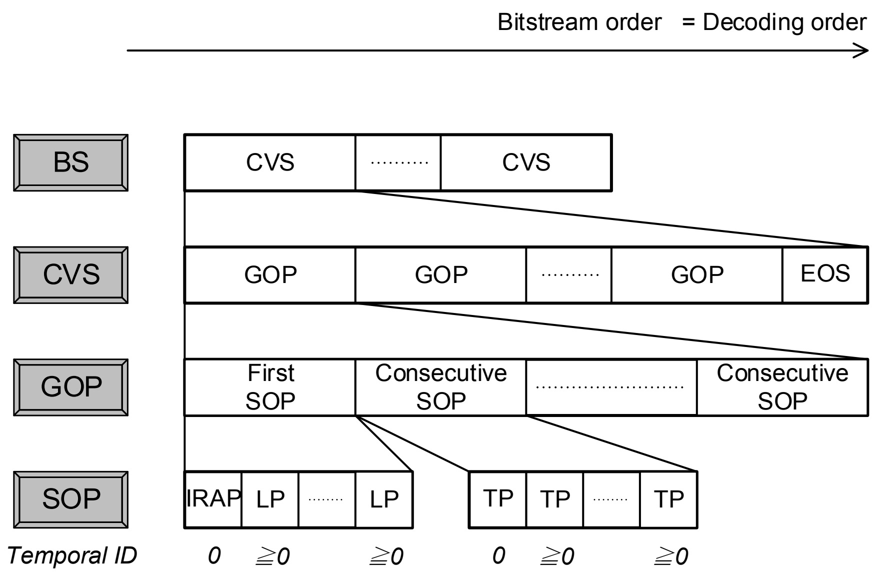


Fig. 4-1: Structure of HEVC bitstream

Table 4-12: Bitstream composition

Composition	Explanation	Restrictions in this guideline
Bitstream (BS)	One Video ES Composed of plural CVS	EOB NALU is not used as a termination of BS.
Coding sequence (CVS)	A set of AU referring to common SPS Composed of plural GOP	Final NALU of CVS shall be EOS.
Group of Pictures (GOP)	A set of AU that the AU which can be re-captured is arranged to the head. Composed of plural SOP	<u>In case of 1080/I</u> AU in the head of GOP shall be the AU which Recovery point SEI is added to, and whose slice_type is I. But only the head of GOP in the head of CVS shall be IRAP AU, and the structure of head SOP in the GOP including this IRAP AU is limited to L=0. <u>In case of other than 1080/I</u> AU in the head of GOP shall be IRAP AU. The AU which can be re-captured shall be only the head AU in GOP.
Picture structure (SOP)	A set of AU which consists of plural AU. The AU whose Temporal ID is 0 is arranged to the head, and another AU whose Temporal ID is more than 0.	Successor SOP in GOP consists of only TP. Only in case of no picture re-ordering, it is permitted that Temporal ID is 0 in AU except at the head of SOP.

(Description)

・BS, CVS

BS and CVS are identical with those defined in HEVC Standard.

・AU which can be re-captured

IRAP AU is defined as the AU which can be always re-captured in HEVC Standard. "AU which can be re-captured" means the AU that when the decoding process is begun from this AU, normal decode is guaranteed about this AU and successive AU in the display order.

In case that video coding format is other than 1080/I, by setting the head of GOP to IRAP AU, all the heads of GOP are arranged to the position where they can be re-captured.

In case that video coding format is 1080/I, according to HEVC Standard, there is a desirable case that IRAP AU is not used as "AU which can be re-captured." (See 4.4.3) In order to give AU in the head of GOP equivalent function as IRAP AU, AU in the head of GOP shall be the AU whose slice_type is I, and Recovery point SEI is added. In Recovery point SEI, the parameters are described which means this AU and all AU displayed after this AU are normally decoded. See Chapter 4, 4.6 about details.

・IRAP, LP, TP

According to HEVC Standard, from 0 to plural LP AU and from 0 to plural TP AU appear between IRAP AU and next IRAP AU.

LP AU is the AU which comes after IRAP AU in coding order, and comes before IRAP AU in display order. TP AU is the AU which comes after IRAP AU in coding order and display order.

When decoding starts from IRAP AU immediately before, LP AU is not always normally decoded. This is because there is a possibility that LP AU refers to the AU whose decoding order is before IRAP AU immediately before.

TP AU can be normally decoded even if decoding starts from IRAP AU immediately before. TP AU comes after LP AU in coding order. And it does not refer to the AU which is decoded before IRAP AU immediately before, and LP AU.

・GOP, SOP

GOP and SOP are definitions which are introduced for these operational guidelines.

The head of GOP is IRAP AU, and by starting decoding from the head of GOP, the receiver can normally decode all AU other than LP AU which appears immediately after.

SOP represents a unit which describes coding order and reference relationship of each AU in case of operating temporal scalable coding which is introduced in HEVC.

Also, in this provision, HandleCraAsBlaFlag described in HEVC Standard, 8.1 shall be reserved. (That is, there is no external means which sets up HandleCraAsBlaFlag.)

4.4 SOP structure

The provision of this clause applies to the coding of video format other than 120/P.

(Note) This provision restricts on decoding order, display order and Temporal ID of each AU. The reference relationship of AU may be set optional in the range which satisfies decoding order to be provided, and restriction on size relationship of Temporal ID. Reference relationship shown in the figures is just an example.

Each SOP in CVS is either SOP which applies to picture re-ordering or SOP which does not apply to picture re-ordering. SOP which applies to picture re-ordering and SOP which does not apply to picture re-ordering may be intermingled in the same CVS. But parameters related to DPB must correspond to SOP which applies to picture re-ordering.

4.4.1 In case of progressive scan video

4.4.1.1 In case of applying picture re-ordering

In the structure shown in Fig. 4-2, any of L=0, L=1, L=2, L=3 shall be used as SOP structure. But only CVS final SOP may take some other structure than these. For these structures (involving CVS final SOP), decoding order and display order as the followings are restricted on each AU.

- AU with the same Temporal ID must not succeed in decoding order. But AU whose Temporal ID is 0 or AU whose Temporal ID is maximum in each SOP are exempt.

  (Supplementary explanation)

  For example, when L=3, AU whose display order is 5 must not succeed immediately after AU whose display order is 1 in decoding order.

  Temporal ID of AU is the subtract 1 from nuh_temporal_id_plus1 of Slice segment NALU in AU.

- AU with the same Temporal ID must not reverse decoding order and display order.

  (Supplementary explanation)

  For example, when L=3, AU whose display order is 2 must not precede to AU whose display order is 0 in decoding order.

Only in case of the structure for L=1, AU whose Temporal ID is 1 may continue N times (N=[1,3]) in decoding order and display order.

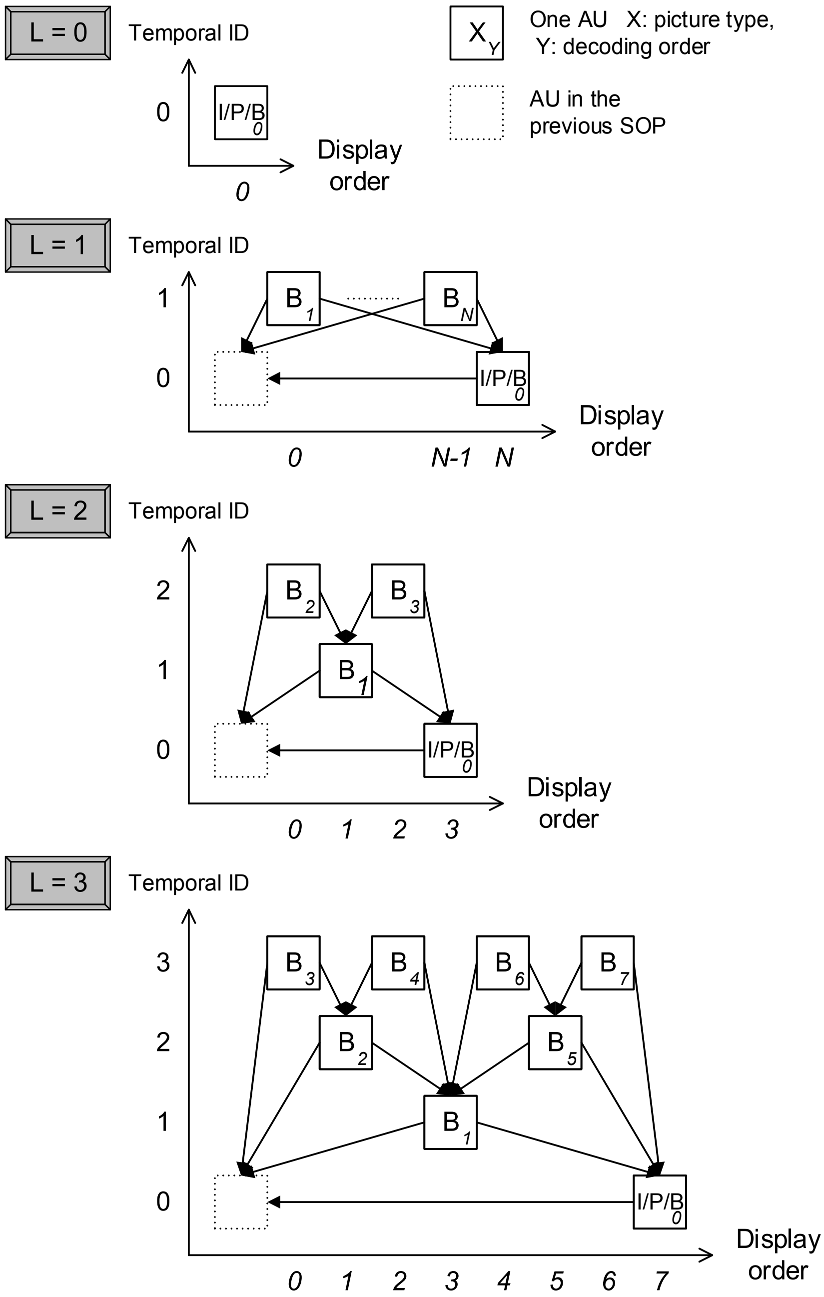


Fig. 4-2: Structure of SOP (in case of applying re-ordering)

Parameters related to DPB described in SPS shall be as shown in Table 4-13.

Table 4-13: Parameters related to DPB (in case of applying picture re-ordering)

SOP structure in CVS	sps_max_sub_layers_minus1	sps_sub_layer_ordering_info_present_flag Note 1	sps_max_dec_pic_buffering_minus1[i] Note 2 i = sps_max_sub_layers_minus1	sps_max_num_reorder_pics[i] i = sps_max_sub_layers_minus1	sps_max_latency_increase_plus1[i] i = sps_max_sub_layers_minus1
L=0	0	0	5	0	0
L=0, L=1	1	0	5	1	1 (N=1) 2 (N=2) 3 (N=3)
L=0, L=1, L=2	2	0	5	2	2
L=0, L=1, L=2, L=3	3	0	5	3	5

Note 1　Only the value in highest layer is described.

Note 2　At most 5 reference pictures are reserved in DPB, irrespective of SOP structure.

4.4.1.2 In case of not applying picture re-ordering

SOP structure shall be the structure shown in Fig. 4-3. L (maximum Temporal ID) takes the value between the range [0, 3]. When L is other than 0, AU of each Temporal ID appears only once in SOP.

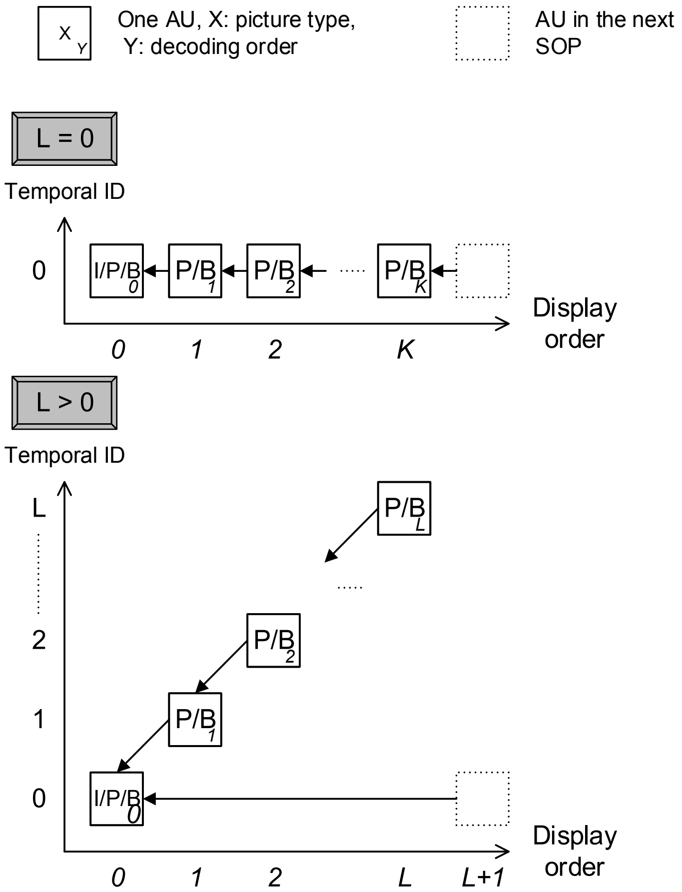


Fig. 4-3: Structure of SOP (in case of not applying re-ordering)

(Supplemental explanation)

Fast forward reproduction in the receiver can be realized easily by setting L to more than or equal to 1. For example, when L is 1, two times fast forward reproduction is realized by decoding and displaying only AU whose Temporal ID is 0.

Parameters related to DBS described in SPS shall be as shown in Table 4-14.

Table 4-14: Parameters related to DPB (in case of not applying picture re-ordering)

SOP structure in CVS	sps_max_sub_layers_minus1	sps_sub_layer_ordering_info_present_flag[Note 1]	sps_max_dec_pic_buffering_minus1[i] i = sps_max_sub_layers_minus1 [Note 2]	sps_max_num_reorder_pics[i] i = sps_max_sub_layers_minus1	sps_max_latency_increase_plus1[i] i = sps_max_sub_layers_minus1
L=0	0	0	5	0	0
L=0, L=1	1	0	5	0	0
L=0, L=1, L=2	2	0	5	0	0
L=0, L=1, L=2, L=3	3	0	5	0	0

Note 1: Only the value in the highest layer is described.

Note 2: Five reference pictures in maximum are reserved in DPB, not depending on SOP structure.

4.4.2 In case of interlaced scan video

Frame coding and field coding are able to be operated by changing them in CVS unit.

A pair of top field (first field) and bottom field (second field) which are successive in display order are called “field pair.”

4.4.2.1 In case of applying frame coding

When frame coding is applied, a field pair is coded as a frame picture. SOP structure shall be the same as the structure provided in Chapter4, 4.4.1.

4.4.2.2 In case of applying field coding and picture re-ordering

When field coding is applied and picture ordering is applied, in the structure shown in Fig. 4-4, any of L=0, L=1, L=2, L=3 shall be used as SOP structure. But only CVS final SOP may take some other structure than these. For these structures (involving CVS final SOP), decoding order and display order as the followings are restricted on each AU.

- Top field and bottom field in the same field pair are successive in decoding order.
- Field pair with the same Temporal ID must not succeed in decoding order. But the field pair whose Temporal ID is 0 or the field pair with maximum Temporal ID are exempt.
- Field pair with the same Temporal ID must not reverse decoding order and display order.

Only SOP in the head of CVS or SOP in bottom CVS takes the structure of L=0.

Only in the structure of L=1, field pair AU whose Temporal ID is 1 may continue N times (N=[1, 3]) in decoding order and display order.

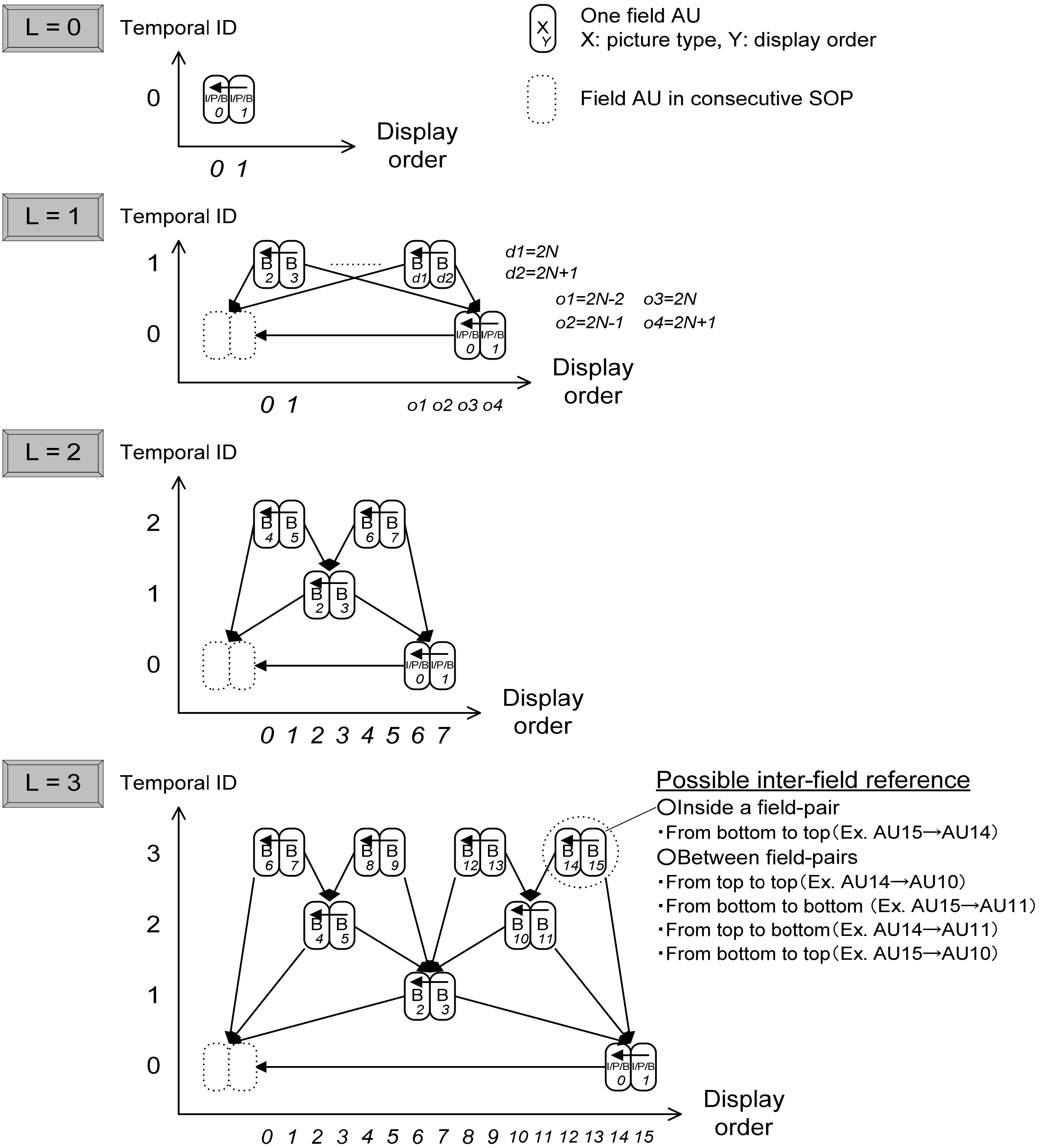


Fig. 4-4: SOP Structure (in case of applying field coding, picture re-ordering)

Moreover in CVS, only in the head SOP of head GOP in CVS (SOP structure L=0), head AU shall be IRAP AU. In the head SOP of other GOP, head AU shall be TP AU whose slice_type is I and to which Recovery point SEI and Buffering period SEI are added. AU except for head AU in

CVS shall be TSA AU in case that Temporal ID is top field other than 0, and shall be TP AU in other case. See Fig. 4-5.

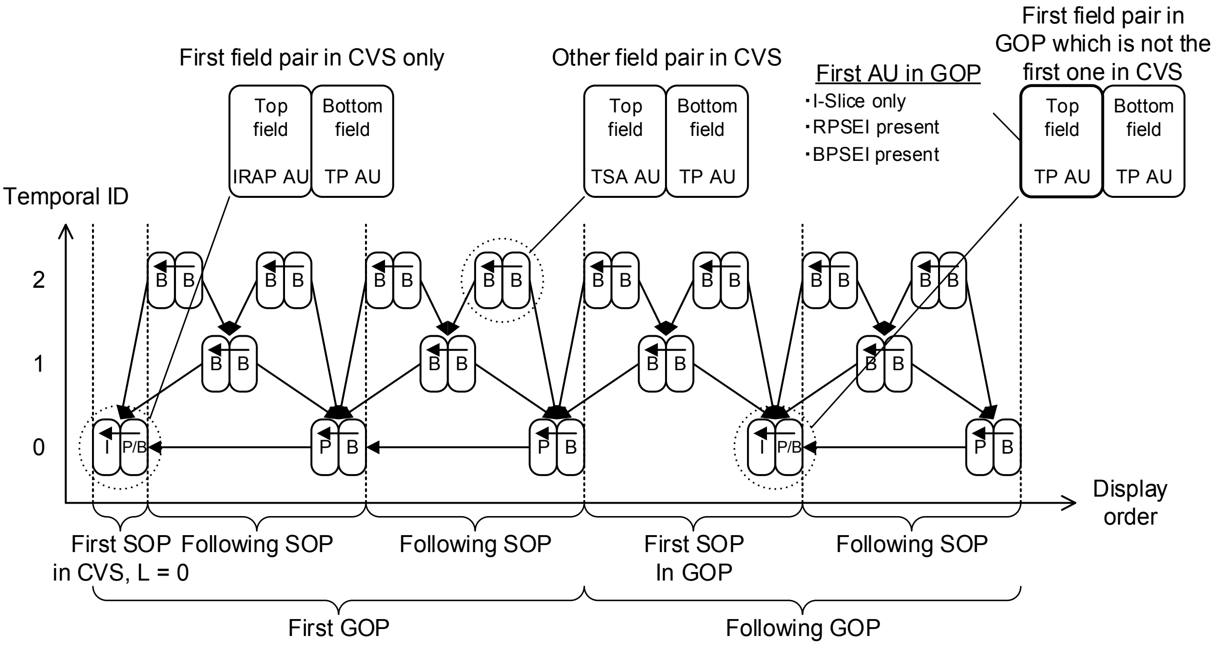


Fig. 4-5: SOP Structure in field coding and Restrictions on AU

Concerning restrictions of reference between AU, the value of sps_temporal_id_nesting_flag shall be 0 in case of field coding and when sps_max_sub_layers_minus1 is other than 0. Moreover, bottom field AU whose Temporal ID is more than 1 does not refer to AU which has the same Temporal ID except for top field in the same field pair.

Parameters related to DPB described in SPS shall be as shown in Table 4-15.

Table 4-15: Parameters related to DPB (in case of applying picture re-ordering)

SOP structure in CVS	sps_max_sub_layers_minus1	sps_sub_layer_ordering_info_present_flag[Note 1]	sps_max_dec_pic_buffering_minus1[i] i = sps_max_sub_layers_minus1 [Note 2]	sps_max_num_reorder_pics[i] i = sps_max_sub_layers_minus1	sps_max_latency_increase_plus1[i] i = sps_max_sub_layers_minus1
L=0	0	0	11	0	0
L=0, L=1	1	0	11	2	1 (N=1) 3 (N=2) 5 (N=3)
L=0, L=1, L=2	2	0	11	4	3
L=0, L=1, L=2, L=3	3	0	11	6	9

Note 1: Only the value in the highest layer is described.

Note 2: 11 reference pictures in maximum are reserved in DPB irrespective of SOP structure.

4.4.2.3 In case of applying field coding and not applying picture re-ordering

When field coding is applied and picture re-ordering is not applied, SOP structure shall be the structure shown in Fig. 4-6. L (maximum Temporal ID) takes the value in the interval [0, 3]. When L is more than or equal to 1, field pair of each Temporal ID appears only once.

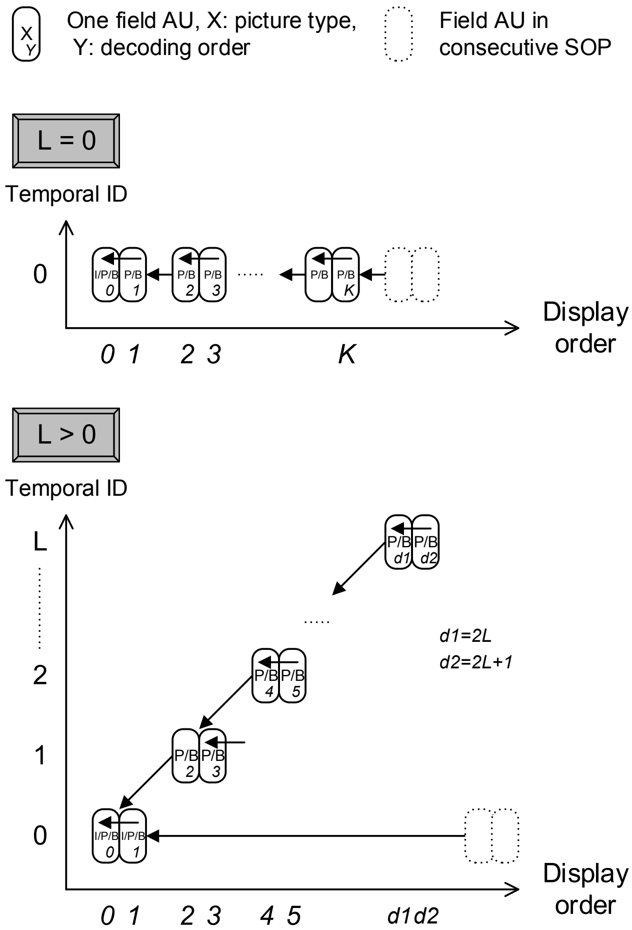


Fig. 4-6: SOP structure (in case of applying field coding and not applying picture re-ordering)

Parameters related to DBS which are described in SPS shall be as shown in Table 4-16.

Table 4-16: Parameters related to DPB (in case of not applying picture re-ordering)

SOP structure in CVS	sps_max_sub_layers_minus1	sps_sub_layer_ordering_info_present_flag Note 1	sps_max_dec_pic_buffering_minus1[i] i = sps_max_sub_layers_minus1 Note 2	sps_max_num_reorder_pics[i] i = sps_max_sub_layers_minus1	sps_max_latency_increase_plus1[i] i = sps_max_sub_layers_minus1
L=0	0	0	11	0	0
L=0, L=1	1	0	11	0	0
L=0, L=1, L=2	2	0	11	0	0
L=0, L=1, L=2, L=3	3	0	11	0	0

Note 1: Only the value of the highest layer is described.

Note 2: 11 reference pictures in maximum are reserved in DPB irrespective of SOP structure.

(Description)

・Items common to progressive scan video (4.4.1) and interlaced scan video (4.4.2)

By making the operation of Temporal ID whose value is more than or equal to 0 mandatory, description of reference picture becomes simple, and fast forward reproduction in the receiver becomes easy.

What is called "IPBB" structure and "IPPP" structure which are generally used in MPEG-2 are dealt with.

By restriction that all SOP structures shall be the same in CVS, in case of re-ordering (and no re-ordering, L>0), N times reproduction (N={2, 4, 8}) can be easily realized looking at only the value of Temporal ID. But it may not be applied in case of CVS boundary (there are the case that different SOP structure is used in order to make specified sequence length) and so on. It is possible to decide Temporal ID of AU to be read and discarded in order to make N times reproduction by referring to PTS value of the head AU in each SOP, so SOP structure in CVS shall be variable.

・Items proper for interlaced scan video (4.4.2)

The coding performance of frame coding and field coding varies according to degrees of motion in video. Sequence unit at frame and field adaptive coding (SAFF) which is permitted in HEVC provision shall be able to be used.

In order to equip receiver easily, when field coding is used, top field and bottom field in the same field are restricted so that they are successive in the decoding order.

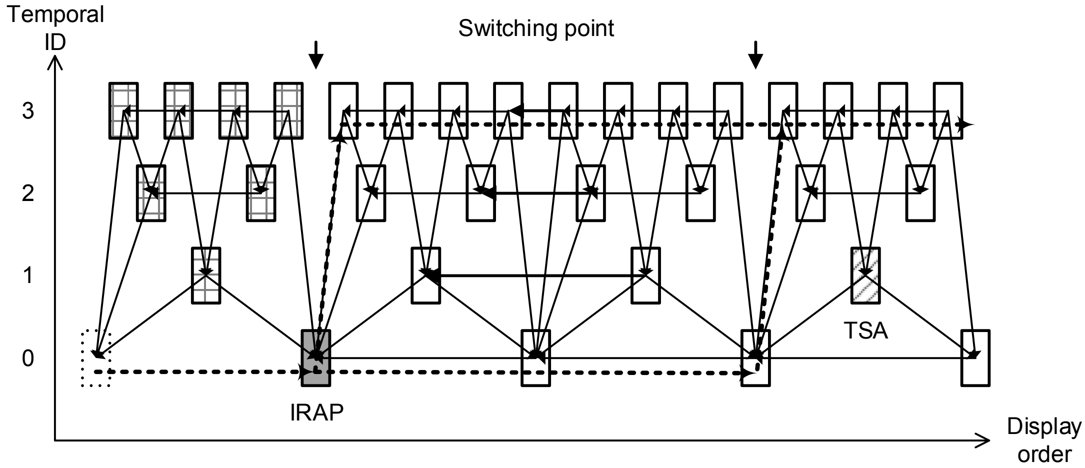
There is a restriction that when TP AU appears immediately after IRAP AU, LP AU must not be applied in HEVC Standard. In this guideline, when field coding is used, both fields in the field pair are always decoded successively. When the top field is IRAP AU, the bottom field will be certainly TP AU, so it is impossible to take other than L=0 as SOP structure. In order to avoid this problem, IRAP AU is not used except at the head AU of CVS. In order to make re-capturing possible, the head AU of GOP shall be intra-coding AU, and the position of re-capturing is expressed by adding Recovery point SEI.

There is a restriction that maximum Temporal ID is 6 in HEVC Standard. So, Temporal ID of the top field and Temporal ID of the bottom field in the same field pair shall be the same. Also, the value of sps_temporal_id_nesting_flag shall be 0 in case of field coding and in the case that sps_max_sub_layers_minus1 is other than 0. This is because that by setting this value to 1, all AU whose Temporal ID is more than or equal to 1 will be TSA AU, and as TSA AU is prohibited to refer between AU of the same Temporal ID, referring to top field from bottom field in the same field pair. But if the restriction is left just as they are, even AU whose Temporal ID is more than or equal to 1 can refer over SOP (adding refer to AU whose Temporal ID is 0). So AU whose Temporal ID is more than or equal to 1, referring between AU whose Temporal ID is the same is limited in the same field pair. Also, there is a restriction in HEVC Standard that when sps_max_sub_layers_minus1 is 0, sps_temporal_id_nesting_flag is set to 1.

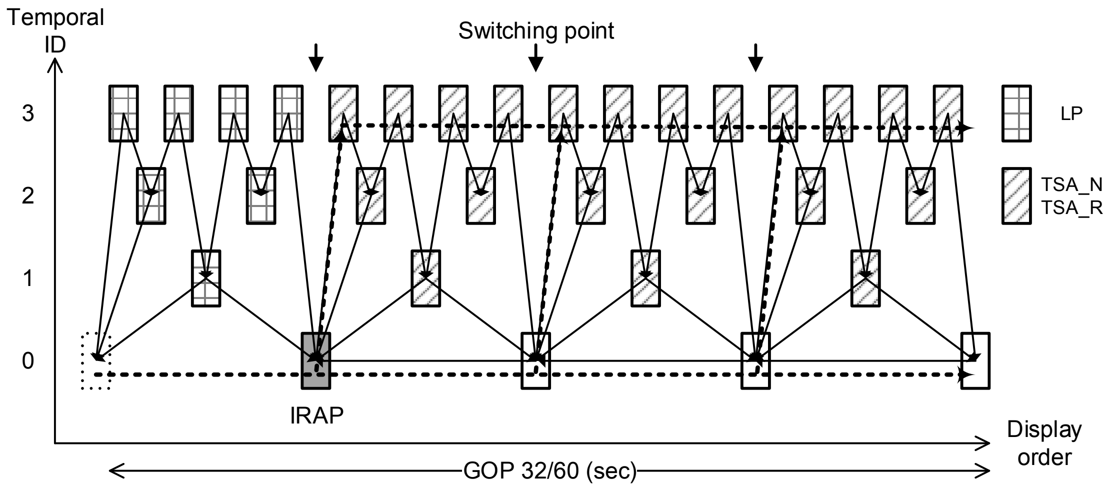
4.4.3 Temporal layer up-switching

This is carried out by setting sps_temporal_id_nesting_flag to 1, or by inserting TSA AU. (See Fig. 4-7)

STSA AU is not used. Also, in case of field coding and when sps_max_sub_layers_minus1 is other than 0, sps_temporal_id_nesting_flag is set to 0.



In the case sps_temporal_id_nesting_flag is equal to 0



In the case sps_temporal_id_nesting_flag is equal to 1

Fig. 4-7: Explanation for sps_temporal_id_nesting_flag

(Description)

When sps_temporal_id_nesting_flag is set to 0, change to normal reproduction in the interval of GOP at least will be possible. By using TSA AU, change to normal reproduction will be possible even in the halfway of GOP.

If sps_temporal_id_nesting_flag is set to 1, change as SOP unit will be possible. But the restriction is added that AU whose Temporal ID is more than or equal to 1 cannot be referred between SOP.

As STSA has a little merit in a practical use, and it makes implemenation of the receiver complicated, so STSA shall be prohibited to use.

4.5 Restrictions on VCL NALU in SOP

Restrictions are as shown in Table 4-17 and Table 4-18.

Table 4-17: Restrictions on VCL NALU (in case of other than 1080/I)

	Explanations	Remarks
In head SOP of head GOP of CVS	Head AU is set to IRAP AU. As NALU type (nal_unit_type), only the following value is used. – IDR (IDR_W_RADL, IDR_N_LP)	BLA (BLA_W_LP, BLA_W_RADL, BLA_N_LP) is prohibited.
	<u>In case of re-ordering</u> Other than head AU is set to LP AU. As nal_unit_type, only the following value is used. – RADL (RADL_N, RADL_R) <u>In case of no re-ordering</u> Other than head AU is set to TP AU. As nal_unit_type, only the following value is used. – TRAIL (TRAIL_N, TRAIL_R)	STSA (STSA_N, STSA_R) is prohibited.
In head SOP of successive GOP	Head AU is set to IRAP AU. As NALU type (nal_unit_type) of IRAP AU, only the following value is used. – CRA (CRA_NUT)	
	<u>In case of re-ordering</u> Other than head AU is set to LP AU. As nal_unit_type, only the following value is used. – RASL (RASL_N, RASL_R) <u>In case of no re-ordering</u> Other than head AU is set to TP AU. As nal_unit_type, only the following value is used. – TRAIL (TRAIL_N, TRAIL_R)	RADL (RADL_N, RADL_R) is prohibited. STSA (STSA_N, STSA_R) is prohibited.
In successive SOP	As nal_unit_type of TP AU, only the following values are used. – TRAIL (TRAIL_N, TRAIL_R) – TSA (TSA_N, TSA_R)	STSA (STSA_N, STSA_R) is prohibited.

Table 4-18: Restrictions on VCL NALU (in case of 1080/I)

	Explanations	Remarks
In head SOP of head GOP of CVS	Head AU is set to IRAP AU. As NALU type (nal_unit_type), only the following value is used. – IDR (IDR_N_LP)	BLA (BLA_W_LP, BLA_W_RADL, BLA_N_LP) is prohibited.
	Head SOP of the head GOP of CVS includes head AU (first field) and one successive AU (second field). Successive AU shall be TP AU. As nal_unit_type of successive AU (TP AU), only the following value is used. – TRAIL (TRAIL_N, TRAIL_R)	
In SOP other than mentioned above	Other than head SOP of the head GOP of CVS includes only TP AU. As nal_unit_type of TP AU, only the following values are used. – TRAIL (TRAIL_N, TRAIL_R) – TSA (TSA_N, TSA_R)	STSA (STSA_N, STSA_R) is prohibited.

(Description)

By prohibiting STSA which has low necessity of use and low frequency, equipping receiver will be simple.

There is a possibility that BLA appears by editing BS, but it is prohibited to use in this operational rule.

4.6 NALU in AU

Restrictions are shown in Table 4-19.

Table 4-19: Restrictions on NALU in AU

NALU	IRAP AU	Re-capturing AU (field coding)	Non IRAP AU
AUD	necessary	necessary	necessary
VPS	necessary	necessary	prohibited
SPS	necessary	necessary	prohibited
PPS	necessary	necessary	possible
Prefix SEI Note 1	The following SEI are necessary. – Buffering period SEI – Picture timing SEI	The following SEI are necessary. – Buffering period SEI – Picture timing SEI – Recovery point SEI	The following SEI is necessary. – Picture timing SEI
Slice segment	necessary All Slice segment in AU must have the same nal_unit_type.	necessary All Slice segment in AU must have the same nal_unit_type.	necessary All Slice segment in AU must have the same nal_unit_type.
Suffix SEI Note 2	prohibited	prohibited	prohibited
Filler	possible Note 3	possible Note 3	possible Note 3
EOS	possible Note 4	possible Note 4	possible Note 4
EOB	prohibited	prohibited	prohibited

Note 1: SEI other than the mentioned above must not appear in bitstream.

Note 2: If Suffix SEI appears in bitstream, the receiver must ignore the SEI.
Filler SEI is not used.

Note 3: In case of operating CBR, Filler NALU is used. And it is permitted to set cbr_flag to 1.

Note 4: Only the case of last AU in CVS

The following is restricted on transmission method of PPS.

— pps_pic_parameter_set_id must take the value in the range [0, 15].

The restrictions shown in Table 4-20 is set on Recovery point SEI parameters.

Table 4-20: Recovery point SEI parameters

Syntax element	Value	Explanations
recovery_poc_cnt	0	Normal decoding of AU which Recovery point SEI is added to and whose display order is after is assured.
exact_match_flag	1	Normal decoding of AU which Recovery point SEI is added to and whose display order is after is assured.
broken_link_flag	0	If decoding starts from immediately before IRAP AU, abnormal decoding will not occur.

(Description)

CBR operation is permitted. There are two cases of Filler insertion: use of Filler SEI, and use of Filler NALU, but the latter shall be possible.

EOB is prohibited to use because there are no reason to use particularly.

Adding method of PPS NALU shall apply to HEVC Standard. PPS which is added to one AU may be 0 or some. As there is an error resistant function of PPS in HEVC Standard, it is expected that the receiver decodes certainly. But in order to avoid to enlarge memories in the receiver unnecessarily, the range of pps_pic_parameter_set_id is limited.

4.7 Restrictions on delay time

Restrictions are shown in Table 4-21.

Table 4-21: Restrictions on delay time

Item	Restrictions
Interval of IRAP AU Interval of inserting RPSEI (in case of field coding)	Within 32/60 second in rule, maximum 1.0 second
CPB size	Within 1.0R [bit] (R is maximum bitrate)
CPB delay	AuNominalRemovalTime[0] must be less than 0.5 second
Maximum DPB size	sps_max_dec_pic_buffering_minus1 must be less than or equal to 5 in case of frame coding, and less than or equal to 11 in case of field coding.

(Description)

"Display delay" in these operational guidelines is defined as the difference (unit is frame time) between decoding time of AU which starts decoding and displayed time of AU which is displayed first (see Fig. 4-8). "Display delay" is equivalent to sps_max_num_reorder_pics regardless of CVS head SOP structure. (In case of progressive video, half value is equivalent.)

Fig. 4-8: Figure of supplemental explanation for display delay

4.8 Picture partitioning

Though picture partitioning is not mandatory in 1080/I, 1080/P and 2160/P, use of WPP in the scope of HEVC Standard is permitted. Slice is also possible to be used including using with WPP, but it must follow the restriction in Table 4-23.

(Supplemental explanation)

WPP means the provided movement in case that entropy_coding_sync_enabled_flag is 1. The value of entropy_coding_sync_flag_enabled_flag is fixed in CVS.

Picture partitioning into four slices shall be mandatory for 4320/P. The method of partitioning is as shown in fig. 4-9. Also, when partitioning into slices, partitioning by WPP or slice may be done in each slice.

Moreover the restrictions on parameters shown in Table 4-22 are applied.

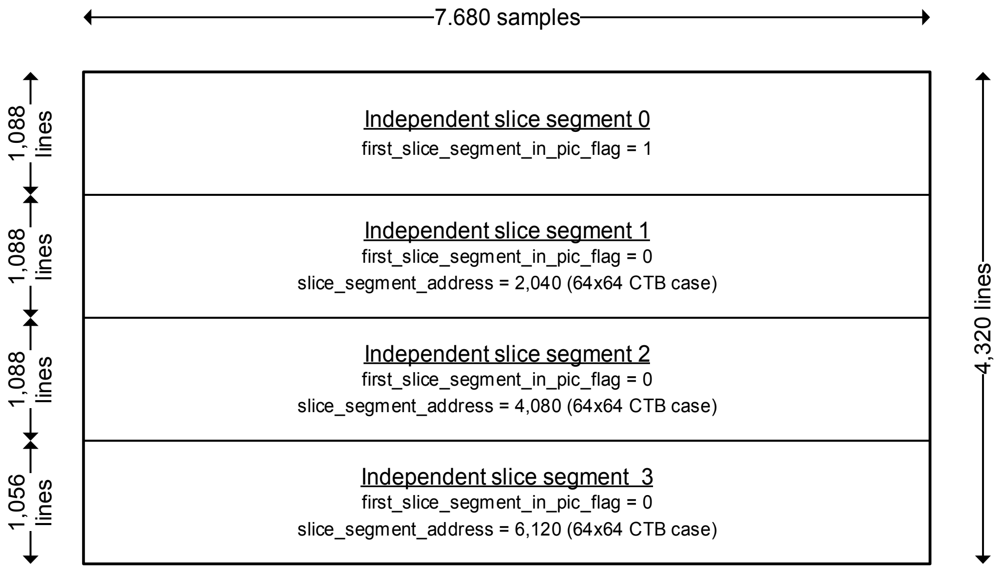


Fig. 4-9: Picture partitioning for 4320/P

Table 4-22: Restrictions on parameters about picture partitioning

Parameter	Restrictions
pic_width_in_luma_samples	7,680
pic_height_in_luma_samples	4,320
first_slice_segment_in_pic_flag slice_segment_address	The value described in Fig. 4-9
pps_loop_filter_across_slices_enabled_flag slice_loop_filter_across_slices_enabled_flag	1
Range of motion vector in vertical direction crossing over Slice boundary	See Fig. 4-10

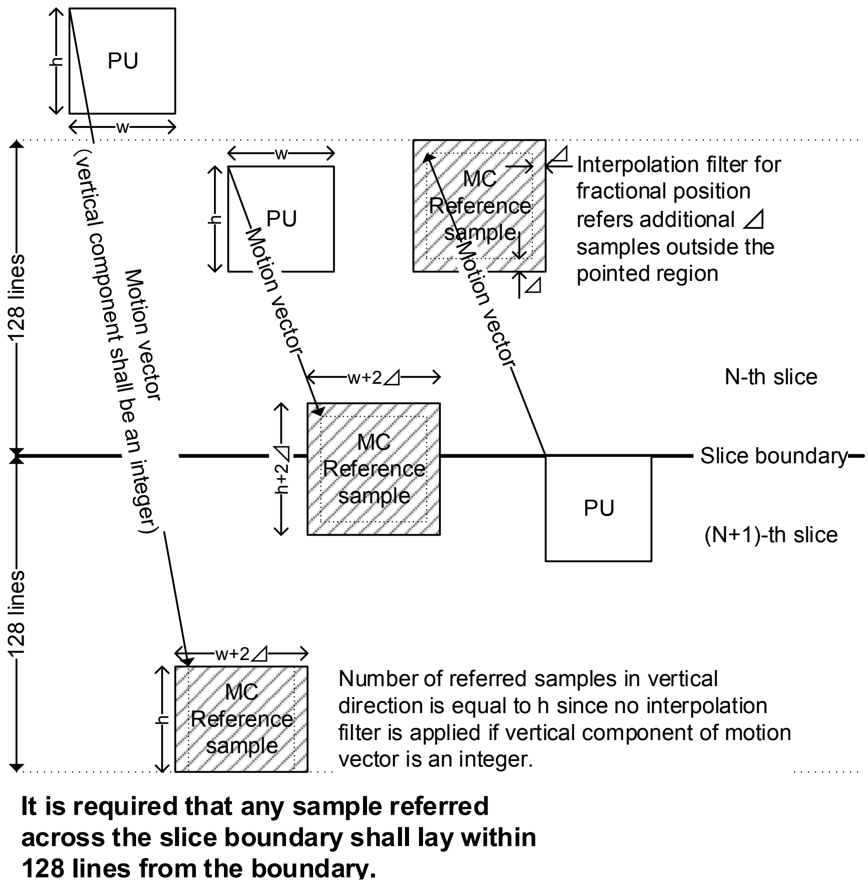


Fig. 4-10: Limitation of motion vector in picture partitioning for 4320/P

(Description)

Tile which is a tool for picture partitioning is not applied regardless of picture size. Because it is worried that picture quality will be decreased at panning by the MV restriction at the boundary of partition in case of 4320/P. In case of 1080/I, 1080/P and 2160/P, there are no MV restriction, but as slice can be used instead, Tile is not applied as 4320/P.

For the coding and decoding of 1080/I, 1080/P and 2160/P, picture partitioning shall not be mandatory because it is very likely to be realized by using 1 core.

For the coding and decoding of 4320/P, picture partitioning shall be mandatory because it is unlikely to be realized by using 1 core for the time being.

In order to avoid that the boundary deterioration is recognized when picture partitioning of 4320/P, filter in loop of slice shall be always available. Also, inter-prediction over slice shall be available. But upper limit of motion vector over boundary is set up in order to suppress bandwidth common to reference picture between cores to a practical value.

4.9 Various coding parameters

Restrictions are shown in Table 4-23.

Table 4-23: Restrictions on various coding parameters

Item	Syntax elements	Restrictions	Remarks
SPS	sps_seq_parameter_set_id seq_parameter_set_rbsp( )	0 Contents of SPS are the same in CVS.	As the contents of SPS in CVS become the same, the value of VUI also becomes the same.
Inter PU size	-	4x4 Inter PU is prohibited In 8x4/4x8 PU, PRED_BI is prohibited.	Restriction in HEVC Standard
Number of reference pictures	num_ref_idx_l0_active_minus1 num_ref_idx_l1_active_minus1	3 (frame coding) 7 (field coding)	Equivalent restriction to AVC restriction in STD-B32
Minimum compression rate	MinCR	See Table 4-24	
Minimum slice unit	-	1 CTU line	Note 1
Motion vector range	-	For 4320/P, the restriction shown in Table 4-22 is applied at the boundary of picture partitioning into four slices.	No restriction for 1080/I, 1080/P and 2160/P.
Dependent slice	dependent_slice_segments_enabled_flag	0	
Slice type	slice_type	All slices in the same AU must have the same value.	

Note 1: Minimum slice unit shall be 1 CTU line (it is equivalent to a rectangle area which is 1 CTU in height and the screen in width on the screen). In case of partitioning into four slices in the horizontal direction by using slice for 4320/P, it is also applied. (That is, in case of CTU whose size of luminance component CTB is 64x64, 120 CTU's is a minimum unit.)

Table 4-24: Restrictions on MinCR

Level	MinCR	Remarks
4	4	
4.1	4	
5	6	
5.1	8	
5.2	8	
6	8	
6.1	8	
6.2	8	Restriction shall to be stronger than HEVC Standard (6).

(Description)

Restriction of size for Inter PU shall follow HEVC Standard.

4.10 Temporal scalable coding

For 2160/120/P video and 4320/120/P video, temporal scalable coding is applied.

120/P bitstream in multiplex layer is transmitted by separating sub-bitstream (hereinafter called as 60/P sub-bitstream) based on Level 5.1 (2160/60/P) or Level 6.1 (4320/60/P) and the other part (hereinafter called as 120/P subset).

The receiver compliant with Level 5.1 (2160/60/P) or Level 6.1 (4320/60/P) can display the frame equivalent to 60Hz in 120/P bitstream by picking up only 60/P sub-bitstream and reproducing.

4.10.1 Summary

Temporal scalable coding is realized by the method provided in HEVC Standard.

After the receiver for 120Hz composes 60/P sub-bitstream and 120/P subset which are transmitted separately, it decodes as 120/P bitstream and outputs decoded video of 120Hz. On the other hand, in the receiver for 60Hz, multiplex decoding block extracts only 60/P sub-bitstream and decodes, and outputs decoded video of 60Hz. A decoding process in the receiver for 120Hz and the receiver for 60Hz is schematically shown in Fig. 4-11.

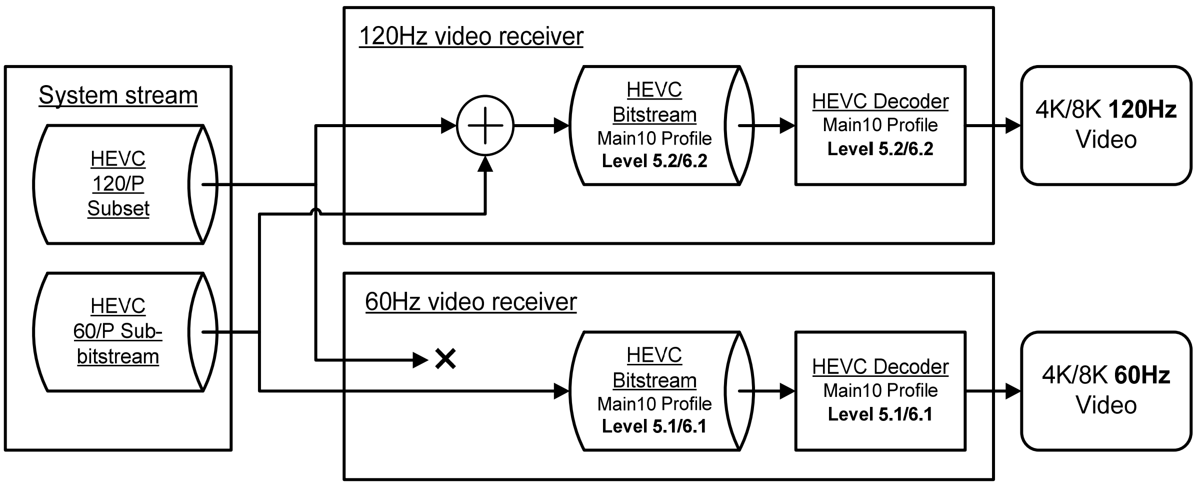


Fig. 4-11: Decoding process of 120Hz/60Hz temporal scalable coded bitstream

Structures of 120/P bitstream, 60/P sub-bitstream and 120/P subset are schematically shown in Fig. 4-12.

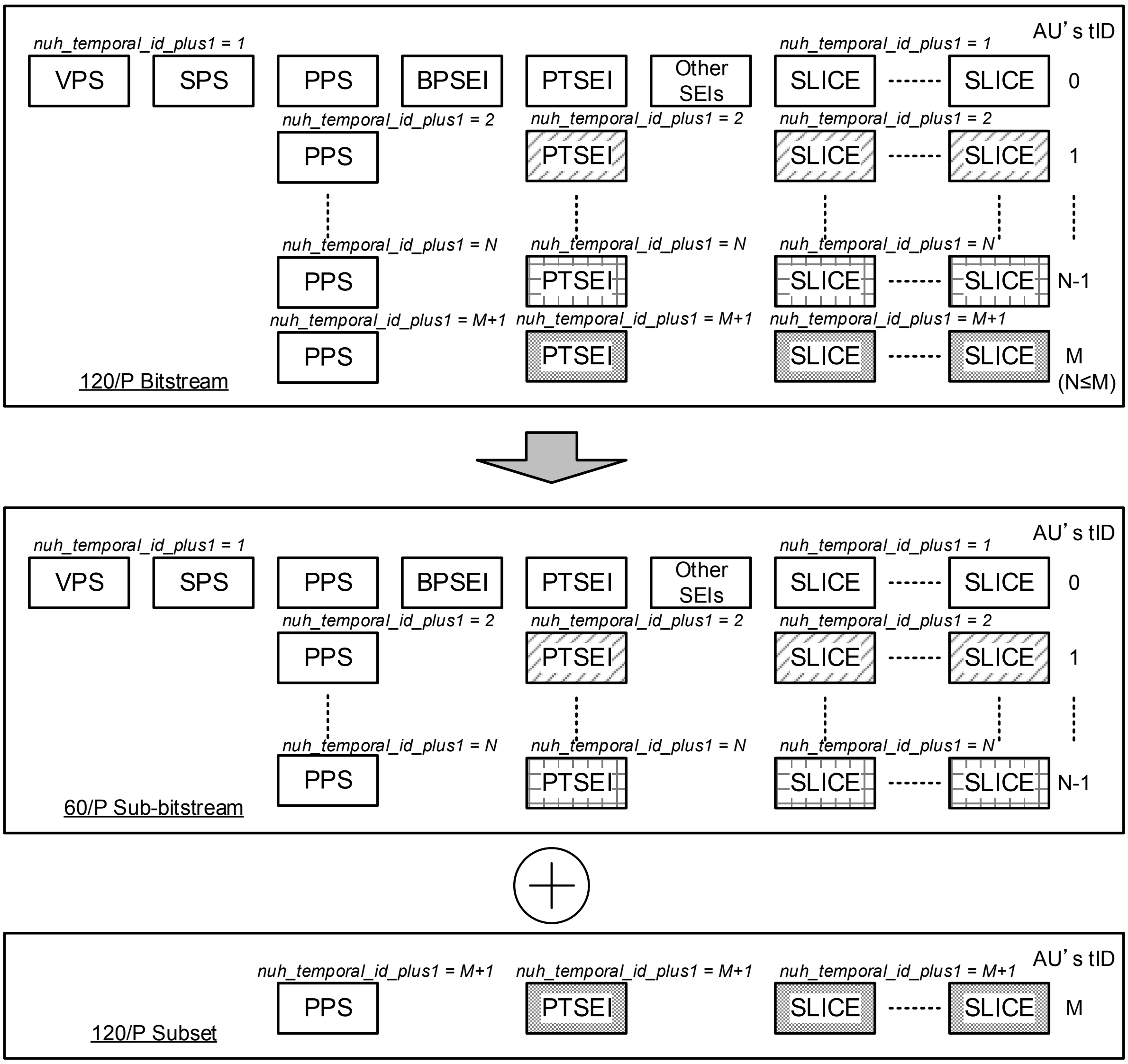


Fig. 4-12: Structure of 120Hz/60Hz temporal scalable coding bitstream

4.10.2 Provisions on bitstream

The provision of 120/P bitstream for temporal scalable coding is shown in the followings.

(1) SOP structure

Picture re-ordering is applied excluding the tail of CVS.

SOP structure shall be any of L=1, L=2, L=3, L=4 by Method 1 or Method 2 in Fig. 4-13, excluding the head and the tail of CVS. SOP structure by the same method shall be usable in the same CVS. Maximum Temporal ID shall be always sps_max_sub_layers_minus1 irrespective of SOP structure (see Fig. 4-14).

At the head of CVS, the same structure as SOP in middle of CVS, or L=0 in Fig. 4-13 is used.

The reference relationship of AU can be set freely in the scope of provided decoding order, and in the scope which satisfies the restriction on various size of Temporal ID.

At the tail of CVS (from last SOP including AU in which Temporal ID in CVS is 0 to final AU in the same CVS), SOP structure excluding L=1, L=2, L=3, L=4 by Method 1 or Method 2 in Fig. 4-13 may be taken. Also, picture re-ordering may not be applied. (See Fig. 5-3)

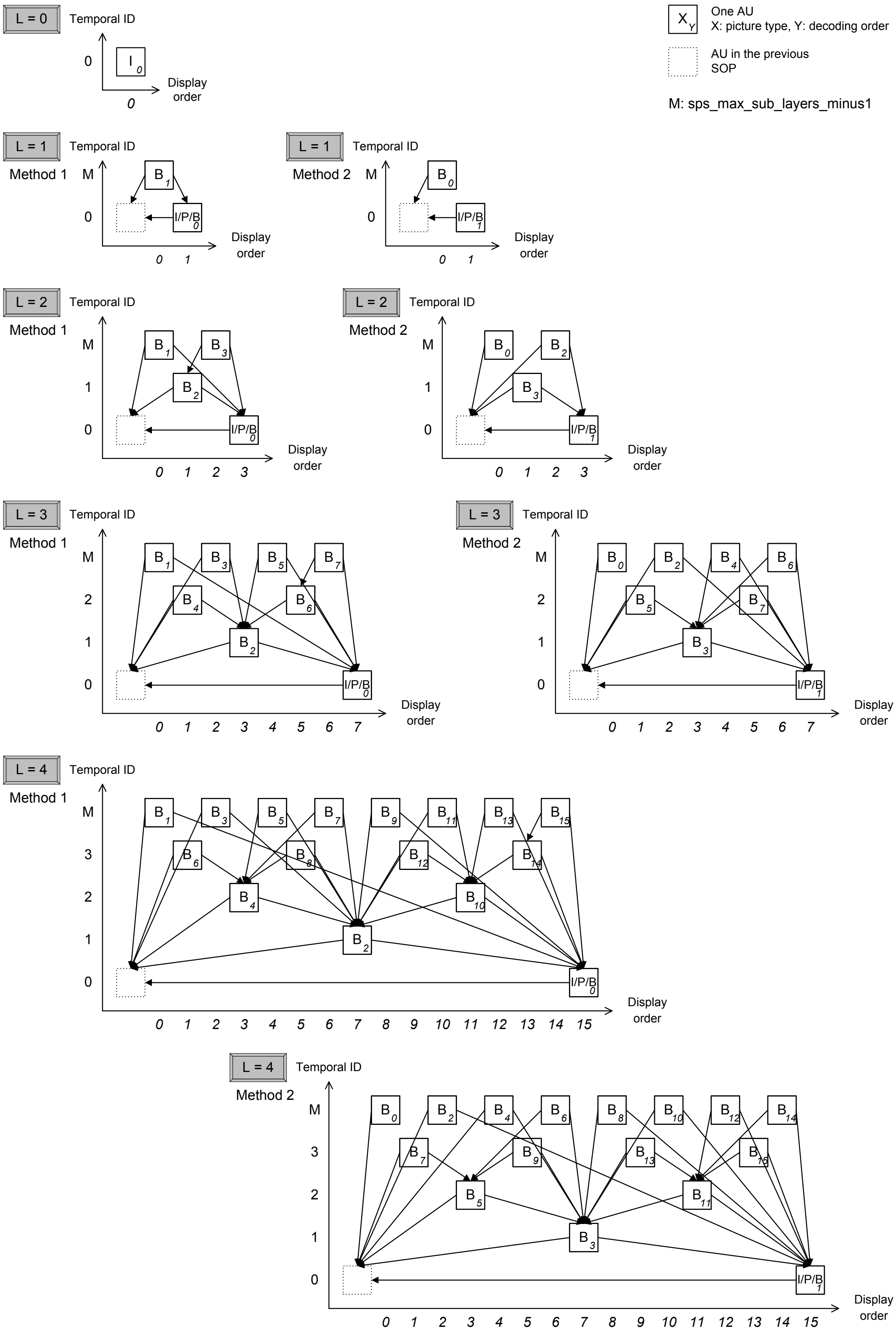

One AU
X: picture type, Y: decoding order
AU in the previous SOP
M: sps_max_sub_layers_minus1
L = 0
L = 1
L = 2
L = 3
L = 4
Method 1
Method 2
Temporal ID
Display order
I/P/B

Fig. 4-13: SOP structure of 120/P bitstream

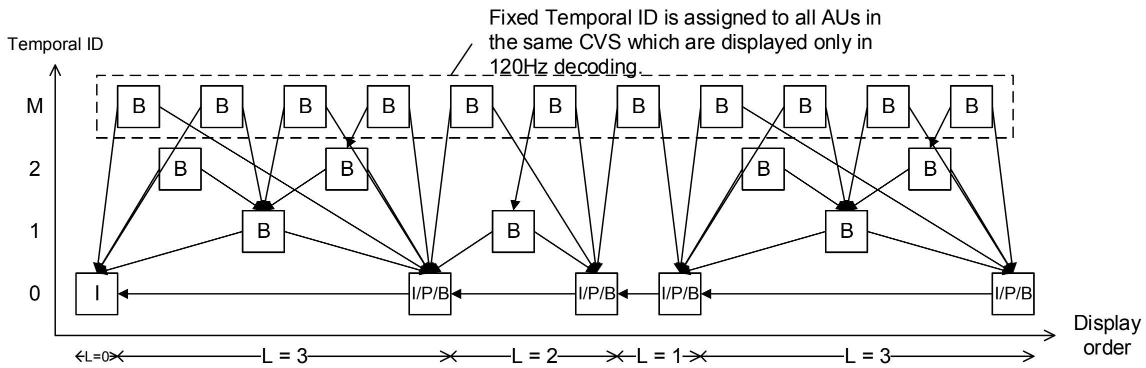


Fig. 4-14: Restrictions in case that SOP structure changes in CVS

(Description)

120/P bitstream is supplied as two bitstreams which are composed of 60/P sub-bitstream and 120/P subset. Then, applying SOP structure shown in Fig.4-13, DTS/PTS of each AU in 60/P sub-bitstream is specified at the interval of 60Hz, and DPS/PTS of each AU in 120/P subset becomes middle value of DTS/PTS of two AU in 60Hz sub-bitstream, and it is also specified at the interval of 60Hz. So when decoding 120Hz, DPS/PTS of 60/P sub-bitstream and 120/P subset can be used just as it is. On the other hand, when decoding 60Hz, DTS/PTS of 60P sub-bitstream can be used just as it is. Here, DTS and PTS are calculated by parameters of Picture timing SEI (au_cpb_removal_delay_minus1, pic_dpb_output_delay) which are added to each AU. (See Fig.4-15.)

SOP structure of 60/P sub-bitstream is the same as the structure in Fig. 4-2.

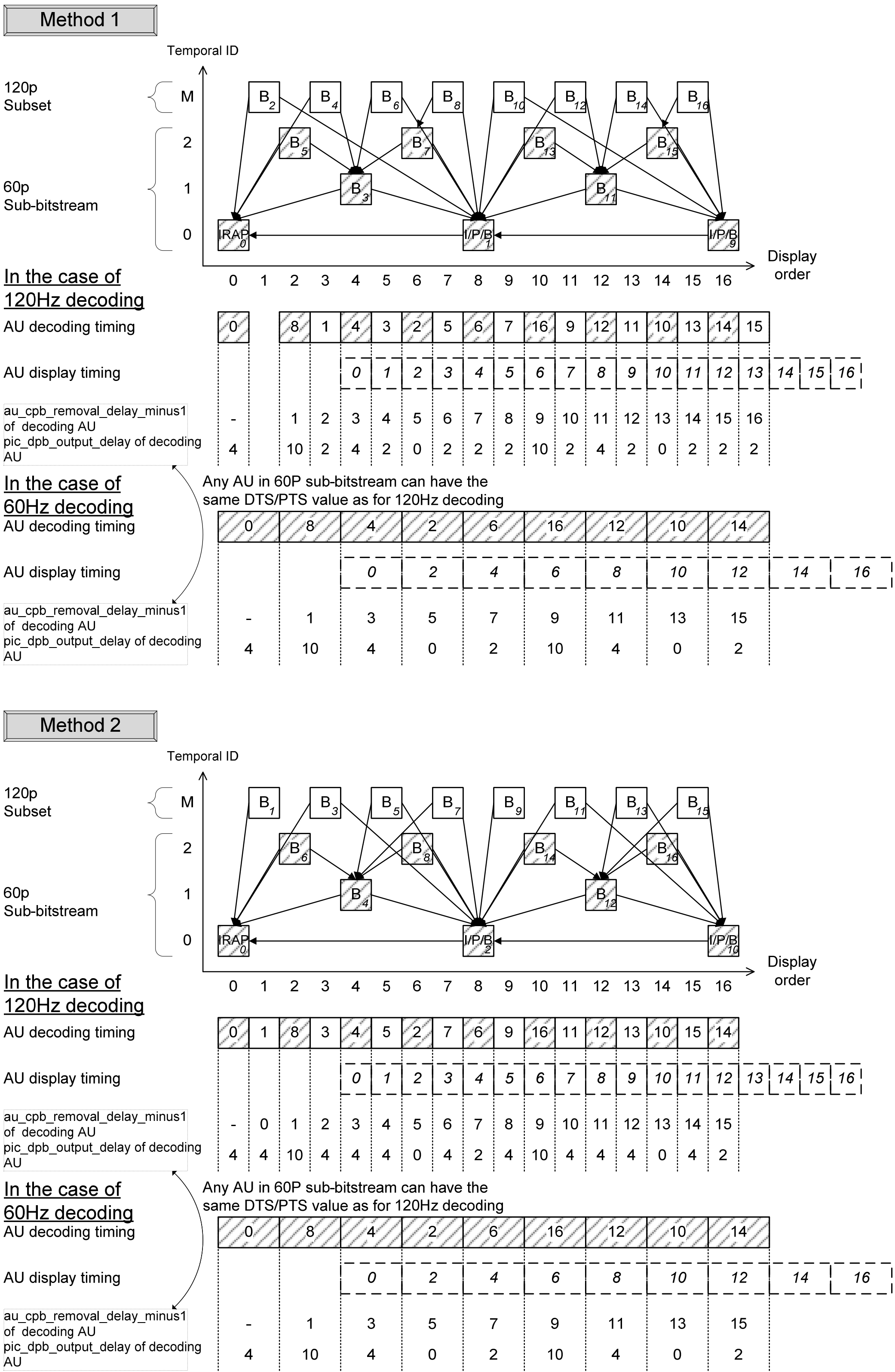


Fig. 4-15: Decode and display timing for each AU in 120/P bitstream

(2) SPS, VUI

The values of each syntax element of SPS, VUI related to temporal scalable coding are as shown in from Table 4-25 to Table 4-28.

Table 4-25: Restrictions on SPS

Syntax element	Value	Explanations
sps_max_sub_layers_minus1	6	Fixed irrespective of maximum Temporal ID in 60Hz sub-bitstream
profile_tier_level()	See Table 4-26	
sps_sub_layer_ordering_info_present_flag	0	As for the values of sps_max_dec_pic_buffering_minus1, sps_max_num_reorder_pics, and sps_max_latency_increase_plus1, only the values when displaying 120Hz decoding are described.
vui_parameters_present_flag	1	Timing information is described by VUI.
vui_parameters()	See Table 4-27	

Table 4-26: Restrictions on profile_tier_level()

Syntax element	Value	Explanations
general_profile_space	0	HEVC standard
general_tier_flag	0	Main tier
general_profile_idc	2	Main10 Profile
general_profile_compatibility_flag[j] (j = [0, 31])	[0, 1]	1: only when j=2 0: otherwise
general_progressive_source_flag	1	
general_interlaced_source_flag	0	
general_frame_only_constraint_flag	1	
general_level_idc	156, 186	Level description of 120Hz bitstream 156: Level 5.2 186: Level 6.2
sub_layer_profile_present_flag[i] (i = [0, sps_max_sub_layers_minus1 − 1])	0	All are the same Profile
sub_layer_level_present_flag[i] (i = [0, sps_max_sub_layers_minus1 − 1])	[0, 1]	1: when i is equal to tID_{max} (maximum Temporal ID of AU in 60Hz sub-bitstream) 0 or 1: when i < tID_{max} 0: otherwise
sub_layer_level_idc[i] (i = [0, sps_max_sub_layers_minus1 − 1])	153, 183	153(2160/ 60/P) or 183(4320/60/P): when i is equal to tID_{max} 153 or 183: when i < tID_{max} and sub_layer_level_present_flag[i] is 1 Undefined: otherwise

Table 4-27: Restrictions of VUI

Syntax element	Value	Explanations
vui_timing_info_present_flag	1	
vui_num_units_in_tick	1,000 or 225,000: for 120.00Hz 1,001 or 225,225: for 119.88Hz	See Table 4-9
vui_time_scale	120,000 or 27,000,000	See Table 4-9
vui_poc_proportional_to_timing_flag	1	
vui_num_ticks_poc_diff_one_minus1	See the right.	vui_num_units_in_tick – 1
vui_hrd_parameters_present_flag	1	
hrd_parameters()	See Table 4-28	

Table 4-28: Restriction on hrd_parameters()

Syntax element	Value	Explanations
nal_hrd_parameters_present_flag	1	
vcl_hrd_parameters_present_flag	0	
sub_pic_hrd_params_present_flag	0	
fixed_pic_rate_general_flag[i] (i = [0, sps_max_sub_layers_minus1])	[0, 1]	1: in case that output time of all AU in bitstream is continuous 0: in case that output time of all AU in bitstream is not continuous
fixed_pic_rate_within_cvs_flag[i] (i = [0, sps_max_sub_layers_minus1])	[0, 1]	1: in case that output time of all AU in CVS is continuous 0: in case that output time of all AU in CVS is not continuous If i is sps_max_sub_layers_minus1 or tID_{max}, the value shall be 1.
elemental_duration_in_tc_minus1[i] (i = sps_max_sub_layers_minus1)	0, 1	0: 120Hz decode displaying 1: 60Hz decode displaying Common ClockTick to 120Hz and 60Hz is used.
elemental_duration_in_tc_minus1[i] (i = tID_{max} and sps_max_sub_layers_minus1 != tID_{max})	1	Setting value for 60Hz substream
cpb_cnt_minus1[i] (i = [0, sps_max_sub_layers_minus1])	0	
bit_rate_value_minus1[i][j] cpb_size_value_minus1[i][j] (i = [0, sps_max_sub_layers_minus1]) (j = [0, cpb_cnt_minus1[i])	See the right.	Maximum bitrate value and CPB size value of sub-layer whose maximum Temporal ID is i. When i is equal to sps_max_sub_layers_minus1 or tID_{max}, effective value is set.
cbr_flag[i][j] (i = [0, sps_max_sub_layers_minus1]) (j = [0, cpb_cnt_minus1[i])	[0, 1]	CBR flag of sub-layer whose maximum Temporal ID is i. When i is equal to sps_max_sub_layers_minus1 or tID_{max}, effective value is set.

(Description)

sps_max_sub_layers_minus1 is set to 6 which is upper limit of HEVC Standard irrespective of maximum Temporal ID in 60Hz sub-bitstream. Temporal ID of 120Hz subset is equal to sps_max_sub_layers_minus1 for both maximum and minimum.

fixed_pic_rate_general_flag[i] is set to 1 when display time of all AU in bitstream is continuous (without freeze or missing) in sub-layer whose maximum Temporal ID is equal to i. When display delay (see 4.7) of each CVS in bitstream is different, it is 0 because display time of AU in around the boundary of CVS is discontinuous.

fixed_pic_rate_within_cvs_flag[i] is set to 1 when display time of all AU in each CVS is continuous in sub-layer whose maximum Temporal ID is equal to i. In this operational provision, when i is equal to sps_max_sub_layers_minus1 or tID_{max}, it is set to 1, and otherwise it is not prescribed. However, when fixed_pic_rate_general_flag[i] is 1, it is considered as 1.

By setting sps_sub_layer_ordering_info_present_flag to 1, it becomes possible to describe DPB parameters of all temporal scales. But it also makes dynamic change of SOP structure in CVS possible, so this operational guideline does not apply.

HRD parameters (bit_rate_value_minus1[][], cpb_size_value_munus1[][], cbr_flag[][]) related to bitrate are described about the case of 120Hz subset＋60Hz sub-bitstream, and the case of 60Hz sub-bitstream, respectively.

DPB related parameters (sps_max_dec_pic_buffering_minus1, sps_max_num_reorder_pics[], sps_max_latency_increase_plus1[]) equivalent to 60Hz sub-bitstream are not described. The values at 120Hz decode and display are succeeded.

4.10.3 Provisions on multiplex

(1) Provisions in MPEG-2 Systems

In case of using MPEG-2 Systems (TS form) for multiplex system, recommended description is explained.

・Transmission form

Fig. 4-16 shows a block diagram of transmission for temporal scalable coding HEVC bitstream in MPEG-2 TS. 60/P sub-bitstream of Level 5.1/6.1 (for 60Hz decode and display) and 120/P subset for 120Hz decode and display are transmitted by separate PID. After buffering in ES buffer of the receiver, they are transmitted to HEVC decoder in DTS order. This transmission data is 120/P bitstream based on Level 5.2/6.2.

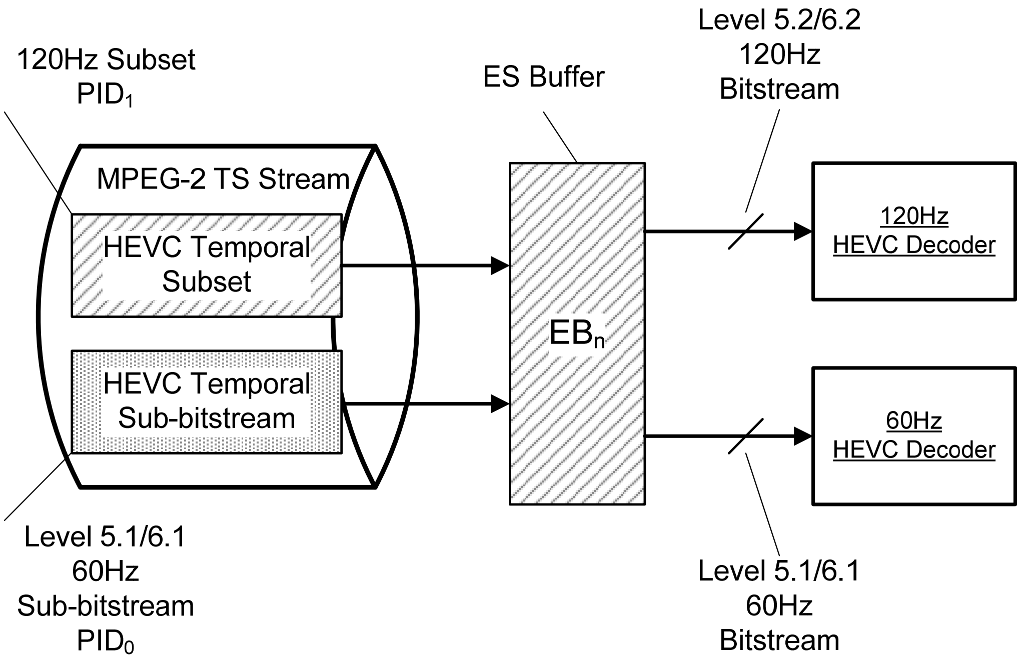


Fig. 4-16: Transmission of temporal scalable coding HEVC bitstream in MPEG-2 TS

・Stream type

stream_type of HEVC temporal sub-bitstream is set to 0x24.

stream_type of HEVC temporal subset is set to 0x25.

・Descriptor

Temporal scalable coding is described about HEVC temporal sub-bitstream and HEVC temporal subset respectively, by using Hierarchy descriptor (descriptor_tag = 4).

Here, when there are no HEVC temporal subset, Hierarchy descriptor is not added to HEVC bitstream.

Table 4-29: Setting Hierarchy descriptor

Syntax element	Value	Explanations
temporal_scalability_flag	0	Temporal scalable coding
spatial_scalability_flag	1	Not spatial scalable coding
quality_scalability_flag	1	Not quality scalable coding
hierarchy_type	3 15	3: HEVC temporal subset (Temporal Scalability) 15: HEVC temporal sub-bitstream
hierarchy_layer_index	0	In case of temporal scalable coding, there are no exchange in the same AU according to hierarchy_layer_index, and the same value is used in subset and sub-bitstream.
tref_present_flag	1	TREF field is not used.
hierarchy_embedded_layer_index	15	undefined not used in temporal scalable coding
hierarchy_channel	0	Temporal subset and temporal sub-bitstream are transmitted in the same channel.

HEVC bitstream is described about HEVC temporal sub-bitstream and HEVC temporal subset respectively, by using HEVC video descriptor (descriptor_tag = 56).

Table 4-30: Setting HEVC video descriptor

Syntax element	Value	Explanations
profile_space tier_flag profile_idc profile_compatibility_indication progressive_source_flag interlaced_source_flag non_packed_constraint_flag frame_only_constraint_flag level_idc	See the right.	In case of HEVC temporal sub-bitstream, the value of Level 5.1/6.1 bitstream is described. In case of HEVC temporal subset, the value of Level 5.2/6.2 bitstream is described.
temporal_layer_subset_flag	1	Layer information is described.
HEVC_still_present_flag	0	
HEVC_24hr_picture_present_flag	0	
temporal_id_min	See the right.	0: HEVC temporal sub-bitstream sps_max_sub_layers_minus1: HEVC temporal subset
temporal_id_max	See the right.	Maximum value of tID_{max} of Temporal ID included in sub-bitstream, in case of HEVC temporal sub-bitstream sps_max_sub_layers_minus1, in case of HEVC temporal subset

(Description)

As level of 120Hz subset is described on general_level_idc in SPS, it is not appropriate to refer to general_level_idc in SPS when only 60Hz sub-bitstream is decoded. So, when multiplexing in MPEG-2 Systems, HEVC video descriptor is always applied, and level of 60Hz sub-bitstream is described on HEVC video descriptor of 60Hz sub-bitstream. Also, level of 60Hz sub-bitstream is described on sub_layer_level_idc[i] ($i = tID_{max}$) in SPS of HEVC.

As for temporal_id_min and temporal_id_max of HEVC temporal subset, sps_max_sub_layers_minus1 is set.

(2) MMT provision

In case of using MMT as multiplex system, recommended description is explained.

・Transmission form

HEVC temporal subset and HEVC temporal sub-bitstream are transmitted as separate assets in the same MMT package.

The same MPU sequence number is added to each MPU in both assets which corresponds to time.

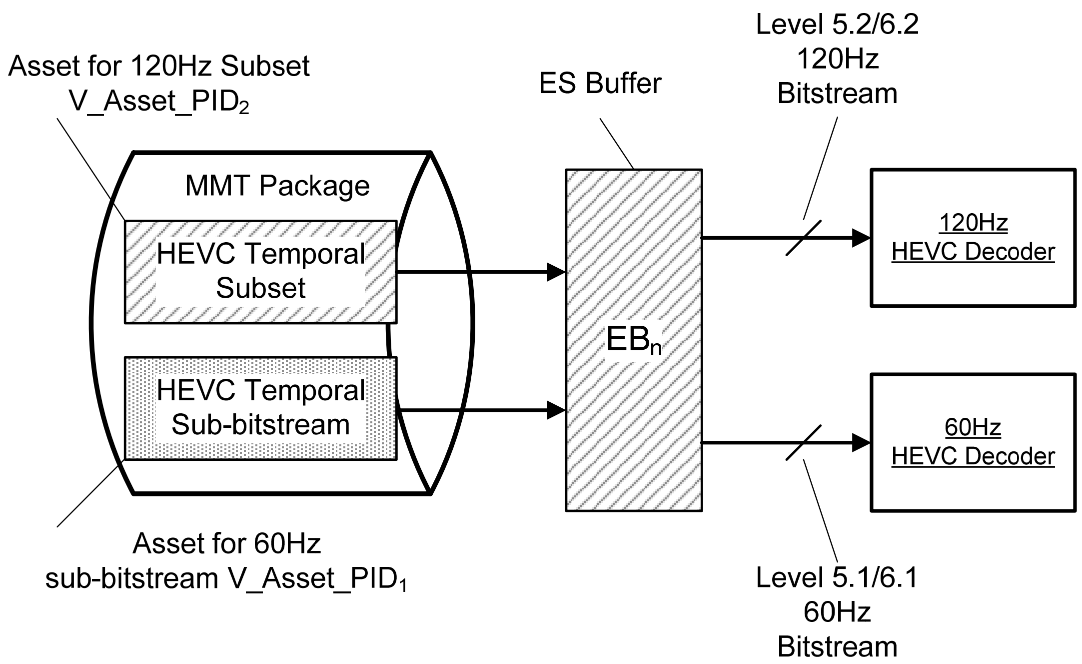


Fig. 4-17: Transmission of temporal scalable coding HEVC bitstream in MMT

・Descriptor

This is used when temporal scalable coding HEVC is multiplexed. Descriptors specified in ARIB STD-B60 are shown in Table 4-31. These descriptors are added to asset of 60Hz sub-bitstream and asset of 120Hz subset, respectively.

Table 4-31: Descriptors of temporal scalable coding HEVC

Name of descriptor	Explanations
MH-HEVC video descriptor (MH-HEVC_Descriptor)	Contents of description are the same as HEVC video descriptor in MPEG-2 multiplex. See Table 4-30.
Video component descriptor (Video_Component_Descriptor)	Video parameter of asset is described. See Table 4-32.
Dependency descriptor (Dependency_Descriptor)	Asset in dependent relationship is described. See Table 4-33.
MH-scalable coding descriptor (MH-Hierarchiy_Descriptor)	Contents of descriptions are the same as Hierarchy descriptor in MPEG-2 TS multiplex. See Table 4-34.

Table 4-32: Setting video component descriptors

Syntax element	Value	Explanations
video_resolution	6, 7	6: 2160/P 7: 4320/P
video_aspect_ratio	3	16:9, without pan-vector
video_scan_flag	1	progressive
video_frame_rate	7, 8, 10, 11	In 60Hz sub-bitstream, 7 (60/1.001 Hz) or 8 (60 Hz) is set. In 120Hz subset, 10 (120/1.001 Hz) or 11 (120 Hz) is set.
video_transfer_characteristics	0, 1, 2, 3, 4, 5	0: Video transfer characteristics is not specified. 1: VUI transfer_characteristics = 1 (Rec. ITU-R BT.709-5) 2: VUI transfer_characteristics = 11 (IEC 61966-2-4) 3: VUI transfer_characteristics = 14 (Rec. ITU-R BT.2020) 4: VUI transfer_characteristics = 16 (Rec. ITU-R BT.2100 PQ) 5: VUI transfer_characteristics = 18 (Rec. ITU-R BT.2100 HLG)

Table 4-33: Setting dependency descriptors

Syntax element	Value	Explanations
num_dependencies	1	60Hz sub-bitstream and 120Hz subset are complementary.
asset_id_scheme asset_id_length asset_id_byte	See the right	In case of 60Hz sub-bitstream asset, asset ID of 120Hz subset asset is described. In case of 120Hz subset asset, asset ID of 60Hz sub-bitstream asset is described.

Table 4-34: Setting Hierarchy descriptors

Syntax element	Value	Explanations
temporal_scalability_flag	0	Temporal scalable coding
spatial_scalability_flag	1	Not spatial scalable coding
quality_scalability_flag	1	Not quality scalable coding
hierarchy_type	3 15	3: HEVC temporal subset (Temporal Scalability) 15: HEVC temporal sub-bitstream
hierarchy_layer_index	0	In case of temporal scalable coding, there are no exchange in the same AU according to hierarchy_layer_index, and the same value is used in subset and sub-bitstream.
tref_present_flag	1	TREF field is not used.
hierarchy_embedded_layer_index	15	undefined not used in temporal scalable coding
hierarchy_channel	0	Temporal subset and temporal sub-bitstream are transmitted in the same channel.

4.11 An example of encoding in case of film material

As a recommended example of encoding 24/1.001Hz film material to 60/1.001Hz, the method of using pic_struct of Picture timing SEI is shown as the following.

When 3:2 pull-down is detected at the transmitting side, in case of interlaced scan system, two fields which have the same display timing are added and coded as a progressive frame. For the field which is equivalent to third field in 3:2 pull-down system, a flag which shows it is the same as the first field is only set, and video data is not transmitted about the field.

On the other hand, in case of progressive scan system, a flag which shows to repeat displaying each frame twice, three times, twice, three times is set, and only 24 frames of video data are transmitted.

Then, by setting pic_struct to the value as shown in Table 4-35, 3:2 sequence is reproduced in the receiver. Setting value at each picture shall be the value shown in Fig. 4-18 and Fig. 4-19.

Table 4-35: Parameters in case of coding film materials

Parameter	Interlace (59.94/I)	Progressive (59.94/P)
general_progressive_source_flag	0	1
general_interlaced_source_flag	1	0
general_frame_only_constraint_flag	0	1
frame_field_info_present_flag	1	1
sps_num_units_in_tick sps_time_scale	1,001 60,000	1,001 60,000
field_seq_flag	1	0
pic_struct	interval=4 5: top-bottom-top 4: bottom-top 6: bottom-top-bottom 3: top-bottom	interval=2 8: frame tripling 7: frame doubling

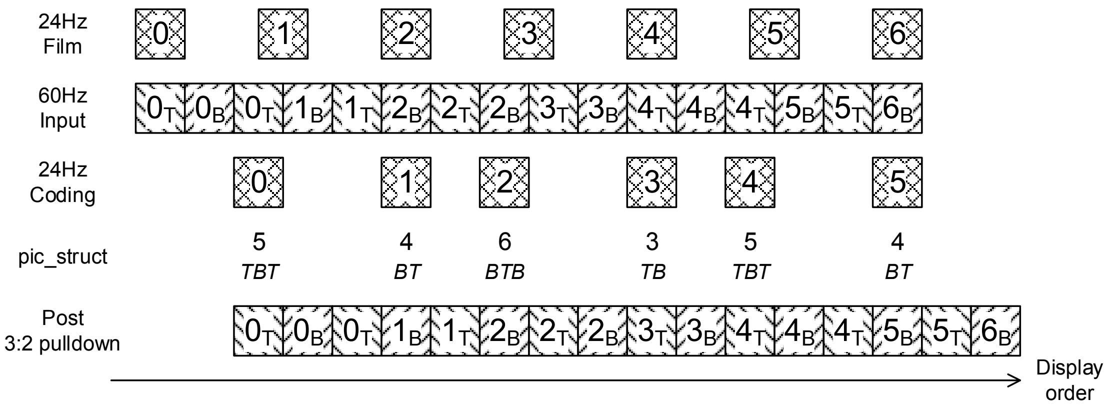


Fig. 4-18: An example of encoding film material (in case of interlace)

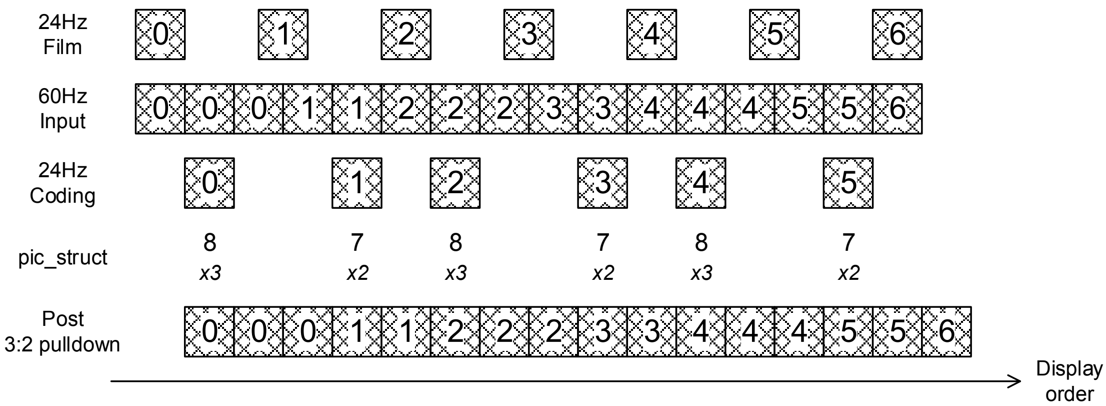


Fig. 4-19: An example of encoding film material (in case of progressive)

4.12 An example of encoding in case of low frame rate material

As an operational example in case of encoding 30/1.001Hz low frame rate material to 60/1.001Hz, a method of using pic_struct of Picture timing SEI is shown as the following.

When low frame rate material is input at the transmission side, video data composed of only 30 frames is transmitted by setting a flag which shows displaying each frame twice.

Then, by setting pic_struct to the value shown in Table 4-36, 60/1.001Hz video is reproduced in the receiver. Setting value of pic_struct at each picture shall be the value shown in Fig.4-20.

Table 4-36: Parameters in case of encoding low frame rate material

Parameter	
general_progressive_source_flag	1
general_interlaced_source_flag	0
general_frame_only_constraint_flag	1
frame_field_info_present_flag	1
sps_num_units_in_tick sps_time_scale	1,001 60,000
field_seq_flag	0
pic_struct	interval=1 7: frame doubling

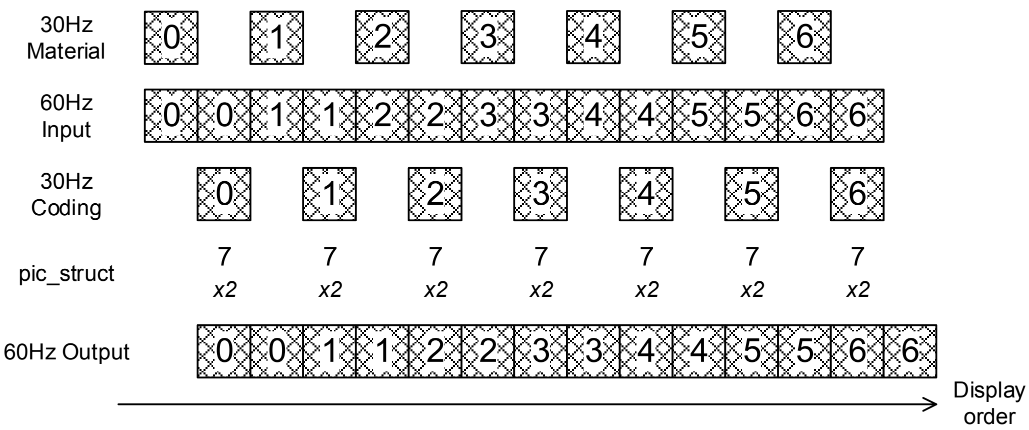


Fig. 4-20: An example of encoding low frame rate material

Chapter 5: Restrictions on system

5.1 Multiplex in general

5.1.1 Multiplex by MPEG-2 Systems Standard

Restrictions are shown in Table 5-1.

Table 5-1: Restrictions on multiplex by MPEG-2 Systems Standard

Composition	Procedure
PES packet	One PES packet always includes 1 AU, and must not include plural AU. In case that video format is 1080/I and field_seq_flag is 1, PES packet includes only AU of 1 field. In PES Header, PTS must be transmitted. In the receiver, after start of decoding and output control shall be operated according to PTS of PES Header and DTS. Also, setting value of PTS_DTS_flag is as the followings. 11b: in case that PES packet includes I-picture, P-picture or B-picture in which PTS and DTS are different. 10b: in case that PES packet includes B-picture in which PTS and DTS are the same.
STD delay	STD delay shall be less than 1 second in normal operation.
Descriptor	In order to make it possible that the receiver knows parameters for decoding before analyzing video stream, it is recommended that basic parameters of HEVC ES (profile, level, maximum and minimum Temporal ID of temporal scalable coding subset, etc.) are described by using HEVC video descriptor defined in MPEG-2 Systems Standard. HEVC timing and HRD descriptor is not included in stream.
HEVC stream format	Byte stream format (Annex B) Note: restrictions in MPEG-2 Systems Standard

(Description)

As vui_timing_info_present_flag is set to 1 in HEVC ES, HEVC timing and HRD descriptor is not necessary.

5.1.2 Multiplex by MMT Standard

Restrictions are shown in Table 5-2.

Table 5-2: Restrictions of multiplex by MMT Standard

Composition	Procedure
MMTP packet	DTS/PTS of each AU is described by using MPU time stamp descriptor and MPU extension time stamp descriptor. In case of 1080/I and field coding, DTS/PTS is added every field.
HRBM delay	Sum of HRBM delay (FEC decoding delay, de-jitter delay, MMTP decapsulation delay) and CPB delay shall be less than 1 second.
Descriptor	In order to make it possible that the receiver knows parameters for decoding before analyzing video stream, it is recommended that basic parameters of HEVC ES (profile, level, maximum and minimum Temporal ID of temporal scalable coding subset, etc.) by using MH-HEVC video descriptor defined in MMT/TLV system. MH-HEVC timing HRD descriptor is not included in stream.
HEVC stream format	Byte stream format (Annex B) is not applied. Note: restriction in ARIB STD-B60 Standard

(Description)

As vui_timing_info_present_flag is set to 1 in HEVC ES, MH-HEVC timing and HRD descriptor is not necessary.

5.2 Channel-hopping time

5.2.1 Restrictions on bitstream

The following operations are recommended in order to suppress channel-hopping time within a fixed time.

AU which can be re-captured is coded once at least 533 ms. In case of field coding, it is the AU which is Recovery point SEI is added to and whose slice_type is I. In case of other than field coding, it is IRAP AU.

(Note) Parameters of video coding related to channel-hopping time are frequency of transmitting the AU which can be re-captured, CPB delay and so on.

(Note) When decoding starts from the AU which can be re-captured, this AU and all AU whose display order is after this AU are able to be decoded normally.

5.2.2 Channel-hopping time in multiplex by MPEG-2 Systems

Fig. 5-1 shows flowchart of each step related for channel-hopping time in case that MPEG-2 Systems is used for multiplex system of Advanced BS digital broadcasting, and Table 5-3 shows an example of each step and delay time.

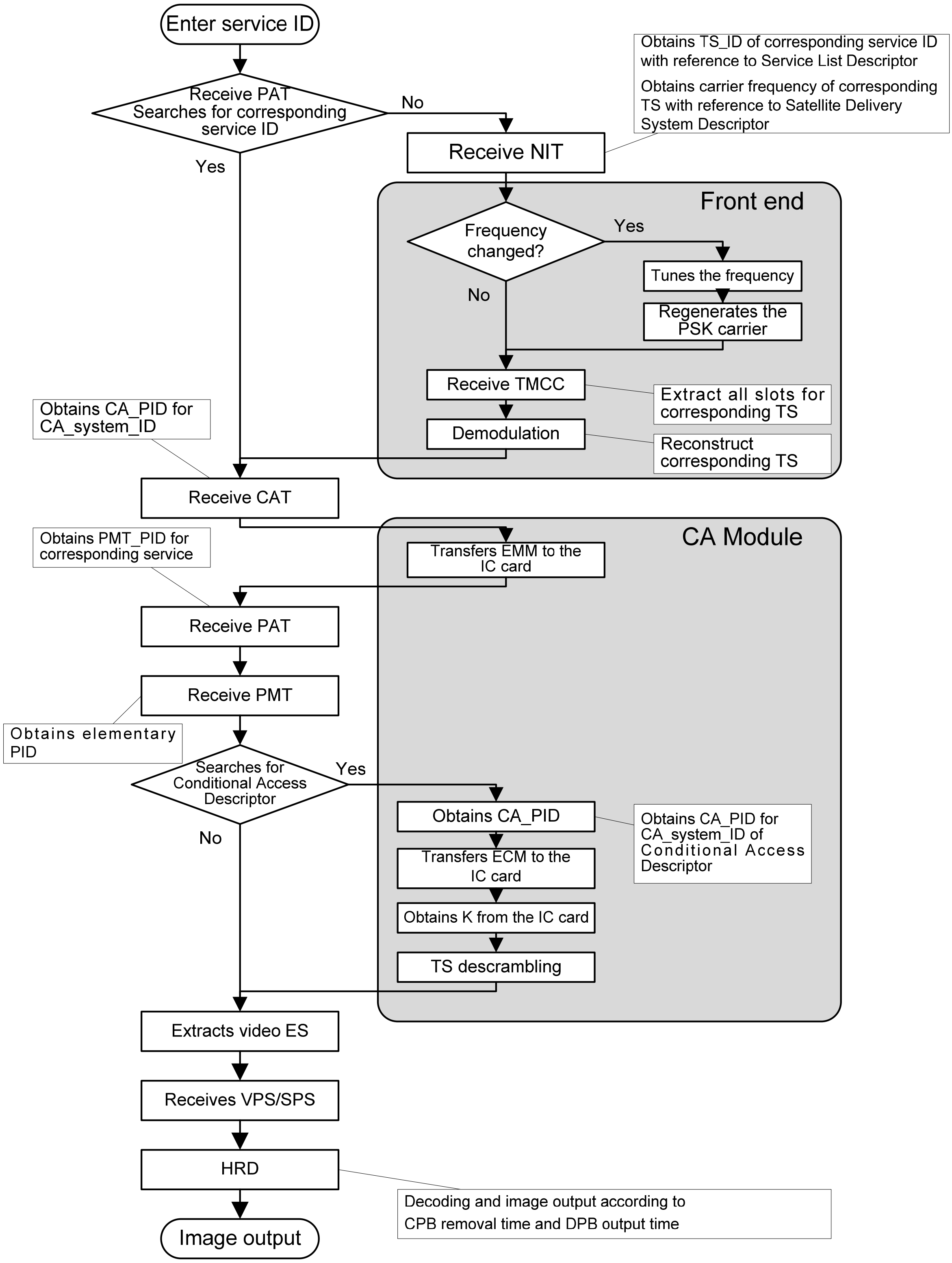


Fig. 5-1: Process flowchart of multiplex by MPEG-2 Systems

Table 5-3: Delay time in multiplex by MPEG-2 Systems

Operation	Maximum delay time	Remarks
PAT receiving	100 ms	0 ms: in case of storing in memory
NIT receiving	10 s	0 s: in case of storing in memory
Frequency tuning	50 ms	
PSK carrier regeneration	600 ms	Measured value by experimental receiver Condition of measurement: Difference of frequency ±1.5 MHz C/N 2dB
extracting slot of concerned MPEG-2 TS from TMCC receiving	34 ms	Value for Advanced BS digital broadcasting
CAT receiving ～ extracting video stream	200 ms	Assumed of updating PSI information and storing EMM in memory every 100 ms
VPS/SPS receiving	533 ms	IRAP interval in case of SOP structure is L=3 (4 layers) 16/30 sec (59.94/I) 32/60 sec (59.94/P) 64/120 sec (119.88/P)
Delay of start decoding head AU	500 ms	AuNominalRemovalTime[0]
Display delay	100 ms (59.94/I) 50 ms (59.94/P) 25 ms (119.88/P)	In case of L=3 (4 layers)
Sum	2,017 ms (59.94/I) 1,967 ms (59.94/P) 1,942 ms (119.88/P)	Assumed that PAT and NIT are stored in memory.

5.2.3 Channel-hopping time in multiplex by MMT

Fig. 5-2 shows flowchart of each step related for channel-hopping time in case that MMT Systems is used for multiplex system of Advanced BS digital broadcasting, and Table 5-4 shows an example of each step and delay time.

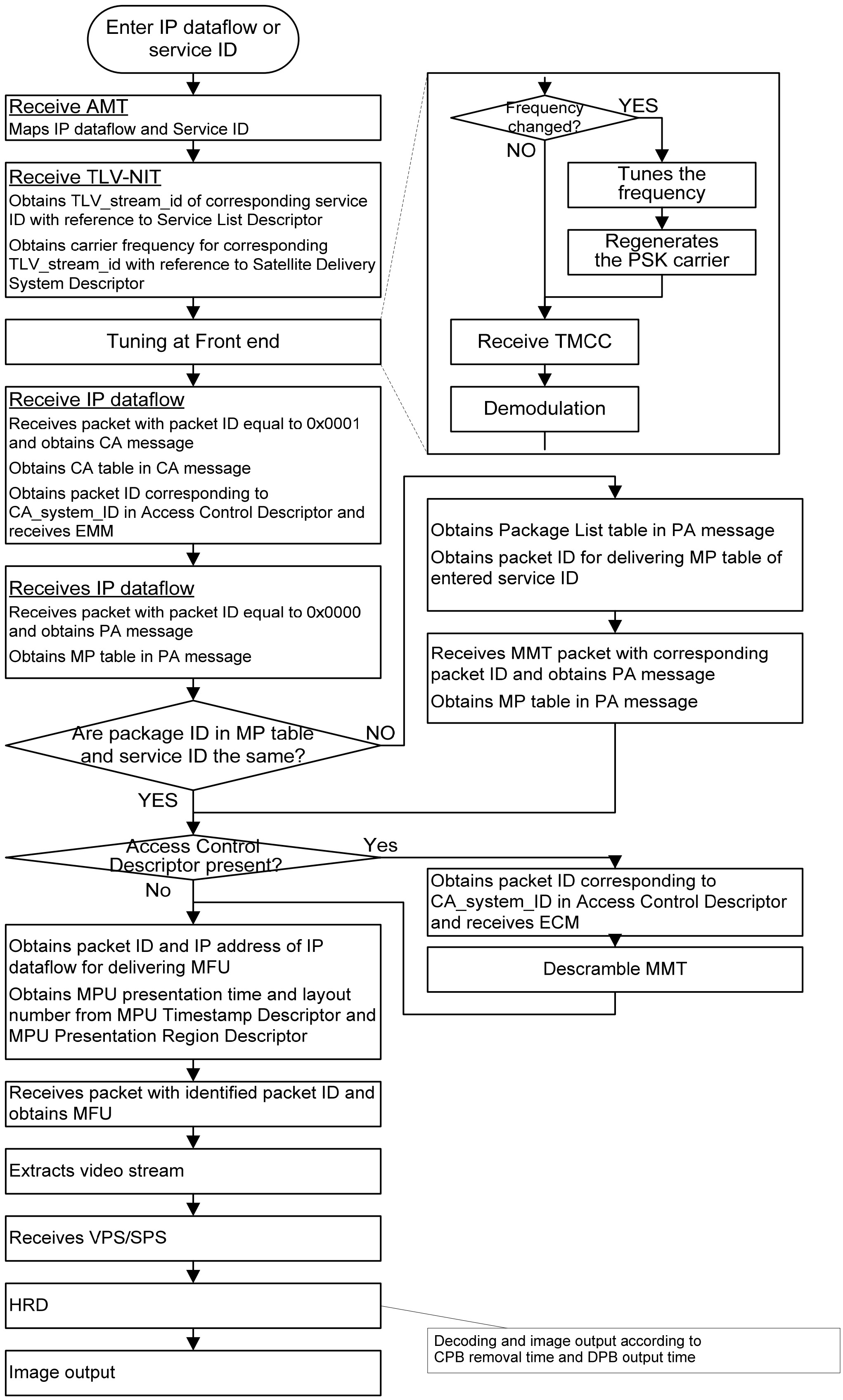


Fig. 5-2: Process flowchart of multiplex by MMT

Table 5-4: Delay time in multiplex by MMT

Operation	Maximum delay time	Remarks
AMT receiving	10 s	0 s: in case of storing in memory
TLV-NIT receiving	10 s	0 s: in case of storing in memory
Frequency tuning	50 ms	
PSK carrier regeneration	600 ms	Measured value by experimental receiver Condition of measurement: Difference of frequency ±1.5 MHz C/N 2dB
Extracting slot of concerned TLV from TMCC receiving	34 ms	Value for Advanced BS digital broadcasting
PA message acquisition (MP table)	100 ms	
PA message acquisition (package list table)	0 ms	Packaged in PA message including MP table
CA message receiving	100 ms	
Acquiring MFU of specific packet ID	533 ms	This is the same as IRAP interval. IRAP interval in case that SOP structure is L=3 (4 layers): 16/30 sec (59.94/I) 32/60 sec (59.94/P) 64/120 sec (119.88/P)
VPS/SPS receiving	0 ms	Receiving VPS/SPS at the head MFU of MPU
Delay of start decoding the head AU	500 ms	AuNominalRemovalTime[0]
Display delay	100 ms (59.94/I) 50 ms (59.94/P) 25 ms (119.88/P)	In case of L=3 (4 layers)
Sum	2,017 ms (59.94/I) 1,967 ms (59.94/P) 1,942 ms (119.88/P)	Assumed that AMT and TLV-NIT are stored in memory.

5.3 Seamless switching

5.3.1 HEVC procedure

(1) Changing the video coding parameters

HEVC procedure is recommended in order to realize the continuity of decoded video output from HEVC decoding block of the receiver when changing the video coding format and so on.

・Video coding parameters as the object

When at least one of the following video coding parameters is changed, CVS before and after the change is restricted. By this restriction, HEVC decoding block of the receiver can output decoded pictures continuously even at CVS boundary in HEVC Standard

- bitrate, or CPB size
- picture resolution
- frame rate
  - ・from 59.94Hz progressive to 119.88Hz progressive, or the reverse
  - ・from 60Hz progressive to 120Hz progressive, or the reverse
  - ・from 59.94Hz interlace to 59.94Hz progressive, or the reverse
- profile (Main/Main10)
- transfer characteristics of video (transfer_characteristics)
  - ・from SDR(1, 11, 14) to HDR(16, 18), or the reverse
  - ・from HDR(16) to HDR(18), or the reverse

(Note) This provision is only the restriction on HEVC bitstream. Continuous output or continuous display from the receiver at the boundary of CVS shall be optional.

(Note) There is a case that level is also changed according to the change of parameters.

Also, when the following video coding parameter is changed, discontinuity of display for several seconds is permitted.

(Note) Discontinuity of display means the situation that black picture is displayed or the picture is frozen for about 1 second just as channel-hopping. The difference of display time between the last display frame before the change and the first display frame after the change becomes longer than 1 frame time.

- frame rate

・from 59.94Hz system to 60Hz system, or the reverse

Restrictions on CVS before and after changing restrictions on CVS are as the following. See Fig.5-3 (change between 60/P and 120/P, an example of the case that SOP structure is L=2).

— CVS before and after the change must have the same display delay.
  ・"Display delay" is the difference between the decoding time for the head AU of CVS and the display time when AU is displayed first in CVS.
  ・As for an example in Fig. 5-3, (ClockTick * 4) is the display delay value.
  ・The value of ClockTick describes precise enough value to describe the decoded delay (au_cpb_removal_delay_minus1) and the output delay (pic_dpb_output_delay) of CVS whose frame rate is large.
  ・An example of Fig. 5-3 shows 0.008333 second (= 1,000 / 120,000) according to the precision of 120Hz.

(Note) It is necessary to follow this restriction in the case that frame rate is changed in CVS before and after the change, and that SOP structure is changed.

(Note) When assuring seamless change, minimum "display delay" is maximum value of re-ordering delay of each CVS included in bitstream. For example, when CVS of 120Hz, whose SOP structure is L=2 and CVS of 60Hz, whose SOP structure is L=2 are included in bitstream, re-ordering delay of the latter (ClockTick * 4) is the maximum value, and this value is considered as "display delay" in all CVS in bitstream.

— The following condition is added to Buffering period SEI which is added to the head AU of CVS after the change
  ・concatenation_flag = 1

(Note) This means HEVC decoder block calculates AuNominalRemovalTime of the head AU in CVS after the change by the value of au_cpb_removal_delay_delta_minus1.

・au_cpb_removal_delay_delta_minus1 is set to the value equivalent to the difference of decoding delay of last AU whose Temporal ID is 0 and which is neither RASL picture nor

RADL picture in immediately before CVS, and AuNominalRemovalTime is made to be continuous before and after the change of CVS.

(Note) It is necessary to follow this restriction when any parameters are changed.

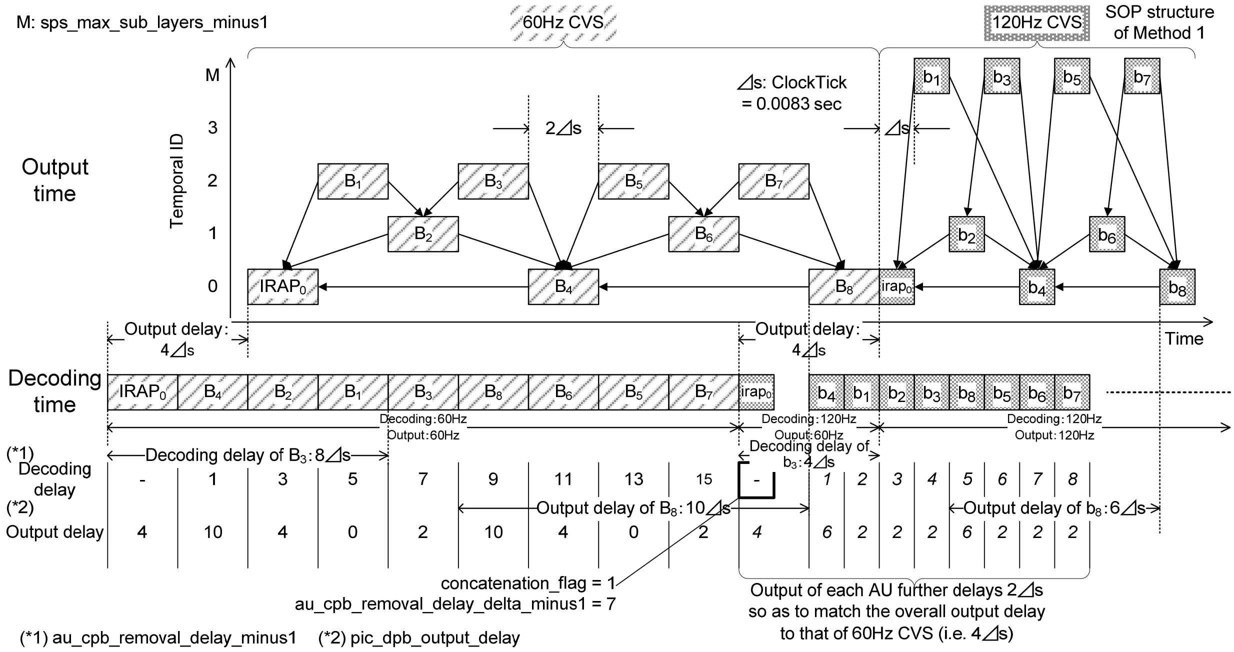
— AuNominalRemovalTime of AU (IRAP AU) in the head of the change of CVS must be equal to AuCpbRemovalTime. That is, CPB underflow must not happen.

(Note) Detail is specified by equation C-11 in HEVC Standard, Annex C.

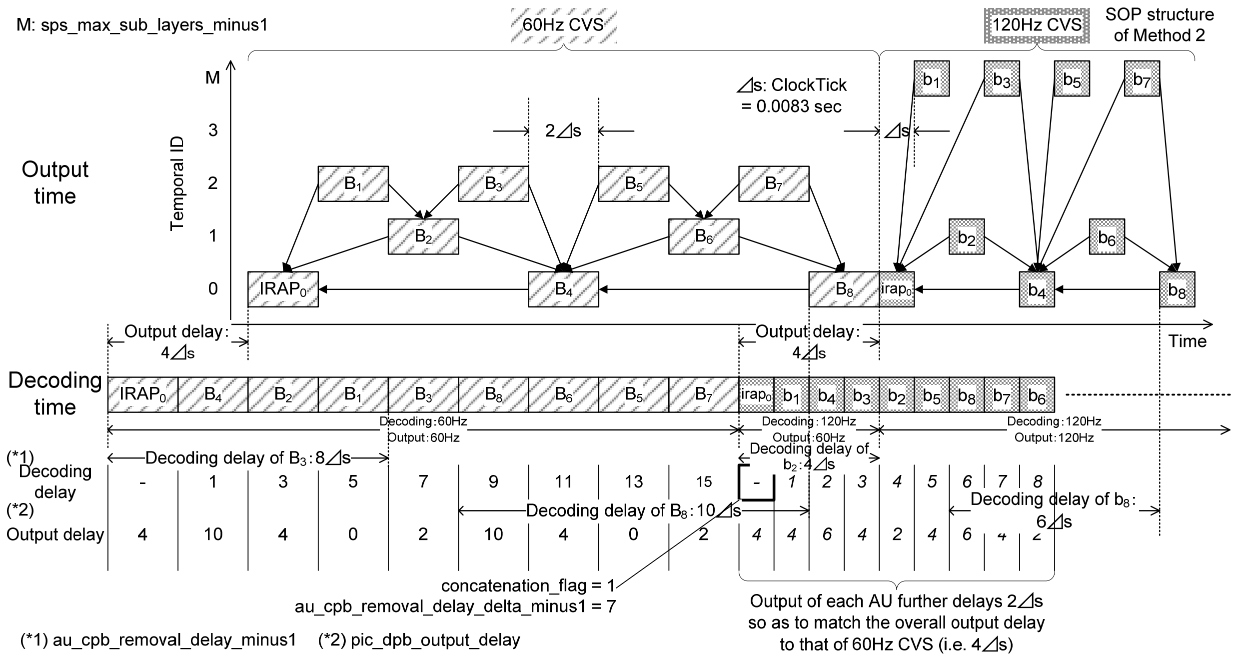
— Decode time of each AU in CVS must be continuous except the head of 120Hz CVS. Display time of each AU in CVS must be constant except the tail of 120Hz CVS.

(Note) Either Method 1 or Method 2 in Fig. 4-13 can be used by changing every CVS as SOP structure for 120Hz CVS. In case of using Method 1, decode time of head IRAP picture and decode time of successive coding picture (b_4 in case of Fig. 5-3(1), (3)) are discontinuous. In other cases, decode time shall be continuous.

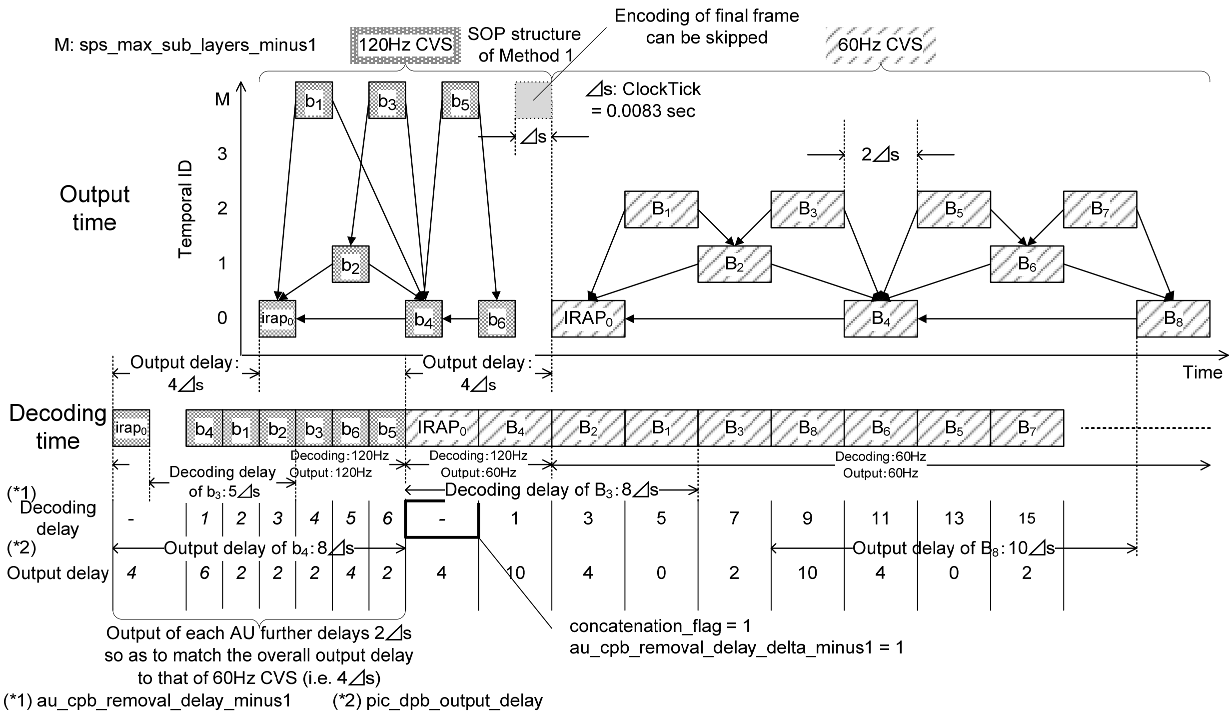
(Note) In case of using Method 1, as decode time of head AU and successive AU in CVS are discontinuous, display delay of CVS tail becomes longer than display delay of CVS head. This increase of display delay can be reduced by skipping last picture in 120Hz program without coding it (an example in Fig. 5-3(3)). But as it is a precondition that each picture in the program is coded without exception, such a coding skip is allowed to occur only at CVS tail. Also, by coding skip of last picture in the program, display time of immediately before picture of last picture in the program (b_6 in case of Fig. 5-3(3)) increases to 2Δs. It is desirable that the receiver displays repeatedly the immediately before picture of last picture in the program.



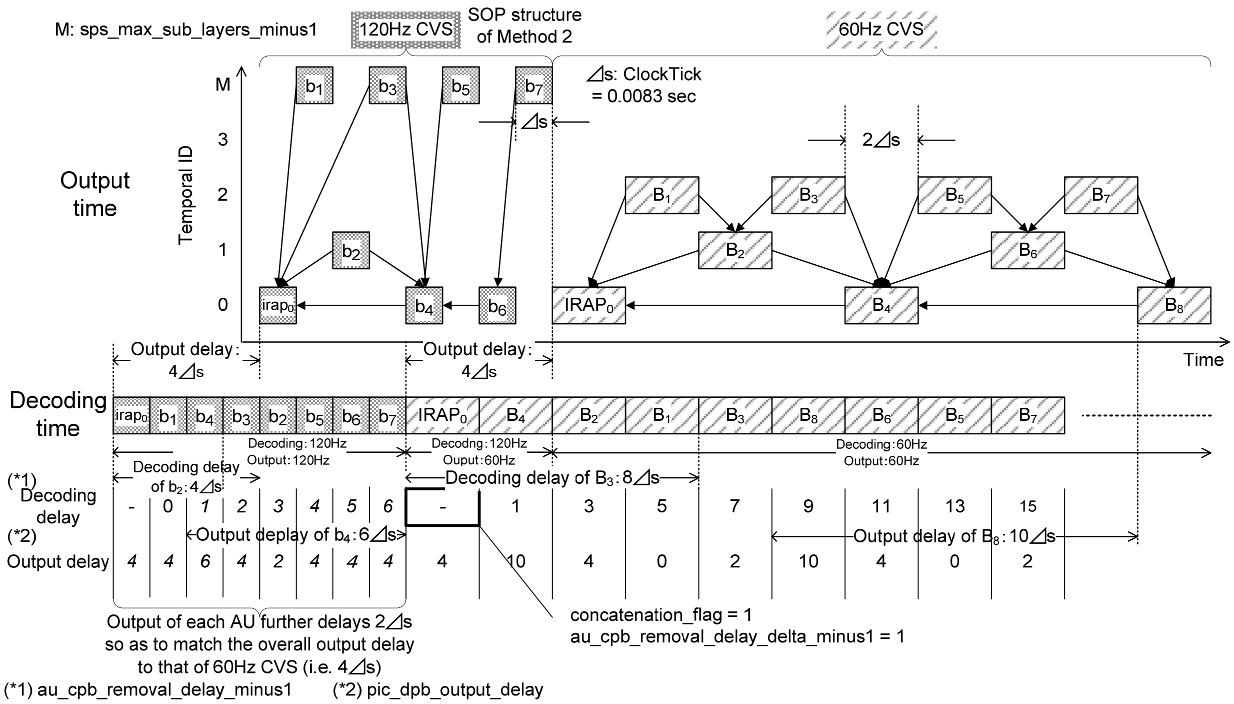
(1) Changing from 60Hz to 120Hz (ⅰ)



(2) Changing from 60Hz to 120Hz (ⅱ)



(3) Changing from 120Hz to 60Hz (ⅰ)



(4) Changing from 120Hz to 60Hz (ⅱ)

Fig. 5-3: Continuous decoding and reproduction in the boundary of CVS

5.3.2 System procedure

(1) Changing the video resolution

When the resolution of video stream (1080/I, 1080/P, 2160/P, 4320/P, etc.) is changed for one service ID, in order to realize seamless display or less incongruous display, procedure at the transmission side and receiver side will be stated.

In order to realize perfect seamless change, it is necessary that both transmitting side and receiving side can deal with seamless change. But it is supposed either transmitting side or receiving side, or both does not deal with it at the beginning of broadcasting. Even in these cases, considering both transmitting side and receiving side can transfer independently to the system in which perfect seamless change is dealt with, by the method displaying still picture or black picture and making bad appearance better, the procedure about changing the video resolution described here is recommended.

・Provision in case of multiplex by MPEG-2 Systems

Procedure by which perfect seamless change is realizable at the transmitting side and the receiving side, and simple procedure at the receiver side is shown in the following (See Fig. 5-4). Here, the case of changing from 3 programs of 4K to 1 program of 8K is explained as an example in this clause, but it can also apply to the case of changing from 3 programs of HDTV to 1 program of 4K. In changing all video resolution for one service ID, ES PID of video stream with different resolution shall use different number before and after changing.

When changing from 3 programs of 4K to 1 program of 8K or changing from 1 program of 8K to 3 programs of 4K, the broadcaster who wants continuous display must transmit PMT which specifies the same service ID and whose number is the same as 4K program in broadcasting 8K program. And he must set ES_PID of 8K program to a different value from PID of all components which are broadcasted at the beginning of transmitting new PMT. Both PMT of 4K program and 8K program shall include video_decode_control_descriptor shown in ARIB STD-B10. On the assumption that the operation is satisfied with these conditions, service ID, ES_PID of each program is set to as the followings for an example.

4K program 1 : service_id=01 ES_PID=101 → 8K program : service_id=01, ES_PID=104
4K program 2 : service_id=02 ES_PID=102 → 8K program : service_id=02, ES_PID=104
4K program 3 : service_id=03 ES_PID=103 → 8K program : service_id=03, ES_PID=104

＜Procedure on the transmitting side＞

Conditions to be satisfied by the transmitting side are shown in the following.

1. Suppose changing time for 4K/8K is T1. Also let PMT of 4K program include video_decode_control_descriptor.
2. 3 encoders for 4K program and 1 encoder for 8K program synchronize PCR and PTS (and DTS) and make PCR continuous at changing time.
3. PMT of 8K program (ES_PID＝104 is specified) starts to be transmitted 1 second before the changing time T1 (nominal value). Transmission of PMT of 4K program is not intermingled with transmission of PMT of 8K program in time-axis, and stops after starting transmission of it. Let PMT of 8K program include video_decode_control_descriptor. (Note 1)
4. As for transmission of 4K bitstream, EOS NAL is added at the end. (Note 1)
5. At changing time, multiplexing of TS for 4K program is stopped in MUX equipment, and multiplexing of TS for 8K program is started. It is desirable that after changing to 8K bitstream, VPS/SPS for 8K program is transmitted as soon as possible. Also let the head AU of HEVC ES in 8K bitstream which is changed be IRAP AU. It may be possible that null data multiplexing is necessary between EOS NAL of 4K bitstream and VPS/SPS of 8K bitstream. (Note 2)

(Note 1): About start timing for transmitting new PMT

- ✓ Considering only free TV, if new PMT is transmitted more than 0.5 second before program change time T1, receiver can respond sufficiently. As the operation at the transmitting side is usually done in the unit of regular second, it shall be standard to transmit new PMT 1 second before T1. If transmitting new PMT is started at the receiver side sometime from 0.5 to 2 seconds before, it is not worth considering.
- ✓ Considering pay TV, when there are many keys to be changed, the case would be supposed that it is not in time even if new ECM is transmitted 2 seconds before because of response time of IC card. But, if new PMT is transmitted more than 2 seconds before, it will be uncomfortable for the person who has selected channel at the time to take a long time to display picture. So, it is desirable that new PMT is transmitted sometime from 0.5 to 2 seconds before, and CAS is operated without any problem at this timing by using various ideas such as making keys common and temporary non-scramble and so on.
- ✓ video_decode_control_descriptor must be certainly corresponded to the bitstream which is transmitted after. For example, before transmitting corresponding bitstream, another video_decode_control_descriptor shall not be transmitted.

(Note 2): By controlling schedule in a unit of second at the broadcast station, timing of the tail GOP does not generally agree with the timing of a unit of second because of GOP length and frame/field frequency 59.94Hz. So the end timing and the start timing of bitstream are a little bit before or after the control timing. The gap between the end of 4K bitstream and the start of 8K bitstream must be so short that decoder buffer in the receiver will not underflow.

＜Receiver operation (in case of the receiver with the seamless decoding mode switching functionality)＞

Working flow at the receiver side is shown in the following.

1. New version PMT is acquired.
2. According to the content of PMT descriptor, if the bitstream is distinguished that EOS NAL is transmitted when changing from 4K to 8K, Demux is set up to input bitstream of ES_PID for both 4K and 8K into AV decoder at the same time. But real data is not inputted redundantly by the transmission timing. After bitstream of 4K ends, bitstream of 8K is stored in the buffer.
3. After video decoder acquires EOS NAL, it makes picture frozen and sound mute if necessary.
4. If VPS/SPS of 8K bitstream is acquired, the decoder follows automatically and decodes appropriately. When the decoder is in a state that normal video and audio can be output, it cancels to freeze picture and to mute sound.
   In order to be displayed seamlessly in appearance, it is necessary that 8K bitstream is received quickly after 4K bitstream is received so that the buffer will not underflow. In this case, frozen picture is not displayed. When the time from the end of 4K bitstream to the start of 8K bitstream is not short enough, and the buffer underflows, the picture transmitted immediately before EOS NAL is displayed in frozen.
5. When confirming 8K decode is started, ES_PID only for 8K is used which is inputted to AV decoder.

(Attention matter)

When there is a kind of receiver without seamless decoding mode switching functionality and it makes picture frozen at a time of receiving new PMT, it is desirable that video whose flicker is not so noticeable even in frozen state is transmitted more than 0.5 second (delay time in buffer) before new PMT transmission starts.

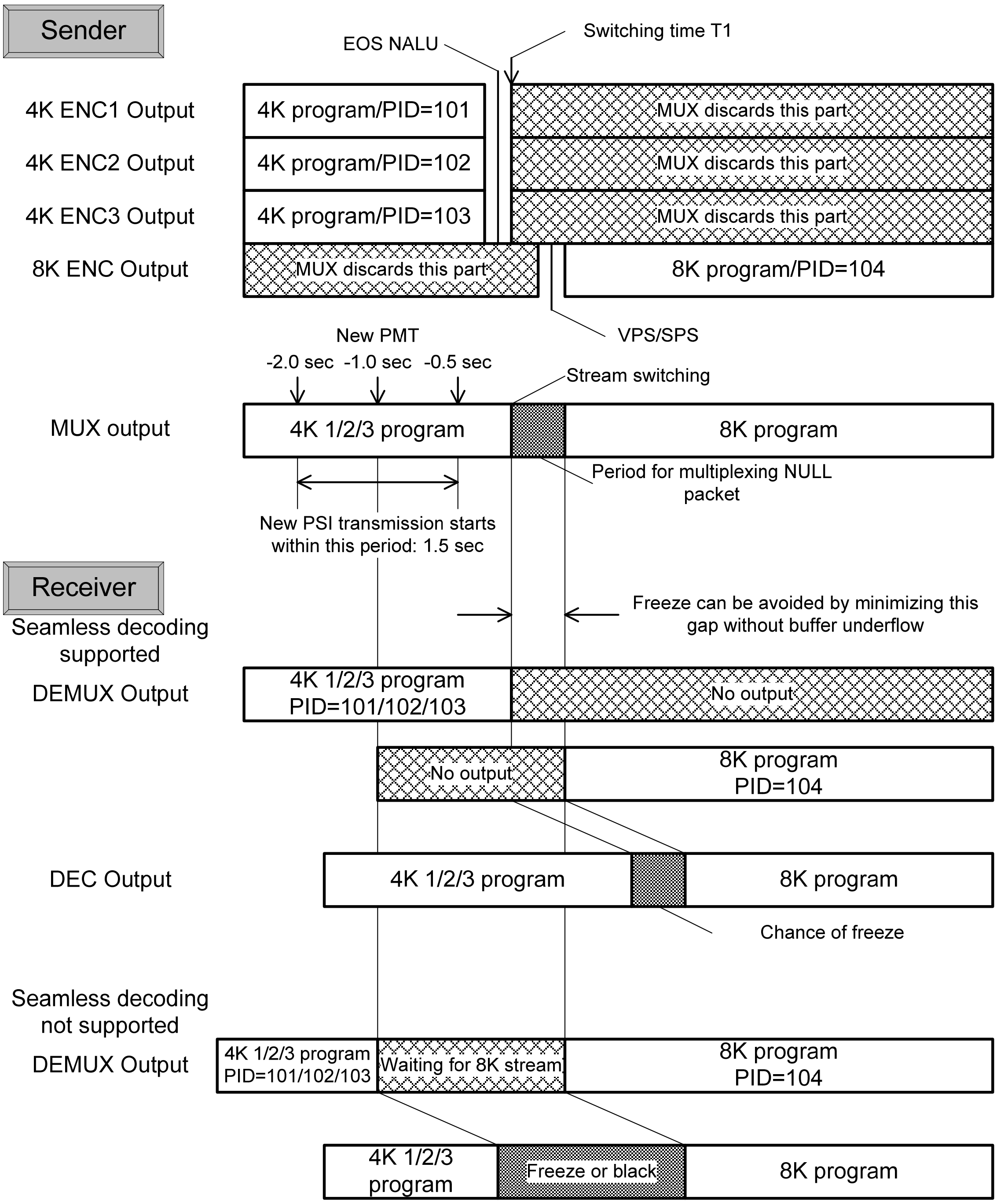


Fig. 5-4: Seamless reproduction (multiplexing by MPEG-2 Systems)

・Provisions in case of multiplex by MMT

Being the same case as multiplex by MPEG-2 Systems, the case of changing from 3 programs of 4K to 1 program of 8K is described as an example. But the case of changing from 3 programs of HDTV to 1 program of 4K can be also applied. See Fig. 5-5. In changing all video resolution for one service ID, packet ID of asset of video stream with different resolution shall use different number before and after changing.

When changing from 3 programs of 4K to 1 program of 8K or changing from 1 program of 8K to 3 programs of 4K, the broadcaster who wants continuous display must transmit MPT which specifies the same service ID as 4K program and whose number is the same as it in broadcasting 8K program, and he must set packet ID of asset of 8K program to a different value from all ID of all assets which are broadcasted at the beginning of transmitting new MPT. On the assumption that the operation is satisfied with these conditions, service ID, packet ID of video asset of each program is set to as the followings for explaining with an example.

4K program 1： service_id=01 V_Asset_PID=101 → 8K program：service_id=01, V_Asset_PID=104
4K program 2： service_id=02 V_Asset_PID=102 → 8K program：service_id=02, V_Asset_PID=104
4K program 3： service_id=03 V_Asset_PID=103 → 8K program：service_id=03, V_Asset_PID=104

＜Procedure on the transmitting side＞

Conditions to be satisfied by the transmitting side are shown in the following.

1. Suppose changing time for 4K/8K is T1. MPT of 4K program shall include Video_Component_Descriptor.
2. PMT of 8K program (ES_PID＝104 is specified) starts to be transmitted 1 second before the changing time T1 (nominal value). Transmission of PMT of 4K program is not intermingled with transmission of PMT of 8K program in time-axis, and stops after starting transmission of it.　(Note 1)
3. As for transmission of 4K bitstream, EOS NAL is added at the end. So last MPU of 4K program shall end at EOS NAL. MPT of 8K program shall include Video_Component_Descriptor.
4. At changing time, multiplexing of TS for 4K program is stopped in MUX equipment, and multiplexing of TS for 8K program is started. As for head MPU of asset of 8K video, set MPU sequence number to 0, and set head AU to IRAP AU. (Note 2)

(Note 1): About start timing for transmitting new PMT

- ✓ Considering only free TV, if new PMT is transmitted more than 0.5 second before program change time T1, receiver can respond sufficiently. It shall be standard to transmit new PMT 1 second before T1. If receiver starts transmitting new PMT some time between 0.5 to 2 seconds before, it is out of the question.
- ✓ Considering pay TV, when there are many keys to be changed, the case would supposed that it is not in time even if new ECM is transmitted 2 seconds before because of response time of IC card. But, if new PMT is transmitted more than 2 seconds before, it will be uncomfortable to take a long time to display picture. So, it is desirable that new PMT is transmitted some time between 0.5 to 2 seconds, and CAS is operated without any problem at this timing by using various ideas such as making keys common and temporary non-scramble and so on.
- ✓ Content of Video_Component_Descriptor described in MPT must be certainly corresponded to the bitstream which is transmitted after. For

example, before transmitting corresponding bitstream, another Video_Component_Descriptor shall not be transmitted.

(Note 2): By controlling schedule every second at the broadcast station, timing of tail GOP is not generally agree with timing of second unit because of GOP length and frame/field frequency 59.94Hz. So end timing and start timing of bitstream are a little before or after control timing. The difference between end of 4K bitstream and start of 8K bitstream must be as short as not to make decoder buffer in the receiver underflowed.

＜Receiver operation (in case of the receiver with the seamless decoding mode switching functionality)＞

Working flow at the receiver side is shown in the following.

1. New PMT is acquired.
2. According to the content of asset descriptor described in MPT, after distinguishing the change from 4K to 8K will occur, Demux is set up so that video assets of both 4K and 8K are inputted to AV decoder at the same time. But real data is not inputted redundantly by transmission timing. After storing 4K bitstream is completed, 8K bitstream is stored in the buffer.
3. When video decoder acquires EOS NAL of 4K video asset, it makes picture frozen and sound mute as necessity.
4. If VPS/SPS of 8K bitstream is acquired, the decoder follows automatically and decodes appropriately. When the decoder is in a state that normal video and audio can be output, it cancels frozen picture and mute sound.
   In order to be displayed seamlessly in appearance, it is necessary that 8K bitstream is received quickly after processing 4K bitstream is completed so that the buffer will not underflow. In this case, frozen picture is not displayed. When the time from the end of processing 4K bitstream to the start of processing 8K bitstream is not short enough, and the buffer underflows, the picture transmitted immediately before EOS NAL is displayed in frozen.
5. When it is confirmed that 8K decoding is started, asset which is inputted to AV decoder shall be only for 8K.

＜Receiver operation (in case of the receiver without the seamless decoding mode switching functionality)＞

Working flow at the receiver side is shown in the following.

1. New PMT is acquired.
2. According to the content of asset descriptor described in MPT, after distinguishing the change from 4K to 8K will occur, video is made to be frozen or black picture is displayed, and sound is made to be mute.
3. Video decoder stops decoding 4K.
4. Demux stops receiving 4K video asset, and bitstream which has packet ID of 8K video asset is set to be inputted to decoder buffer.
5. Monitoring VPS/SPS monitor register of video decoder by host CPU, it is waited that 8K bitstream is inputted.
6. When VPS/SPS of 8K bitstream is acquired, the decoder starts decoding 8K, and when the decoder is in a state that normal video and audio can be output, it cancels frozen or black picture of video and mute sound.

(Attention matter)

In case that there is a kind of receiver without seamless decoding mode switching functionality and it makes picture frozen at the time of receiving new MPT, it is desirable that video whose flicker is less noticeable should be transmitted more than 0.5 second (delay time in buffer) before start of transmitting new MPT.

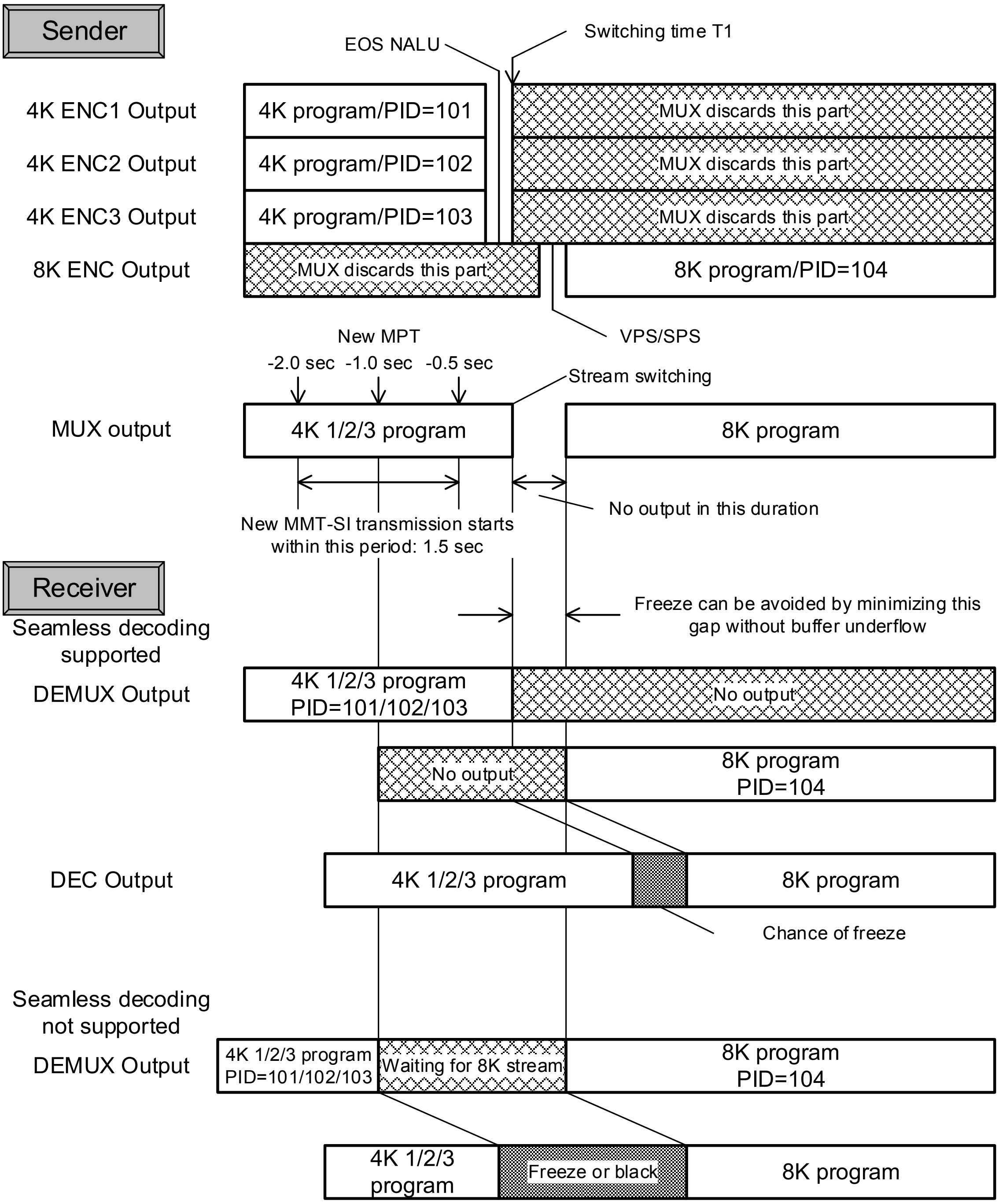


Fig. 5-5: Seamless reproduction (in MMT multiplexing)

(2) Changing frame rate

Procedure at the transmitting side and receiving side will be mentioned, when the frame rate of video bitstream is changed from 60Hz (or 60/1.001Hz) to 120Hz (or 120/1.001Hz) , or changed in the reverse for a service ID.

When the receiver only for 60Hz receives 120Hz video bitstream which is scalable coded, it reads and throws 120Hz subset and decodes only 60Hz sub-bitstream and display. So, when frame rate of video bitstream is changed, PID (ES_PID in case of MPEG-2 Systems multiplexing, and V_Asset_PID in case of MMT multiplexing) of 60Hz sub-bitstream shall be agreed with PID of 60Hz bitstream, in order to make decoded video output of 60Hz receiver seamless every possibility.

There are two cases supposed at the transmitting side. Regarding HEVC coding block of 120Hz coding and HEVC coding block of 60Hz coding, one case is that the same block is used, and the other is that the different blocks are used. In the former case, it is easy to make PID agree with. Considering these, if the transmission side can deal with changing frame rate, procedure mentioned here is recommended.

Also, in case of changing video resolution and frame rate at the same time, the procedure of changing video resolution shall be applied. That is, PID of 60Hz sub-bitstream shall be different value from PID of 60Hz bitstream.

・Provision in case of multiplex by MPEG-2 Systems

When changing from 60Hz (or 60/1.001Hz) program to 120Hz (or 120/1.001Hz) program, or changing in the reverse, broadcast station who wants to display continuously must transmit PMT specifying the same service ID in broadcasting 120Hz as in broadcasting 60Hz. Both PMT in 60Hz program and PMT in 120Hz program shall include video_decode_control_descriptor and hierarchy_descriptor shown in ARIB STD-B10. On condition that the operation which satisfies these conditions is performed, service ID and ES_PID of each program are set to as the following by explanation.

60Hz program (60Hz bitstream) service_id=01　ES_PID=101	⇔	120Hz program (60Hz sub-bitstream) service_id=01, ES_PID=101 120Hz program (120Hz subset) service_id=01, ES_PID=102

Further, the notation of aaaa(bbbb) hereafter means that “aaaa” is applied in changing 60Hz/120Hz and “bbbb” is applied in changing 120Hz/60Hz, respectively. Fig. 5-6 shows the movement in changing 120Hz/60Hz after changing 60Hz/120Hz.

＜Procedure on the transmitting side＞

Conditions to be satisfied by the transmitting side are shown in the following.

1. Suppose changing time for 60Hz/120Hz (120Hz/60Hz) is time T1. Let PMT of 60Hz program (120Hz program) include video_decode_control_descriptor.
2. Encoder makes PTS/DTS and PCR continuous before and after the change.
3. PMT of 120Hz program (60Hz program) starts to be transmitted from 1 second before changing time T1. Transmission of PMT of 60Hz program (120Hz program) is not intermingled with transmission of PMT of 120Hz program (60Hz program) in time-axis, and stops after starting transmission of it. PMT of 120Hz program (60Hz program) shall include video_decode_control_descriptor and hierarchy_descriptor. Timing for transmitting new PMT shall be equivalent to that for changing video resolution.
4. When transmitting 60Hz bitstream (120Hz scalable coding bitstream), EOS NAL is added at the end.
5. If it is time to change, multiplex of TS for 60Hz bitstream (120Hz scalable coding bitstream) stops in MUX equipment, and multiplex of TS for 120Hz scalable coding bitstream (60Hz bitstream) starts. After changing to 120Hz scalable coding bitstream (60Hz bitstream), it is desirable VPS/SPS of 120Hz scalable coding bitstream (60Hz bitstream) is transmitted as soon as possible. Also, in changed 120Hz scalable coding bitstream (60Hz bitstream), the head AU shall be IRAP AU. Multiplex of null data may be necessary between EOS NAL of 60Hz bitstream (120Hz scalable coding bitstream) and VPS/SPS of 120Hz scalable coding bitstream (60Hz bitstream). (Note 1)

(Note 1): By controlling schedule every second in the broadcast station, the timing of tail GOP does not generally agree with the timing of second unit because of GOP length and frame/field frequency 60Hz. So the end timing and the start timing of bitstream deviate a little from the control timing. The difference between the end of 60Hz bitstream (120Hz scalable coding bitstream) and the start of 120Hz scalable coding bitstream (60Hz bitstream) must be so little enough as decoder buffer in the receiver does not underflow. Also, in case of changing from 120Hz scalable coding bitstream to 60Hz bitstream, the number of frame which is coded in 120Hz scalable coding bitstream must be a multiple of 2.

＜Receiver operation＞

Working flow on the receiver side is shown in the following.

1. New version PMT is acquired.
2. According to the content of PMT descriptor (video_decode_control_descriptor and hierarchy_descriptor), after distinguishing the bitstream is that EOS NAL is transmitted by changing from 60Hz program to 120Hz program (from 120Hz program to 60Hz program), Demux is set up so that 60Hz bitstream, 60Hz sub-bitstream, and stream of ES_PID of 120Hz subset are inputted to AV decoder at the same time. But 60Hz bitstream and 60Hz sub-bitstream are not inputted redundantly by transmitting timing, and after 60Hz bitstream data (60Hz sub-bitstream data) stops to be inputted, 60Hz sub-bitstream data

(60Hz bitstream data) is stored in buffer.

3. When video decoder acquires EOS NAL, it makes picture frozen and sound mute as necessity.
4. If VPS/SPS of 60Hz sub-bitstream (60Hz bitstream) is acquired, the decoder follows automatically and decodes appropriately. When the decoder is in a state that normal video and audio can be output, it cancels frozen picture and mute sound.
   In order to be displayed seamlessly in appearance, it is necessary that 60Hz sub-bitstream (60Hz bitstream) is received quickly after 60Hzbitstream (60Hz sub-bitstream) stops so that the buffer does not underflow. In this case, frozen picture is not displayed. When the time from the end of 60Hz bitstream (60Hz sub- bitstream) to the start of 60Hz sub-bitstream (60Hz bitstream) is not short enough, and the buffer underflows, the picture transmitted immediately before EOS NAL is displayed in frozen.
5. In case of changing from 120Hz program to 60Hz program, when confirming decode of 60Hz bitstream is started, ES_PID to be inputted to AV decoder shall be only for 60Hz bitstream.

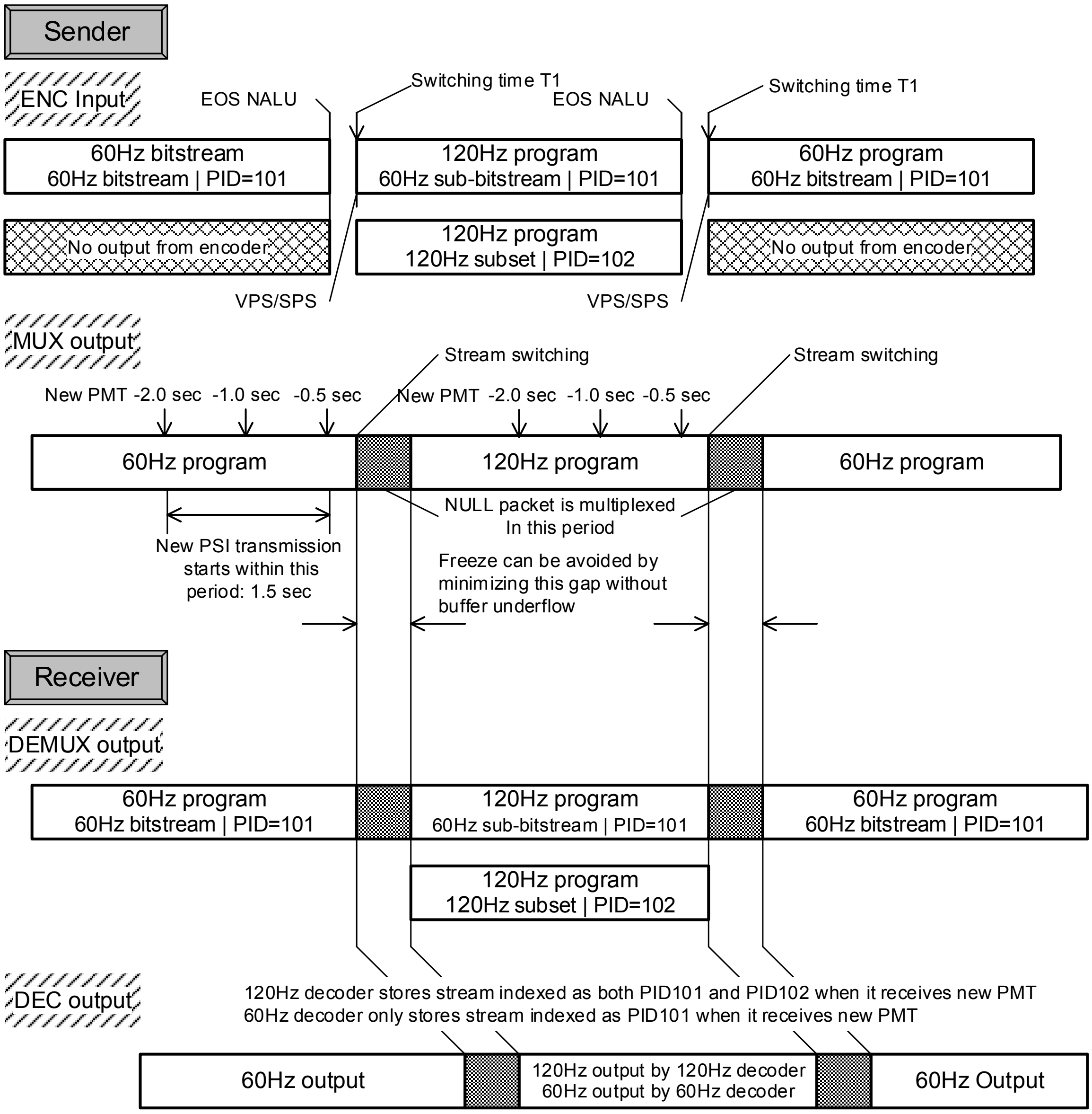


Fig. 5-6 Changing frame rate (in MPEG-2 Systems multiplexing)

・Provision in case of multiplex by MMT

When changing from 60Hz (or 60/1.001Hz) program to 120Hz (or 120/1.001Hz) program, or changing in the reverse, broadcast station who wants to display continuously shall transmit PMT specifying the same service ID in broadcasting 120Hz as in broadcasting 60Hz. Both MPT of 60Hz program and MPT of 120Hz program shall include Video_Component_Descriptor and MH-Hiarachy_Descriptor written in ARIB STD-B60. On condition that the operation which satisfies these conditions is performed, service ID of each program and packet ID of video asset are set to as the following by explanation.

60Hz program (60Hz bitstream) service_id=01 V_Asset_PID=101	⇔	120Hz program (60Hz sub-bitstream) service_id=01, V_Asset_PID=101 120Hz program (120Hz subset) service_id=01, V_Asset_PID=102

Further, the notation of aaaa(bbbb) hereafter means that “aaaa” is applied in changing 60Hz/120Hz and “bbbb” is applied in changing 120Hz/60Hz, respectively. Fig. 5-7 shows the movement in case of changing 120Hz/60Hz after changing 60Hz/120Hz.

＜Procedure on the transmitting side＞

Conditions to be satisfied by the transmitting side are shown in the following.

1. Suppose changing time for 60Hz/120Hz (120Hz/60Hz) is time T1. MPT of 60Hz program (120Hz program) shall include Video_Component_Descriptor.
2. MPT of 120Hz program (60Hz program) starts to be transmitted from 1 second (nominal value) before changing time T1. Transmission of MPT of 60Hz program (120Hz program) is not intermingled with transmission of MPT of 120Hz program (60Hz program) in time-axis, and stops after starting transmission of it.
3. On transmitting 60Hz bitstream (120Hz scalable coding bitstream), EOS NAL is added at the end. MPT of 120Hz program (60Hz program) shall include Video_Component_Descriptor and MH-Hierarchy_Descriptor.
4. If it is time to change, multiplex of asset for 60Hz video (120Hz video) stops in MUX equipment, and multiplex of asset for 120Hz video (60Hz video) starts. After changing to 120Hz scalable coding bitstream (60Hz bitstream), it is desirable VPS/SPS of 120Hz scalable coding bitstream (60Hz bitstream) is transmitted as soon as possible. Also, in changed 120Hz scalable coding bitstream (60Hz bitstream), the head AU shall be IRAP AU. (Note 1)

(Note 1): By controlling schedule every second at the broadcast station, the timing of tail GOP does not generally agree with the timing of second unit because of GOP length and frame/field frequency 60Hz. So the end timing and the start timing of bitstream deviate a little from control timing. The difference between the end of 60Hz bitstream (120Hz scalable coding bitstream) and start of 120Hz scalable coding bitstream (60Hz bitstream) must be so little enough as decoder buffer in the receiver does not underflow. Also, in case of changing from 120Hz scalable coding bitstream to 60Hz bitstream, the number of frame which is coded in 120Hz scalable coding bitstream must be a multiple of 2.

＜Receiver operation＞

Working flow on the receiver side is shown in the following.

1. New MPT is acquired.
2. According to the content of asset descriptor described in MPT (Video_Component_Descriptor and MH-Hierarchy_Descriptor), after distinguishing that changing from 60Hz program to 120Hz program (from 120Hz program to 60Hz program) will occur, Demux is set up so that 60Hz bitstream, 60Hz sub-bitstream, and the asset equivalent to 120Hz subset are inputted to AV decoder at the same time. But 60Hz bitstream and 60Hz sub-bitstream are not inputted redundantly by transmitting timing, and after 60Hz bitstream data (60Hz sub-bitstream data) stops to be inputted, 60Hz sub-bitstream data (60Hz bitstream data) is stored in buffer.
3. When video decoder acquires EOS NAL of 60Hz sub-bitstream (60Hz bitstream), it makes picture frozen and sound mute as necessity.
4. If VPS/SPS of 60Hz sub-bitstream (60Hz bitstream) is acquired, the decoder follows automatically and decodes appropriately. When the decoder is in a state that normal video and audio can be output, it cancels frozen picture and mute sound.
   In order to be displayed seamlessly in appearance, it is necessary that 60Hz sub-bitstream (60Hz bitstream) is received quickly after 60Hzbitstream (60Hz sub-bitstream) stops so that the buffer does not underflow. In this case, frozen picture is not displayed. When the time from the end of 60Hz bitstream (60Hz sub- bitstream) to the start of 60Hz sub-bitstream (60Hz bitstream) is not short enough, and the buffer underflows, the picture transmitted immediately before EOS NAL is displayed in frozen.
5. In case of changing from 120Hz program to 60Hz program, when confirming decode of 60Hz bitstream is started, asset to be inputted to AV decoder shall be only for 60Hz bitstream.

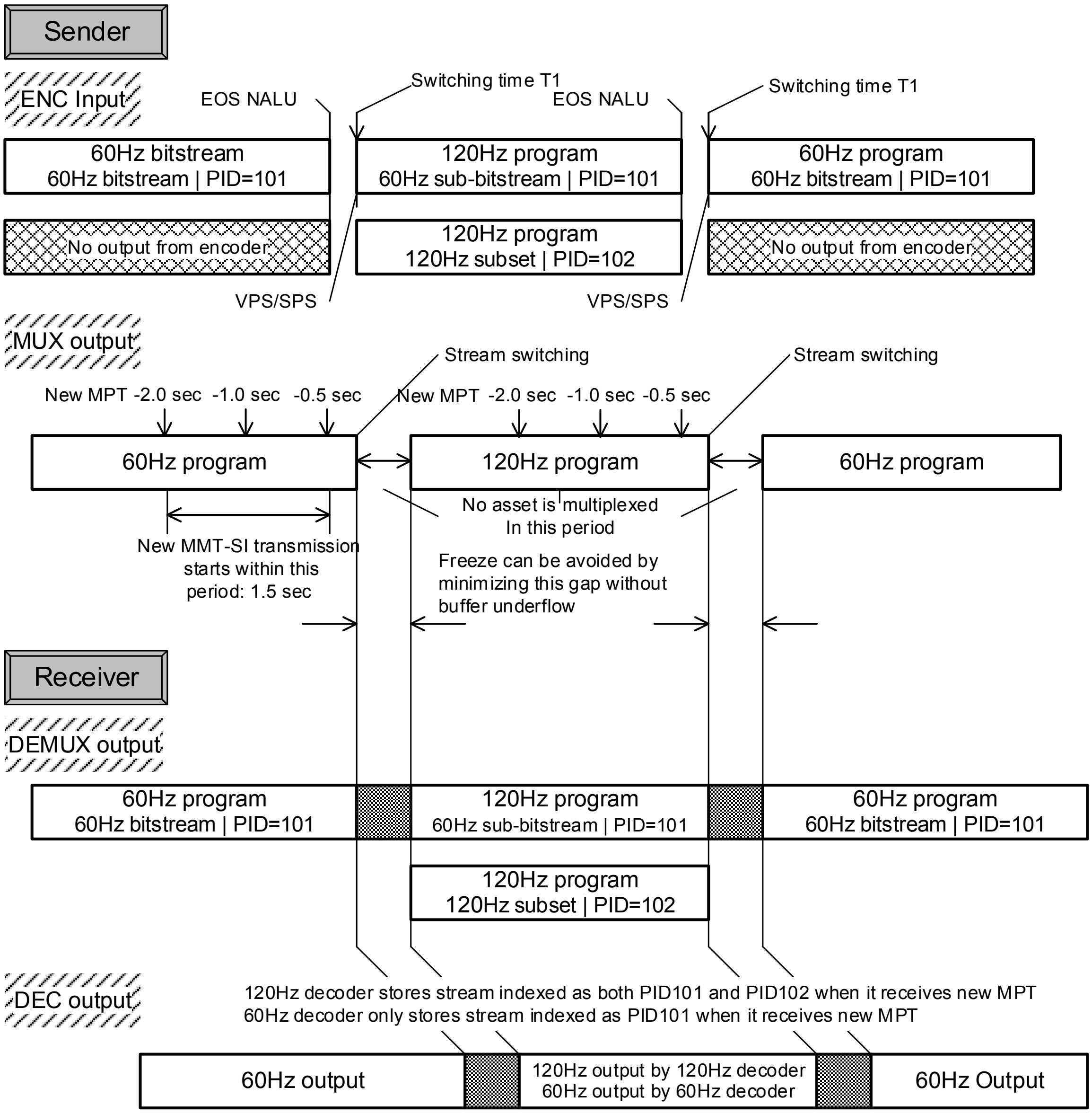


Fig. 5-7: Changing frame rate (in MMT multiplexing)

(3) Changing the video transfer characteristics

When the transfer characteristics of display video (SDR, HDR) is changed for one service ID, procedure at the transmission side and the receiver side will be stated for realizing seamless display or less incongruous display,.

In order to realize seamless change, it is necessary that both transmitting side and receiving side can deal with seamless change. At the transmitting side, as the difference of SDR signal and HDR signal is small from a view of coding, and it is possible to transmit both signals seamlessly on the same channel, the possibility of realizing seamless switching is high.

On the other hand, at the receiving side, though decoding process is no matter, it is supposed that the display does not handle seamless switching. Even in such a case, the procedure mentioned here is recommended about changing the video transfer characteristics, as the method that can reduce artifact in appearance small at the time of switching.

・Provision in case of multiplex by MPEG-2 Systems

Procedure by which seamless change is realizable at the transmitting side and the receiving side is shown in the following. (See Fig. 5-8.) In changing all video transfer characteristics for one service ID, the same number shall be used for ES PID of video stream with different video transfer characteristics before and after changing.

At the time of switching between SDR program and HDR program or at the time of switching between HDR programs which have different values of transfer_characteristics (hereinafter generally called as the time of switching transfer characteristics), both PMTs of program before and after switching shall include video_decode_control_descriptor shown in ARIB STD-B10 respectively.

In this provision, it is a precondition that a stream becomes continuous for both ES and multiplex at the time of switching transfer characteristics, and the transfer characteristics is switched with the frame precision for video input to the encoder.

＜Procedure on the transmitting side＞

Conditions to be satisfied at the transmitting side are shown in the following.

$\Delta t1$ is equivalent to the time that is necessary for the receiver to receive PMT surely, and suppose transmitting interval of PMT is 0.1 second at the maximum, it shall be twice, that is 0.2 second.

$\Delta t2$ is equivalent to the time till the movement becomes stable after the display switches the transfer characteristics, and it shall be 0.5 second.

1. Suppose changing time for transfer characteristics is T1.
2. PMT of the program after change (new PMT) starts to be transmitted from $\Delta t1$ before changing time T1. Transmission of PMT of the program before change (old PMT) is not intermingled with transmission of <u>new</u> PMT on the time-axis, and stops after starting it. The arriving time of new PMT is the time of changing transfer characteristics. (Note 1) From just before new PMT is transmitted, video of the program is made to be visually no matter, even if it is displayed by different transfer characteristics from described in <u>new</u> PMT, such as a black screen. (Note 2)
3. As for transmission of the program before changing, EOS NAL is added at the end.
4. Video of the program after changing is returned to usual at the time ($T1+\Delta t2$).

(Note1): About start timing for transmitting new PMT

✓ If new PMT is received more than $\Delta t2$ before the time ($T1+\Delta t2$) to start normal display of the program after changing (that is before the time T1), the receiver can sufficiently respond the seamless change. It is desirable to transmit by changing from old PMT to new PMT at the time (T1-Δt1) so that the receiver can surely receive new PMT before the time T1.

- ✓ For the time from (T1-Δt1) to T1, the transfer characteristics described by PMT and video ES happen to disagree. During this interval, video which is no matter for displaying with any movement mode of the display is coded.
- ✓ video_decode_control_descriptor must certainly correspond to bitstream which is transmitted after. For example, before transmitting corresponding bitstream, another video_decode_control_descriptor shall not to be transmitted.

(Note2): In case that transfer characteristics of the program and resolution or frame rate are changed at the same time, PID is also changed. Even in such a case, the picture with no matter for displaying is coded from the time (T1-Δt1), when the display is in working mode of any transfer characteristics.

＜Receiver operation＞

Working flow at the receiver side is shown in the following.

1. New PMT is detected.
2. If the video decoder displays decoded screen of AU just before EOS NAL, it changes transfer characteristics of the video display output. Then, the display changes the working mode, and it is desirable that changing time is shorter than Δt2. Also, the display may independently display by reducing the turbulence of video until the screen becomes stable after change. As the coded video arrives continuously after EOS NAL, decoding continues. (Note 1)
3. In case of displaying independently after change, if sufficient time passes to make display stable, output of display is changed to decoded video. Timing for display follows the time described in stream.

(Note1): According to the timing of switching channel in the receiver, there is some possibility that the transfer characteristics between PMT which is received first and video ES which is received immediately after the PMT happen to be different. Even in such a case, the difference disappears after Δt1 at the latest.

In case of the receiver which changes transfer characteristics by watching information in system layer, without watching EOS NAL of video ES, the operation is as the following.

1. PMT of new version is detected.
2. If the receiver receives new PMT, it changes transfer characteristics for video display output. Then, the display changes working mode, and it is desirable that the changing time is shorter than Δt2. Also, the display may independently display by reducing the turbulence of video until the screen becomes stable after change.
3. In case of displaying independently after change, if sufficient time passes to make display stable, output of display is changed to decoded video. Timing for display follows the time described in stream.

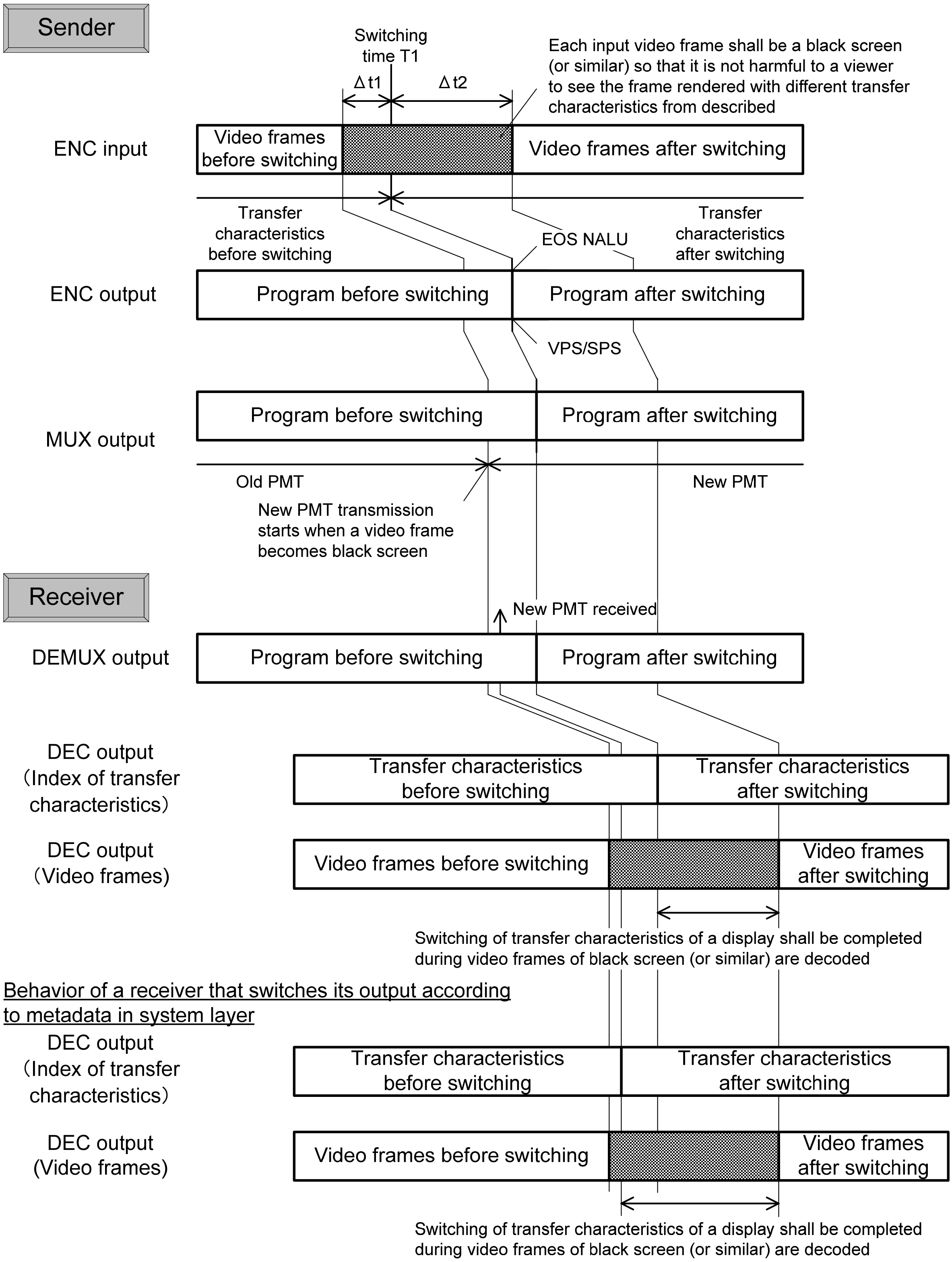


Fig. 5-8: Changing transfer characteristics (in MPEG-2 Systems multiplexing)

・Provisions in case of multiplex by MMT

In the same way as the case for multiplex by MPEG-2 Systems, in changing all video transfer characteristics for one service ID, for packet ID of asset of video stream with different video transfer characteristics, the same number shall be used before and after changing. See Fig. 5-9.

At the time of switching between SDR program and HDR program or at the time of switching between HDR programs which have different values of transfer_characteristics (hereinafter generally called as the time of switching transfer characteristics), MPT of program before and after switching shall designate the same service ID.

＜Procedure on the transmitting side＞

Conditions to be satisfied at the transmitting side are shown in the following.

Δt1 is equivalent to the time that is necessary for the receiver to receive MPT surely, and suppose transmitting interval of MPT is maximum 0.1 second, it shall be twice, that is 0.2 second.

Δt2 is equivalent to the time till the movement becomes stable after the display switches the transfer characteristics, and shall be 0.5 second.

1. Suppose changing time for transfer characteristics is T1.
2. MPT of the program after change (new MPT) starts to be transmitted from Δt1 before changing time T1. Transmission of MPT of the program before change (old PMT) is not intermingled with transmission of new MPT on time-axis, and stops after starting it. The arriving time of new MPT is the time of changing transfer characteristics. (Note 1)
   From just before new MPT is transmitted, video of the program is made to be visually no matter, even if it is displayed by different transfer characteristics from described in new MPT, such as a black screen. (Note 2)
3. As for transmission of the program before changing, EOS NAL is added at the end. Final MPU before changing shall stop at EOS NAL.
4. Video of the program after changing is returned to usual at the time (T1+Δt2).

(Note1): About start timing for transmitting new MPT

- ✓ If new MPT is received more than Δt2 before the time (T1+Δt2) to start normal display of the program after changing (that is before the time T1), the receiver can sufficiently respond the seamless change. It is desirable to transmit by changing from old MPT to new MPT at the time (T1-Δt1) so that the receiver can surely receive new MPT before the time T1.
- ✓ For the time from (T1-Δt1) to T1, the transfer characteristics described by MPT and MPU happen to disagree. During this interval, video which is no matter for displaying with any working mode of the display is coded.
- ✓ The content of Video_Component_Descriptor described in MPT must certainly correspond to bitstream which is transmitted after. For example,

before transmitting the corresponding, another Video_Component_Descriptor shall not to be transmitted.

(Note2): In case that transfer characteristics of the program and resolution or frame rate are changed at the same time, PID is also changed. Even in such a case, the picture with no matter for displaying is coded from the time (T1-Δt1), when the display is in working mode of any transfer characteristics.

＜Receiver operation＞

Working flow at the receiver side is shown in the following.

1. MPT of new version is acquired.
2. If the video decoder displays decoded screen of AU just before EOS NAL, it changes transfer characteristics of video display output. Then, the display changes the working mode, and it is desirable that changing time is shorter than Δt2. Also, the display may independently display by reducing the turbulence of video until the screen becomes stable. As the decoded video arrives continuously after EOS NAL, decoding continues. (Note 1)
3. In case of displaying independently after change, if sufficient time passes to make display stable, output of display is changed to decoded video. Timing for display follows the time described in stream.

(Note1): According to the timing of changing channel in the receiver, there is some possibility that the transfer characteristics between MPT which is received first and MPU which is received immediately after the MPT happen to be different. Even in such a case, the difference disappears after Δt1 at the latest.

In case of the receiver which changes transfer characteristics by watching information in system layer, without watching EOS NAL in MPU, the operation is as the following.

1. MPT of new version is detected.
2. If the receiver receives new MPT, it changes transfer characteristics for video display output. Then, the display changes the working mode, and it is desirable that changing time is shorter than Δt2. Also, the display may independently display by reducing the turbulence of video until the screen becomes stable after change.
3. In case of displaying independently after change, if sufficient time passes to make display stable, output of display is changed to decoded video. Timing for display follows the time described in stream.

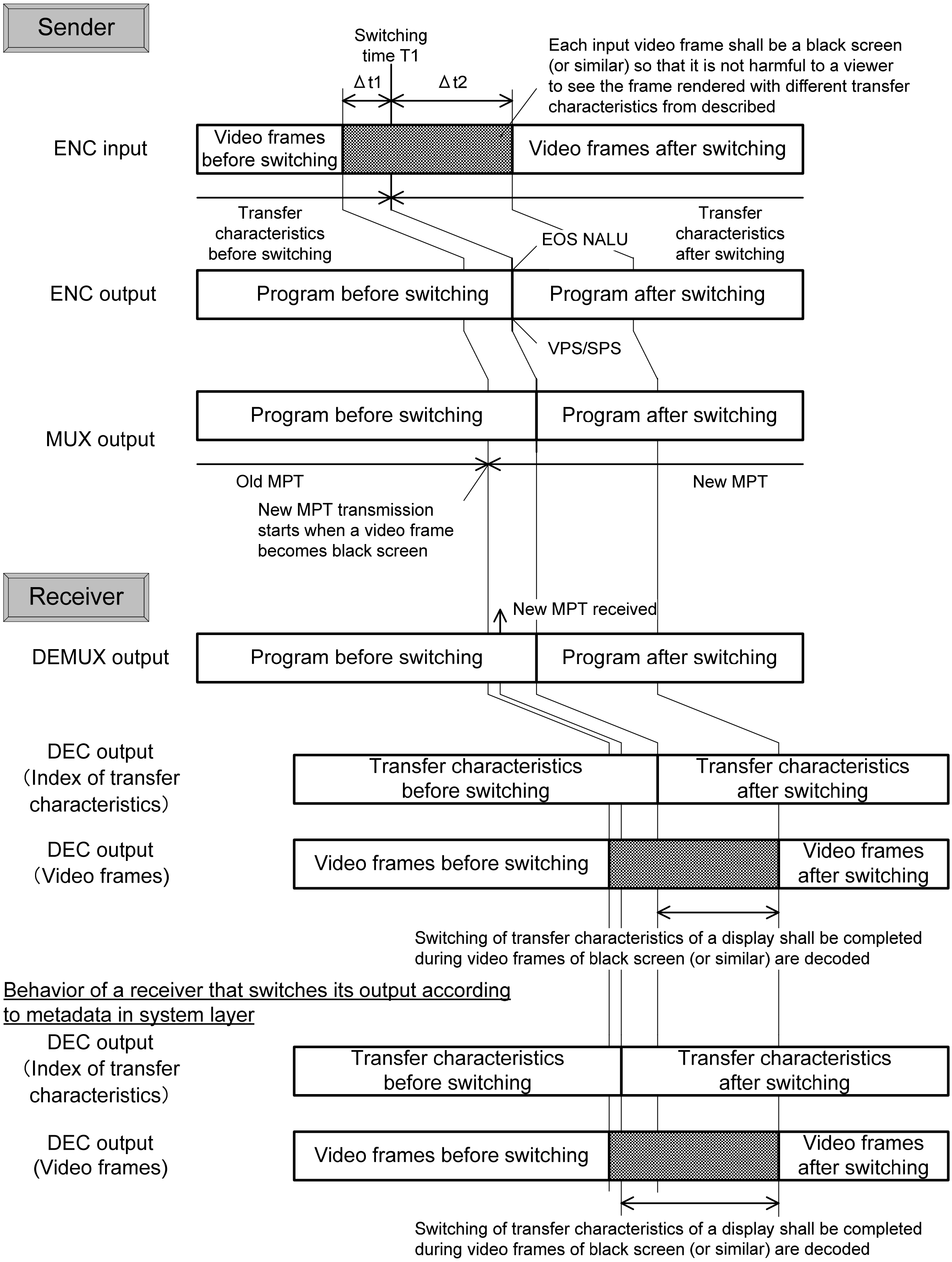


Fig. 5-9: Changing transfer characteristics (in MMT multiplexing)

Attachment 6: Operational Guidelines for HEVC Standard on low definition video services

Chapter 1: General Terms

1.1 Objective

The purpose of these operational guidelines is to present technical recommendations of HEVC Standard in practical operation concerning to video signals and video coding systems for low definition video services. (See main clause, Chapter 5, 5.2.)

1.2 Scope

These operating guidelines apply to video signals based on HEVC Standard among the video signals in low definition video services transmitted by digital broadcasting that comply with the "Standard transmission system for digital broadcasting among standard television broadcasting and the like" (Ordinance).

1.3 References

1.3.1 Normative References

(1) Rec. ITU-T H.265 (2013) | ISO/IEC 23008-2:2013 - Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 2: High efficiency video coding (hereinafter referred to as "HEVC Standard")

Chapter 2: Restrictions on coding parameters

These restrictions are applied to the restrictions provided in Attachment 5: "Operational guidelines for HEVC Standard in television services" by addition.

2.1 Profile, Level and Tier

Video coding system shall be based on Main profile or Main 10 profile and Main tier provided in HEVC Standard. When the number of coding pixel bit is 10, Main 10 profile is applied, and when the number of coding pixel bit is 8, Main profile is applied, respectively. Level shall be any of 2, 2.1, 3, 3.1, 4 and 4.1.

Maximum picture size and frame rate (number of pixels per unit time) are provided according to Level, and it is desirable that level to be operated and video coding format are decided considering format of resources, display device of the receiver and the processing and so on.

2.2 Video format and corresponding parameters

2.2.1 Assumed video format

Assumed video format and corresponding parameters are shown in Table 2-1.

Table 2-1: Assumed video format

Picture size (Horizontal×Vertical)	Picture aspect ratio (Horizontal : Vertical)	Picture coding system	pic_width_in_luma_samples	pic_height_in_luma_samples	aspect_ratio_idc
320×180※	16:9	frame	320	184	1
320×240	4:3	frame	320	240	1
352×240	4:3	frame	352	240	3
352×240	16:9	frame	352	240	5
640×480	4:3	frame	640	480	1
640×480	4:3	field	640	240	1
720×480	4:3	field	720	240	3
720×480	16:9	field	720	240	5
720×480	4:3	frame	720	480	3

720×480	16:9	frame	720	480	5
1280×720	16:9	frame	1280	720	1
1440×1080※	16:9	field	1440	544	14
1440×1080	16:9	frame	1440	1080	14

※ When vertical number is not divided by 8 without remainder, the encoder adds fictional video data (dummy data) below the active line and lines of video data with a multiple of 8 are coded. The decoder discards the dummy data, outputting only active lines.

2.2.2 Frame rate

Frame rate shall be integer times of 1/1.001 or 1/27000000, calculating as frame rate = time_scale/num_units_in_tick by using variables of VUI parameters. But maximum frame rate [Hz] in each level shall be as shown in Table 2-2 for the video format in operation.

Table 2-2: Maximum frame rate at each level [Hz]

Picture size	Level 1	Level 2	Level 2.1	Level 3	Level 3.1	Level 4	Level 4.1
320×240	-	-	60.00	60.00	60.00	60.00	60.00
320×180	-	60.00	60.00	60.00	60.00	60.00	60.00
352×240	-	-	60.00	60.00	60.00	60.00	60.00
640×480	-	-	-	30.00	60.00	60.00	60.00
720×480	-	-	-	30.00	60.00	60.00	60.00
1280×720	-	-	-	-	-	60.00	60.00
1440×1080	-	-	-	-	-	30.00	60.00

2.3 Syntax

This shall be the value shown in Attachment 5 "Operational guidelines for HEVC standard in television service," Chapter 4.2. But syntax elements described in Table 2-3 to Table 2-7 shall be the values shown in these tables.

Table 2-3: Profile level (Profile, Tier and Level)

Syntax element	Value	Remarks
general_progressive_source_flag	0, 1	0: interlaced 1: progressive
general_interlaced_source_flag	0, 1	0: progressive 1: interlaced
general_frame_only_constraint_flag	0, 1	0: interlaced 1: progressive
general_level_idc	60, 63, 90, 93,	60: Level 2

Syntax element	Value	Remarks
	120, 123	63: Level 2.1 90: Level 3 93: Level 3.1 120: Level 4 123: Level 4.1 (Note) See Table 2-2 about combination of video format

Table 2-4: Sequence parameter set (SPS)

Syntax element	Value	Remarks
pic_width_in_luma_samples pic_height_in_luma_samples	See Table 2-6	MinCbSizeY(=8) times
conformance_window_flag conf_win_left_offset conf_win_right_offset conf_win_top_offset conf_win_bottom_offset	See table 2-6	

Table 2-5: VUI parameters

Syntax element	Value	Remarks
aspect_ratio_idc	See Table 2-1	Sample aspect ratio
video_format	0	Component
colour_primaries	1	1: Rec. ITU-R BT.709, IEC 61966-2-4
transfer_characteristics	1, 11	1: Rec. ITU-R BT.709 (HDTV conventional color gamut) 11: IEC 61966-2-4 (HDTV wide color gamut)
matrix_coefficients	1	1: Rec. ITU-R BT.709, IEC 61966-2-4
chroma_loc_info_present_flag	1	The position of 4:2:0 color difference signals is described.
chroma_sample_loc_type_top_field chroma_sample_loc_type_bottom_field	0, 2	0: This lies on the middle of 2 luminance lines in vertical direction (in case of interlace) 2: This agrees with luminance line in vertical direction (in case of progressive)
field_seq_flag	See Table 2-6	
vui_num_units_in_tick vui_time_scale	See Table 2-7	Either “case 1” or “case 2” is applied.

Table 2-6: Parameters representing picture size

Input video format (“i” represents interlace)	field_seq_flag	general_progressive_source_flag	general_interlace_source_flag	pic_width_in_luma_samples	pic_height_in_luma_samples	conformance_window_flag	conf_win_left_offset	conf_win_right_offset	conf_win_top_offset	conf_win_bottom_offset
320×180	0	1	0	320	184	1	0	0	0	2
320×240	0	1	0	320	240	0	0	0	0	0
352×240	0	1	0	352	240	0	0	0	0	0
640×480i	1	0	1	640	240	0	0	0	0	0
640×480i (Note1)	0	0	1	640	480	0	0	0	0	0
640×480	0	1	0	640	480	0	0	0	0	0
720×480i	1	0	1	720	240	0	0	0	0	0
720×480i (Note1)	0	0	1	720	480	0	0	0	0	0
720×480	0	1	0	720	240	0	0	0	0	0
1280×720	0	1	0	1280	720	0	0	0	0	0
1440×1080i	1	0	1	1440	544	1	0	0	0	2
1440×1080i (Note1)	0	0	1	1440	1080	0	0	0	0	0
1440×1080	0	1	0	1440	1080	0	0	0	0	0

(Note 1) In case of coding in frame unit

Table 2-7: Time scale

Frame/field frequency and scanning	vui_time_scale		vui_num_units_in_tick	
	Case 1	Case 2	Case 1	Case 2
59.94/I, 59.94/P	60,000	27,000,000	1,001	450,450
60.00/I, 60.00/P	60,000	27,000,000	1,000	450,000

2.4 Restrictions on delay time

Restrictions are shown in Table 2-8.

Table 2-8: Restrictions on delay time

Item	Restrictions
IRAP AU interval RPSEI insert interval (in case of field coding)	Less than 1 second in rule, and maximum 2 seconds
CPB size	Within 1.0R [bit]　(R is maximum bitrate.)
CPB delay	AuNominalRemovalTime[0] must be less than 0.5 second.
Maximum DPB size	sps_max_dec_pic_buffering_minus1 must be less than 5 for frame coding, and less than 11 for field coding.

2.5 Picture partitioning

Picture partitioning is not be mandatory, but usage of WPP in scope of HEVC Standard is permitted. Slice including usage of WPP shall be also permitted to use, but minimum slice unit shall be 1 CTU line. The value of entropy_coding_sync_flag_enabled_flag shall be fixed in CVS.

2.6 Various coding parameters

The value of MinCR is restricted as shown in Table 2-9.

Table 2-9: Restrictions of MinCR

Level	MinCR	Remarks
1	2	
2	2	
2.1	2	
3	2	
3.1	2	
4	4	
4.1	4	

<Blank Page>

Part 2: Audio Signal and Coding System

Part 2: Audio Signal and Coding System

Contents

<Blank Page>

Chapter 1: General

1.1 Objective

The purpose of this standard is to set specific parameters for audio signal and audio coding systems in digital broadcasting.

1.2 Scope

This standard applies to digital broadcasting that comply with the "Standard transmission system for digital broadcasting among standard TV broadcasting and the like" (Ordinance of the Ministry of Internal Affairs and Communications, No. 87, 2011) and digital broadcasting that comply with the "Standard transmission system for satellite general broadcasting" (Ordinance of the Ministry of Internal Affairs and Communications, No. 94, 2011).

1.3 References

1.3.1 Normative references

The followings are those documents that a part of the items, provided in the following documents, is quoted in this standard:

(1) Ordinance of the Ministry of Internal Affairs and Communications, No. 87, 2011 "Standard transmission system for digital broadcasting among standard TV broadcasting and the like" (Partial Amendment: Dec. 10, 2013, July 3, 2014, Oct. 21, 2014. Hereinafter referred to as "Ordinance". But as for the number of Ordinance specified, it shall be followed.)
(2) Ordinance of the Ministry of Internal Affairs and Communications, No. 94, 2011 "Standard transmission system for satellite general broadcasting" (Partial Amendmet: Dec. 10, 2013, July 3, 2014. Hereinafter referred to as "Ordinance No. 94".)
(3) Notification of the Ministry of Internal Affairs and Communications, No. 234, 2014 "Defining compression and transmission procedures for a video signal and audio signals" (Partial Amendment: Oct. 21, 2014. Hereinafter referred to as "Notification".)
(4) ISO/IEC 13818-7:2006 Information technology -- Generic coding of moving pictures and associated audio information: Advanced Audio Coding (AAC)
(5) ISO/IEC 13818-7:2006/Cor.1:2009 Information technology -- Generic coding of moving pictures and associated audio information -- Part 7: Advanced Audio Coding (AAC), TECHNICAL CORRIGENDUM 1
(the above mentioned standards (4) and (5) are hereinafter referred to as "MPEG-2 AAC Standard")
(6) ISO/IEC 13818-3:1998 Information technology -- Generic coding of moving pictures and associated audio information: Audio (hereinafter referred to as "MPEG-2 BS Standard".)
(7) ISO/IEC 14496-3:2009 Information technology -- Coding of audio-visual objects – Part 3: Audio
(8) ISO/IEC 14496-3:2009/Cor.1:2009 Information technology -- Coding of audio-visual objects -- Part 3: Audio
(9) ISO/IEC 14496-3:2009/AMD 2:2010 Information technology -- Coding of audio-visual objects -- Part 3: Audio
(10) ISO/IEC 14496-3:2009/cor.2:2011 Information technology -- Coding of audio-visual objects -- Part 3: Audio
(11) ISO/IEC 14496-3:2009/AMD 4:2013 Information technology -- Coding of audio-visual objects -- Part 3: Audio

(12) ISO/IEC 23003-1:2007 Information technology -- MPEG audio technologies -- Part 1: MPEG Surround

1.3.2 Informative references

(1) ARIB STD-B21 "Receiver for Digital Broadcasting (desirable specifications)"
(2) ARIB STD-B59 "Three-dimensional Multichannel Stereophonic Sound System for Programme Production"
(3) ARIB-STD-B60 "MMT-based Media Transport Scheme in Digital Broadcasting Systems"

1.4 Terms

1.4.1 Definitions

(1) Digital Terrestrial Sound Broadcasting:

Digital broadcasting among very high frequency broadcasting which are operated by key terrestrial broadcasting stations that are provided in Ordinance, Chapter 2.

(2) Digital Terrestrial Television Broadcasting:

Digital broadcasting and high definition television broadcasting among standard television broadcasting which are operated by key terrestrial broadcasting stations that are provided in Ordinance, Chapter 3.

(3) Multimedia Broadcasting:

Television broadcasting and multimedia broadcasting which are operated by key terrestrial broadcasting stations that are provided in Ordinance, Chapter 4. Among these, it is V-Low multimedia broadcasting by connected segment system that is provided in Chapter 4, Section 1. And it is V-High multimedia broadcasting by connected segment system that is provided in Chapter 4, Section 2.

(4) BS Digital Broadcasting:

Digital broadcasting among standard television broadcasting, high definition television broadcasting, very high frequency broadcasting and data broadcasting which are operated by key satellite broadcasting stations using radio wave whose frequency is from 11.7GHz to 12.2GHz that is provided in Ordinance, Chapter 5, Section 2.

(5) Advanced BS Digital Broadcasting:

Digital broadcasting among standard television broadcasting, high definition television broadcasting, ultra-high definition television broadcasting, very high frequency broadcasting and data broadcasting by advanced wide band transmission system which are operated by key satellite broadcasting stations using radio wave whose frequency is from 11.7GHz to 12.2GHz that is provided in Ordinance, Chapter 5, Section 3.

(6) Narrow band CS Digital Broadcasting:

Standard television broadcasting, high definition television broadcasting, very high frequency broadcasting and data broadcasting by narrow band transmission system which are operated as general satellite broadcasting by satellite stations using radio wave whose frequency is from 12.2GHz to 12.75GHz that is provided in Ordinance No.94, Article 3, Paragraph 1.

(7) Wide band CS Digital Broadcasting:

Standard television broadcasting, high definition television broadcasting, very high frequency broadcasting and data broadcasting by wide band transmission system which are operated by key satellite stations using radio wave whose frequency is from 12.2GHz to 12.75GHz that is provided in Ordinance, Chapter 6, Section 3.

(8) Advanced Narrow band CS Digital Broadcasting:

Standard television broadcasting, high definition television broadcasting, ultra-high definition television broadcasting, very high frequency broadcasting and data broadcasting by advanced narrow band transmission system which are operated as general satellite broadcasting by satellite stations using radio wave whose frequency is from 12.2GHz to 12.75GHz that is provided in Ordinance No.94, Article 3, Paragraph 1.

(9) Advanced Wide band CS Digital Broadcasting:

Standard television broadcasting, high definition television broadcasting, ultra-high definition television broadcasting, very high frequency broadcasting and data broadcasting by advanced wide band transmission system which are operated by key satellite stations using radio wave whose frequency is from 12.2GHz to 12.75GHz that is provided in Ordinance, Chapter 6, Section 5.

1.4.2 Abbreviations

AAC:	Advanced Audio Coding
ADTS:	Audio Data Transport Stream
ALS	Audio Lossless Coding
BC	Backward Compatible
CPE:	Channel Pair Element
CRC:	Cyclic Redundancy Check
DSE:	Data Stream Element
HE-AAC	High Efficiency Advanced Audio Coding
LATM	Low Overhead Audio Transport Multiplex
LC:	Low Complexity
LFE:	Low Frequency Effects
LOAS	Low Overhead Audio Stream
MPEG:	Moving Picture Experts Group
PCE:	Program Configuration Element
SBR	Spectral Band Replication
SCE:	Single Channel Element
TNS	Temporal Noise Shaping

<Blank Page>

Chapter 2: Audio Input Signal

(1) The sampling frequency for audio signals shall be 32 kHz, 44.1 kHz, or 48 kHz. But the sampling frequency for advanced BS digital broadcasting and advanced wide band CS digital broadcasting shall be 48 kHz, and the sampling frequency for V-Low multimedia broadcasting by connected segment system shall be 32 kHz or more.

(2) To configure stereophonic signals (consisting of two or more audio signals to achieve a three-dimensional reproduction of sound), the sampling timing for all signals shall be the same.

(3) The number of quantization bits for the input signal shall be 16 or more.

(4) The maximum number of audio input channels shall be five, in addition to one channel used to enhance low frequencies. But the maximum number of audio input channels for advanced BS digital broadcasting, advanced narrow band CS digital broadcasting and advanced wide band CS digital broadcasting shall be 22, in addition to two channels used to enhance low frequencies.

(Ordinance Article 7, Article 24-8, Article 45, Article 65, Article 81-4)

<Blank Page>

Chapter 3: Audio Coding System

3.1 System based on MPEG-2 AAC Standard

The system shall be a combination of time-frequency transform coding system and psycho-acoustic weighted bit assignment system, and the audio compression and transmission procedures shall comply with the other Notification by the Minister of Internal Affairs and Communications (refer to Chapter 4.1).

(Ordinance Article 5, Article 44, Article 81-3)

3.2 System based on MPEG-2 BC Standard

The system shall be a combination of band division coding system and psycho-acoustic weighted bit assignment system, and the audio compression and transmission procedures shall comply with the other Notification by the Minister of Internal Affairs and Communications (refer to Chapter 4.2).

(Ordinance Article 72)

3.3 System based on MPEG-4 AAC Standard

The system shall be a combination of time-frequency transform coding system and psychoacoustic coding system, and the audio compression and transmission procedures shall comply with the other Notification of the Minister of Internal Affairs and Communications (refer to Chapter 4.3).

(Ordinance Article 64, Article 81-3)

3.4 System based on MPEG-4 ALS Standard

The system shall be a combination of linear predictive coding system and variable length coding system, and the audio compression and transmission procedures shall comply with the other Notification of the Minister of Internal Affairs and Communications (refer to Chapter 4.4).

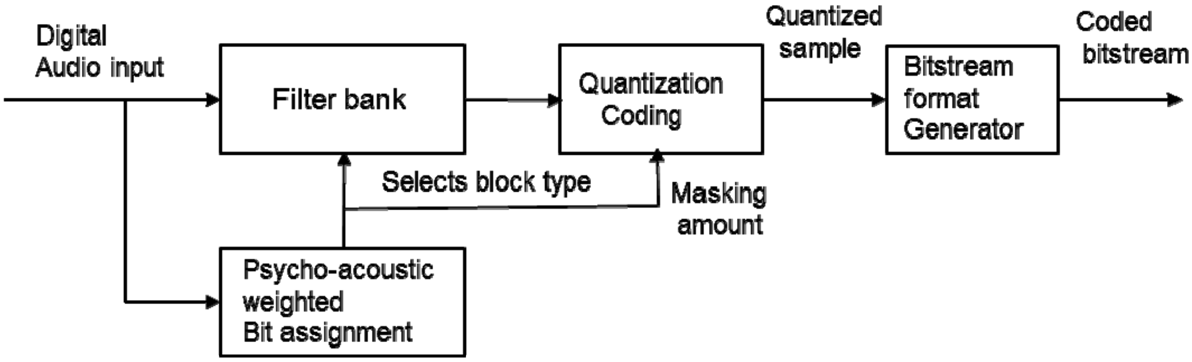
(Ordinance Article 64, Article 81-3)

<Blank Page>

Chapter 4: Audio Compression and Transmission Procedures

4.1 System based on MPEG-2 AAC Standard

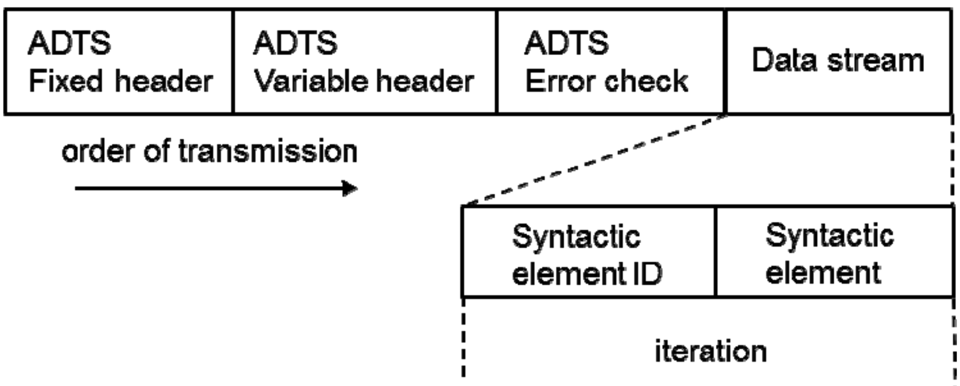
Audio compression and transmission procedures shall be as specified in the following.



Notes:

1. The filterbank converts a digital audio input signal from time-axis over to frequency-axis by modified discrete cosine transform. At this time, the filterbank selects block type input to modified discrete cosine transform and window function according to psychoacoustic characteristics of the input signal.
2. Psycho-acoustic weighted bit assignment calculates masking amount (limits of differentiating a specific audio signal from other audio signals) and block type input to the filterbank.
3. Quantization and coding allows a quantized sample to be output after quantizing and coding the output signal from the filterbank based on the masking amount calculated by psychoacoustic weighted bit assignment so that the total number of bits that can be used by each block is not exceeded.
4. The maximum number of channel modes for coding the bitstream shall be five channels, plus one channel used to enhance low frequencies (*).
5. The bitstream shall be configured as shown below.

(Bitstream configuration)



Notes:

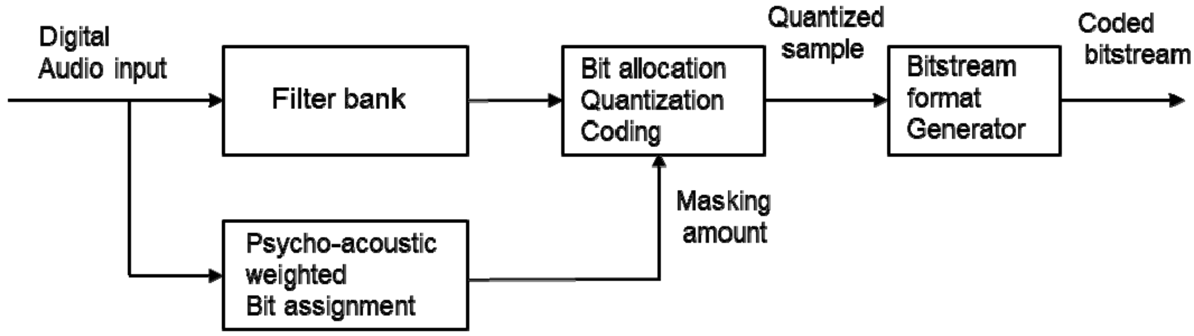
1. The ADTS fixed header consists of synchronization and audio coded information defined in ISO/IEC 13818-7.
   But for multimedia broadcasting, it consists of synchronization and audio coded information defined in ISO/IEC 13818-7, ISO/IEC 23003-1, ISO/IEC 14496-3, ISO/IEC 14496-3:2001/Amd.1, and ISO/IEC 14496-3:2005/Amd.2:2006.
2. The ADTS variable header consists of audio coded information defined in ISO/IEC 13818-7.
   But for multimedia broadcasting, it consists of audio coded information defined in ISO/IEC 13818-7, ISO/IEC 23003-1, ISO/IEC 14496-3, ISO/IEC 14496-3:2001/Amd.1, and ISO/IEC 14496-3:2005/Amd.2:2006.
3. ADTS error check consists of error detection information.
4. The data stream consists of audio data coded according to ISO/IEC 13818-7.
   But for multimedia broadcasting, it consists of audio data coded according to ISO/IEC 13818-7, ISO/IEC 23003-1, ISO/IEC 14496-3, ISO/IEC 14496-3:2001/Amd.1, and ISO/IEC 14496-3:2005/Amd.2:2006.
5. The syntactic element ID indicates the type of syntactic element that follows this ID or end of the data stream.
6. The syntactic element consists of various components of audio data coded according to ISO/IEC 13818-7. It is iterated the number of times specified in the ADTS variable header.
   But for multimedia broadcasting, it consists of each component of audio data coded according to ISO/IEC 13818-7, ISO/IEC 23003-1, ISO/IEC 14496-3, ISO/IEC 14496-3:2001/Amd.1, and ISO/IEC 14496-3:2005/Amd.2:2006, and it is iterated the number of times specified in the ADTS variable header.

(Notification, Appended Table No.5, Appendx 1)

(*) Though the maximum number of audio input channels for advanced narrow band CS digital broadcasting is 22.2 in Ordinance, the maximum number of coded channels for MPEG-2 AAC System is limited to 5.1 in Chapter 5 of this standard.

4.2 System based on MPEG-2 BC Standard

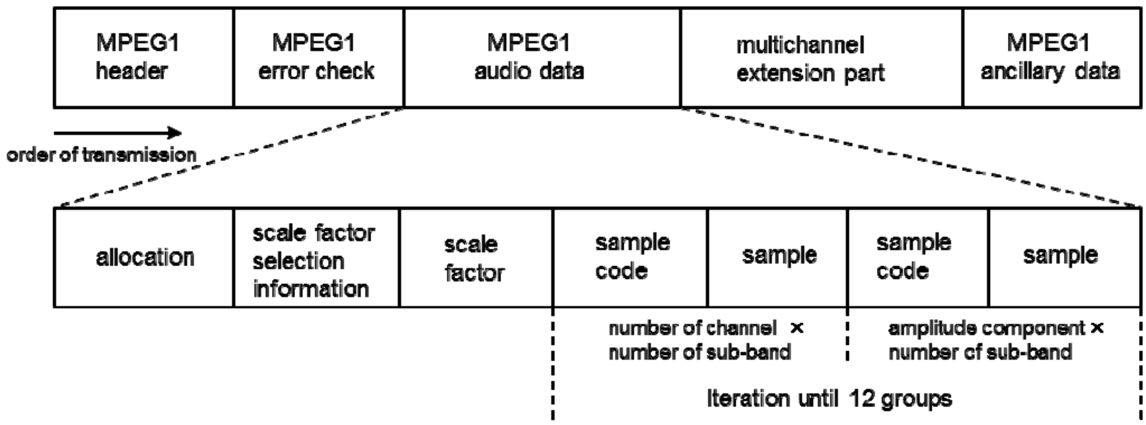
Audio compression and transmission procedures shall be as specified in the following



Notes:

1 The filter bank transforms digital audio input signal from time-axis to frequency-axis and processes band division. The filter bank is divided into 32 bands, and processed according to the provisions in ISO/IEC 11172-3 and 13818-3.
2 Psycho-acoustic weighted bit assignment calculates masking amount for each band of the filter bank.
3 Bit allocation decides the number of quantized bits for each sub-band (one of the filter bank which is divided into 32 bands. Hereinafter the same.) in the range less than the number of total bits used in frame. Quantization and Coding quantizes and codes signal in each band which is output from the filter bank, by using the number of quantized bit decided by bit allocation, and outputs quantized sample.
4 Coded bitstream shall be configured as Appendix No.1.
5 Coded bitstream shall take any channel mode shown in Appendix No. 2.

Appendix No.1 Coded bitstream configuration



Notes:

1 MPEG-1 header consists of synchronization and audio coding information which is specified in ISO/IEC 11172-3.
2 MPEG-1 error check consists of error detection information.
3 MPEG-1 audio data consists of audio data which is coded according to ISO/IEC 11172-3.
4 Multichannel extension part consists of the data which extends audio data which is coded according to ISO/IEC 13818-3.

5 Allocation consists of information which indicates the order of coding sub-band.
6 Scale factor consists of information which indicates a magnifying power when waveform in each sub-band is normalized.
7 Sample code and sample consist of coded audio data, and are iterated until 12 group in maximum. When using joint stereo mode, they consist of sample code of amplitude component with high frequencies and sample.

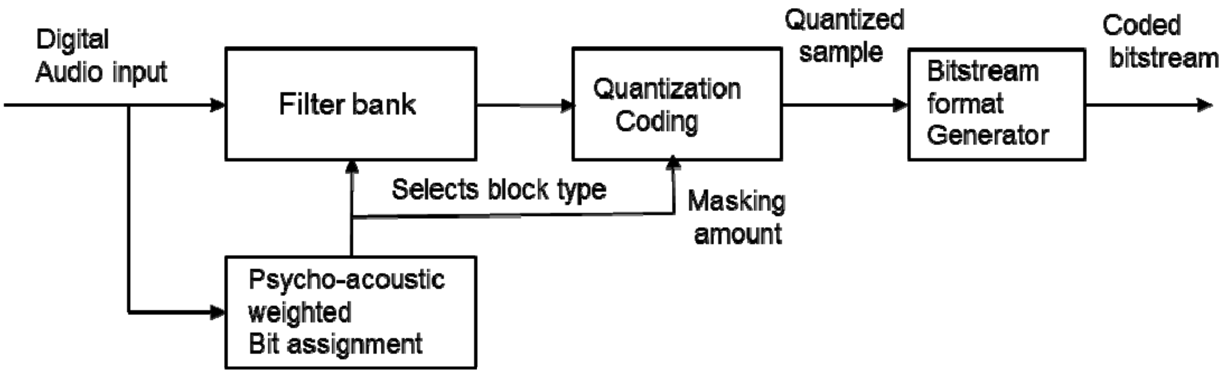
Appendix No.2 Channel mode

Channel mode	Contents
Stereo	Those which are coded by left signal and right signal, in order to achieve a three-dimensional reproduction.
Joint stereo	Those which achieve a three-dimensional reproduction, and those which are coded by only amplitude component for high frequency components among left and right signals, or those which are coded by the sum and the difference signals of left and right signals, in order to enhance the efficiency of audio compression.
Dual channel	Those which are coded by two independent audio signals.
Single channel	Those which are coded by one audio signal.
3 front/ 0 rear channel	Those which are coded by left signal, right signal, and center signal.
2 front/ 1 rear channel	Those which are coded by left signal, right signal, and surround signal (this is generated by the left rear signal and the right rear signal).
Dual stereo channel	Those which are coded by left signal and right signal of the first program and left signal and right signal of the second program.
2 front/ 2 rear channel	Those which are coded by left signal, right signal, left rear signal and right rear signal.
3 front/ 1 rear channel	Those which are coded by left signal, right signal, center signal and surround signal.
3 front/ 0 rear channel + stereo	Those which are coded by left signal, right signal, and center signal of the first program, and left signal and right signal of the second program.
3 front/ 2 rear channel	Those which are coded by left signal, right signal, center signal, left rear signal and right rear signal.

(Notification, Appended Table No.7)

4.3 System based on MEPG-4 AAC Standard

Audio compression and transmission procedures shall be as specified in the following



Notes:

1 The filter bank transforms digital audio input signal from time domain to frequency domain by modified discrete cosine transform. At this time, the filter bank selects input block type for modified discrete cosine transform and window function according to the psychoacoustic characteristics of the input signal.

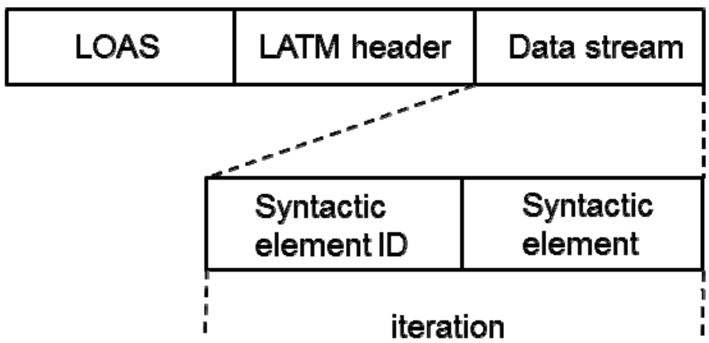
2 Psychoacoustic model calculates masking amount (limits of differentiating a specific audio signal from other audio signals) and block type input to the filter bank.

3 Quantization and coding allows a quantized sample to be output after quantizing and coding the output signal from the filter bank based on the masking amount calculated by psychoacoustic model so that the total number of bits that can be used by each block is not exceeded.

4 The maximum number of channel modes for coded bitstream shall be 22, plus two channels used to enhance low frequencies.

5 Bitstream configuration shall be either LATM/LOAS format as the following, or the other format (*).

(Bitstream configuration of LATM/LOAS format)



Notes:

1 LOAS shall consist of synchronization and audio coding information which is specified in ISO/IEC 14496-3.

2 LATM header shall consist of audio coding information which is specified in ISO/IEC 14496-3.

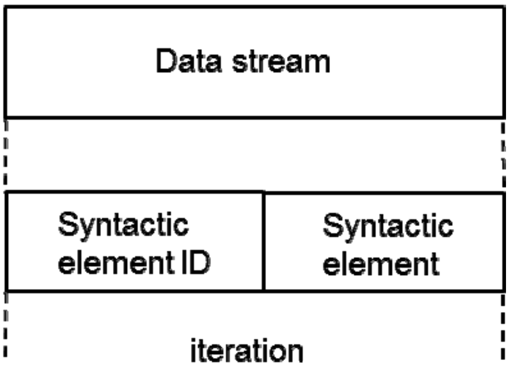
3 Data stream shall consist of audio data coded according to ISO/IEC 14496-3.

4 The syntactic element ID shall indicate the type of syntactic element that follows this ID

or the end of the data stream.

5 The syntactic element shall consist of each component of audio data coded according to ISO/IEC 14496-3, and it shall be iterated the number of times addressed in the LATM header.

(Bitstream configuration of the other format)



Notes:

1 Data stream shall consist of audio data coded according to ISO/IEC 14496-3.

2 The syntactic element ID shall indicate the type of syntactic element that follows this ID or the end of the data stream.

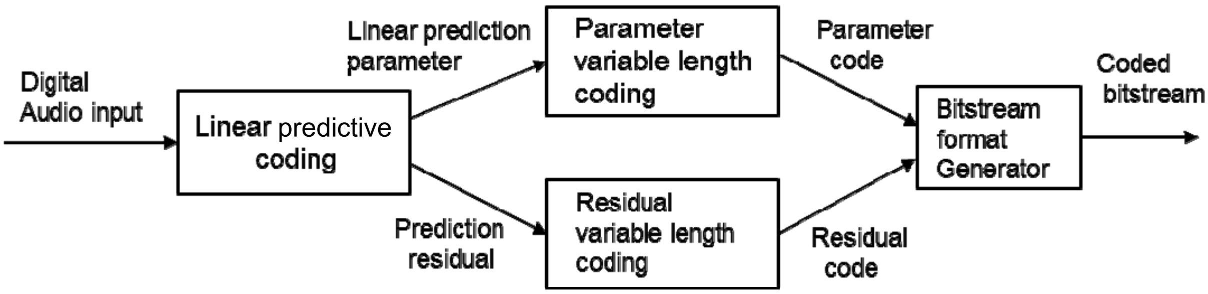
3 The syntactic element consists of each component of audio data coded according to ISO/IEC 14496-3.

(Notification Appended Table No.5, Appendix 2, Appendix 3)

(*) For V-Low multimedia broadcasting by connected segment system, bitstream by ADTS format described in Chapter 4.1 can be used.

4.4 System based on MPEG-4 ALS Standard

Audio compression and transmission procedures shall be as specified in the following.



Notes:

1. Linear predictive coding shall analyze digital audio input, and calculating linear prediction parameter and prediction residual.
2. Variable length coding for parameter shall encode linear prediction parameter to variable length code, and then provide parameter code.
3. Variable length coding for residual shall encode prediction residual (which is the differential between input value and predicted value) to variable length code, and then provide residual code.
4. Bitstream format generator shall provide coded bitstream as the following by combining parameter codes and residual codes.

(Bitstream configuration)

Coding information	Parameter code (variable length code)	Residual code (variable length code)

Note: Coding information, parameter code and residual code shall comply with audio lossless coding specified in ISO/IEC 14496-3.

(Notification Appended Table No.6)

<Blank Page>

Chapter 5: Restrictions on MPEG-2 AAC Audio Coding Parameters

This chapter defines operational restrictions regarding audio coding systems for digital broadcasting based on MPEG-2 AAC System, in addition to the provisions of Ordinances and Notifications given in Chapters 2 through 4.

5.1 Input audio format based on MPEG-2 AAC System

The input audio format for digital broadcasting is subject to the following restrictions:

Parameter		Restriction
Audio mode	Possible audio mode	mono, stereo, multichannel stereo (3/0, 2/1, 3/1, 2/2, 3/2, 3/2+LFE (3/2.1)) (Note), 2-audio signals (dual mono), multi-audio (3 or more audio signals) and combinations of the above
	Recommended audio mode	mono, stereo, multichannel stereo (3/1, 3/2, 3/2+LFE (3/2.1)) (Note), 2-audio signals (dual mono)
Emphasis		None

(Note) Notation for audio mode of multichannel stereo: Audio mode of multichannel streo is denoted as "front/rear.LEF".

There is a case to denote "+ LFE" when the assigned channel for LFE (low frequency enhance effect channel) is one.

There is a related record about notation for audio mode in Description 2.

5.2 Audio coding system based on MPEG-2 AAC System

MPEG-2 AAC is stipulated in the Ordinance as the audio coding system for digital broadcasting. (See Chapter 3.1.) However, this chapter defines additional operational restrictions applicable to digital broadcasting services.

See Appendix 3 for the references of MPEG-2 AAC System.

5.2.1 Main parameters

Parameter	Restriction
Bitstream format	AAC Audio Data Transport Stream (ADTS)
Profile	Low Complexity (LC) profile
Max. number of coded channels	5.1 channels (Note) per ADTS
Max. bitrate	Compliant to ISO/IEC 13818-7

(Note) 5 channels + LFE channel

5.2.2 Restrictions on MPEG-2 AAC ADTS coding parameters

(1) Fixed header of ADTS

Parameter	Restriction
protection_absent	'0' (CRC error check is always presented)
profile	1 (LC profile)
sampling_frequency_index	0x0 (96kHz) (Note 1), 0x3 (48kHz), 0x4 (44.1kHz), 0x5 (32kHz), 0x6 (24kHz) (Note 2), 0x7 (22.05kHz) (Note 2), 0x8 (16kHz) (note 2)
channel_configuration	See Chapter 5.2.3.

(Note 1) 0x0 (96kHz) can be used only for V-Low multimedia broadcasting by connected segment system.

(Note 2) 0x6 to 0x8 (24 k, 22.05 k, 16 kHz) are not used for BS/wide band CS digital broadcasting.

(2) Variable header of ADTS

Parameter	Restriction
adts_buffer_fullness	Use of 0x7FF (indicating variable rate) is not permitted.
number_of_raw_data_blocks_in_frame	0 (number of raw_data_blocks per frame = 1)

(3) Raw data stream

Parameter	Restriction
Coding mode in a single ADTS and raw_data_block configuration (order of transmission)	See Chapter 5.2.3.
Handling of Coupling Channel option	Use of Coupling Channel option is not permitted.
Handling of Program Configuration Element (PCE)	See Chapter 5.2.3.
Handling of Data Stream Element (DSE)	See Chapter 5.2.3.
Handling of Fill Element (FIL)	See Chapter 5.2.3.

5.2.3 Detailed provisions regarding audio stream configuration and multiplexing

(1) Provisions regarding input audio mode and ADTS configuration and multiplexing

Input audio mode	ADTS configuration and multiplexing
mono, stereo	Comprises one ADTS.
Multichannel stereo (3/0, 2/1, 3/1, 2/2, 3/2, 3/2+LFE (3/2.1))	Comprises one ADTS.
2-audio signals (dual mono) (Note)	Comprises one ADTS.
Multiple audio signals other than dual mono (2/0+2/0, etc.)	Comprises the same number of ADTSs as that of audio streams (languages) and is multiplexed with the MPEG-2 systems layer.

(Note) Dual mono is defined as "two monophonic audio channels that can be simultaneously reproduced by a single ADTS."

(2) Detailed provisions regarding coding mode in a single ADTS and ADTS configuration (order of transmission)

Coding mode stipulated as default in the AAC Standard

Coding mode	channel_configuration (adts_fixed_header)	SE configuration (order of transmission) (Note 1) (Transmission shall occur in the following order) (Note) "1" and "2" to the right of SCE and CPE are the numbers assigned to both for convenience in identifying the order of transmission within the same frame.	Default element to speaker mapping (Note 2)
mono (1/0)	1	<SCE1><TERM>	SCE1 = C
stereo (2/0)	2	<CPE1><TERM>	CPE1 = L and R
3/0	3	<SCE1><CPE1><TERM>	SCE1 = C , CPE1 = L and R
3/1	4	<SCE1><CPE1><SCE2><TERM>	SCE1 = C, CPE1 = L and R, SCE2 = MS
3/2	5	<SCE1><CPE1><CPE2><TERM>	SCE1 = C, CPE1 = L and R, CPE2 = LS and RS
3/2+LFE (3/2.1)	6	<SCE1><CPE1><CPE2><LFE><TERM>	SCE1 = C, CPE1 = L and R, CPE2 = LS and RS, LFE = LFE

Coding mode other than AAC default provision

Coding mode	channel_configuration (adts_fixed_header)	SE configuration (order of transmission) (Note 1)	Default element to speaker mapping (Note 2)
2/1	0	<CPE1>＜SCE1><TERM>	CPE1 = L and R, SCE1=MS
2/2	0	<CPE1>＜CPE2><TERM>	CPE1 = L and R, CPE2=LS and RS
2-audio signals (1/0+1/0)	0	<SCE1><SCE2><TERM>	SCE1 = Main, SCE2 = Subordinate

(Note 1) Abbreviations in relation to Syntactic Element (SE):
SCE: Single Channel Element, CPE: Channel Pair Element, LFE: LFE Channel Element, TERM: Terminator

(Note 2) Abbreviations in relation to speaker arrangement:
L: Left front speaker / R: Right front speaker / C: Center front speaker /
LFE: Low frequency emphasis / LS: Left surround speaker / RS: Right surround speaker /
MS: Monophonic surround speaker

(3) Detailed provisions regarding transmission of PCE (Program Configuration Element)

(a) During continuous service using the same service ID, PCE shall be transmitted when switching between audio modes (2/1, 2/2, 1/0+1/0) for which channel_configuration (parameter within adts_fixed_header) = 0. At this time, the PCE parameter value shall match that included in the ADTS header.

(b) When downmix coefficient is transmitted in audio mode for channel_configuration = 5 or 6, PCE shall be transmitted at an interval of less than 550 ms for that purpose. When performing this operation, PCE shall always be transmitted during the period in which channel_configuration = 5 or 6 is in continuous service.

(c) While PCE may be included in every ADTS frame, any modification of parameters other than changes made (for example) to channels and downmix coefficients is prohibited.

(d) The following operational provisions are established for bits comprising PCE. Note that provisions (1) through (3) described above apply to bits not specifically mentioned.

- The same value shall be assigned to Profile and Sampling_frequency_index as the header.
- No specific provisions are established for Num_assoc_data_elements.
- Num_valid_cc_elements shall be 0.
  Therefore, the following flags do not exist:
  cc_element_is_ind_sw
  valid_cc_element_tag_select
- Mono_mixdown_present shall be 0.
  Therefore, mono_mixdown_element_number does not exist.
- Stereo_mixdown_present shall be 0.
  Therefore, stereo_mixdown_element_number does not exist.
- Comment_field_bytes shall be treated according to the AAC standard. Its content is meaningless as far as the system is concerned.
  It is treated as an option for using bitstream control.
  (Note)
  The decoder needs not decode this area. However, it shall be ensured that decoding is not seriously affected.

(4) Detailed provisions regarding configuration of Fill Element (FIL)

(a) When the value of coding parameter sampling_frequency_index in the ADTS Fixed Header is in the range of 0x6 to 0x8 (24k, 22.05k, 16kHz), EXT_SBR_DATA ('1101') and EXT_SBR_DATA_CRC ('1110') can be used in Fill Element (FIL). For V-Low multimedia broadcasting by connected segment system, even when sampling_frequency_index is 0x3 (48k), EXT_SBR_DATA ('1101') and EXT_SBR_DATA_CRC ('1110') can be also used.

(Note) For BS / wide band CS digital broadcasting, the value of sampling_frequency_index does not fall within the range of 0x6 to 0x8, therefore, EXT_SBR_DATA ('1101') and EXT_SBR_DATA_CRC ('1110') are not used.

(b) For multimedia broadcasting, EXT_SAC_DATA ('1100') can be used in Fill Element (FIL).

5.2.4 Operational provisions regarding downmixing when multichannel stereo service is provided

This section defines the conditions and lists considerations in relation to compatibility with 2-channel stereo-capable receiver when multichannel stereo service of 5.1-channel stereo or less is provided.

(1) Two-channel stereo simulcasting is not obligatory when multichannel stereo service of 5.1-channel stereo (3/2+LFE (3/2.1)) or less is provided. Basically, 2-channel stereo-capable receiver shall handle the service by downmixing.

(2) It shall be possible to transmit downmix coefficient using PCE according to the AAC Standard when 5-channel stereo (3/2) and 5.1-channel stereo (3/2+LFE (3/2.1)) services are provided. For the detailed provisions regarding transmission of PCE, refer to the section 5.2.3 (3).

(3) It shall be also possible to provide 2-channel stereo simulcasting service at the request of broadcasting stations. In this case, two streams shall be treated as different ADTSs, multiplexed, and stream-controlled by the systems layer.

(4) For more information on downmixing operations of a 2-channel stereo-capable receiver other than the above mentioned cases (2) and (3), refer to the ARIB STD-B21 section 6.2.1(7), “Down mixing function from multi-channel to 2-channel stereo”.

<Blank Page>

Chapter 6: Restrictions on MPEG-4 AAC Audio Coding Parameters

This chapter specifies restrictions on operations related to audio coding system of digital broadcasting based on MPEG-4 AAC System. Input audio format is described in Chapter 6.1, restrictions on coding parameters by MPEG-4 AAC System is described in Chapter 6.2, restrictions on stream format by MPEG-4 AAC System is described in Chapter 6.3, and restrictions on operation for multichannel stereo service is described in Chapter 6.4..

6.1 Input Audio Format based on MPEG-4 AAC System

Restrictions on input audio format for digital broadcasting shall be as the following.

Item	Restriction
Audio mode	・mono ・stereo ・multichannel stereo (Note) 3/0, 2/1, 3/1, 2/2, 3/2, 3/2.1, 5/2.1, 3/3.1, 3/2/2.1, 2/0/0-3/0/2-0.1. 3/3/3-5/2/3-3/0/0.2 ・2-audio signals (dual mono) (1/0+1/0)
Emphasis	none

(Note) Notation of audio mode in multichannel stereo:

The number of channel is represented as

"upper layer (front/side/back)-middle layer (front/side/back)-lower layer (front/side/back).LFE".

But the layer which does not have any allocated channel is denoted as 0. Also, audio mode by only middle layer is denoted as "middle layer (front/side/back).LFE", and multichannel stereo which is only by middle layer without side channel is simply denoted as "middle layer (front/back).LFE". When the allocated channel to LFE (low frequency effect channel) is one, there is a case that it is denoted as "+LFE".
There is a related record in Description 2 about notation for audio mode.

6.2 Coding parameters for MPEG-4 AAC System

MPEG-4 AAC System as an audio coding system for digital broadcasting is provided in Ordinance (refer to Chapter 3.3). But in this section, more restrictions on operations for realizing digital broadcasting services.

6.2.1 Main parameters

Item	Restriction
Profile	AAC Profile, HE-AAC profile (Note 1)
Audio object type	2 (AAC LC) (Note 2) 5 (SBR) (in case of HE-AAC profile)
Maximum number of coding channels	22.2 channels per 1 raw_data_block (Note 3)
Maximum bitrate	based on ISO/IEC 14496-3

(Note 1) It shall be possible that HE-AAC profile is operated in V-Low multimedia broadcasting by connected segment system.
(Note 2) The meaning of profile differs for MPEG-2 and MPEG-4. Audio object type of MPEG-4 and profile of MPEG-2 are the same in the meaning.
(Note 3) 22 channels＋2 LFE channels

6.2.2 Restrictions on MPEG-4 Audio parameters

For MPEG-4 Audio, parameters in the coding system to be used are set by using AudioSpecificConfig(). When using MPEG-4 AAC System, restrictions are specified for setting parameters.

Here, "Not used" in the table represents that the item is not recorded in the bitstream for any setting value of the other parameters.

AudioSpecificConfig()

Item	Restriction
samplingFrequencyIndex	0: 96000Hz (Note) 3: 48000Hz 6: 24000Hz (Note)
samplingFrequency	Not used
channelConfiguration	1: 1ch (1/0) 2: 2ch (2/0) 3: 3ch (3/0) 4: 4ch (3/1) 5: 5ch (3/2) 6: 5.1ch (3/2.1) 7: 7.1ch (5/2.1) 11: 6.1ch (3/0/3.1) 12: 7.1ch (3/2/2.1) 13: 22.2ch (3/3/3-5/2/3-3/0/0+2) 14: 7.1ch (2/0/0-3/0/2-0/0/0+1) 0: program_config_element() is used. (in case of 3ch(2/1), 4ch(2/2) and 2-audio signals (dual mono) (1/0+1/0))
extensionSamplingFrequencyIndex	Not used
extensionSamplingFrequency	Not used
extensionChannelConfiguration	Not used
CelpSpecificConfig()	Not used
HvxcSpecificConfig()	Not used
TTSSpecificConfig()	Not used
StructuredAudioSpecificConfig()	Not used
ErrorResilientCelpSpecificConfig()	Not used
ErrorResilientHvxcSpecificConfig()	Not used
ParametricSpecificConfig()	Not used
SSCSpecificConfig()	Not used
sacPayloadEmbedding	Not used
SpatialSpecificConfig()	Not used
MPEG_1_2_SpecificConfig()	Not used
DSTSpecificConfig()	Not used
fillBits	Not used
ALSSpecificConfig()	Not used
SLSSpecificConfig()	Not used
ELDSpecificConfig()	Not used
SymbolicMusicSpecificConfig()	Not used
epConfig	Not used
ErrorOrotectonSpecificConfig()	Not used
directMapping	Not used
syncExtensionType	Not used
sbrPresentFlag	-1 (in case of HE-AAC profile)
extensionSamplingFrequencyIndex	Not used

extensionSamplingFrequency	Not used
syncExtensionType	Not used
psPresentFlag	Not used
extensionChannelConfiguration	Not used

(Note) Only for V-Low multimedia broadcasting by connected segment system, it shall be possible to operate 0:96000Hz, 6:24000Hz.

GetAudioObjectType()

Item	Restriction
audioObjectType	2 (AAC LC)
audioObjectTypeExt	Not used

GASpecificConfig()

Item	Restriction
frameLengthFlag	0 (frameLength = 1024)
dependsOnCoreCoder	0
coreCoderDelay	Not used
extensionFlag	0
program_config_element()	This is used only for audio mode of 2/1, 2/2, and 2-audio signals (dual mono) (1/0+1/0)
layerNr	Not used
numOfSubFrame	Not used
layer_length	Not used
aacSectionDataResilienceFlag	Not used
aacScalefactorDataResilienceFlag	Not used
aacSpectralDataResilienceFlag	Not used
extensionFlag3	Not used

PayloadLengthInfo()

Item	Restriction
MuxSlotLengthCoded[]	Not used
numChunk	Not used
streamIndx	Not used
AuEndFlag[]	Not used
MuxSlotLengthCoded[]	Not used

PayloadMux()

Item	Restriction
payload[0]	This stores Raw Data Stream

Raw Data Stream

Item	Restriction
Configuration of Raw Data Stream	Comprises 1 raw_data_block
Coding mode and configuration in 1 raw_data_block (order of transmission)	Refer to Chapter 6.2.3, (1)
Handling of Coupling Channel Element	Prohibited from using.
Handling of Program Configuration Element (PCE)	Refer to Chapter 6.2.3, (2)
Handling of Data Stream Element (DSE)	Refer to Chapter 6.2.3, (3)

6.2.3 Detailed provisions regarding audio stream configuration and multiplexing

(1) Detailed provisions regarding coding mode in 1 raw_data_block and configuration of raw_data_block (order of transmission)

Coding mode based on ISO/IEC 14496-3:2009 and ISO/IEC 14496-3:2009/AMD 4 is used.

Coding mode provided as default in AAC standard

coding mode	channel_configuration	SE configuration (order of transmission) (Note 1)	Default element to speaker mapping (Note 2) (Note 3)	index mapping for dialogue_src_index
mono (1/0)	1	<SCE1> <TERM>	SCE1 = C	1 : C
stereo (2/0)	2	<CPE1> <TERM>	CPE1 = L and R	1 : L 2 : R
3/0	3	<SCE1> <CPE1> <TERM>	SCE1 = C, CPE1 = L and R	1: C 2 : L 3 : R
3/1	4	<SCE1> <CPE1> <SCE2> <TERM>	SCE1 = C, CPE1 = L and R, SCE2 = MS	1 : C 2 : L 3 : R 4 : MS
3/2	5	<SCE1> <CPE1> <CPE2> <TERM>	SCE1 = C, CPE1 = L and R, CPE2 = LS and RS	1 : C 2 : L 3 : R 4 : LS 5 : RS
3/2.1	6	<SCE1> <CPE1> <CPE2> <LFE> <TERM>	SCE1 = C, CPE1 = L and R, CPE2 = LS and RS, LFE = LFE	1 : C 2 : L 3 : R 4 : LS 5 : RS 6 : LFE
5/2.1	7	<SCE1> <CPE1> <CPE2> <CPE3> <LFE> <TERM>	SCE1 = FC, CPE1 = FLc and FRc, CPE2 = FL and FR, CPE3 = BL and BR, LFE = LFE	1 : FC 2 : FLc 3 : FRc 4 : FL 5 : FR 6 : BL 7 : BR 8 : LFE
3/3.1	11	<SCE1> <CPE1> <CPE2> <SCE2> <LFE> <TERM>	SCE1 = FC, CPE1 = FL and FR, CPE2 = BL and BR, SCE2 = BC, LFE = LFE	1 : FC 2 : FL 3 : FR 4 : BL 5 : BR 6 : BC 7 : LFE
3/2/2.1	12	<SCE1> <CPE1> <CPE2> <CPE3> <LFE> <TERM>	SCE1 = FC, CPE1 = FL and FR, CPE2 = SiL and SiR, CPE3 = BL and BR, LFE = LFE	1 : FC 2 : FL 3 : FR 4 : SiL 5 : SiR 6 : BL 7 : BR

				8 : LFE
3/3/3-5/2/ 3-3/0/0.2	13	<SCE1> <CPE1> <CPE2> <CPE3> <CPE4> <SCE2> <LFE1> <LFE2> <SCE3> <CPE5> <CPE6> <SCE4> <CPE7> <SCE5> <SCE6> <CPE8> <TERM>	SCE1 = FC, CPE1 = FLc and FRc, CPE2 = FL and FR, CPE3 = SiL and SiR, CPE4 = BL and BR, SCE2 = BC, LFE1 = LFE1, LFE2 = LFE2, SCE3 = TpFC, CPE5 = TpFL and TpFR, CPE6 = TpSiL and TpSiR, SCE4 = TpC, CPE7 = TpBL and TpBR, SCE5 = TpBC, SCE6 = BtFC, CPE8 = BtFL and BtFR	1 : FC 2 : FLc 3 : FRc 4 : FL 5 : FR 6 : SiL 7 : SiR 8 : BL 9 : BR 10 : BC 11 : LFE1 12 : LFE2 13 : TpFC 14 : TpFL 15 : TpFR 16 : TpSiL 17 : TpSiR 18 : TpC 19 : TpBL 20 : TpBR 21 : TpBC 22 : BtFC 23 : BtFL 24 : BtFR
2/0/0-3/0/ 2-0.1	14	<SCE1> <CPE1> <CPE2> <LFE> <CPE3> <TERM>	SCE1 = FC, CPE1 = FL and FR, CPE2 = LS and RS, LFE = LFE, CPE3 = TpFL and TpFR	1 : FC 2 : FL 3 : FR 4 : LS 5 : RS 6 : LFE 7 : TpFL 8: TpFR

Coding mode other than AAC default provision

coding mode	channel_configuration	SE configuration (order of transmission) (Note 1)	Default element to speaker mapping (Note 2)	index mapping for dialogue_src_index
2/1	0	<CPE1> <SCE1> <TERM>	CPE1 = L and R, SCE1 = MS	1 : L 2 : R 3 : MS
2/2	0	<CPE1> <CPE2> <TERM>	CPE1 = L and R, CPE2 = LS and RS	1 : L 2 : R 3 : LS 4 : RS
2-audio signals (1/0+1/0)	0	<SCE1> <SCE2> <TERM>	SCE1 = main, SCE2 = sub	1 : main, 2 : sub

(Note 1) Abbreviations in relation to Syntactic Element (SE):
SCE: Single Channel Element, CPE: Channel Pair Element, LFE: LFE Channel Element, TERM: Terminator

(Note 2) Abbreviations in relation to speaker arrangement: channel_configuration＝1~6
L: Left front speaker / R: Right front speaker / C: Center front speaker / LFE: Low frequency effect / LS: Left surround speaker / RS: Right surround speaker / MS: Mono surround speaker

(Note 3) Abbreviations in relation to speaker arrangement: channel_configuration＝7, 11~14
This is based on audio channel label in ARIB STD-B59 "Three-dimensional Multichannel Stereophonic Sound System for Programme Production"

(Note 4) Abbreviations in relation to index mapping for dialogue_src_index: Though index is also assigned to LFE, it is not used as a dialogue channel.

(2) Detailed provision regarding transmission of PCE (Program Configuration Element)

(a) PCE in raw_data_block() is transmitted at an interval of less than 550 ms in order to transmit audio mode. Also, in case that downmixing coefficient into 2 ch stereo is transmitted in audio mode whose channel_configuration=5 or 6, the downmixing coefficient is transmitted by using this PCE. In case that downmixing coefficient is transmitted in audio mode whose channel_configuration=7, 11, 12, 13 or 14, DSE which is specified in ISO/IEC 14496-3:2009/AMD 4 is used. (Refer to Chapter 6.2.3, (3))

(b) Information regarding elements in PCE for each audio mode is specified as the following.

Element configuration information in case of audio mode for channelConfiguration＝1

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	1
num_side_channel_elements	-	0
num_back_channel_elements	-	0
num_lfe_channel_elements	-	0
num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	C	0
front_element_tag_select[0]		0

Element configuration information in case of audio mode for channelConfiguration＝2

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	1
num_side_channel_elements	-	0
num_back_channel_elements	-	0
num_lfe_channel_elements	-	0
num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	L and R	1
front_element_tag_select[0]		0

Element configuration information in case of audio mode for channelConfiguration＝3

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	2
num_side_channel_elements	-	0
num_back_channel_elements	-	0
num_lfe_channel_elements	-	0

num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	C	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	L and R	1
front_element_tag_select[1]		0

Element configuration information in case of audio mode for channelConfiguration＝4

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	2
num_side_channel_elements	-	0
num_back_channel_elements	-	1
num_lfe_channel_elements	-	0
num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	C	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	L and R	1
front_element_tag_select[1]		0
back_element_is_cpe[0]	MS	0
back_element_tag_select[0]		1

Element configuration information in case of audio mode for channelConfiguration＝5

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	2
num_side_channel_elements	-	0
num_back_channel_elements	-	1
num_lfe_channel_elements	-	0
num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	C	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	L and R	1
front_element_tag_select[1]		0
back_element_is_cpe[0]	LS and RS	1
back_element_tag_select[0]		1

Element configuration information in case of audio mode for channelConfiguration＝6

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	2
num_side_channel_elements	-	0

num_back_channel_elements	-	1
num_lfe_channel_elements	-	1
num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	C	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	L and R	1
front_element_tag_select[1]		0
back_element_is_cpe[0]	LS and RS	1
back_element_tag_select[0]		1
lfe_element_tag_select[0]	LFE	0

Element configuration information in case of audio mode for channelConfiguration＝7

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	3
num_side_channel_elements	-	0
num_back_channel_elements	-	1
num_lfe_channel_elements	-	1
num_assoc_data_elements	-	1: in case of transmitting either downmixing coefficient or dialogue information, or transmitting both by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	FC	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	FLc and FRc	1
front_element_tag_select[1]		0
front_element_is_cpe[2]	FL and FR	1
front_element_tag_select[2]		1
back_element_is_cpe[0]	BL and BR	1
back_element_tag_select[0]		2
lfe_element_tag_select[0]	BL and BR	0

Element configuration information in case of audio mode for channelConfiguration＝11

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	2
num_side_channel_elements	-	0
num_back_channel_elements	-	2
num_lfe_channel_elements	-	1
num_assoc_data_elements	-	1: in case of transmitting either downmixing coefficient or dialogue information, or transmitting both by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	FC	0

front_element_tag_select[0]		0
front_element_is_cpe[1]	FL and FR	1
front_element_tag_select[1]		0
back_element_is_cpe[0]	BL and BR	1
back_element_tag_select[0]		1
back_element_is_cpe[1]	BC	0
back_element_tag_select[1]		1
lfe_element_tag_select[0]	LFE	0

Element configuration information in case of audio mode for channelConfiguration＝12

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	2
num_side_channel_elements	-	1
num_back_channel_elements	-	1
num_lfe_channel_elements	-	1
num_assoc_data_elements	-	1: in case of transmitting either downmixing coefficient or dialogue information, or transmitting both by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	FC	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	FL and FR	1
front_element_tag_select[1]		0
side_element_is_cpe[0]	SiL and SiR	1
side_element_tag_select[0]		1
back_element_is_cpe[0]	BL and BR	1
back_element_tag_select[0]		2
lfe_element_tag_select[0]	LFE	0

Element configuration information in case of audio mode for channelConfiguration＝13

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	7
num_side_channel_elements	-	3
num_back_channel_elements	-	4
num_lfe_channel_elements	-	2
num_assoc_data_elements	-	1: in case of transmitting either downmixing coefficient or dialogue information, or transmitting both by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	FC	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	FLc and FRc	1
front_element_tag_select[1]		0
front_element_is_cpe[2]	FL and FR	1

front_element_tag_select[2]		1
front_element_is_cpe[3]	TpFC	0
front_element_tag_select[3]		1
front_element_is_cpe[4]	TpFL and TpFR	1
front_element_tag_select[4]		2
front_element_is_cpe[5]	BtFC	0
front_element_tag_select[5]		2
front_element_is_cpe[6]	BtFL and BtFR	1
front_element_tag_select[6]		3
side_element_is_cpe[0]	SiL and SiR	1
side_element_tag_select[0]		4
side_element_is_cpe[1]	TpSiL and TpSiR	1
side_element_tag_select[1]		5
side_element_is_cpe[2]	TpC	0
side_element_tag_select[2]		3
back_element_is_cpe[0]	BL and BR	1
back_element_tag_select[0]		6
back_element_is_cpe[1]	BC	0
back_element_tag_select[1]		4
back_element_is_cpe[2]	TpBL and TpBR	1
back_element_tag_select[2]		7
back_element_is_cpe[3]	TpBC	0
back_element_tag_select[3]		5
lfe_element_tag_select[0]	LFE1	0
lfe_element_tag_select[1]	LFE2	1

Element configuration information in case of audio mode for channelConfiguration＝14

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	3
num_side_channel_elements	-	0
num_back_channel_elements	-	1
num_lfe_channel_elements	-	1
num_assoc_data_elements	-	1: in case of transmitting either downmixing coefficient or dialogue information, or transmitting both by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	FC	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	FL and FR	1
front_element_tag_select[1]		0
front_element_is_cpe[2]	TpFL and TpFR	1
front_element_tag_select[2]		1
back_element_is_cpe[0]	LS and RS	1
back_element_tag_select[0]		2
lfe_element_tag_select[0]	LFE	0

Element configuration information in case of audio mode for channelConfiguration＝0(2/1)

Data Elements	Default element to	Restriction

	speaker mapping	
num_front_channel_elements	-	1
num_side_channel_elements	-	0
num_back_channel_elements	-	1
num_lfe_channel_elements	-	0
num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	L and R	1
front_element_tag_select[0]		0
back_element_is_cpe[0]	MS	0
back_element_tag_select[0]		0

Element configuration information in case of audio mode for channelConfiguration＝0(2/2)

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	1
num_side_channel_elements	-	0
num_back_channel_elements	-	1
num_lfe_channel_elements	-	0
num_assoc_data_elements	-	1: in case of transmitting dialogue information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	L and R	1
front_element_tag_select[0]		0
back_element_is_cpe[0]	LS and RS	1
back_element_tag_select[0]		1

Element configuration information in case of audio mode for channelConfiguration＝0(2 audio(1/0+1/0))

Data Elements	Default element to speaker mapping	Restriction
num_front_channel_elements	-	2
num_side_channel_elements	--	0
num_back_channel_elements	-	0
num_lfe_channel_elements	-	0
num_assoc_data_elements	-	1: in case of transmitting dialog information by DSE, 0: otherwise
num_valid_cc_elements	-	0
front_element_is_cpe[0]	main	0
front_element_tag_select[0]		0
front_element_is_cpe[1]	sub	0
front_element_tag_select[1]		1

--In case of audio mode for channelConfiguration=13 and 14, height_extension_element which is specified in ISO/IEC 14496-3:2009/AMD 4 shall be handled as the following.

height_extension_element in case of audio mode for channelConfiguration=13

Data elements	Default element to speaker mapping	Restriction
front_element_height_info[0]	FC	“0”
front_element_height_info[1]	FLc and FRc	“0”
front_element_height_info[2]	FL and FR	“0”
front_element_height_info[3]	TpFC	“1”
front_element_height_info[4]	TpFL and TpFR	“1”
front_element_height_info[5]	BtFC	“2”
front_element_height_info[6]	BtFL and BtFR	“2”
side_element_height_info[0]	SiL and SiR	“0”
side_element_height_info[1]	TpSiL and TpSiR	“1”
side_element_height_info[2]	TpC	“1”
back_element_height_info[0]	BL and BR	“0”
back_element_height_info[1]	BC	“0”
back_element_height_info[2]	TpBL and TpBR	“1”
back_element_height_info[3]	TpBC	“1”
height_info_crc_check	-	“101”

height_extension_element in case of audio mode for channelConfiguration=14

Data elements	Default element to speaker mapping	Restrictions
front_element_height_info[0]	FC	“0”
front_element_height_info[1]	FL and FR	“0”
front_element_height_info[2]	TpFL and TpFR	“1”
back_element_height_info[0]	LS and RS	“0”
height_info_crc_check	-	“47”

(c) Regarding PCE configuration bit except mentioned-above, it is specified as the following.
--Num_valid_cc_elements shall be 0.
Therefore, the following two flags do not exist.
cc_element_is_ind_sw
valid_cc_element_tag_select
--Mono_mixdown_present shall be 0.
Therefore, mono_mixdown_element_number does not exist.
--Stereo_mixdown_present shall be0.
Therefore, stereo_mixdown_element_number does not exist.

(3) Detailed provision regarding transmission of DSE (Data Stream Element)

(a) When transmitting downmixing coefficients in audio mode for channelConfiguration=7, 11, 12, 13 or 14, DSE specified in ISO/IEC 14496-3:2009/AMD 4 is transmitted by the interval less than 550 ms. In case of this operation, in the period of service when audio mode for channel_configuration=7, 11, 12, 13 or 14 continues, DSE must be always transmitted. And, DSE is transmitted as the following syntactic element to PCE in frame which transmits PCE in raw_data_block(). Also, in case of audio mode for channelConfiguration=5 or 6, downmixing coefficients are not transmitted by DSE.

When transmitting dialogue information, DSE specified in ISO/IEC 14496-3:2009/AMD 4 is also transmitted by the interval less than 550 ms. In case of this operation, in the period when audio mode continues, DSE must be always transmitted. As for detail of dialogue information, refer to (d).

(b) MPEG4_ancillary_data() specified in ISO/IEC 14496-3:2009/AMD 4 shall be handled as the following for audio mode for channelConfiguration=7, 11, 12, 13 and 14. In the table, “Not used” represents that the item is not recorded in bitstream for any setting value of the other parameter.

Restrictions on coding parameters in MPEG4_ancillary_data()

Item	Restriction
ancillary_data_sync;	"0xBC"
mpeg_audio_type	"11"
dolby_surround_mode	"00"
drc_presentation_mode	"00"
stereo_downmix_mode;	"0"
downmixing_levels_MPEG4_status	No restriction
ancillary_data_extension_status;	No restriction
audio_coding_and_compression_status	"0"
coarse_grain_timecode_status	"0"
fine_grain_timecode_status	"0"
center_mix_level_on	No restriction
center_mix_level_value	No restriction
surround_mix_level_on	No restriction
surround_mix_level_value	No restriction
audio_coding_mode_reserved	Not used
compression_on	Not used
compression_value	Not used
coarse_grain_timecode	Not used
fine_grain_timecode	Not used
ext_downmixing_levels_status	In case of audio mode for channelConfiguration=7, 11, 12 and 14, "1" or "0". Otherwise "0".
ext_downmixing_global_gains_status	No restriction
ext_downmixing_lfe_level_status	No restriction
dmix_a_idx	No restriction
dmix_b_idx	No restriction
dmx_gain_5_sign	No restriction
dmx_gain_5_idx	No restriction
dmx_gain_2_sign	No restriction
dmx_gain_2_idx	No restriction
dmix_lfe_idx	No restriction

(c) In case of audio mode for channelConfiguration=7, 11, 12 and 14, the following value specified in ISO/IEC 14496-3:2009/AMD 4 Table AMD4.12 shall have been transmitted, until downmixing coefficients to 5.1 channels are transmitted for the first time.

Default value of downmixing coefficient for audio mode for channelConfiguration=7, 11, 12 and 14

Data field	Default value
dmix_a _idx	"010" (=-3 dB)
dmix_b_idx	"010" (=-3 dB)

(d) Bitstream syntax shown in the following is added for downmixing of audio mode for channelConfiguration=13, and transmitting dialogue information. Bitstream specified in the following is transmitted to the end of MPEG4_ancillary_data() specified in ISO/IEC 14496-3:2009/AMD 4. In case of audio mode except channelConfiguration=13, ext_downmixing_level_status2 is set to "0", and downmixing coefficients (dmix_c_idx, dmix_d_idx, dmix_e_idx, dmix_f_idx, dmix_g_idx, and dmix_l_idx) are not transmitted.

Bitstream syntax which is added to the end of MPEG4_ancillary_data()

Syntax	No. of Bits	Mnemonic
ancillary_data_sync2;	8	bslbf
ext_downmixing_level_status2;	1	bslbf
if (ext_downmixing_level_status2 == 1) {		
dmix_c_idx;	3	bslbf
dmix_d_idx;	3	bslbf
dmix_e_idx;	3	bslbf
dmix_f_idx;	3	bslbf
dmix_g_idx;	3	bslbf
dmix_l_idx	4	bslbf
reserved, set to "0000"	4	bslbf
} else {		
reserved, set to "0000000"	7	bslbf
}		
ext_dialogue_status;	1	bslbf
if (ext_dialogue_status == 1) {		
chans = get_audio_chans(channelConfiguration);		
chn_bits = max(ceil(log(chans)/log(2)),1);		
num_dialogue_chans;	chn_bits	bslbf
sn_dialogue_plus_index;	3	bslbf
sn_dialogue_minus_index;	3	bslbf
dialogue_main_lang_code;	24	uimsbf
dialogue_main_lang_comment_bytes;	8	uimsbf
for(i = 0; i < dialogue_main_lang_comment_bytes; i++){		
dialogue_main_lang_comment_data[i];	8	uimsbf
}		
for(i = 0; i < num_dialogue_chans; i++){		
dialogue_src_index[i];	chn_bits	bslbf
dialogue_gain_index[i];	4	bslbf
}		
num_additional_lang_chans;	4	bslbf
for(i = 0; i < num_additional_lang_chans; i++){		
dialogue_additional_lang_code[i];	24	uimsbf
dialogue_additional_lang_comment_bytes[i];	8	uimsbf
for(j = 0;		
j < dialogue_ additional _lang_comment_bytes; j++){		
dialogue_additional_lang_comment_data[i][j];	8	uimsbf
}		
}		
}		
byte_alignment();		

* ceil() is a helper function that returns the smallest integer which is bigger than a decimal given by argument.
* max(a, b) is a helper function that returns maximum value of a and b given by arguments.

byte_alignment() is a function for adjusting data length to byte unit (a multiple of 8 bits), whose start point shall be at ext_dialogue_status.

Bitstream syntax of DSE where additional dialogue channel is stored

Syntax	No. of Bits	Mnemonic
additional_dialogue_data () {		
additional_dialogue_data_sync;	16	bslbf
additional_dialogue_index;	4	bslbf
single_channel_element();		
byte_alignment();		

}		

Start point of byte_alignment shall be at additional_dialogue_data_sync.

DSE in which additional dialogue channel data is stored is transmitted after PCE in raw_data_block(), DSE in which MPEG4_ancillary_data() is stored, all SCEs and CPEs. Here, element_instance_tag of single_channel_element() which is involved in DSE in which additional dialogue channel data is stored is not specified. Also, not depending on the number of DSE in which single_channel_element() is stored, the value of num_assoc_data_elements is always transmitted as only the number of DSE in which MPEG4_ancillary_data() is stored, that is "1" or "0". For detail, refer to the table about element configuration information recorded in Chapter 6.2.3, (2).

Bitstream syntax of helper function get_audio_chans() which acquires the number of main audio channels

Syntax	No. of Bits	Mnemonic

get_audio_chans(channelConfiguration){
  return audio_chans_table[channelConfiguraion];
}		

Corresponding table of audio_chans_table

channelConfiguration	Number of audio_chans
1	1
2	2
3	3
4	4
5	5
6	6
7	8
11	7
12	8
13	24
14	8

Terms in bitstream syntax mentioned-above are explained in the following.

ancillary_data_sync2
    This shall be "0xBD".

ext_downmixing_levels_status2
    This represents whether downmixing coefficient exists or not
    in case of audio mode for channelConfiguration=13.
    This shall be "1" or "0".

dmix_c_idx, dmix_d_idx, dmix_e_idx, dmix_f_idx, dmix_g_idx
    These represent index of downmixing coefficients from 22.2 ch to 5.1 ch. Then, as index, Table AMD4.8 specified in ISO/IEC 14496-3:2009/AMD 4 is used.

dmix_l_idx
    This represents index of LFE downmixing coefficient from 22.2 ch to 5.1 ch. Then, as index, Table AMD4.9 specified in ISO/IEC 14496-3:2009/AMD 4 is used.

ext_dialogue_status
    This represents whether dialogue information exists or not. This shall be "1" or "0".

num_dialogue_chans
    This represents the number of the dialog channel.

num_additional_lang_chans;
    This represents the number of additional dialogue.

dialogue_src_index[i]
    This represents index of the dialogue channel. The value of subtracting 1 from the value of index mapping for dialogue_src_index specified in Chapter 6.2.3, (1) is used.

dialogue_main_lang_comment_bytes

This represents the number of bytes of character string information for representing the contents of main dialogue.

dialogue_main_lang_comment_data

This represents byte data of character string information for representing the contents of main dialogue.

dialogue_main_lang_code

This represents language code of main dialogue. Code value is based on ISO 639-2, and the value defined in ISO/IEC 8859-1 is used for character.

An example of languages

language	Code which is stored in dialogue_main_lang_code and dialogue_additional_lang_code[i]
Japanese	“jpn”(0x6A,0x70,0x6E)
English	“eng”(0x65,0x6E,0x67)
French	“fre”(0x66,0x72,0x65) or ”fra”(0x66,0x72,0x61)
Germany	“ger”(0x67,0x65,0x72) or ”deu”(0x64,0x65,0x75)

dialogue_additional_lang_code[i]

This represents language code of additional dialogue. Code value is based on ISO 639-2, and the value defined by ISO/IEC 8859-1 is used for character.

dialogue_additional_lang_comment_bytes[i]

This represents the number of bytes of character string information for representing the contents of i th additional dialogue.

dialogue_additional_lang_comment_data[i]

This represents the byte data of character string information for representing the contents of i th additional dialog.

dialogue_gain_index[i]

This represents index of gain compensation value for additional dialogue.

dialogue_gain_index	Multiplication factor
0000	1 (0dB)
0001	0.891 (-1dB)
0010	0.794 (-2dB)
0011	0.708 (-3dB)
0100	0.631 (-4dB)
0101	0.562 (-5dB)
0110	0.501 (-6dB)
0111	0.447 (-7dB)
1000	0.398 (-8dB)
1001	0.355 (-9dB)
1010	0.316 (-10dB)
1011	0.282 (-11dB)
1100	0.251 (-12dB)
1101	0.224 (-13dB)
1110	0.200 (-14dB)
1111	0.000 (-∞dB)

sn_dialogue_plus_index

This represents upper limit for gain control in the receiver.

sn_dialogue_plus_index	Multiplication factor
000	1 (0dB)
001	1.413 (+3dB)
010	1.995 (+6dB)
011	2.818 (+9dB)
100	3.981 (+12dB)
101	5.623 (+15dB)
110	7.943 (+18dB)
111	+∞ (+∞dB)

sn_dialogue_minus_index

This represents lower limit for gain control in the receiver

sn_dialogue_minus_index	Multiplication factor
000	1 (0dB)
001	0.708 (-3dB)
010	0.501 (-6dB)
011	0.355 (-9dB)
100	0.251 (-12dB)
101	0.178 (-15dB)
110	0.126 (-18dB)
111	0.000 (-∞dB)

additional_dialogue_data_sync

This represents DSE in which additional dialogue data is stored. The value shall be "0xED01".

additional_dialogue_index

This represents index for identifying additional dialogue. additional_dialogue_index of data corresponding to dialogue_additional_lang_code[0] shall be "0", additional_dialogue_index of data corresponding to dialogue_additional_lang_code[1] shall be "1", and the value of x in dialogue_additional_lang_code[x] shall be the value of additional_dialogue_index.

Downmixing from 22.2 ch to 5.1 ch using coefficient index mentioned-above is specified as the following.

C' = FC+g1*FLc+g1*FRc+g3*(TpFC+g4*TpC+BtFC)

L' = FL+g1*FLc+g2*SiL+g3*(TpFL+g2*TpSiL+BtFL)

R' = FR+g1*FRc+g2*SiR+g3*(TpFR+g2*TpSiR+BtFR)

Ls' = BL+g5*BC+g2*SiL+g3*(TpBL+g5*TpBC+g2*TpSiL+g4*TpC)

Rs' = BR+g5*BC+g2*SiR+g3*(TpBR+g5*TpBC+g2*TpSiR+g4*TpC)

LFE' = g6*(LFE1+LFE2)

Here, g1, g2, g3, g4, and g5 are obtained from dmix_c_idx, dmix_d_idx, dmix_e_idx, dmix_f_idx, and dmix_g_idx respectively, by using Table AMD4.8 specified in ISO/IEC 14496-3:2009/AMD 4. Also, g6 is obtained from dmix_l_idx by using Table AMD4.9 specified in ISO/IEC 14496-3:2009/AMD 4. Until these downmixing coefficients from 22.2 ch to 5.1 ch are transmitted for the first time, the following values shall have been transmitted.

Default value of downmix coefficient from 22.2 ch to 5.1 ch

Data field	Default value

dmix_c _idx	"011" (=-4.5dB)
dmix_d_idx	"011" (=-4.5dB)
dmix_e_idx	"000" (=0dB)
dmix_f_idx	"100" (=-6dB)
dmix_g_idx	"010" (=-3dB)
dmix_l_idx	"0111" (=-3dB)

(e) About transmitting downmixing coefficients to 2 ch stereo

In case of 2 ch stereo reproduction from multichannel stereo more than 5.1 ch stereo (audio mode for channelConfiguration=7, 11, 12, 13 or 14) by downmixing, after downmixing to 5.1 ch stereo once, the 5.1 ch stereo shall be downmixed to 2 ch stereo. When transmitting downmixing coefficients from 5.1 ch stereo to 2 ch stereo, DSE specified in ISO/IEC 14496-3:2009/AMD 4 is used. Also, in case of audio mode for channelConfiguration=7, 11, 12, 13 or 14, downmixing coefficients is not transmitted by PCE.

(f) About transmitting dialogue information

Dialogue information is transmitted as ext_dialogue_status=1 for transmitting dialogue information.
Until dialogue information is transmitted for the first time, the following values shall have been transmitted.

Default value of dialogue information

Data field	Default value
ext_dialogue_status	0
num_dialogue_chans	0
num_additional_lang_chans	0

An example of dialogue channel control for 22.2 ch sound using above-mentioned data is shown. On the assumption that FC and BtFC of 22.2 ch sound are the dialogue channel for Japanese, the distribution level of additional dialogue is FC: -3dB, BtFC: -0dB, the adjustment range of dialogue level is from +12dB to -∞dB, and the additional dialogues are English and French, the following values shall be used.

num_dialogue_chans = 2

dialogue_main_lang_code = "jpn" (0x6A,0x70,0x6E)

dialogue_src_index[0] =0

dialogue_src_index[1] =19

dialogue_gain_index [0] =-3dB

dialogue_gain_index[1]= 0dB

num_additional_lang_chans =2;

dialogue_additional_lang_code[0] = "eng"(0x65,0x6E,0x67)

dialogue_additional_lang_code[1] = "fre"(0x66,0x72,0x65)

sn_dialogue_plus_index = "100"(+12dB)

sn_dialogue_minus_index = "010"(-∞dB)

(f).1 Level control of dialogue

The receiver adjusts level of 22.2 ch audio signal by receiving command of changing the sound volume of dialogue given from the outside. When receiving command of raising the dialogue channels FC and BtFC x dB from the initial level, the receiver lowers each level of 20.2 ch except FC and BtFC by x dB within the range of $0dB \leqq x \leqq +12dB$ which sn_dialogue_plus_index indicates. On the other hand, when receiving command of lowering the dialogue channels FC and BtFC by x dB from the initial level, the receiver lowers each level of FC and BtFC by xdB within the range of $-\infty dB \leqq -x \leqq 0dB$ which sn_dialogue_minus_index indicates.

(f).2 Replacement of dialogue

The receiver replaces Japanese dialogue which was initially in FC and BtFC with English or French dialogue by receiving command of replacement of dialogue given from the outside. When receiving command of replacement of English dialogue, the receiver assigns English dialogue which is lowered 3 dB level to FC, and English dialogue which is lowered 0 dB level to BtFC in stead of Japanese dialogue, by referring dialogue_gain_index [0](-3 dB) which indicates assign level to FC and dialogue_gain_index [1](0 dB) which indicates assign level to BtFC. Also, about level control of dialogue, the mentioned-above procedure is carried out after processing of dialogue replacement.

6.3 Stream format for MPEG-4 AAC System

More restrictions on operations are provided about LATM/LOAS stream format and data stream format which are specified as stream format for transmitting audio coding information by MPEG-4 AAC System. Also, restrictions on operations are provided about ADTS stream format which can be used for V-Low multimedia broadcasting by connected segment system.

6.3.1 Restrictions on LATM/LOAS stream format

LATM/LOAS frame comprises 1 raw_data_block which is specified in section 6.2.

Here, "Not used" in the table represents that the item is not recorded in bitstream for any setting value of the other parameters.

(1) Provisions on input audio mode and method of configuration and multiplex for LATM/LOAS

Input audio mode	Method of configuration and multiplex for LATM/LOAS
mono, stereo	Comprises 1 LATM/LOAS
multichannel stereo	Comprises 1 LATM/LOAS
multiple audio signals (2/0+2/0, etc.)	Comprises the same number of LATM/LOAS as the number of audio streams (languages) and is multiplexed with the MPEG-4 systems layer.
2 audio signals (dual mono) (Note)	Comprises 1 LATM/LOAS

(Note) Dual mono is defined as "two monophonic audio channels that can be simultaneously reproduced by a single LATM/LOAS."

(2) Header of LATM/LOAS

Item	Restriction
Synchronization Layer	AudioSyncStream() is used
Multiplex Layer	AudioMuxElement() is used

AudioMuxElement()

Item	Restriction
useSameStreamMux	0 (StreamMuxConfig() is transmitted every frame)
otherDataBit	Not used

StreamMuxConfig()

Item	Restriction
audioMuxVersion	0
allStreamsSameTimeFraming	1
numSubFrames	0 (number of subframe in one frame＝1)
numProgram	0 (number of program in one frame＝1)
numLayer	0 (number of layer in one frame＝1)
fillBits	Not used
frameLengthType[0]	0 (Payload with variable frame length)
latmBufferFullness[0]	0xFF (represents variable rate) is prohibited to use
coreFrameOffset	Not used
frameLength[]	Not used
CELPframeLengthTableIndex[]	Not used
HVXCframeLengthTableIndex[]	Not used
otherDataPresent	0 (otherDataBit is not used)
otherDataLenEsc	Not used
otherDataLenTmp	Not used
crcCheckPresent	1 (CRC error check is carried out)

(3) Detailed provisions on transmitting PCE (Program Configuration Element)

(a) In case of audio mode with channelConfiguration=0 (2/1, 2/2, 2-audio signals (dual mono) (1/0+1/0)) in LATM/LOAS header, PCE in LATM/LOAS header is transmitted every frame for transmitting the audio mode. When the audio mode is except channelConfiguration=0 in LATM/LOAS header, as PCE cannot be transmitted by LATM/LOAS header, the audio mode is transmitted by PCE in raw_data_block(). (Refer to Chapter 6.2.3, (2).)

(b) In case of audio mode with channelConfiguration=0 (2/1, 2/2, 2-audio signals (dual mono) (1/0+1/0)), PCE in LATM/LOAS header shall agree with PCE in raw_data_block().

(c) sampling_frequency_index of PCE in raw_data_block() shall agree with samplingFrequencyIndex in LATM/LOAS header.

(d) Audio mode of PCE in raw_data_block() shall agree with audio mode in LATM/LOAS header. (Refer to Chapter 6.2.3, (2) about audio mode of PCE in raw_data_block())

(e) It is permitted to put PCE in every LATM frame, but change of parameter values is prohibited except in necessary (such as change of audio mode, change of coefficients, etc.)

6.3.2 Restrictions on data stream format

Raw Data Stream which is specified in Chapter 6.2 shall be output.

Provisions on input audio mode and method of configuration and multiplex for data stream

Input audio mode	Method of configuration and multiplex for data stream
mono, stereo	Comprises one Raw Data Stream
multichannel stereo	Comprises one Raw Data Stream
multiple audio signals (such as 2/0+2/0)	Comprises the same number of Raw Data Streams as the number of audio streams (languages) and is multiplexed with the MPEG-4 systems layer.
2-audio signals (dual mono) (Note)	Comprises one Raw Data Stream

(Note) Dual mono is defined as “two monophonic audio channels that can be simultaneously reproduced by a single Raw Data Stream.”

6.3.3 Restrictions on ADTS stream format

ADTS frame (Note) specified in ISO/IEC 14496-3 Annex1.A is composed of 1 raw_data_block which is specified in Chapter 6.2. ADTS can be used in V-Low multimedia broadcasting by connected segment system.

(Note) ADTS configuration in ISO/IEC 14496-3 is fundamentally the same as ADTS configuration specified in ISO/IEC13818-7.

(1) Provisions on input audio mode and method of construction and multiplex for ADTS

mono, stereo	Comprises one ADTS
multi-channel stereo	Comprises one ADTS
multiple audio signals (such as 2/0+2/0)	Comprises the same number of ADTS as the number of audio streams (languages) and is multiplexed with the MPEG-4 systems layer.
2-audio signals (dual mono) (Note)	Comprises one ADTS

(Note) Dual mono is defined as "two monophonic audio channels that can be simultaneously reproduced by a single ADTS."

(2) Fixed header of ADTS

protection_absent	0 (CRC check is attached.)
ID	1 (MPEG-4 AAC)
Profile_ObjectType	1 (AAC LC object)
Sampling_frequency_index	0 (96kHz), 3 (48kHz), 6 (24kHz)
Channel_configuration	Refer to Chapter 6.2.3

(3) Variable header of ADTS

adts_buffer_fullness	0x7FF (which represents variable rate) is prohibited to use
number_of_raw_data_block_in_frame	0 (raw_data_block number = 1 in 1 frame)

(4) Detailed provisions on construction of Fill Element (FIL)

(a) In case that sampling_frequency_index in ADTS fixed header is 0x3 (48kHz) and 0x6 (24kHz), EXT_SBR_DATA ('1101') and EXT_SBR_DATA_CRC ('1110') can be used in Fill Element (FIL).

(b) EXT_SAC_DATA('1100') can be used in Fill Element (FIL).

6.4 Compatibility with the receiver when multichannel stereo service is provided

6.4.1 Compatibility in multichannel stereo service of below 5.1 ch stereo

When multichannel stereo service of 5.1 ch stereo (3/2+LTE (3/2.1)) or less is provided, consideration about the compatibility with the receiver for 2 ch stereo is as the following.

(1) When 5 ch stereo or 5.1 ch stereo is provided, according to AAC standard, it shall be possible that downmixing coefficients are transmitted by using PCE. Refer to 6.2.3 (2) about detailed provision on transmission of PCE.

(2) About downmixing in the receiver for 2 ch stereo which does not depend on item (1) mentioned-above, refer to ARIB STD-B21.

(3) When multichannel stereo below 5.1 ch stereo is provided, it shall be possible that 2 ch stereo simulcast service is provided. In this case, the format shall be those which is provided in Chapter 6.3, and multiplexed in system layer and managed as a stream.

6.4.2 Compatibility in multichannel stereo service of more than 5.1 ch stereo

When multichannel stereo service of more than 5.1 ch stereo (3/2.1) is provided, consideration about the compatibility with the receiver for 5.1 ch stereo and 2 ch stereo is as the following.

(1) When multichannel stereo of more than 5.1 ch stereo is provided, it shall be possible that downmixing coefficients to 5.1 ch stereo are transmitted by using DSE, according to Chapter 6.2.3 (3).

(2) About downmixing in the receiver for 5.1 ch stereo in case that it does not depend on item (1) mentioned-above, the default value specified in ISO/IEC 14496-3:2009/AMD 4 or the default value recorded in Chapter 6.2.3 (3) (c) and (d) is used.

(3) About downmixing of the receiver for 2 ch stereo, after downmixing to 5.1 ch stereo by item (1) or (2) mentioned-above, downmixing to 2 ch stereo is carried out according to ISO/IEC 14496-3:2009/AMD 4 or ARIB STD-B21.

(4) When multichannel stereo of more than 5.1 ch stereo (3/2.1) is provided, it shall be possible that simulcast service by both 5.1 ch stereo and 2 ch stereo, or either of them is carried out. In this case, the format shall be those which is provided in Chapter 6.3, and multiplexed in system layer and managed as streams.

Chapter 7: Restrictions on MPEG-4 ALS Lossless Audio Coding Parameters

This chapter specifies lossless audio coding system for digital broadcasting based on MPEG-4 ALS system. Audio input format in Chapter 7.1, restrictions on coding parameters of MPEG-4 ALS system in Chapter 7.2, restrictions on stream format of MPEG-4 ALS system in Chapter 7.3, and transmission procedure of stream format in Chapter 7.4 are described.

7.1 Input Audio format based on MPEG-4 ALS System

Restrictions on audio input format for digital broadcasting shall be as the following.

Item	Restriction
Audio mode	・mono ・stereo ・multichannel stereo (Note) 3/0, 2/1, 3/1, 2/2, 3/2, 3/2.1, 5/2.1, 3/3.1, 3/2/2.1, 2/0/0-3/0/2-0.1, 3/3/3-5/2/3-3/0/0.2 ・2-audio signals (dual mono)
Emphasis	none

(Note) Notation of audio mode for multichannel stereo:

The number of channel is represented as

"upper layer (front/side/back)-middle layer (front/side/back)-lower layer (front/side/back).LFE".

But the layer which does not have any allocated channel is denoted as 0. Also, audio mode by only middle layer is denoted as "middle layer (front/side/back).LFE", and multichannel stereo which is only by middle layer, without side channel is simply denoted as "middle layer (front/back).LFE".

When the allocated channel to LFE (low frequency effect channel) is one, there is a case that it is denoted as "+LFE".

There is a related record in Description 2 about notation for audio mode.

7.2 Coding parameters for MPEG-4 ALS System

MPEG-4 ALS System is provided in Ordinance as audio coding system for digital broadcasting (refer to Chapter 3.4), but in this section, for realizing digital broadcasting service, more restrictions on operation are provided.

Also, MPEG-4 ALS system is provided as MPEG-4 ALS (Audio Lossless Coding) in ISO/IEC 14496-3:2009 (Information technology -- Coding of audio-visual objects -- Part 3: Audio).

7.2.1 Main parameters

Item	Restriction
Profile	Usable tool is used in ALS Simple Profile (Refer to the following.)
Audio object type	36 (ALS)
Maximum number of channels	Max. 22.2 channels (Note) per 1 frame_data()

(Note) 22 channels + 2 LFE channels

Allowed tools are described in ALS Simple Profile which is defined in ISO/IEC 14496-3:2009 Amd 2:2010 (Information technology – Coding of audio-visual objects – Part 3: Audio AMENDMENT2: ALS simple profile and transport of SAOC).

Allowed tools in ALS Simple Profile shall be defined as the following.

Item	Restriction
Maximum number of samples per frame	4096
Maximum prediction order	15
Maximum number of stages for BS(Block switching) tool	3
Maximum number of stages for MCC(Multi-channel coding) tool	1
BGMC tool	Not used
RLS-LMS tool	Not used
sampling frequency, number of quantizing bits, and number of audio channels	These comply with audio input signal described in Chapter 2. (But the number of quantized bits shall be 32 in maximum.)

7.2.2 Restrictions on MPEG-4 Audio parameters

For MPEG-4 Audio, parameters in the coding system are set by using AudioSpecificConfig(). For using MPEG-4 ALS System, restrictions are provided on setting parameters.

Also, “Not used” in the table represents that the item is not recorded in the bitstream depending on the other parameters.

AudioSpecificConfig()

Item	Restriction
samplingFrequencyIndex	0: 96000Hz (Note 1) 3: 48000Hz
samplingFrequency	Not used
channelConfiguration	1: 1ch (1/0) 2: 2ch (2/0) 3: 3ch (3/0) 4: 4ch (3/1) 5: 5ch (3/2) 6: 5.1ch (3/2.1) 7: 7.1ch (5/2.1) 11: 6.1ch (3/0/3.1) 12: 7.1ch (3/2/2.1) 13: 22.2ch (3/3/3-5/2/3-3/0/0+2)

	14: 7.1ch (2/0/0-3/0/2-0/0/0+1) 0: Coding mode corresponding to the specified value in channels of ALSSpecificConfig() is used. (3ch(2/1), 4ch(2/2) or 2-audio signals (dual mono; when (1/0+1/0)))
extensionSamplingFrequencyIndex	Not used
extensionSamplingFrequency	Not used
extensionChannelConfiguration	Not used
GASpecificConfig()	Not used
CelpSpecificConfig()	Not used
HvxcSpecificConfig()	Not used
TTSSpecificConfig()	Not used
StructuredAudioSpecificConfig()	Not used
ErrorResilientCelpSpecificConfig()	Not used
ErrorResilientHvxcSpecificConfig()	Not used
ParametricSpecificConfig()	Not used
SSCSpecificConfig()	Not used
sacPayloadEmbedding	Not used
SpatialSpecificConfig()	Not used
MPEG_1_2_SpecificConfig()	Not used
DSTSpecificConfig()	Not used
fillBits	Used (Note 2)
SLSSpecificConfig()	Not used
ELDSpecificConfig()	Not used
SymbolicMusicSpecificConfig()	Not used
epConfig	Not used
ErrorProtectionSpecificConfig()	Not used
directMapping	Not used
syncExtensionType	Not used
sbrPresentFlag	Not used
extensionSamplingFrequencyIndex	Not used
extensionSamplingFrequency	Not used
syncExtensionType	Not used
psPresentFlag	Not used
extensionChannelConfiguration	Not used

(Note 1) Only for V-Low multimedia broadcasting by connected segment system, it shall be possible to operate by 0:96000Hz.

(Note 2) fillBits shall be used for byte alignment (adjusting data length to byte unit (multiple of 8 bit)) of ALSSpecificConfig(), and start point shall be at AudioSpecificConfig().

Return value of GetAudioObjectType() is 36 (ALS)

Item	Restriction
audioObjectType	31
audioObjectTypeExt	4

ALSSpecificConfig()

Item	Restriction
als_id	0x414C5300
samp_freq	48000
samples	0xFFFFFFFF
channels	The number of input channel-1 (The value corresponding to coding mode specified in Chapter 7.2.3 is set.)
file_type	000 (unknown/raw file)
resolution	Any of 001, 010, 011
floating	0 = integer
msb_first	0 or 1
frame_length	Frame length -1
random_access	1 (Every frame shall be random accessable.)
ra_flag	00 or 01
adapt_order	0 or 1
coef_table	0 or 1
long_term_prediction	0 or 1
max_order	Less than or equal to 15
block_switching	00 or 01
bgmc_mode	0
sb_part	0 or 1
joint_stereo	0 or 1
mc_coding	0 or 1
chan_config	0
chan_sort	0 or 1
crc_enabled	0
RLSLMS	0
aux_data_enabled	0 or 1
chan_config_info	Not used
chan_pos[]	This depends on the value of chan_sort.
header_size	0
trailer_size	0
orig_header[]	Not used
orig_trailer[]	Not used
crc	Not used
ra_unit_size[]	Not used
aux_size	This depends on the value of aux_data_enabled.
aux_data	This depends on the value of aux_data_enabled.

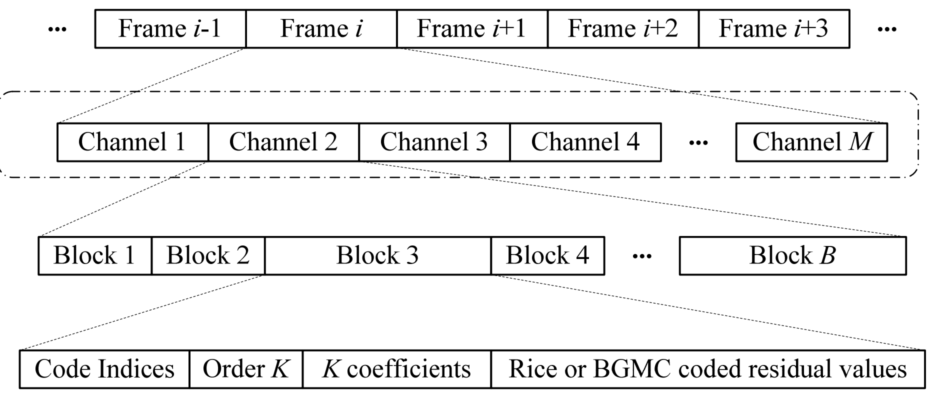
Raw Data Stream

Item	Restriction
Configuration of Raw Data Stream	Comprises frame_data() defined as ALS top level payload in ISO/IEC 14496-3:2009 Subpart 11

7.2.3 Detailed provisions on Channel Configuration and Speaker Mapping configuration

Coding mode based on MPEG-4 Audio standard ISO/IEC 14496-3:2009 is used.

Configuration of coding bitstream by MPEG-4 ALS, especially correspondence to logic channel number is shown as the following. In the figure, Channel 1 to Channel M enclosed by chain line with dot represents each section of bitstream corresponding to logic channel number 1 to M.



Configuration of coded bitstream of MPEG-4 ALS

Coding mode specified in MPEG-4 Audio standard and the value of channels designated in ALSSpecificConfig(), and correspondence between logic channel number (Channel no) in ALS coded bitstream and speaker mapping are shown in the following. When chan_sort is enabled, the logic channel number after restoration to the order of input channel by referring to chan_pos[] shall be corresponded to speaker mapping.

Coding mode which is provided in MPEG-4 Audio Standard as default

Coding mode	channel_configuration	Value of channels in ALSSpecificConfig() (number of channel -1 is designated.)	Correspondence between logic channel number in ALS coded bitstream and speaker mapping (Note 1) (Note 2) (Note 3)
mono (1/0)	1	0	1:C
stereo (2/0)	2	1	1:L 2:R
3/0	3	2	1:C 2:L 3:R
3/1	4	3	1:C 2:L 3:R 4:MS
3/2	5	4	1:C 2:L 3:R 4:LS 5:RS
3/2.1	6	5	1:C 2:L 3:R 4:LS 5:RS 6:LFE

5/2.1	7	7	1:FC 2:FLc 3:FRc 4:FL 5:FR 6:BL 7:BR 8:LFE
3/3.1	11	6	1:FC 2:FL 3:FR 4:BL 5:BR 6:BC 7:LFE
3/2/2.1	12	7	1:FC 2:FL 3:FR 4:SiL 5:SiR 6:BL 7:BR 8:LFE
3/3/3-5/2/3-3/0/0.2	13	23	1:FC 2:FLc 3:FRc 4:FL 5:FR 6:SiL 7:SiR 8:BL 9:BR 10:BC 11:LFE1 12:LFE2 13:TpFC 14:TpFL 15:TpFR 16:TpSiL 17:TpSiR 18:TpC 19:TpBL 20:TpBR 21:TpBC 22:BtFC 23:BtFL 24:BtFR
2/0/0-3/0/2-0.1	14	7	1:FC 2:FL 3:FR 4:LS 5:RS 6:LFE 7:TpFL 8:TpFR

Coding mode other than MPEG-4 Audio default provision

Coding mode	channel_configuration	Value of channels (Note 1) in ALSSpecificConfig()	Correspondence between logic channel number in ALS coded bitstream and speaker mapping (Note 2) (Note 3)
2/1	0	2	1:L 2:R 3:MS
2/2	0	3	1:L 2:R 3:LS 4:RS
2-audio signals (1/0+1/0)	0	1	1:main 2:sub

(Note 1) value of channels: actual channel number-1 is set.

(Note 2) Notation of speaker arrangement: channel_configuration＝1~6
L: Left front speaker / R: Right front speaker / C: Center front speaker / LFE: Low frequency effects / LS: Left surround speaker / RS: Right surround speaker / MS: Mono surround speaker

(Note 3) Notation of speaker arrangement: channel_configuration＝7、11~14
Based on acoustic channel label in ARIB STD-B59 “Three dimensional Multichannel Stereophonic Sound System for Programme Production”

7.3 Restrictions on stream format for MPEG-4 ALS System

Restrictions for operations on LATM/LOAS stream format and data stream format, which are defined as a stream format to transmit audio coding information of MPEG-4 ALS, are provided.

7.3.1 Restrictions on LATM/LOAS stream format

LATM/LOAS frame is composed of frame_data() specified in this standard, section 7.2.

Here, “Not used” in the table represents that the item is not recorded in bitstream according to the setting value of the other parameters.

(1) Provision on input audio mode and method of configuration and multiplex for LATM/LOAS

Input audio mode	Method of configuration and multiplex for LATM/LOAS
mono, stereo	Comprises one LATM/LOAS
multi-channel stereo	Comprises one LATM/LOAS
multiple audio signals (such as 2/0+2/0)	Comprises the same number of LATMs/LOASs as the number of audio streams (languages) and is multiplexed with the MPEG-4 systems layer.
2-audio signals (dual mono) (Note)	Comprises one LATM/LOAS

(Note) Dual mono is defined as “two monophonic audio channels that can be simultaneously reproduced from a single LATM/LOAS.” Here, as only a part of channels cannot be selectively decoded, in case of 2-audio signals (dual mono), 2 channels are decoded simultaneously and one channel is used.

(2) Header of LATM/LOAS

Item	Restriction
Synchronization Layer	AudioSyncStream() is used.
Multiplex Layer	AudioMuxElement() is used.

AudioMuxElement()

Item	Restriction
useSameStreamMux	0 or 1 (StreamMuxConfig() is transmitted only for the first part of random-accessible AudioMuxElement().)
otherDataBit	Not used

StreamMuxConfig()

Item	Restriction
audioMuxVersion	0
allStreamsSameTimeFraming	1 or 0 (When frame_data() is large and the size of AudioMuxElement() exceeds 8191 bytes, frame_data() is divided into multiple Payload[]. This field shall be 0 to indicate that one frame_data() is transmitted in multiple AudioSyncStream().)
numSubFrames	0 (number of sub frame in 1 frame＝1)
numProgram	0 (number of program in 1 frame＝1)
numLayer	0 (number of layer in 1 frame＝1)
fillBits	Not used
frameLengthType[0]	0 (Payload with variable frame length)
latmBufferFullness[0]	0xFF (representing variable rate) is used.
coreFrameOffset	Not used
frameLength[]	Not used
CELPframeLengthTableIndex[]	Not used
HVXCframeLengthTableIndex[]	Not used
otherDataPresent	0 (otherDataBit is not used.)
otherDataLenEsc	Not used
otherDataLenTmp	Not used
crcCheckPresent	1 (CRC error check is done.)

PayloadLengthInfo()

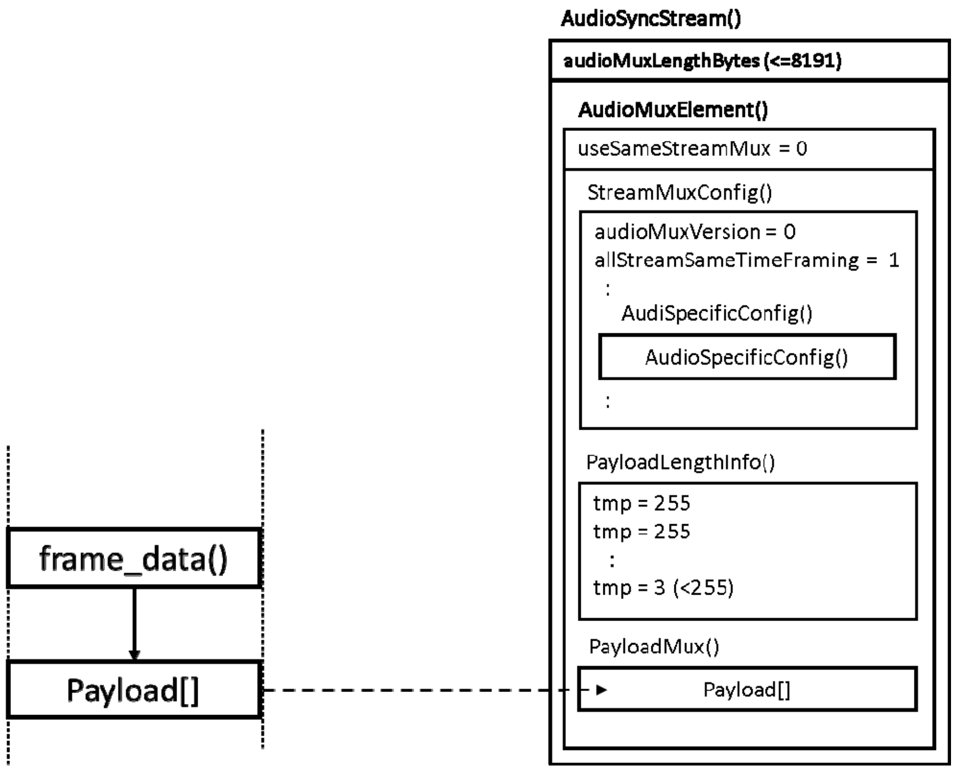
Item	Restriction
tmp	The value which is added to MuxSlotLengthBytes[0]
MuxSlotLengthBytes[0]	Size of PayloadMux()
MuxSlotLengthCoded[]	Not used
numChunk	0 (number of chunks =1) or not used.
streamIndx	0 or not used.
AuEndFlag[]	0, 1, or not used.
MuxSlotLengthCoded[]	Not used

PayloadMux()

Item	Restriction
payload[0]	A part or all of Raw Data Stream is stored (Refer to Chapter 7.2.2 and next section)

(3) An example of composing LATM/LOAS frame

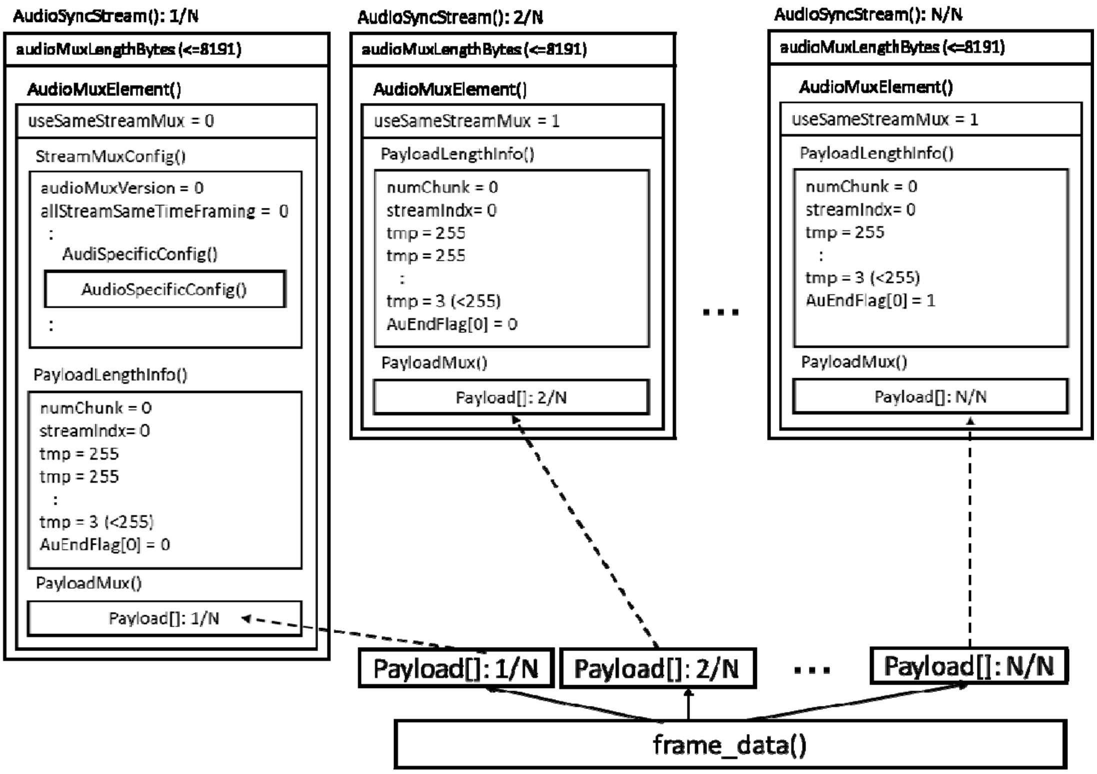
In case that the size of AudioMuxElement() is less than or equal to 8191 bytes on account of the size of Raw Data Stream (=frame_data()), LATM/LOAS frame is configured by making one frame_data() correspond to one AudioSyncStream() as the following figure.



An example of configuring one LATM/LOAS frame by one frame_data()

Depending on the size of Raw Data Stream (=frame_data()), the size of AudioMuxElement(), which includeds frame_data() may exceeds 8191 bytes. In this case, one frame_data() shall be divided into multiple Payload[], which make multiple AudioSyncStream() in LATM/LOAS frames. Only the AudioSyncStream() which includes the first byte of the frame_data() shall have useSameStreamMux = 0, and StreamMuxConfig() and AudioSpecificConfig() shall be included. in addition, allStreamSameTimeFraming shall be 0, which means the Payload[] is a part of divided frame_data(). Only the AudoSyncStream() which includes the last byte of the frame_data() shall have AuEndFlag[] = 1, and all others shall have AuEndFlag[] = 0.

The decoder shall concatenate all divided Payload[] in transmission order to reconstruct the frame_data(). When one or more parts are lost in the transmission, the whole frame_data() shall be dropped.



An example of configuring multiple LATM/LOAS frames by dividing one frame_data()

7.3.2 Restrictions on data stream format

Raw Data Stream provided in Chapter 7.2 must be transmitted.

(1) Provision on input audio mode and method of configuration and multiplex for data stream

Input audio mode	Method of configuration and multiplex for data stream
mono, stereo	Comprises one Raw Data Stream
multi-channel stereo	Comprises one Raw Data Stream
multiple audio signals (such as 2/0+2/0)	Comprises the same number of Raw Data Streams as the number of audio streams (languages) and is multiplexed with the MPEG-4 systems layer.
2-audio signals (dual mono) (Note)	Comprises one Raw Data Stream

(Note) Dual mono is defined as “two monophonic audio channels that can be simultaneously reproduced by a single Raw Data Stream.” Here, as only a part of channel cannot be selectively decoded in ALS, in case of 2-audio signals (dual mono), 2 channels are decoded simultaneously and one channel is used.

<Blank Page>

Annex A Technical system applied to Digital Broadcasting

Technical system applied to each standard system for digital broadcasting which is provided in Ordinance (Ordinance of the Ministry of Internal Affairs and Communications No.87, 2011 or Ordinance of the Ministry of Internal Affairs and Communications No.94, 2011) is shown in Table A-1.

Table A-1 Technical system applied to standard system (○: applied)

	Digital Broadcasting	Digital Terrestrial Sound Broadcasting	Digital Terrestrial Television Broadcasting	V-High Multimedia Broadcasting (Note1)	V-Low Multimedia Broadcasting (Note1)	BS Digital Broadcasting	Advanced BS Digital Broadcasting	Narrow band CS Digital Broadcasting	Wide band CS Digital Broadcasting	Advanced Narrow band CS Digital Broadcasting	Advanced Wide band CS Digital Broadcasting
Input audio format	Sampling frequency	32 kHz, 44.1kHz, 48 kHz	32 kHz, 44.1kHz, 48 kHz	32 kHz, 44.1kHz, 48 kHz	more than or equal to 32 kHz	32 kHz, 44.1kHz, 48 kHz	48kHz	32 kHz, 44.1kHz, 48 kHz	32 kHz, 44.1kHz, 48 kHz	32 kHz, 44.1kHz, 48 kHz	48kHz
	Maximum audio input channels	5.1ch	5.1ch	5.1ch	5.1ch	5.1ch	22.2ch	5.1ch	5.1ch	22.2ch (Note 2)	22.2ch
Audio coding system	MPEG-2 AAC	○	○	○	○	○		○	○	○	
	MPEG-2 BC							○			
	MPEG-4 AAC				○		○			○	○
	MPEG-4 ALS				○		○			○	○

(Note 1) Multimedia broadcasting by connected segment system
(Note 2) In MPEG-2 AAC, maximum number of audio input channel is limited to 5.1 ch according to restrictions on coding parameters. (Refer to Chapter 5)

<Blank Page>

Attachment: Operational Guidelines

Attachment: Operational Guidelines

Chapter 1: General

1.1 Objective

The purpose of these guidelines is to present recommended technical requirements for practical operations regarding audio signal and audio coding systems for digital broadcasting.

1.2 Scope

These guidelines apply to digital broadcasting that comply with the "Standard transmission system for digital broadcasting among standard TV broadcasting and the like" (Ordinance of the Minister of Internal Affairs and Communications No.87, 2011) and "Standard transmission systemfor satellite general broadcasting" (Ordinance of the Minister of Internal Affairs and Communications No.94, 2011).

1.3 References

1.3.1 Normative references

(1) ISO/IEC 13818-7:2006 Information technology—Generic coding of moving pictures and associated audio information: Advanced Audio Coding (AAC)
(2) ISO/IEC 13818-7 2006/Cor.1:2009 Information technology—Generic coding of moving pictures and associated audio information -- Part 7: Advanced Audio Coding (AAC), TECHNICAL CORRIGENDUM 1 ((1) and (2) are hereinafter referred to as "MPEG-2 AAC Standard")
(3) ISO/IEC 14496-3:2009 Information technology -- Coding of audio-visual objects -- Part 3: Audio
(4) ISO/IEC 14496-3:2009/Cor.1:2009 Information technology -- Coding of audio-visual objects -- Part 3: Audio
(5) ISO/IEC 14496-3:2009/AMD 2:2010 Information technology -- coding of audio-visual objects -- Part 3: Audio
(6) ISO/IEC 14496-3:2009/Cor.2:2011 Information technology -- Coding of audio-visual objects -- Part 3: Audio
(7) ISO/IEC 14496-3:2009/AMD 4:2013 Information technology -- Coding of audio-visual objects -- Part 3: Audio
(8) ISO/IEC 13818-1:2013 | ITU-T Rec. H.222: Information technology --Generic coding of moving pictures and associated audio information: Systems (hereinafter referred to as "MPEG-2 Systems Standard")
(9) ISO/IEC 23008-1:2014 Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 1: MPEG media transport (MMT) (hereinafter referred to as "MMT Standard")

1.4 Terms

1.4.1 Abbreviations

AAC:	Advanced Audio Coding
ADTS:	Audio Data Transport Stream
ALS	Audio Lossless Coding
CCE:	Coupling Channel Element

CPE:	Channel Pair Element
CRC:	Cyclic Redundancy Check
ICS:	Individual Channel Stream
LATM	Low Overhead Audio Transport Multiplex
LC:	Low Complexity
LFE:	Low Frequency Effects
LOAS	Low Overhead Audio Stream
MMT	MPEG Media Transport
MPEG:	Moving Picture Experts Group
MPT	MMT package Table
MPU	Media Processing Unit
SCE:	Single Channel Element
SSR:	Scalable Sampling Rate
PCE:	Program Configuration Element
PTS:	Program Time Stamp

Chapter 2: Switching Audio Parameters

This provision applies to switches made to audio stream parameters within the same service ID to be transmitted from local station. More specifically, this provision applies the following parameters:

- Sampling frequency
- Bitrate
- Channel configuration
- Audio mode
- Audio coding system

The followings are taken into account with regard to this provision:

- A switch must be made to any of the audio parameters with at least 0.5 seconds of mute input to the audio encoder. The future potential reduction of mute time must be considered as well.
- The specifics of implementing the audio encoder are unspecified.
- The audio decoder shall have buffer capacity sufficient for the maximum number of channels to be handled by that decoder. Switch to any of the audio parameters must be made by controlling the entire buffer. However, note that control and monitoring of the buffer capacity stipulated by that parameter (e.g., overflow, underflow) is performed under a steady-state condition.
- Provision regarding receiver
  - The buffer may underflow.
  - A signal for mute is output if the buffer becomes empty. (If necessary, the audio level will start fading out immediately before the buffer becomes empty.)
  - After the buffer is empty, decoding will resume when the predetermined coded audio data is received.

2.1 Switching Audio Parameters in MPEG-2 AAC Standard

(1) Switching sampling frequency

When the sampling frequency is altered, the decoder will change its reference clock. Therefore, a transient and unstable condition occurs for a specific period of time. Since there is some question as to whether inserting 0.5 seconds of mute is sufficient, caution shall be exercised during operation.

(2) Switching bitrate

It is possible to ensure seamless changes in bitrate by appropriately controlling the buffer at the encoder side. If it is possible that the buffer may not be properly controlled, due (for example) to change in coding delay caused by switch to bitrate, it is necessary to abide by the rules indicated in the next section, "(3) Switching to other parameters."

(3) Switching to other parameters

(a) The encoder waits until there is no more stream data stored in the encoder and decoder buffers. Then, the encoder changes the target audio parameter and resumes encoding. After encoding resumes, the preset amount of coded audio stream data is stored in the encoder buffer. Finally, audio stream data is sent to the decoder.

Since stream data is transmitted using MPEG-2 Systems, PTS shall be added to the first frame of stream data encoded after any interruption. Also note that to ensure that the decoder can find that a change has been made to a parameter, there shall be a gap of at least three frames between the PTS of the stream (stream introduced on the assumption that it occurs after the stream with the previous parameter) and the PTS added to the stream that is actually transmitted.

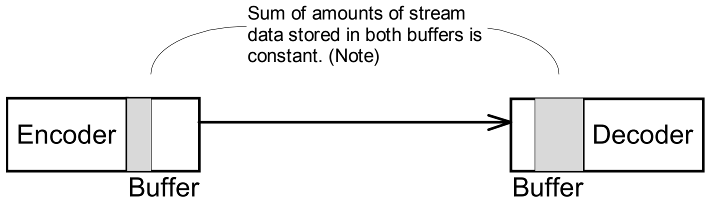
(b) The decoder halts decoding and mutes the audio when no audio stream data is found in the decoder input buffer. If audio stream data remains in the decoder input buffer and if the ADTS (Audio_Data_Transport_Stream) frame header is found, the decoder waits until the amount of stream data specified by the adts_buffer_fullness field is stored in the input buffer and resumes decoding based on the new audio parameter information.

The decoder cancels audio muting and outputs decoded audio signals when this signal is requested (at any time after completion of decoding of two frames because overlapping occurs.)

However, note that streams generated by the above model are in practice transmitted through the MPEG-2 Systems, and that the decoder performs buffer control using system buffer and PTS. In this case, the decoder may not always be able to find that the decoder buffer is empty, despite the assumption made above at the elementary stream level. Under such circumstances, the decoder can determine that a change has been made to a parameter by finding that streams are not in succession based on the PTS added to the first audio frame after parameter change and also based on system clock information.

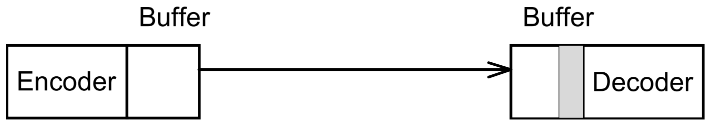
To facilitate the comprehension of audio parameter change sequence, Figs. 2-1 and 2-2, respectively, show the flow diagram and the timing diagram for switching audio parameters.

1. Steady-state condition

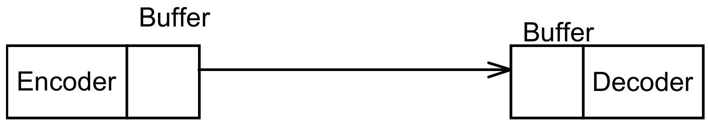


(Note) The sum of stream data stored in buffers is set at 6144 bits/channel or less.

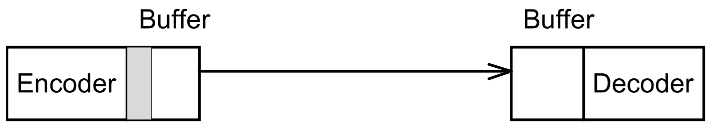
2. The encoder stops encoding.
   -> The amount of stream data stored in the encoder and decoder buffers decreases.



3. The encoder and decoder buffers become empty.



4. The encoder resumes encoding using new parameter information. It sends audio stream data to the decoder when the predetermined amount of stream data is stored in the encoder buffer. (During this period, the decoder does not perform decoding and mutes the audio.)



5. The amount of stream data stored in the encoder and decoder buffers reaches a constant level.
   -> The decoder resumes decoding based on new parameter information.
   -> The decoder cancels audio muting and outputs decoded audio signals at the time specified by PTS.

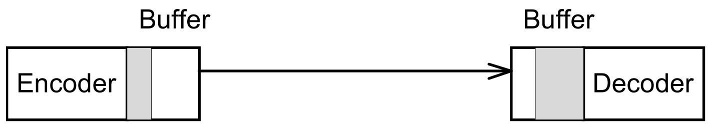


Fig. 2-1: Flow diagram for switching audio parameters

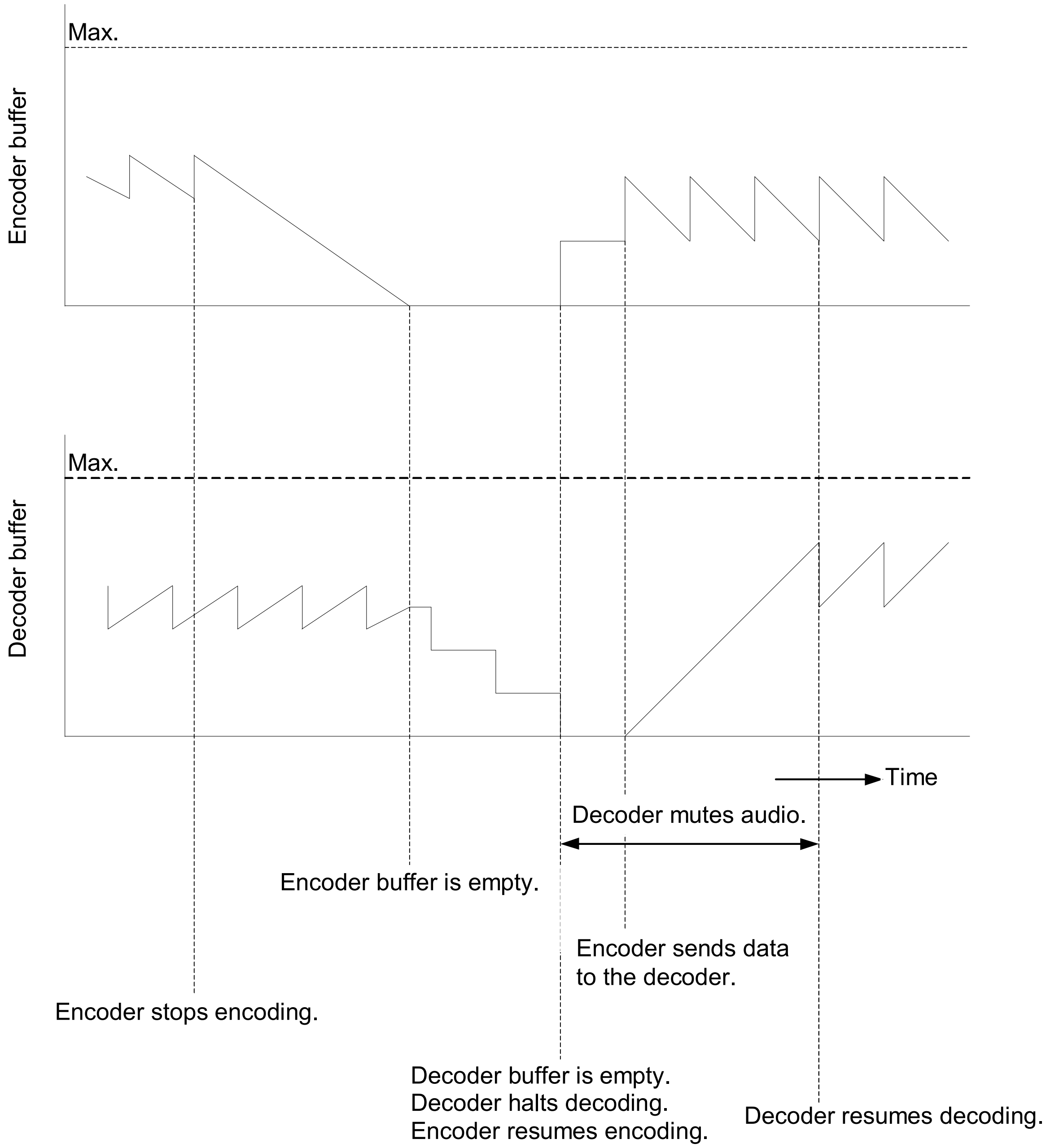


Fig. 2-2: Timing diagram for changing audio parameters

2.2 Swtching Audio Parameters in MPEG-4 AAC Standard

As for switching audio parameters in MPEG-4 AAC standard, it is basically carried out according to switching audio parameters in MPEG-4 AAC standard recorded in 2.1. As for switching sampling frequency and bitrate, 2.1 (1) and (2) shall be referred to, and change of the other parameters are provided as the following.

(1) Switching parameters other than sampling frequency and bitrate

(a) The encoder waits until the amount of stream data which is stored in encoder buffer and decoder buffer becomes zero, then it changes audio parameter and resumes encode processing. After resume, it stores coded audio stream with the amount of data which is set in encoder buffer, then it transmits audio stream to the decoder.

In case of transmission by using MMT, presentation time (PTS) of first audio frame of the resumed stream should be described by MPU time stamp descriptor which is arranged in asset descriptor region of MPT and MPU extended time stamp descriptor. (About calculation method of presentation time, refer to ARIB STD-B60, Description 2. As for audio stream, audio frame corresponds to access unit.) Also, as the decoder certainly recognizes switch of parameters, the gap between the presentation time of following stream which is introduced on the assumption of the continuity from the stream with previous parameters and the presentation time which is put to the stream transmitted actually should be more than three frames.

(b) The decoder stops decoding process when the amount of audio stream in decoder input buffer becomes zero, and processes muting. When there is audio stream in the decoder input buffer and LATM/LOAS frame header is found, it waits until the amount of stream represented in latmBufferFullness field is stored in input buffer, then resumes decoding process according to new audio parameter information. (As the process is overlapped, after processing 2 frames decoding) at the time decoded audio signal is required, it stops muting and outputs.

Though, in case of transmission by using MMT, the stream generated by above-mentioned model is controlled for receiving buffer by using system buffer and presentation time in the decoder. (Refer to ARIB STD-B60, Description 2.) In this case, the decoder cannot always detect emptiness of the decoder buffer which is supposed by the level of audio stream mentioned-above. In such a case, the decoder can recognize change of parameters by detecting discontinuity of stream according to information about presentation time which is put to first audio frame after change of parameter, and system clock.

2.3 Switching Audio Parameters in MPEG-4 ALS Standard

Switching audio parameters in MPEG-4 ALS Standard is basically performed by analyzing AudioSpecificConfig() and ALSSpecificConfig() which were transmitted by LATM/LOAS or MMT. Since every frame is random-accessable as defined in 7.2.2, MPEG-4 ALS standard allows to switch audio parameters in every frame in an extreme case. But as sufficient muting time (around 0.5 second) is necessary for switching on the reason mentioned below, the operation in actual time is desired. Here, as bitrate depends on input audio signal, it cannot be controlled.

(1) Sampling frequency

As the same case of MPEG-2 AAC/MPEG-4 AAC, when switching sampling frequency, as the decoder changes reference clock, a transient unstable state occurs for a certain period. It is not clear whether inserting silence for 0.5 second is sufficient or not in this case under the present conditions, so the operation must be paid attention to.

(2) Channel configuration

The number of speakers which can be output varies with implementation of the receiver. Audio input signal which is sufficiently processed for fade-in and fade-out is expected for derating the burdens of viewers.

(3) Number of input quantized bits

As there is some possibility that DA converter varies because of implementation of the receiver, the time for changing DA convertor and additional process must be considered.

(4) Frame length

Though frame length can be selected within a fixed range in MPEG-4 ALS, it is desirable to fix the value which frame length of AAC divided by integer, in order to synchronize MPEG-2 AAC/MPEG-4 AAC. Also, as the size of payload which can be described to audioMuxLengthBytes of LATM/LOAS is limited to less than 8192, in case of not dividing frame_data(), it is desirable to set 1024 samples for Level 1 and Level 2, 512 samples for Level 3, 256 samples for Level 4 in Simple Profile.

In case of transmission by using MMT, presentation time (PTS) of first audio frame in resumed stream should be described by MPU time stamp descriptor which is arranged in the region of asset descriptor of MPT, and MPU extension time stamp descriptor. (About calculation method of presentation time, refer to ARIB STD-B60, Description 2. About audio stream, audio frame corresponds to access unit.) Also, as the decoder certainly identifies the change of parameters, the difference between the presentation time of successive stream which is introduced on the assumption of the continuity with stream of former parameters and the presentation time which is added to actually transmitted stream should be more than three frames.

For audio stream transmitted by MMT, the buffer for receiving is controlled by using system buffer and presentation time in the decoder. (Refer to ARIB STD-B60 Description 2.) The decoder can recognize the change of parameters by detecting discontinuity of stream, by using presentation time which is added to first audio frame after the change of parameters and information of system clock.

2.4 Switching between Audio Coding System standard

In this section, the case is handled that audio coding system for audio stream in the same service ID is changed.

The case accompanying the change of audio coding system is that multiple audio coding systems are able to be applied to one standard system, and the use of audio stream by different audio coding system in the same service ID is supposed. For example, switching of audio stream between MPEG-4 AAC Standard and MPEG-4 ALS Standard for advanced wide band digital satellite broadcasting corresponds to it. This case involves the case that multiple simulcast audio stream by different audio coding system in the same service ID is transmitted, and audio stream by different audio coding system is automatically changed in the receiver according to the change of stream configuration, etc.

(1) Basic consideration

- Preceding change of audio coding system, the change of audio parameters in multiplex level must be able to be discriminated. For example, it is realized by updating contents of asset descriptor which is described in MPT about 0.5 second ahead.
- Before and after change of audio coding system, all related audio stream must follow the guideline of switching audio parameters for each audio coding system.

(2) Precautionary matters

- When multiple simulcast audio streams are transmitted, there is a case that audio stream to be decoded and reproduced is different by the audio decoding function of the receiver. Even in such a case, in order to discriminate certainly the change of audio coding system in the receiver, it must follow the guideline of switching audio parameters for each audio coding system. For example, even in the case that for a certain audio stream, audio parameters are not changed alone before and after concerned change, all audio stream involving the stream must follow concerned guideline.

<Blank Page>

Chapter 3: Audio Quality Indication

A quality indicator (“quality_indicator”) is assigned as an audio component descriptor for multiplexing systems. It can be used to transmit and indicate audio quality signals. The quality indicator shall be transmitted in accordance with the quality of an AAC coding stream that meets the audio quality criteria. Two bits are assigned to the quality indicator so that audio quality can be classified into up to four different groups.

With BS digital broadcasting, it is assumed that two types of audio quality will be available: one equivalent to B mode in BS analog television broadcasting (Note); and other type such as A mode. Conversely, terrestrial digital television and terrestrial digital audio broadcasting requires the use of sampling frequencies below 32 kHz. Therefore, three audio quality indications are assumed to be available.

Table 1 lists the quality indicator assignments and correspondence between the content of quality indicators and coded audio quality.

Mode 1 represents the high audio quality equivalent to B mode in BS analog television broadcasting. Audio quality by MPEG-4 ALS lossless audio coding system is higher than the quality equivalent to B mode in BS analog television broadcasting, and placed to mode 1. Audio coding system by MPEG-2 AAC and MPEG-4 AAC is the irreversible audio coding system, and the audio quality changes by bitrate, etc. The expected bitrate for 2-channel stereo transmission (to be used for the time being), is shown for reference purposes in the table, based on the results of tests including a subjective assessment test. Also note that mode 2 represents standard audio quality that is not classified as mode 1. The bitrate available when applying the audio quality criteria in the ITU-R standard is shown for reference purposes.

These reference values have been introduced based on the tests using a properly adjusted encoder that can handle off-line processing. It will be necessary to check these results with a practical real-time encoder for broadcasting. Conversely, it is expected that advances in encoder technology that occur at the start of broadcasting and later will ensure that specified audio quality criteria are met at lower bitrates than the reference values.

In the meantime, the main purpose of mode 3 is to inform viewers that this mode offers limited audio quality compared to modes 1 and 2. Therefore, no quantitative guidelines are established. Instead, it is assumed that broadcasters will choose whether to offer this mode, based on agreed upon rules. Note that mode 3 will not be used for BS or wide band CS digital broadcasting for the reasons mentioned above.

(Note) Standard television broadcasting (except digital broadcasting) which is specified in the Ordinance of the Ministry of Internal Affairs and Communications, No. 88, 2011 “Standard transmission system (except digital broadcasting) on standard television broadcasting,” Chapter 3, and is operated by the key satellite broadcasting stations using radio wave whose frequency is from 11.7 GHz to 12.2GHz.
(This was repealed by the Ordinance of the Ministry of Internal Affairs and Communications, No.7, 2013.)

Table 1: Quality indicator assignments and coded audio qualities

Quality indicator	Audio quality name (Note 1)	Coded audio quality criterion	Remarks
00	Reserve		
01	Mode 1	Audio quality equivalent to B mode available in BS analog television broadcasting	Reference bitrate by ALS lossless audio coding system and AAC audio coding system: 192 to 256 kbps/stereo or more (Note 2)
10	Mode 2	Audio quality (Note 3) other than mode 1 that is not classified in mode 3	
11	Mode 3	Mode with limited audio quality compared to modes 1 and 2	It is assumed that broadcasters will choose whether to offer this mode at their own discretion based on agreed upon rules. (ex. Sampling frequency below 32 kHz) This mode is not used in BS or wide band CS digital broadcasting.

(Note 1) In this table, audio quality is referred to as mode 1, 2, or 3 audio quality for the sake of convenience. Note that audio quality may be referred to as something else when actual services are provided.

(Note 2) Mode 1 audio quality will be offered at 192 to 256 kbps/stereo or more for the following reasons:

- Following subjective assessment tests conducted by ARIB in June 1998, we can say that the following holds true:
  The higher the bitrate, the better the audio quality. The audio quality available with 192 kbps/stereo is barely distinguishable from the original sound.
- It is appropriate to examine the possible application of a bitrate that is approximately 1.5 times that used for broadcasting, given the relationship of codec bitrate for broadcasting and material transmission (MPEG-1 layer 2 coding) in the Rec. ITU-R.

(Note 3) The reference bitrate is as shown below when applying the following audio quality criterion:

Audio quality criterion	Reference bitrate
Audio quality for broadcasting given in ITU-R	144 kbps/stereo or more

Description: Considerations of Developing Operating Conditions

Description 1: Considerations in Developing Operational Conditions for MPEG-2 AAC Standard

The following lists the topics considered before restrictions in relation to audio coding based on MPEG-2 AAC Standard were established:

(1) Input audio format

There is a provision for audio mode (channel mode) of the system based on MPEG-2 BC Standard in Notification of the Ministry of Internal Affairs and Communications (No. 300, 2011) regarding audio modes. However, no provisions are provided for audio mode of the system based on MPEG-2 AAC Standard. This is because the MPEG-2 AAC Standard provides no provisions that clearly specify audio modes. However, considering the continuity of operation for MPEG-2 AAC Standard and MPEG-2 BC Standard, and following a review of expected services, we have listed as possible modes the audio modes shown in Section 5.1. Note that compliance is required with the provision given in Section 5.2.3 regarding the relationship between audio modes and coding modes.

However, note that it seems advantageous in terms of operations (broadcasters), cost (receivers), and services (viewers) to trim this list of audio modes to some extent after examining real-world needs. After having reviewed needs for services for the time being, we have decided to define the audio modes shown in Section 5.1 as recommended modes.

(2) Main parameters

The ADTS format — a format with a header in each frame — has been adopted as the bitstream format, since it will be used for broadcasting purposes. Restrictions on ADTS header will be given later.

The LC profile was initially adopted for use with BS/wide band CS digital broadcasting based on the following factors:

(a) As a result of the AAC audio quality assessment test conducted by ARIB in June 1998, we found that the LC and SSR profiles met the ITU-R broadcasting quality criteria or the criteria required by BS/wide band CS digital broadcasting at 144 kbps/2 channels or more.
(b) It was pointed out that SSR profile-specific features were not effective for BS/wide band CS digital broadcasting.
(c) It was pointed out that the LC profile could improve audio quality as a result of optimization and technical advance of encoders beyond year 2000 when BS digital broadcasting would begin.
(d) Based on the premise that BS digital broadcasting shall begin in 2000, it was pointed out that it would be possible to develop encoders and receivers for the LC profile, but would be difficult to do so for the MAIN profile.
(e) There is a significant difference in chip costs between MAIN and LC profiles.
(f) There are technical problems to be solved for MAIN profile.

We have decided to adopt the LC profile for digital terrestrial television broadcasting and digital terrestrial audio broadcasting as well for the above reasons and in view of consistency with BS/wide band CS digital broadcasting.

No restrictions have been introduced especially for digital broadcasting in relation to the maximum bitrate. In terms of the standard, the maximum bitrate for AAC format is 288 kbps/channel when the sampling frequency is 48 kHz.

(3) Restrictions on AAC ADTS coding parameters

To improve the error tolerance of ADTS, Cyclic Redundancy Check (CRC) data or adts_error_check must be added after ADTS header. This requires that protection_absent be 0.

The CRC processing procedure is defined in the AAC Standard. For clarification, this procedure is shown in Appendix 1.

As for Sampling_frequency_index, so-called low sampling frequencies — 24, 22.05, and 16 kHz — have been introduced in addition to three frequencies defined in Chapter 2. (However, note that only the three frequencies given in Chapter 2 are used for BS/wide band CS digital broadcasting.)

The need for partial reception in terrestrial digital television and for audio transmission at low bitrates in terrestrial digital audio broadcasting was pointed out, involving transmission line restrictions. For this reason, a study entitled "Audio quality assessment test at low bitrates coding by MPEG-2 AAC" was carried out by ARIB in March 1999. As a result of the test, it has been suggested that audio services are feasible at bitrates lower than 144 kbps/2 channels (LC profile) and sampling rate lower than 32 kHz. A test was also conducted by MPEG for the same purpose. Low sampling frequencies have been added given the findings from these tests.

To maintain the average bitrate fixed, 0x7FF (indicating variable bitrate) is prohibited for use as adts_buffer_fullness value.

With ADTS format, a single header can control up to four pieces of raw_data_block(). However, one would encounter the following problems when attempting to control many pieces of raw_data_block() by a single header: (1) seriously adverse impact due to header loss, (2) seriously adverse impact in the event of even a single error because the number of pieces of raw_data_block() controlled by a single header also represents the number of CRCs.
For these reasons, only a single piece of raw_data_block() shall be controlled by a single header.

(4) Audio stream configuration and multiplexing

It is necessary to clearly define the correspondence between input modes and coding modes in relation to audio modes. For this reason, we have decided, based on the AAC Standard, to establish some provisions specific to digital broadcasting regarding ADTS configuration.

Determination as to whether to use a single or multiple ADTSs for different input audio modes was made based on potential need for simultaneous reproduction. With 2-audio transmission (for example), dual mono mode with one ADTS is used when simultaneous reproduction is requested. However, dual mono mode with two ADTSs can be used when simultaneous reproduction is not requested.

Program Configuration Element (PCE) shall be used only to transmit channel configuration and downmix coefficients. It is also necessary to ensure that the PCE is consistent with ADTS header. When the channel_configuration bit in ADTS header is 0, it is possible to accurately represent the intended state of reproduction by decoding PCE.

(5) Compatibility between multi-channel stereo and 2-channel stereo

There is a strong likelihood that not only digital terrestrial television broadcasting, BS/wide band CS digital broadcasting and advanced BS/wide band CS digital broadcasting receivers capable of reproducing multi-channel stereo, but also even those receivers capable of reproducing two-channel stereo will be commercially available. Thus, full compatibility with two-channel stereo-capable receiver shall be accounted for when multi-channel stereo service is provided.

There are two basic possible approaches to ensuring compatibility between multi-channel stereo and two-channel stereo: (1) multi-channel stereo/two-channel stereo simulcasting and (2) downmixing from multi-channel stereo to two-channel stereo at the receiver. We decided to adopt approach (2), because with conventional digital broadcasting services up to 5.1 channels, the transmitting side need only transmit a single stream, ensuring improved efficiency in bitrate, although this places a slightly greater burden on the receiver (decoder). Note that we have decided that simulcasting may also be implemented if requested by program producers. Here, for advanced BS/wide band CS digital broadcasting, as audio mode is 24 channels (3/3/3-5/2/3-3/0/0.2) in maximum, it is supposed that the burden of the receiver will be relatively heavy. So it is desired that the operation of 2-channel stereo simulcasting would be investigated considering technical trend at the beginning of service.

The AAC standard stipulates that PCE can transmit downmix coefficients only for five channels (3/2). Therefore, we have decided to adopt this approach as is. On the other hand, for advanced BS/advanced wide band CS digital broadcasting, downmixing of multichannel audio with more than 5.1 channels needs to be taken into account. As a result of investigation, about three dimensional multichannel audio (audio mode which assigned channels exist in upper layer and lower layer, among multichannel audio with more than 5.1 channels), transmission procedure of downmix coefficients using DSE has been provided.

In case of downmixing multichannel with more than 5.1 channels to two channels, the process such as downmixing to 2-channel stereo after downmixing multichannel to 5.1 channels may cause to vary front standpoint of audio. So such a processing should be avoided.

Description 2: Notation for Audio Mode

When providing multi-channel sound more than 5.1 channel, the notation for sound mode has been investigated which cannot represent speaker system of 3 dimensional arrangement, considering the continuity of conventional notation. This description explains the notation for sound mode used in this standard.

(1) Notation for audio mode

upper layer (front/side/rear) - middle layer (front/side/rear) - lower layer (front/side/rear).LFE

Transcribing is that allocated the number of channels for front, side, rear in each layer is connected by " / ", and each layer is connected by " - ". If there is an allocated channel for LFE, it is transcribed as ".number of LFE" at the end. But no allocated channel is transcribed as 0.

Example: 2/0/0-2/0/2-0.1 = 2 upper front＋2 middle front＋2 middle rear＋1 LFE

A) In case of audio mode comprising only middle layer, notation is as the following for simplification.

(i) In case that there are no allocated channel for upper layer and lower layer;

This is transcribed as middle layer (front/side/rear).LFE.

Example: 3/2/2.1 = 3 middle front＋2 middle side＋2 middle rear＋1 LFE

(ii) Besides the case that there is only allocated channel for middle layer, also in case that there are no allocation for side channel;

This is transcribed as middle layer (front/rear).LFE.

Example: 3/1 ＝ 3 middle front＋1 middle rear

3/2.1 ＝ 3 middle front＋2 middle rear＋1 LFE

B) Notation for LFE

In case that the number of allocated channels to LFE is one, there is a case of transcribing "＋LFE".

Example: 3/2.1 ＝ 3/2+LFE ＝ 3 middle front＋2 middle rear＋1 LFE

(2) Notation for audio mode which is used in a sentence

There is a case of transcribing:
number of all channels. number of LFE

Example: 5/0/2.1 = 7.1 or 7.1 ch

Example: 3/2+LFE ＝ 3/0/2.1 = 5.1 or 5.1 ch

Description 3: Overview of MPEG-4 ALS System

MPEG-4 ALS (Audio Lossless Coding) system is a high efficiency audio lossless coding system which is provided in MPEG-4 Audio (ISO/IEC 14496-3).

<Overview of technical system>

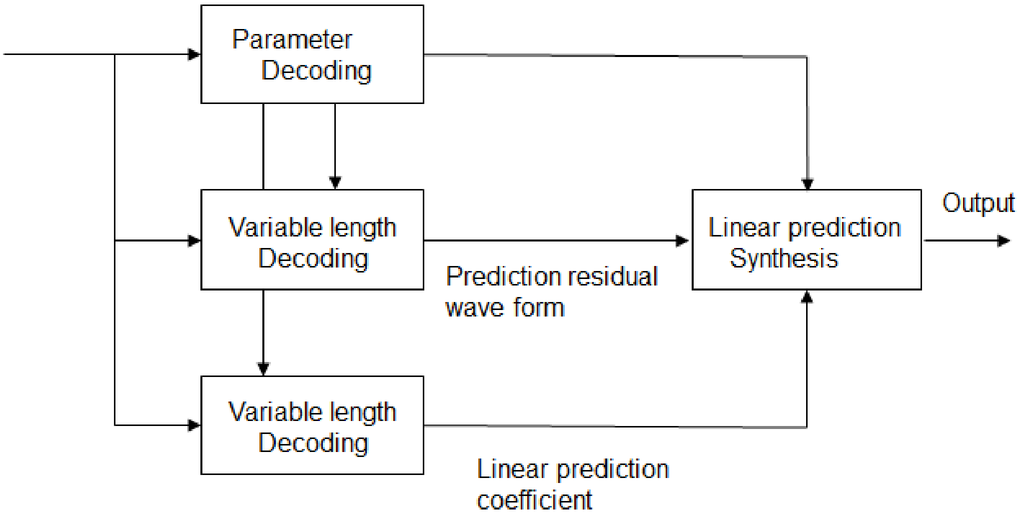


Fig. 1 Configuration of MPEG-4 ALS decoder

- Based on compression coding system MPEG-4 ALS (Audio Lossless Coding) which does not cause distortion.
- Theoretically assured perfectly reconstructing input wave form at decoding, by deterministic integer operation which is perfectly consistent between encoder and decoder.
- Information compression by deleting redundancy between adjacent samples by linear prediction.
- Information compression by reducing redundancy of bias in amplitude distribution by coding amplitude of predicted error wave form into variable length code.

<An example of application to broadcasting system>

- MPEG-4 ALS stream is transmitted as MPEG-2 TS after being made to PES as a unit of frame.
- MPEG-4 ALS stream is identified by MPEG-4_audio_extension_descriptor which is allocated in PMT in MPEG-2 TS.

<Level of ALS Simple Profile>

- Level is defined as the following for Simple Profile of MPEG-4 ALS. The appropriate level is desired to be selected according to the restriction of input audio signal which is transmitted and reproduced.

Level	Max. number of channels	Max. sampling rate [kHz]	Max. word length [bit]	Max. number of samples per frame	Max. prediction order	Max. BS* stages	Max. MCC** stages
1	2	48	16	4096	15	3	1
2	2	48	24	4096	15	3	1
3	6	48	16	4096	15	3	1
4	6	48	24	4096	15	3	1

- "audioProfileLevelIndication values" corresponding to the level mentioned-above is as the following.

0x3C		ALS Simple Profile	L1
…		…	…
0x5A		ALS Simple Profile	L2
0x5B		ALS Simple Profile	L3
0x5C		ALS Simple Profile	L4

Appendix

Appendix 1: CRC (Cyclic Redundancy Check) Processing Procedures for MPEG-2 AAC ADTS (Audio Data Transport Stream)

This appendix is intended to clarify the CRC processing procedure in the MPEG-2 AAC Standard (ISO/IEC 13818-7) ADTS. That the description given in this appendix does not pose any problems has been confirmed in the MPEG Beijing Conference (July, 2000), and is spelled out in Section 2.5.9 of the Resolution of the Conference.

The MPEG-2 AAC Standard includes the following as CRC processing procedure:

adts_error_check() CRC error detection data generated as described in ISO/IEC 11172-3, subclause 2.4.3.1 (table 1.7)
The following bits are protected and fed into the CRC algorithm in order of appearance:
all bits of the headers
first 192 bits of any
single_channel_element (SCE)
channel_pair_element (CPE)
coupling_channel_element (CCE)
low frequency enhancement channel (LFE)
In addition, the first 128 bits of the second individual_channel_stream in the channel_pair_element shall be protected. All information in any program configuration element or data element shall be protected.
For any element where the specified protection length of 128 or 192 bits exceeds its actual length, the element is zero padded to the specified protection length for CRC calculation.

An example of cases in which description is difficult is zero padding when the CPE length is less than 192 bits and when the second ICS of the same CPE is less than 128 bits.

We have reached the following conclusions:

- First, the CPE is processed from the beginning. If the CPE is less than 192 bits in length, it will be zero padded to 192 bits.
- Next, the second ICS of the same CPE is processed from the beginning. If this ICS is less than 128 bits in length, it will be zero padded to 128 bits.

That is, the total number of 0s padded is as follows:

(192 - CPE length) + (128 - 2nd ICS length) bits

```
  ---              +-------+  <-- start point of CPE
   |               |       |
   |               |       |
   |      ---      +-------+  <-- start point of 2nd individual_channel_stream
   |       |       |       |
   |       |       |       |
   |       |       |       |
   |       |       +-------+  <-- end point of this CPE
192 bit    |       |   0   |
   |       |       |   0   |
   |   128 bit     |   0   |
   |       |       |   0   |
   |       |       |   0   |
   |       |       |   0   |
   |      ---      |   0   |
  ---              |   0   |
                   |       |
```

Fig. 1: Example in which CPE length < 192 bits and 2nd ICS length < 128 bits

Appendix 2: Overview of ISO/IEC 13818-1 AMD 6 (Related to AAC System Buffer)

The MPEG-2 Systems (ISO/IEC 13818-1) defines the system buffer size needed to decode audio and video data. This buffer size corresponds to the audio/video coding system. However, at the MPEG Dublin Conference held in July 1998, it was pointed out that the provision regarding AAC system buffer was unclear. The MPEG-2 Systems AMD 6 (Amendment 6 of the International Standard) was published in response.

(Note: At present, AMD 6 is an integral part of the MPEG-2 Systems Standard as ISO/IEC 13818-1: 2000. For the sake of convenience, it is called AMD 6 in this appendix.)

The MPEG-2 Systems AMD 6 defines four AAC system buffer sizes and leak rates for varying numbers of channels (up to 2, 8, 12 and 48 channels). The specific buffer sizes and leak rates are given below:

Leak rate (Rxn):	Number of channels	Rxn [bps]
	1 – 2	2,000,000
	3 – 8	5,529,600
	9 – 12	8,294,400
	13 – 48	33,177,600
Buffer size (BSn):	Number of channels	BSn [bytes]
	1 – 2	3,584
	3 – 8	8,976
	9 – 12	12,804
	9 – 48	51,216

Channels: Channels that require their own decoder buffer in this elementary stream n.

With terrestrial and BS/wide band CS digital broadcasting systems in Japan, 5.1 channels are defined as the maximum number of AAC audio stream channels. According to the AMD 6 provision, the system layer must have the buffer size appropriate for the number of channels (3 to 8 channels). More specifically, the total buffer size shall be 9,488 bytes (BSn 8,976 bytes + transport buffer 512 bytes).

Appendix 3: Precautions associated with revision to ISO/IEC 13818-7:2003

The first version of the MPEG-2 AAC standard (ISO/IEC 13818-7) was issued in 1997. Initially, the ISO/IEC 13818-7:1997 was supposed to be referenced when developing audio coding systems in digital terrestrial sound broadcasting, digital terrestrial television broadcasting, BS digital broadcasting, and wide band CS digital broadcasting.

Later in the revised version of ARIB STD-B32 Ver. 1.6 issued in May 2004, ISO/IEC 13818-7:2003 was referenced entirely. The following precautions should be taken as a result. (Note: At present, the reference standard is revised to ISO/IEC 13818-7:2006, but this item is still valid.)

(1) Differences between ISO/IEC 13818-7:1997 and ISO/IEC 13818-7:2003 (excerption)

(a) Meaning of adts_buffer_fullness

No specification is given in ISO/IEC 13818-7:1997, but ISO/IEC 13818-7:2003 stipulates that adts_buffer_fullness is the "amount of remained equivalent buffer per channel (6,144 bits per channel)."

(b) Meaning of "Minimum Decoder Input Buffer"

No specification is given in ISO/IEC 13818-7:1997, but ISO/IEC 13818-7:2003 stipulates that there is no LFE component in the decoder buffer.

(2) Precautions

Some transmission devices manufactured and used on the basis of any ARIB STD-B32 version before Ver. 1.5 may not comply with the explanation given in ISO/IEC 13818-7:2003. Some of such devices are difficult to modify to make them compatible with ISO/IEC 13818-7:2003. The present standard is thus applied as follows in view of the revisions made to the referenced MPEG-2 AAC standard.

- The present standard does not apply to devices manufactured and used on the basis of any ARIB STD-B32 version before Ver. 1.5.
- Devices manufactured and used on the basis of ARIB STD-B32 version after Ver. 1.6 shall comply with the the explanation given in the referenced international standard ISO/IEC 13818-7:2003.

When designing digital broadcasting receivers (defined by ARIB STD-B21) that comply with the present standard, due consideration must be given to the presence of streams based on different explanations associated with the revisions made to the MPEG-2 AAC standard.

Appendix 4: Precautions associated with implementation of MPEG-2 AAC Standard

This appendix explains the characteristics of coding tools used in the MPEG-2 AAC standard and presents precautions associated with the implementation of the standard.

(1) Treatment of AAC coding tools

The MPEG-2 AAC standard specifies three coding tools that can be used in the AAC LC profile: M/S Stereo, Intensity stereo, and TNS. (Note) Prediction and Gain Control, which are beyond the usable range in the LC profile, cannot be used and the use of Coupling Channel is prohibited in Section 5.2.2 of Part 2 of this standard.

The AAC standard stipulates that these coding tools shall be treated in the decoder according to bit streams. Namely, it is specified that decoding shall be carried out according to the bit stream no matter which coding tool is being used. This specification must be taken into account when designing and implementing decoders.

(2) TNS

When designing and implementing decoders, care must be taken on TNS, which may require a large number of steps for decoding. Although the highest TNS filter order in long window mode is restricted to 12 in the LC profile, compared with 20 in the Main profile, it is still possible that a large number of processing steps will be required. The factors that directly affect the number of steps for TNS decoding include (1) the number of filters, (2) filter order, (3) filter length, and (4) the number of channels.

In the case of a decoder DSP that performs fixed-point calculation, the required precision of operation may not be ensured when a high scaling level is adopted to avoid overflow or underflow caused by filtering (i.e., TNS may not be effective in improving sound quality). When designing and implementing decoder systems, therefore, care must be taken on overflow and underflow caused by filtering.

(3) Huffman decoding

The processing load of Huffman decoding tends to increase with bit rate. When designing and implementing decoders, theoretical maximum instantaneous rate must be taken into account in reference to the buffer model of the AAC standard. In the case of encoders, care must be taken to avoid excessively high maximum instantaneous rate when bit rate is relatively high.

<Blank Page>

Part 3: Signal Multiplexing System

<Blank Page>

Part 3: Signal Multiplexing System

Contents

Chapter 1: General Terms

1.1 Objective

The purpose of this standard is to define the signal multiplexing system for digital broadcasting.

1.2 Scope

This standard applies to “Standard transmission system for digital broadcasting among standard television broadcasting and the like” (Ordinance No.87 of the Ministry of Internal Affairs and Communications, 2011) and “Standard transmission system for satellite general broadcasting (Ordinance No.94 of the Ministry of Internal Affairs and Communications, 2011), excluding multimedia broadcasting beased on selective band transmission system among multimedia broadcastings..

1.3 References

1.3.1 Normative documents

This standard contains excerpts from the matters provided for in the following documents:

(1) Ordinance No.87 of the Ministry of Internal Affairs and Communications, 2011 (Partial revision: December 10, 2013, July 3, 2014) “Standard transmission system for digital broadcasting among standard TV broadcasting and the like” (hereinafter referred to as “Ordinance”)

(2) Ordinance No.94 of the Ministry of Internal Affairs and Communications, 2011 (Partial revision: December 10, 2013, July 3, 2014) “Standard transmission system for satellite general broadcasting (hereinafter referred to as “Ordinance No.94”).

(3) Notification No.233 of the Ministry of Internal Affairs and Communications, 2014 “Defining configuration and transmission procedure of the conditional access related information , transmission procedure for PES packet, Section, TS packet, IP packet, ULE packet, MMTP packet, compressed IP packet and TLV packet,, configuration of transmission control signal and identifier, and configuration of emergency information identifier and emergency alarm broadcasting message” (hereinafter referred to as “Notification”)

(4) Rec. ITU-T H.222.0 (2012) | ISO/IEC 13818-1:2013: Information technology – Generic coding of moving pictures and assoiated audio information: Systems (Including Amendment 1 to 4)

(5) ISO/IEC 23008-1: 2014: Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 1: MPEG media transport (MMT)

(6) IETF RFC 768: User Datagram Protocol, Aug. 1980

(7) IETF RFC 791: Internet Protocol, Sep. 1981

(8) IETF RFC 2460: Internet Protocol, Version 6 (IPv6) Specification, Dec. 1998

(9) IETF RFC 3095: RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed, Jul. 2001.

(10) IETF RFC 4326: Unidirectional Lightweight Encapsulation (ULE) for Transmission of IP Datagrams over an MPEG-2 Transport Stream (TS), Dec. 2005.

(11) IETF RFC 4815: RObust Header Compression (ROHC): Corrections and Clarifications to RFC 3095, Feb. 2007.

(12) IETF RFC 5905: Network Time Protocol Version 4: Protocol and Algorithms Specification, Jun. 2010.

1.4 Terminology

1.4.1 Definitions

(1) Digital terrestrial sound broadcasting:
Digital broadcasting among various types of ultra-short wave broadcasting carried out by terrestrial basic broadcast stations defined in Chapter 2, Ordinance.

(2) Digital terrestrial television broadcasting:
Digital broadcasting and high-definition television broadcasting among various types of standard television broadcasting carried out by terrestrial basic broadcast stations defined in Chapter 3, Ordinance.

(3) Multimedia broadcasting:
Television broadcasting and multimedia broadcasting carried out by terrestrial broadcast stations defined in Chapter 4, Ordinance.
- V-Low multimedia broadcasting based on connected segment system:
  Multimedia broadcasting defined in Section 1 of Chapter 4, Ordinance
- V-High multimedia broadcasting based on connected segment system:
  Television broadcasting and multimedia broadcasting defined in Section 2 of Chapter 4, Ordinance
- Multimedia broadcasting based on selective band transmission system:
  Television broadcasting and multimedia broadcasting defined in Section 3 of Chapter 4, Ordinance

(4) BS digital broadcasting:
Digital broadcasting among various types of standard television broadcasting, high-definition television broadcasting, ultra-short wave broadcasting and data broadcasting carried out by the wide band transmission system using satellite basic broadcast stations (including satellite basic broadcasting experimental test stations and development test stations for satellite basic broadcasting) that employ radio waves of frequencies higher than 11.7 GHz and equal to or lower than 12.2 GHz as defined in Section 2 of Chapter 5, Ordinance.

(5) Advanced BS digital broadcasting:
Digital broadcasting among various types of standard television broadcasting, high-definition television broadcasting, ultra high-definition television broadcasting, ultra-short wave broadcasting and data broadcasting carried out by the advanced wide band transmission system using satellite basic broadcast stations (including satellite basic broadcasting experimental test stations and development test stations for satellite basic broadcasting) that employ radio waves of frequencies higher than 11.7 GHz and equal to or lower than 12.2 GHz as defined in Section 3 of Chapter 5, Ordinance.

(6) Narrowband CS digital broadcasting:
Standard television broadcasting, high-definition television broadcasting, ultra-short wave broadcasting and data broadcasting carried out as a satellite general broadcasting by the narrowband transmission system using satellite broadcast stations that employ radio waves of frequencies higher than 12.2 GHz and equal to or lower than 12.75GHz as defined in Paragraph 1 of Article 3, Ordinance No.94.

(7) Wide band CS digital broadcasting:

Standard television broadcasting, high-definition television broadcasting, ultra-short wave broadcasting and data broadcasting carried out by the wide band transmission system using satellite basic broadcast stations that employ radio waves of frequencies higher than 12.2 GHz and equal to or lower than 12.75 GHz as defined in Section 3 of Chapter 6, Ordinance.

(8) Advanced narrowband CS digital broadcasting:

Standard television broadcasting, high-definition television broadcasting, ultra-short wave broadcasting and data broadcasting carried out as satellite general broadcasting by the advanced narrowband transmission system using satellite broadcast stations that employ radio waves of frequencies higher than 12.2 GHz and equal to or lower than 12.75 GHz as defined in Paragraph 1 of Article 3, Ordinance No.94.

(9) Advanced wide band CS digital broadcasting:

Standard television broadcasting, high-definition television broadcasting, ultra-short wave broadcasting and data broadcasting carried out by the advanced wide band transmission system using satellite basic broadcast stations that employ radio waves of frequencies higher than 12.2 GHz and equal to or lower than 12.75 GHz as defined in Section 5 of Chapter 6, Ordinance.

1.4.2 Abbreviations

ACI: Account Control Information

AL-FEC: Application-Layer Forward Error correction

AMT: Address Map Table

BCD: Binary Coded Decimal

CA: Conditional Access

CAT: Conditional Access Table

CRC: Cyclic Redundancy Check

DSM-CC: Digital Storage Media Command and Control

ECM: Entitlement Control Message

EMM: Entitlement Management Message

ES: Elementary Stream

FEC: Forward Error Correction

HCfB: Header Compression for Broadcasting

HEVC: High Efficiency Video Coding

IEC: International Electrotechnical Commission

IETF: Internet Engineering Task Force

INT: IP/MAC Notification Table

ISO: International Organization for Standardization

ITU-T: International Telecommunication Union, Telecommunication Standardization Sector

IP: Internet Protocol

IPMP:	Intellectual Property Management and Protection
MAC:	Media Access Control
MFU:	Media Fragment Unit
MHEG:	Multimedia Hypermedia Expert Group
MMT:	MPEG Media Transport
MMTP:	MMT Protocol
MPU:	Media Processing Unit
MP:	MMT Package
NIT:	Network Information Table
NTP:	Network Time Protocol
PA:	Package Access
PAT:	Program Association Table
PCR:	Program Clock Reference
PES:	Packetized Elementary Stream
PID:	Packet Identifier
PMT:	Program Map Table
RAP:	Random Access Point
RFC:	Request For Comment (IETF standard)
ROHC:	RObust Header Compression
SL:	Sync Layer
TLV:	Type Length Value
TREF:	Timestamp Reference
TS:	Transport Stream
UDP:	User Datagram Protocol
ULE:	Unidirectional Lightweight Encapsulation

Chapter 2: Multiplexing System

2.1 Transmission by TS packet

2.1.1 Coded signals

Transmission of coded video and audio signals, data signals, metadata signals, the related information (necessary information for domestic subscribers to receive pay boradcasting services or for broadcasters to collect charges for the services, necessary information for broadcasters to make their broadcast programs to be received only by receivers that protect their rights of the programs, and other information notified separately by the Minister for Internal Affairs and Communications), and information indicating the right of broadcast programs (hereinafter referred to as "coded signals") shall comply with the following rules:

1. Coded signals shall be multiplexed by packets.
2. Coded signals shall be grouped to an arbitrary length. Their structures shall comply with PES packet and Section shown in Table No. 1.
3. PES packet and Section shall be transmitted by TS packet shown in Table No. 2.

Table No. 1: Structure of PES packet and Section

PES packet

Header	Extension header	Payload
48 bits		

Notes: 1. The header is used to identify the type of PES packet.
2. The extension header is used to transmit additional infromation of header.
3. The payload is used to transmit data.

Section

(1) General format

Header	Payload
24 bits	8 × N bits

(2) Extended format

Header	Payload	CRC
64 bits	8 × N bits	32 bits

Notes: 1. N represents a positive integer.
2. The header is used to identify the type of Section.
3. The payload is used to transmit data.
4. The CRC is a code for detecting error.

Table No. 2: Structure of TS packet

Header	Adaptation field and payload
4 bytes	184 bytes

Notes: 1. One byte represents eight bits.
2. The header is used to identify the type of TS packet.
3. The adaptation field is used to transmit additional information of the header.
4. The payload is used to transmit PES packet and Section information.

(Paragraph 1 of Article 3, Ordinance)

2.1.2 Transmission control signal

(1) Structure of transmission control signal

Tranmission control for the coded signals transmitted by TS packets shall be performed by the following transmission control signals:

1. PAT specifying the PIDs (packet identifier) of the TS packets that carry the PMTs for the broadcast programs.
2. PMTs specifying the PIDs of the TS packets that carry coded signals comprising broadcast programs (excluding conditional access rerated information) and conditional access common information defined separately by the Minister for Internal Affairs and Communications.
3. CAT specifying the PID of the TS packets that carry conditional access individual information defined separately by the Minister for Internal Affairs and Communications, among the related information.
4. NIT that carries information correlating modulation frequencies with other information on transmission channel with broadcast programs.
5. Program arrangement information that indicates the arrangement sequence of broadcast programs on transmission channel.

(Paragraph 2 of Article 3, Ordinance)

(2) Transmission of transmission control signals

The structures of transmission control signals defined above shall comply with the applicable specified section format.

The transmission procedures for PES packet, section format and TS packet, and the structures of transmission control signals and identifiers shown in Table No. 3 shall comply with the notifications separately given by the Minister for Internal Affairs and Communications.

Table No. 3: Identifiers and their functions

Identifier	Function
Table id	Identifies section type
Descriptor tag	Identifies descriptor type
Stream type id	Identifies coded signal type
Service type id	Identifies service type
Broadcast program number id	Identifies broadcast program number
Service id	Identifies broadcast program number
Network id	Identifies network
Transport stream id	Identifies transport stream
Conditional access system for reception id	Identifies conditional access system for reception
System management id	Identifies broadcasting or non-broadcasting and broadcasting signal standard
Hierachical coding id	Identifies hierachical coding
Scrambling method id	Identifies scrambling method

(Paragraph 3 and Paragraph 4 of Article 3, Ordinance)

(For more information on the transmission procedures for PES packet, section format and TS packet, and the structures of transmission control signals and identifiers shown in Table No. 3, refer to Chapter 3.)

2. 1.3 Emergency alarm signal

The emergency alarm signal shall be transmitted by the emergency information descriptor, and the structure of this descriptor shall comply with the notification separately given by the Minister for Internal Affairs and Communications.

(Article 17, Ordinance)

(For more information on the structure of the emergency information descriptor, refer to Figure No. 11 inSection 3.11.1.)

2.2 IP packet transmission by TS packet

2.2.1 Coded signals

Transmission of coded signals shall comply with the following rules:

1. Coded signals shall be multiplexed by packets.
2. Coded signals shall be grouped to an arbitrary length. Their structures shall comply with IP packet and compressed IP packet (hereinafter referred to as “IP packet or the like”) shown in Table No. 22.
3. IP packet or the like shall be transmitted by ULE packet shown in Table No. 23.
4. ULE packet shall be transmitted by TS packet.

Table No. 22: Structure of IP packet

1. IPv4 packet

IPv4 header	UDP header	Payload
	64 bits	8xN bits

Notes: 1. IPv4 header and UDP header are used to identify the type of IPv4 packet.
2. The payload is used to transmit data.
3. N represents a positive integer.

2. IPv6 packet

IPv6 header	UDP header	Payload
	64 bits	8xN bits

Notes: 1. IPv6 header and UDP header are used to identify the type of IPv6 packet.
2. The payload is used to transmit data.
3. N represents a positive integer.

Table No. 23: Structure of ULE packet

Header	Payload	CRC
		32 bits

Notes: 1. The header is used to identify the type of ULE packet.
2. The payload is used to transmit data.
3. The CRC is a code for detecting error.

(Paragraph 1 of Article 24-3, Paragrapgh 1 of Article 27, Ordinance)

2.2.2 Transmission control signal

(1) Structure of transmission control signal

Transmission control for the coded signals transmitted by TS packets shall be performed by the transmission control signal defined in Section 2.1.2, and AMT (Transmission control signal correlating service id to identify broadcast program number with IP packet or the like) or INT (Transmission control signal correlating service id to identify broadcast program number with IP packet or the like).

(Paragraph 2 of Article 24-3, Paragrapgh 2 of Article 27, Ordinance)

(2) Transmission of transmission control signals

The structures of AMT and INT prescribed above shall comply with the applicable specified section format.

The transmission procedures for IP packet and ULE packet, and the structures of AMT and INT shall comply with the notifications separately given by the Minister for Internal Affairs and Communications.

(Paragraph 3 and 4 of Article 24-3, Paragrapgh 3 and 4 of Article 27, Ordinance)

(AMT and INT are defined in Article 24-3, Ordinance (V-Low multimedia broadcasting based on connected segment tranmission) and Article 27, Ordinance (V-High multimedia broadcasting based on connected segment tranmission), respectively.

2.2.3 Emergency alarm signal

The provision prescribed in Section 2.1.3 is applied.

(Paragraph 7 of Article 24, Article 32, Ordinance)

2.3 Transmission by TLV packet

2.3.1 Coded signals

Transmission of coded signals shall comply with the following rules:

1. Coded signals shall be multiplexed by packets.
2. Coded signals shall be grouped to an arbitrary length. Their structures shall comply with MMTP packet shown in Table No. 59-2.
3. MMTP packet shall be transmitted by IP packet or compressed IP packet shown in Table No. 60.
4. IP packet or compressed IP packet shall be transmitted by TLV packet shown in Table No.61.

Table No. 59-2: Structure of MMTP packet

Header	Payload

→ Order of transmission

Notes: 1. The header is used to identify the type of MMTP packet.
2. The payload is used to transmit data.

Table No. 60: Structure of compressed IP packet

Header	Payload
	8xN bits

Notes: 1. The header is used to identify the type of the compressed IP packet.
2. The payload is used to transmit data.
3. N represents a positive integer.

Table No. 61: Structure of TLV packet

Header	Payload
32 bits	8xN bits

Notes: 1. The header is used to identify the type of TLV packet.
2. The payload is used to transmit data.

(Paragraph 1 of Article 58, Ordinance)

2.3.2 Transmission control signal

(1) Structure of transmission control signal of TLV packet

Transmission control for the coded signals tranmitted by TLV packets shall be perfromed by the following transmission control signals:

1. TLV-NIT that carries information correlating modulation frequencies and other information on transmission channel with broadcast programs.
2. AMT that correlates service identifier for identifing broadcast numbers and IP packet or compressed IP packet.

(Paragraph 2 of Article 58, Ordinance)

The structures of transmission control signal shall comply with the applicable specified section format.

(Paragraph 3 of Article 58, Ordinance)

(2) Structure of transmission control signal of MMTP packet

Transmission control for the coded signals tranmitted by MMTP packets shall be perfromed by the following transmission control signals:

1. PA message that carries broadcast programs table.
2. M2 section message that carries section format.
3. CA message that carries information on identifiers for scrambling method type.

(Paragraph 4 of Article 58, Ordinance)

(3) Transmission of tranmission control signal

The transmission procedures for MMTP packet, compressed IP packet and TLV packet, and the structures of transmission control signals and identifiers shown in Table No. 61-2 shall comply with the notification separately given by the Minister for Internal Affairs and Communications.

Table No. 61-2: Identifiers

Identifier	Function
Description tag	Identifies description type
Conditional access system for reception id	Identifies conditional access system for reception id
Scrambling method type id	Identifies scrambling method type
Service id	Identifies broadcast program number

(Paragraph 3 and 5 of Article 58, Ordinance)

(For more information on the transmission procedures for MMTP packet, compressed IP packet and TLV packet, the structures of transmission control signals and identifiers shown in Table No. 61-2, refer to Chapter 3.)

2.3.3 Emergency alarm signal

The emergency alarm signal shall be transmitted by the emergency information descriptors, and the structure of this descriptor shall comply with the notification separately given by the Minister for Internal Affairs and Communications.

(Article 17 and Article 66, Ordinance)

(For more information on the structure of emergency information descriptor, refer to Figure No.5 in Section 3.11.2.)

<Blank Page>

Chapter 3: Multiplexed Signal Format

3.1 PES packet

Header			Extension header		Payload
Packet start code prefix	Stream id	PES packet length	Optional PES header	Stuffing bytes	Data
24	8	16			

Order of transmission →

Notes:

1. The value of a packet start code prefix shall be set to 0x000001 representing the start of PES packet.
2. The stream id shall be a field used for identifying the type and number of elementary stream (coded signals; the same shall apply hereinafter) and the assignments are given in the table below.
3. The PES packet length shall be a field that writes the number of bytes in the PES packet following this field..
   Note that the value is '0' when the payload of PES packet is video elementary stream, and the packet length is not specified and the boundary is not fixed.
4. The optional PES header shall comply with ITU-T Rec. H.222.0.
5. The value of the stuffing bytes shall be set to 0xFF and shall not exceed 32 bytes in length.

Table: Stream id

Value	Assignment
0xBC	Program stream map
0xBD	Private stream 1
0xBE	Padding stream
0xBF	Private stream 2
'110xxxxx'	Audio stream number 'xxxxx' in ISO/IEC 13818-3, ISO/IEC 11172-3, ISO/IEC 13818-7 or ISO/IEC 14496-3
'1110xxxx'	Video stream number 'xxxx' in ITU-T Rec. H.262, ISO/IEC 11172-2, ISO/IEC 14496-2, or ITU-T Rec. H.264 or ITU-T Rec. H265
0xF0	ECM stream
0xF1	EMM stream
0xF2	DSMCC stream defined in ITU-T Rec. H.222.0 Annex A or ISO/IEC 13818-6
0xF3	Stream defined in ISO/IEC 13522
0xF4	ITU-T Rec. H.222.1 type A
0xF5	ITU-T Rec. H.222.1 type B
0xF6	ITU-T Rec. H.222.1 type C
0xF7	ITU-T Rec. H.222.1 type D
0xF8	ITU-T Rec. H.222.1 type E
0xF9	Auxiliary stream
0xFA	SL-packetized stream defined in ISO/IEC 14496-1
0xFB	FlexMux stream defined in ISO/IEC 14496-1
0xFC	Meta data stream
0xFD	Extended stream ID

0xFE	Undefined
0xFF	Program stream directory

Numbers enclosed in ‘ ’ represents binary numbers.The same shall apply hereinafter.

(Table No.4, Notification)

3.2 Section

1. Genaral format

← Header →

Table id	Section syntax indicator	‘1’	‘11’	Section length	Data
8	1	1	2	12	8 × N

→ Order of transmission

2. Extended format

← Header →

Table id	Section syntax indicator	‘1’	‘11’	Section length	Table id extension	‘11’	Version number	Current next indica-tor	Section number	Last section number	Data	CRC
8	1	1	2	12	16	2	5	1	8	8	8 × N	32

→ Order of transmission

Notes:

1. The table id shall be a field used for identifying the table to which the section belongs.
2. The section syntax indicator shall be a field determining whether normal or extension format is used and shall be ‘0’ for normal format and ‘1’ for extension format.’
3. The section length shall be a field that writes the number of data bytes following this field and shall not exceed 4093.
4. The table id extension shall be a field extending the table identifier.
5. The version number shall be a field that writes the table version number.
6. The current next indicator shall contain ‘1’ and ‘0,’ respectively, when the table is currently used and when the table cannot be used at present but will be valid next.
7. The section number shall be a field that writes the number of the section comprising the table.
8. The last section number shall be a field that writes the number of the last section comprising the table.
9. The CRC shall comply with ITU-T Rec. H.222.0.

(Table No.5, Notification)

3.3 TS packet

← 188 bytes →

← Header →

Sync byte	Transport error indicator	Payload unit start indicator	Transport priority	PID	Transport scrambling control	Adaptation field control	Continuity index	Adaptation field/payload
8	1	1	1	13	2	2	4	

→ Order of transmission

Notes:

1. The sync byte shall be 0x47.
2. The transport error indicator shall be a flag that indicates whether there is any bit error or not in the TS packet. If this flag contains '1,' it indicates that the TS packet has an uncorrectable error of at least one bit.
3. The payload unit start indicator shall indicate that the payload of this TS packet starts at the PES packet start or pointer when it contains '1.'
4. The transport priority shall be a flag that indicates priority among packets with the same PID. The packet is given priority if this flag contains '1.'
5. The PID shall be a field used for identifying the payload data type and the assignment shall be as shown in Table No. 1.
6. The transport scrambling control shall be a field identifying the payload scrambling mode for TS packet. The value of this field shall be as shown in Table No. 2.
7. The adaptation field control shall be a field used for indicating the configuration of the adaptation field/payload, the value of which shall be as shown in Table No. 3.
8. The continuity index shall be a field specifying the sequence of TS packets with the same PID. The value of this field shall start with '0000' and be incremented by 1. The value shall change back to '0000' after '1111.'
   However, note that it shall be ensured that the same TS packet is transmitted only up to twice in a row and that in this case the value of this field shall not be incremented.
9. The adaptation field shall comply with ITU-T Rec. H.222.0.

Table No. 1: PID assignments

Value	Description
0x0000	PAT
0x0001	CAT
0x0002 – 0x000F	Reserved
0x0010	NIT
0x0011 – 0x1FFE	May be assigned to other than PAT, CAT, NIT and Null packet
0x1FFF	Null packet

Table No. 2: Scrambling control value

Value	Description
‘00’	Not scrambling
’01’	Reserved
’10’	Scrambled by Even key
‘11’	Scrambled by Odd key

Table No. 3: Adaptation field control value

Value	Description
‘00’	Reserved
’01’	No adaptation field, payload only
’10’	Adaptation field only, no payload
‘11’	Adaptation field followd by payload

(Table No.6, Notification)

The usage criteria for PIDs shown in Table No. 1 shall be as follows (specified in ARIB STD-B10):

Type	Value range	Notes
Specified by the the Ministry of Internal Affairs and Communications	0x0000 – 0x0010, 0x1FFFF	Specified in the Notification
Specified by the standardization organization	0x0011 – 0x002F(Note)	Used after deliberations
Specified by companies	ranges that do not interfere with the above	Registration and release
Used by companies	ranges that do not interfere with the above	Indirect designation by PMT

(Note) 0x0015 and 0x0016 are specified and used by companies before formulation of ARIB STD-B10.

3.4 IP packet

1. IPv4 packet

IPv4 header	UDP header	Payload
	64 bits	8xN bits

→ Order of transmission

Notes: 1. The structure of IPv4 header is shown in the following No.1.
2. The structure of UDP headee is shown in the following No.2.

2. IPv6 packet

IPv6 header	UDP header	Payload
	64 bits	8xN bits

→ Order of transmission

Notes: 1. The structure of IPv6 header is shown in the following No.3.
2. The structure of UDP headee is shown in the following No.2.

No1: Structure of IPv4 header

Version	Heasder length	Service type	Packet length	Identifier	Flag	Fragment offset	Lifetime	Protocol
4	4	8	16	16	3	13	8	8

→ Order of transmission

Header cheksum	Source address	Destination address	Extension information
16	32	32	32xN

Note 1: Each item shall comply with IETF RFC 791.

No2: Structure of UDP header

Source port	Destination port	Data length	Checksum
16	16	16	16

→ Order of transmission

Note 1: Each item shall comply with IETF RFC 768.

No3: Structure of IPv6 header

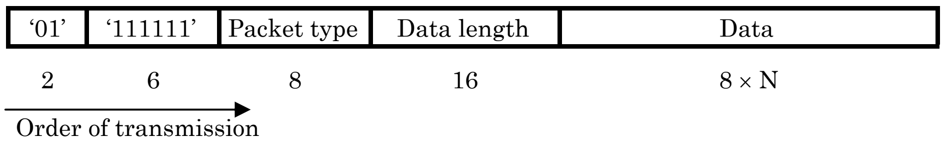
Version	Traffic class	Flow label	Payload length	Next header	Hop limit	Sourece address	Destination address	Extension header
4	8	20	16	8	8	128	128	0 or 8xN

→ Order of transmission

Note1: Each item shall comply with IETF RFC 2460.

(Table No.7, Notification)

3.5 TLV packet



Note 1: The packet type is a field used for identifying the type of packets stored in TLV and its assignment shall be shown as the table below.

2: The data length shall be a field that writes the number of data bytes following this field.

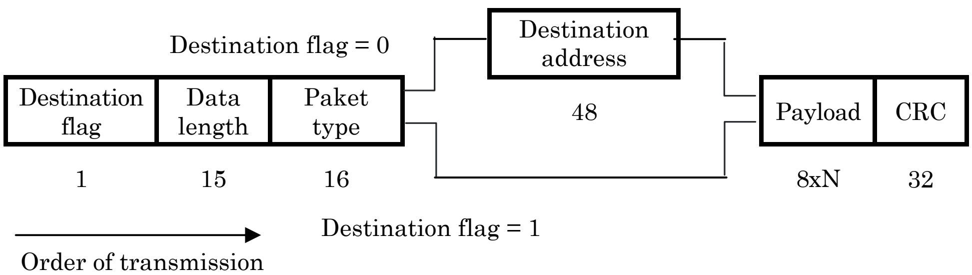
Packet type assignment

Value of packet type	Assignment
0x00	Undefined
0x01	IPv4 packet
0x02	IPv6 packet
0x03	Header compressed IP packet
0x04 - 0xFD	Undefined
0xFE	Transmission control signal packet
0xFF	Null packet

Note: Null packet shall be the packet that writes the byte sequence of 0xFF having the length indicated in the data length on the data.

(Table No.11, Notification)

3.6 ULE packet



Notes:

1. The destination flag is a field used for identifying whether there is a destination address or not and its value shall be shown below.
2. The data length shall be a field that writes the number of data bytes following the packet type.
3. The packet type shall be a field used for identifying the type of packets stored in

the payload.

4. The destination address shall be a field that writes the ULE packet destination address.
5. The CRC shall comply with IETF RFC 4326.

Destination flag value

Value	Description
0	Destination address field is present
1	Destination address field is absent

(Table No.8, Notification)

The packet type shall comply with the following assignment.

Value	Description	Remarks
0x0001	Bridged frame	RFC 4326
0x0800	IPv4 packet	Ether Type, RFC 4326
0x22F1	Compressed IP packet by ROHC	Ether Type
0x22F2	Compressed IP packet by HCfB	Ether Type
0x86DD	IPv6 packet	Ether Type, RFC 4326

3.7 Compressed IP packet

(1) HCfB

Header	Payload
	8xN bits

→ Order of transmission

Notes: 1. The structure of the header is shown in the following No.1.

No.1 Structure of the compressed IP packet header

Context id	Sequence number	Header type of context identification	Compressed header
12	4	8	

→ Order of transmission

Notes:

1. The context identifier shall represent a header compressed IP packet flow (an aggregate IP packet having the same combination of the value for five fields: IPv4 packet protocol or IPv6 packet next header, source address, destination address, source port and destination port).
2. The sequence number shall represent a sequence of the header compressed IP packet having the same context identifier.

3. The header type of context identification shall be a field used for identifying compressed header type and its assignment shall be shown in the following No. 2.
4. The compressed header shall be a field that writes IPv4 header or IPv6 header and UDP header as compressed information by the designated method specified in the context identification header type.

No.2 Assignment for the header type of context identification

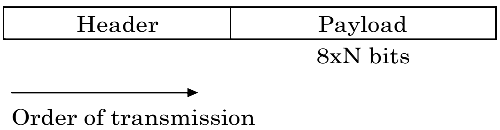
Header type value of context identification	Assignment
0x20	Partial IPv4 header and partial UDP header
0x21	IPv4 header identifier
0x60	Partial IPv6 header and partial UDP header
0x61	No compressed header
Other than the above	Undefined

Notes:

1. The partial IPv4 header shall be the header excluding packet length, header checksum and extension information from the IPv4 header.
2. The partial IPv6 header shall be the header excluding payload length from IPv6 header.
3. The partial UDP header shall be the header excluding data length and checksum from UDP header.

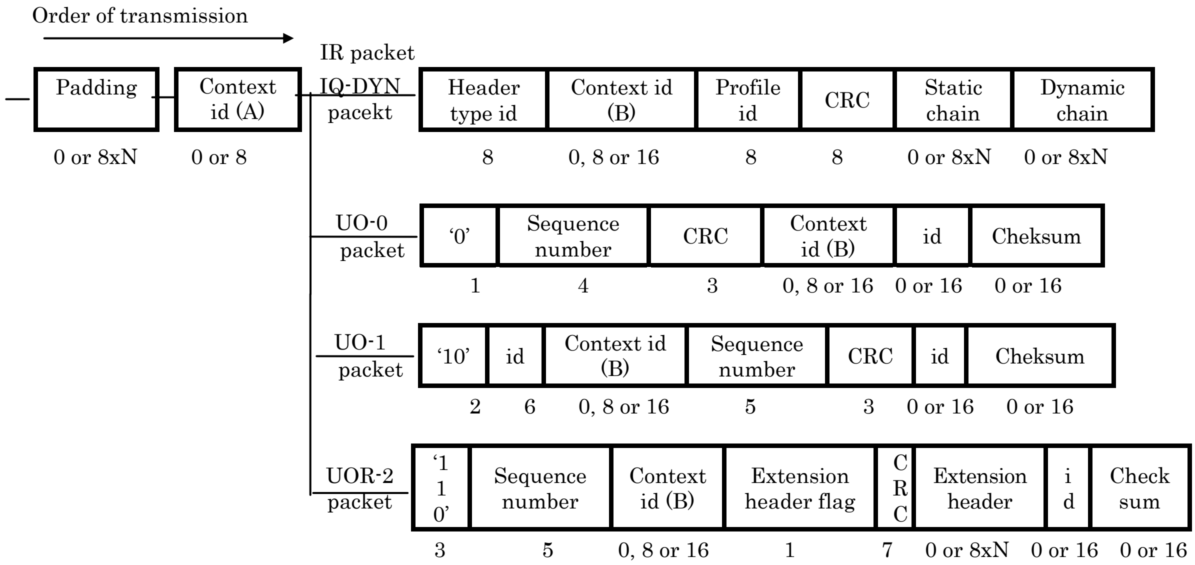
(Table No. 10, Notification)

(2) ROHC



Note 1: The structure of the header shall be shown below.

Structure of the ROHC header



Notes:

1. The context id (A) and context id (B) shall repesent the header compressed IP packet flow (an aggregate IP packet having the same combination of the value for five fields: IPv4 packet protocol or IPv6 packet next header, source address, destination address, source port and destination port). The context id (A) is used only when the number of IP packet flow is 15 and less. The context id (B) is used only when the number of IP packet flow is 16 and more.
2. The IR packet is used to initialize a context and its header type id shall be set to '11111101'.
3. The IR-DYN packet is used to update a part of the header information on the equal to or upper than IP layer and its header type id shall be set to '1111 1000'.
4. When the header information of UO-0packet, UO-1packet or UOR-2packet is not updated on the the equal to or upper tha IP layer, any of the packets transmit data.
5. The static chain is used only for IR packet. The dynamic chain is used for IR packet or IR-DYN packet.
6. The profile id shall be set to 0x0002 (UDP profile).
7. The padding, sequence number, CRC, identifier, checksum, static chain, dynamic chain,extension header flag and extension header shall comply with RFC 3095.

3.8 Bridged frame

Bridged frame header	Payload	CRC
Variable length	8xN bits	32 bits

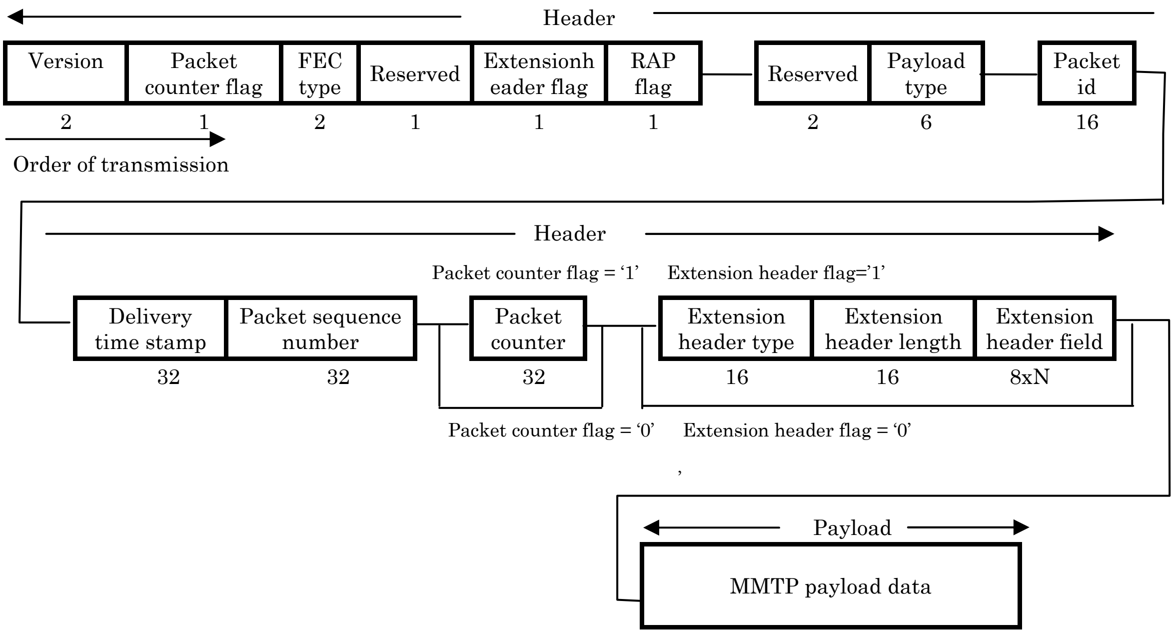
Order of transmission →

Notes:

1. The bridged frame header is used to identify the type of bridged frame. The structure of the bridged frame header shall comply with IETF RFC 4326. The details shall be set by the opration rules specified by operators.

2. The payload is used to transmit data.
3. N represents a positive integer.
4. The CRC is a code for detecting error.

3.9 MMTP packet



Notes:

1. The version shall be '00'.
2. The packet counter flag shall be '0' when the packet counter does not exist, and '1' when the counter exsits.
3. The FEC type shall represent information on AL-FEC (Decoding extracted information by using pre-generated information. The same shall apply hereinafter.) of this packet. The FEC type assignment shall be shown in the following table.

Value	Assignment
0	Non-protected MMTP packet by AL-FEC
1	Source packet among MMTP packets protected by AL-FEC
2	Repair packet among MMTP packets protected by AL-FEC
3	Reserved

4. The value of the extension header flag shall be '0' when the MMTP packet header is not extended and '1' when the header is extended.
5. The value of the RAP flag shall be '1' when the MMTP payload transmitted by MMTP packet contains the head of random access point and '0' otherwise.
6. The payload type represents a data type of MMTP payload and its assignment shall be shown in the following table.

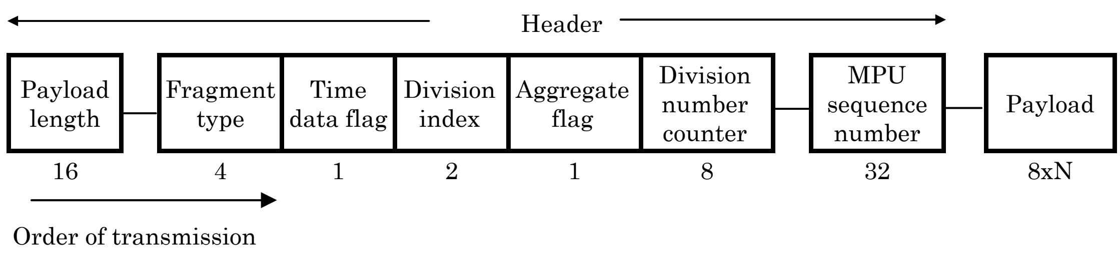
Value	Assignment
0x00	MPU
0x01	Undefined
0x02	Contains one or more control messages
0x03 - 0x3F	Undefined

7. The packet id shall be a field used for identifying payload data type. The assignment shall be shown in the following table.

Value	Assignment
0x0000	PA message
0x0001	CA message
0x0002 - 0xFFFF	Undefined

8. The delivery time stamp shall represent the time when the head byte of this MMTP packet is outputted from the broadcasting station by the short NTP time stamp described in IETF RFC 5905.
9. The packet sequence number shall be the sequence of MMTP packets having the same packet identifier.
10. The packet counter shall be the sequence of MMTP packets in the same IP data flow, regardless of the packet id value.
11. The extension header type shall be an extension type in the extension header field.
12. The extension header length shall be a field that writes the number of data byte in the extension field.
13. The extension header field shall be data byte for extending header.
14. The MMT payload data shall be shown in the following Figure No.1 or Figure No.2, respectively, when the payload type is 0x00 or 0x02.

Figure No. 1



Notes:

1. The payload length shall be a field that writes the number of data bytes following this field.
2. The fragment type represents a fragment type of information stored in the MMTP payload and its assignment shall be shown in the following table.

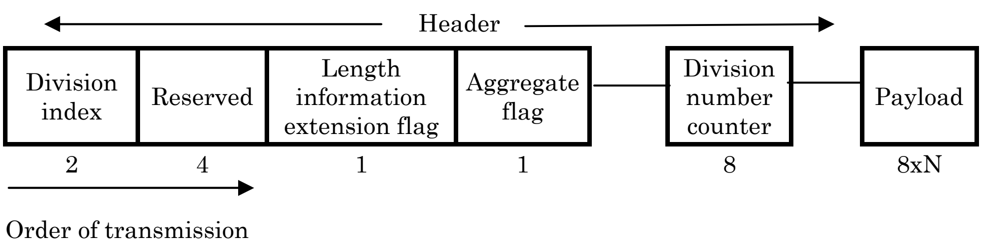
Value	Assignment
0	MPU metadata
1	Movie fragment metadata
2	MFU (Part of coded video and audio signals, and etc. The same shall apply hereinafter.)
3 - 15	Reserved

3. The time data flag shall be ‘0’ when the data stored in the MMTP payload does not specify the presentation time and ‘1’ when the data specify the time.
4. The division index represents a divided state of the data stored in the MMTP payload and its assignment shall be shown in the following table.

Value	Assignment
‘00’	Undivided
‘01’	Divided, Including the head part of the data before division
‘10’	Divided, Not including the head part and end part of the data before division
‘11’	Divided, Including the end part of the data before division

5. The aggregate flag shall be ‘0’ when one data is stored in the MMTP payload and ‘1’ when two or more data are stored in the payload.
6. The division number counter shall indicate a number of divided data after the data part stored in this MMTP payload, when the data are divided. The counter is restarted from ‘0’ when this value exceeds 255. Note that the value is set to ‘0’ when the aggregate flag is ‘1’.
7. The MPU sequence number shall represent a sequence number of MPU to which MPU metadata, movie fragment metadata and MFU belong in the case of storing them.

Figure No.2



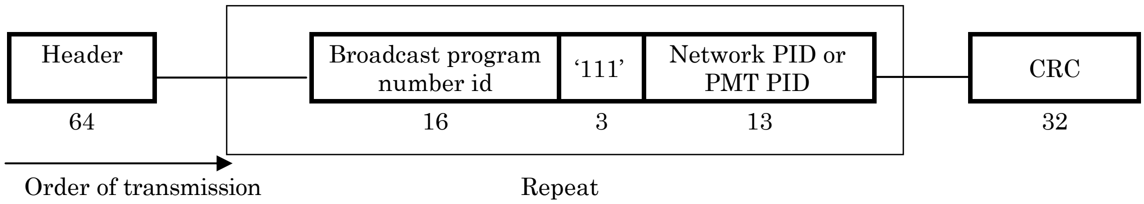
Notes:

1. The division index shall be a divided state of the data stored in the MMTP payload and be shown in the table of Note 4 in the above Figure No.1.
2. The length information extension flag shall be ‘0’ or ‘1’, respectively, when the message data length indicating message size is 16 bits or 32bits.

(Table No.9, Notification)

3.10 Transmission control signal

(1) PAT

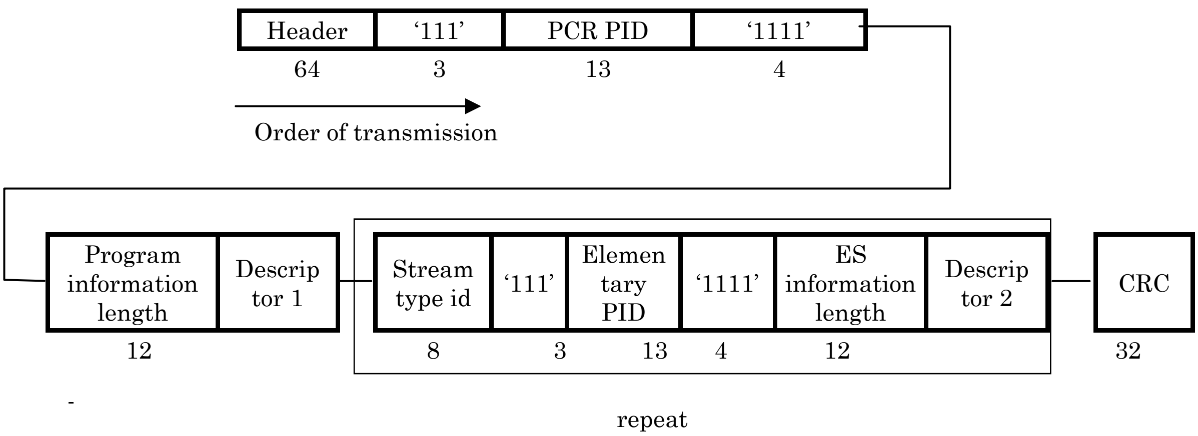


Notes:

1. The header and CRC shall be the same as those for extended section format shown in Section 3.2.
   Note that the content of the bit that follows the "section syntax indicator" shall be '0.'
2. The value of the "table id" within the header shall be 0x00, representing the PAT. The "table id extension" shall be used to transmit the transport stream id.
3. The broadcast program id number shall be used to identify broadcast program number. '0' shall be used for NIT.
4. The network PID or PMT PID represents NIT PID when the program number is '0,' and the value of this field shall be 0x0010. The network PID or PMT PID represents PMT PID when the program number is any number other than '0.'

(Notification)

(2) PMT



Notes:

1. The header and CRC shall be the same as those for extended section format shown in Section 3.2.
   Note that the content of the bit that follows the "section syntax indication" shall be '0.'
2. The value of the "table id" within the header shall be 0x02, representing the PMT. The "table id extension" shall be used to transmit the program number.
3. The PCR PID represents the PID of the TS packet that transmits the valid PCR field for the broadcast program specified by the broadcast program number id.

4. The value of the first two bits of the program information length shall be ‘00.’ The remaining 10 bits shall be a field that writes the number of bytes in the descriptor that follows the program information length.
5. Descriptor 1 shall be a field that writes the descriptor related to the applicable broadcast program and descriptor 2 shall be a field that writes the descriptor related to the applicable elementary stream.
6. The stream type id shall be used to identify broadcast program element type and its assignment shall be shown below.
7. The elementary PID represents the PID for the TS packet that transmits related broadcast program element.
8. The value of the first two bits of the ES information length shall be ‘00.’ The remaining 10 bits shall be a field that writes the number of bytes in the descriptor that follows the ES information length.

Table: Stream type identifier

Value	Assignment
0x00	Undefined
0x01	Video stream defined in ISO/IEC 11172-2
0x02	Video stream prescribed in ITU-T Rec. H.262 or constrained parameter video stream defined in ISO/IEC 11172-2
0x03	Audio stream defined in ISO/IEC 11172
0x04	Audio stream defined in ISO/IEC 13818-3
0x05	Private sections prescribed in ITU-T Rec. H.222.0
0x06	PES packets containing private data prescribed in ITU-T Rec. H.222.0
0x07	MHEG defined in ISO/IEC 13522
0x08	DSM-CC prescribed in ITU-T Rec. H.222.0 Annex A
0x09	ITU-T Rec. H.222.1
0x0A – 0x0D	ISO/IEC 13818-6 (type A – D)
0x0E	Data type prescribed in ITU-T Rec. H.222.0 other than the above
0x0F	Audio stream defined in ISO/IEC 13818-7
0x10	Video stream defined in ISO/IEC 14496-2
0x11	Audio stream defined in ISO/IEC 14496-3
0x12	SL-packetized stream or FlexMux stream carried in PES packets defined in ISO/IEC 14496-1
0x13	SL-packetized stream or FlexMux stream (defined in ISO/IEC 14496-1) carried in the sections (defined in ISO/IEC 14496)
0x14	Synchronized download protocol defiend in ISO/IEC 13818-6
0x15	Metadata carried in PES packets
0x16	Metadata carried in metadata_sections
0x17	Metadata carried in Data carousel defined in ISO/IEC 13818-6
0x18	Metadata carried in Object carousel defined in ISO/IEC 13818-6
0x19	Metadata carried in Synchronized download protocol defined in ISO/IEC 13818-6
0x1A	IPMP stream defined in ISO/IEC 13818-11
0x1B	AVC video stream prescribed in ITU-T Rec. H.264
0x1C	Audio stream defined in ISO/IEC 14496-3 (not using additional transport structure)
0x1D – 0x23	Undefined

Value	Assignment
0x24	HEVC video stream or HEVC video sub bits stream in the time direction prescribed in ITU-T Rec. H265
0x25	Vido subset of HEVC video stream in the time direction based on one or more profiles prescribed in ITU-T Rec. H265 Annex
0x26 – 0x7E	Undefined
0x7F	IPMP stream

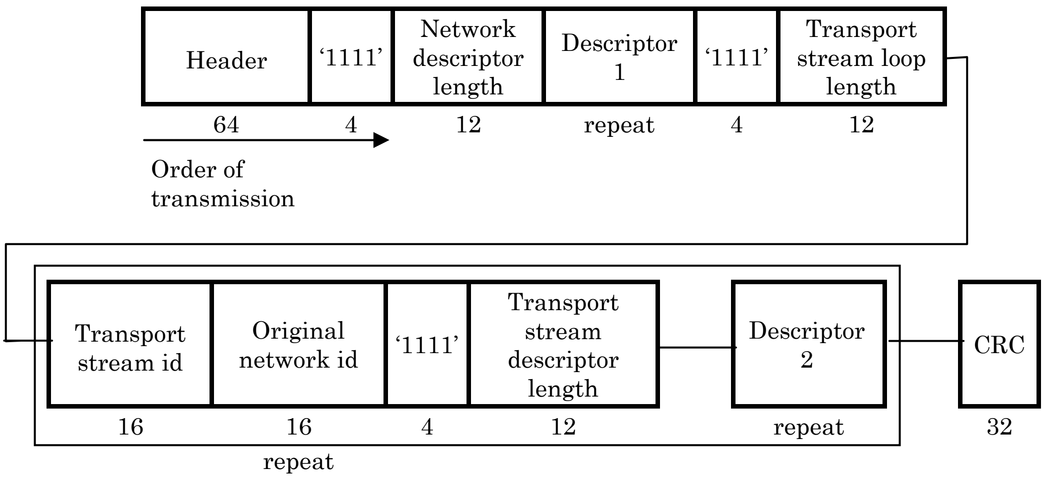
(Table No.16, Notification)

(3) CAT

Notes:
1. The header and CRC shall be the same as those for the extended section format shown in Section 3.2.
   Note that the content of the bit that follows the "section syntax indication" shall be '0.'
2. The value of the "table id" within the header shall be 0x01, representing the CAT.
   The "table id extension" field is undefined.

(Table No.17 Notification)

(4) NIT

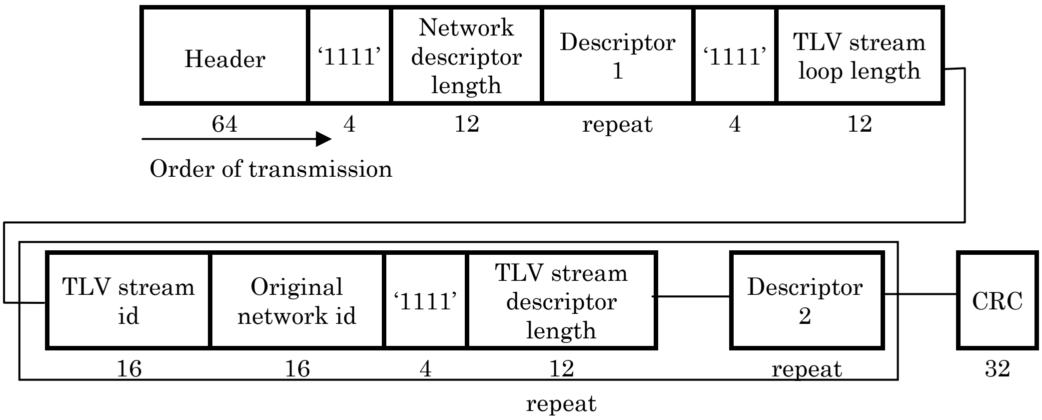


Notes:
1. The header and CRC shall be the same as those for section extension format shown in Section 3.2.
2. The value of the "table id" within the header shall be 0x40 for actual network and 0x41 for any other network. The "table id extension" shall be used to transmit network id.
3. The network id shall be a field used for identifying the network number.

4. The value of the first two bits of the network descriptor length shall be '00.' The remaining 10 bits shall be a field that writes the number of bytes in the descriptor that follows the network descriptor length.
5. The descriptor 1 or descriptor 2 shall be a field for writing the descriptor related to the applicable network.
6. The value of the first two bits of the transport stream loop length shall be '00.' The remaining 10 bits shall be a field that writes the number of data bytes just before the CRC following this field.
7. The transport stream id represents the identification number of the applicable transport stream.
8. The original network id represents the identification number of the original network of the applicable transport stream.
9. The transport stream descriptor length represents the number of bytes in all descriptors of the applicable transport stream just behind this field. Note that the value of the first two bits shall be '00.'

(Table No.18-1, Notification)

(5) TLV-NIT



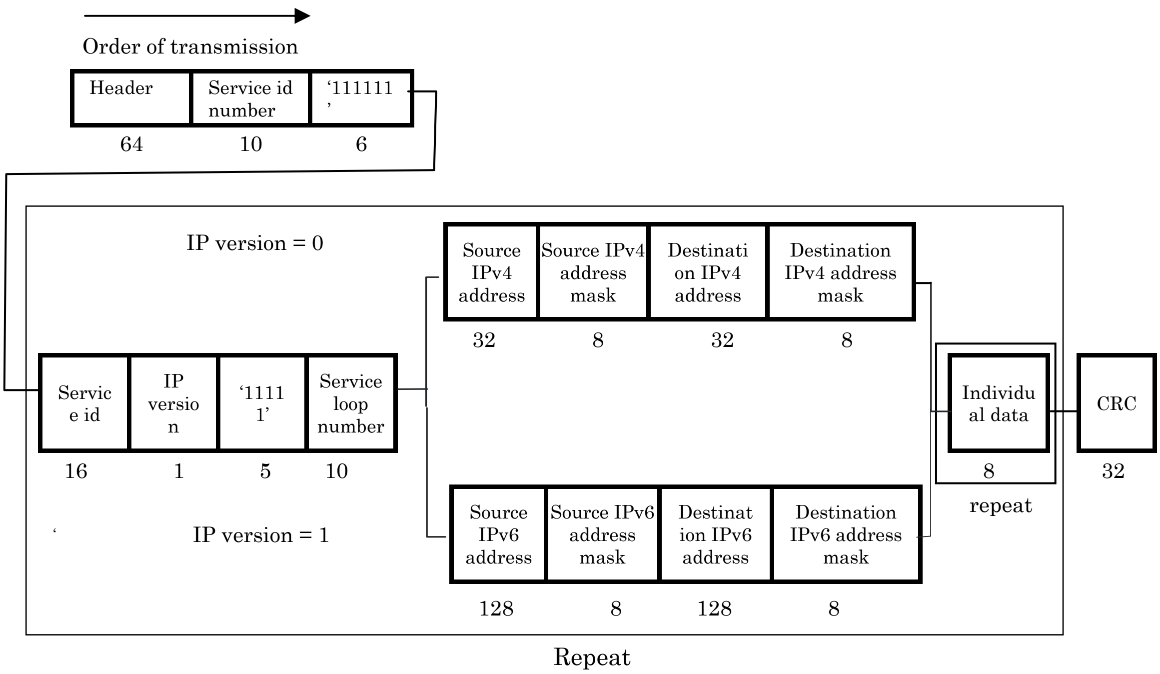
Notes:

1. The header and CRC shall be the same as those for section extension format shown in Section 3.2.
2. The value of the "table id" within the header shall be 0x40 for actual network and 0x41 for any other network. The "table id extension" shall be a field used to transmit network id.
3. The network id shall be a field used for identifying the network number.
4. The value of the first two bits of "the section length" within the header shall be '00.' The remaining 10 bits shall be a field that writes the number of data bytes following this field. Note that this vakue shall not exceed 1021.
5. The value of the first two bits of the network descriptor length shall be '00.' The remaining 10 bits shall be a field that writes the number of bytes in the descriptor that follows the network descriptor length.
6. The descriptor 1 or 2 shall be a field for writing the descriptor related to the applicable network.

7. The value of the first two bits of the TLV stream loop length shall be '00.' The remaining 10 bits shall be a field that writes the number of data bytes just before CRC following this field.
8. The TLV stream id shall be a field that writes the identification number of TLV stream.
9. The original network id shall be a field that writes identification number of the original network of the applicable TLV stream.
10. The TLV stream descriptor length represents the number of bytes in all descriptors of the applicable TLV stream just behind this field. Note that the value of the first two bits shall be '00.'

(Table 18-2, Notification)

(6) AMT



Notes:

1. The header and CRC shall be same as those for the extension format shown in Section 3.2.
2. The value of the "table id" within the header shall be 0xFE. The "table id extension" shall be be 0x0000 repersenting AMT.
3. The service id number shall represent the number of service id described in this AMT.
4. The service id shall be a field to be used for identifying broadcast program numbers.
5. The IP version shall represent the version and its assignemnt shall be shown in the following table.

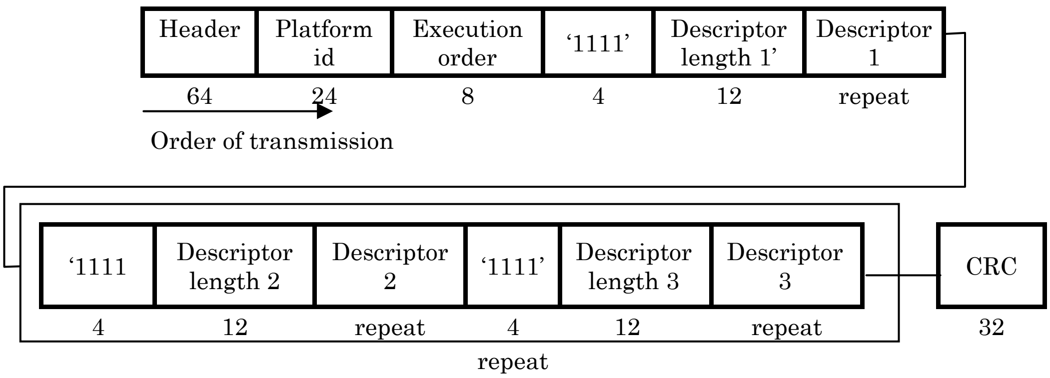
Value	Assignment
0	IPv4
1	IPv6

6. The service loop length shall represent the number of bytes from the field just behind this field to the next listed service id field or to the field just before CRC.

7. The source IPv4 address shall be a field that writes the source IP address for the IPv4 packet comprising broadcasting program.
8. The source IPv4 address mask shall be a field used for specifying the number of bits from the valid most significant bit for the IP address designated to the source IPv4 address.
9. The destination IPv4 address shall be a field that writes the destination IP address for the IPv4 packet comprising broadcast program.
10. The destination IPv4 address mask shall be a field used for specifying the number of bits from the valid most significant bit for the IP address designated to the destination IPv4 address.
11. The source IPv6 address shall be a field that writes the source IP address for the IPv6 packet comprising broadcast program.
12. The source IPv6 address mask shall be a field used for specifying the number of bits from the valid most significant bit for the IP address designated to the source IPv6 address.
13. The destination IPv6 address shall be a field that writes the destination IP address for the IPv6 packet comprising broadcast program.
14. The destination IPv6 address mask shall be a field used for specifying the number of bits from the valid most significant bit for the IP address designated to the destination IPv6 address.
15. The individual data shall be a field that writes the data defined individually.

(Table No.19, Notification)

(7) INT

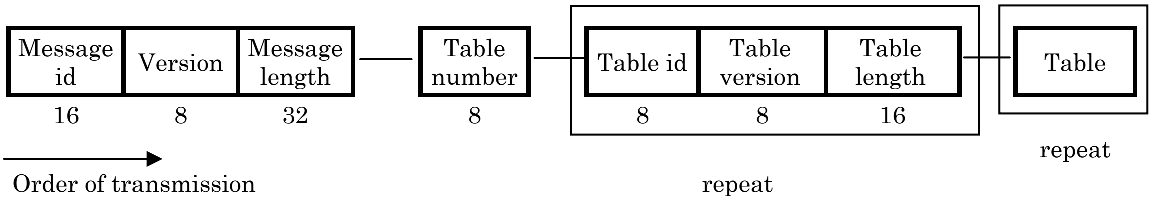


Notes:
1. The header and CRC shall be the same as those for section extension format shown in Section 3.2.
2. The value of the "table id" within the header shall be 0x4C. The "table id extension" is used to transmit auxiliary infromation to identify the execution type and platform id.
3. The platform id shall be a field used for identifying the platform type.
4. The descriptor length 1, descriptor length 2 or descriptor length 3 shall represent the byte length of each descriptor, respectively.
5. The descriptor 1 shall be a field that writes an applicable platform descriptor.
6. The descriptor 2 shall be a field that writes a descriptor identifying a receiver.

7. The descriptor 3 shall be a field that writes a descriptor for the receiver id specified by the descriptor 2.

(Table No.20, Notification)

(8) PA message



Notes:

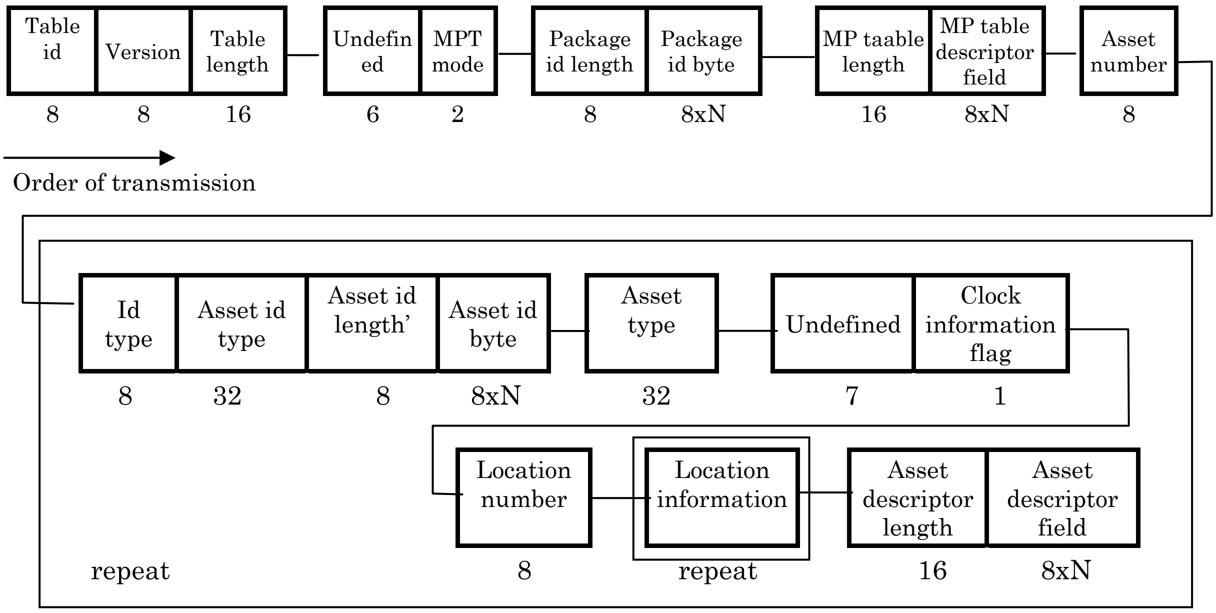
1. The message id shall be 0x0000 representing PA message.
2. The version shall be a field that writes the version number of PA message.
3. The message length shall be a field that writes the number of data bytes following this field.
4. The table number shall be a field that writes the number of table stored in this PA message.
5. The table id shall be a field used for identifying the table stored in this PA message. The assignment shall be shown in the following table.

Value	Assignment
0x00 - 0x10	Undefined
0x11	Subset 0 MP table
0x12	Subset 1 MP table
0x13	Subset 2 MP table
0x14	Subset 3 MP table
0x15	Subset 4 MP table
0x16	Subset 5 MP table
0x17	Subset 6 MP table
0x18	Subset 7 MP table
0x19	Subset 8 MP table
0x1A	Subset 9 MP table
0x1B	Subset 10 MP table
0x1C	Subset 11 MP table
0x1D	Subset 12 MP table
0x1E	Subset 13 MP table
0x1F	Subset 14 MP table
0x20	Complete MP table
0x21 - 0x81	Undefined
0x82, 0x83	ECM
0x84, 0x85	EMM

0x86	CA table
0x87 - 0xFF	Undefined

6. The table version shall be a version of the table stored in PA message.
7. The table length shall be a field that writes the number of data bytes of the table stored in this PA message.
8. The table shall represent control information and be shown below.

MP TABLE



Notes:
1. The value of the table id shall be set to 0x20 when this MP table is a complete MP table and be 0x11 to 0x1F corresponding to the subset 0 to subset 14 when the one package configuration is described by some MP tables.
2. The version shall be a field that writes the version number for a table.
3. The table length shall be a field that writes the number of data bytes following this field.
4. The MPT mode represents the action when MP table is divided into subset. The mode assignment shall be shown in the following tabel.

Value	Assignment
'00'	Processing in the order of subset
'01'	Processing of any subset having the same version after receiving MP table with subset 0
'10'	Optional processing of MP table of subsets
'11'	Undefined

5. The IP package length shall be a field that writes the number of data bytes for the package ID byte.

6. The package ID byte represents the package ID.
7. The MP table descriptor length shall be a field that writes the number of data bytes in the MP table descriptor field.
8. The MP table descriptor field shall be a field that stores MP table descriptor.
9. The asset number shall be a field that writes the number of asset providing this table information.
10. The value of id type shall be set to 0x00 representing asset id in the ID system of the MMTP packet flow.
11. The asset ID format shall be a field used for identifying program elements signal such as video, audio and data.
12. The asset ID length shall be a field that writes the number of data bytes of the asset ID byte.
13. The asset ID byte represents asset ID.
14. The asset type represents the type of asset and its assignment shall be shown in the following table.

Code	Assignment
hvc1	HEVC video stream prescribed in ITU-T Rec. H.265
mp4a	Audio stream defined in ISO/IEC 14496-3

15. The clock information flag shall be ‘1’ or ‘0’ whether clock information id flag field and time scalable flag field exist or not.
16. The location number shall represent the number of asset location information.
17. The asset descriptor length shall be a field that writes the number of data bytes having the asset descriptor field size.
18. The asset descriptor field shall be a field that stores the descriptor indicating asset information.

(Table No.27, Notification)

(9) M2 section message

Message id	Version	Message length	Table id	Section syntax direction	‘1’	‘11’	Section length	Table identification extension
16	8	16	8	1	1	2	12	16

Order of transmission →

‘11	Version number	Current next indicator	Section number	Last section number	Data	CRC
2	5	1	8	8	8xN	32

Notes:

1. The value of the message id shall be set to 0x8000 representing M2 section message.
2. The version shall be a field that writes a version number of M2 section message.
3. The message length shall be a field that writes a number of data bytes following this field.
4. The table id shall be a field used for identifying the table to which the section belongs.
5 The value of section syntax indicator shall be ‘1’ representing the extension format among the formats shown in Section 3.2.
6. The section length shall be a field that writes a number of data bytes following this field.
7. The version number shall be a field that writes a table version number.
8. The current next indicator shall be ‘0’ when the table is not currently available and will be valid next, and ‘1’ when the table is available at present.
9. The section number shall be a field that writes a section number comprising table.
10. The last section number shall be a field that writes the last section number comprising table.
11. The CRC shall comply with ITU-T Recommendation H.222.0.

(Table No.28, Notification)

(10) CA message

Message identifier	Version	Message length	Table
16	8	16	

Order of transmission →

Notes: 1. The value of message id shall be 0x8001 representing CA message.
2. The table represents control information and shall be shown below.

CA Table

Table id	Version	Table length	Descriptor
8	8	16	repeat

Order of transmission →

Note: The value of table id shall be 0x86 representing CA table.

(Table No.29, Notification)

3.11 Descriptors

3.11.1 Transmission by TS packet and TLV packet

Descriptor	Configuration	Transmission system*	
		TS	TLV
Conditional access descriptor	As per Figure No. 1	○	
Conditional playback descriptor	As per Figure No. 2	○	
Partial reception descriptor	As per Figure No. 3	○	
Terrestrial delivery system descriptor	As per Figure No. 4	○	
Satellite delivery system descriptor	As per Figure No. 5	○	○
Service list descriptor	As per Figure No. 6	○	○
System management descriptor	As per Figure No. 7	○	○
Data component descriptor	As per Figure No. 8	○	
Carousel compatible composite descriptor	As per Figure No. 9	○	
Copyright descriptor	As per Figure No. 10	○	
Emergency information descriptor	As per Figure No. 11	○	
IP/MAC stream arrangement descriptor	As per Figure No. 12	○(Note1)	
Access control descriptor	As per Figure No. 13	○	
Hierarchical coding descriptor	As per Figure No. 14	○	
Scrambling method descriptor	As per Figure No. 15	○	

* Refer to Section 2. Tranmission system availabe for each descriptor is shown by ○.
Note 1: Only used for INT.

Figure No. 1: Conditional access descriptor

Descriptor tag	Descriptor length	Conditional access id	'111'	Conditional access PID	Data
8	8	16	3	13	8 × N

Order of transmission →

Notes:

1. The value of the descriptor tag shall be 0x09, representing the conditional access descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The conditional access identifier shall be a field used for identifying the conditional access.
4. The conditional access PID shall be a field that writes the PID of the TS packet that contains information related to the conditional access.
5. This descriptor may be tranmitted in the descriptor field of CAT or in the descriptor 1 or 2 field of PMT.

Figure No. 2: Conditional playback descriptor

Descriptor tag	Descriptor length	Conditional playback id	'111'	Conditional playback PID	Data
8	8	16	3	13	8 × N

Order of transmission →

Notes:
1. The value of the descriptor tag shall be 0xF8, representing the conditional playback descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The conditional playback identifier shall be a field used for identifying the conditional playback.
4. The conditional playback PID shall be a field that writes the PID of the TS packet that contains information related to the conditional playback.
5. This descriptor may be transmitted in the descriptor field of CAT or in the descriptor 1 or 2 field of PMT.

Figure No. 3: Partial reception descriptor

Descriptor tag	Descriptor length	Service id (repeat)
8	8	16

Order of transmission →

Notes:
1. The value of the descriptor tag shall be 0xFB, representing the partial reception descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The service id shall be a field used for identifying the program number of the broadcast program transmitted in the partial reception segment.
4. This descriptor shall be used only when there is a partial reception segment in digital terrestrial sound broadcasting, digital terrestrial television broadcasting, V-Low

multimedia broadcasting based on connected segment system or V-High multimedia broadcasting based on connected segment system. The descriptor may be transmitted in the descriptor 2 field of NIT.

Figure No. 4: Terrestrial delivery system descriptor

Descriptor tag	Descript or length	Area code	Guard interval ratio	Mode information	Center frequency
8	8	12	2	2	16

Order of transmissio →

repeat

Notes:

1. The value of the descriptor tag shall be 0xFA, representing the terrestrial delivery system descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The area code shall be a field used for identifying the coverage area.
4. The guard interval ratio is a field used for identifying the ratio of guard interval to valid symbol length. ‘00,’ ‘01,’ ‘10’ and ‘11’ shall represent 1/32, 1/16, 1/8 and 1/4, respectively.
5. The mode information shall represent modes 1, 2, and 3 when its value is ‘00,’ ‘01’ and ‘10,’ respectively. ‘11’ shall be undefined.
6. The center frequency shall be a center frequency of the frequency band used to transmit a broadcast program. The value shall be expressed in 1/7 MHz units.
7. This descriptor shall be used only for digital terrestrial sound broadcasting, digital terrestrial television broadcasting, V-Low multimedia broadcasting based on connected segment system or V-High multimedia broadcasting based on connected segment system. The descriptor may be transmitted in the descriptor 2 field of NIT.

Figure No. 5: Satellite delivery system descriptor

Descriptor tag	Descriptor length	Freque ncy	Orbit	Longitude flag	Polari zation	Modula tion	Transmission rate (symbol rate)	FEC (inner code)
8	8	32	16	1	2	5	28	4

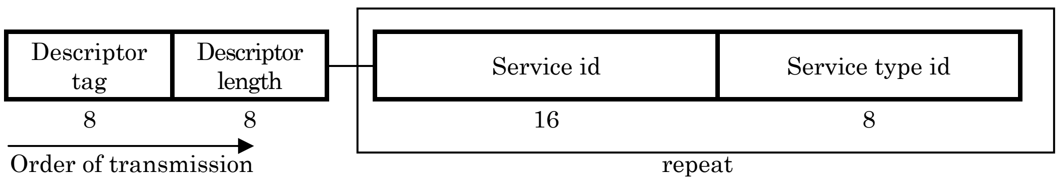
Order of transmission →

Notes:

1. The value of the descriptor tag shall be 0x43, representing the satellite delivery system descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The frequency shall be a field that writes a frequency (GHz). An 8-digit number, each digit of which consists of a 4-bit BCD code, shall be used to write a frequency. The lower four digits represent the fractional part.

4. The orbit shall be a field that writes an orbital position (degrees). A 4-digit number, each digit of which consists of a 4-bit BCD code, shall be used to write a position. The lower four digits represent the fractional part.
5. The longitude flag shall represent west and east longitude when its value is ‘0’ and ‘1,’ respectively.
6. The polarization is a field used for identifying the polarization type. It shall represent horizontally, vertically, left-handed and right-handed polarized waves when its value is ‘00,’ ‘01,’ ‘10’ and ’11, respectively.’
7. The modulation is a field used for identifying the modulation type. It shall represent 4-phase modulation, undefined, modulation for BS digital broadcasting and wide band CS digital broadcasting, modulation for advanced narrowband CS digital broadcasting and modulation for advanced BS digital broadcasting and advnaced wide band CS digital broadcasting, when its value is ‘00001’, ‘01001’, ‘01000’, ‘01010’ and ‘01011’, respectively.
8. The transmission rate is a field that writes symbols transmitted per second (Mbaud). A 7-digit number, each digit of which consists of a 4-bit BCD code, shall be used to write a speed. The lower four digits represent the fractional part.
9. The FEC is a field identifying the coding rate of inner code. It shall represent coding rates of 1/2, 2/3, 3/4, 5/6 and 7/8 when its value is ‘0001,’ ‘0010,’ ‘0011,’ ‘0100’, and ‘0101’, respectively. It also shall represent the coding rate of inner code for BS digital broadcasting and wide band CS digital broadcasting, for advanced narrowband CS digital broadcasting and for advanced BS digital broadcasting and advnaced wide band CS digital broadcasting, when its value is ‘1000’, ‘1010’ and ‘1011’, respectively. ‘1111’ indicates that there is no inner code.
10. This descriptor shall be used only for satellite digital audio broadcasting, BS digital or broadband CS digital broadcasting. The descriptor may be transmitted in the descriptor 2 field of NIT or TLV-NIT.

Figure No. 6: Service list descriptor



Notes:
1. The value of the descriptor tag shall be 0x41, representing the service list descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The service id shall be a field used for identifying the broadcast program number.
4. The service type id shall be a field used for identifying type of broadcasting. Service type assignment shall be shown below.

Value	Description
0x00	Undefined
0x01	Television broadcasting
0x02	Ultra-short wave broadcasting
0x03 – 0x7F	Undefined
0xC0	Data broadcasting
0xC1	Storage broadcasting using TLV
0xC2	Multimedia broadcasting
0xC3 – 0xFF	Undefined

5. The descriptor shall be transmitted in the descriptor 2 field of NIT or TLV-NIT.

Figure No. 7: System management descriptor

Descriptor tag	Descriptor length	System management id	Additional id information (extension of id number)
8	8	16	8 × N

Order of transmission →

Notes:

1. The value of the descriptor tag shall be 0xFE, representing the system management descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The system management id is used to identify the type (such as broadcasting or non-broadcasting). Its structure shall be as shown below.

System management identifier

Broadcasting or non-Broadcasting	Broadcasting signal standard	Detailed identification
2	6	8

Broadcasting or non-Broadcasting type

Value	Assignment
‘00’	Broadcasting
’01’, ’10’	Non-broadcasting
’11’	Undefined

Broadcasting standard type

Value	Assignment
'000000'	Undefied
'000001'	Narrowband CS digital broadcasting
'000010'	BS digital broadcasting
'000011'	Digital terrestrial television broadcasting
'000100'	Wide band CS digital broadcasting
'000101'	Digital terrestrial sound broadcasting
'000110'	Undefined
'000111'	Advanced narrowband CS digital broadcasting
'001000'	Advanced BS digital broadcasting
'001001'	Advanced wide band CS digital broadcasting
'001010'	V-High multimedia broadcasting based on connected segment transmission system
'001011'	V-Low multimedia broadcasting based on connected segment transmission system
'001100' – '111111'	Undefined

4. The additional identifier information shall be a field used for extending idetification number.
5. For digital terrestrial sound, digital terrestrial television, V-High multimedia and V-Low multimedia based on connected segment transmission system, BS digital, wide band CS digital, advanced BS digital or advanced wide band CS digital broadcasting transmitted by TS packets, this descriptor shall be transmitted in the descriptor 1 field of PMT or in the descriptor 1 or 2 field of NIT. If this descriptor is transmitted in plural fields, the priority is given to descriptor 1 of PMT, followed by descriptor 2 of NIT and then descriptor 1 of NIT.
6. For narrowband CS digital broadcasting, this descriptor shall be transmitted in the descriptor 1 field of PMT.
7. For advanced BS digital broadcasting and advanced wide band CS digital broadcasting transmitted by TS packets, this descriptor shall be transmitted in the descriptor 1 or 2 field of TLV-NIT. If this descriptor is transmitted in plural fields, the priority is given to descriptor 1 and then descriptor 2.

Figure No. 8: Data component descriptor

Descriptor tag	Descriptor length	Data coding sytem id	Additional id information (id number extension)
8	8	16	8 × N

Order of transmission →

Notes:

1. The value of the descriptor tag shall be 0xFD, representing the data coding system descriptor.

2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The data coding system id shall be a field used for identifying the data coding system standard.
4. The additional id information shall be a field that writes additional infromation on the data coding system standard shown in the data coding system id..
5. This descriptor may be transmitted in the descriptor 2 field of PMT.

Figure No. 9: Data structure of carousel compatible composite descriptor

Descriptor tag	Descriptor length	Subdescriptor
8	8	8×N
Order of transmission →		Repeat

Notes:
1. The value of the descriptor tag shall be 0xF7, representing the carousel compatible composite descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The subdescriptor shall be a filed to write information including the subdescriptors described in (1) to (3) below.

(1) Accumulation route subdescriptor

Subdescriptor tag	Subdescriptor length	Data	Directory name
8	8	8	8×N

Order of transmission →

a. The value of the subdescriptor tag shall be 0xC5, representing the accumulation route subdescriptor.
b. The subdescriptor length shall be a field that writes the number of data bytes following this field.
c. The directory name shall be a field that describes the name in text format of the uppermost directory in the directory structure used when accumulating programs in the receiving equipment.

(2) Subdirectory subdescriptor

Subdescriptor tag	Subdescriptor length	Subdirectory name
8	8	8×N

→

Order of transmission

a. The value of the subdescriptor tag shall be 0xC6, representing the subdirectory subdescriptor.
b. The subdescriptor length shall be a field that writes the number of data bytes following this field.
c. The subdirectory name shall be a field that describes a directory structure in text formatused for accumulating programs in the receiving equipment , excliding the structure specified by the accumulation route subdescriptor.

(3) Accumulation name subdescriptor

Subdescriptor tag	Subdescriptor length	Accumulation name
8	8	8×N

Order of transmission →

a. The value of the subdescriptor tag shall be 0x02, representing the accumulation name subdescriptor.
b. The subdescriptor length shall be a field that writes the number of data bytes following this field.
c. The accumulation name shall be a field that describes a name in text formatused for accumulating programs in the receiving equipment.

Figure No. 10: Copyright descriptor

Descriptor tag	Descriptor length	Copyright idr	Additional copyright information
8	8	32	8 × N

Order of transmission →

Notes:
1. The value of the descriptor tag shall be 0x0D, representing the copyright descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The copyright identifier shall be a field identifying the copyright.

Figure No. 11: Emergency information descriptor

Descriptor tag	Descriptor length	Service id	Start/end flag	Signal type	Undefined	Area code length	Area code	Undefined
8	8	16	1	1	6	8	12	4
		repeat					repeat	

Order of transmission →

Notes:

1. The value of the descriptor tag shall be 0xFC, representing the emergency information descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The service id shall be used to identify the broadcast program number.
4. The value of the start/end flag shall be '1' when transmission of emergency information signal starts or is in progress and '0' when the transmission ends.
5. The value of the signal type shall be '0' and '1,' respectively, when transmitting Class 1 and 2 start signals defined in Paragraph 1 of Article 138-2 of the Regulations for Operating Radio Station, (Radio Regulatory Commission Rule No. 17, 1950).
6. The area code length shall be a field that writes the number of data bytes following this field.
7. The area code shall be a field transmitting the area code defined in Table No. 1 in (Notification No. 405 (Defining the structure of emergency alarm signal by the provisions in Article 9-3, Radio Equipment Regulations) of the Ministry of Posts and Telecommunications 1985.
8. This descriptor may be used only for digital terrestrial sound, digital terrestrial television, mutimedia, BS digital, wide band CS digital, advanced BS digital or advanced wide band CS digital broadcasting transmitted by TS packets. The descriptor may be tranmsitted in the descriptor 1 field of PMT or in the descriptor 1 or 2 field of NIT.

Figure No. 12: IP/MAC stream arrangement descriptor

INT descriptor tag	Descriptor length	Network id	Original network id	Transport stream id	Service id	Component tag
8	8	16	16	16	16	8

Order of transmission →

Notes:
1. The value of INT descriptor tag shall be set to 0x13 representing IP/MAC stream arrangement descriptor.
2. The descriptor length shall be a field that writes the number of data bytes after this field.
3. The network id shall be a field used for identifying the network number.
4. The original network id shall represent an identification number of the original network for the relevant transport stream.
5. The transport stream id shall represent the identification number of transport stream.
6. The service id shall be a field used for identifying the broadcast program number.
7. The component tag shall be a field that writes the tag value of the stream transmitting the object ULE packet.
8. This descriptor may be transmitted in the descriptor 1 or descriptor 3 field of INT.

Figure No.13: Structure of access control descriptor

Descriptor tag	Descriptor length	Conditional access system id	Transmission information	PID	Payload
8	8	16	3	13	8xN

Order of transmission →

Notes:

1. The value of the descriptor tag shall be set to 0xF6 representing access control descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. Conditional access system for reception id shall be a field used for identifying the type of conditional access system.
4. Transmission information shall be a field used for identifying the relevant information transmission.
5. PID shall be a field that writes PID of TS packet including the related information.
6. This descriptor may be transmitted in the descriptor field of CAT, or in the descriptor 1 or descriptor 2 field of PMT.

Figure No14: Structure of hierarchical coding descriptor

Descriptor tag	Descriptor length	Undefin ed	Time directional hierarchical coding flag	Space directional hierarchical coding flag	Image quality hierarchical coding flag	Hierachical coding id
8	8	1	1	1	1	4

Undefin ed	Hierachical index	TREF presence flag	Undefin ed	Basic hierachical index	Undefin ed	Hierarchical channel
2	6	1	1	6	2	6

Notes:

1. The value of the descriptor tag shall be 0x04, representing hierachical coding descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The time directional hierarchical coding flag shall be ‘0’ when frame rates of program elements bit stream are improved by the basic hierarchical index and ‘1’ when the frame rates are not improved.
4. The space directional hierarchical coding flag shall be ‘0’ when spatial resolution of program elements bit stream is improved by the basic hierarchical index and ‘1’ when the resolution is not improved
5. The image quality hierarchical coding flag shall be ‘0’ when SNR quality or fidelity of program elements bit stream is improved by the basic hierarchical index and ‘1’ when the quality or fidelity is not improved.

6. The hierarchical coding identifier is a field used for identifying the hierarchical coding type that has been used and its assignment shall be shown in the following table.

Value	Assignment
0	Undefined
1	Space directional hierarchical coding
2	Image quality hierarchical coding
3	Time directional hierarchical coding
4 - 7	Undefined
8	Mixed hierarchical coding
9 - 14	Undefined
15	Video stream specified in ITU-T Rec. H.265 or time directional video sub bit stream specified in ITU-T Rec.265

7. The hierarchical index shall be a unique value used for the relevant program in the coding hierarchical table.
8. TREF existing flag shall be '0' and '1', respectively, when there is TREF field or not in the PES packet header of the relevant elementary stream.
9. The basic hierarchical index shall be a hierarchical program elements index.
10. The hierarchical channel shall represent an object channel number for the related program element among a series of transmission channels having ranking order.
11. This descriptor shall be transmitted in the descriptor 2 field of PMT.

Figure No. 15: Structure of scrambling method descriptor

Descriptor tag	Descriptor length	Scrambling method id	Payload
8	8	8	8×N

Order of transmission →

Notes:
1. The descriptor tag value shall be set to 0xF5 indicating the scrambling method description.
2. The descriptor length shall be a field that writes the number of data bytes after this field.
3. The scrambling method id shall be a field used for identifying encryption algorithm in the case of scrambling. The assignment shall be shown in the following table.

Value	Assignment
'00000000'	Undefined
'00000001'	AES (limited to a key length of 128 bits)
'00000010'	Camellia (limited to a key length of 128 bits)
'00000011'-'11111111'	Undefined

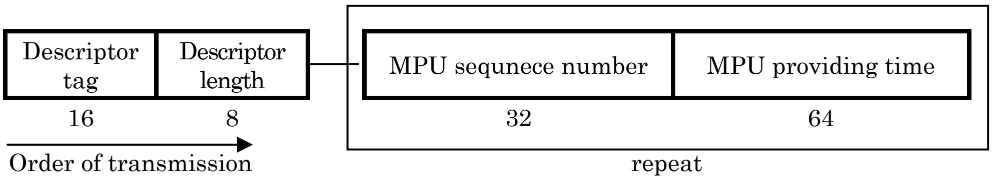
4. This descriptor may be transmitted in the descriptor field of CAT, or the descriptor 1 or descriptor 2 field of PMT.

(Table No.21, Notification)

3.11.2 Transmission by MMTP packet

Descriptor	Configuration
MPU time stamp descriptor	As shown in No.1 below
Dependency descriptor	As shown in No.2 below
Access control descriptor	As shown in No.3 below
Scrambling method descriptor	As shown in No.4 below
Emergency information descriptor	As shown in No.5 below

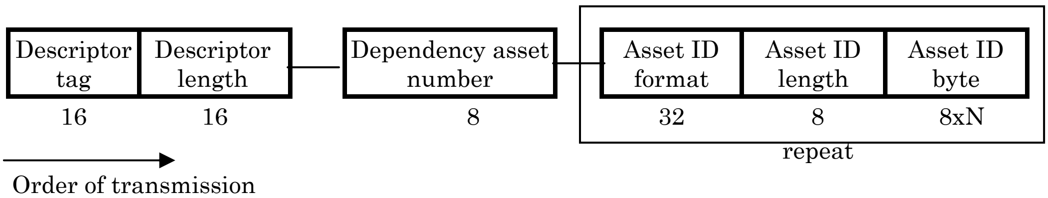
No. 1: Structure of MPU time stamp descriptor



Notes:

1. The value of descriptor tag shall be 0x0001 representing MPU time stamp descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The MPU sequence number shall be a field that writes the sequence number of MPU describing time stamp.
4. The MPU providing time shall be a 64 bits NTP timestamp defined in IETF RFC 5905.
5. This descriptor may be transmitted in the field of asset descriptor in the MPU table.

No. 2: Structure of dependency relation descriptor



Notes:

1. The value of descriptor tag shall be 0x0002 representing dependency relation descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3 The dependency asset shall represent the asset in which this descriptor is inserted and the number of asset.
4. The asset ID format shall represent the asset ID format for the complementary asset.
5. The asset ID length shall be a field that writes the data bytes of asset ID byte for the complementary asset.
6. The asset ID byte shall represent the asset ID for complementary asset.
7. This descriptor may be transmitted in the field of asset descriptor in MPU table.

No. 3: Structure of access control descriptor

Descriptor tag	Descriptor length	Conditional access system id	Location information	Payload
16	8	16	8xN	8xN

→ Order of transmission

Notes:

1. The value of descriptor tag shall be 0x8004 representing access control descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The conditional access system id shall represent the type of conditional access system.
4. The location information shall represent the location of MMTP packet including the related information.
5. This descriptor may be transmitted in CA table descriptor field in the CA message, MP table descriptor field in MP table or asset descriptor field.

No. 4: Structure of access scrambling method descriptor

Descriptor tag	Descriptor length	Target layer id	'111111'	Scrambling methods id	Payload
16	8	2	6	8xN	8xN

→ Order of transmission

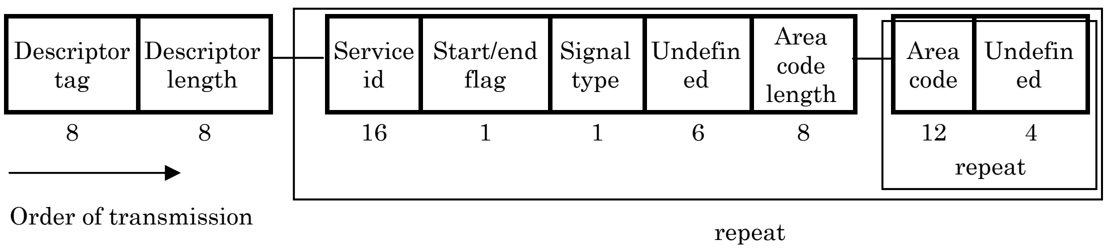
Notes:

1. The value of the descriptor tag shall be 0x8005 representing scrambling method descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3 The target layer id shall represent the encryption target at the time of scrambling (IP packet and MMTP packet).
4. The scrambling method id represents an encryption algorithm type at the time of scrambling and its assignment shall be shown in the following table.

Value	Assignment
'00000000'	Undefined
'00000001'	AES (limited to 128 bit key length)
'00000010'	Camellia (limited to 128 bit key length)
'00000011'-'11111111'	Undefined

5. This descriptor may be transmitted in CA table descriptor field in the CA message, MP table descriptor field in MP table or asset descriptor field.

No. 5: Structure of emergency information descriptor



Notes:

1. The value of descriptor tag shall be 0x8007 representing emergency information descriptor.
2. The descriptor length shall be a field that writes the number of data bytes following this field.
3. The service id shall be used for identifying the number of broadcast program.
4. The value of the start/end flag shall be ‘1’ when transmission of emergency information signal starts or is in progress and ‘0’ when the transmission ends.
5. The value of the signal type shall be ‘0’ and ‘1,’ respectively, when transmitting Class 1 and 2 start signals defined in Paragraph 1 of Article 138-2 of the Regulations for Operating Radio Station, Radio Regulatory Commision Rule No.17, 1950).
6. The area code length shall be a field that writes the number of data bytes following this field.
7. The area code shall be a field that transmits the area code defined in the Appended table No.1, Notification No. 405 of the Ministry of Post and Telecommunications, 1985 (Defining the configuration of emergency alarm signal pursuant to the provisions of Article 9-3, Radio Equipment Regulations).
8. This descriptor may be transmitted in the MP table descriptor field in MP table.

(Table No.30, Notification)

3.12 Identifiers

3.12.1 Transmission by TS packet or TLV packet

Identifier	Configuration
Table id	As shown in Section 3.10 (1) to (7) and Section 3.13
Descriptor tag	As per Section 3.11.1 and ITU-T Rec. H.222.0
Stream type id	As shown in Section 3.10 (2)
Service type id	As shown in Section 3.11.1
Broadcast program number id	As shown in Section 3.10 (1)
Service id	As shown in Section 3.11.1
Network id	As shown in Section 3. 10 (4) and (5)
Transport stream id	As shown in Section 3.10(4) and (5)
CA system id	As shown in Section 3.11.1
System management id	As shown in Section 3.11.1
Hierarchical coding id	As shown in Section 3.11.1
Scrambling method id	As shown in Section 3.11.1

(Table No.22, Notification)

3.12.2 Transmission by MMTP packet

Identifier	Configuration
Descriptor tag	As per Section 3.11.2 and ISO/IEC 23008-1
Conditional access system id	As shown in Section 3.11.2
Scramble system id	As shown in Section 3.11.2
Service id	As shown in Section 3.11.2

(Table No.31, Notofication)

3.13 Structure and transmission procedure of conditional access related information

1. Among conditional access common information, ECM, whose scope of scrambling is the TS packet (excluding that for sending transmission control signal and conditional access related infromation) payload in the standard transmission system of digital broadcasting (hereinafter referred to as “standard system”) among standard television broadcasting and the like, shall contain program information, key information for de-scrambling, and control information instructing a forced switching of the receiver's de-scrambling function. The structure and transmission procedure of ECM shall be as shown in Table No. 1.
2. Among conditional access common information, ACI, whose scope of scrambling is limited to the section format signals in the standard system, shall contain program information, key information for de-scrambling, and control information which

instruct a forced switching of the receiver's de-scrambling function. ACI shall include a protocol number showing the ACI structure, an entity id to identify the entity who performs scrambling, and an encryption key id to identify the encryption key used for encrypting the information contained in ACI and be transmitted as modules defined in Paragraph 2, Notification No. 301 of the Ministry of Internal Affairs and Communications (Defining transmission procedure for video signal and audio signal by section format.

3. The conditional access individual information (hereinafter referred to as "EMM") shall contain domestic subscribers' specific contract information and key information for decrypting ECM. The structure and transmission procedure of EMM shall be as shown in Table No. 2.

Table No. 1: Structure and transmission procedure of ECM

Header	ECM	CRC
64	8 × N	32

Order of transmission →

Notes:

1. Each number without a unit shall represent the number of bits for that field. The same shall apply hereinafter.
2. Numbers following "0x" shall represent hexadecimal numbers. The same shall apply hereinafter.
3. Each field shall be transmitted from MSB (most significant bit) to LSB (least significant bit). The same shall apply hereinafter.
4. ECM shall be transmitted in the extended section format given in Section 3.2.
5. The value of the "table id" within the header shall be 0x82 or 0x83, representing the ECM. The "table id extension" shall be a field to identify type of information contained in ECM.
6. ECM shall consist of information including those listed below. Howevev, note that for advanced BS broadcasting and advanced wide band CS digital broadcasting, Protocol number, Entity id, and Data and time may not be contained In addition note that for narrowband CS digital broadcasting Entity id may not be contained. Information other than the protocol number, entity id, and encryption key id can be encrypted using the key identified by the encryption key id.

Items
Protocol number
Entity id
Encryption key id
De-scrambling key
Date and time

Table No. 2: Structure and transmission procedure of EMM

Header	EMM1	EMM2	…	EMMn	CRC
64		8 × N			32

Order of transmission →

Notes:

1. EMM shall be transmitted in the extended section format shown in Section 3.2. Multiple EMMs may be multiplexed within that extent mentioned above.
2. The value of the “table id” within the header shall be 0x84 or 0x85, representing the EMM. The “table id extension” shall be a field used to identify type of information contained in EMM.
3. For digital terrestrial sound, digital terrestrial television, multimedia, BS digital, advanced BS dgital, wide band CS digital, or advanced CS digital broadcasting, the value of the “table id” within the header shall be 0x85 when information to send message information to the receiver (referred to as an “EMM message”) is contained in EMM,. The value of the “table identifier extension” shall be 0x0000 and 0x0001 through 0xFFFF respectively when EMM message is transmitted to specific receivers and to all receivers notwithstanding Note.2..
4. EMM shall consist of EMM messages or information including those listed below. Note that decoder id number shall be id number and protocol number may not be contained for advanced BS digital broadcasting and advanced wide band CS digital broadcasting.
   In addition the following information can be encrypted:
   - Information other than the protocol number for broadcasting except for advanced BS digital broadcasting and advanced wide band CS digital broadcasting
   - Information other than id number for advanced BS digital broadcasting and advanced wide band CS digital broadcasting.

Items
Decoder id
Protocol number

(Notification)

Annex A: Technical methods applied to digital broadcasting

Table A-1 and Table A-2 show technical methods applied to the digital broadcasting standard systems defined in the Ordinance (Ordinance of the Ministry of Internal Affairs and Communications No.87, 2011 or No.94, 2011)

Table A-1 Technical methods applied to standard systems (Coded signals transmission) (○: Applied)

	Digital broadcasting	Digital terrestrial sound broadcasting	Digital terrestrial television broadcasting	V-Low multimedia broadcasting (Note 1)	V-High multimedia broadcasting (Note 1)	BS digital broadcasting	Advanced BS digital broadcasting		Narrowband CS digital broadcasting	Advanced narrowband CS digital broadcasting	Wide band CS digital broadcasting	Advanced CS digital broadcasting	
							TS	TLV				TS	TLV
Coded signals transmission	PES packet	○	○	○	○	○	○		○	○	○	○	
	Section	○	○	○	○	○	○		○	○	○	○	
	TS packet	○	○	○	○	○	○		○	○	○	○	
	IP packet			○	○			○					○
	HCfB			○ (Note 2)	○ (Note 2)			○					○
	ROHC			○ (Note 2)	○ (Note 2)								
	ULE packet			○	○								
	TLV packet							○					○
	MMTP packet							○					○

(Note 1) Based on connected segment transmission system

(Note 2) Not specified in the Ministerial Ordinance and Notification

Table A-2 Technical methods applied to standard systems (Transmission control signal and Emergency alarm signal) (○: Applied)

	Digital broadcasting	Digital terrestrial sound broadcasting	Digital terrestrial television broadcasting	V-Low multimedia broadcasting (Note 1)	V-High multimedia broadcasting (Note 1)	BS digital broadcasting	Advanced BS digital broadcasting		Narrowband CS digital broadcasting	Advanced narrowband CS digital broadcasting	Wide band CS digital broadcasting	Advanced CS digital broadcasting	
							TS	TLV				TS	TLV
Transmission control signal	PAT	○	○	○	○	○	○		○	○	○	○	
	PMT	○	○	○	○	○	○		○	○	○	○	
	CAT	○	○	○	○	○	○		○	○	○	○	
	NIT	○	○	○	○	○	○		○	○	○	○	
	TLV-NIT							○					○
	AMT			○				○					○
	INT				○								
	PA message							○					○
	M2 section message							○					○
	CA message							○					○
Emergency alarm signal		○	○	○	○	○	○	○	○	○	○	○	○

(Note 1) Based on connected segment transmission system

VIDEO CODING, AUDIO CODING, AND
MULTIPLEXING SPECIFICATIONS FOR
DIGITAL BROADCASTING

ARIB STANDARD

ARIB STD-B32 VERSION 3.9-E1
(December 2016)

This Document is based on the ARIB standard of "Video Coding, Audio Coding, and Multiplexing Specifications for Digital Broadcasting" in Japanese edition and translated into English in August 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