ENGLISH TRANSLATION

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

ARIB STANDARD

ARIB STD-B32 Version 3.11
(Fascicle 2)

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
Revised April 12, 2018	Version 3.10
Revised July 26, 2018	Version 3.11

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)

<table>
<tr><th>Patent Applicant/ Holder</th><th>Name of Patent</th><th>Registration No./ Application No.</th><th>Remarks</th></tr>
<tr><td rowspan="2">Japan Broadcasting Corporation (NHK)</td><td>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</td><td>特願平 05-65183
特開平 06-276169</td><td>Japan</td></tr>
<tr><td colspan="2">Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.1.*[15]</td><td></td></tr>
<tr><td rowspan="8">NEC Corporation</td><td>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</td><td>特許 1890887</td><td>Japan</td></tr>
<tr><td>画像の圧縮記録システム</td><td>特許 2036887</td><td>Japan,
United States,
United Kingdom,
Germany, France,
Netherlands,
Canada</td></tr>
<tr><td>適応変換符号化の方法及び装置</td><td>特許 2569842</td><td>Japan,
United States,
United Kingdom,
Germany, France,
Netherlands</td></tr>
<tr><td>適応変換符号化の方法及び装置</td><td>特許 2778161</td><td>Japan,
United States,
United Kingdom,
Germany, France,
Netherlands</td></tr>
<tr><td>適応変換符号化の方法及び装置</td><td>特許 2569849</td><td>Japan,
United States,
United Kingdom,
Germany, France,
Netherlands</td></tr>
<tr><td>適応変換符号化復号化の方法及び装置</td><td>特許 2638208</td><td>Japan,
United States,
United Kingdom,
Germany, France,
Netherlands</td></tr>
<tr><td>符号化方式及び復号方式</td><td>特許 2820096</td><td>Japan, Korea,
Australia</td></tr>
<tr><td>改良 DCT の順変換計算装置および逆変換計算装置</td><td>特許 3185214</td><td>Japan,
United States,
United Kingdom,
Germany, France,
Netherlands,
Canada</td></tr>
</table>

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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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		
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		
	Submitted comprehensive confirmation of patents for ARIB STD-B32 Ver3.0.*21		

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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]		
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
	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

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
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
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	

Patent Applicant/ Holder	Name of Patent	Registration No./ Application No.	Remarks
	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)
*21 : Valid for the revised parts of ARIB STD-B32 Ver3.0 (received on April 26, 2018)

TOTAL CONTENTS

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).

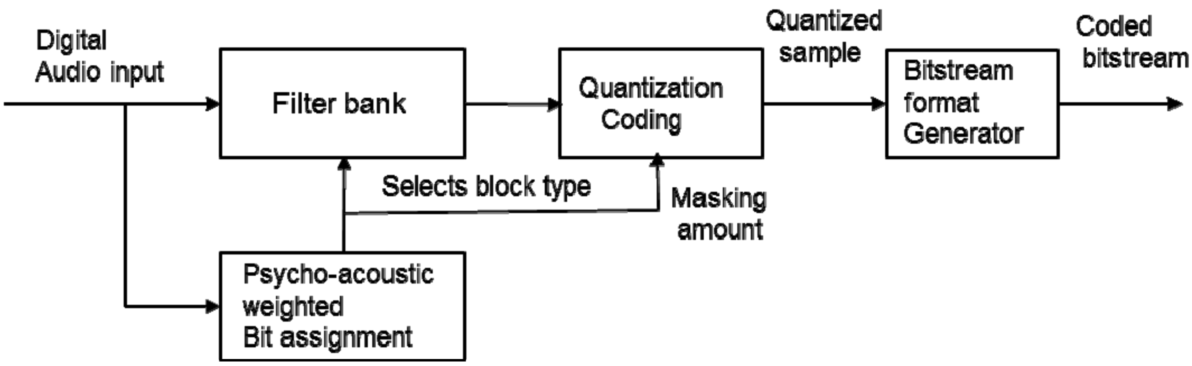
(Ordinance Article 64, Article 81-3)

<Blank Page>

Chapter 4: Audio Compression and Transmission Procedures

4.1 System based on MPEG-2 AAC Standard

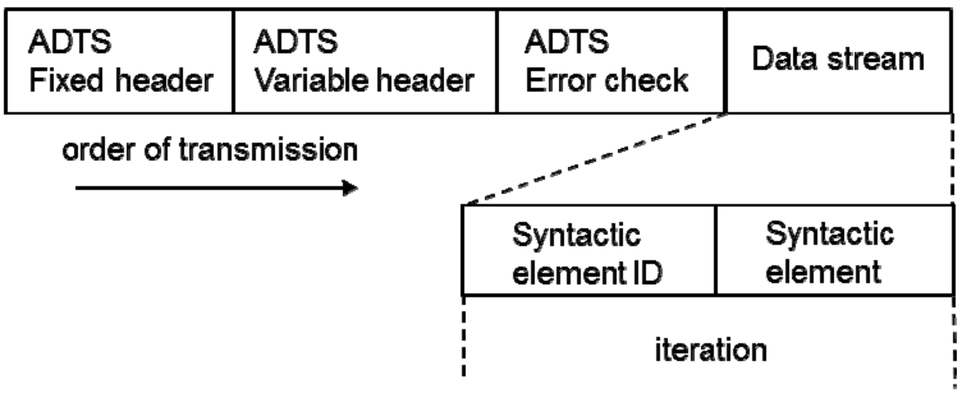
Audio compression and transmission procedures shall be as specified in the following.



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:

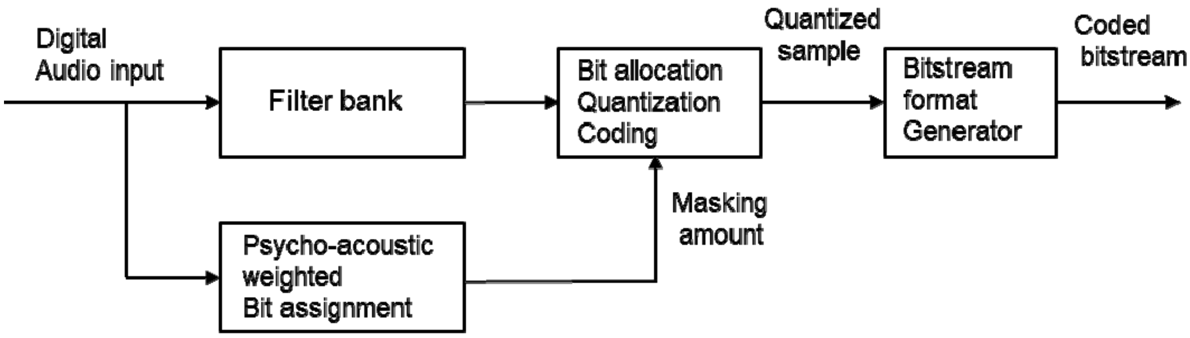
1. The ADTS fixed header consists of synchronization and audio coded information defined in ISO/IEC 13818-7.
   But for multimedia broadcasting, it consists of synchronization and audio coded information defined in ISO/IEC 13818-7, ISO/IEC 23003-1, ISO/IEC 14496-3, ISO/IEC 14496-3:2001/Amd.1, and ISO/IEC 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.

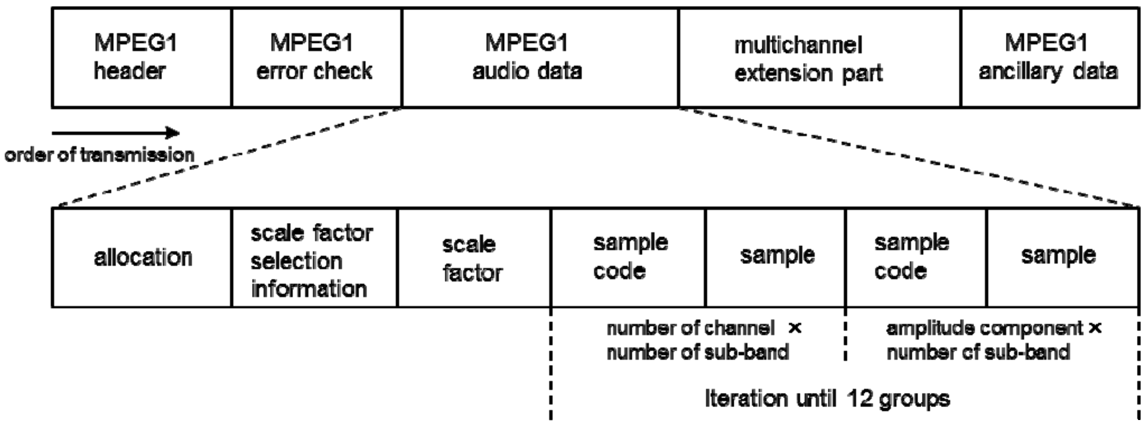
2 Psycho-acoustic weighted bit assignment calculates masking amount for each band of the filter bank.

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.

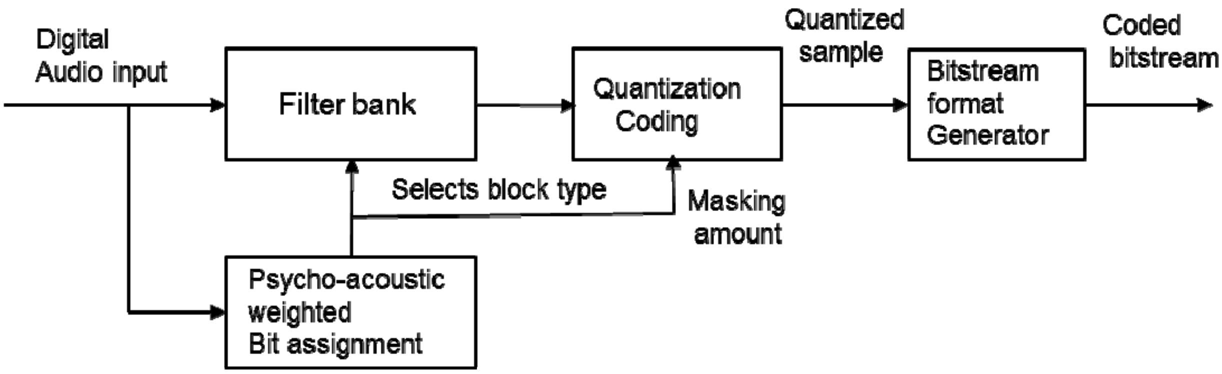
Appendix No.2 Channel mode

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.

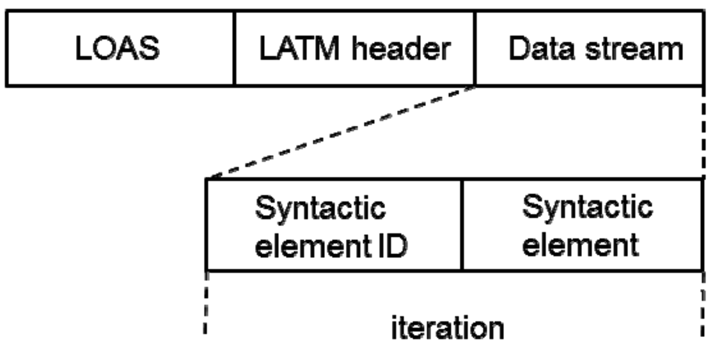
2 Psychoacoustic model calculates masking amount (limits of differentiating a specific audio signal from other audio signals) and block type input to the filter bank.

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.

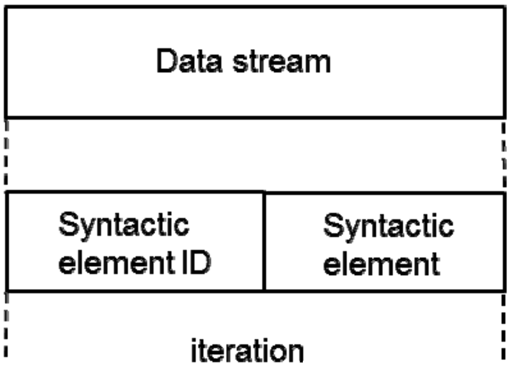
3 Data stream shall consist of audio data coded according to ISO/IEC 14496-3.

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.

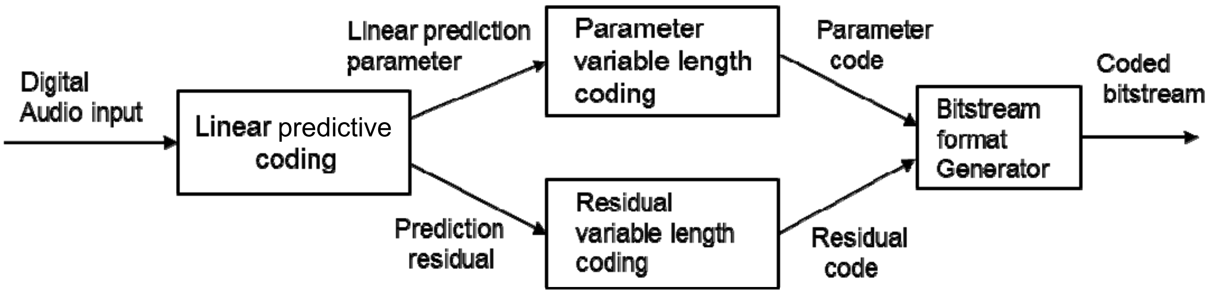
3 The syntactic element consists of each component of audio data coded according to ISO/IEC 14496-3.

(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 $0\text{dB} \leqq x \leqq +12\text{dB}$ 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 $-\infty\text{dB} \leqq -x \leqq 0\text{dB}$ 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.

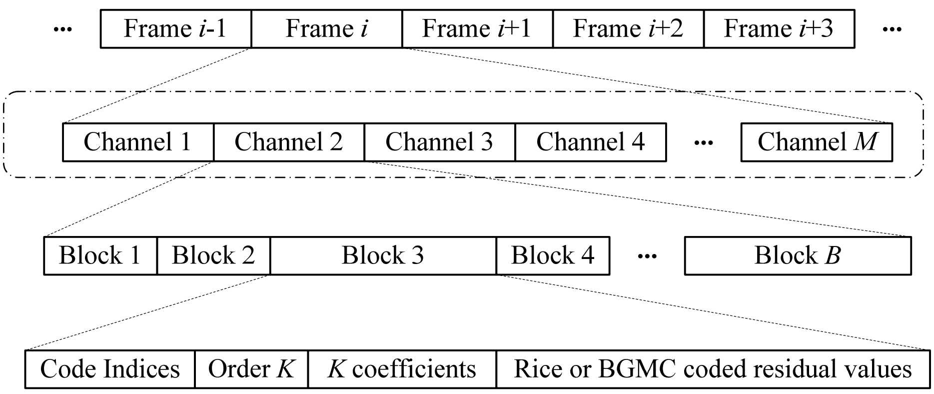
Raw Data Stream

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()

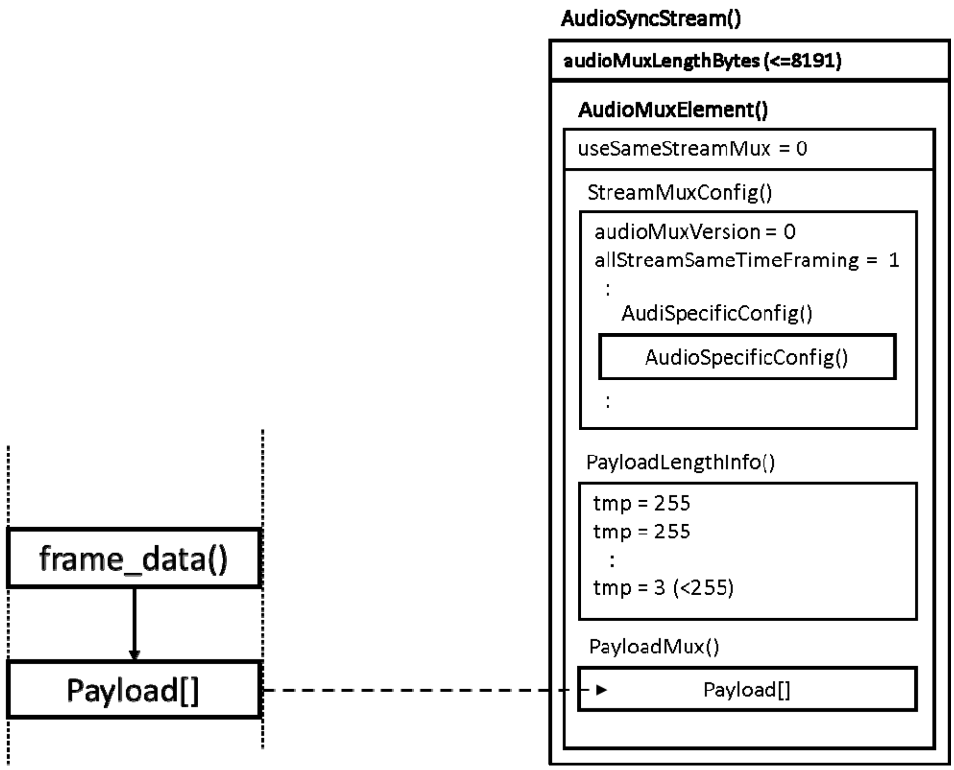
Item	Restriction
tmp	The value which is added to MuxSlotLengthBytes[0]
MuxSlotLengthBytes[0]	Size of PayloadMux()
MuxSlotLengthCoded[]	Not used
numChunk	0 (number of chunks =1) or not used.
streamIndx	0 or not used.
AuEndFlag[]	0, 1, or not used.
MuxSlotLengthCoded[]	Not used

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

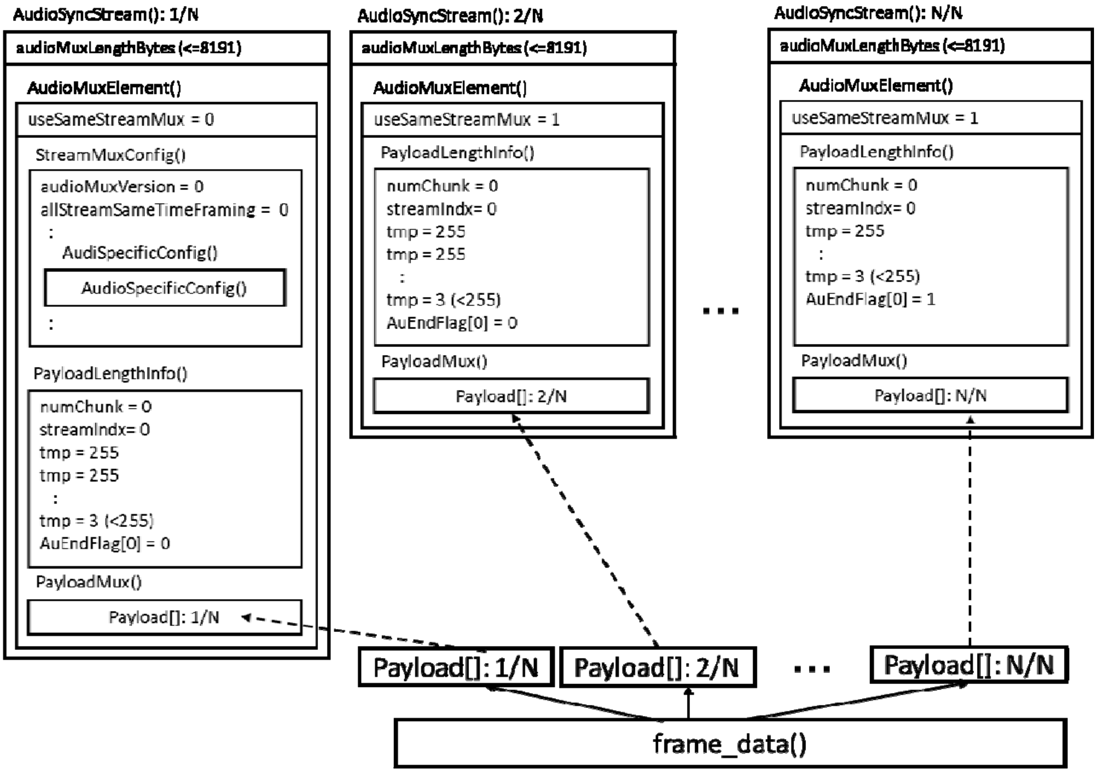
In case that the size of AudioMuxElement() is less than or equal to 8191 bytes on account of the size of Raw Data Stream (=frame_data()), LATM/LOAS frame is configured by making one frame_data() correspond to one AudioSyncStream() as the following figure.



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.

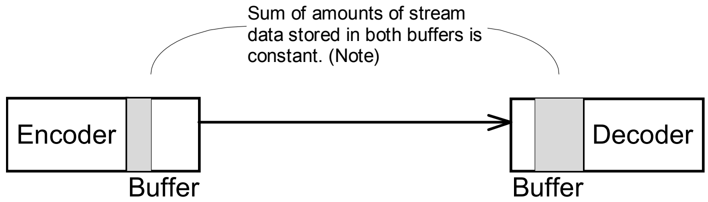
(b) The decoder halts decoding and mutes the audio when no audio stream data is found in the decoder input buffer. If audio stream data remains in the decoder input buffer and if the ADTS (Audio_Data_Transport_Stream) frame header is found, the decoder waits until the amount of stream data specified 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.

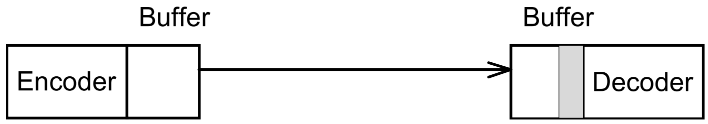
To facilitate the comprehension of audio parameter change sequence, Figs. 2-1 and 2-2, respectively, show the flow diagram and the timing diagram for switching audio parameters.

1. Steady-state condition

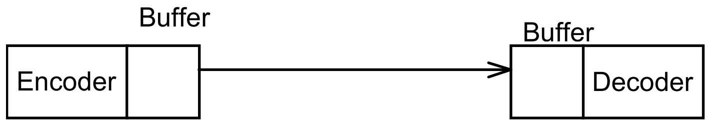


(Note) The sum of stream data stored in buffers is set at 6144 bits/channel or less.

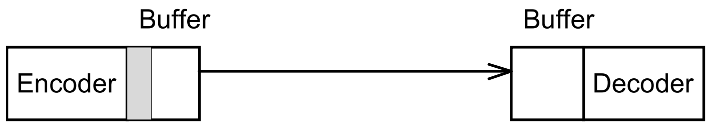
2. The encoder stops encoding.
-> The amount of stream data stored in the encoder and decoder buffers decreases.



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.

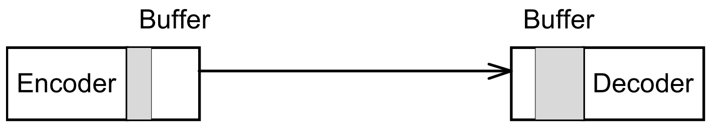


Fig. 2-1: Flow diagram for switching audio parameters

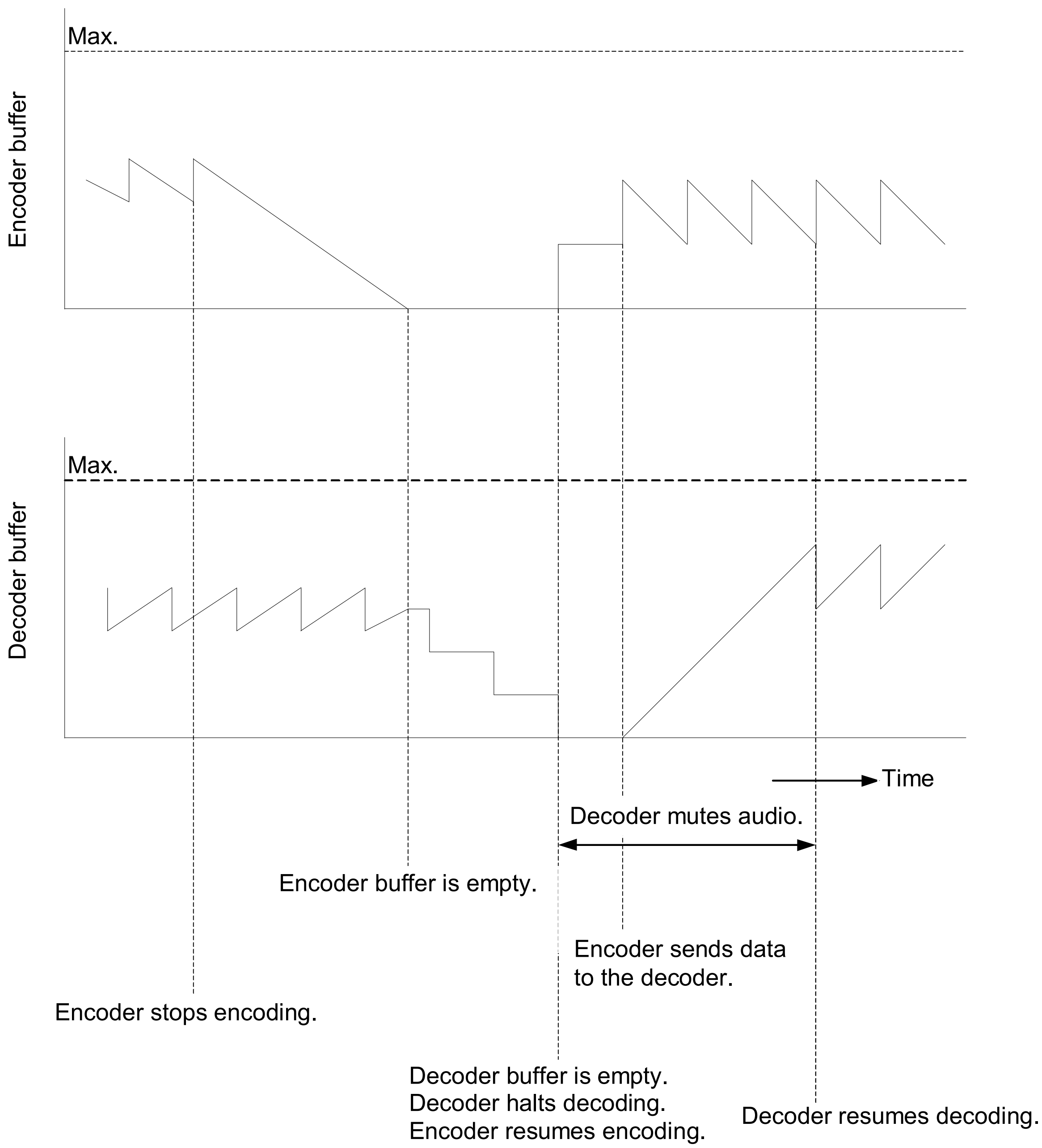


Fig. 2-2: Timing diagram for changing audio parameters

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>

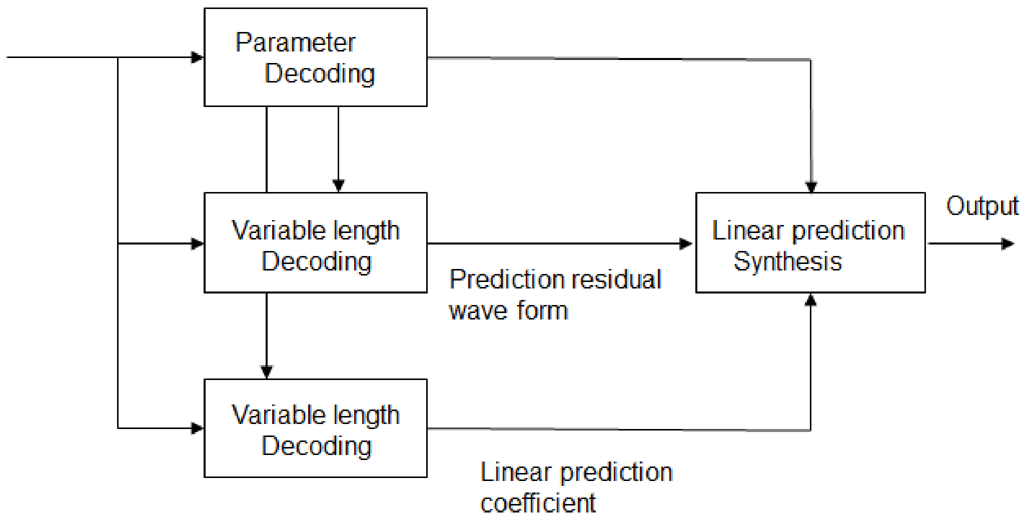


Fig. 1 Configuration of MPEG-4 ALS decoder

- 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.

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

ARIB STANDARD

ARIB STD-B32　VERSION 3.11-E1 (Fascicle2)
(July 2018)

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 November 2018.

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