
Information technology — Coding of
audio-visual objects —

Part 3:
Audio

AMENDMENT 1: Bandwidth extension

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Technologies de l'information — Codage des objets audiovisuels —

Partie 3: Codage audio

ISO/IEC 14496-3:2001/Amd.1:2003

<https://standards.iteh.ai/catalog/standards/sist/8761741e91c1b178472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003>
AMENDEMENT 1: Extension de largeur de bande

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC 14496-3:2001/Amd 1:2003](https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003>

© ISO/IEC 2003

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Amendment 1 to ISO/IEC 14496-3:2001 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

(standards.iteh.ai)

[ISO/IEC 14496-3:2001/Amd 1:2003](https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003>

Introduction

This document specifies the first Amendment to the ISO/IEC 14496-3:2001 standard. The document specifies the normative syntax of the SBR tool and the decoding process. An informative encoder description is given as well. Furthermore, this document specifies two new profiles, one based on the AAC LC Audio Object Type and one based on AAC in combination with SBR.

iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC 14496-3:2001/Amd 1:2003](https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003>

Information technology — Coding of audio-visual objects —

Part 3: Audio

AMENDMENT 1: Bandwidth extension

In ISO/IEC 14496-3:2001, Introduction, MPEG-4 general audio coding tools, add:

“

MPEG-4 SBR, (Spectral Band Replication) is a bandwidth extension tool used in combination with the AAC general audio codec. When integrated into the MPEG AAC codec, a significant improvement of the performance is available, which can be used to lower the bitrate or improve the audio quality. This is achieved by replicating the highband, i.e. the high frequency part of the spectrum. A small amount of data representing a parametric description of the highband is encoded and used in the decoding process. The data rate is by far below the data rate required when using conventional AAC coding of the highband.

“

(standards.iteh.ai)

[ISO/IEC 14496-3:2001/Amd 1:2003](https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003)

Amendment Subpart 1

standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003

In Part 3: Audio, Subpart 1, in subclause 1.3 Terms and Definitions, add:

“

206. **SBR**: Spectral Band Replication.

”

and increase the index-number of subsequent entries.

In Part 3: Audio, Subpart 1, in subclause 1.5.1.1 Audio object type definition, replace table 1.1 with the following table:

Table 1.1 – Audio object definition

Tools/ Modules	gain control	block switching	window shapes - standard	window shapes – AAC LD	filterbank - standard	filterbank – SSR	TNS	LTP	intensity	coupling	MPEG-2 prediction	PNS	MS	SIAQ	FSS	upsampling filter tool	quantisation&coding - AAC	quantisation&coding - TwinVQ	quantisation&coding - BSAC	AAC ER Tools	ER payload syntax	EP Tool 1)	CELP	Silence Compression	HVXC	HVXC 4kbs VR	SA tools	SASBF	MIDI	HILN	TTSI	SBR	Remark	Object Type ID	
Null																																	0		
AAC main		X	X	X	X	X	X	X	X	X	X	X	X				X																2)	1	
AAC LC		X	X	X	X	X	X	X	X	X	X	X	X				X																	2	
AAC SSR	X	X	X		X	X	X	X	X	X	X	X	X				X																	3	
AAC LTP		X	X	X	X	X	X	X	X	X	X	X	X				X																2)	4	
SBR																																X	5		
AAC Scalable		X	X	X	X	X	X	X				X	X	X	X	X	X																6)	6	
TwinVQ		X	X	X	X	X							X					X																7	
CELP																							X											8	
HVXC																								X										9	
(Reserved)																																		10	
(Reserved)																																			11
TTSI																															X			12	
Main synthetic																											X	X	X				3)	13	
Wavetable synthesis																											X	X					4)	14	
General MIDI																													X					15	
Algorithmic Synthesis and Audio FX																										X								16	
ER AAC LC		X	X	X	X	X	X	X				X	X				X				X	X	X											17	
(Reserved)																																			18
ER AAC LTP		X	X	X	X	X	X	X				X	X				X				X	X	X											5)	19
ER AAC scalable		X	X	X	X	X	X	X				X	X	X	X	X	X				X	X	X											6)	20
ER TwinVQ		X	X	X	X	X							X				X				X	X												21	
ER BSAC		X	X	X	X	X	X	X				X	X					X			X	X													22
ER AAC LD			X	X	X	X	X	X				X	X				X				X	X	X												23
ER CELP																					X	X	X	X										24	
ER HVXC																								X	X	X	X								25
ER HILN																								X	X						X				26
ER Parametric																							X	X	X	X				X					27
(Reserved)																																			28
(Reserved)																																			29
(Reserved)																																			30
(Reserved)																																			31

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC 14496-3:2001/Amd.1:2003
<https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-e911715e8e7/iso-iec-14496-3-2001-amd-1-2003>

In Part 3: Audio, Subpart 1, in subclause 1.5.1.2 Description, add after 1.5.1.2.5, add:

“

1.5.1.2.6 SBR-Object

The SBR Object contains the SBR-Tool and can be combined with the audio object types indicated in Table 1.2A

Table 1.2A – Audio object types that can be combined with the SBR Tool

Audio Object Type	Combination with SBR Tool permitted	Object Type ID
Null		0
AAC main	X	1
AAC LC	X	2
AAC SSR	X	3
AAC LTP	X	4
SBR		5
AAC Scalable	X	6
TwinVQ		7
CELP		8
HVXC		9
(Reserved)		10
(Reserved)		11
TTSI		12
Main synthetic		13
Wavetable synthesis		14
General MIDI		15
Algorithmic Synthesis and Audio FX		16
ER AAC LC	X	17
(Reserved)		18
ER AAC LTP	X	19
ER AAC scalable	X	20
ER TwinVQ		21
ER BSAC		22
ER AAC LD		23
ER CELP		24
ER HVXC		25
ER HILN		26
ER Parametric		27
(Reserved)		28
(Reserved)		29
(Reserved)		30
(Reserved)		31

”

In Part 3: Audio, Subpart 1, subclause 1.5.2.1 (Profiles), replace:

“

Eight Audio Profiles have been defined:

”

with

“

Ten Audio Profiles have been defined:

”

and add after item 8:

“

9. The **AAC Profile** contains the audio object type 2 (AAC-LC).
10. The **High Efficiency AAC Profile** contains the audio object types 5 (SBR) and 2 (AAC LC) The High Efficiency AAC Profile is a superset of the AAC Profile

”

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC 14496-3:2001/Amd 1:2003](https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003>

In Part 3: Audio, Subpart 1, replace Table 1.2 (Audio Profiles definition) with the following table:

“

Table 1.2 – Audio Profiles definition

Audio Object Type	Main Audio Profile	Scalable Audio Profile	Speech Audio Profile	Synthetic Audio Profile	High Quality Audio Profile	Low Delay Audio Profile	Natural Audio Profile	Mobile Audio Inter-networking Profile	AAC Profile	High Efficiency AAC Profile	Object Type ID
Null											0
AAC main	X						X				1
AAC LC	X	X			X		X		X	X	2
AAC SSR	X						X				3
AAC LTP	X	X			X		X				4
SBR										X	5
AAC Scalable	X	X			X		X				6
TwinVQ	X	X					X				7
CELP	X	X	X		X	X	X				8
HVXC	X	X	X			X	X				9
(reserved)											10
(reserved)											11
TTSI	X	X	X	X		X	X				12
Main synthetic	X			X							13
Wavetable synthesis											14
General MIDI											15
Algorithmic Synthesis and Audio FX											16
ER AAC LC					X		X	X			17
(reserved)											18
ER AAC LTP					X		X				19
ER AAC Scalable					X		X	X			20
ER TwinVQ							X	X			21
ER BSAC							X	X			22
ER AAC LD						X	X	X			23
ER CELP					X	X	X				24
ER HVXC						X	X				25
ER HILN							X				26
ER Parametric							X				27
(reserved)											28
(reserved)											29
(reserved)											30
(reserved)											31

”

In Part 3: Audio, Subpart 1, subclause 1.5.2.2 (Complexity units), replace table 1.3 by the table below:

Table 1.3 – Complexity of Audio Object Types and SR conversion

Object Type	Parameters	PCU (MOPS)	RCU	Remarks
AAC Main	fs = 48 kHz	5	5	1)
AAC LC	fs = 48 kHz	3	3	1)
AAC SSR	fs = 48 kHz	4	3	1)
AAC LTP	fs = 48 kHz	4	4	1)
SBR	fs = 24/48 kHz (in/out) (SBR tool)	3	2.5	1)
	fs = 24/48 kHz (in/out) (Low Power SBR tool)	2	1.5	1)
	fs = 48/48 kHz (in/out) (Down Sampled SBR tool)	4.5	2.5	1)
	fs = 48/48 kHz (in/out) (Low Power Down Sampled SBR tool)	3	1.5	1)
AAC Scalable	fs = 48 kHz	5	4	1), 2)
TwinVQ	fs = 24 kHz	2	3	1)
CELP	fs = 8 kHz	1	1	
CELP	fs = 16 kHz	2	1	
CELP	fs = 8/16 kHz (bandwidth scalable)	3	1	
HVXC	fs = 8 kHz	2	1	
TTSI			-	4)
General MIDI				
Wavetable Synthesis	fs = 22.05 kHz	depends on bitstreams (3)	depends on bitstreams (3)	
Main Synthetic		depends on bitstreams (3)	depends on bitstreams (3)	
Algorithmic Synthesis and AudioFX		depends on bitstreams (3)	depends on bitstreams (3)	
Sampling Rate Conversion	rf = 2, 3, 4, 6	2	0.5	7)
ER AAC LC	fs = 48 kHz	3	3	1)
ER AAC LTP	fs = 48 kHz	4	4	1)
ER AAC Scalable	fs = 48 kHz	5	4	1), 2)
ER TwinVQ	fs = 24 kHz	2	3	1)
ER BSAC	fs = 48 kHz (input buffer size=26000bits)	4	4	1)
	fs = 48 kHz (input buffer size=106000bits)	4	8	
ER AAC LD	fs = 48 kHz	3	2	1)
ER CELP	fs = 8 kHz	2	1	
	fs = 16 kHz	3	1	
ER HVXC	fs = 8 kHz	2	1	
ER HILN	fs = 16 kHz, ns=93	15	2	6)
	fs = 16 kHz, ns=47	8	2	
ER Parametric	fs = 8 kHz, ns=47	4	2	5),6)

In Part 3: Audio, Subpart 1, subclause 1.5.2.3 (Levels within the profiles), add at the end:

“

- **Levels for the AAC Profile**

Table 1.7A - Levels for the AAC Profile

Level	Max. channels/object	Max. sampling rate [kHz]	Max. PCU	Max. RCU
1	2	24	3	5
2	2	48	6	5
3	NA	NA	NA	NA
4	5	48	19	15
5	5	96	38	15

For the audio object type 2 (AAC LC), mono or stereo mixdown elements are not permitted.

The NA (Not Applicable) levels are introduced to emphasize the hierarchical structure of the AAC Profile and the High Efficiency AAC Profile. Hence, a decoder supporting the High Efficiency AAC Profile at a given level can decode an AAC Profile stream of the same or a lower level. The NA levels are not indicated in the audioProfileLevelIndication table (Table 1.7z).

<https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003>

- **Levels for the High Efficiency AAC Profile**

Table 1.8A - Levels for the High Efficiency AAC Profile

Level	Max. channels/object	Max. AAC sampling rate, SBR not present [kHz]	Max. AAC sampling rate, SBR present [kHz]	Max. SBR sampling rate [kHz] (in/out)	Max. PCU	Max. RCU	Max. PCU Low power SBR	Max. RCU Low power SBR
1	NA	NA	NA	NA	NA	NA	NA	NA
2	2	48	24	24/48	9	10	7	8
3	2	48	48	48/48 (Note 1)	15	10	12	8
4	5	48	24/48 (Note 2)	48/48 (Note 1)	25	28	20	23
5	5	96	48	48/96	49	28	39	23

Note 1: For level 3 and level 4 decoders, it is mandatory to operate the SBR tool in downsampled mode if the sampling rate of the AAC core is higher than 24kHz. Hence, if the SBR tool operates on a 48kHz AAC signal, the internal sampling rate of the SBR tool will be 96kHz, however, the output signal will be downsampled by the SBR tool to 48kHz.

Note 2: For one or two channels the maximum AAC sampling rate, with SBR present, is 48kHz. For more than two channels the maximum AAC sampling rate, with SBR present, is 24kHz.

For the audio object type 2 (AAC LC), mono or stereo mixdown elements are not permitted.

”

STANDARD PREVIEW
(standards.iteh.ai)

In Part 3: Audio, Subpart 1, subclause 1.5.2.4 (Table 1.7z - audioProfileLevelIndication Values), replace the row:

“

ISO/IEC 14496-3:2001/Amd 1:2003

0x28-0x7F	reserved for ISO use
-----------	----------------------

”

with

“

0x28	AAC Profile	L1
0x29	AAC Profile	L2
0x2A	AAC Profile	L4
0x2B	AAC Profile	L5
0x2C	High Efficiency AAC Profile	L2
0x2D	High Efficiency AAC Profile	L3
0x2E	High Efficiency AAC Profile	L4
0x2F	High Efficiency AAC Profile	L5
0x30-0x7F	reserved for ISO use	-

”

In Part 3: Audio, Subpart 1, in subclause 1.6.2.1 AudioSpecificConfig, replace table 1.8 with the following table:

“

Table 1.8 – Syntax of AudioSpecificConfig()

Syntax	No. of bits	Mnemonic
AudioSpecificConfig ()		
{		
audioObjectType;	5	uimsbf
samplingFrequencyIndex;	4	uimsbf
if (samplingFrequencyIndex==0xf)		
samplingFrequency;	24	uimsbf
channelConfiguration;	4	uimsbf
sbrPresentFlag = -1;		
if (audioObjectType == 5) {		
extensionAudioObjectType = audioObjectType;		
sbrPresentFlag = 1;		
extensionSamplingFrequencyIndex;	4	uimsbf
if (extensionSamplingFrequencyIndex==0xf)		
extensionSamplingFrequency;	24	uimsbf
audioObjectType;	5	uimsbf
}		
else {		
extensionAudioObjectType = 0;		
}		
if (audioObjectType == 1 audioObjectType == 2		
audioObjectType == 3 audioObjectType == 4		
audioObjectType == 6 audioObjectType == 7)		
GASpecificConfig();		
if (audioObjectType == 8)		
CelpSpecificConfig();		
if (audioObjectType == 9)		
HvxcSpecificConfig();		
if (audioObjectType == 12)		
TTSSpecificConfig();		
if (audioObjectType == 13 audioObjectType == 14		
audioObjectType == 15 audioObjectType==16)		
StructuredAudioSpecificConfig();		
/* the following Objects are Amendment 1 Objects */		
if (audioObjectType == 17 audioObjectType == 19		
audioObjectType == 20 audioObjectType == 21		
audioObjectType == 22 audioObjectType == 23)		
GASpecificConfig();		
if (audioObjectType == 24)		
ErrorResilientCelpSpecificConfig();		
if (audioObjectType == 25)		
ErrorResilientHvxcSpecificConfig();		
if (audioObjectType == 26 audioObjectType == 27)		
ParametricSpecificConfig();		
if (audioObjectType == 17 audioObjectType == 19		
audioObjectType == 20 audioObjectType == 21		
audioObjectType == 22 audioObjectType == 23		
audioObjectType == 24 audioObjectType == 25		
audioObjectType == 26 audioObjectType == 27) {		
epConfig;	2	uimsbf
if (epConfig == 2 epConfig == 3) {		
ErrorProtectionSpecificConfig();		
}		
if (epConfig == 3) {		

<pre> directMapping; if (! directMapping) { /* tbd */ } } } } </pre>	<p>1</p>	<p>uimsbf</p>
<pre> if (extensionAudioObjectType != 5 && bits_to_decode() >= 16) { syncExtensionType; if (syncExtensionType == 0x2b7) { extensionAudioObjectType; if (extensionAudioObjectType == 5) { sbrPresentFlag; if (sbrPresentFlag == 1) { extensionSamplingFrequencyIndex; if (extensionSamplingFrequencyIndex == 0xf) extensionSamplingFrequency; } } } } </pre>	<p>11</p> <p>5</p> <p>1</p> <p>4</p> <p>24</p>	<p>bslbf</p> <p>uimsbf</p> <p>uimsbf</p> <p>uimsbf</p> <p>uimsbf</p>

iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC 14496-3:2001/Amd 1:2003](https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003)
<https://standards.iteh.ai/catalog/standards/sist/78761174-b9e1-4bd7-8472-ce9f4775e8e7/iso-iec-14496-3-2001-amd-1-2003>

In Part 3: Audio, Subpart 1, in subclause 1.6.2.2.1 Overview, replace table 1.9 by the following table:

“

Table 1.9 – Audio Object Types

Audio Object Type	Object Type ID	definition of elementary stream payloads and detailed syntax	Mapping of audio payloads to access units and elementary streams
AAC MAIN	1	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.2
AAC LC	2	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.2
AAC SSR	3	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.2
AAC LTP	4	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.2
SBR	5	ISO/IEC 14496-3 subpart 4	
AAC scalable	6	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.3
TwinVQ	7	ISO/IEC 14496-3 subpart 4	
CELP	8	ISO/IEC 14496-3 subpart 3	
HVXC	9	ISO/IEC 14496-3 subpart 2	
TTSI	12	ISO/IEC 14496-3 subpart 6	
Main synthetic	13	ISO/IEC 14496-3 subpart 5	
Wavetable synthesis	14	ISO/IEC 14496-3 subpart 5	
General MIDI	15	ISO/IEC 14496-3 subpart 5	
Algorithmic Synthesis and Audio FX	16	ISO/IEC 14496-3 subpart 5	
ER AAC LC	17	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.4
ER AAC LTP	19	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.4
ER AAC scalable	20	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.4
ER Twin VQ	21	ISO/IEC 14496-3 subpart 4	
ER BSAC	22	ISO/IEC 14496-3 subpart 4	
ER AAC LD	23	ISO/IEC 14496-3 subpart 4	see subclause 1.6.2.2.2.1.4
ER CELP	24	ISO/IEC 14496-3 subpart 3	
ER HVXC	25	ISO/IEC 14496-3 subpart 2	
ER HILN	26	ISO/IEC 14496-3 subpart 7	
ER Parametric	27	ISO/IEC 14496-3 subpart 2 and 7	

”

In Part 3: Audio, Subpart 1, under 1.6.3 Semantics, after 1.6.3.6 Direct Mapping add:

“

1.6.3.7 extensionSamplingFrequencyIndex

A four bit field indicating the output sampling frequency of the extension tool corresponding to the extensionAudioObjectType, according to Table 1.10.

1.6.3.8 extensionSamplingFrequency

The output sampling frequency of the extension tool corresponding to the extensionAudioObjectType. Either transmitted directly, or coded in the form of extensionSamplingFrequencyIndex.