INTERNATIONAL STANDARD

ISO/IEC 14496-4

Second edition
2004-12-15

AMENDMENT 8
2005-06-01

Information technology — Coding of audio-visual objects —

Part 4: Conformance testing

AMENDMENT 8: High Efficiency Advanced Audio Coding, audio BIFS, and structured audio conformance

Technologies de l'information — Codage des objets audiovisuels —

Partie 4: Essai de conformité

AMENDEMENT 8: Codage sonore avancé à haute efficacité, BIFS sonore et conformité sonore structurée

Reference number
ISO/IEC 14496-4:2004/Amd.8:2005(E)

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 8 to ISO/IEC 14496-4:2004 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Introduction

This document specifies the eighth amendment to the ISO/IEC 14496-4:2004 standard. The amendment adds the conformance testing for the SBR audio object type defined in 14496-3. It also specifies conformance sequences for BIFS and Structured Audio.

Information technology — Coding of audio-visual objects —

Part 4:
Conformance testing

AMENDMENT 8: High Efficiency Advanced Audio Coding, audio BIFS, and structured audio conformance

In subclause 6.5.1 File name conventions, insert the following row into Table 29 after the row for AAC LTP:

Table 29 – File name conventions

SBR (+AAC LC)	al_sbr_<tool>_<fs>_<nchan>[_fsaac<fs>][_sig<sig>]	al_sbr_<mode>_<tool>_<fs>_<nchan>[_fsaac<fs>][_sig<sig>][_<chan>]

And add:

<tool> indicates the SBR module mainly targeted by the test sequence. Possible values are "e" for testing the envelope adjuster "s" for testing sine addition, "gh" for testing time-grid transitions in combination with changes of SBR header data, "i" for testing inverse filtering, "qmf" for testing the QMF implementation, “cm” for testing various channel modes, “sig” for testing SBR signaling, "twi" for QMF identification, and “sr” for testing various combinations of sampling rates.

<nchan> corresponds to the number of channels present in the conformance test sequence. It is either a single integer, in which case it refers to the number of main audio channels, or two integers separated by a ‘.’, in which case the first integer equals the number of main audio channels, while the second number equals the number of low frequency enhancement channels.

fsaac<fs> corresponds to the sampling rate of the underlying AAC-LC data. If it is omitted, it is half the sampling rate given as output sampling rate.

<sig> is an integer describing the kind of signalling used according to the table below. If this value is omitted, backwards compatible explicit signalling of SBR is used.

Table 29A – File name conventions

sig	Signalling method used
0	Implicit signalling of SBR
1	Hierarchical explicit signalling of SBR
2	Backwards compatible explicit signalling of SBR

<mode> is either "hq" or "lp" for the high quality or the low power version of the SBR decoding algorithm respectively.

In subclause 6.6.1.2.2 (Test procedure), replace:

...testing can be done by comparing the output of a decoder under test with a reference output also supplied by ISO/IEC JTC 1/SC 29/WG 11.

with:

...testing can be done by comparing the output of a decoder under test with a reference output also supplied by ISO/IEC JTC 1/SC 29/WG 11. In cases where the decoder under test is followed by additional operations (e.g. quantizing a signal to a 16 bit output signal) the conformance point is prior to such additional operations, i.e. it is permitted to use the actual decoder output (e.g. with more than 16 bit) for conformance testing.

In subclause 6.6.3.1.2.3 (Bitstream payload), add:

6.6.3.1.2.3.2 aac_scalable_main_header()

max_sfb: Shall not be smaller than last_max_sfb (helper variable specified in ISO/IEC 14496-3).

6.6.3.1.2.3.3 aac_scalable_extension_header()

max_sfb: Shall not be smaller than last_max_sfb (helper variable specified in ISO/IEC 14496-3).

In subclause 6.6.4.1.2.1.1 AudioSpecificConfig(), add:

extensionAudioObjectType: Shall be the Audio Object Type SBR (AOT == 5).

extensionSamplingFrequency: Shall be encoded with a value listed in Table 34, and the value shall be the same as samplingFrequency, or twice the value of samplingFrequency.

extensionSamplingFrequencyIndex: Shall be encoded with a value listed in Table 34, and the value shall indicate an extensionSamplingFrequency being the same as samplingFrequency as indicated by samplingFrequencyIndex, or the value shall indicate an extensionSamplingFrequency being twice the value of samplingFrequency.

sbrPresentFlag: Shall be encoded with the value zero if no SBR data is contained in the compressed MPEG-4 data. If SBR data is present in the compressed MPEG-4 data the parameter shall be encoded with the value one.

In subclause 6.6.4.1.2.1.1 AudioSpecificConfig(), add the following entries to Table 34:

Table 34 — Specification of samplingFrequencyIndex and samplingFrequency

SamplingFrequencyIndex / SamplingFrequency	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
AAC Profile	0x6..0xc, 0xf / <= 24000	0x3..0xc, 0xf / <= 48000	NA	0x3..0xc, 0xf / <= 48000	0x0..0xc, 0xf / <= 96000

samplingFrequencyIndex / samplingFrequency		Level 1	Level 2	Level 3	Level 4	Level 5
High Efficiency AAC Profile	**SBR present**	NA	0x6..0xc, 0xf / <= 24000	0x3..0xc, 0xf / <= 48000	0x3..0xc, 0xf / <= 48000 (Note 1)	0x3..0xc, 0xf / <= 48000
	SBR not present	NA	0x3..0xc, 0xf / <= 48000	0x3..0xc, 0xf / <= 48000	0x3..0xc, 0xf / <= 48000	0x0..0xc, 0xf / <= 96000
Note 1: For Level 4, for one or two channels the maximum AAC sampling rate, with SBR present, is 48 kHz. For more than two channels the maximum AAC sampling rate, with SBR present, is 24 kHz. (0x6..0xc, 0xf / <= 24000)						

extensionSamplingFrequencyIndex / extensionSamplingFrequency	Level 1	Level 2	Level 3,4	Level 5
High Efficiency AAC Profile	NA	0x6..0xc, 0xf / <= 24000	0x3..0xc, 0xf / <= 48000	0x0..0xc, 0xf / <= 96000

In subclause 6.6.4.1.2.1.1 AudioSpecificConfig(), add the following entries to Table 35:

Table 35 — Specification of ChannelConfiguration

ChannelConfiguration	Level 1	Level 2	Level 3	Level 4	Level 5
AAC Profile	0..2	0..2	NA	0..6	0..6
High Efficiency AAC Profile	NA	0..2	0..2	0..6	0..6

In subclause 6.6.4.1.2.2.4 (ics_info()), replace:

Test bitstreams al03 and as17 are provided respectively for Main and Low-Complexity profiles to test decoder performance on non-meaningful transitions

with:

Test sequences al03 and as17 are provided respectively for AOT 2 (AAC LC) and AOT 3 (AAC SSR) to test decoder performance on non-meaningful window sequence transitions (note that AOT 1 (AAC Main) and AOT 4 (AAC LTP) decoders also need to fulfil conformance for AOT 2)

In subclause 6.6.4.1.2.2.9 fill_element(), add:

Fill elements containing an extension_payload with an extension_type of EXT_SBR_DATA or EXT_SBR_DATA_CRC shall not contain any other extension_payload of any other extension_type. For fill elements containing an extension_payload with an extension_type of EXT_SBR_DATA or EXT_SBR_DATA_CRC, the fill_element count field shall be set equal to the total length in bytes, including the SBR enhancement data plus the extension_type field.

In subclause 6.6.15.5 Procedure to Test Decoder Conformance:

At the end of the second paragraph, replace the last sentence:

Conformant decoders must use the RMS Measurement criterion for sequences SY001 through SY004

with:

Conformant decoders must use the RMS Measurement criterion for sequences SY001 through SY004 and SY016 through SY019

In the same subclause, before the last paragraph (Testing of non-normative effects...) add:

Testing of the bus width calculation and send/route mechanism shall be performed using test sequence SY014 and SY015. A reference output is provided by ISO/IEC as example. To be called an ISO/IEC 14496-3 audio decoder, the decoder shall provide two identical outputs for the two sequences, and in particular this output shall be composed of two channels characterized by a reverberated sound in the first and a dry sound in the second

In subclause 6.6.15.6 Description of Conformance sequences, add the following sequence descriptions after SY013:

Sequence SY014 "clarinet1.mp4"

A clarinet synthesized at 44.1 kHz in FM with linear interpolation is routed to a reverberation instrument. The bus width is not set in the send statement, instead the sound is output twice to the bus, creating a two-channel bus. In the reverberation instrument only the first channel is reverberated, the second is output as is.

Sequence SY015 "clarinet2.mp4"

A clarinet synthesized at 44.1 kHz in FM with linear interpolation is routed to a reverberation instrument. The bus width is set to 2 in the send statement, the sound is output in mono to the bus, and then it must be replicated identical on the second channel. In the reverberation instrument only the first channel is reverberated, the second is output as is.

Sequence SY016 "fir.mp4"

In this sequence a sound synthesized at 32kHz is low pass filtered using both fir and firt core opcodes. The two filters have 16 coefficients and normalized cutoff frequency of 0.5 and 0.25 respectively. They are used to filter the left (fir) and right (firt) channels of the sound. This sequence also test pitch converters and other mathematical operators.

Sequence SY017 "vtone.mp4"

A sequence of monophonic sinusoidal tones with sampling rate frequency at 44.1 kHz is shaped according to attack and release time. This test sequence experiences mathematical operators, pitch converters and kline signal generator. Interpolation is linear.

Sequence SY018 "ttone.mp4"

A sequence of monophonic sinusoidal tones with sampling rate frequency at 44.1 kHz is shaped according to attack and release time. It is similar to SY017 but in this case tones are generated through tables instead that by mathematical functions. This test sequence experiences table generators, mathematical operators, table access and oscillators. Interpolation is linear.

Sequence SY019 "otone.mp4"

A sine tone is played in the left channel, with its octave tone in the right channel. Interpolation is linear, sampling rate is 44.1 kHz. This sequence especially experiences user defined core opcodes (including rate polymorphic), parameter passing, array variables.

In the same subclause, after Table 92, add:

Table AMD8-1 — Algorithmic Synthesis and Audio Fx Object Type Test Sequences (continued)

File Name	SY014	SY015	SY016	SY017	SY018	SY019
Content	clarinet1	clarinet2	fir	vtone	ttone	otone
Processing level	All	All	≥Med	All	All	All
RCU - RAM (KB)	< 4	< 4	< 4	< 4	< 4	< 4

After subclause 6.6.16 Main Synthetic, add the following subclauses:

6.6.17 SBR

6.6.17.1 Compressed data

6.6.17.1.1 Characteristics

For all applicable Audio Object Types the SBR extension_payload() elements should be placed last among the extension_payload() elements, i.e. if another type of extension_payload() element is present it should be placed prior to the SBR extension_payload() elements.

If the Audio Object Type SBR is used in combination with either of the Audio Object Types AAC main, AAC LC, AC SSR or AAC LTP, the compressed data shall be stored as outlined in ISO/IEC 14496-3, subclause 4.5.2.8.2.2 SBR Extension Payload for the Audio Object Types AAC main, AAC SSR, AAC LC and AAC LTP.

If the Audio Object Type SBR is used in combination with either of the Audio Object Types ER AAC LC or ER AAC LTP, the compressed data shall be stored as outlined in ISO/IEC 14496-3, subclause 4.5.2.8.2.3 SBR Extension Payload for the Audio Object Types ER AAC LC and ER AAC LTP. For these AOTs, DRC extension_payload() elements are not permitted simultaneously with SBR extension_payload() elements within one er_raw_data_block(). Moreover, SBR extension_payload() elements of the type EXT_SBR_DATA_CRC shall not be used with these AOTs.

For the scalable AOTs (AAC scalable and ER AAC scalable), the SBR data should be transmitted and devised according to ISO/IEC 14496-3, subclause 4.5.2.8.2.4 SBR extension payload for the Audio Object Types AAC scalable and ER AAC scalable. Restrictions are here put on the frequency range of the SBR data and in what layers of the scalable stream the SBR data is stored. Furthermore, SBR extension_payload() elements of the type EXT_SBR_DATA_CRC shall not be used with the audio object types ER AAC scalable.

6.6.17.1.2 Test procedure

Each compressed data shall meet the syntactic and semantic requirements specified in ISO/IEC 14496-3. The decoded data shall also meet the requirements defined in ISO/IEC 14496-3 subclause 4.6.18.3.6 Requirements. If a syntactic element is not listed below, no restrictions apply to that element. The **bs_reserved** elements shall be encoded with the value zero.

6.6.17.1.2.1 Compressed MPEG-4 data payload

6.6.17.1.2.1.1 sbr_header()

The following parameters shall be encoded with values subsequently used in defining a frequency range, a number of noise bands, a number of limiter bands, and a number of patches:

bs_start_freq
bs_stop_freq
bs_xover_band
bs_alter_scale
bs_noise_bands
bs_limiter_bands

The above parameters are used (in ISO/IEC 14496-3) to calculate the variables below:

k_2
k_0
k_x
M
N_Q
numPatches
numBands
numBands0
vDk0
vDk1

Conformant compressed MPEG-4 data shall have values for the above parameters that subsequently evaluate to values of the above variables that satisfy the requirements outlined in ISO/IEC 14496-3, subclause 4.6.18.3.6 Requirements.

6.6.17.1.2.1.2 sbr_channel_pair_base_element()

bs_coupling: Shall be encoded with the value of 1

6.6.17.1.2.1.3 sbr_grid()

The following compressed MPEG-4 data elements shall be encoded so that a value of the number of SBR envelopes for a SBR frame, for a given frame class, is within the limits defined in ISO/IEC 14496-3, subclause 4.6.18.3.6 Requirements:

bs_rel_bord_0
bs_rel_bord_1
bs_num_env
bs_var_bord
bs_num_rel_0
bs_var_bord_0
bs_var_bord_1
bs_num_rel_0
bs_num_rel_1

Conformant compressed MPEG-4 data shall have the above parameters chosen so that the leading border of a given SBR frame (the frame boundary) coincides with the trailing border of the previous SBR frame (the frame boundary of the previous frame). Furthermore, the above parameters shall be chosen so that the envelope borders of the SBR envelopes in a given frame fall within the boundaries of the SBR frame. The above parameters shall also be chosen so that every SBR envelope within the SBR frame has a duration larger than zero.

6.6.17.1.2.1.4 sbr_dtdf()

bs_df_env[]: Shall be encoded with the value 0 for the first envelope of the present frame, if the compressed MPEG-4 data element bs_header_flag has the value one (i.e. a new sbr_header is available), or if the amp_res value has changed from the previous frame due to the rule specifying amp_res = 0 for a frame of frame class FIXFIX with only one envelope.

bs_df_noise[]: Shall be encoded with the value 0 for the first noise floor of the present frame, if the compressed MPEG-4 data element bs_header_flag has the value one, i.e. a new sbr_header is available.

6.6.17.1.2.1.5 sbr_envelope()

bs_codeword: Shall be encoded with the values listed in the corresponding Huffman table, defined in ISO/IEC 14496-3, Annex 4.A.6.1

Conformant compressed MPEG-4 data shall have coded envelope scalefactors based on quantized envelopes scalefactors that satisfy the requirements outlined in ISO/IEC 14496-3, subclause 4.6.18.3.6 Requirements.

The quantised envelope scale factors $\mathbf{E}$ for single channel elements and $\mathbf{E}_0$ and $\mathbf{E}_1$ for channel pair elements shall be encoded with values that are within the following limits:

- For single channel elements: $0 \le \mathbf{E}(i,l) < 2^{7-amp\_res}$
- For channel pair elements: $\begin{cases} 0 \le \mathbf{E}_0(i,l) < 2^{7-amp\_res} \\ 0 \le \mathbf{E}_1(i,l) < 2^{7-amp\_res-bs\_coupling} \end{cases}$

where

$$amp\_res = \begin{cases} 0 & \text{, if } bs\_num\_env = 1 \text{ and } frame\_class = FIXFIX \\ bs\_amp\_res & \text{, otherwise} \end{cases}$$

where subscript zero indicates the firstly encoded channel in the channel pair element and subscript one indicates the secondly encoded channel in the channel pair element.

6.6.17.1.2.1.6 sbr_noise()

bs_codeword: Shall be encoded with the values listed in the corresponding Huffman table, defined in ISO/IEC 14496-3, Annex 4.A.6.1

Conformant compressed MPEG-4 data shall have coded noise floor scalefactors based on quantised noise floor scalefactors that satisfy the requirements outlined in ISO/IEC 14496-3, subclause 4.6.18.3.6 Requirements.

6.6.17.2 Decoders

6.6.17.2.1 Characteristics

The object type SBR has the Object Type ID 5, and the compressed MPEG-4 data syntax is defined in ISO/IEC 14496-3. The Audio Object Type SBR contains the SBR Tool. The SBR Tool can be implemented in two different versions:

- High-Quality SBR Tool
- Low-Power SBR Tool

The different versions can also be operated in down-sampled SBR-mode.