INTERNATIONAL STANDARD

ISO/IEC 14496-26

First edition
2010-05-01
AMENDMENT 3
2014-01-15

Information technology — Coding of audio-visual objects —

Part 26:
Audio conformance

AMENDMENT 3: Conformance for Low Delay AAC v2 profile

Technologies de l'information — Codage des objets audiovisuels —

Partie 26: Conformité audio

AMENDEMENT 3: Conformité pour profil Low Delay AAC v2

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 3 to ISO/IEC 14496-26:2010 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

It adds the conformance testing for the Low Delay AAC v2 profile defined in ISO/IEC 14496-3.

Information technology — Coding of audio-visual objects —

Part 26:
Audio conformance

AMENDMENT 3: Conformance for Low Delay AAC v2 profile

In 2, Normative references, add:

ISO/IEC 23003-1, *Information technology — MPEG audio technologies — Part 1: MPEG Surround*

ISO/IEC 23003-2, *Information technology — MPEG audio technologies — Part 2: Spatial Audio Object Coding (SAOC)*

In 6.1, File name conventions, insert the following rows at the end of Table 2:

"

LD MPEG Surround (+ER AAC LD)	er_ad_ldmps_<coreSetup>	er_ad_ldmps_<coreSetup>
LD MPEG Surround (+ER AAC ELD)	er_eld_ldmps_<coreSetup>	er_eld_ldmps_<coreSetup>

"

After 7.25, Low Delay SBR, add:

7.26 Low Delay MPEG Surround

7.26.1 Compressed data

7.26.1.1   Characteristics

The Low Delay MPEG Surround tool can be used in combination with the ER AAC LD AOT and the ER AAC ELD AOT.

The compressed data shall be stored as outlined in ISO/IEC 14496-3:2009, 4.5.2.13. The extension_type for Low Delay MPS is EXT_LDSAC_DATA. The transport and signalling of LD MPS in an MPEG-4 Audio/Systems Environment is done in the same way as for the SAOC object type, as described in 8.2.2 of ISO/IEC 23003-2. LD MPS uses the AOT 44. The ldsac_extension_data() is used to carry a LDSacDataFrame() as described in B.4 of ISO/IEC 23003-2.

7.26.1.2   Test procedure

Each compressed data shall meet the syntactic and semantic requirements specified in ISO/IEC 23003-2:2010, Annex B. If a syntactic element is not listed below, no restrictions apply to that element.

7.26.1.2.1 Compressed MPEG-4 data payload

7.26.1.2.1.1 LDSpatialSpecificConfig()

bsSamplingFrequencyIndex

	For further restrictions, see 7.26.1.2.2
bsSamplingFrequency	For restrictions, see 7.26.1.2.2
bsFreqRes	Shall not be encoded with a value of 0
bsFrameLength	For restrictions, see 7.26.1.2.2
bsTreeConfig	For restrictions, see 7.26.1.2.2
bsQuantMode	Shall not be encoded with a value of 3
bsOneIcc	Not present if **bsTreeConfig** = 7. No further restrictions apply
bsArbitraryDownmix	No restrictions apply
bsFixedGainsSur	Not present if **bsTreeConfig** = 7. Shall be in the range 0…4
bsFixedGainsLFE	Not present if **bsTreeConfig** = 7. Shall be in the range 0…4
bsFixedGainsDMX	No restrictions apply
bsMatrixMode	Not present if **bsTreeConfig** = 7. No further restrictions apply
bsTempShapeConfig	Shall not be encoded with a value of 3
bsDecorrConfig	Shall not be encoded with a value of 3
bs3DaudioMode	Not present if **bsTreeConfig** = 7. No further restrictions apply
bsEnvQuantMode	Shall be 0
bs3DaudioHRTFset	Shall be 0

7.26.1.2.1.2 OttConfig()

bsOttBands	Shall not be encoded with a value larger than numBands given in table Table B.9 of ISO/IEC 23003-2:2010

7.26.1.2.1.3 TttConfig()

bsTttDualMode	No restrictions apply
bsTttModeLow	Shall be in the range of 0..5
bsTttModeHigh	Shall be in the range of 0..5
bsTttBandsLow	Shall not be encoded with a value larger than numBands given in table Table B.9 of ISO/IEC 23003-2:2010

7.26.1.2.1.4 ParamHRTFset()

bsHRTFfreqRes	Shall not be encoded with a value of 0
bsHRTFasymmetric	No restrictions apply
bsHRTFlevelLeft	No restrictions apply
bsHRTFlevelRight	No restrictions apply
bsHRTFphase	No restrictions apply
bsHRTFphaseLR	No restrictions apply
bsHRTFicc	No restrictions apply
bsHRTFiccLR	No restrictions apply

7.26.1.2.1.5   SpatialExtensionConfig()

bsSacExtType	Shall be in the range of 2..15. Note that in the case of values indicated as "reserved" in Table 54 of ISO/IEC 23003-1:2007, the parsing function SpatialExtensionConfigData(bsSacExtType) shall return the value 0, such that possibly present data is read as bsFillBits (i.e., skipped) and correct parsing of the bitstream can continue
bsSacExtLen	No restrictions apply
bsSacExtLenAdd	No restrictions apply
bsSacExtLenAddAdd	No restrictions apply
bsFillBits	No restrictions apply

7.26.1.2.1.6   SpatialExtensionConfigData(0)

None

7.26.1.2.1.7   SpatialExtensionConfigData(1)

None

7.26.1.2.1.8   TreeConfig()

bsOttBoxPresent	No restrictions apply
bsOttDefaultCld	No restrictions apply
bsOttModeLFE	No restrictions apply
bsOttBands	Shall not be encoded with a value larger than numBands given in Table B.9 of ISO/IEC 23003-2:2010
bsOutputChannelPos	Shall be in the range 0..26

For restrictions, see 7.26.1.2.2

7.26.1.2.1.9   LDSacDataFrame()

ldsacHeaderFlag	No restrictions apply
ldsacHeaderLen	No restrictions apply
ldsacHeaderLenAdd	No restrictions apply
bsFillBits	No restrictions apply
ldsacTimeAlignFlag	No restrictions apply
ldsacTimeAlign	Shall be 0

7.26.1.2.1.10   LDSpatialFrame()

bsIndependencyFlag	No restrictions apply

7.26.1.2.1.11   FramingInfo()

bsFramingType	No restrictions apply
bsNumParamSets	For restrictions, see 7.26.1.2.2
bsParamSlot[ps]	Shall be in the range 0..**bsFrameLength**

7.26.1.2.1.12   OttData()

No restrictions apply

7.26.1.2.1.13 TttData()

ICC values of a certain TTT box shall not be encoded with a value of 0 if (**bsTTTModeLow** < 2 || (**bsTTTDualMode**==1&&**bsTTTModeHigh**<2))

7.26.1.2.1.14 SmgData()

bsSmoothMode	No restrictions apply
bsSmoothTime	No restrictions apply
bsFreqResStrideSmg	No restrictions apply
bsSmgData	No restrictions apply

7.26.1.2.1.15 TempShapeData()

bsTempShapeEnable	No restrictions apply
bsTempShapeEnableChannel[ch]	No restrictions apply

7.26.1.2.1.16 EnvelopeReshapeHuff()

Hcod2D_EnvRes   **bsCodeW** shall have a value of a set of values as defined by column 'codeword' of Table A.25 of ISO/IEC 23003-1:2007 and shall have a length as defined by the corresponding entry in column 'length'

7.26.1.2.1.17 ArbitraryDownmixData()

No restrictions apply

7.26.1.2.1.18 EcData()

bsXXXdataMode	Shall fulfil the requirements outlined in ISO/IEC 23003-1:2007, 6.1.13
bsDataPair	Shall have the value 0 if setIdx == datasets-1. No further restrictions apply
bsQuantCoarseXXX	No restrictions apply
bsFreqResStrideXXX	No restrictions apply

7.26.1.2.1.19 EcDataPair()

bsPcmCodingXXX	No restrictions apply

7.26.1.2.1.20 GroupedPcmData()

bsPcmWord	No restrictions apply

7.26.1.2.1.21 DiffHuffData()

bsDiffType	No restrictions apply
bsCodingScheme	No restrictions apply

7.26.1.2.1.22 HuffData1D()

hcodFirstBand_XXX   **bsCodeW** shall have a value out of a set of values as defined by column 'codeword' of Tables A.2, A.3 or A.4 of ISO/IEC 23003-1:2007, respectively, and shall have a length as defined by the corresponding entry in column 'length'

hcod1D_XXX_YY   **bsCodeW** shall have a value out of a set of values as defined by column 'codeword' of Tables A.5, A.6 or A.7 of ISO/IEC 23003-1:2007, respectively, and shall have a length as defined by the corresponding entry in column 'length'

bsSign   No restrictions apply

7.26.1.2.1.23   HuffData2DFreqPair()

hcodLavIdx   **bsCodeW** shall have a value out of a set of values as defined by column 'codeword' of Table A.24 of ISO/IEC 23003-1:2007, and shall have a length as defined by the corresponding entry in column 'length'

hcod2D_XXX_YY_FP_LL

bsCodeW shall have a value out of a set of values as defined by column 'codeword' of the applicable table out of Tables A.8, A.9 or A.10 of ISO/IEC 23003-1:2007, and shall have a length as defined by the corresponding entry in column 'length'

hcod1D_XXX_YY   **bsCodeW** shall have a value out of a set of values as defined by column 'codeword' of Tables A.5, A.6 or A.7 of ISO/IEC 23003-1:2007, respectively, and shall have a length as defined by the corresponding entry in column 'length'

bsSign   No restrictions apply

7.26.1.2.1.24   SymmetryData()

bsSymBit[i]   No restrictions apply

7.26.1.2.1.25   LsbData()

bsLsb   For restrictions see ISO/IEC 23003-1:2007, 8.4.2.3.24

7.26.1.2.1.26   SpatialExtensionFrame()

Shall be in the range of 2…15. Note that in case of bsSacExtType having values indicated as "reserved" in Table 54 ISO/IEC of 23003-1:2007, the parsing function SpatialExtensionFrameData(bsSacExtType) shall return the value 0, such that possibly present data is read as bsFillBits (i.e., skipped) and correct parsing of the bitstream can continue

7.26.1.2.1.27   SpatialExtensionFrameData(0)

None

7.26.1.2.1.28   SpatialExtensionFrameData(1)

None

7.26.1.2.1.29   SpatialExtensionFrameData(2)

No further restrictions apply

7.26.1.2.1.30   ArbitraryTreeData()

No further restrictions apply

7.26.1.2.1.31   Restrictions applying to decoded parameters

See ISO/IEC 23003-1:2007, 8.4.2.3.24

7.26.1.2.2   Low Delay AAC v2 profile

For LD MPEG Surround in the Low Delay AAC v2 profile, the following further restrictions apply.

bsSamplingFrequencyIndex
Shall be encoded with a value listed in **Table 1**

bsSamplingFrequency Shall be encoded with a value listed in **Table 1**

bsFrameLength   Shall be in the range 3..31

bsTreeConfig   Shall be encoded with a value listed in **Table 1**

Table 1 - Restrictions for the Low Delay AAC v2 profile

	Level 1	Level 2	Level 3	Level 4
bsSamplingFrequencyIndex	n/a	0x3..0xc, 0xf	0x3..0xc, 0xf	0x3..0xc, 0xf
bsSamplingFrequency	n/a	<= 48000	<= 48000	<= 48000
bsTreeConfig	n/a	7	7	0,1,2,7

TreeConfig()   After decoding of the syntactic element TreeConfig(), the helper variable numOutChanAT shall have a value not larger than 6

SpatialExtensionConfigData(0)

This syntactic element shall not be present

SpatialExtensionConfigData(1)

This syntactic element shall not be present

bsNumParameterSets   Shall have a value not larger than 2

7.26.2 Decoders

7.26.2.1   Characteristics

The object type LD MPEG Surround has the Object Type ID 44, and the compressed MPEG-4 data syntax is defined in ISO/IEC 14496-3:2009. The Audio Object Type LD MPEG Surround contains the LD MPEG Surround tool. The LD MPEG Surround decoder can only be implemented in one flavour, it does not allow for Low Power decoding.

7.26.2.2   Low Delay AAC v2 profile

According to the Low Delay AAC v2 profile as defined in ISO/IEC 14496-3:2009, LD MPEG Surround object type can be used in combination with ER AAC LD, and ER AAC ELD object types. In this case, the aforementioned downmix decoders shall fulfil the conformance criteria that are applicable to each of them.

7.26.2.3   LD MPEG Surround conformance test procedure

The test sequences specified in Table 4 and Table 5 have to be applied. The conformance of the underlying downmix decoder shall be tested before conformance testing is done for the LD MPEG Surround decoder. The following figure gives an overview of the conformance process:
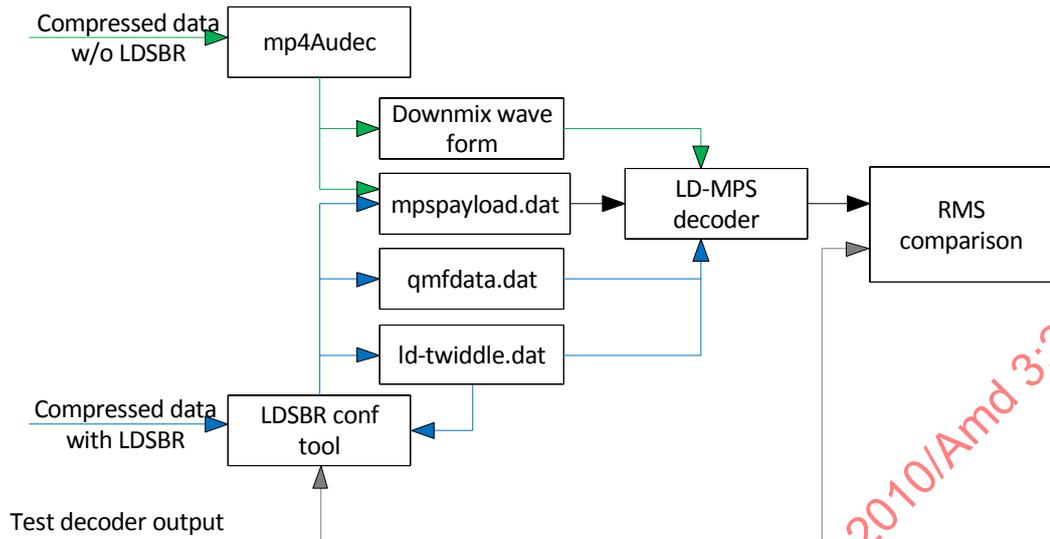
Figure 1: Conformance process of Low delay AAC v2 bit streams including LD-MPS

The decoding process is separated into a downmix decoding step and a LD-MPS decoding step. Besides the downmix data, the LD-MPS payload data needs to be extracted from the bit stream. For all bit streams not using the LD-SBR tool, the interface between downmix core coder and LD-MPS decoder is in the time domain. For stream where LD-SBR is enabled, the interface is in the QMF domain.

For those streams where LD-SBR is enabled, the twiddle factors of the QMF can vary depending on the implementation. For analysing and testing the twiddle factors, the downmix needs to be decoded using the LD-SBR conformance tool (see ISO/IEC 14496-26, Section 7.25.2.1.1). Note, that in order to analyse the QMF twiddle factors, the bit streams er_eld_ldmps_ldqmf_x need to be decoded first. These estimated twiddle factors will be used for the remaining bit streams.

7.26.2.4 Test sequences

The naming convention for the test sequences is defined in

Table 2 and

Table 3. In **Table 4** and **Table 5** the test sequences are defined for ER AAC LD and ER AAC ELD downmix coders respectively.

Table 2 - Mapping between file prefix and downmix coder

File prefix	Downmix coder
er_ad	ER AAC LD
er_eld	ER AAC ELD

Table 3 - Mapping between third part of file name and bitstream properties

File basename	Bitstream Properties
_ldqmf_X	Test the low delay qmf implementation; estimation of twiddle factors. X=[1, 2 (downsampled SBR)]
_param_X	Test various numbers of parameter bands. X=[4, 5, 7, 9, 12, 15, 23]
_X_ts	Test various numbers of time slots. X=[8, 15, 16]
_Xkhz	Tests various sampling rates. X=[16, 22, 24, 32, 44]
_tree_X	Test various tree configurations. X=[5151, 5152, 525]
_one_ICC	Test one ICC mode.