DRAFT AMENDMENT
ISO/IEC 23003-3:2012/DAM 3

ISO/IEC JTC **1**/SC **29**

Secretariat: **JISC**

Voting begins on:
2015-09-28

Voting terminates on:
2015-12-28

Information technology — MPEG audio technologies —

Part 3:
Unified speech and audio coding

AMENDMENT 3: Support of MPEG-D DRC

Technologies de l'information — Technologies audio MPEG —

Partie 3: Discours unifié et codage audio

AMENDEMENT 3: .

ICS: 35.040

Reference number
ISO/IEC 23003-3:2012/DAM 3:2015(E)

© ISO/IEC 2015

COPYRIGHT PROTECTED DOCUMENT

Contents

Page

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 23003-3:2012 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Information technology — MPEG audio technologies — Part 3: Unified speech and audio coding, AMENDMENT 3: Support of MPEG-D DRC

1   Changes to the text of ISO/IEC 23003-3:2012

In Clause 2 Normative References add:

ISO/IEC 23003-4, "Information technology — MPEG audio technologies — Part 4: Dynamic Range Control"

In Clause 4 add new sub-clause:

4.xx Combination of USAC with MPEG-D DRC

The output of the USAC decoder can be further processed by MPEG-D DRC (ISO/IEC 23003-4). If the SBR tool in USAC is active, a USAC decoder can typically be efficiently combined with a subsequent MPEG-D DRC decoder by connecting them in the QMF domain in the same way as it is described in ISO/IEC 23003-4. If a connection in the QMF domain is not possible they shall be connected in the time domain.

The MPEG-D DRC payload shall be embedded into a USAC bitstream by means of the usacExtElement mechanism, with usacExtElementType of type ID_EXT_ELE_UNI_DRC. The loudness metadata shall be embedded by means of the usacConfigExt mechanism with usacConfigExtType of type ID_CONFIG_EXT_LOUDNESS_INFO. The time-alignment between the USAC data and the MPEG-D DRC data assumes the most efficient connection between the USAC decoder and the MPEG-D DRC decoder. If the SBR tool in USAC is active, the most efficient connection is in the QMF domain. Otherwise, the most efficient connection is in the time domain. The DRC tool is operated in regular delay mode and the DRC frame size has the same duration as the USAC frame size. The same holds for the DRC sampling rate, which is synchronized to the USAC sampling rate.

The time resolution of the DRC tool is specified by *deltaTmin* in units of the audio sample interval. It is calculated as specified in ISO/IEC 23003-4. Specific values are provided here as examples based on the following formula:

$$\text{deltaTmin} = 2^M.$$

The applicable exponent *M* is found by looking up the audio sample rate range that fulfills:

$$f_{s,\min} \le f_s < f_{s,\max}.$$

Table AMD3.1 --- Lookup table for the exponent *M*

fs,min [Hz]	*fs*,max [Hz]	*M*
8000	16000	3
16000	32000	4
32000	64000	5

64000	128000	6

Given the codec frame size N_{Codec} (==outputFrameLength), the DRC frame size in units of DRC samples at a rate of *deltaTmin* is:

$$N_{DRC} = N_{Codec} 2^{-M}.$$

For USAC, MPEG-D DRC offers mandatory decoding capability of up to four DRC subbands using the time-domain DRC filter bank. More DRC subbands can be supported by operating in the QMF-domain. DRC sets that contain more than four DRC subbands must contain gain sequences that are all aligned with the QMF-domain used for SBR. If the SBR tool in USAC is active, MPEG-D DRC shall always operate in the QMF-domain. The gain sequences are all aligned with the QMF domain in that case.

If no additional filter bank is required for the application of multiband DRC gains, MPEG-D DRC doesn't introduce any additional decoding delay.

The drcLocation parameter shall be encoded according to Table AMD3.2.

Table AMD3.2 - Encoding of drcLocation parameter

drcLocation n	Payload
1	uniDrcConfig() / uniDrcGain() (see ISO/IEC 23003-4)
2	*reserved*
3	*reserved*
4	*reserved*

In Clause 5, replace Table 14 with:

Table 14 — Syntax of UsacExtElementConfig()

Syntax	No. of bits	Mnemonic
UsacExtElementConfig()		
{		
usacExtElementType = escapedValue(4,8,16)**;**		
usacExtElementConfigLength = escapedValue(4,8,16);		
usacExtElementDefaultLengthPresent;	1	uimsbf
if (usacExtElementDefaultLengthPresent) {		
usacExtElementDefaultLength = escapedValue(8,16,0) + 1;		
} else {		
usacExtElementDefaultLength = 0;		
}		
usacExtElementPayloadFrag;	1	uimsbf
switch (usacExtElementType) {		
case ID_EXT_ELE_FILL:		
break;		
case ID_EXT_ELE_MPEGS:		
SpatialSpecificConfig();		
break;		
case ID_EXT_ELE_SAOC:		

```
        SaocSpecificConfig();
        break;
    case ID_EXT_ELE_AUDIOPREROLL:
        /* No configuration element */
        break;
    case ID_EXT_ELE_UNI_DRC:
        uniDrcConfig();
        break;
    default:                                                NOTE
        while (usacExtElementConfigLength--) {
            tmp;                                            8           uimsbf
        }
        break;
    }
}
```

NOTE: The default entry for the usacExtElementType is used for unknown extElementTypes so that legacy decoders can cope with future extensions.		

In Clause 5 replace Table 15 with:

Table 15 — Syntax of UsacConfigExtension()

Syntax	No. of bits	Mnemonic
UsacConfigExtension()		
{		
numConfigExtensions = escapedValue(2,4,8) + 1;		
for (confExtIdx=0; confExtIdx<numConfigExtensions; confExtIdx++) {		
usacConfigExtType[confExtIdx] = escapedValue(4,8,16);		
usacConfigExtLength[confExtIdx] = escapedValue(4,8,16);		
switch (usacConfigExtType[confExtIdx]) {		
case ID_CONFIG_EXT_FILL:		
while (usacConfigExtLength[confExtIdx]--) {		
fill_byte[i]; /* should be '10100101' */	**8**	**uimsbf**
}		
break;		
case ID_CONFIG_EXT_LOUDNESS_INFO:		
loudnessInfoSet()		
break;		
default:		
while (usacConfigExtLength[confExtIdx]--) {		
tmp;	**8**	**uimsbf**
}		
break;		
}		
}		
}		

Add Subclause "5.3.5 Payload of extension elements":

5.3.5 Payload of extension elements

Table AMD3.3 — Syntax of AudioPreRoll()

Syntax	No. of bits	Mnemonic

AudioPreRoll()
{
    configLen = escapedValue(4,4,8);
    Config()
 	
    numPreRollFrames = escapedValue(2,4,0);
 	
    for (frameIdx=0; frameIdx < numPreRollFrames; ++frameIdx) {
        auLen = escapedValued(16,16,0)
        AccessUnit()
    }
}
	
	4..16
	8*configLen
	
	2..6
	
	16..32
	8*auLen
		
		uimsbf

In Clause 6 replace Table 73 with:

Table 73 — Value of usacExtElementType

usacExtElementType	Value
ID_EXT_ELE_FILL	0
ID_EXT_ELE_MPEGS	1
ID_EXT_ELE_SAOC	2
ID_EXT_ELE_AUDIOPREROLL	3
ID_EXT_ELE_UNI_DRC	4
/* reserved for ISO use */	5-127
/* reserved for use outside of ISO scope */	128 and higher
NOTE: Application-specific usacExtElementType values are mandated to be in the space reserved for use outside of ISO scope. These are skipped by a decoder as a minimum of structure is required by the decoder to skip these extensions.	

In Clause 6 replace Table 74 with:

Table 74 — Value of usacConfigExtType

usacConfigExtType	Value
ID_CONFIG_EXT_FILL	0
/* reserved for ISO use */	1
ID_CONFIG_EXT_LOUDNESS_INFO	2
/* reserved for ISO use */	3-127
/* reserved for use outside of ISO scope */	128 and higher

In Clause 6 replace Table 81 with:

Table 81 — Interpretation of data blocks for USAC extension payload decoding

usacExtElementType	The concatenated usacExtElementSegmentData represents:
ID_EXT_ELE_FIL	Series of **fill_byte**
ID_EXT_ELE_MPEGS	SpatialFrame()
ID_EXT_ELE_SAOC	SaocFrame()
ID_EXT_ELE_AUDIOPREROLL	AudioPreRoll()
ID_EXT_ELE_UNI_DRC	uniDrcGain() as defined in ISO/IEC 23003-4
unknown	unknown data. The data block shall be discarded.

Add Subclause "7.18 Audio Pre-Roll":

7.18 Audio Pre-Roll

7.18.1 General

The *AudioPreRoll()* syntax element is used to transmit audio information of previous frames along with the data of the present frame. The additional audio data can be used to compensate the decoder startup delay (pre-roll), thus enabling random access at stream access points (SAP) that make use of *AudioPreRoll().*

A *UsacExtElement()* with the usacExtElementType of ID_EXT_ELE_AUDIOPREROLL shall be used to transmit the *AudioPreRoll().*

7.18.2 Semantics

configLen   Size of the configuration syntax element in bytes.

Config()   The decoder configuration syntax element. In the context of this standard this shall be the UsacConfig() as defined in 5.2. The Config() field may be transmitted to be able to respond to changes in the audio configuration (e.g. switching of streams).

numPreRollFrames   The number of pre-roll access units (AUs) transmitted as audio pre-roll data. The reasonable number of AUs depends on the decoder start-up delay.

auLen   AU length in bytes.

AccessUnit()   The pre-roll AU(s).

NOTE   The pre-roll data carried in the extension element may be excluded from buffer requirement restrictions, i. e. the buffer requirements may not be satisfied

In order to use *AudioPreRoll()* for both random access and bitrate adaptation the following restrictions apply:

— The first element of every frame shall be an extension element (UsacExtElement) of type ID_EXT_ELE_AUDIOPREROLL.

— The corresponding UsacExtElement() shall be configured as specified in Table AMD3.4.