INTERNATIONAL STANDARD

ISO/IEC 23003-1

First edition
2007-02-15

AMENDMENT 3
2016-12-15

Information technology — MPEG audio technologies —

Part 1:
MPEG Surround

AMENDMENT 3: MPEG Surround extension for 3D Audio

Technologies de l'information — Technologies audio MPEG —

Partie 1: Ambiance MPEG

AMENDEMENT 3: Extension de l'ambiance MPEG pour audio 3D

Reference number
ISO/IEC 23003-1:2007/Amd.3:2016(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Information technology — MPEG audio technologies —

Part 1:
MPEG Surround

AMENDMENT 3: MPEG Surround extension for 3D Audio

Page 3, 3.1

Add the following after 3.1.12:

3.1.13

N-N/2-N configuration

configuration of MPEG Surround coding system that recreated N channels from half of N downmixed channels with the corresponding spatial parameters

Pages 3 and 4, 3.1

Renumber the terms 3.1.13 to 3.1.26 as 3.1.14 to 3.1.27.

Page 6, 3.5

Add the following variables:

$\mathbf{I}_N$ is unity matrix and subscript index indicate matrix dimension, e.g. N by N unity matrix.

$\mathbf{O}_N$ is null matrix and subscript index indicate matrix dimension, e.g. N by N null matrix.

Add a new Clause 10

10 Outline

10.1 General

The decoding process for N-N/2-N is described in the following clause.

10.2 Syntax

10.2.1 Payloads for N-N/2-N Extension

Table 10.1 — Syntax of SpatialSpecificConfig()

Syntax	No. of bits	Mnemonic
`SpatialSpecificConfig()`		
`{`		
`  bsSamplingFrequencyIndex;`	4	uimsbf
`  if (bsSamplingFrequencyIndex == 0xf ) {`		
NOTE 1 SpeakerConfig3d() is defined in ISO/IEC 23008-3:2015, Table 5.		
NOTE 2 numOttBoxes and numTttBoxes are defined by Table 10.2 dependent on bsTreeConfig.		

Table 10.1 *(continued)*

Syntax	No. of bits	Mnemonic
bsSamplingFrequency;	**24**	**uimsbf**
}		
bsFrameLength;	**7**	**uimsbf**
bsFreqRes;	**3**	**uimsbf**
bsTreeConfig;	**4**	**uimsbf**
if (bsTreeConfig == '0111') {		
bsNumInCh;	**4**	**uimsbf**
bsNumLFE	**2**	**uimsbf**
bsHasSpeakerConfig	**1**	**uimsbf**
if (bsHasSpeakerConfig == 1) {		
audioChannelLayout = SpeakerConfig3d();		Note 1
}		
}		
bsQuantMode;	**2**	**uimsbf**
bsOneIcc;	**1**	**uimsbf**
bsArbitraryDownmix;	**1**	**uimsbf**
bsFixedGainSur;	**3**	**uimsbf**
bsFixedGainLFE;	**3**	**uimsbf**
bsFixedGainDMX;	**3**	**uimsbf**
bsMatrixMode;	**1**	**uimsbf**
bsTempShapeConfig;	**2**	**uimsbf**
bsDecorrConfig;	**2**	**uimsbf**
bs3DaudioMode;	**1**	**uimsbf**
if (bsTreeConfig == '0111') {		
for (i=0; i< NumInCh - NumLfe; i++) {		
defaultCld[i] = 1;		
ottModelfe[i] = 0;		
}		
for (i= NumInCh - NumLfe; i< NumInCh; i++) {		
defaultCld[i] = 1;		
ottModelfe[i] = 1;		
}		
}		
for (i=0; i<numOttBoxes; i++) {		Note 2
OttConfig(i);		
}		
for (i=0; i<numTttBoxes; i++) {		Note 2
TttConfig(i);		
}		

NOTE 1 SpeakerConfig3d() is defined in ISO/IEC 23008-3:2015, Table 5.

NOTE 2 numOttBoxes and numTttBoxes are defined by Table 10.2 dependent on bsTreeConfig.

Table 10.1 *(continued)*

Syntax	No. of bits	Mnemonic
`if (bsTempShapeConfig == 2) {`		
`    bsEnvQuantMode`	1	uimsbf
`}`		
`if (bs3DaudioMode) {`		
`    bs3DaudioHRTFset;`	2	uimsbf
`if (bs3DaudioHRTFset==0) {`		
`    ParamHRTFset();`		
`  }`		
`}`		
`ByteAlign();`		
`SpatialExtensionConfig();`		
`}`		
NOTE 1 SpeakerConfig3d() is defined in ISO/IEC 23008-3:2015, Table 5.		
NOTE 2 numOttBoxes and numTttBoxes are defined by Table 10.2 dependent on bsTreeConfig.		

Table 10.2 — bsTreeConfig

bsTreeConfig	Meaning
0,1,2,3,4,5,6	Identical meaning in ISO/IEC 20003-1:2007, Table 40
7	N-N/2-N configuration numOttBoxes = NumInCh numTttBoxes = 0 numInChan = NumInCh numOutChan = NumOutCh output channel ordering is according to Table 10.5
8...15	Reserved

bsNumInCh Defines number of input DMX channels for N-N/2-N configuration according to:

Table 10.3 — bsNumInCh

bsNumInCh	NumInCh	NumOutCh
0	12	24
1	7	14
2	5	10
3	6	12
4	8	16
5	9	18
6	10	20
7	11	22
8	13	26
9	14	28

Table 10.3 *(continued)*

bsNumInCh	NumInCh	NumOutCh
10	15	30
11	16	32
12,...,15	Reserved	Reserved

bsNumLfe Defines number $\mathbf{N_{LFE}}$ of output Lfe channels for N-N/2-N configuration

Table 10.4 — bsNumLFE

bsNumLFE	NumLfe
0	0
1	1
2	2
3	Reserved

Table 10.5 — Output channel ordering for N-N/2-N configuration

NumOutCh	NumLfe	Output channel ordering
24	2	Rv,Rb,Lv,Lb,Rs,Rvr,Lsr,Lvr,Rss,Rvss,Lss,Lvss,Rc,R, Lc,L,Ts,Cs,Cb,Cvr, C,LFE,Cv,LFE2,
14	0	L,Ls,R,Rs,Lbs,Lvs,Rbs,Rvs,Lv,Rv, Cv,Ts, C,LFE
12	1	L,Lv,R,Rv,Lsr,Lvr,Rsr,Rvr,Lss,Rss,C,LFE
12	2	L,Lv,R,Rv,Ls,Lss,Rs,Rss,C,LFE,Cvr,LFE2
10	1	L,Lv,R,Rv,Lsr,Lvr,Rsr,Rvr,C,LFE

NOTE 1 All of Names and layouts of loudspeaker follows the naming and position in ISO/IEC 23001-8:2013/FDAM1, Table 8.

NOTE 2 Output channel ordering for the case of 16, 20, 22, 26, 30 and 32 is following the arbitrary order from 1 to N without any specific naming of speaker layouts.

NOTE 3 Output channel ordering for the case when bsHasSpeakerConfig == 1 follows the order from 1 to N with associated naming of speaker layouts as specified in ISO/IEC 23008-3:2015, Table 94.

bsHasSpeakerConfig This flag indicates whether the output channels have a different layout than the output channel ordering specified in Table 10.5. If present (bsHasSpeaker-Config == 1), the loudspeaker layout of the output configuration "audioChan-nelLayout" can be used for rendering if the N-N/2-N system is used together with other MPEG standards (e.g. ISO/IEC 23008-3:2015).

audioChannelLayout This structure describes the loudspeaker layout of the output configuration. If the output configuration contains LFE channels, the LFE channels shall be ordered such that each LFE channel is processed together with one non-LFE channel using one OTT box and shall be positioned at the end of the channel list (e.g. L, Lv, R, Rv, Ls, Lss, Rs, Rss, C, LFE, Cvr, LFE2).

10.3 The N-N/2-N configuration

10.3.1 Introduction

In the following subclauses, the general structure for the N-N/2-N system is outlined. For this configuration, N/2 is identical to the number of downmix signals (*NumInCh* = N/2), denoted x_0 to $x_{NumInCh-1}$. Therefore, the number of output signals (i.e. N) should be an even number in order to process N/2 downmix signals, since the number of OTT boxes is equal to N/2.

The input vector to be multiplied by $\mathbf{M}_1^{n,k}$ is a vector containing the N/2 downmix channels. A maximum number of N/2 decorrelators can be used when LFE channels are not included in output channels. However, if the number of output channels exceeds twenty channels, the de-correlation filters are reused according to 10.7. Some of the decorrelator indices are repeated because the number of available decorrelators that ensure orthogonal decorrelated output signals is limited to 10, as defined in ISO/IEC 23003-1:2007. Therefore, the recommended number of output channels for the N-N/2-N configuration is less than 20 (or 24 with two Lfe channels).

The outputs of the decorrelators can be replaced by residual signals for certain frequency regions, depending on the bitstream. No decorrelation is used for the case of OTT based upmix when a LFE channel is one output of the OTT box. No residual signal can be inserted for these OTT boxes.

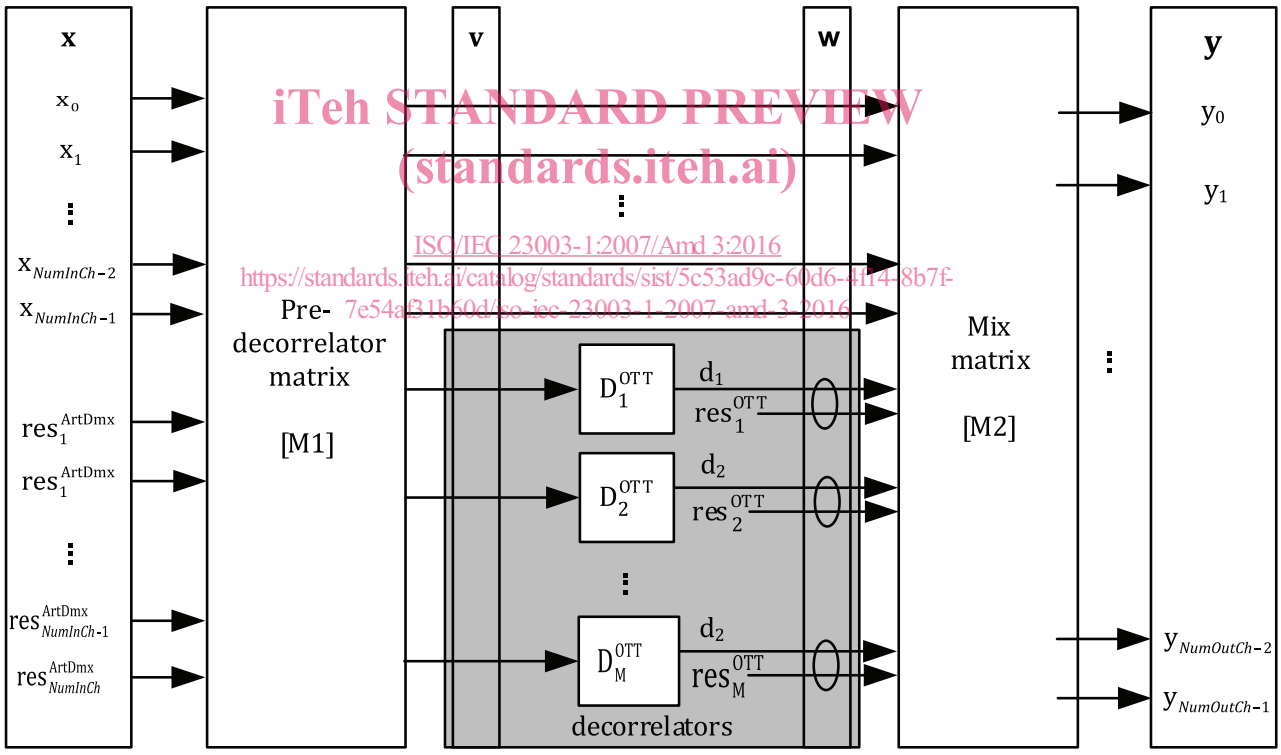

Figure 10.1 — Matrix view of the spatial audio processing for the N-N/2-N configuration

The decorrelators, decorrelated signals and residual signals in Figure 10.1 (labelled "1" to "M (i.e. NumInCh-NumLfe)") correspond to different OTT boxes depending on configuration.

The multi-channel reconstruction for the N-N/2-N configuration can also be visualized by means of a tree-structure. This is outlined in Figure 10.2. In Figure 10.2, every OTT box re-creates two channels based on one input channel, the corresponding CLD and ICC parameters, and residual signal. The OTT boxes and the corresponding data are numbered corresponding to the order they appear in the bitstream.