INTERNATIONAL STANDARD

ISO/IEC 14496-3

Second edition
2001-12-15

AMENDMENT 6
2005-07-01

Information technology — Coding of audio-visual objects —

Part 3: **Audio**

AMENDMENT 6: Lossless coding of oversampled audio

Technologies de l'information — Codage des objets audiovisuels —
Partie 3: Codage audio
AMENDEMENT 6: Codage sans perte d'audio suréchantillonné

Reference number
ISO/IEC 14496-3:2001/Amd.6:2005(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 6 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*.

Introduction

This document specifies the 6-th Amendment to the ISO/IEC 14496-3:2001 standard. It contains the text for the Final Draft Amendment on lossless coding of 1-bit oversampled audio signals. It contains the DSD and DST definitions as described in the Super Audio CD Specification Version 1.3. Note that in the context of SA-CD, only an oversampling ratio of 64 is defined. In this description, also oversampling ratios other than 64 are supported.

Information technology — Coding of audio-visual objects —

Part 3:
Audio

AMENDMENT 6: Lossless coding of oversampled audio

In ISO/IEC 14496-3:2001, Introduction, add:

MPEG-4 DST, Direct Stream Transfer for lossless coding of oversampled audio signals.

Amendment subpart 1

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

DST: Direct Stream Transfer

and increase the index-number of subsequent entries.

In Part 3: Audio, Subpart 1, in subclause 1.5.1.1 Audio object type definition, amend table 1.1 according to the update in the table below:

Table 1.1 — Audio object definition

Tools/ Modules Audio Object Type		SBR	SSC	DST	Remark	Object Type ID
Null						0
AAC main					2)	1
AAC LC						2
AAC SSR						3
AAC LTP					2)	4
SBR		X				5
AAC Scalable					6)	6
TwinVQ						7
CELP						8
HVXC						9
(Reserved)						10
(Reserved)						11
TTSI						12
Main synthetic					3)	13
Wavetable synthesis					4)	14
General MIDI						15
Algorithmic Synthesis and Audio FX						16
ER AAC LC						17
(Reserved)						18
ER AAC LTP					5)	19
ER AAC scalable					6)	20
ER TwinVQ						21
ER BSAC						22
ER AAC LD						23
ER CELP						24
ER HVXC						25
ER HILN						26
ER Parametric						27
SSC			X			28
(Reserved)						29
(Reserved)						30
(Reserved)						31
(Reserved)						32
(Reserved)						33
DST				X		35

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 Internet-working 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
SSC											28
(reserved)											29
(reserved)											30
(reserved)											31
(reserved)											32
(reserved)											33
DST											35

In Part 3: Audio, Subpart 1, in subclause 1.6.2.1 AudioSpecificConfig, adapt Table 1.8 according to the modification in the table below:

Table 1.8 — Syntax of AudioSpecificConfig()

Syntax	No. of bits	Mnemonic

AudioSpecificConfig ()
{
 	
---- 	
 	
if (audioObjectType == 26 || audioObjectType == 27)
ParametricSpecificConfig();
 	
if (audioObjectType == 35)
DSTSpecificConfig();
 	
if (audioObjectType == 17 || audioObjectType == 19 ||
audioObjectType == 20 || audioObjectType == 21 ||
audioObjectType == 22 || audioObjectType == 23 ||
audioObjectType == 24 || audioObjectType == 25 ||
audioObjectType == 26 || audioObjectType == 27) {
 	
---- 	
 	
}		

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
	18		
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	
SSC	28	ISO/IEC 14496-3 subpart 8	
…			
DST	35	ISO/IEC 14496-3 subpart 10	

Create Part 3: Audio, Subpart 10:

Subpart 10: Technical description of lossless coding of oversampled audio

10.1 Scope

This part of ISO/IEC 14496 describes the MPEG-4 lossless coding algorithm for oversampled audio signals.

10.2 Terms and definitions

The following definitions are used in this document.

Audio Channel	The stream of DSD bits intended for one loudspeaker.
Audio Frame	A Frame containing Audio data.
Audio Channel Number	The sequence number assigned to an Audio Channel. Audio Channel Numbers are contiguously numbered starting with one.
Frame	A block of data belonging to a certain Time Code. The playing time of a Frame is 1/75 Sec.
Reserved	All fields labelled Reserved are reserved for future standardization. All Reserved fields must be set to zero.
Silence Pattern	A digitally generated DSD pattern with the following properties:

- All Audio Bytes (see 10.6.1.1) have the same value.
- Each Audio Byte must contain 4 bits equal to zero and 4 bits equal to one.

Direct Stream Digital	A one bit oversampled representation of the audio signal.
Direct Stream Transfer	The lossless coding technique used for DSD signals in Super Audio CD.
DSD	See Direct Stream Digital.
DST	See Direct Stream Transfer.
Half Probability	Half Probability defines for each Audio Channel in an Audio Frame whether the first DSD bits are arithmetically encoded using the Ptable values, or using a probability equal to ½.
Mapping	Mapping defines, for each Segment, the Prediction Filter and Probability Table that is used.
Prediction Filter	A Prediction Filter is a transversal filter used to predict the value of the next DSD bit. A Prediction Filter is characterized by a prediction order and by coefficients.
Probability Table	A Probability Table contains the probability that the value of a DSD bit is predicted erroneously for a given output of the prediction filter.
Ptable	See Probability Table.

Sampling Frequency	The sampling frequency of the DSD signal shall be 64 * 44.1 kHz, 128 * 44.1 kHz or 256 * 44.1 kHz.
Segmentation	Each Audio Channel in an Audio Frame can be partitioned into Segments.

10.3 Conventions

In this document the conventions as described in this subclause are used.

10.3.1 Arithmetic and bit operations

a>>b	Right shift a over b bits. The new msb bits are set to '0'.
a<<b	Left shift a over b bits. The new lsb bits are set to '0'.
a | b	Bitwise OR of a and b.
a & b	Bitwise AND of a and b.
min(a,b)	Minimum value of a and b.
max(a,b)	Maximum value of a and b.
a mod b	Value of a modulo b.
trunc(a)	Value of a, rounded downwards.
| a |	Absolute value of a.
a==b	Evaluate if a is equal to b.
a!=b	Evaluate if a is not equal to b.
a=b	Variable a is set to the value of b.
a++	a = a + 1.
a -= b	a = a – b.
a += b	a = a + b.

10.3.2 Bit ordering

The graphical representation of all multiple-bit quantities is such that the most significant bit (msb) is on the left, and the least significant bit (lsb) is on the right. Figure 10.1 defines the bit position in a Byte.

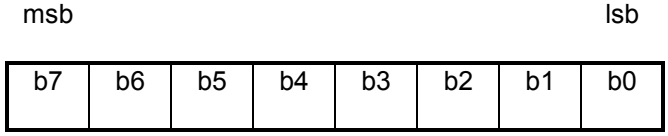

Figure 10.1 — Bit ordering in a Byte

10.3.3 Bit sequence

In all places where a bit sequence is used, a most significant bit first notation is used.

10.3.4 Decimal notation

All Decimal values are preceded by a blank space or the range indicator (..) when included in a range. The most significant digit is on the left, the least significant digit is on the right.

10.3.5 DSD bit order

The first sampled DSD bit is stored in the most significant bit of a byte. See subclause 10.6.1.1.

10.3.6 DSD Polarity

A DSD bit equal to one means "plus". A DSD bit equal to zero means "minus".

10.3.7 Hex notation

All Hexadecimal values are preceded by a $. The most significant nibble is on the left, the least significant nibble is on the right.

10.3.8 Range

Constant_1..Constant_2 denotes the range from and including Constant_1 up to and including Constant_2, in increments of 1.

10.3.9 Until

Until is used in figures to indicate that for a structure Byte Positions are used upto (not including) a given value.

At Byte Position B1, the expression “until B2” specifies B2-B1 bytes. At Byte Position B1, the expression “until eos” specifies the number of bytes from B1 up to and including the last byte of the current Sector. Note that Byte Position is specified relative to the start of the current, or a previous, Sector.

10.4 Basic Types

10.4.1 BsMsbf

Bit Sequence, Most Significant Bit First, must be interpreted as a Bit String.

10.4.2 Char

A one-byte character, encoded according to ISO 646. The NUL ($00) character is not allowed for Char.

10.4.3 SiMsbf

Bit sequence, Most Significant Bit First, must be interpreted as Signed Integer using two’s complement notation.

10.4.4 UiMsbf

Bit sequence, Most Significant Bit First, must be interpreted as Unsigned Integer.

10.4.5 Uintn

An n bit, binary encoded, unsigned numerical value.

10.4.6 Uint8

An 8 bit, binary encoded, unsigned numerical value. A Uint8 value must be recorded in a one-byte field.

10.4.7 Uint16

A 16-bit, binary encoded, unsigned numerical value. A Uint16 value, represented by the hexadecimal representation $wxyz, must be recorded in a two-byte field as $wx $yz (most significant byte first).

10.4.8 Uint32

A 32-bit, binary encoded, unsigned numerical value. A Uint32 value, represented by the hexadecimal representation $stuvwxyz, must be recorded in a four-byte field as $st $uv $wx $yz (most significant byte first).

10.5 Payloads for the audio object

Table 10.1 — Syntax of Audio_Frame()

Syntax	Bits	Mnemonics

DSTSpecificConfig(channelConfiguration) {
if (DSD_Coded)
{
DSD()
}
if (DST_Coded)
{
DST()
}
}
		
		DSD
		
		DST

Table 10.2 – Syntax of DSD

Syntax	Bits	Mnemonics

DSD() {
For (Byte_Nr=0; Byte_Nr<Frame_Length; Byte_Nr++)
{
For (Channel_Nr=1; Channel_Nr<=N_Channels;
Channel_Nr++)
{
DSD_Byte[Channel_Nr][Byte_Nr]
}
}
}
	
	1
		
		Audio_Byte