

## Content for Subpart 1

1.1	Scope .....	3
1.2	Normative references .....	3
1.3	Terms and definitions .....	4
1.4	Symbols and abbreviations .....	16
1.4.1	Arithmetic operators .....	16
1.4.2	Logical operators .....	17
1.4.3	Relational operators .....	17
1.4.4	Bitwise operators .....	18
1.4.5	Assignment .....	18
1.4.6	Mnemonics .....	18
1.4.7	Constants .....	18
1.4.8	Method of describing bitstream payload syntax .....	19
1.4.9	Arithmetic data types .....	20
1.5	Technical overview .....	21
1.5.1	MPEG-4 audio object types .....	21
1.5.2	Audio profiles and levels .....	26
1.6	Interface to ISO/IEC 14496-1 (MPEG-4 Systems) .....	38
1.6.1	Introduction .....	38
1.6.2	Syntax .....	38
1.6.3	Semantics .....	44
1.6.4	Upstream .....	48
1.6.5	Signaling of SBR .....	50
1.6.6	Signaling of Parametric Stereo (PS) .....	53
1.6.7	Interface between Audio and Systems .....	57
1.6.8	Signalling of BSAC extension payloads .....	58
1.6.9	Original Audio File Information for the ISO Base Media file format .....	59
1.7	MPEG-4 Audio transport stream .....	63
1.7.1	Overview .....	63
1.7.2	Synchronization Layer .....	64
1.7.3	Multiplex Layer .....	67
1.8	Error protection .....	78
1.8.1	Overview of the tools .....	78
1.8.2	Syntax .....	82
1.8.3	General information .....	85
1.8.4	Tool description .....	88
Annex 1.A	(informative) Audio Interchange Formats .....	106
1.A.1	Introduction .....	106
1.A.2	AAC Interchange formats .....	106
1.A.3	Syntax .....	106
1.A.4	Semantic .....	109
Annex 1.B	(informative) Error protection tool .....	111
1.B.1	Example of out-of-band information .....	111
1.B.2	Example of error concealment .....	114
1.B.3	Example of EP tool setting and error concealment for HVXC .....	119
Annex 1.C	(informative) Patent statements .....	135

## Content for Subpart 2

2.1	Scope .....	3
2.2	Terms and definitions .....	3
2.3	Bitstream syntax .....	3
2.3.1	Decoder configuration (HvxcSpecificConfig) .....	4
2.3.2	Bitstream frame (alPduPayload) .....	5
2.3.3	Decoder configuration (ErrorResilientHvxcSpecificConfig) .....	8
2.3.4	Bitstream frame (alPduPayload) .....	9
2.4	Bitstream semantics .....	21
2.4.1	Decoder configuration (HvxcSpecificConfig,ErrorResilientHvxcSpecificConfig) .....	21
2.4.2	Bitstream frame (alPduPayload) .....	21
2.5	HVXC decoder tools .....	22
2.5.1	Overview .....	22
2.5.2	LSP decoder .....	25
2.5.3	Harmonic VQ decoder .....	30
2.5.4	Time domain decoder .....	35
2.5.5	Parameter interpolation for speed control .....	37
2.5.6	Voiced component synthesizer .....	41
2.5.7	Unvoiced component synthesizer .....	54
2.5.8	Variable rate decoder .....	57
2.5.9	Extension of HVXC variable rate mode .....	59
Annex 2.A	(informative) HVXC Encoder tools .....	63
2.A.1	Overview of encoder tools .....	63
2.A.2	Normalization .....	63
2.A.3	Pitch estimation .....	68
2.A.4	Harmonic magnitudes extraction .....	70
2.A.5	Perceptual weighting .....	71
2.A.6	Harmonic VQ encoder .....	71
2.A.7	V/UV decision .....	73
2.A.8	Time domain encoder .....	74
2.A.9	Variable rate encoder .....	76
2.A.10	Extension of HVXC variable rate encoder .....	79
Annex 2.B	(informative) HVXC Decoder tools .....	84
2.B.1	Postfilter .....	84
2.B.2	Post processing .....	86
Annex 2.C	(informative) System layer definitions .....	88
2.C.1	Random access point .....	88
Annex 2.D	(informative) Example of EP tool setting and error concealment for HVXC .....	89
2.D.1	Overview .....	89
2.D.2	EP tool setting .....	89
2.D.3	Error concealment .....	92
Annex 2.E	(normative) VQ codebooks for HVXC .....	95
2.E.1	List of the VQ codebooks .....	95
2.E.2	CbAm .....	95
2.E.3	CbAm4k .....	102
2.E.4	CbCelp .....	133
2.E.5	CbCelp4k .....	143
2.E.6	CbLsp .....	146
2.E.7	CbLsp4k .....	150

## Contents for Subpart 3

3.1	Scope .....	2
3.1.1	General description of the CELP decoder.....	2
3.1.2	Functionality of MPEG-4 CELP.....	2
3.2	Terms and Definitions .....	5
3.3	Bitstream syntax .....	6
3.3.1	CELP object type .....	6
3.3.2	ER-CELP object type .....	11
3.4	Semantics .....	25
3.4.1	Header semantics .....	25
3.4.2	Frame semantics.....	28
3.5	MPEG-4 CELP Decoder tools .....	33
3.5.1	General Introduction to the MPEG-4 CELP decoder tool-set .....	34
3.5.2	AAC/CELP scalable configuration.....	35
3.5.3	Helping variables .....	35
3.5.4	Bitstream elements for the MPEG-4 CELP decoder tool-set.....	36
3.5.5	CELP bitstream demultiplexer.....	36
3.5.6	CELP LPC decoder and interpolator.....	37
3.5.7	CELP excitation generator .....	57
3.5.8	CELP LPC synthesis filter.....	80
3.5.9	CELP silence compression tool .....	81
Annex 3.A	(informative) MPEG-4 CELP decoder tools .....	89
3.A.1	CELP post-processor .....	89
Annex 3.B	(informative) MPEG-4 CELP encoder tools .....	92
3.B.1	General Introduction to the MPEG-4 CELP encoder tool-set.....	92
3.B.2	Helping variables .....	92
3.B.3	Bitstream elements for the MPEG-4 CELP encoder tool-set.....	94
3.B.4	CELP preprocessing.....	94
3.B.5	CELP LPC analysis .....	95
3.B.6	CELP LPC quantizer and interpolator.....	97
3.B.7	CELP LPC analysis filter .....	106
3.B.8	CELP weighting module.....	107
3.B.9	CELP excitation analysis .....	108
3.B.10	CELP bitstream multiplexer .....	123
3.B.11	CELP silence compression tool .....	123
Annex 3.C	(normative) Tables.....	129
3.C.1	LSP VQ tables and gain VQ tables for 8 kHz sampling rate .....	129
3.C.2	LSP VQ tables and gain VQ tables for the 16 kHz sampling rate.....	137
3.C.3	Gain tables for the bitrate scalable tool .....	153
3.C.4	LSP VQ tables and gain VQ tables for the bandwidth scalable tool.....	153
Annex 3.D	(informative) Tables .....	165
3.D.1	Bandwidth expansion tables in LPC analysis of the mode II coder .....	165
3.D.2	Downsampling filter coefficients for the bandwidth scalable tool .....	165
Annex 3.E	(informative) Example of a simple CELP transport stream.....	167
Annex 3.F	(informative) Random access points.....	169

## Contents for Subpart 4

4.1	Scope .....	3
4.1.1	Technical Overview .....	3
4.2	Normative references .....	11
4.3	Terms and definitions .....	11
4.4	Syntax .....	11
4.4.1	Decoder configuration (GASpecificConfig) .....	11
4.4.2	GA bitstream payloads .....	13
4.5	Overall data structure .....	53
4.5.1	Decoding of the GA specific configuration .....	53
4.5.2	Decoding of the GA bitstream payloads .....	56
4.5.3	Buffer requirements .....	123
4.5.4	Tables .....	124
4.5.5	Figures .....	138
4.6	GA-Tool Descriptions .....	139
4.6.1	Quantization .....	139
4.6.2	Scalefactors .....	140
4.6.3	Noiseless coding .....	142
4.6.4	Noiseless coding for the fine grain scalability .....	149
4.6.5	Interleaved vector quantization .....	158
4.6.6	Frequency domain prediction .....	163
4.6.7	Long term prediction (LTP) .....	163
4.6.8	Joint Coding .....	166
4.6.9	Temporal noise shaping (TNS) .....	173
4.6.10	Spectrum normalization .....	177
4.6.11	Filterbank and block switching .....	188
4.6.12	Gain Control .....	193
4.6.13	Perceptual noise substitution (PNS) .....	200
4.6.14	Frequency selective switch (FSS) Module .....	203
4.6.15	Upsampling filter tool .....	206
4.6.16	Tools for AAC error resilience .....	207
4.6.17	Low delay codec .....	217
4.6.18	SBR tool .....	220
4.6.19	Low Delay SBR .....	269
4.6.20	Enhanced Low Delay Codec .....	276
Annex 4.A (normative)	Normative Tables .....	280
4.A.1	Huffman codebook tables for AAC-type noiseless coding .....	280
4.A.2	Window tables .....	293
4.A.3	Differential scalefactor to index tables .....	317
4.A.4	Tables for TwinVQ .....	317
4.A.5	Tables for ER BSAC .....	339
4.A.6	Tables for SBR .....	348

<b>Annex 4.B (informative) Encoder tools .....</b>	<b>373</b>
4.B.1 Psychoacoustic model.....	373
4.B.2 Gain control.....	373
4.B.3 Filterbank and block switching .....	373
4.B.4 Frequency domain prediction .....	373
4.B.5 Temporal noise shaping (TNS).....	373
<b>4.B.6 Contents for Subpart 5.....</b>	<b>373</b>
4.B.7 Quantization .....	373
4.B.8.1 Noise shaping.....	373...2
4.B.8.2 Feature Normalization and dynamic range control.....	373...2
4.B.9 Long term predictions.....	373...2
4.B.10 Perceptual Noise Substitution (PNS).....	375...3
4.B.11 Random access points for GA coded bitstreams payloads.....	375...4
4.B.12 Weighted interleaved vector quantization.....	376...9
4.B.13 Object types.....	376...9
4.B.14 Spectrum normalization.....	378...10
4.B.15 Scalable AAC with core coder.....	382...17
4.B.16 SAOL syntax and semantics.....	384...58
4.B.17 SAOL core opcode definitions and semantics.....	384...98
4.B.18 Fine-grain scalability: FSAC (Bit-Sliced Arithmetic Coding) .....	394...105
4.B.19 Informative SBR encoder description .....	400...108
4.B.20 SAOL/SASL tokenisation .....	400...110
4.B.21 Enhanced Low Delay Codec.....	401...112
4.B.22 Sample Bark syntax and semantics.....	401...112
4.B.23 Delayless mixing using AAC ELD .....	401...112
4.B.24 MIDI semantics.....	401...112
4.B.25 Input sounds and relationship with AudioBIFS.....	401...112
5.15 Input sounds and relationship with AudioBIFS.....	117
<b>Annex 5.A (normative) Coding tables.....</b>	<b>120</b>
<b>Annex 5.B (informative) Encoding .....</b>	<b>123</b>
<b>Annex 5.C (informative) lex/yacc grammars for SAOL .....</b>	<b>126</b>
<b>Annex 5.D (informative) PICOLA Speed change algorithm .....</b>	<b>133</b>
<b>Annex 5.E (informative) Random access to Structured audio bitstreams.....</b>	<b>136</b>
<b>Annex 5.F (informative) Directly-connected MIDI and microphone control of the orchestra .....</b>	<b>141</b>
<b>Bibliography .....</b>	<b>143</b>
<b>Alphabetical Index to Subpart 5 of ISO/IEC 14496-3 .....</b>	<b>144</b>

## Contents for Subpart 5

5.1	Scope .....	2
5.2	Normative references .....	2
5.3	Terms and definitions .....	2
5.4	Symbols and abbreviations .....	3
5.5	Bitstream syntax and semantics .....	4
5.6	Object types .....	9
5.7	Decoding process .....	10
5.8	SAOL syntax and semantics .....	17
5.9	SAOL core opcode definitions and semantics .....	58
5.10	SAOL core wavetable generators .....	98
5.11	SASL syntax and semantics .....	105
5.12	SAOL/SASL tokenisation .....	108
5.13	Sample Bank syntax and semantics .....	110
5.14	MIDI semantics .....	112
5.15	Input sounds and relationship with AudioBIFS .....	117
Annex 5.A (normative) Coding tables .....		120
Annex 5.B (informative) Encoding .....		123
Annex 5.C (informative) lex/yacc grammars for SAOL .....		126
Annex 5.D (informative) PICOLA Speed change algorithm .....		133
Annex 5.E (informative) Random access to Structured audio bitstreams .....		136
Annex 5.F (informative) Directly-connected MIDI and microphone control of the orchestra .....		141
Bibliography .....		143
Alphabetical Index to Subpart 5 of ISO/IEC 14496-3 .....		144

## Contents for Subpart 6

6.1	Scope .....	2
6.2	Terms and definitions .....	2
6.3	Symbols and abbreviations .....	2
6.4	MPEG-4 audio text-to-speech bitstream syntax .....	2
6.4.1	MPEG-4 audio TTSSpecificConfig .....	2
6.4.2	MPEG-4 audio text-to-speech payload .....	3
6.5	MPEG-4 audio text-to-speech bitstream semantics .....	5
6.5.1	MPEG-4 audio TTSSpecificConfig .....	5
6.5.2	MPEG-4 audio text-to-speech payload .....	5
6.6	MPEG-4 audio text-to-speech decoding process .....	7
6.6.1	Interface between DEMUX and syntactic decoder .....	7
6.6.2	Interface between syntactic decoder and speech synthesizer .....	8
6.6.3	Interface from speech synthesizer to compositor .....	8
6.6.4	Interface from compositor to speech synthesizer .....	8
6.6.5	Interface between speech synthesizer and phoneme/bookmark-to-FAP converter .....	9
Annex 6.A (informative)	Applications of MPEG-4 audio text-to-speech decoder .....	10
6.A.1	General .....	10
6.A.2	Application scenario: MPEG-4 Story Teller on Demand (STOD) .....	10
6.A.3	Application scenario: MPEG-4 audio text-to-speech with moving picture .....	10
6.A.4	MPEG-4 audio TTS and face animation using bookmarks appropriate for trick mode .....	10
6.A.5	Random access unit .....	10

## Contents for Subpart 7

7.1	Scope .....	2
7.1.1	Technical Overview .....	2
7.2	Terms and definitions .....	3
7.3	Bitstream syntax.....	3
7.3.1	Decoder configuration (ParametricSpecificConfig).....	3
7.3.1.1	Parametric Audio decoder configuration.....	4
7.3.1.2	HILN decoder configuration .....	4
7.3.2	Bitstream frame (sIPacketPayload) .....	6
7.3.2.1	Parametric Audio bitstream frame .....	7
7.3.2.2	HILN bitstream frame .....	9
7.3.2.3	HILN codebooks .....	20
7.3.2.4	HILN SubDivisionCode (SDC).....	25
7.4	Bitstream semantics.....	26
7.4.1	Decoder configuration (ParametricSpecificConfig).....	26
7.4.1.1	Parametric Audio decoder configuration.....	26
7.4.1.2	HILN decoder configuration .....	26
7.4.2	Bitstream frame (sIPacketPayload) .....	26
7.4.2.1	Parametric Audio bitstream frame .....	26
7.4.2.2	HILN bitstream frame .....	26
7.5	Parametric decoder tools.....	29
7.5.1	HILN decoder tools .....	29
7.5.1.1	Harmonic line decoder.....	31
7.5.1.2	Individual line decoder.....	34
7.5.1.3	Noise decoder .....	37
7.5.1.4	Harmonic and individual line synthesizer .....	38
7.5.1.5	Noise synthesizer .....	45
7.5.2	Integrated parametric coder .....	49
7.5.2.1	Integrated parametric decoder.....	49
7.6	Error resilient bitstream payloads .....	49
7.6.1	Overview of the tools .....	49
7.6.2	ER HILN.....	50
Annex 7.A (informative)	Parametric audio encoder .....	52
7.A.1	Overview of the encoder tools .....	52
7.A.2	HILN encoder tools .....	52
7.A.2.1	HILN parameter extraction.....	52
7.A.2.2	HILN parameter encoder .....	56
7.A.2.3	HILN bitrate scalability .....	59
7.A.3	Music/Speech Mixed Encoder tool.....	59
7.A.3.1	Music/Speech classification tool .....	60
7.A.3.2	Integrated parametric coder .....	61

## Contents for Subpart 8

8.1	Scope .....	2
8.2	Terms and definitions .....	2
8.3	Symbols and abbreviations .....	4
8.3.1	Arithmetic operators .....	4
8.3.2	Relation operators .....	4
8.3.3	Mnemonics .....	4
8.3.4	Ranges .....	4
8.3.5	Number notation .....	4
8.3.6	Definitions .....	5
8.4	Payloads for the audio object type SSC .....	5
8.4.1	Decoder configuration (SSCSpecificConfig) .....	5
8.4.2	SSC Bitstream Payload .....	5
8.5	Semantics .....	13
8.5.1	SSCSpecificConfig .....	13
8.5.2	Decoding of SSC Bitstream Payload .....	14
8.5.3	Indexing of sub-frames .....	25
8.6	Decoding processes .....	25
8.6.1	Transients .....	26
8.6.2	Sinusoids .....	28
8.6.3	Noise .....	35
8.6.4	Parametric stereo .....	42
8.6.5	Start/stop situations for decoding .....	63
8.7	References .....	64
Annex 8.A (normative)	Combination of the SBR tool with the parametric stereo tool .....	65
8.A.1	Overview .....	65
8.A.2	Bitstream syntax and semantics .....	65
8.A.3	Decoding process .....	65
8.A.4	Baseline version of the parametric stereo coding tool .....	66
Annex 8.B (normative)	Normative Tables .....	68
8.B.1	Huffman tables for SSC .....	68
Annex 8.C (informative)	Encoder description .....	93
8.C.1	Technical overview .....	93
8.C.2	Audio read and filtering .....	95
8.C.3	Transients .....	95
8.C.4	Sinusoids .....	97
8.C.5	Noise .....	102
8.C.6	Parametric stereo .....	106
8.C.7	Bit-Stream Formatter .....	109
Annex 8.D (informative)	Tempo and pitch scaling in decoder .....	114
8.D.1	Continuous phase .....	114
8.D.2	Tempo scaling for stereo .....	115

## Contents for Subpart 9

9.1	Scope .....	2
9.2	MPEG_1_2_SpecificConfig .....	2
9.3	Channel Mapping .....	2
9.4	Access Unit Format .....	2
9.4.1	Layer 1 and 2 .....	2
9.4.2	Layer 3 .....	3
9.5	Sampling rate extension for Layer 3 .....	3
9.5.1	Bitrates .....	3
9.5.2	Sampling frequency .....	4
9.5.3	Padding .....	4
9.5.4	Scalefactor bands .....	4
Annex 9.A	(normative) Scalefactor band tables .....	5
Annex 9.B	(informative) Converting MPEG-1/2 Layer 3 bitstreams into mp3_channel_elements .....	8
Annex 9.C	(informative) Converting mp3_channel_elements into MPEG-1/2 Layer 3 bitstreams .....	9
9.C.1	Overview .....	9
9.C.2	Sampling rate signaling .....	9
9.C.3	Reconstruction instructions .....	9
Annex 9.D	(informative) Legacy MPEG-4 Systems Interface to MPEG-1/2 Audio .....	12
9.D.1	Overview .....	12
9.D.2	Decoder Specific Information .....	12
9.D.3	Access Units .....	12

## Contents for Subpart 10

10.1	Scope .....	2
10.2	Terms and definitions .....	2
10.3	Conventions .....	3
10.3.1	Arithmetic and bit operations.....	3
10.3.2	Bit ordering .....	3
10.3.3	Bit sequence .....	3
10.3.4	Decimal notation.....	4
10.3.5	DSD bit order.....	4
10.3.6	DSD Polarity.....	4
10.3.7	Hex notation.....	4
10.3.8	Range.....	4
10.3.9	Until.....	4
10.4	Basic Types.....	4
10.4.1	BsMsbf.....	4
10.4.2	Char.....	4
10.4.3	SiMsbf.....	4
10.4.4	UiMsbf.....	4
10.4.5	Uintn.....	4
10.4.6	Uint8.....	5
10.4.7	Uint16.....	5
10.4.8	Uint32.....	5
10.5	Payloads for the audio object .....	5
10.6	Semantics.....	10
10.6.1	Audio Streams .....	10
10.6.1.1	DSD Sampled Bit Stream Data.....	11
10.6.1.2	Structure of the Audio Stream.....	11
10.6.1.3	Audio_Frame .....	11
10.7	DST Decoder Reference Model (Normative).....	24
10.7.1	DST Decoder Block Diagrams.....	24
10.7.2	DST Decoding Processes .....	25
10.7.2.1	Introduction .....	25
10.7.2.2	Arithmetic Decoder.....	25
10.7.2.3	Source Model.....	29
10.7.2.4	Multiplexing/Demultiplexing .....	31
10.7.3	Restrictions to DST coded Audio_Frames (Normative) .....	32
10.7.3.1	Limited number of erroneously predicted samples .....	32
10.7.3.2	Probability table design requirement .....	32
Annex 10.A	(informative) Encoder description.....	34
10.A.1	Technical overview.....	34
10.A.1.1	Framing .....	34
10.A.1.2	Prediction.....	35
10.A.1.3	Arithmetic encoding .....	35
10.A.1.4	Channel multiplexing.....	36
10.A.1.5	References.....	36

## Contents for Subpart 11

11.1	Scope .....	2
11.2	Technical Overview .....	2
11.2.1	Encoder and Decoder Structure .....	2
11.2.2	Floating-Point Extensions .....	3
11.3	Terms and Definitions .....	3
11.3.1	Definitions .....	3
11.3.2	Mnemonics .....	4
11.3.3	Data Types.....	4
11.4	Syntax .....	4
11.4.1	Decoder Configuration.....	4
11.4.2	Bitstream Payloads .....	5
11.4.3	Payloads for Floating-Point Data .....	10
11.5	Semantics.....	12
11.5.1	General Semantics .....	12
11.5.2	Semantics for Floating-Point Data.....	17
11.6	ALS Tools .....	19
11.6.1	Overview.....	19
11.6.2	Block Switching .....	23
11.6.3	Prediction .....	25
11.6.4	Long-term prediction (LTP) .....	32
11.6.5	RLS-LMS Predictor.....	34
11.6.6	Coded Residual.....	49
11.6.7	Joint Coding of Channel Pairs .....	62
11.6.8	Multi-Channel Coding (MCC).....	62
11.6.9	Extension for Floating-Point Data.....	66

# Contents for Subpart 12

12.1	Scope .....	2
12.2	Terms and definitions .....	2
12.2.1	Definitions .....	2
12.2.2	Notations .....	3
12.2.3	Definitions .....	3
12.3	Payloads for the audio object .....	3
12.4	Semantics .....	6
12.5	SLS decoder tool .....	7
12.5.1	Overview .....	7
12.5.2	Oversampling technique .....	8
12.5.3	SLS with Scalable AAC Core .....	9
12.5.4	Decoding of lle_single_channel_element (LLE_SCE) and lle_channel_pair_element (LLE_CPE) .....	10
12.5.5	Decoding of lle_data .....	14
12.5.6	Compensation for IntMDCT residual for early terminating BPGC/CBAC decoding .....	23
12.5.7	Inverse Error Mapping .....	23
12.5.8	Integer Mid/Side process .....	24
12.5.9	Integer Temporal Noise Shaping (IntTNS) .....	25
12.5.10	IntMDCT and Inverse IntMDCT .....	27
12.5.11	Computation of table values based on compact tables .....	44
Annex 12.A	(Informative) Encoder Description .....	47
12.A.1	Overview .....	47
12.A.2	Integer MDCT .....	48
12.A.3	Grouping and Interleaving .....	48
12.A.4	Integer Mid/Side .....	49
12.A.5	Normalize before AAC Coding .....	49
12.A.6	Error Mapping .....	50
12.A.7	BPGC/CBAC Encoder .....	50
12.A.8	Method of bitstream truncation by re-parsing the bitstream .....	53
Annex 12.B	(Normative) Tables .....	55
12.B.1	Tables for pre-defined fixed-point coefficients .....	55