ISO/IEC 23003-3:2020-06 (E)

Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding

Contents

Page

Foreword vii

Introduction viii

1 Scope 1

2 Normative references 1

3 Terms, definitions, symbols and abbreviated terms 1
3.1 Terms and definitions 1
3.2 Symbols and abbreviated terms 3

4 Technical overview 4
4.1 Decoder block diagram 4
4.2 Overview of the decoder tools 5
4.3 Combination of USAC with MPEG Surround and SAOC 9
4.4 Interface between USAC and systems 9
4.4.1 Decoder behaviour 10
4.5 USAC profiles and levels 10
4.5.1 General 10
4.5.2 MPEG-4 HE AACv2 compatibility 11
4.5.3 Baseline USAC profile 12
4.5.4 Extended high efficiency AAC profile 13
4.6 Combination of USAC with MPEG-D DRC 14

5 Syntax 15
5.1 General 15
5.2 Decoder configuration (UsacConfig) 15
5.3 USAC bitstream payloads 20
5.3.1 Payloads for audio object type USAC 20
5.3.2 Subsidiary payloads 23
5.3.3 Payloads for enhanced SBR 34
5.3.4 Payloads for MPEG Surround 43
5.3.5 Payload of extension elements 53

6 Data structure 53
6.1 USAC configuration 53
6.1.1 Definition of elements 53
6.1.2 UsacConfig() 63
6.1.3 Usac Output Sampling Frequency 63
6.1.4 UsacChannelConfig() 63
6.1.5 UsacDecoderConfig() 64
6.1.6 UsacSingleChannelElementConfig() 64
6.1.7 UsacChannelPairElementConfig() 64
6.1.8 UsacLfeElementConfig() 64
6.1.9 UsacCoreConfig() 64
6.1.10 SbrConfig() 65
6.1.11 SbrDfltHeader() 65
6.1.12 Mps212Config() 65
6.1.13 UsacExtElementConfig() 65
6.1.14 UsacConfigExtension() 65

6.1.15 Unique stream identifier (Stream ID) 66
6.2 USAC payload 66
6.2.1 Definition of elements 66
6.2.2 UsacFrame() 68
6.2.3 UsacSingleChannelElement() 69
6.2.4 UsacExtElement() 69
6.2.5 UsacChannelPairElement() 70
6.2.6 Low frequency enhancement (LFE) channel element, UsacLfeElement() 70
6.2.7 UsacCoreCoderData() 71

6.2.8 StereoCoreToolInfo() 71
6.2.9 fd_channel_stream() and ics_info() 72
6.2.10 lpd_channel_stream() 76
6.2.11 Spectral noiseless coder 79
6.2.12 Enhanced SBR 80
6.2.13 Definition of MPEG Surround 2-1-2 payloads 82
6.2.14 Buffer requirements 84

7 Tool descriptions 85
7.1 Quantization 85
7.1.1 Tool description 85
7.1.2 Definition of elements 85
7.1.3 Decoding process 85
7.2 Noise filling 85
7.2.1 Tool description 85
7.2.2 Definition of elements 86
7.2.3 Decoding process 86
7.2.4 Generation of random signs for spectral noise filling 87
7.3 Scale factors 87
7.4 Spectral noiseless coding 87
7.4.1 Tool description 87
7.4.2 Definition of elements 88
7.4.3 Decoding process 89
7.5 enhanced SBR tool (eSBR) 93
7.5.1 Modifications to SBR tool 93
7.5.2 Additional pre-processing in the MPEG-4 SBR within USAC 108
7.5.3 DFT based harmonic transposer 110
7.5.4 QMF based harmonic transposer 120
7.5.5 4:1 Structure for SBR in USAC 128
7.5.6 Predictive vector coding (PVC) decoding process 138
7.6 Inter-subband-sample temporal envelope shaping (inter-TES) 141
7.6.1 Tool Description 141
7.6.2 Definition of elements 142
7.6.3 Inter-TES 142
7.7 Joint stereo coding 144
7.7.1 M/S stereo 144
7.7.2 Complex stereo prediction 144
7.8 TNS 151
7.8.1 General 151
7.8.2 Definition of elements 151
7.8.3 Decoding process 152
7.8.4 Maximum TNS bandwidth 152
7.9 Filterbank and block switching 153
7.9.1 Tool description 153
7.9.2 Definition of elements 153
7.9.3 Decoding process 153
7.10 Time-warped filterbank and blockswitching 161
7.10.1 Tools description 161
7.10.2 Definition of elements 161
7.10.3 Decoding process 163
7.11 MPEG Surround for mono to stereo upmixing 169
7.11.1 Tool description 169

7.11.2 Decoding process 170
7.12 AVQ decoding 182
7.13 LPC-filter 189
7.13.1 Tool description 189
7.13.2 Definition of elements 189
7.13.3 Number of LPC filters 189
7.13.4 General principle of the inverse quantizer 189
7.13.5 Decoding of the LPC quantization mode 190
7.13.6 First-stage approximation 191
7.13.7 AVQ refinement 191
7.13.8 Reordering of quantized LSFs 193
7.13.9 Conversion into LSP parameters 193
7.13.10 Interpolation of LSP parameters 194
7.13.11 LSP to LP conversion 194
7.13.12 LPC initialization at decoder start-up 195
7.14 ACELP 196
7.14.1 General 196
7.14.2 Definition of elements 196
7.14.3 ACELP initialization at USAC decoder start-up 197
7.14.4 Setting of the ACELP excitation buffer using the past FD synthesis and LPC0 197
7.14.5 Decoding of CELP excitation 197
7.14.6 Excitation postprocessing 203
7.14.7 Synthesis 204
7.14.8 Writing in the output buffer 204
7.15 MDCT based TCX 205
7.15.1 Tool description 205
7.15.2 Decoding process 205
7.16 Forward aliasing cancellation (FAC) tool 209
7.16.1 Tool description 209
7.16.2 Definition of elements 209
7.16.3 Decoding process 210
7.16.4 Writing in the output buffer 211
7.17 Post-processing of the synthesis signal 212
7.18 Audio pre-roll 214
7.18.1 General 214
7.18.2 Semantics 214
7.18.3 Decoding process 215

8 Conformance testing 217
8.1 General 217
8.2 USAC conformance testing 217
8.2.1 Profiles 217
8.2.2 Conformance tools and test procedure 218
8.3 USAC bitstreams 222
8.3.1 General 222
8.3.2 USAC configuration 222
8.3.3 Framework 225
8.3.4 Frequency domain coding (FD mode) 226
8.3.5 Linear predictive domain coding (LPD mode) 228
8.3.6 Common core coding tools 229
8.3.7 Enhanced spectral band replication (eSBR) 230
8.3.8 eSBR - Predictive vector coding (PVC) 232
8.3.9 eSBR - Inter temporal envelope shaping (inter-TES) 233
8.3.10 MPEG Surround 2-1-2 233
8.3.11 Configuration Extensions 235
8.3.12 AudioPreRoll 235
8.3.13 DRC 236
8.3.14 Restrictions depending on profiles and levels 236
8.4 USAC decoders 238
8.4.1 General 238
8.4.2 FD core mode tests 238
8.4.3 LPD core mode tests 245

8.4.4 Combined core coding tests 250
8.4.5 eSBR tests 251
8.4.6 MPEG Surround 212 tests 260
8.4.7 Bitstream extensions 262
8.5 Decoder settings 265
8.5.1 General 265
8.5.2 Target loudness [Lou-<x>] 265
8.5.3 DRC effect type request [Eff-<x>] 265
8.6 Decoding of MPEG-4 file format parameters to support exact time alignment in file-to-file coding 265

9 Reference software 266
9.1 Reference software structure 266
9.1.1 General 266
9.1.2 Copyright disclaimer for software modules 266
9.2 Bitstream decoding software 268
9.2.1 General 268
9.2.2 USAC decoding software 268

Annex A (normative) Tables 269

Annex B (informative) Encoder tools 274

Annex C (normative) Tables for arithmetic decoder 314

Annex D (normative) Tables for predictive vector coding 320

Annex E (informative) Adaptive time/frequency post-processing 329

Annex F (informative) Audio/systems interaction 335

Annex G (informative) Reference software 337

Annex H (normative) Carriage of MPEG-D USAC in ISO base media file format 338

Bibliography 339