

# ETSI TS 126 253 V18.8.0 (2026-03)



TECHNICAL SPECIFICATION

**LTE; 5G;  
Codec for Immersive Voice and Audio Services (IVAS);  
Detailed Algorithmic Description including RTP payload format  
and SDP parameter definitions  
(3GPP TS 26.253 version 18.8.0 Release 18)**



---

**Reference**

RTS/TSGS-0426253vi80

---

**Keywords**

5G,LTE

**ETSI**

650 Route des Lucioles  
F-06921 Sophia Antipolis Cedex - FRANCE

---

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B  
Association à but non lucratif enregistrée à la  
Sous-Préfecture de Grasse (06) N° w061004871

---

**Important notice**

The present document can be downloaded from the  
[ETSI Search & Browse Standards](#) application.

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format on [ETSI deliver](#) repository.

Users should be aware that the present document may be revised or have its status changed, this information is available in the [Milestones listing](#).

If you find errors in the present document, please send your comments to the relevant service listed under [Committee Support Staff](#).

If you find a security vulnerability in the present document, please report it through our [Coordinated Vulnerability Disclosure \(CVD\)](#) program.

---

**Notice of disclaimer & limitation of liability**

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

---

**Copyright Notification**

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2026.  
All rights reserved.

---

# Intellectual Property Rights

## Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: "*Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards*", which is available from the ETSI Secretariat. Latest updates are available on the [ETSI IPR online database](#).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

## Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

**DECT™**, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™**, **LTE™** and **5G™** logo are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

---

# Legal Notice

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found at [3GPP to ETSI numbering cross-referencing](#).

---

# Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

# Contents

Intellectual Property Rights .....	2
Legal Notice .....	2
Modal verbs terminology.....	2
Foreword.....	24
1 Scope .....	26
2 References .....	26
3 Definitions of terms, symbols and abbreviations .....	27
3.1 Terms.....	27
3.2 Symbols.....	27
3.3 Abbreviations .....	28
4 General description of the coder .....	29
4.1 Introduction .....	29
4.2 IVAS codec overview .....	30
4.2.1 General.....	30
4.2.2 Mono (EVS-compatible) Operation.....	31
4.2.3 Stereo Operation .....	31
4.2.4 Scene-based Audio (SBA, Ambisonics) Operation .....	32
4.2.5 Metadata-assisted Spatial Audio (MASA) Operation.....	32
4.2.6 Objects (Independent Streams with Metadata, ISM) Operation .....	32
4.2.7 Multi-Channel (MC) Operation .....	33
4.2.8 Combined Objects and SBA (OSBA) Operation .....	33
4.2.9 Combined Objects and MASA (OMASA) Operation .....	33
4.2.10 Discontinuous Transmission (DTX) Operation .....	33
4.3 Input/output audio configurations .....	34
4.3.1 Input/output audio sampling rate .....	34
4.3.2 Input/output audio formats.....	34
4.4 Algorithmic delay.....	35
4.5 Rendering overview .....	36
4.5.1 Rendering introduction .....	36
4.5.2 Internal IVAS renderer .....	37
4.5.3 External IVAS renderer .....	37
4.5.4 Interface for external rendering .....	38
4.5.5 Split rendering .....	38
4.6 Organization of the rest of the Technical Standard .....	38
5 Functional description of the encoder .....	39
5.1 Encoder overview.....	39
5.2 Common processing and coding tools.....	40
5.2.1 High-pass filtering .....	40
5.2.2 Core-coder processing .....	40
5.2.2.1 Overview.....	40
5.2.2.2 Core-coder front pre-processing.....	41
5.2.2.2.1 General .....	41
5.2.2.2.2 Sample rate conversion.....	41
5.2.2.2.3 Pre-emphasis .....	42
5.2.2.2.4 Spectral analysis .....	42
5.2.2.2.5 Signal activity detection .....	42
5.2.2.2.6 Bandwidth detector.....	42
5.2.2.2.7 Time domain transient detection .....	44
5.2.2.2.8 Linear prediction analysis.....	45
5.2.2.2.9 Open-loop pitch analysis .....	45
5.2.2.2.10 Coding mode determination .....	45
5.2.2.2.11 Speech/Music classification .....	45
5.2.2.2.12 Core-coder technology pre-selection.....	56

5.2.2.3	Core-coder modules .....	63
5.2.2.3.1	Core-coder pre-processing .....	63
5.2.2.3.2	LP based coding .....	68
5.2.2.3.3	MDCT based coding .....	96
5.2.2.3.4	Switching coding modes .....	111
5.2.2.3.5	DTX/CNG operation .....	114
5.2.3	Common audio coding tools .....	116
5.2.3.1	General .....	116
5.2.3.2	Single Channel Element (SCE) .....	117
5.2.3.3	Channel Pair Element (CPE) .....	117
5.2.3.4	Multichannel Coding Tool (MCT) .....	119
5.2.3.4.1	Summary of the system .....	119
5.2.3.4.2	MCT configurations .....	120
5.2.3.4.3	Encoder single channel preprocessing up to whitened spectrum .....	120
5.2.3.4.4	Joint channel Encoding System Description .....	121
5.2.3.4.5	IGF for multi-channel processing .....	126
5.2.3.4.6	Bitrate distribution .....	126
5.2.3.4.7	Quantization and coding of each channel .....	127
5.2.4	Common spatial metadata coding tools .....	127
5.2.4.1	General .....	127
5.2.4.2	Spatial metadata composition .....	127
5.2.4.3	Direction metadata coding tools .....	128
5.2.4.3.1	Overview .....	128
5.2.4.3.2	Direction metadata quantization .....	128
5.2.4.3.3	Direction metadata raw encoding .....	131
5.2.4.3.4	Direction metadata entropy encoding tools .....	131
5.2.4.4	Diffuseness and energy ratio coding methods .....	132
5.2.4.4.1	Diffuseness and energy ratio definitions .....	132
5.2.4.4.2	Diffuseness parameter quantization .....	132
5.2.4.4.3	Diffuseness and energy ratio indices coding .....	133
5.2.4.4.4	Diffuseness and energy ratio coding with two concurrent directions .....	134
5.2.4.5	Direction metadata coding methods .....	136
5.2.4.5.1	Entropy coding 1 (EC1) .....	136
5.2.4.5.2	Entropy coding 2 (EC2) .....	141
5.2.4.5.3	Entropy coding 3 (EC3) .....	143
5.2.4.6	Coherence coding .....	147
5.2.4.6.1	Spread coherence coding .....	147
5.2.4.6.2	Surround coherence coding .....	149
5.2.4.7	DTX coding .....	150
5.2.5	Modified Discrete Fourier Transform (MDFT) Analysis Filter Bank .....	151
5.2.5.1	General .....	151
5.2.5.2	Modified Discrete Fourier Transform (MDFT) .....	151
5.2.5.3	Filter bank responses .....	151
5.2.5.3.1	General .....	151
5.2.5.3.2	Magnitude responses for 5 ms stride operation .....	152
5.2.5.3.3	Responses for sampling rates below 48 kHz .....	153
5.3	Stereo audio operation .....	154
5.3.1	Stereo format overview .....	154
5.3.2	Unified stereo .....	155
5.3.2.1	Overview .....	155
5.3.2.2	Common unified stereo tools .....	155
5.3.2.2.1	Inter-Channel Bandwidth Extension (IC-BWE) .....	155
5.3.2.2.2	Front VAD .....	166
5.3.2.3	TD-based stereo .....	167
5.3.2.3.1	Inter-channel Alignment (ICA) .....	167
5.3.2.3.2	TD stereo coder overview .....	179
5.3.2.3.3	Time-domain stereo downmix .....	180
5.3.2.3.4	Near out-of-phase operation (NOOP) .....	182
5.3.2.3.4.1	NOOP signal detection .....	182
5.3.2.3.4.2	NOOP sub-mode selection .....	184
5.3.2.3.4.3	NOOP signal coding .....	185
5.3.2.3.4.4	Adaptive mixing ratio for the NOOP signal .....	187

5.3.2.3.5	LP filter coherence .....	187
5.3.2.3.6	Open-loop pitch coherence .....	190
5.3.2.3.7	Excitation coding in the secondary channel using two or four subframes .....	191
5.3.2.3.8	Bit allocation .....	193
5.3.2.3.9	Quantization of LP coefficients for the secondary channel .....	194
5.3.2.4	DFT-based stereo .....	197
5.3.2.4.1	General .....	197
5.3.2.4.2	Time Domain ITD compensation .....	197
5.3.2.4.3	STFT analysis .....	199
5.3.2.4.4	GCC-PHAT ITD estimation .....	199
5.3.2.4.5	FD circular time-shift .....	205
5.3.2.4.6	Stereo parameters estimation .....	205
5.3.2.4.7	IPD calculation, stabilization and encoding scheme .....	206
5.3.2.4.8	Calculation of the side and residual prediction gains .....	210
5.3.2.4.9	Stereo parameter coding .....	211
5.3.2.4.10	Active Downmix .....	215
5.3.2.4.11	STFT synthesis .....	216
5.3.2.4.12	Residual coding .....	218
5.3.2.4.12.1	Overview .....	218
5.3.2.4.12.2	Adaptive residual signal encoding .....	221
5.3.2.4.12.2.1	Adaptive residual signal encoding parameter .....	221
5.3.2.4.12.2.2	Adaptive downmix for stereo coding .....	221
5.3.2.4.12.2.3	Calculation of downmixed signal and residual signal during transitional frame .....	222
5.3.2.4.12.2.4	Adaptive downmix .....	223
5.3.2.4.13	Reverberation gain parameter determination .....	223
5.3.2.5	Stereo classifier .....	223
5.3.2.5.1	General .....	223
5.3.2.5.2	Classification of uncorrelated content .....	224
5.3.2.5.3	Cross-talk detection using logistic regression .....	227
5.3.2.5.4	Cross-talk detection based on the GCC-PHAT .....	233
5.3.2.5.5	Stereo mode selection .....	236
5.3.3	MDCT-based stereo .....	239
5.3.3.1	General Overview .....	239
5.3.3.2	ITD compensation .....	239
5.3.3.3	MDCT core pre-processing .....	240
5.3.3.3.1	General .....	240
5.3.3.3.2	TCX-LTP parameter estimation .....	240
5.3.3.3.3	Time-to-frequency transformations .....	240
5.3.3.3.4	Mid-high bitrate pre-processing .....	243
5.3.3.3.5	Temporal noise shaping (TNS) .....	249
5.3.3.3.6	Spectral Noise Shaping (SNS) .....	250
5.3.3.4	Stereo processing .....	260
5.3.3.4.1	General .....	260
5.3.3.4.2	Stereo spectral bands .....	260
5.3.3.4.3	Global ILD normalization .....	262
5.3.3.4.4	Conditional ILD correction .....	262
5.3.3.4.5	Band-wise M/S decision .....	263
5.3.3.4.6	M/S decision for IGF .....	264
5.3.3.4.7	M/S Transformation .....	265
5.3.3.4.8	Encoding of the stereo parameters .....	266
5.3.3.5	Power spectrum calculation and spectrum noise measure .....	266
5.3.3.6	Intelligent gap filling (IGF) .....	266
5.3.3.6.1	General overview .....	266
5.3.3.6.2	Stereo IGF encoding .....	267
5.3.3.7	Bitrate distribution .....	268
5.3.3.8	Quantization and encoding .....	268
5.3.3.8.1	Spectrum quantization and coding .....	268
5.3.3.8.2	Global gain estimator .....	269
5.3.3.8.3	Rate-loop for constant bit rate and global gain .....	269
5.3.3.8.4	Stereo noise level estimation .....	269
5.3.4	Switching between stereo modes .....	269
5.3.4.1	Overview .....	269

5.3.4.2	General .....	270
5.3.4.3	Memory handling .....	272
5.3.4.4	Switching between DFT and TD stereo modes .....	273
5.3.4.4.1	Switching from TD stereo to DFT stereo .....	273
5.3.4.4.2	Switching from DFT stereo to TD stereo .....	274
5.3.4.5	Switching between DFT and MDCT stereo modes .....	277
5.3.4.5.1	Switching from DFT stereo to MDCT stereo .....	277
5.3.4.5.2	Switching from MDCT stereo to DFT stereo .....	277
5.3.4.6	Switching between TD and MDCT stereo modes .....	277
5.3.4.6.1	Switching from TD stereo to MDCT stereo .....	277
5.3.4.6.2	Switching from MDCT stereo to TD stereo .....	278
5.3.5	DTX operation .....	278
5.3.5.1	DTX in Unified stereo .....	278
5.3.5.1.1	Signal activity detection in Unified stereo .....	278
5.3.5.1.2	General .....	279
5.3.5.1.3	Stereo CNG side gain encoding .....	280
5.3.5.1.4	Stereo CNG ITD and IPD encoding .....	281
5.3.5.1.5	Stereo CNG coherence encoding .....	281
5.3.5.2	DTX in MDCT-based stereo .....	284
5.3.5.2.1	Overview .....	284
5.3.5.2.2	VAD in MDCT-based Stereo .....	284
5.3.5.2.3	Coherence estimation .....	285
5.3.5.2.4	Noise parameter estimation and encoding .....	286
5.4	Scene-based audio (SBA) operation .....	286
5.4.1	SBA format overview .....	286
5.4.2	Combined DirAC and SPAR based SBA coding .....	287
5.4.2.1	Combined DirAC and SPAR based SBA coding overview .....	287
5.4.2.2	FOA signal coding at all bitrates and HOA2, HOA3 signal coding at bitrates below 256 kbps .....	289
5.4.2.2.1	Overview .....	289
5.4.2.2.2	Metadata parameter analysis .....	289
5.4.2.2.3	Downmixing and core-coding .....	289
5.4.2.3	HOA2 and HOA3 signal coding overview at 256 kbps .....	290
5.4.2.3.1	Overview .....	290
5.4.2.3.2	Metadata parameter analysis .....	290
5.4.2.3.3	Downmixing and core-coding .....	290
5.4.2.4	HOA2 and HOA3 signal coding overview at 384 kbps .....	291
5.4.2.4.1	Overview .....	291
5.4.2.4.2	Metadata parameter generation .....	291
5.4.2.4.3	Downmixing and core-coding .....	291
5.4.2.5	HOA2 and HOA3 signal coding overview at 512 kbps .....	292
5.4.2.5.1	Overview .....	292
5.4.2.5.2	Metadata parameter generation .....	292
5.4.2.5.3	Downmixing and core-coding .....	292
5.4.3	SBA parameter estimation .....	293
5.4.3.1	SBA front VAD .....	293
5.4.3.2	Transient Detection .....	293
5.4.3.3	Analysis windowing and MDFT .....	293
5.4.3.4	DirAC parameter estimation .....	295
5.4.3.4.1	General .....	295
5.4.3.4.2	High-order operation mode .....	295
5.4.3.4.3	Low-order operation mode .....	299
5.4.3.5	Hybrid encoder .....	300
5.4.3.6	Mono signal detection .....	300
5.4.3.6.1	Overview .....	300
5.4.3.6.2	Signalling of Mono through the bitstream .....	304
5.4.3.7	SPAR Metadata computation .....	305
5.4.3.7.1	Overview .....	305
5.4.3.7.2	Banded covariance computation .....	306
5.4.3.7.3	Active W downmix detector .....	307
5.4.3.7.4	Banded covariance re-mixing .....	308
5.4.3.7.5	SPAR parameter estimation .....	309
5.4.3.7.6	Passive-W channel prediction coefficients computation .....	309

5.4.3.7.7	Active-W channel prediction coefficients computation .....	309
5.4.3.7.8	Cross-Prediction coefficients computation .....	312
5.4.3.7.9	Decorrelation coefficients computation .....	313
5.4.3.7.10	Overview of quantization, bitrate distribution and coding of SPAR parameters .....	313
5.4.3.7.11	SPAR bitrate distribution table .....	314
5.4.3.7.12	Quantization and coding process .....	317
5.4.3.7.13	Spectral filling in the side channels at 13.2 kbps and 16.4 kbps .....	320
5.4.3.7.14	Time differential coding .....	320
5.4.3.7.15	Band interleaving and 40ms MD update rate at 13.2 kbps and 16.4 kbps .....	321
5.4.3.7.16	Entropy coding with Arithmetic coders .....	322
5.4.3.7.17	Base2 coding with Huffman coder .....	324
5.4.3.8	DirAC and SPAR parameter merge .....	325
5.4.3.8.1	General .....	325
5.4.3.8.2	DirAC to SPAR parameter conversion .....	325
5.4.3.8.3	Covariance estimation from DirAC parameters .....	325
5.4.3.8.4	Directional diffuseness .....	327
5.4.3.8.5	SPAR parameter estimation from DirAC parameters .....	327
5.4.3.8.6	Active W prediction from DirAC parameters .....	327
5.4.3.8.7	Passive W prediction from DirAC parameters .....	328
5.4.4	Downmix matrix calculation .....	328
5.4.4.1	SPAR parameters to downmix matrix conversion .....	328
5.4.4.2	Energy compensation in DirAC bands .....	329
5.4.4.3	Band Interleaving of downmix matrix at 13.2 and 16.4 kbps .....	329
5.4.5	Downmix generation .....	329
5.4.5.1	Overview .....	329
5.4.5.2	Time domain to MDFT conversion .....	329
5.4.5.3	Filterbank mixing .....	330
5.4.5.4	MDFT to Time domain conversion .....	330
5.4.5.5	Windowing and crossfading .....	330
5.4.5.6	Time domain mixing .....	331
5.4.6	Principal Component Analysis (PCA) .....	331
5.4.7	Automatic Gain Control (AGC) .....	337
5.4.8	Core-coder encoding .....	339
5.4.8.1	General .....	339
5.4.8.2	MCT bitrate distribution in SBA high bitrate mode .....	339
5.4.9	DTX operation .....	340
5.4.9.1	Overview .....	340
5.4.9.2	DirAC parameter estimation .....	340
5.4.9.3	SPAR parameter estimation .....	340
5.4.9.3.1	General .....	340
5.4.9.3.2	Banded covariance smoothing .....	341
5.4.9.3.3	Quantization and coding in VAD inactive frames .....	341
5.4.9.4	SPAR and DirAC parameter merge .....	341
5.4.9.5	CNG parameters encoding .....	342
5.4.10	SBA bitrate switching .....	342
5.5	Metadata-assisted spatial audio (MASA) operation .....	343
5.5.1	MASA format overview .....	343
5.5.2	MASA format metadata input .....	344
5.5.2.1	Obtaining the MASA format metadata .....	344
5.5.2.2	Direction index deindexing .....	344
5.5.2.3	TF tile based energy calculation .....	346
5.5.3	Coding of MASA format input data .....	346
5.5.3.1	Overview .....	346
5.5.3.2	MASA metadata pre-encoding configuration, processing, and signalling .....	346
5.5.3.2.1	Overview .....	346
5.5.3.2.2	Bitrate dependent parameter settings .....	347
5.5.3.2.3	Metadata composition detection, alignment, and check for validity .....	348
5.5.3.2.4	Common MASA metadata codec configuration .....	354
5.5.3.2.5	Energy ratio compensation .....	356
5.5.3.2.6	Merging of MASA spatial parameter metadata across MASA frequency bands and subframes .....	357
5.5.3.2.7	Combining of MASA spatial audio metadata across multiple directions .....	359
5.5.3.2.8	Metadata reductions for low rate .....	361

5.5.3.2.9	Reordering of directional data with two concurrent directions.....	363
5.5.3.2.10	Signalling bits.....	363
5.5.3.2.11	Adjustment of highest transmitted MASA metadata band .....	363
5.5.3.3	MASA metadata quantization and encoding.....	364
5.5.3.3.1	Overview .....	364
5.5.3.3.2	Energy ratio encoding.....	365
5.5.3.3.3	Direction encoding .....	365
5.5.3.3.4	Spread coherence encoding .....	366
5.5.3.3.5	Surround coherence encoding .....	366
5.5.3.3.6	Coding of the second direction parameters.....	366
5.5.4	Encoding of MASA audio transport channels .....	367
5.5.5	DTX operation .....	367
5.5.6	Bitstream structure.....	368
5.5.7	MASA bitrate switching .....	368
5.6	Object-based audio (ISM) operation .....	368
5.6.1	ISM format overview.....	368
5.6.2	Discrete ISM coding mode .....	369
5.6.2.1	Overview .....	369
5.6.2.2	DiscISM encoding system.....	369
5.6.2.3	Bitrates distribution between audio streams.....	370
5.6.2.3.1	Bitrate distribution algorithm .....	370
5.6.2.3.2	Bitrate adaptation based on ISM importance.....	371
5.6.3	Parametric ISM coding mode .....	372
5.6.3.1	General .....	372
5.6.3.2	Parameters .....	373
5.6.3.3	Noisy Speech.....	374
5.6.3.4	Downmix.....	374
5.6.3.5	Encoded Data .....	376
5.6.4	ISM metadata coding .....	377
5.6.4.1	General .....	377
5.6.4.2	Direction metadata encoding and quantization .....	377
5.6.4.2.1	General .....	377
5.6.4.2.2	Intra-object metadata coding logic .....	379
5.6.4.2.3	Inter-object metadata coding logic .....	379
5.6.4.3	Extended metadata encoding and quantization .....	379
5.6.4.4	Panning gain in non-diegetic rendering.....	380
5.6.5	Bitstream structure.....	380
5.6.5.1	Overview.....	380
5.6.5.2	ISM signalling.....	380
5.6.5.3	Coded metadata payload .....	381
5.6.5.4	Audio streams payload.....	381
5.6.6	DTX operation .....	381
5.6.6.1	Overview.....	381
5.6.6.2	ISM DTX operation overview.....	381
5.6.6.3	Classification of Inactive Frames in ISM Encoder.....	382
5.6.6.3.1	General .....	382
5.6.6.3.2	Global SID Counter.....	383
5.6.6.4	SID Metadata Analysis, Quantization and Coding.....	384
5.6.6.5	CNG parameters encoding .....	385
5.6.6.6	SID bitstream structure .....	386
5.6.6.6.1	Overview .....	386
5.6.6.6.2	ISM Common Signalling in SID Frame .....	387
5.6.6.6.3	SID Data Payload .....	387
5.6.6.6.4	SID Audio Streams Payload.....	387
5.6.7	ISM bitrate switching.....	388
5.7	Multi-channel audio (MC) operation.....	388
5.7.1	MC format overview.....	388
5.7.2	LFE channel encoding .....	388
5.7.2.1	Low-pass filtering .....	388
5.7.2.2	MDCT based LFE encoding .....	389
5.7.2.2.1	Overview .....	389
5.7.2.2.2	Windowing and MDCT.....	389

5.7.2.2.3	Quantization .....	390
5.7.2.2.4	Coding .....	391
5.7.3	Multi-channel MASA (McMASA) coding mode .....	392
5.7.3.1	McMASA coding mode overview .....	392
5.7.3.2	LFE energy computation .....	394
5.7.3.2.1	LFE energy computation in normal sub-mode .....	394
5.7.3.2.2	LFE energy computation in separate-channel sub-mode .....	395
5.7.3.3	McMASA spatial audio parameter estimation .....	395
5.7.3.4	Even loudspeaker setup determination .....	398
5.7.3.5	McMASA transport audio signal generation .....	399
5.7.3.6	McMASA metadata encoding .....	400
5.7.3.6.1	McMASA metadata pre-encoding configuration and processing .....	400
5.7.3.6.2	Spatial metadata encoding .....	400
5.7.3.6.3	LFE-to-total energy ratio encoding .....	401
5.7.3.7	McMASA transport audio signal encoding .....	402
5.7.3.8	Bitstream structure .....	403
5.7.4	Parametric MC coding mode .....	403
5.7.4.1	ParamMC coding mode overview .....	403
5.7.4.2	ParamMC parameter frame index initialization and update .....	404
5.7.4.3	ParamMC transport channel and parameter band configuration .....	404
5.7.4.4	ParamMC transport audio signal generation .....	405
5.7.4.5	ParamMC time domain transient detection .....	406
5.7.4.6	ParamMC analysis windowing and MDFT .....	406
5.7.4.7	ParamMC covariance estimation .....	407
5.7.4.8	ParamMC default settings for the LFE channel .....	408
5.7.4.9	ParamMC parameter based transient detection .....	408
5.7.4.10	ParamMC band combining in case of transients .....	408
5.7.4.11	ParamMC complex valued to real valued covariance conversion .....	408
5.7.4.12	ParamMC LFE activity detection .....	409
5.7.4.13	ParamMC parameter quantization .....	409
5.7.4.13.1	Parameter quantizers .....	409
5.7.4.13.2	Interchannel Level differences (ICLDs) .....	409
5.7.4.13.3	Inter-channel coherences (ICCs) .....	411
5.7.4.14	ParamMC parameter encoding .....	411
5.7.4.14.1	Common ParamMC parameter encoding .....	411
5.7.4.14.2	ICC and ICLD Parameter quantization indices encoding .....	412
5.7.4.15	ParamMC transport audio signal encoding .....	415
5.7.4.16	ParamMC bit rate switching .....	415
5.7.5	Parametric upmix MC coding mode .....	416
5.7.5.1	General .....	416
5.7.5.2	Sectioning and downmixing .....	416
5.7.5.3	Metadata parameter calculation .....	416
5.7.5.4	Quantization .....	418
5.7.5.5	Entropy Coding .....	419
5.7.6	Discrete MC coding mode .....	420
5.7.7	MC bitrate switching .....	420
5.8	Combined Object-based audio and SBA (OSBA) operation .....	420
5.8.1	OSBA format overview .....	420
5.8.2	Low-bitrate pre-rendering OSBA coding mode .....	421
5.8.3	High-bitrate discrete OSBA coding mode .....	421
5.8.4	OSBA bitrate switching .....	421
5.9	Combined Object-based audio and MASA (OMASA) operation .....	422
5.9.1	OMASA format overview .....	422
5.9.2	OMASA format configurations .....	422
5.9.3	OMASA pre-coding processing tools .....	423
5.9.3.1	OMASA spatial audio parameter analysis .....	423
5.9.3.2	OMASA MASA metadata combining .....	424
5.9.3.3	OMASA audio signals downmix .....	425
5.9.3.4	Determination of an object to be separated .....	426
5.9.3.5	Separation of an object from other objects .....	427
5.9.4	Low-bitrate pre-rendering (Rend OMASA) coding mode .....	428
5.9.4.1	Overview .....	428

5.9.4.2	Low-bitrate pre-rendering coding method .....	428
5.9.5	One object with MASA representation (One OMASA) coding mode.....	429
5.9.5.1	Overview .....	429
5.9.5.2	One object with MASA representation coding method.....	429
5.9.6	Parametric one object (Param OMASA) coding mode.....	429
5.9.6.1	Overview .....	429
5.9.6.2	Determination of OMASA parametric information .....	430
5.9.6.3	One object with parametric representation coding method .....	430
5.9.6.3.1	Coding method overview.....	430
5.9.6.3.2	Encoding of MASA-to-total energy ratios .....	431
5.9.6.3.3	Encoding of ISM energy ratios.....	432
5.9.6.3.4	Encoding of ISM metadata .....	435
5.9.7	Discrete (Disc OMASA) coding mode .....	436
5.9.8	Inter-format bitrate adaptation .....	437
5.9.9	OMASA bitrate switching .....	439
5.9.10	OMASA bitstream structure .....	439
5.10	EVS-compatible mono audio operation .....	440
5.11	Stereo downmix operation for EVS mono coding.....	440
5.11.1	Overview .....	440
5.11.2	Phase-only correlation (POC) mode .....	440
5.11.2.1	Overview .....	440
5.11.2.2	ITD analysis .....	440
5.11.2.2.1	Cross talk adding procedure in the frequency domain.....	440
5.11.2.2.2	Approximated phase spectrum .....	441
5.11.2.2.3	Calculation of phase only correlation .....	441
5.11.2.2.4	Calculation of ITD and weighting coefficients.....	442
5.11.2.3	Downmix process.....	443
5.11.3	Phase compensation (PHA) mode .....	444
5.11.3.1	Overview .....	444
5.11.3.2	Inter spectral difference.....	444
5.11.3.3	Inter-channel phase difference (IPD) and realigned inter-channel correlation (ICCr).....	445
5.11.3.4	Selection rule for PHA sub-mode .....	446
5.11.3.5	Determination of filter taps .....	446
5.11.3.5.1	General .....	446
5.11.3.5.2	Filter taps for IPD .....	446
5.11.3.5.3	Filter taps for IPD2.....	447
5.11.3.5.4	Truncation and normalization.....	448
5.11.3.6	Downmix process.....	448
5.11.4	Mode selection rule.....	449
5.11.5	Handling of mode switching .....	449
5.11.5.1	Introduction.....	449
5.11.5.2	Level .....	449
5.11.5.3	Crossfade.....	450
6	Functional description of the decoder .....	451
6.1	Decoder overview.....	451
6.2	Common processing and decoding tools .....	452
6.2.1	Common decoder processing overview .....	452
6.2.1.1	High-pass filtering.....	452
6.2.1.2	Limiter.....	452
6.2.1.2.1	Overview .....	452
6.2.1.2.2	Detailed algorithmic description.....	452
6.2.1.2.3	Handling of strong saturations.....	453
6.2.1.3	Time-domain decorrelator.....	455
6.2.2	Core-decoder processing.....	456
6.2.2.1	Overview .....	456
6.2.2.2	LP based decoding .....	457
6.2.2.2.1	Variable bitrate ACELP decoding .....	457
6.2.2.2.2	Fast algebraic codebook decoding.....	457
6.2.2.2.3	AVQ Bit Savings Decoder .....	457
6.2.2.2.4	Bass Post-Filter.....	461
6.2.2.2.5	GSC core decoding.....	461

6.2.2.2.6	Bandwidth extension in time domain .....	464
6.2.2.2.7	Multimode FD bandwidth extension coding .....	464
6.2.2.3	MDCT based decoding .....	464
6.2.2.3.1	General .....	464
6.2.2.3.2	Variable bitrate MDCT decoding .....	464
6.2.2.3.3	TCX entropy decoding .....	464
6.2.2.3.4	PLC method selection.....	466
6.2.2.3.5	Phase ECU enhancements .....	467
6.2.2.3.6	Long term prediction processing .....	468
6.2.2.4	Background noise estimation .....	470
6.2.2.4.1	General .....	470
6.2.2.4.2	Power spectrum compression .....	470
6.2.2.4.3	Compensation for the loss of variance.....	472
6.2.2.4.4	Spectral smoothing .....	472
6.2.2.4.5	Initialization period .....	473
6.2.2.4.6	Power spectrum expansion .....	475
6.2.2.5	Switching coding modes .....	475
6.2.2.5.1	Improved handling of envelope stability parameter in HQ MDCT .....	475
6.2.2.6	DTX/CNG operation.....	476
6.2.2.6.1	General .....	476
6.2.2.6.2	FD-CNG VQ global gain dequantization .....	476
6.2.2.6.3	First stage VQ decoding for FD-CNG.....	476
6.2.3	Common audio decoding tools .....	477
6.2.3.1	General .....	477
6.2.3.2	Single Channel Element (SCE) decoder .....	477
6.2.3.3	Channel Pair Element (CPE) decoder .....	478
6.2.3.4	Multichannel coding tool (MCT) decoder.....	480
6.2.3.4.1	General overview.....	480
6.2.3.4.2	Core and SNS parameter decoding.....	480
6.2.3.4.3	MCT parameter decoding.....	480
6.2.3.4.4	Bitrate distribution.....	480
6.2.3.4.5	Spectral data decoding and noise filling.....	481
6.2.3.4.6	Application of IGF and stereo IGF.....	481
6.2.3.4.7	MCT decoding.....	481
6.2.3.4.8	Noise shaping .....	481
6.2.3.4.9	Inverse MDCT transform to time domain .....	482
6.2.3.4.10	MCT PLC.....	482
6.2.4	Common spatial metadata decoding tools .....	482
6.2.4.1	Direction metadata decoding tools.....	482
6.2.4.1.1	Overview .....	482
6.2.4.1.2	Direction dequantization tools.....	482
6.2.4.1.3	Direction metadata raw decoding .....	483
6.2.4.1.4	Direction metadata entropy decoding tools .....	483
6.2.4.2	Diffuseness and energy ratio decoding methods .....	484
6.2.4.2.1	Diffuseness decoding.....	484
6.2.4.2.2	Diffuseness and energy ratio decoding with two concurrent directions .....	484
6.2.4.3	Direction metadata decoding methods .....	485
6.2.4.3.1	Direction metadata decoding overview .....	485
6.2.4.3.2	Decoding EC1 .....	485
6.2.4.3.3	Decoding EC2 .....	486
6.2.4.3.4	Decoding EC3 .....	486
6.2.4.4	Coherence decoding .....	488
6.2.4.4.1	Spread coherence decoding .....	488
6.2.4.4.2	Surround coherence decoding .....	489
6.2.4.5	DTX decoding.....	489
6.2.5	Common filter bank operation tools .....	490
6.2.5.1	Complex Low-delay Filter Bank (CLDFB) analysis.....	490
6.2.5.2	Complex Low-delay Filter Bank (CLDFB) synthesis.....	490
6.2.6	CLDFB domain decorrelator .....	491
6.2.6.1	Overview .....	491
6.2.6.2	Onset detection and removal .....	491
6.2.6.3	De-correlation .....	492

6.2.6.3.1	Overview .....	492
6.2.6.3.2	Derivation of filter coefficients from lattice coefficients .....	492
6.2.6.3.3	Decorrelator IIR Filtering .....	495
6.2.6.4	Energy adjustment .....	496
6.2.6.5	Final output .....	496
6.2.7	Common tools for decoding with time scale modification (JBM) .....	497
6.2.7.1	Overview .....	497
6.2.7.2	Transport channel buffer management .....	497
6.2.7.3	Time scale modification .....	498
6.2.7.3.1	Extension to multi-channel processing .....	498
6.2.7.3.2	Energy estimation .....	498
6.2.7.4	Metadata and processing parameter adaption .....	499
6.2.7.4.1	Overview .....	499
6.2.7.4.2	Metadata adaption .....	499
6.2.7.4.3	Processing parameter adaption .....	501
6.2.7.5	Bitrate switching .....	503
6.2.7.5.1	Renderer flushing .....	503
6.2.7.5.2	Transport channel reconfiguration .....	503
6.2.7.5.3	Renderer discard samples .....	503
6.3	Stereo audio decoding .....	504
6.3.1	Common stereo decoding tools .....	504
6.3.1.1	Common Stereo postprocessing .....	504
6.3.1.2	ITD compensation .....	504
6.3.1.2.1	Overview .....	504
6.3.1.2.2	Inter-channel Alignment (ICA) decoder .....	504
6.3.2	Unified stereo decoding .....	505
6.3.2.1	Inter-channel BWE (IC-BWE) decoder .....	505
6.3.2.1.1	Inter-Channel BWE Decoder overview .....	505
6.3.2.1.2	Inter-Channel BWE Decoder Block .....	506
6.3.2.2	TD-based stereo decoding .....	507
6.3.2.2.1	TD stereo decoder overview .....	507
6.3.2.2.2	Reusing close-loop pitch and adaptive codebook of the primary channel .....	508
6.3.2.2.3	Reusing LP filter coefficients of the primary channel and dequantization of LSF parameters for the secondary channel .....	509
6.3.2.2.4	Estimation of spatial parameters of the background noise .....	510
6.3.2.3	DFT-based stereo decoding .....	510
6.3.2.3.1	General .....	510
6.3.2.3.2	STFT analysis .....	510
6.3.2.3.3	Stereo parameter decoding .....	511
6.3.2.3.4	Residual signal decoding .....	512
6.3.2.3.5	Bass post filter application .....	513
6.3.2.3.6	Residual predicted signal generation .....	515
6.3.2.3.7	Stereo upmix .....	522
6.3.2.3.8	Stereo comfort noise injection .....	522
6.3.2.3.9	STFT synthesis .....	524
6.3.2.3.10	DFT-based stereo PLC .....	525
6.3.3	MDCT-based stereo decoding .....	528
6.3.3.1	General Overview .....	528
6.3.3.2	Decoding of parameters .....	528
6.3.3.2.1	TCX block configuration .....	528
6.3.3.2.2	Core parameters decoding .....	528
6.3.3.2.3	SNS parameters decoding .....	528
6.3.3.2.4	Stereo parameters decoding .....	530
6.3.3.3	Decoding process .....	531
6.3.3.3.1	Spectral data decoding and noise filling .....	531
6.3.3.3.2	Application of stereo IGF .....	531
6.3.3.4	Stereo decoding process .....	532
6.3.3.4.1	Inverse mid-side transform .....	532
6.3.3.4.2	De-normalization to global ILD .....	534
6.3.3.5	Post-stereo decoding process .....	534
6.3.3.5.1	General .....	534
6.3.3.5.2	Spectral noise shaping .....	534

6.3.3.5.3	Temporal noise shaping.....	535
6.3.3.6	Frequency-to-time domain transformations .....	535
6.3.3.6.1	General .....	535
6.3.3.6.2	Kernel switching.....	535
6.3.3.6.3	Transition between long and short blocks with double overlap.....	535
6.3.3.7	PLC in MDCT-based stereo.....	537
6.3.3.7.1	Frequency-domain concealment.....	537
6.3.3.7.2	Time-domain concealment .....	541
6.3.3.7.3	Fading to zero output.....	542
6.3.4	Switching between stereo modes .....	542
6.3.4.1	Overview.....	542
6.3.4.2	General .....	542
6.3.4.3	Memory handling .....	545
6.3.4.4	Switching between DFT and TD stereo modes.....	546
6.3.4.4.1	Switching from TD stereo to DFT stereo .....	546
6.3.4.4.2	Switching from DFT stereo to TD stereo .....	549
6.3.4.5	Switching between DFT and MDCT stereo modes.....	552
6.3.4.5.1	Switching from DFT stereo to MDCT stereo .....	552
6.3.4.5.2	Switching from MDCT stereo to DFT stereo .....	552
6.3.4.6	Switching between TD and MDCT stereo modes.....	552
6.3.4.6.1	Switching from TD stereo to MDCT stereo .....	552
6.3.4.6.2	Switching from MDCT stereo to TD stereo .....	553
6.3.5	DTX operation .....	553
6.3.5.1	DTX in Unified stereo.....	553
6.3.5.1.1	General .....	553
6.3.5.1.2	Stereo CNG spectral shape extraction .....	553
6.3.5.1.3	Stereo CNG side gain, ITD and IPD decoding .....	556
6.3.5.1.4	Stereo CNG coherence decoding.....	557
6.3.5.1.5	Stereo CNG synthesis.....	557
6.3.5.2	DTX in MDCT-based stereo .....	558
6.3.5.2.1	General overview.....	558
6.3.5.2.2	Decoding of parametric noise data .....	558
6.3.5.2.3	Stereo CNG .....	558
6.3.5.2.4	Smoothing of transitions from DTX frame to active TCX .....	559
6.3.6	Stereo output format conversion.....	560
6.3.6.1	Overview.....	560
6.3.6.2	Stereo-to-Mono output.....	560
6.3.6.2.1	DFT stereo mono output.....	560
6.3.6.2.2	TD stereo mono output.....	561
6.3.6.2.3	MDCT stereo mono output.....	561
6.3.6.3	Stereo-to-MC output .....	565
6.3.7	Stereo audio decoding with TSM .....	565
6.4	Scene-based audio (SBA) decoding .....	565
6.4.1	Combined DirAC and SPAR based SBA decoding overview .....	565
6.4.1.1	Overview.....	565
6.4.1.2	FOA signal decoding at all bitrates and HOA2, HOA3 signal decoding at bitrates below 256 kbps.....	566
6.4.1.2.1	Overview .....	566
6.4.1.2.2	Metadata decoding.....	567
6.4.1.2.3	Core-coder decoding .....	567
6.4.1.2.4	SBA upmix and rendering .....	567
6.4.1.3	HOA2 and HOA3 signal decoding at bitrate 256 kbps .....	568
6.4.1.3.1	Overview .....	568
6.4.1.3.2	Metadata decoding.....	568
6.4.1.3.3	Core-coder decoding .....	568
6.4.1.3.4	SBA upmix and rendering .....	569
6.4.1.4	HOA2 and HOA3 signal decoding at 384 kbps .....	569
6.4.1.4.1	Overview .....	569
6.4.1.4.2	Metadata decoding.....	570
6.4.1.4.3	Core-coder decoding .....	570
6.4.1.4.4	SBA upmix and rendering .....	570
6.4.1.5	HOA2 and HOA3 signal decoding at 512 kbps .....	571

6.4.1.5.1	Overview .....	571
6.4.1.5.2	Metadata decoding.....	571
6.4.1.5.3	Core-coder decoding .....	571
6.4.1.5.4	SBA upmix and rendering .....	571
6.4.2	DirAC parameter decoding .....	572
6.4.3	SPAR parameter decoding.....	572
6.4.3.1	General .....	572
6.4.3.2	Set active W prediction flag.....	572
6.4.3.3	Decode quantization and coding strategy bits.....	572
6.4.3.4	Time differential decoding.....	573
6.4.3.4.1	General .....	573
6.4.3.4.2	Time differential decoding across same quantization strategies.....	573
6.4.3.4.3	Time differential decoding across different quantization strategies .....	574
6.4.3.5	Band interleaved decoding at 13.2 kbps and 16.4 kbps.....	574
6.4.3.6	Arithmetic decoder.....	574
6.4.3.7	Huffman decoder.....	574
6.4.4	SPAR and DirAC parameter merge .....	574
6.4.4.1	SBA mono handling .....	574
6.4.4.2	SPAR to DirAC parameter conversion .....	575
6.4.4.3	DirAC model for the SPAR parameters.....	576
6.4.4.3.1	Overview .....	576
6.4.4.3.2	DirAC to SPAR conversion for parameters corresponding to downmix channels.....	576
6.4.4.3.3	DirAC to SPAR conversion for PR, CP and D coefficients corresponding to parametric channels .....	576
6.4.5	Upmix matrix calculation .....	577
6.4.5.1	Overview.....	577
6.4.5.2	SPAR matrix generation (P and C matrices).....	577
6.4.5.3	Upmix matrix generation .....	578
6.4.6	Decoded audio processing .....	578
6.4.6.1	Downmix signal decoding (core-decoding) .....	578
6.4.6.2	AGC .....	578
6.4.6.3	PCA.....	579
6.4.6.4	SPAR upmix processing .....	581
6.4.6.4.1	General .....	581
6.4.6.4.2	Time domain decorrelator .....	581
6.4.6.4.3	CLDFB analysis .....	581
6.4.6.4.4	CLDFB upmixing.....	581
6.4.6.4.5	MDFT filterbank banded upmix matrix to CLDFB banded upmix matrix conversion .....	583
6.4.6.4.6	Crossfading in CLDFB.....	584
6.4.6.4.7	Additional VLBR smoothing.....	584
6.4.6.5	DirAC synthesis and rendering .....	584
6.4.6.5.1	Overview .....	584
6.4.6.5.2	Mode selection.....	585
6.4.6.5.3	High-order operation mode .....	585
6.4.6.5.4	Low-order operation mode .....	588
6.4.6.5.5	Bitrate switching.....	591
6.4.6.5.6	Binaural rendering.....	591
6.4.6.5.7	Loudspeaker rendering .....	592
6.4.6.5.8	Stereo and mono output.....	594
6.4.6.6	CLDFB synthesis .....	598
6.4.7	DTX operation .....	599
6.4.7.1	Overview.....	599
6.4.7.2	DTX in DirAC .....	599
6.4.7.3	DTX in SPAR .....	599
6.4.8	SBA PLC .....	600
6.4.8.1	PLC in DirAC .....	600
6.4.8.2	PLC in SPAR .....	600
6.4.8.2.1	SPAR processing of erroneous frames .....	600
6.4.8.2.2	Recovery after one or more lost frames.....	601
6.4.8.2.3	SPAR parameter recovery for 24.4 kbps and above.....	601
6.4.8.2.4	SPAR parameter recovery for 13.2 kbps and 16.4 kbps.....	601
6.4.8.2.5	Interpolation of missing SPAR parameters from current existing SPAR parameters.....	601