3GPP TS 26.401 V7.0.0 (2006-06)

Technical Specification

**3rd Generation Partnership Project;
Technical Specification Group Services and System Aspects;
General audio codec audio processing functions;
Enhanced aacPlus general audio codec;
General description
(Release 7)**

Keywords
UMTS, codec

3GPP

Postal address

3GPP support office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet

http://www.3gpp.org

3GPP

Contents

Foreword

The present document describes the Enhanced aacPlus general audio codec within the 3GPP system.

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of this TS, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x   the first digit:

1   presented to TSG for information;

2   presented to TSG for approval;

3   Indicates TSG approved document under change control.

y   the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z   the third digit is incremented when editorial only changes have been incorporated in the specification;

1 Scope

This Telecommunication Standard (TS) describes the detailed mapping from an MPEG-4 bitstream containing Enhanced aacPlus coded audio to PCM sample output. The Enhanced aacPlus audio codec is based on the AAC, SBR and Parametric Stereo coding tools defined in the MPEG-4 Audio standard [5][6][7]. In addition it includes further tools such as error concealment, spline resampler, and stereo-to-mono downmix.

This Telecommunication Standard (TS) also describes the detailed mapping from a PCM sample input to an MPEG-4 bitstream containing Enhanced aacPlus coded audio.

2 Normative references

This TS incorporates by dated and undated reference, provisions from other publications. These normative references are cited in the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this TS only when incorporated in it by amendment or revision. For undated references, the latest edition of the publication referred to applies.

[1]   3GPP TS 26.410 : Enhanced aacPlus general audio codec; Floating-point ANSI-C Code.

[2]   3GPP TS 26.403 : Enhanced aacPlus general audio codec; Encoder Specification AAC part.

[3]   3GPP TS 26.404 : Enhanced aacPlus general audio codec; Encoder Specification SBR part.

[4]   3GPP TS 26.405 : Enhanced aacPlus general audio codec; Encoder Specification Parametric Stereo part.

[5]   ISO/IEC 14496-3:2001, Information technology - Coding of audio-visual objects - Part 3: Audio.

[6]   ISO/IEC 14496-3:2001/Amd.1:2003, Bandwidth Extension.

[7]   ISO/IEC 14496-3:2001/Amd.1:2003/DCOR1.

[8]   ISO/IEC 14496-3:2001/Amd.2:2004, Parametric Coding for High Quality Audio.

[9]   3GPP TS 26.402: Enhanced aacPlus general audio codec; Additional Decoder Tools.

[10]   3GPP TS 26.411 : Enhanced aacPlus general audio codec; Fixed-point ANSI-C Code.

[11]   3GPP TS 26.234 : Transparent end-to-end Packet-switched Streaming Service (PSS) ; Protocols and codecs.

[12]   ISO/IEC 14496-3:2001/Amd.2:2004/DCOR 1.

3 Abbreviations

For the purposes of this TS, the following abbreviations apply.

AAC   Advanced Audio Coding
AAC-LC   Advanced Audio Coding Low Complexity Object Type
AAC-LTP   Advanced Audio Coding Long Term Predictor Object Type
aacPlus   MPEG-4 High Efficiency AAC, the combination of MPEG-4 AAC and MPEG-4 Bandwidth extension (SBR)
Enhanced aacPlus   MPEG-4 High Efficiency AAC plus MPEG-4 Parametric StereoMDCT   Modified Discrete Cosine Transform
QMF   Quadrature Mirror Filter
SBR   Spectral Band Replication

# 4	Outline description

This TS is structured as follows:

Section 5 gives a general overview of the parts in the Enhanced aacPlus codec. It further specifies what parts of the cited ISO standards apply.

Section 7 gives a more detailed overview of the Enhanced aacPlus encoder, and references the relevant detailed technical description documents.

Section 8 gives a more detailed overview of the ISO standardised parts of the Enhanced aacPlus decoder, and references the relevant ISO standards.

Section 9 gives a more detailed overview of the additional tools present in the Enhanced aacPlus decoder that are not part of the cited ISO standards, and references the relevant detailed technical description documents.

# 5	General

The Enhanced aacPlus general audio codec consists of MPEG-4 AAC, MPEG-4 SBR and MPEG-4 Parametric Stereo. The AAC is a general audio codec, SBR is a bandwidth extension technique offering substantial coding gain in combination with AAC, and Parametric Stereo enables stereo coding at very low bitrates. In addition to the above parts of the Enhanced aacPlus codec that are specified in ISO standards [5][6][7][8][12] there are 3 additional tools included in the Enhanced aacPlus decoder:

- Error concealment tools for AAC, SBR, and Parametric Stereo make the decoder robust against transmission errors like frame loss. These tools mitigate audible effects of such errors.

- The stereo-to-mono downmix tool enables a decoder only capable of mono output to downmix a stereo bitstream. For the AAC part this is done in the time domain after the stereo decoding but for SBR this is done on the SBR parameters and thus saving complexity since only a mono decoding of SBR is needed.

- The Spline resampler tool gives the possibility to resample the output to a sampling frequency different than what was supplied in the bitstream. This gives for example handsets with a D/A converter only capable of 16 kHz sampling frequency the possibility to play bit streams encoded with 22.05 kHz sampling frequency.

The 3GPP Enhanced aacPlus general audio codec offers monophonic and stereophonic coding. For stereophonic coding two stereo modes are used: parametric stereo for low bitrates and M/S stereo for high bitrates. The codec is based on the MPEG-4 Audio ISO standard. The cited ISO standards define several profiles and levels of which not all are applicable in the 3GPP context. From the ISO standards the following subset shall be used:

The Enhanced aacPlus general audio codec implements the High Efficiency AAC Profile at Level 2[1] as defined in [6]. In addition, the following restriction applies:

- frameLengthFlag in GASpecificConfig() shall be 0 (i.e., 960 framing is not supported);

For terminals supporting stereophonic output the following additional statements apply:

- for mono and parametric stereo bitstreams, the Enhanced aacPlus decoder operates the SBR tool in HQ mode, thus the SBR HQ tool is required;

- the parametric stereo enhancement implements the baseline version of the parametric stereo coding tool in direct combination with the SBR tool, as defined in [8].

- for M/S stereo bitstreams, it is recommended that the Enhanced aacPlus decoder operates the SBR tool in Low Power mode.

For terminals that are only capable of producing monophonic output the following additional statements apply:

---

[1] The HE-AAC Profile combines the AAC Low Complexity object type plus the SBR tool. The AAC LC object type does not implement the Long Term Predictor (LTP) tool. The Level 2 implies a restriction to a maximum of two channels. Furthermore in case of SBR being used, the maximum AAC sampling rate is restricted to 24 kHz whereas if SBR is not used the maximum AAC sampling rate is restricted to 48 kHz.

- implementation of the parametric stereo tool is not required. The decoder would skip the parametric stereo data and only decode the mono portion of the signal.

- the stereo-to-mono-downmix tool is required in order to be able to decode M/S stereo bitstreams.

- implementation of the SBR HQ tool is not required. Instead it is recommended to only implement the SBR Low Power tool since it allows for reduced computational complexity and lower memory requirements

Figure 1 illustrates how the AAC, SBR and the Parametric Stereo tools are combined to form the enhanced aacPlus codec: aacPlus consists of AAC and SBR. Enhanced aacPlus consists of aacPlus and the additional Parametric Stereo tool. Enhanced aacPlus is thus a true superset of aacPlus and AAC.
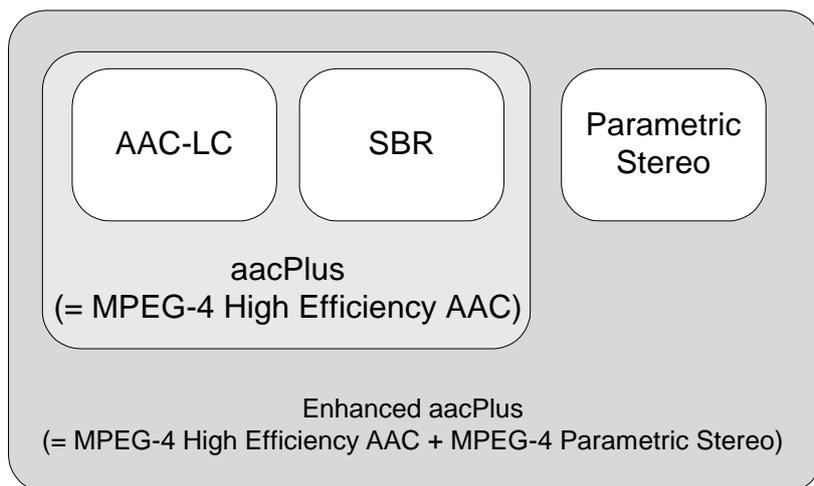

Figure 1: MPEG tools used to form the Enhanced aacPlus codec

6 Enhanced aacPlus general audio codec: ANSI-C code

The Floating-point ANSI –C-code of the general audio codec Enhanced aacPlus is described in [1]. The Fixed-point ANSI –C-code of the general audio codec Enhanced aacPlus is described in [10].

7 Enhanced aacPlus general audio codec: Enhanced aacPlus encoder

Figure 2 shows a block diagram of the Enhanced aacPlus encoder. The input PCM time domain signal is first fed to a stereo-to-mono downmix unit, which is only applied if the input signal is stereo but the chosen audio encoding mode is selected to be mono.

Next, the (mono or stereo) input time domain signal is fed to an IIR resampling filter in order to adjust the input sampling rate fs_{in} to the best-suited sampling rate fs_{enc} for the encoding process. The usage of the IIR resampler is only applied if the input signal sampling rate differs from the encoding sampling rate. The IIR resampler may either be run as a 3:2 downsampler (e.g. to downsample from 48 kHz to 32 kHz) or as a 1:2 upsampler (e.g. to upsample from 16 to 32 kHz).

The Enhanced aacPlus encoder basically consists of the well-known AAC[2] (Advanced Audio Coding) waveform encoder, the SBR (Spectral Band Replication) high frequency reconstruction encoding tool and the Parametric Stereo encoding tool. The Enhanced aacPlus encoder is operating in a dual rate mode, whereas the SBR encoder operates at the encoding sampling rate fs_{enc} as delivered from the IIR resampler and the AAC encoder at half of this sampling rate

---

[2] AAC has been standardized as recommended audio codec in 3GPP, Release 5

$fs_{enc}/2$. Consequently a 2:1 downsampler is present at the input to the AAC encoder. For an efficient implementation an IIR (Infinite Impulse Response) filter algorithm is used. The Parametric Stereo tool is used for low-bitrate stereo coding, i.e. at and below a bitrate of 44 kbit/s. The AAC encoder implementation complies with the AAC Low Complexity Object Type [5].
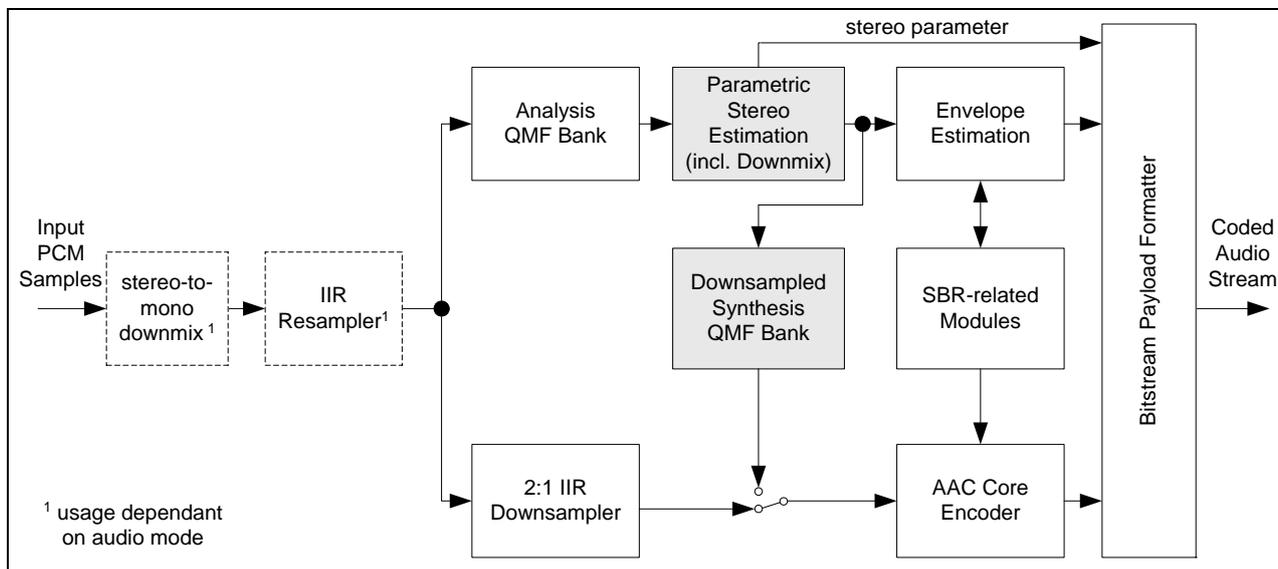

Figure 2: Enhanced aacPlus Encoder overview

The SBR encoder consists of a QMF (Quadrature Mirror Filter) analysis filter bank, which is used to derive the spectral envelope of the original input signal. Furthermore the SBR related modules control the selection of a input signal adaptive grid partitioning of the QMF samples on the time axis (i.e. control the framing), analyse the relation of noise floor to tonal components in the high band, collect guidance information for the transposition process in the decoder and detect missing harmonic components which could not be reconstructed by pure transposition. This gathered information about the characteristics of the input signal, together with the spectral envelope data forms the SBR stream. The amount of bits for the SBR stream is subtracted from the bits available to the AAC encoder in order to achieve a constant bitrate encoding of the multiplexed Enhanced aacPlus stream.

For stereo bitrates at and below 44 kbit/s, the Parametric Stereo encoding tool in the Enhanced aacPlus encoder is used. For stereo bitrates above 44 kbit/s, normal stereo operation is performed. The Parametric Stereo encoding tool estimates parameters characterizing the perceived stereo image of the input signal. These stereo parameters are embedded in the SBR stream. At the same time, a signal-adaptive mono downmix of the input signal is generated in the QMF domain and fed into the SBR encoder operating in mono. This downmix is also processed by a downsampled QMF synthesis filterbank to obtain the time domain input signal for the AAC core encoder with the sampling rate $fs_{enc}/2$. In this case, the 2:1 IIR downsampler is not active.

The embedding of the SBR stream (including the Parametric Stereo data) into the AAC stream is done in a backwards compatible way, i.e. a legacy Release 5 AAC decoder is able to parse the Enhanced aacPlus stream and decode the AAC core part.

The Enhanced aacPlus encoder is described in detail in [2], [3] and [4].

8 Enhanced aacPlus general audio codec: Enhanced aacPlus decoder

Figure 3 shows a block diagram of the Enhanced aacPlus decoder. In the decoder the bitstream is de-multiplexed into the AAC and the SBR stream. Error concealment, e.g. in case of frame loss, is achieved by designated algorithms in the decoder for AAC, SBR and Parametric Stereo: the AAC core decoder employs signal-adaptive spectrally shaped noise generation for error concealment, in the SBR and Parametric Stereo decoders, error concealment is based on extrapolation of guidance, envelope, and stereo information.

For the SBR processing, the Low-Power tool of SBR as described in [6] is used for full stereo decoding in order to keep the peak computational complexity as low as possible over all channel modes. Usage of the SBR Low-Power tool provides a computational complexity of an aacPlus stereo decoder in the same range as plain AAC stereo decoders.

The lowband AAC time domain signal, sampled at $fs_{enc}/2$, is first fed to a 32-channel QMF analysis filter bank. The QMF lowband samples are then used to generate a highband signal, whereas the transmitted transposition guidance information is used to best match the original input signal characteristics.

The transposed highband signal is then adjusted according to the transmitted spectral envelope signal to best match the original's spectral envelope. Also, missing components that could not be reconstructed by the transposition process are introduced. Finally, the lowband and the reconstructed highband are combined to obtain the complete output signal in the QMF domain.

In case of a stream using parametric stereo, the mono output signal from the underlying aacPlus decoder is converted into a stereo signal. This processing is carried out in the QMF domain and is controlled by the parametric stereo parameters embedded in the SBR stream. The relevant blocks for the Parametric Stereo operation are highlighted using grey background colour in Figure 3.

A 64-channel QMF synthesis filter bank is used to obtain the time domain output signal, sampled at the encoding sampling rate fs_{enc}. The synthesis filter bank may also be used to apply an implicit downsampling by a factor of 2, resulting in an output sampling rate of $fs_{enc}/2$.
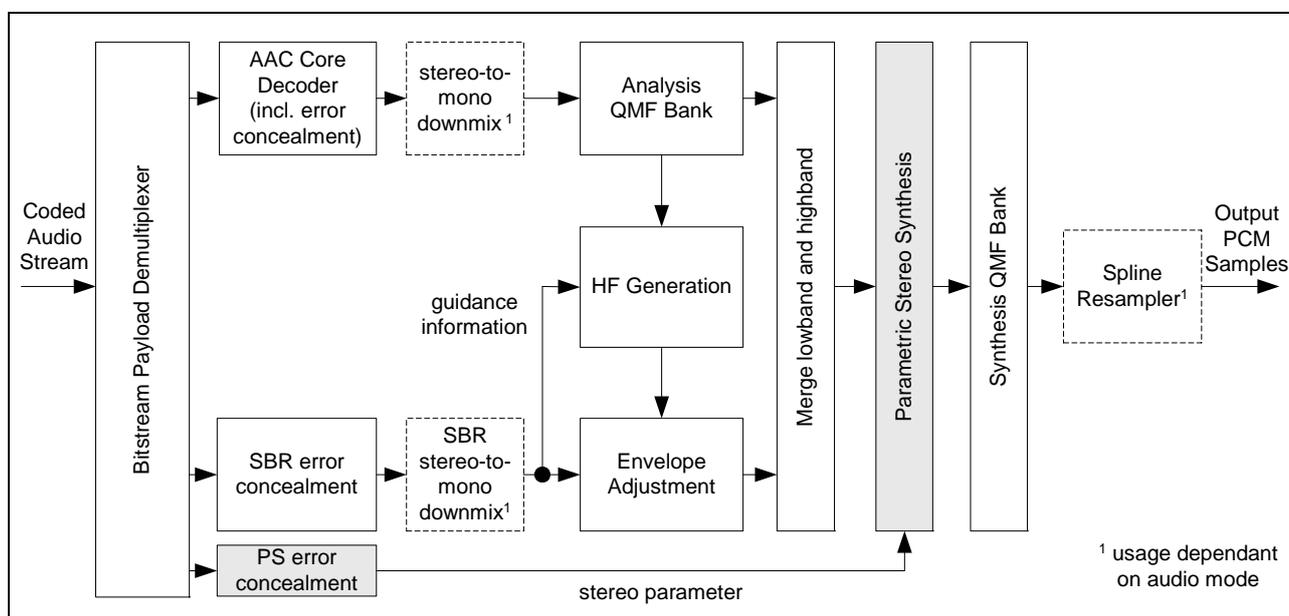

Figure 3: Enhanced aacPlus Decoder overview

The Enhanced aacPlus decoder is described in [5], [6], [7] and [8]. This description does not cover the additional decoder tools; error-concealment, SBR stereo-to-mono downmix and the spline resampler, that are not part of the cited ISO standards and therefore explained in section 9.

9 Enhanced aacPlus general audio codec: Additional Decoder Tools

Three additional tools are incorporated in the Enhanced aacPlus that are not part of the cited ISO standards. These are a error concealment algorithm, stereo-to-mono downmix, and a spline resampler.

The error concealment, e.g. in case of frame loss, is achieved by designated algorithms in the decoder for AAC, SBR andParametric Stereo: the AAC core decoder employs signal-adaptive spectrally shaped noise generation for error concealment, in the SBR and Parametric Stereo decoders, error concealment is based on extrapolation of guidance, envelope, and stereo information.

If the transmitted stream is a M/S stereo stream, but a monophonic output is requested, for each of the two components a stereo-to-mono downmix tool is available. In case of AAC the downmix is applied in the time-domain after AAC decoding. In case of SBR the stereo SBR stream is mapped to a mono SBR stream, thus resulting in low computational complexity since all further processing is then done on one channel only. If the transmitted stream uses parametric stereo, but a monophonic output is requested, the Parametric Stereo decoder is deactivated.

Finally a spline resampler algorithm is used to match the Enhanced aacPlus decoder output sampling rate to any arbitrary sampling rate. The spline resampler is only used if the handset requires any other specific output sampling rate different from fs_{enc} or $fs_{enc}/2$, e.g. 8 or 16 kHz if fs_{enc} is 44.1 kHz.

The additional decoder tools are described in [9].

10 Enhanced aacPlus general audio codec: Compatibility

Due to the modular approach the enhanced aacPlus encoder also includes a fully featured aacPlus and AAC-LC encoder.

A further consequence of the modular approach is decoder backwards compatibility: a Release-5 AAC decoder is able to decode the AAC part of an Enhanced aacPlus bitstream which contains both, SBR and Parametric Stereo data. However playback quality will be significantly limited. It is therefore recommended to not use this kind of compatibility unless it is specifically desired. Restricted backward compatibility can be accomplished by selecting the appropriate signalling as described in [6].

Table 1 illustrates the bitstream and decoding compatibilities as outlined above.

Table 1: Playback capabilities of decoders

		Decoder			
		Enhanced aacPlus		Release-5 AAC-LC, backwards compatible signalling	Release 5 AAC-LC, non backwards compatible signalling
Encoder mode	Enhanced aacPlus stereo, at and below 44 kbit/s	Yes		Mono only, no SBR	No
	Enhanced aacPlus mono or stereo above 44 kbit/s	Yes		No SBR	No
	Release 5 AAC-LC	Yes		Yes	Yes

Note: the table does not contain information on AAC-LTP, which is an optional audio codec since Release 5.

11 SBR Signalling in Payload formats

The decoder shall support all three signalling types defined in [6]. If implicit signalling is used, AAC-LC shall be signalled as described in Rel.5 of TS 26.234, in order to maintain backwards compatibility. If, in such a case, the sampling rate as indicated by the AAC object type descriptor (in the SDP) is 24 kHz or below and "SBR-enabled" (see [11]) is not specified in the SDP (i.e. the content does not originate from a Rel.6 server), the decoder output shall be configured to twice the AAC sampling rate.

Annex A (informative): Change history

Change history							
Date	**TSG SA#**	**TSG Doc.**	**CR**	**Rev**	**Subject/Comment**	**Old**	**New**
2004-09	25	SP-040633			Approved at TSG SA#25	2.0.0	6.0.0
2004-12	26	SP-040840	001	1	Alignment with C-code: Clarification on SBR mode to be used for mono only capable decoders	6.0.0	6.1.0
2005-03	27	SP-050095	002	1	Correction to written specification: add missing reference to MPEG corrigendum	6.1.0	6.2.0
2006-06	32	SP-060360	0004	1	Modification of written specification: Change of encoder bitrate border for Parametric Stereo usage	6.2.0	7.0.0