AMR-WB+
audio codec for low bit rate
streaming/MMS services

Pasi Ojala
Nokia Research Center
AMR-WB+ content

- Justification
- AMR-WB extension
- Functionalities
- Performance
- 3GPP Standardisation
AMR-WB+ justification

- 3GPP Release-5 (Year 2002) Messaging and Streaming specifications have speech codec and audio codecs
  - AMR for narrow band speech
  - AMR-WB for wideband speech
  - MPEG-4 AAC LC and LTP for audio (recommended)

- There was no low bit rate audio in 3GPP
  - AMR-WB was originally proposed for audio as well
  - Reasonable bit rates for AAC are well above 32 kbit/s

- Possibility to extend AMR-WB for audio
  - Keep speech performance
  - Add audio performance
  - Add new functionality

- Use cases only in streaming and messaging
  - Relaxed delay and complexity requirements

- AMR-WB+ standardized by 3GPP for Release 6 specifications
AMR-WB+ justification

Existing audio codecs

- AMR-WB+ for 16 kHz speech
- 16 kHz speech and audio
- 24 kHz speech and audio

Open standards, codecs available from 3GPP

- 4 kbps
- 8 kbps
- 12 kbps
- 16 kbps
- 24 kbps
- 32 kbps
- 64 kbps

44 kHz high quality audio
24 kHz high quality audio
16 kHz low end audio
16 kHz wideband Speech
8 kHz speech
AMR-WB and AMR-WB extension
AMR-WB+ functionality

- 3GPP set the design constraints and performance requirements
  - Limits to complexity and memory consumption
  - Performance as good as or better than AMR-WB and AAC-LC
- Functionality
  - Wide audio band
  - Stereo output needed
  - Low complexity encoder for terminal base content creation (MMS)
AMR-WB+ functionality

- Audio performance with hybrid time and frequency domain coding
  - ACELP for speech and transients
  - Transform codec excitation (TCX) for stationary signals
    - Fourier transform with lattice quantisation
- Higher audio bandwidth
  - Bandwidth extension: from 16 kHz up to 24 kHz
  - Overclocking to increase rate to 48 kHz
- Stereo
  - Parametric stereo coding
- Delay
  - Increased delay for transform coding
  - 80 ms superframe (20 ms transport)
- Bit allocation
  - Based on AMR-WB
- Bit rate
  - 10 kbit/s mono -> 24 kbit/s stereo -> 48 kbit/s stereo
AMR-WB+ encoder

Input signal L
Input signal R
Input signal M

preprocessing and analysis filterbank

LHF

HF signals folded in 0-6400 Hz band

RHF

HF encoding

HF parameters

Mode

Mono LF parameters

ACELP/TCX encoding

Stereo parameters

Mono operation

MUX

L parameters

R parameters

M parameters

S parameters

Down Mixing (L,R) to (M,S)

Stereo encoding

HF parameters
AMR-WB+ decoder

- **HF parameters**
  - HF decoding
  - HF decoding
  - Synthesis filterbank and postprocessing

- **Mono LF parameters**
  - ACELP/TCX decoding
  - Stereo decoding

- **Stereo parameters**
  - Stereo decoding

- **DEMUX**
  - Mono operation

- **Output signals**
  - L
  - M
  - R

- **HF signals folded in 0-6400 Hz band**
  - L<sub>HF</sub>
  - R<sub>HF</sub>
  - M<sub>HF</sub>
  - M<sub>LF</sub>
  - L<sub>LF</sub>
  - R<sub>LF</sub>
AMR-WB+ performance

- 3GPP conducted subjective listening tests to assess the performance
  - Evaluation against performance requirements
  - Competition against MPEG-4 HE AAC and Enhanced aacPlus
  - Separate tests for low and high bit rate codecs

- Low bit rate experiments
  - 14 kbit/s mono (use case A and bandwidth limited use case B)
  - 18 kbit/s stereo (use case A and B)
  - 24 kbit/s mono
  - 24 kbit/s stereo
  - Channel errors

- Audio content
  - Speech, music and mixed content

- MUSHRA methodology
Low-rate test global results

MUSHRA score

Condition

A1: 14kbs, mono, use case A
A2: 18kbs, stereo, use case A
A3: 24kbs, mono, use case A
A4: 24kbs, stereo, use case A
B1: 14kbs, mono, use case B, 16kHz I/O
B2: 18kbs, stereo, use case B
B3: 14kbs, mono, use case A, 3% FER
B4: 24kbs, stereo, use case A, 3% FER
AMR-WB+ performance

- Comparison of the candidates
  - AAC+
  - MPEG-4 HE AAC
  - CT
  - MPEG-4 Enhanced aacPlus (eAAC+)
  - AMR-WB+

<table>
<thead>
<tr>
<th>Test</th>
<th>Operating condition</th>
<th>AAC+</th>
<th>AMR-WB+</th>
<th>eAAC+</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>14 kbps, mono, use case A (PSS)</td>
<td>50.8</td>
<td>62.6</td>
<td>51.5</td>
</tr>
<tr>
<td>A2</td>
<td>18 kbps, stereo, use case A (PSS)</td>
<td>37.5</td>
<td>55.6</td>
<td>53.3</td>
</tr>
<tr>
<td>A3</td>
<td>24 kbps, mono, use case A (PSS)</td>
<td>75.0</td>
<td>67.4</td>
<td>75.8</td>
</tr>
<tr>
<td>A4</td>
<td>24 kbps, stereo, use case A (PSS)</td>
<td>55.3</td>
<td>61.3</td>
<td>67.1</td>
</tr>
<tr>
<td>B1</td>
<td>14 kbps, mono, use case B (MMS), 16 kHz inp. and outp. sampling rate</td>
<td>45.5</td>
<td>50.7</td>
<td>44.4</td>
</tr>
<tr>
<td>B2</td>
<td>18 kbps, stereo, use case B (MMS)</td>
<td>43.3</td>
<td>50.7</td>
<td>55.7</td>
</tr>
<tr>
<td>B3</td>
<td>14 kbps, mono, use case A (PSS), 3% FER</td>
<td>43.1</td>
<td>52.5</td>
<td>44.3</td>
</tr>
<tr>
<td>B4</td>
<td>24 kbps, stereo, use case A (PSS), 3% FER</td>
<td>48.9</td>
<td>53.3</td>
<td>58.0</td>
</tr>
</tbody>
</table>