Source:	**TSG_SA WG4**
Title:	**AMR-WB speech codec; test sequences (Release 4)**
Document for:	**Approval**

Presentation of Specification to TSG-SA

Presentation to:	**TSG SA Meeting #11**
Document for presentation:	**TS 26.174 Version 2.0.0**
Presented for:	**Approval**

Abstract of document:

The present document specifies the digital test sequences for the adaptive multi-rate wideband (AMR-WB) speech codec. These sequences test for a bit-exact implementation of the adaptive multi-rate wideband (AMR-WB) speech transcoder (TS 26.190), voice activity detection (TS 26.194), comfort noise (TS 26.192), and source controlled rate operation (TS 26.193).

Changes since last presentation to TSG Meeting #10:

Editorial changes.

Outstanding Issues:

None.

Contentious Issues:

None.

3GPP TS 26.174 V2.0.0 (2001-03)

Technical Specification

**3rd Generation Partnership Project;
Technical Specification Group Services and System Aspects;
Speech codec speech processing functions;
AMR Wideband Speech Codec test sequences
(Release 4)**

**GLOBAL SYSTEM FOR
MOBILE COMMUNICATIONS**

Keywords

<keyword[, keyword]>

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

Contents

Foreword

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

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 the present document, 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 or greater 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 document.

1 Scope

The present document specifies the digital test sequences for the adaptive multi-rate wideband (AMR-WB) speech codec. These sequences test for a bit-exact implementation of the adaptive multi-rate wideband (AMR-WB) speech transcoder (TS 26.190 [2]), voice activity detection (TS 26.194 [5]), comfort noise (TS 26.192 [3]), and source controlled rate operation (TS 26.193 [4]).

2 Normative references

This document 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 document only when incorporated in it by amendment or revision. For undated references, the latest edition of the publication referred to applies.

[1]   3GPP TS 26.201: "AMR wideband speech codec; Frame structure".

[2]   3GPP TS 26.202: "AMR Wideband Speech Codec; Interface to RAN".

[3]   3GPP TS 26.190: "AMR Wideband Speech Codec; Transcoding functions".

[4]   3GPP TS 26.193: "AMR Wideband Speech Codec; Source Controlled Rate operation".

[5]   3GPP TS 26.194: "AMR wideband speech codec; Voice Activity Detection (VAD)".

[6]   3GPP TS 26.201: "AMR Wideband Speech Codec; Frame structure".

[7]   3GPP TS 26.173: "AMR wideband speech codec; ANSI-C code".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in TS 26.190 [2], TS 26.091 [6], TS 26.192 [3], TS 26.193 [4] and TS 26.194 [5] apply.

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

4 General

Digital test sequences are necessary to test for a bit exact implementation of the adaptive multi-rate wideband (AMR-WB) speech transcoder (TS 26.190 [2]), voice activity detection (TS 26.194 [5]), comfort noise generation (TS 26.192 [3]), and source controlled rate operation (TS 26.193 [4]).

The test sequences may also be used to verify installations of the ANSI C code in TS 26.173 [7].

Clause 5 describes the format of the files which contain the digital test sequences. Clause 6 describes the test sequences for the speech transcoder. Clause 7 describes the test sequences for the VAD, comfort noise and source controlled rate operation.

Clause 8 describes the method by which synchronisation is obtained between the test sequences and the speech codec under test.

5 Test sequence format

This clause provides information on the format of the digital test sequences for the adaptive multi-rate wideband (AMR-WB) speech transcoder (TS 26.190 [3]), voice activity detection (TS 26.194 [5]), comfort noise generation (TS 26.192 [3]), and source controlled rate operation (TS 26.193 [4]).

5.1 File format

The test sequence files in PC (little-endian) byte order are provided in archive files (ZIP format) which accompany the present document.

Following decompression, three types of file are provided:

- Files for input to the speech encoder: *.INP

- Files for comparison with the encoder output and for input to the speech decoder: *.COD

- Files for comparison with the decoder output: *.OUT

- One mode control file for the mode switching test   T22.MOD

All file formats are described in TS 26.173 [7].

5.2 Codec homing

Each *.INP file includes two homing frames (see TS 26.173 [7]) at the start of the test sequence. The function of these frames is to reset the speech encoder state variables to their initial value. In the case of a correct installation of the ANSI-C simulation (TS 26.173 [7]), all speech encoder output frames shall be identical to the corresponding frame in the *.COD file. In the case of a correct hardware implementation undergoing testing, the first speech encoder output frame is undefined and need not be identical to the first frame in the *.COD file, but all remaining speech encoder output frames shall be identical to the corresponding frames in the *.COD file.

The function of the two homing frames in the *.COD files is to reset the speech decoder state variables to their initial value. In the case of a correct installation of the ANSI-C simulation (TS 26.173 [7]), all speech decoder output frames shall be identical to the corresponding frame in the *.OUT file. In the case of a correct hardware implementation undergoing testing, the first speech decoder output frame is undefined and need not be identical to first frame in the *.OUT file, but all remaining speech decoder output frames shall be identical to the corresponding frames in the *.OUT file.

6 Speech codec test sequences

This clause describes the test sequences designed to exercise the adaptive multi-rate wideband (AMR-WB) speech transcoder (TS 26.190 [3]).

6.1 Codec configuration

The speech encoder shall be configured not to operate in the source controlled rate mode.

6.2 Speech codec test sequences

6.2.1 Speech encoder test sequences

Twenty-three encoder input sequences are provided. Note that for the input sequences T00.INP to T03.INP, the amplitude figures are given in 14-bit precision. The active speech levels are given in dBov.

- T00.INP - Synthetic harmonic signal. The pitch delay varies slowly from 34 to 231 samples. The minimum and maximum amplitudes are -1475 and +5952.

- T01.INP - Synthetic harmonic signal. The pitch delay varies slowly from 231 down to 34 samples. Amplitudes at saturation point -5386 and +21707.

- T02.INP - Square sweep varying from 50 Hz to 7000 Hz. Amplitudes ± 32767.

- T03.INP - Sinusoidal sweep varying from 50 Hz to 7000 Hz. Amplitudes ± 6217.

- T04.INP - Female speech, ambient noise, active speech level: -22.5 dBov, P.341 filtered.

- T05.INP - Male speech, ambient noise, active speech level: -29.9 dBov, P.341 filtered.

- T06.INP – Female and male speech, ambient noise, active speech level: -36.1 dBov, P.341 filtered.

- T07.INP - Female and male speech, ambient noise, active speech level: -45.8 dBov, P.341 filtered.

- T08.INP - Female and male speech, ambient noise, active speech level: -7.7 dBov, P.341 filtered.

- T09.INP - Female and male speech, Hoth noise, active speech level: -37.4 dBov, P.341 filtered.

- T10.INP - Female and male speech, Hoth noise, active speech level: -27.3 dBov, P.341 filtered.

- T11.INP - Female and male speech, Hoth noise, active speech level: -16.9 dBov, P.341 filtered.

- T12.INP - Female and male speech, ambient noise, active speech level: -46.0 dBov, P.341 filtered.

- T13.INP - Speech, very high and low car noise, P.341 filtered.

- T14.INP - Female and male speech, ambient noise, active speech level: -26.0 dBov, P.341 filtered.

- T15.INP - Female and male speech, rain noise, active speech level: -37.2 dBov, P.341 filtered.

- T16.INP - Female and male speech, rain noise, active speech level: -26.5 dBov, P.341 filtered.

- T17.INP - Female and male speech, rain noise, active speech level: -16.4 dBov, P.341 filtered. This file includes homing frame test.

- T18.INP - Male speech, active speech level: -29.7 dBov, P.341 filtered, with many zero frames.

- T19.INP - Child speech, ambient noise, active speech level: -34.7 dBov, P.341 filtered.

- T20.INP - Sequence for exercising the LPC vector quantisation codebooks and ROM tables of the codec.

- T21.INP - Zero signal sequence.

- T22.INP - Speech sequence for mode switching test.

The output using these input sequences will be different depending on the tested adaptive multi-rate mode. In the notation used below <mode> should be changed to the number of the tested mode, i.e. one of 2385, 2305, 1985, 1825, 1585, 1425, 1265, 885 or 660.

The T00.INP and T01.INP sequences were designed to test the pitch lag of the adaptive multi-rate wideband speech encoder. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T00_<mode>.COD and T01_<mode>.COD sequences, respectively.

The T02.INP and T03.INP sequences are particularly suited for testing the LPC analysis, as well as for finding saturation problems. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T02_<mode>.COD and T03_<mode>.COD sequences, respectively.

The T04.INP and T05.INP sequences contain a lot of low-frequency components. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T04_<mode>.COD and T05_<mode>.COD sequences, respectively.

The T18.INP and T21.INP sequences contain "all zeros" frames (silence) in between segments of speech. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T18_<mode>.COD and T21_<mode>.COD sequences, respectively.

The T20.INP sequence was designed to exercise the LPC code indices and the ROM table indices of the codec.

The sequences T06.INP to T17.INP and T19.INP were selected on the basis of bringing various input characteristics (background noise) and levels to the test sequence set. Homing frame test is also included in T17.INP. T17.INP has homing frames with length 320 smp, 640 smp and 960 smp starting from 32000 smp, 16000 smp and 48000 smp in a respective order. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T06_<mode>.COD to T17_<mode>.COD sequences, respectively.

The T22.INP sequence was designed to test mode switching in the encoder. For testing mode switching this sequence is used together with the mode control file T22.MOD. See TS 26.173 [7] for the format of the mode control file. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the sequence T22.COD. Note that T22.COD contains parameter frames in different codec modes.

6.2.2 Speech decoder test sequences

Twenty-two times nine speech decoder input sequences TXX_<mode>.COD (XX = 00..21, <mode> = {2385, 2305, 1985, 1825, 1585, 1425, 1265, 885 or 660}) are provided for the static mode tests. These are the output of the corresponding TXX.INP sequences, one set per mode. In a correct implementation, the resulting speech decoder output shall be identical to the corresponding TXX_<mode>.OUT sequences.

The switching test decoder input T22.COD shall result in decoder output identical to the T22.OUT sequence. For the decoder switching test no special mode control file is needed since the mode information is included in the .COD file according to the file format (see TS 26.173 [7]).

6.2.3 Codec homing sequence

In addition to the test sequences described above, the homing sequences are provided to assist in codec testing. T23.INP contains one encoder-homing-frame. The sequences T23_<mode>.COD (<mode> = {2385, 2305, 1985, 1825, 1585, 1425, 1265, 885 or 660}) contain one decoder-homing-frame each for the corresponding mode. The use of these sequences is described in TS 26.171 [1].

All files are contained in the archive T.zip which accompanies the present document.

7 Test sequences for source controlled rate operation

This clause describes the test sequences designed to exercise the VAD algorithm (TS 26.194 [5]), comfort noise (TS 26.192 [3]), and source controlled rate operation (TS 26.193 [4]).

Test sequences DTX1.*, DT2.*, DTX4.* and DTX5.* shall be run only with speech codec 23.85 kbit/s. Test sequence DTX3.* shall be run for all the speech codec modes.

7.1 Codec configuration

The VAD, comfort noise and source controlled rate operation shall be tested in conjunction with the speech coder (TS 26.190 [2]). The speech encoder shall be configured to operate in the source controlled rate mode, with VAD.

7.2 Test Sequences

Each DTX test sequence consists of three files:

- Files for input to the speech encoder: *.INP

- Files for comparison with the encoder output and input to the speech decoder: *.COD

- Files for comparison with the decoder output: *.OUT

The *.COD and *.OUT file names has the format DTXA_<mode>.*, "A" is the test case number (1, 2, 3, 4 or 5) and <mode> is the speech codec mode.

In a correct implementation, the speech encoder parameters generated by the *.INP file shall be identical to those specified in the *.COD file; and the speech decoder output generated by the *.COD file shall be identical to that specified in the *.OUT file.

7.2.1 Test sequences for background noise estimation

Background noise estimation algorithm is tested by the following test sequences:

DTX1.*

DTX2.*

7.2.2 Test sequences for tone signal detection

Tone signal detection algorithm is tested by the following test sequence:

DTX3.*

7.2.3 Real speech and tones

This test sequence consists of very clean speech, barely detectable speech and a swept frequency tone.

DTX4.*

7.2.4 Test sequence for signal-to-noise ratio estimation

The full range of SNR estimates are tested by the following test sequence:

DTX5.*

8 Sequences for finding the 20 ms framing of the adaptive multi-rate speech encoder

When testing the decoder, alignment of the test sequences used to the decoder framing is achieved by the air interface (testing of MS) or can be reached easily on the Abis-interface (testing on network side).

When testing the encoder, usually there is no information available about where the encoder starts its 20 ms segments of speech input to the encoder.

In the following, a procedure is described to find the 20 ms framing of the encoder using special synchronisation sequences. This procedure can be used for MS as well as for network side.

Synchronisation can be achieved in two steps. First, bit synchronisation has to be found. In a second step, frame synchronisation can be determined. This procedure takes advantage of the codec homing feature of the adaptive multi-rate codec, which puts the codec in a defined home state after the reception of the first homing frame. On the reception of further homing frames, the output of the codec is predefined and can be triggered to.

8.1 Bit synchronisation

The input to the speech encoder is a series of 14 bit long words (224 kbit/s, 14 bit linear PCM). When starting to test the speech encoder, no knowledge is available on bit synchronisation, i.e., where the encoder expects its least significant bits, and where it expects the most significant bits.

The encoder homing frame consists of 320 samples, all set to 0x0008 hex. If two such encoder homing frames are input to the encoder consecutively, the corresponding decoder homing frame of the used codec mode is expected at the output as a reaction of the second encoder homing frame.

Since there are only 14 possibilities for bit synchronisation, after a maximum of 14 trials bit synchronisation can be reached for each codec mode. In each trial three consecutive encoder homing frames are input to the encoder. If the corresponding decoder homing frame is not detected at the output, the relative bit position of the three input frames is shifted by one and another trial is performed. As soon as the decoder homing frame of the used codec mode is detected at the output, bit synchronisation is found, and the first step can be terminated.

The reason why three consecutive encoder homing frames are needed is that frame synchronisation is not known at this stage. To be sure that the encoder reads two complete homing frames, three frames have to be input. Wherever the encoder has its 20 ms segmentation, it will always read at least two complete encoder homing frames.

An example of the 14 different frame triplets is given in sequence BITSYNC.INP.

8.2 Frame synchronisation

Once bit synchronisation is found, frame synchronisation can be found by inputting two identical frames consecutively to the encoder. There exist 320 different output sequences depending on the 320 different positions that the beginning of this sequence of frames can possibly have with respect to the encoder framing.

Before inputting this special synchronisation sequence to the encoder, again the encoder has to be reset by one encoder homing frame. A second encoder homing frame is needed to provoke a decoder homing frame at the output that can be triggered to. And since the framing of the encoder is not known at that stage, three encoder homing frames have to precede the special synchronisation sequence to ensure that the encoder reads at least two homing frames, and at least one decoder homing frame is produced at the output, serving as a trigger for recording.

After the last decoder homing frame of the used codec mode it is required to detect two consecutive output frames that are different from the preceding decoder homing frame.

The special synchronisation sequence preceded by three encoder homing frames are given in SEQSYNC.INP.

Generally, the output sequences will be different depending on the tested adaptive multi-rate wideband mode. In the notation below <mode> should be changed to the number of the tested mode, i.e. one of 2385, 2305, 1985, 1825, 1585, 1425, 1265, 885 or 660.

In all 320 output sequences only the second frame after the last decoder homing frame is given in SYNC000_<mode>.COD through SYNC319_<mode>.COD. These output frames were calculated by shifting the sequence SEQSYNC.INP through the positions 0 to 319, where the samples at the beginning were set to zero. For each codec mode it was finally verified that the last frame in each of the 320 output sequences is different to all other last frames.

The three digit number in the filenames above indicates the number of samples by which the input was retarded with respect to the encoder framing. By a corresponding shift in the opposite direction, alignment with the encoder framing for the used codec mode can be reached.

8.3 Formats and sizes of the synchronisation sequences

BITSYNC.INP:

This sequence consists of 14 frame triplets. It has the format of the speech encoder input test sequences.

The size of it is therefore:

SIZE (BITSYNC.INP) = 14 * 3 * 320 * 2 bytes = 26880 bytes

SYNCXXX_<mode>.COD:

These sequences consists of 1 encoder output frame each. They have the format of the speech encoder output test sequences. In these frames the values of the TX/RX_TYPE is fixed to indicate transmit frame type and FRAME_TYPE and MODE_INFO fields are set to the transmit frame type and to the corresponding encoding mode information [3].

The sizes of them are therefore:

SIZE (SYNCXXX_2385.COD) = (477 + 3) * 2 bytes = 960 bytes
SIZE (SYNCXXX_2305.COD) = (461 + 3) * 2 bytes = 928 bytes
SIZE (SYNCXXX_1985.COD) = (397 + 3) * 2 bytes = 800 bytes
SIZE (SYNCXXX_1825.COD) = (365 + 3) * 2 bytes = 736 bytes
SIZE (SYNCXXX_1585.COD) = (317 + 3) * 2 bytes = 640 bytes
SIZE (SYNCXXX_1425.COD) = (285 + 3) * 2 bytes = 576 bytes
SIZE (SYNCXXX_1265.COD) = (253 + 3) * 2 bytes = 512 bytes
SIZE (SYNCXXX_885.COD) = (177 + 3) * 2 bytes = 360 bytes
SIZE (SYNCXXX_660.COD) = (132 + 3) * 2 bytes = 270 bytes

All files are contained in the archive S.zip which accompanies the present document.

Annex A (informative):
Change history

Change history							
Date	TSG #	TSG Doc.	CR	Rev	Subject/Comment	Old	New