INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.722.2

Annex D
(01/2002)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Digital terminal equipments – Coding of analogue signals by methods other than PCM

Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)

Annex D: Digital test sequences

ITU-T Recommendation G.722.2 – Annex D

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TESTING EQUIPMENTS	G.500–G.599
TRANSMISSION MEDIA CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
General	G.700–G.709
Coding of analogue signals by pulse code modulation	G.710–G.719
Coding of analogue signals by methods other than PCM	**G.720–G.729**
Principal characteristics of primary multiplex equipment	G.730–G.739
Principal characteristics of second order multiplex equipment	G.740–G.749
Principal characteristics of higher order multiplex equipment	G.750–G.759
Principal characteristics of transcoder and digital multiplication equipment	G.760–G.769
Operations, administration and maintenance features of transmission equipment	G.770–G.779
Principal characteristics of multiplexing equipment for the synchronous digital hierarchy	G.780–G.789
Other terminal equipment	G.790–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
QUALITY OF SERVICE AND PERFORMANCE	G.1000–G.1999
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DIGITAL TERMINAL EQUIPMENTS	G.7000–G.7999
DIGITAL NETWORKS	G.8000–G.8999

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation G.722.2

Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)

Annex D

Digital test sequences

Summary

This annex specifies the bit-exact test sequences for the verification of the implementation of G.722.2 AMR-WB codec, voice activity detection, comfort noise generation and source controlled rate operation.

The test sequences specified in this annex were also adopted by 3GPP in 3GPP specification TS 26.174.

These test sequences are freely available on the ITU-T website. They are also available for a fee on a CD-ROM from the ITU sales department at sales@itu.int.

Source

Annex D to ITU-T Recommendation G.722.2 was prepared by ITU-T Study Group 16 (2001-2004) and approved under the WTSA Resolution 1 procedure on 13 January 2002.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

		Page
D.1	Scope	1
D.2	General	1
D.3	Test sequence format	1
D.3.1	File format	1
D.3.2	Codec homing	1
D.4	Speech codec test sequences	2
D.4.1	Codec configuration	2
D.4.2	Speech codec test sequences	2
D.5	Test sequences for source controlled rate operation	4
D.5.1	Codec configuration	4
D.5.2	Test Sequences	4
D.6	Sequences for finding the 20 ms framing of the adaptive multi-rate speech encoder	5
D.6.1	Bit synchronisation	5
D.6.2	Frame synchronisation	5
D.6.3	Formats and sizes of the synchronisation sequences	6

Electronic attachment:

SyncVertors

TestVectors

DTX_TestVectors

ITU-T Recommendation G.722.2

Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)

Annex D

Digital test sequences

D.1 Scope

This annex specifies the digital test sequences for the adaptive multi-rate wideband (AMR-WB) speech codec specified in ITU-T Rec. G.722.2, its Annexes A and B, and its Appendix I. These sequences test for a bit-exact implementation of the adaptive multi-rate wideband speech transcoder, voice activity detection, comfort noise, and source controlled rate operation.

D.2 General

Digital test sequences are necessary to test for a bit exact implementation of the adaptive multi-rate wideband (AMR-WB) speech transcoder, voice activity detection, comfort noise generation, and source controlled rate operation.

The test sequences may also be used to verify installations of the ANSI C code.

Clause D.3 describes the format of the files which contain the digital test sequences. Clause D.4 describes the test sequences for the speech transcoder. Clause D.5 describes the test sequences for the VAD, comfort noise and source controlled rate operation.

Clause D.7 describes the method by which synchronisation is obtained between the test sequences and the speech codec under test.

D.3 Test sequence format

This clause provides information on the format of the digital test sequences for the adaptive multi-rate wideband (AMR-WB) speech, voice activity detection, comfort noise generation, and source controlled rate operation.

D.3.1 File format

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

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 Annex C/G.722.2.

D.3.2 Codec homing

Each *.INP file includes two homing frames (see Annex C/G.722.2) 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, 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, 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.

D.4 Speech codec test sequences

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

D.4.1 Codec configuration

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

D.4.2 Speech codec test sequences

D.4.2.1 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 +21 707.
– 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 32 000 smp, 16 000 smp and 48 000 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 Annex C/G.722.2 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.

D.4.2.2 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.

D.4.2.3 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.

All files are contained in the archive T.zip which accompanies this annex.

D.5 Test sequences for source controlled rate operation

This clause describes the test sequences designed to exercise the VAD algorithm, comfort noise, and source controlled rate operation.

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.

D.5.1 Codec configuration

The VAD, comfort noise and source controlled rate operation shall be tested in conjunction with the speech coder. The speech encoder shall be configured to operate in the source controlled rate mode, with VAD.

D.5.2 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.

D.5.2.1 Test sequences for background noise estimation

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

DTX1.*

DTX2.*

D.5.2.2 Test sequences for tone signal detection

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

DTX3.*

D.5.2.3 Real speech and tones

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

DTX4.*

D.5.2.4 Test sequence for signal-to-noise ratio estimation

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

DTX5.*

D.6 Sequences for finding the 20 ms framing of the adaptive multi-rate speech encoder

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. 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. This output can be used to trigger other actions.

D.6.1 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.

D.6.2 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.

D.6.3 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 consist 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.

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 this annex.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

22175

Printed in Switzerland
Geneva, 2002