



INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.862

Amendment 1
(03/2003)

SERIES P: TELEPHONE TRANSMISSION QUALITY,
TELEPHONE INSTALLATIONS, LOCAL LINE
NETWORKS

Methods for objective and subjective assessment of
quality

Perceptual evaluation of speech quality (PESQ): An
objective method for end-to-end speech quality
assessment of narrow-band telephone networks
and speech codecs

**Amendment 1: Revised Annex A: Source code
for reference implementation and conformance
tests**

ITU-T Recommendation P.862 (2001) – Amendment 1

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Subscribers' lines and sets	Series	P.30 P.300
Transmission standards	Series	P.40
Objective measuring apparatus	Series	P.50 P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of quality	Series	P.80 P.800
Audiovisual quality in multimedia services	Series	P.900

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

ITU-T Recommendation P.862

Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs

Amendment 1

Revised Annex A: Source code for reference implementation and conformance tests

Summary

This Annex A replaces the previous Annex A to ITU-T Rec. P.862 (February 2001), which defines the Perceptual Evaluation of Speech Quality (PESQ) model for objective speech quality measurement for telephone networks and speech codecs. It clarifies the sampling rate that must be offered by implementations of PESQ, and sets out an extended set of conformance tests that support this.

It is not anticipated that this Annex A will affect any existing implementations of PESQ. Future implementations of PESQ that only offer 8 kHz sampling rate are accepted as valid and must meet the new conformance test 1(b) proposed below. Future implementations at 16 kHz sampling rate must also meet the new conformance test 2(a).

The ANSI-C reference implementation of PESQ remains unchanged.

The only files that are modified relative to ITU-T Rec. P.862 (02/2001) are:

- `supp23_16k.txt`, `supp23_16k.bat`, `supp23_8k.txt`, `supp23_8k.bat`;
- `voipref_16k.txt`, `voipref_16k.bat`, `voipref_8k.txt`, `voipref_8k.bat`;
- `process.bat`.

These files are available for free on ITU Publications website at www.itu.int, under P.862 as "P.862 (2001) Amendment 1".

Source

Amendment 1 to ITU-T Recommendation P.862 was approved by ITU-T Study Group 12 (2001-2004) under the ITU-T Recommendation A.8 procedure on 16 March 2003.

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.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

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.

© ITU 2003

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

CONTENTS

	Page
Amendment 1 – Revised Annex A: Source code for reference implementation and conformance tests	1
A.1 Files provided	1
A.2 Sampling rate	2
A.3 Conformance tests	2

ITU-T Recommendation P.862

Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs

Amendment 1

Revised Annex A: Source code for reference implementation and conformance tests

A.1 Files provided

A.1.1 List of files provided for the ANSI-C reference implementation

The ANSI-C reference implementation of PESQ is contained in the following text files which are provided in the `source` sub-directory of the CD-ROM distribution:

- `dsp.c` *Basic DSP routines*
- `dsp.h` *Header file for dsp.c*
- `pesq.h` *General header file*
- `pesqdsp.c` *PESQ DSP routines*
- `pesqio.c` *File input/output*
- `pesqmain.c` *Main program*
- `pesqmod.c` *PESQ high-level model*
- `pesqpar.h` *PESQ perceptual model definitions*

The ANSI-C reference implementation provided in separate files and forms is an integral part of this Recommendation. The ANSI-C reference implementation shall take precedence in case of conflicts between the high level description as given in this amendment and the ANSI-C reference implementation.

These files form an integral part of this Annex A.

A.1.2 List of files provided for conformance validation

The conformance validation process described below makes reference to the following files, which are provided in the `conform` sub-directory of the CD-ROM distribution:

- `supp23_16k.txt` *File pairs and PESQ scores for test 1(a)*
- `supp23_16k.bat` *Batch script to assist with test 1(a)*
- `supp23_8k.txt` *File pairs and PESQ scores for test 1(b)*
- `supp23_8k.bat` *Batch script to assist with test 1(b)*
- `voipref_16k.txt` *File pairs and PESQ scores for test 2(a)*
- `voipref_16k.bat` *Batch script to assist with test 2(a)*
- `voipref_8k.txt` *File pairs and PESQ scores for test 2(b)*
- `voipref_8k.bat` *Batch script to assist with test 2(b)*
- `process.bat` *Sample batch script to assist with preparing material for tests 1(b) and 2(a)*

- *Speech files provided for validation with variable delay:*

or105.wav or109.wav or114.wav or129.wav or134.wav or137.wav
 or145.wav or149.wav or152.wav or154.wav or155.wav or161.wav
 or164.wav or166.wav or170.wav or179.wav or221.wav or229.wav
 or246.wav or272.wav dg105.wav dg109.wav dg114.wav dg129.wav
 dg134.wav dg137.wav dg145.wav dg149.wav dg152.wav dg154.wav
 dg155.wav dg161.wav dg164.wav dg166.wav dg170.wav dg179.wav
 dg221.wav dg229.wav dg246.wav dg272.wav
 u_ams01.wav u_ams02.wav u_ams03.wav
 u_ams01b1c1.wav u_ams01b1c7.wav u_ams01b1c15.wav u_ams02b1c9.wav
 u_ams03b1c16.wav u_ams03b1c18.wav u_ams01b2c1.wav u_ams01b2c8.wav
 u_ams02b2c4.wav u_ams02b2c5.wav u_ams02b2c14.wav u_ams03b2c5.wav
 u_ams03b2c6.wav u_ams03b2c7.wav u_ams03b2c11.wav u_ams03b2c18.wav
 u_af1s01.wav u_af1s02.wav u_af1s03.wav
 u_af1s01b2c16.wav u_af1s03b2c16.wav u_af1s02b2c17.wav u_af1s03b2c17.wav

These speech files are in Wave format (16-bit linear PCM, Intel byte ordering, 44 byte header), at 8 kHz sample rate.

These files form an integral part of this Annex A.

A.2 Sampling rate

An implementation of ITU-T Rec. P.862 may, at the implementer's discretion, operate at 8 kHz sampling rate, 16 kHz sampling rate, or both.

However, the implementation must pass the conformance tests set out in ITU-T Rec. P.862 and its annexes for all sampling rates that are offered by the implementation.

A.3 Conformance tests

A.3.1 Conformance data sets

The data sets for the conformance tests are as follows.

Test	Number of file pairs	(a) 16 kHz data set	(b) 8 kHz data set	Type of test
1	1736	ITU-T P-series Supplement 23	Downsampled from Supplement 23 using ITU-T Software Tools Library (version 2000, release 3 ¹) and <code>process.bat</code> .	Mandatory
2	40	Upsampled from P.862 VoIP variable delay data using Software Tools Library (version 2000, release 3) and <code>process.bat</code> .	ITU-T Rec. P.862 Annex A VoIP variable delay data.	Mandatory
3	No data set defined. This test is open-ended, based on general, unknown data.			Mandatory

¹ ITU-T Rec. G.191 (2000), *Software tools for speech and audio coding standardization*.

A.3.2 Conformance requirements

The test requirements are summarized in the following table and are set out in detail below. The requirements are based on the absolute difference in PESQ score between the implementation under test and the ANSI-C reference implementation, calculated for each reference and degraded file pair. For the conformance tests defined in Annex A/P.862 (February 2001), there is no change.

Test	Number of file pairs	Lower threshold	Upper threshold	Type of test
1(a)	1736	Difference may not exceed 0.05 in any situation.	Not applicable	Mandatory
1(b)	1736	Difference may exceed 0.05 in not more than 2 file pairs (approx. 0.1% of cases).	Difference may not exceed 0.1 in any case.	Mandatory
2(a)	40	Difference may exceed 0.05 in not more than 1 file pair (2.5% of cases).	Difference may not exceed 0.5 in any case.	Mandatory
2(b)	40	Difference may exceed 0.05 in not more than 1 file pair (2.5% of cases).	Difference may not exceed 0.5 in any case.	Mandatory
3	No data set defined	Difference may exceed 0.05 in not more than 0.5% of cases.	Difference may exceed 0.05 in not more than 5% of cases.	Lower threshold is advisory. Upper threshold is mandatory.

A.3.2.1(a) Conformance test 1(a) (16 kHz sampling rate)

In this test, all files from all ten experiments as released with ITU-T P-series Supplement 23 are used, on a file-by-file basis. The Supplement 23 data is all at 16 kHz sampling rate. The original and degraded file names, and the PESQ score given by the reference implementation, are provided in the files listed above.

An implementation passes this test when the absolute difference in PESQ score compared to the reference implementation is not greater than 0.05 in all cases.

This conformance test is mandatory for all implementations of PESQ at 16 kHz sampling rate.

Supplement 23 to the P-series Recommendations can be obtained separately from the ITU.

A.3.2.1(b) Conformance test 1(b) (8 kHz sampling rate)

In this test, 8 kHz resampled versions of the Supplement 23 files are used, on a file-by-file basis. The original and degraded files must be downsampled using the ITU-T Software Tools Library 2000 release 3, program `filter`, using the following command:

```
filter -down HQ2 inputfile.raw outputfile.raw
```

This assumes that the 16 kHz input speech file is called `inputfile.raw` and the 8 kHz output file is called `outputfile.raw`.

A batch script to assist with this, and the original and degraded file names, and the PESQ score given by the reference implementation, are provided in the files listed above.

An implementation passes this test when the absolute difference in PESQ score compared to the reference implementation is not greater than 0.05 in more than 2 file pairs (these may be any two of the file pairs), and not greater than 0.1 in all cases.

This conformance test is mandatory for all implementations of PESQ at 8 kHz sampling rate.

Supplement 23 to the P-series Recommendations can be obtained separately from the ITU.

A.3.2.2(a) Conformance test 2(a) (16 kHz sampling rate)

This test is based on data provided with PESQ and described in this Annex A. In this test, 16 kHz resampled versions of the Annex A/P.862 VoIP test files are used, on a file-by-file basis. The original and degraded files must be upsampled using the ITU-T Software Tools Library 2000 release 3, program `filter`, using the following command:

```
filter -up HQ2 inputfile.raw outputfile.raw
```

This assumes that the 8 kHz input speech file is called `inputfile.raw` and the 16 kHz output file is called `outputfile.raw`.

A batch script to assist with this, and the original and degraded file names, and the PESQ score given by the reference implementation, are provided in the files listed above.

An implementation passes this test when the absolute difference in PESQ score compared to the reference implementation is not greater than 0.05 in more than 1 file pair (this may be any one of the file pairs), and not greater than 0.5 in all cases.

This conformance test is mandatory for all implementations of PESQ at 16 kHz sampling rate.

A.3.2.2(b) Conformance test 2(b) (8 kHz sampling rate)

A composite database was constructed for Annex A/P.862 from 40 conditions (file pairs) from two subjective tests covering real and simulated VoIP connections that exhibit time-varying delay. Many of these file pairs also trigger the bad interval realignment process. This data is provided at 8 kHz sampling rate as the Annex A/P.862 VoIP test files, and these are used on a file-by-file basis.

The original and degraded file names, and the PESQ score given by the reference implementation, are provided in the files listed above.

An implementation passes this test when the absolute difference in PESQ score compared to the reference implementation is not greater than 0.05 in more than 1 file pair (this may be any one of the file pairs), and not greater than 0.5 in all cases.

This conformance test is mandatory for all implementations of PESQ at 8 kHz sampling rate.

A.3.2.3 Conformance test 3 (8 kHz or 16 kHz sampling rate) – Additional comparisons

To prevent implementers from specifically tailoring an algorithm to conform to requirements for the files described above, a further test is available. An implementation of PESQ that conforms to ITU-T Rec. P.862 must in at least 95% of cases give an output score that is within 0.05 of the PESQ score given by the ANSI-C reference implementation. These cases must be based on speech files covering a representative sample of reasonable telephone network conditions, and must lie within the scope of ITU-T Rec. P.862.

In practice it has been found that this is a much wider margin than required for most implementations. Users should expect that, in at least 99.5% of cases, an implementation should give an output score that is within 0.05 of the PESQ score given by the ANSI-C reference implementation. This should be considered to be a desirable level of accuracy, but it is not mandatory.

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