

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.862.2

(11/2007)

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

Methods for objective and subjective assessment of
quality

**Wideband extension to Recommendation P.862
for the assessment of wideband telephone
networks and speech codecs**

ITU-T Recommendation P.862.2



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
Transmission performance and QoS aspects of IP end-points	Series	P.1000

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

ITU-T Recommendation P.862.2

Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs

Summary

ITU-T Recommendation P.862.2 describes a simple extension to the perceptual evaluation of listening speech quality (PESQ) algorithm defined in ITU-T Recommendation P.862. It allows ITU-T Recommendation P.862 to be applied to the evaluation of conditions, such as speech codecs, where the listener uses wideband headphones. (In contrast, ITU-T Recommendation P.862 assumes a standard IRS-type narrow-band telephone handset which attenuates strongly below 300 Hz and above 3100 Hz.) This Recommendation is mainly intended for use with wideband audio systems (50-7000 Hz), although it may also be applied to systems with a narrower bandwidth.

Source

ITU-T Recommendation P.862.2 was approved on 13 November 2007 by ITU-T Study Group 12 (2005-2008) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). 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, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2008

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

CONTENTS

	Page
1 Scope	1
2 References.....	1
3 Definitions	2
4 Abbreviations and acronyms	2
5 Conventions	2
6 Description of wideband extension to ITU-T Rec. P.862	2
6.1 Input filter.....	2
6.2 Output mapping	3
7 ANSI-C reference implementation	3
8 Conformance	3

Introduction

This Recommendation describes a simple extension to the perceptual evaluation of listening speech quality (PESQ) algorithm defined in [ITU-T P.862]. It allows this algorithm to be applied to the evaluation of conditions, such as speech codecs, where the listener uses wideband headphones. (In contrast, [ITU-T P.862] assumes a standard IRS-type narrow-band telephone handset which attenuates strongly below 300 Hz and above 3100 Hz.) This Recommendation is mainly intended for use with wideband audio systems (50-7000 Hz), although it may also be applied to systems with a narrower bandwidth.

ITU-T Recommendation P.862.2

Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs

1 Scope

It is assumed that the reader is familiar with [ITU-T P.862].

The wideband extension to [ITU-T P.862] described in this Recommendation is subject to the limitations and applications that are described in the scope of [ITU-T P.862]. Further guidance on the limitations and applications of the wideband extension can be found in [ITU-T P.862.3].

Use of the wideband extension with systems that include noise suppression algorithms between the signal insertion point and signal capture point is not recommended. Additionally, clean speech samples should be employed because noisy speech samples, i.e., those with a poor signal-to-noise ratio, may lead to errors in prediction. The user should also be aware that the relative ranking of different distortion classes in wideband speech subjective experiments can vary slightly as a function of language. In particular, it should be noted that the wideband extension may overestimate MOS scores for ITU-T Rec. G.722 in experiments conducted in the Japanese and Korean languages.

When using the wideband extension to compare the performance of systems that may band-limit the audio signal, it is recommended that a wideband (50-7000 Hz audio bandwidth) version of the signal is used as the original reference signal for all measurements¹. Substantial bandwidth limitation by the system under test will be treated as a degradation and reduce the output score in the same way as other audible impairments. Such bandwidth limitation of the degraded signal may reduce prediction accuracy. Severe bandwidth limitation of the degraded signal, i.e., narrower than the traditional telephone bandwidth (300-3400 Hz) is not recommended.

It should be emphasized that the wideband extension predicts subjective opinion in the context of a subjective experiment that includes wideband speech conditions, i.e., signals with an audio bandwidth extending from 50 to 7000 Hz. This means that direct comparisons between scores produced by the wideband extension and scores produced by baseline [ITU-T P.862] or [ITU-T P.862.1] are not possible, due to the different experimental context.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T G.191] ITU-T Recommendation G.191 (2005), *Software tools for speech and audio coding standardization*.

[ITU-T P.341] ITU-T Recommendation P.341 (2005), *Transmission characteristics for wideband (150-7000 Hz) digital hands-free telephony terminals*.

¹ [ITU-T P.341] specifies a send filter mask for wideband speech systems. A filter implementation meeting this mask is included in the ITU-T Software Tool Library filter program [ITU-T G.191]. The pass-band of this filter extends from 50 Hz to 7 kHz.

- [ITU-T P.800] ITU-T Recommendation P.800 (1996), *Methods for subjective determination of transmission quality*.
- [ITU-T P.862] ITU-T Recommendation P.862 (2001), *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs plus Amendment 2* (2005), Revised Annex A – *Reference implementations and conformance testing for ITU-T Recs P.862, P.862.1 and P.862.2*.
- [ITU-T P.862.1] ITU-T Recommendation P.862.1 (2003), *Mapping function for transforming P.862 raw result scores to MOS-LQO*.
- [ITU-T P.862.3] ITU-T Recommendation P.862.3 (2007), *Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2*.

3 Definitions

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ACR	Absolute Category Rating
CCR	Comparison Category Rating
DCR	Degradation Category Rating
IRS	Intermediate Reference System
MOS	Mean Opinion Score

5 Conventions

This Recommendation is intended to provide an objective measure of quality that is comparable to ACR listening-only tests conducted according to [ITU-T P.800] using:

- a number of conditions with a wideband audio bandwidth (50-7000 Hz);
- listening quality opinion scale;
- naive listeners;
- quiet listening environment;
- binaural or monaural wideband headphone presentation with a frequency response that is either flat or equalized to be flat (as opposed to a telephone handset);
- speech material;
- an overall listening level of approximately 79 dB SPL.

The comparison of results produced by the wideband extension and subjective data using the DCR or CCR scales for wideband speech quality assessment is for further study.

6 Description of wideband extension to ITU-T Rec. P.862

6.1 Input filter

The input filter that is applied to both the reference and degraded files is replaced by an IIR filter. This is achieved in the function `pesq_measure()`, by changing the call to `apply_filter()` to a call to `IIRFilt()` with the appropriate filter definition, along with some pre-processing to reduce the effects of transients at the start or end of the file.

See the ANSI-C reference implementation for the filter coefficients and other implementation details. The new filter has a flat response above 100 Hz and a gentle roll-off below this point, modelling the attenuation of the headphones and ear at low frequencies.

Separate filter coefficients are supplied for use at 16 kHz and at 8 kHz sample rates, to ensure that both implementations have the same gain (within 0.1 dB) in the 10 Hz-4 kHz range.

6.2 Output mapping

The basic P.862 model provides raw scores in the range -0.5 to 4.5 . The wideband extension to [ITU-T P.862] includes a mapping function that allows linear comparisons with MOS values produced from subjective experiments that include wideband speech conditions with an audio bandwidth of 50-7000 Hz. This means that direct comparisons between scores produced by the wideband extension and scores produced by baseline [ITU-T P.862] or [ITU-T P.862.1] are not possible, due to the different experimental context. The output mapping function used in the wideband extension is defined as follows:

$$y = 0.999 + \frac{4.999 - 0.999}{1 + e^{-1.3669 \times x + 3.8224}} \quad (6-1)$$

where:

x is the raw model output

The mapping function was derived from data from a number of subjective experiments; some of these experiments contained only wideband speech conditions, others contained a mixture of narrow-band, wideband, and intermediate bandwidth speech.

For calculating the mapping function of the P.862.2 raw outcome to the MOS-LQO domain, a set of seven (7) provided databases were used. These databases were not only focused on pure wideband context but rather also contained databases from a so-called mixed content where various amounts of narrow-band conditions were presented along with wideband conditions as well. Out of the seven (7) databases five (5) were pure wideband data sets and two (2) contained narrow-band conditions as well scored on a so-called mixed scale.

It should be noted that no data for this use were derived in real-field measurements. The databases cover only simulated data. Note that the mapping function introduced here is only driven by use of databases containing in majority wideband conditions in a simulated context.

NOTE – The reference C code automatically includes this mapping when the wideband extension is selected.

7 ANSI-C reference implementation

The ANSI-C reference implementation of the wideband extension to [ITU-T P.862] is specified in Annex A of [ITU-T P.862].

8 Conformance

Implementations of the wideband extension to [ITU-T P.862] must meet the conformance criteria defined in Annex A of [ITU-T P.862].

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
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	Telecommunication management, including TMN and network maintenance
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, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems