International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.833.1

(07/2008)

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

Methods for objective and subjective assessment of quality

Methodology for the derivation of equipment impairment factors from subjective listening-only tests for wideband speech codecs

Recommendation ITU-T P.833.1

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
Transmission performance and QoS aspects of IP end-points	Series	P.1000
Communications between cars	Series	P.1100

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

Recommendation ITU-T P.833.1

Methodology for the derivation of equipment impairment factors from subjective listening-only tests for wideband speech codecs

Summary

Recommendation ITU-T P.833.1 describes an extension of the methodology for deriving equipment impairment factors from subjective listening-only tests. It is intended that it primarily be applied to determining wideband equipment impairment factors *Ie,wb*, capturing the degradation introduced by wideband speech codecs. The resulting wideband equipment impairment factors derived by this methodology are intended to be used on the extended transmission rating scale underlying the E-model, see Appendix II in Amendment 1 to Recommendation ITU-T G.107. They will reflect the auditory impairments of the corresponding equipment in a listening-only mode.

Source

Recommendation ITU-T P.833.1 was approved on 14 July 2008 by ITU-T Study Group 12 (2005-2008) under Recommendation ITU-T A.8 procedure.

Keywords

Codec, E-model, impairment factor method, wideband speech transmission.

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 not 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/.

CONTENTS

Introduction

ITU-T has introduced the impairment factor method in order to describe the joint effect of different types of degradations on overall transmission quality. According to this method, transmission impairments can be transformed into so-called "psychological factors", which are assumed to be additive on a psychological scale. The transmission rating scale (*R*-scale) underlying the E-model (see Recommendation ITU-T G.107) is assumed to reflect such a psychological scale.

The impairment factor method allocates a value of impairment to each parameter describing the transmission channel, and then allows the simple addition of these impairments to determine the overall impairment. The overall transmission rating *R* is finally calculated by subtracting the overall impairment from the maximum value R_{max}. Whereas the *R*-scale and the E-model have originally been formulated for narrow-band (300-3400 Hz) speech transmission only, Appendix II of Recommendation ITU-T G.107 presents an extension of this scale to wideband (50-7000 Hz) speech transmission. Although a complete wideband version of the E-model is not yet available, Appendix IV of Recommendation ITU-T G.113 provides values for wideband equipment impairment factors *Ie,wb*, which describe the degradation resulting from selected wideband speech codecs on the extended *R*-scale. This Recommendation describes a methodology for how *Ie,wb* values can de derived for other wideband speech codecs, from the results of listening-only tests.

Recommendation ITU-T P.833.1

Methodology for the derivation of equipment impairment factors from subjective listening-only tests for wideband speech codecs

1 Scope

It is assumed that users of this Recommendation are familiar with [ITU-T P.833].

The extension of the methodology described in this Recommendation is intended to derive wideband equipment impairment factors *Ie,wb* quantifying the impairment introduced by wideband (50-7000 Hz) speech codecs, including or not the effects of transmission errors, such as random bit errors, random packet loss or discard, or bursty packet loss or discard. Such wideband equipment impairment factors have been introduced as a simplified measure of the degradation introduced by wideband speech codecs on the integral transmission quality from mouth to ear. They are in no way an exact description of the effects related to each individual codec or codec tandem, which may be very diverse in their perceptual nature. Instead, they represent the relative degradation in comparison to other impairments occurring in a connection.

In order to provide guidance on the quantitative amount of impairment introduced by such codecs, a framework of wideband equipment impairment factor values for several wideband speech codecs has been derived by ITU-T, see Appendix IV of [ITU-T G.113]. The derivation of these values was based on many subjective listening-only tests carried out in different test laboratories, so as to guarantee stable values for all the codecs under investigation so far, as well as their relations to each other. If new equipment impairment factor values for different codecs have to be derived, then the overall consistency with the established framework is of primary importance if results are to be obtained that are valid for network planning. The methodology described in this Recommendation was designed to fulfil this requirement.

So far, wideband speech codecs are the only wideband elements for which impairment factors have been derived. For asynchronous tandems of multiple codecs of the same type, or of multiple codecs of different types, it is assumed that individual equipment impairment factors are additive. The overall equipment impairment factor for the chain of codecs is then calculated as a simple sum of all the individual *Ie,wb*s. Experimental test data collected provide evidence that this simple additivity is not satisfied for all the potential combinations of codecs [b-Möller]. In some cases, order effects exist, i.e., the tandem of codec A followed by codec B results in a different degradation than codec B followed by codec A. These deviations from the pure additivity property are an item for further study. Impairment factor values for other wideband network elements, resulting in other types of degradations and to be used in a wideband version of the E-model, have not yet been defined and are for further study as well.

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.107] Recommendation ITU-T G.107 (2008), *The E-model, a computational model for use in transmission planning.*

[ITU-T G.113] Recommendation ITU-T G.113 (2007), *Transmission impairments due to speech processing.*

[ITU-T G.722] Recommendation ITU-T G.722 (1988), *7 kHz audio-coding within 64 kbit/s.*

[ITU-T G.722.1] Recommendation ITU-T G.722.1 (2005), *Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss.*

[ITU-T G.722.2] Recommendation ITU-T G.722.2 (2003), *Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB).*

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

[ITU-T P.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality.*

[ITU-T P.830] Recommendation ITU-T P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs.*

[ITU-T P.833] Recommendation ITU-T P.833 (2001), *Methodology for derivation of equipment impairment factors from subjective listening-only tests.*

3 Definitions

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ADPCM Adaptive Differential Pulse Code Modulation

CELP Code Excited Linear Prediction

MLTC Modified Lapped Transform Coding

MOS Mean Opinion Score

5 Conventions

None.

6 Selection of experiment parameters

The listening-only test from which *Ie,wb* values are to be derived should fulfil the general requirements for listening-only tests given in [ITU-T P.800] and [ITU-T P.830], as well as the requirements set in [ITU-T P.833]. However, in contrast to [ITU-T P.833], the test should be carried out in a wideband or mixed-band mode, implying several changes given hereafter.

In contrast to narrow-band, there is no standard wideband terminal which would be representative for most wideband connections. For the given purpose, the test should be carried out either with diotic or monotic headphone presentation. The frequency response of the terminal should represent a wideband characteristic in the range 50 to 7000 Hz, as it is defined in [ITU-T P.341].

For reasons given in [ITU-T P.833], subjective tests aiming at deriving equipment impairment factors should include a number of reference conditions. These references are necessary for anchoring impairment factor values, and they will guarantee that new equipment impairment factor values fit into the existing system given in Appendix IV of [ITU-T G.113].

It is recommended that different types of wideband speech codecs be used as a reference for experiments deriving wideband equipment impairment factors (i.e., waveform codec as G.722 or hybrid codec as G.722.2). The exact reference conditions to be used vary depending on whether transmission errors are to be considered, and whether an additivity check has to be performed.

In addition, a narrow-band codec may be used in case of mixed-band test.

6.1 Reference conditions for wideband speech codecs without transmission errors

When wideband equipment impairment factors for codecs disregarding transmission errors are determined, the set of 12 reference codec conditions given in Table 1 should be included in the subjective test conditions. This list has been chosen from the codecs for which values are already defined in Appendix IV to [ITU-T G.113], and they have been selected to cover the whole range of *Ie,wb* values.

Table 1 – Reference conditions for wideband speech codecs without transmission errors

No.	Abbreviation	Codec type	Reference	Operating rate (kbit/s)	*Ie,wb* value
1	Clean	Linear PCM, 16 bit	--	--	0
2	G.722.2@23.05	CELP	ITU-T G.722.2	23.05	1
3	G.722.2@19.85	CELP	ITU-T G.722.2	19.85	3
4	G.722.2@15.85	CELP	ITU-T G.722.2	15.85	7
5	G.722.2@14.25	CELP	ITU-T G.722.2	14.25	10
6	G.722@64	ADPCM	ITU-T G.722	64	13
7	G.722.1@32	MLTC	ITU-T G.722.1	32	13
8	G.722.1@24	MLTC	ITU-T G.722.1	24	19
9	G.722@56	ADPCM	ITU-T G.722	56	20
10	G.722.2@8.85	CELP	ITU-T G.722.2	8.85	26
11	G.722@48	ADPCM	ITU-T G.722	48	31
12	G.722.2@6.6	CELP	ITU-T G.722.2	6.6	41

It is important to check the additivity of the newly derived wideband equipment impairment factor in the framework of other equipment impairment factor values defined so far. If such an additivity check is not performed, the property of a simple summation of equipment impairment factors in order to cater for codec tandems should not be regarded as valid. Table 2 gives a minimum number of additional reference conditions which should, in any case, be included in the test set to allow for a rough additivity check. It is preferable, however, to test inter-codec tandem operations with a larger set of similar conditions, including triple tandems in different codec orders.

Table 2 – Reference conditions for the additivity check in tandem operation of wideband speech codecs without transmission errors

No.	Tandem operation	Reference codec type	Operating rate (kbit/s)	*Ie,wb* value
13	G.722.2@19.85*(new codec)	CELP	19.85	3 + *Ie,wb*(new codec)
14	G.722.2@14.25*(new codec)	CELP	14.25	10 + *Ie,wb*(new codec)
15	G.722@64*(new codec)	ADPCM	64	13 + *Ie,wb*(new codec)
16	G.722.1@32*(new codec)	MLTC	32	13 + *Ie,wb*(new codec)
17	G.722.1@24*(new codec)	MLTC	24	19 + *Ie,wb*(new codec)
18	G.722@48*(new codec)	ADPCM	48	31 + *Ie,wb*(new codec)
19	(new codec)*G.722.2@19.85	CELP	19.85	*Ie,wb*(new codec) + 3
20	(new codec)*G.722.2@14.25	CELP	14.25	*Ie,wb*(new codec) + 10
21	(new codec)*G.722@64	ADPCM	64	*Ie,wb*(new codec) + 13
22	(new codec)*G.722.1@32	MLTC	32	*Ie,wb*(new codec) + 13
23	(new codec)*G.722.1@24	MLTC	24	*Ie,wb*(new codec) + 19
24	(new codec)*G.722@48	ADPCM	48	*Ie,wb*(new codec) + 31
NOTE – A*B designates asynchronous tandeming of codecs A and B, B followed by A.				

6.2 Reference conditions for wideband speech codecs with transmission errors

When wideband equipment impairment factors for codecs under the effects of transmission errors are determined, the same reference conditions as given in Table 1 should be applied. In addition to these conditions, supplementary conditions may be included in the test from the codecs listed in Appendix IV of [ITU-T G.113], including perceptively noticeable degradation due to transmission errors (random bit errors, random packet loss, bursty packet loss, or propagation errors in terms of error patterns). However, no *Ie,wb* values have yet been agreed upon for these codecs, so that these conditions cannot be considered as references. Preliminary values can be found, e.g., in [b-Möller], and the *Ie,wb* values obtained for these codecs may be compared to these values. Larger deviations from the values will have to be considered when interpreting the newly-derived *Ie,wb* values for codecs including transmission errors.

7 Test method

The test method generally follows the recommendations given in [ITU-T P.833]. Due to the lack of available references for codecs under transmission errors, the number of mandatory reference conditions is slightly different. Table 3 summarizes the test conditions to be included in the different parts of the experiment.

Table 3 – Overview of test conditions for the different parts of the experiment

Part	Purpose	Test conditions	Mandatory/ optional	Min. overall Σ test cond.
A	Determination of *Ie,wb* for the new codec in error-free conditions	References 1-12	Mandatory	13
		New codec in single operation	Mandatory	
		New codec in single operation, at 2 alternative speech input levels	Optional	
		Additional wideband codec references	Optional	
B	Additivity check	References 13-24	Mandatory	14
		New codec alone in double and triple tandem operation	Mandatory	
		New codec in double and triple tandem operation with other wideband codecs	Optional	
C	Determination of *Ie,wb* for the new codec in transmission error conditions	New codec in single operation in different transmission error conditions (*m* conditions)	Mandatory	*m*
		Additional references according to 4.2.2 of [ITU-T P.833]	Optional	

8 Derivation of wideband equipment impairment factors

The methodology for deriving wideband equipment impairment factors follows mostly [ITU-T P.833], however with some modifications of the calculations. It consists of three to five steps, depending on whether transmission errors are taken into account or not:

Step 1: Scale transformation of the subjective test data.

Step 2: Derivation of a stable *Ie,wb* value for the codec under test, in single codec operation without transmission errors, via a linear interpolation of the test results.

Step 3: Additivity check.

If transmission errors are under investigation, the following additional steps are to be taken:

Step 4: Derivation of stable *Ie,wb* values for different transmission error conditions in single codec operation.

Step 5: Additivity check.

These steps are described below.

8.1 Scale transformation (step 1)

Mean opinion scores (MOSs) are determined for all test conditions, including the references listed in Table 3. These MOS results have to be transformed to the extended *R*-scale. For this purpose, they are first transformed to the non-extended, narrow-band R_{NB}-scale (range [0;100], subscript *NB*), using either a numeric or graphic inversion of the relationship between MOS and R_{NB} given in the E-model:

$$\begin{aligned} &\text{for } MOS = 1.0: && R_{NB} = 0 \\ &\text{for } 1.0 < MOS < 4.5: && MOS = 1 + 0.035 \cdot R_{NB} + R_{NB} \cdot (R_{NB} - 60) \cdot (100 - R_{NB}) \cdot 7 \cdot 10^{-6} \\ &\text{for } MOS \geq 4.5: && R_{NB} = 100 \end{aligned} \quad (1)$$

In case that the maximum MOS_{max} value of the test is higher than 4.5, the range of MOS values of that test first has to be linearly transformed to a maximum value of 4.5, so that equation 1 can be applied:

$$MOS_{norm,i} = \frac{MOS_i - 1}{MOS_{max} - 1} \cdot 3.5 + 1 \qquad (2)$$

These values have to be transformed to the extended *R*-scale (range [0;129]). For this purpose, a linear expansion of the obtained R_{NB} values to the wideband R_{WB} values (subscript *WB*) is carried out:

$$R_{WB} = 1.29 \cdot R_{NB} \qquad (3)$$

From the resulting values for R_{WB}, the corresponding raw *Ie,wb,sub* values (subscript *sub* for "subjective test") can be calculated by defining the R_{WB}-value for reference condition no. 1 (see Table 1) as an anchor, thus:

$$Ie,wb,sub = R_{WB}(\text{condition No.1}) - R_{WB}(\text{test condition}) \qquad (4)$$

This equation results in *Ie,wb,sub* for the reference condition no. 1 always set to 0. The outcome of step 1 is an *Ie,wb,sub* value for each test condition. It reflects the specific test condition, and it is not necessarily consistent with wideband equipment impairment factors defined in Appendix IV to [ITU-T G.113].

8.2 Linear interpolation of the test results (step 2)

For all 12 reference conditions of Table 1, as well as possibly for all supplementary reference conditions involving only codecs for which *Ie,wb* values have already been defined, pairs of expected equipment impairment factors *Ie,wb,exp* and observed values *Ie,wb,sub* are now available. These pairs can be represented as a scatter plot. A linear interpolation using a straight line

$$Ie,wb,sub = a \cdot Ie,wb,exp + b \qquad (5)$$

can now be made. The coefficients *a* and *b* are determined numerically, approximating all the reference pairs in a least-squares sense. From this approximation, a stable equipment impairment factor value for the codec under test (*Ie,wb* = *Ie,wb,exp*) can be derived.

In rare cases, the linear transformation may result in a negative *Ie,wb* value for the codec under investigation. This might occur if the related subjective ratings are close to or better than the one for reference condition no. 1 of Table 1. In this case, *Ie,wb* should be set to zero instead.

8.3 Additivity check (step 3)

The equipment impairment factor derived in step 2 does not necessarily satisfy the additivity property of *Ie,wb*s. This has to be checked for both tandems of the new codec alone and mixed tandems with codecs for which *Ie,wb* values have already been defined. The procedure is the same as described in [ITU-T P.833]. If more than 4 out of 14 tandem conditions (2 pure tandems of the codec under investigation and 12 reference tandem conditions nos 13-24, see Table 2) show major deviations from the interpolation line, the additivity property should not be regarded as having been satisfied. In this case, the wideband equipment impairment factor derived from the experiment will not properly represent the degradations occurring in tandem operations of the new codec.

8.4 Derivation of *Ie,wb* values for transmission error conditions (step 4)

Using the interpolation line and the procedure of step 2, *Ie,wb* values can be derived for all conditions of the new codec including transmission errors, as well as for all conditions of the codecs given in Appendix IV of [ITU-T G.113] which include transmission errors. Because no reference *Ie,wb* values are defined for the latter codecs, it cannot be checked whether the inclusion of transmission errors may have an effect on the judgments, and consequently also on the interpolation line. However, the obtained values may be compared to the ones given in [b-Möller]. Any large deviations may indicate that the rating behaviour changed for the error conditions in comparison to the error-free case; the resulting *Ie,wb* values for codecs under transmission errors should then be considered with care.

Graphs or tables of *Ie,wb* values for transmission error conditions can be used for a plausibility check. The minimum consistency to be reached is to have non-decreasing wideband equipment impairment factor values for increasing transmission error rates.

8.5 Additivity check (step 5)

A similar additivity check as in step 3 should be carried out on the wideband equipment impairment factor values derived in step 4, using all the available tandem conditions of the codec under investigation with transmission errors introduced and other codecs for which impairment factors have already been derived.

9 Application of the derived wideband equipment impairment factor values

Due to the inherent variability of subjective ratings, *Ie,wb* values derived using this methodology cannot be expected to represent very exact quantitative measures of impairment that a subject in a specific situation would experience. Instead, they represent the additional degradation expected to be introduced by the respective wideband speech codec, expressed on the extended *R*-scale. A value of *Ie,wb* = 0 signifies that the corresponding codec introduces no additional impairment compared to the clean (16 kHz sampling frequency, 16-bit quantization, linear PCM) channel. Unless additivity has been proven in steps 3 and 5, it cannot be concluded from the results how large the impairment would be when the codec is combined with other codecs.

In addition, it is unclear how the codec degradation combines with other degradations which are currently covered by the (narrow-band) E-model, like ambient noise, circuit noise, non-optimal loudness, sidetone, talker or listener echo, or delay. These items may be covered by future wideband extensions of the E-model and are under study in ITU-T SG 12. Unless such a model is available, *Ie,wb* values cannot be combined with other types of impairment factors in order to obtain an estimate of the integral wideband speech quality mouth-to-ear.

Bibliography

[b-Möller] Möller, S., Raake, A., Kitawaki, N., Takahashi, A., Wältermann, M., (2006), *Impairment Factor Framework for Wide-band Speech Codecs*, IEEE Transactions on Audio, Speech, and Language Processing, Vol. 14, No. 6, pp. 1969-1976.

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

Printed in Switzerland
Geneva, 2009