I n t e r n a t i o n a l   T e l e c o m m u n i c a t i o n   U n i o n

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.834.1

(06/2015)

SERIES P: TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Methods for objective and subjective assessment of speech quality

---

Extension of the methodology for the derivation of equipment impairment factors from instrumental models for wideband speech codecs

Recommendation ITU-T P.834.1

ITU-T

ITU-T P-SERIES RECOMMENDATIONS

TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Voice terminal characteristics	Series	P.30 P.300
Reference systems	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 speech 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 involving vehicles	Series	P.1100
Models and tools for quality assessment of streamed media	Series	P.1200
Telemeeting assessment	Series	P.1300
Statistical analysis, evaluation and reporting guidelines of quality measurements	Series	P.1400
Methods for objective and subjective assessment of quality of services other than voice services	Series	P.1500

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

Recommendation ITU-T P.834.1

Extension of the methodology for the derivation of equipment impairment factors from instrumental models for wideband speech codecs

Summary

Recommendation ITU-T P.834.1 describes an extension of the methodology for deriving equipment impairment factors from instrumental models of Recommendation ITU-T P.834. It is intended that it be primarily applied to determine 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 of Recommendation ITU-T G.107). They reflect the auditory impairments of the corresponding equipment in a listening-only mode.

The methodology of Recommendation ITU-T P.834.1 makes use of instrumental models (so called "objective methods"), e.g., the model defined in Recommendation ITU-T P.862.2. It is to be considered as supplementary to the methodology based on auditory listening-only tests, described in Recommendation ITU-T P.833.1. It provides valid *Ie,wb* values only for those codecs for which the instrumental model used produces meaningful estimations.

Speech material associated with Recommendation ITU-T P.834.1 is freely available on the test signal database at www.itu.int/net/ITU-T/sigdb.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.834.1	2009-04-29	12	11.1002/1000/9729
2.0	ITU-T P.834.1	2015-06-29	12	11.1002/1000/12517

Keywords

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

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

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/.

Table of Contents

Introduction

ITU-T has introduced the impairment factor method in order to describe the joint effect of different types of degradations on the 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 [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 *Rmax*. Whereas the *R* scale and the E-model were originally formulated for narrowband (300-3400 Hz) speech transmission only, Appendix II of [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 to [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 the derivation of *Ie,wb* values for other wideband speech codecs, based on the results of so-called "objective methods".

Recommendation ITU-T P.834.1

Extension of the methodology for the derivation of equipment impairment factors from instrumental models for wideband speech codecs[1]

1 Scope

This Recommendation describes an extension of the ITU-T P.834 methodology and is intended to derive wideband equipment impairment factors *Ie,wb* quantifying the impairment introduced by wideband (50-7000 Hz) speech codecs. It is assumed that the reader is familiar with [ITU-T P.834]. 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 to [ITU-T G.113]. The derivation of these values is 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 relationship to each other. If new equipment impairment factor values for different codecs require derivation, 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 has been 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. 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 also for further study.

ITU-T currently recommends two methodologies for the derivation of wideband equipment impairment factors: this Recommendation and [ITU-T P.833.1]. The methodology described in [ITU-T P.833.1] is based on the results of auditory listening-only tests. The approach in this Recommendation, which is based on instrumental models (so-called "objective methods"), requires that the models used provide valid estimations of auditory judgements. Only in this case will the derived wideband equipment impairment factors be valid. The methodology described has been tested with the model defined in [ITU-T P.862.2] and results have been found to be satisfactory in most cases.

---

[1] This Recommendation includes speech material which is freely available on the test signal database at www.itu.int/net/ITU-T/sigdb. The speech material is also available in a zipped file associated with this Recommendation.

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 (2015), *The E-model: A computational model for use in transmission planning*.

[ITU-T G.107.1] Recommendation ITU-T G.107.1 (2015), *Wideband E-model*.

[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 (2012), *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.56] Recommendation ITU-T P.56 (2011), *Objective measurement of active speech level*.

[ITU-T P.341] Recommendation ITU-T P.341 (2011), *Transmission characteristics for wideband digital loudspeaking and hands-free telephony terminals*.

[ITU-T P.501] Recommendation ITU-T P.501 (2012), *Test signals for use in telephonometry*.

[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.833.1] Recommendation ITU-T P.833.1 (2009), *Methodology for the derivation of equipment impairment factors from subjective listening-only tests for wideband speech codecs*.

[ITU-T P.834] Recommendation ITU-T P.834 (2015), *Methodology for the derivation of equipment impairment factors from instrumental models*.

[ITU-T P.862.2] Recommendation ITU-T P.862.2 (2007), *Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs*.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ADPCM Adaptive Differential Pulse Code Modulation

CELP Codebook-Excited Linear Prediction

MLTC Modified Lapped Transform Coding

MOS Mean Opinion Score

5 Conventions

None.

6 Experimental set-up and data processing

The instrumental methodology in this Recommendation is mainly based on the auditory equivalent specified in [ITU-T P.833.1], substituting the auditory test by an instrumental model, e.g., that described in [ITU-T P.862.2]. Apart from this substitution, both methodologies are very similar.

Input to the methodology is the speech files given in Annex A (see clause 6.1). Alternatively, other wideband-recorded speech files can be used, e.g., addressing languages not covered by the set provided in Annex A, provided that they have been recorded according to the recommendations given in [ITU-T P.800]. The speech files of Annex A are available in an original, uncoded version, as well as processed through standard implementations of 12 reference codecs defined in clause 6.2. For the new codec under investigation, the original source material has also to be processed, in the way described in clause 6.3. Processing is necessary for the new codec alone, in double and triple tandem configuration with itself, as well as in mixed tandem configurations with other codecs defined in Table 2.

Input and output files of each codec or codec tandem are then used as an input to the instrumental model. The instrumental model will provide an estimated mean opinion score (MOS) value for each speech file pair (see clause 6.4). These MOS estimations are then used to derive an *Ie,wb* for the codec under test, as described in clause 7.

6.1 Input speech material

The material that can be found at the URL provided in Annex A contains uncoded source files taken from [ITU-T P.501], as well as coded output files processed with the 12 reference codecs of Table 1. All speech material was recorded in 16-bit linear PCM (binary) files with a PC format (i.e., low byte first). The filename convention is listed in Table A.2. Users of this Recommendation may process the files themselves, starting from the source material, provided that they respect the same processing steps described in clause 6.3.

The material for the 12 reference conditions has already been processed in the way that is required for the ITU-T P.834.1 methodology, see the files at the URL given in Annex A. For the codec under investigation – alone, in double and triple tandems with itself, and in mixed tandems as indicated in Table 2 – the processing has to be carried out on the pre-processed speech files (extension `.axp`). It is recommended that all speech files available for this purpose be used and then that the mean be calculated (see clause 7, step 3). This processing procedure is described in clause 6.3.

6.2 Reference conditions

Two sets of reference conditions are to be used in steps 1 and 4 of the methodology. For step 1, the set of 12 reference codec conditions given in Table 1 should be processed by the instrumental model. This set has been chosen from the codecs for which values are already defined in Appendix IV to [ITU-T G.113] and have been selected to cover the whole range of *Ie,wb* values.

Table 1 – Reference conditions for step 1

No.	Abbreviation	Codec type	Reference	Operating rate (kbit/s)	*Ie,wb,def* value
1	Clean	linear PCM, 16 bits	–	–	0
2	G.722.2@23.05	codebook-excited linear prediction (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	adaptive differential pulse code modulation (ADPCM)	[ITU-T G.722]	64	13
7	G.722.1@32	modified lapped transform coding (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 equipment impairment factor in the framework of other equipment impairment factor values previously defined. This is done in step 4 of the methodology. 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 10 additional reference conditions (Nos. 13-22) that should, in any case, be processed by the instrumental model 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. In addition to the mixed codec tandems of Table 2, the new codec under investigation should be checked in double and triple tandem operation with itself.

Table 2 – Reference conditions for the additivity check in step 4

No.	Tandem operation	Reference codec type	Reference codec operating rate (kbit/s)	*Ie,wb,def* 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)

Table 2 – Reference conditions for the additivity check in step 4

No.	Tandem operation	Reference codec type	Reference codec operating rate (kbit/s)	*Ie,wb,def* value
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, A followed by B.				

6.3 Processing of speech material

In addition to the processed speech material made available here, the source speech material requires processing:

– through the codec under investigation alone;

– through the codec under investigation in double and triple tandem with itself;

– through mixed tandems of the codec under investigation with the reference codecs, as given in Table 2.

This processing of the input speech data has to follow the procedure, which has been used for the provision of the reference speech material. The procedure consists of four steps.

1) The source speech file is filtered to the standard wideband bandwidth (50-7000 Hz). For this purpose the filter according to [ITU-T P.341] should be used.

2) The filtered signal is level aligned to an active speech level of −26 dB below the overload point of the digital system, using the procedure defined in [ITU-T P.56].

   These two first steps result in the pre-processed speech files (`.axp`).

3) The pre-processed speech file is then processed through the codec under test.

4) In tandem operation, the resulting processed speech file from step 3 is processed through the second (or any following) speech codec.

6.4 Calculation of MOS estimations using the instrumental model

Instrumental models which are usable for the derivation of equipment impairment factors as described here perform a comparison between an undistorted source signal and a processed and probably distorted signal (the degradation of which is quantified). The instrumental model requires both signals, the source signal and the distorted signal, as input in each case, as well as the corresponding sampling frequency. Some models also require information about the pre-processing status of the reference signal (e.g., pre-filtered or not). For the data material made available in this Recommendation, the unfiltered sources (extension `.src`) as well as the pre-filtered speech files (extension `.axp`) are provided. It is recommended that the pre-processed sources be used as reference signals for the instrumental models. The pre-filtered speech data should be used as input signals for processing of codec conditions (see clause 6.3).

The instrumental model predicts the degradation due to the coding process in terms of an MOS, reflecting the perceived listening quality as it would have been determined in a listening-only test carried out according to [ITU-T P.800] and [ITU-T P.830]. Because the *Ie,wb* values are expected to describe the effect of wideband speech codecs, the instrumental model should provide valid predictions for wideband transmission scenarios. Under this provision, the procedure described can in principle be used with all instrumental models that provide an estimated value for listening quality on the MOS scale from 1 to 5 reflecting a wideband, mixed-band, super-wideband or fullband test scenario, e.g., the model according to [ITU-T P.862.2].

These estimated MOS values are the input needed for the four-step procedure described in clause 7.

7 Derivation of wideband equipment impairment factors

Based on the instrumental MOS estimations described in clause 6, a wideband equipment impairment factor for a new wideband speech codec can be derived using the four-step procedure described in the following.

Step 1: Determination of raw K values for the reference conditions

Input and output speech material of the 12 reference conditions of Table 1 is processed by the instrumental model, producing an MOS estimate for each speech file pair (see clause 6). Mean MOS values are then calculated over all speech files processed with one specific condition as described in Table 1, for each of the 12 reference circuits. The mean MOS estimations are first transformed to the non-extended, narrowband R_{NB} scale (range [0;100], subscript *NB*), using the relationship between MOS and R_{NB} given in the E-model:

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

These R_{NB} values still reflect the narrowband use of the MOS scale assumed by the E-model, and not a wideband use scenario, which would be encountered in a subjective test with wideband connections. In order to reflect the superior quality of wideband compared to narrowband transmission, the R_{NB} values have to be transformed to R_{WB} values (range [0;129]), using the following relationship taken from [b-Möller]:

$$R_{WB} = 1.29 \cdot R_{NB} \qquad (2)$$

Taking the R_{WB} value for the "clean" condition (No. 1 of Table 1) as the reference for $K = Ie,wb = 0$, raw estimates for wideband equipment impairment factors (*K*) can be calculated with:

$$K = R_{WB}(\text{clean}) - R_{WB}(\text{test condition}) \qquad (3)$$

Step 2: Calculation of interpolation line parameters

From the raw *K* values and the respective defined *Ie,wb,def* values (see Appendix I to [ITU-T G.113]) of the 12 reference conditions of Table 1, an interpolation line is calculated in the *Ie,wb* space, using a straight line in the scatter plot of [*K; Ie,wb,def*] pairs:

$$K = a \cdot Ie,wb,def + b \qquad (4)$$

The coefficients *a* and *b* are determined numerically, approximating all the reference *K–Ie,wb,def* pairs in a least-squares sense. Alternatively, but with less precision, the approximation can also be made graphically on the scatter plot.

Step 3: Determination of a stable Ie,wb value for the codec under test

Starting from this interpolation line, an instrumentally derived *Ie,wb* value for the new codec can be determined using the speech material from Annex A that has been processed through the codec under investigation. Input and output speech material are evaluated by means of the instrumental model, producing MOS estimations for each speech file pair (see clause 6). A mean estimated MOS value is then calculated for the new codec. The mean estimated MOS is transformed into a *K* value, using formulae (1) to (3). The *K* value is then transformed into a raw estimate of *Ie,wb* using the parameters *a* and *b* of the interpolation line:

$$Ie,wb = \frac{K - b}{a} \tag{5}$$

Equation (5) or its graphical representation leads to a wideband equipment impairment factor value for the codec under test, which can be regarded as stable. This *Ie* value will normally fit into the framework of wideband equipment impairment factors the interpolation line has been derived from, namely for the codecs included in reference conditions Nos 1 to 12. However, it does not necessarily satisfy the additivity property underlying the impairment factor principle.

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 instrumental MOS estimates are close to or better than the one for the "clean" condition (No. 1 of Table 1). In this case, *Ie,wb* should be set to zero instead.

Step 4: Additivity check

Additivity of the newly derived *Ie,wb* value 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 in Appendix IV to [ITU-T G.113]. For all the tandems of the new codec alone, as well as for all the reference conditions of Table 2, pairs of observed *K* values and expected *Ie,wb,def* values are made available, using the instrumental model. *Ie,wb,def* for tandems of the new codec alone and for reference conditions of Table 2 are computed using the *Ie,wb* value derived in step 3. These pairs can be represented in the same scatter plot of step 2. All major deviations from the interpolation line should be noted and investigated, as they may question the applicability of the additivity property of impairment factors.

If more than three out of 14 tandem conditions (two pure tandems of the codec under investigation and 12 reference tandem conditions Nos 13-24, see Table 2) show major deviations from the interpolation lines, the additivity property should not be regarded as being satisfied. In this case, the equipment impairment factor derived from the experiment does not properly represent the degradations occurring in tandem operations of the new codec.

8 Limitations of the instrumentally derived wideband equipment impairment factor values

The methodology described in this Recommendation, which is based on instrumental models, requires that the models used provide valid estimations of auditory judgements. When the validity of the model has been verified for transmission error conditions, such as frame erasures or packet loss, the methodology using the model is applicable to such conditions. The models defined in [ITU-T P.863] and [ITU-T P.862.2] have been tested for transmission error conditions, and results were found to be satisfactory.

If *Ie,wb* is derived directly by using the instrumental method recommended in this Recommendation, it already reflects the effect of packet loss introduced in the preparation of speech materials under test. Therefore, the *Ie,wb* value derived as *Ie-effWB* in [ITU-T G.107.1] should be used, avoiding double counting of the packet-loss effect by calculating *Ie-effWB* using the regular formula from equation (7-20) in [ITU-T G.107.1]. However, if the values for *Ie,wb* and *Bpl* given in [ITU-T G.113] and values for *Ppl* are available, then the standard formulae of the wideband E-model in [ITU-T G.107.1] should be used instead of determining a new *Ie-effWB* value using this Recommendation.

Annex A

Speech material

(This annex forms an integral part of this Recommendation.)

The speech material consists of the files that can be found in the test signal database at http://www.itu.int/net/ITU-T/sigdb/genaudio/Pseries.htm#P.834.1. Table A.1 summarizes the contents of the downloadable file, and Table A.2 describes the file name convention for each language sub-directory.

Table A.1 – Directory structure of the speech data

Directory			Language	# of Files	Total Mbytes
\data	\original	\fr	French	4	1.3
		\jp	Japanese	4	1.0
		\ae	American English	4	1.0
	\pre_proc	\fr	French	4	1.3
		\jp	Japanese	4	1.0
		\ae	American English	4	1.0
	\coded	\fr	French	48	15.6
		\jp	Japanese	48	11.8
		\ae	American English	48	11.9

Table A.2 – Correspondence list for individual speech files

	\fr, \jp and \ae, respectively		
	Source	**Pre-processed**	**Processed**
Clean	f1.src	f1.axp	f1.c01
	f2.src	f2.axp	f2.c01
	m1.src	m1.axp	m1.c01
	m2.src	m2.axp	m2.c01
G.722.2@23.05	f1.src	f1.axp	f1.c02
	f2.src	f2.axp	f2.c02
	m1.src	m1.axp	m1.c02
	m2.src	m2.axp	m2.c02
G.722.2@19.85	f1.src	f1.axp	f1.c03
	f2.src	f2.axp	f2.c03
	m1.src	m1.axp	m1.c03
	m2.src	m2.axp	m2.c03

Table A.2 – Correspondence list for individual speech files

	\fr, \jp and \ae, respectively		
	Source	**Pre-processed**	**Processed**
G.722.2@15.85	f1.src	f1.axp	f1.c04
	f2.src	f2.axp	f2.c04
	m1.src	m1.axp	m1.c04
	m2.src	m2.axp	m2.c04
G.722.2@14.25	f1.src	f1.axp	f1.c05
	f2.src	f2.axp	f2.c05
	m1.src	m1.axp	m1.c05
	m2.src	m2.axp	m2.c05
G.722@64	f1.src	f1.axp	f1.c06
	f2.src	f2.axp	f2.c06
	m1.src	m1.axp	m1.c06
	m2.src	m2.axp	m2.c06
G.722.1@32	f1.src	f1.axp	f1.c07
	f2.src	f2.axp	f2.c07
	m1.src	m1.axp	m1.c07
	m2.src	m2.axp	m2.c07
G.722.1@24	f1.src	f1.axp	f1.c08
	f2.src	f2.axp	f2.c08
	m1.src	m1.axp	m1.c08
	m2.src	m2.axp	m2.c08
G.722@56	f1.src	f1.axp	f1.c09
	f2.src	f2.axp	f2.c09
	m1.src	m1.axp	m1.c09
	m2.src	m2.axp	m2.c09
G.722.2@8.85	f1.src	f1.axp	f1.c10
	f2.src	f2.axp	f2.c10
	m1.src	m1.axp	m1.c10
	m2.src	m2.axp	m2.c10
G.722@48	f1.src	f1.axp	f1.c11
	f2.src	f2.axp	f2.c11
	m1.src	m1.axp	m1.c11
	m2.src	m2.axp	m2.c11
G.722.2@6.6	f1.src	f1.axp	f1.c12
	f2.src	f2.axp	f2.c12
	m1.src	m1.axp	m1.c12
	m2.src	m2.axp	m2.c12

Appendix I

Calculation of equipment impairment factors from scores provided by the instrumental model described in ITU-T P.862.2

(This appendix does not form an integral part of this Recommendation.)

If the instrumental model defined in [ITU-T P.862.2] is used for the derivation of *Ie,wb* values, the following interpolation line results from the processing of the 12 reference conditions of Table 1:

$$K = a \cdot Ie{,}wb{,}def + b \qquad \text{(I-1)}$$

with the values a = 0.8720 and b = 19.9487.

The resulting scatter plot is depicted in Figure I.1

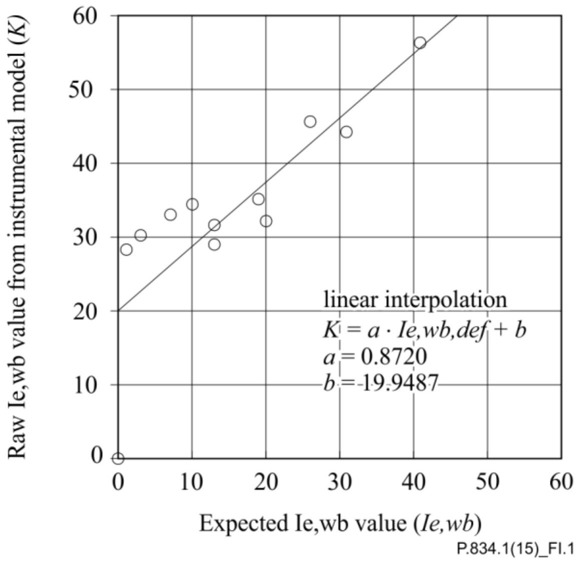

Figure I.1 – Linear interpolation line for the instrumental model according to ITU-T P.862.2

For each stimulus, an estimated MOS value is computed by [ITU-T P.862.2], v_i .

Mean MOS values are then calculated over all speech files processed with one specific condition, for each of the 12 reference circuits

$$MOS = \frac{1}{I}\sum_{i=1}^{I} v_i$$

I denotes the number of files per test condition.

The mean estimated MOS is transformed into a K value, using equations (1) to (3).

Then, an *Ie,wb* value for the new codec can be calculated as described in steps 3 and 4, using the relationship

$$Ie{,}wb = \frac{K - b}{a} \qquad \text{(I-2)}$$

with the values a = 0.8720 and b = 19.9487, and using the value:

$$R_{WB}(\text{clean}) = 129.0 \qquad \text{(I-3)}$$

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**, Issue 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	Environment and ICTs, climate change, e-waste, energy efficiency; 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	**Terminals and subjective and objective assessment methods**
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, 2015