International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.863

Amendment 1

(04/2020)

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

Methods for objective and subjective assessment of speech and video quality

---

Perceptual objective listening quality prediction

Amendment 1: Revised Appendix III – Prediction of acoustically recorded narrowband speech

Recommendation ITU-T P.863 (2018) – 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	P.10–P.19
Voice terminal characteristics	P.30–P.39
Reference systems	P.40–P.49
Objective measuring apparatus	P.50–P.59
Objective electro-acoustical measurements	P.60–P.69
Measurements related to speech loudness	P.70–P.79
Methods for objective and subjective assessment of speech quality	P.80–P.89
Voice terminal characteristics	P.300–P.399
Objective measuring apparatus	P.500–P.599
Measurements related to speech loudness	P.700–P.709
Methods for objective and subjective assessment of speech and video quality	**P.800–P.899**
Audiovisual quality in multimedia services	P.900–P.999
Transmission performance and QoS aspects of IP end-points	P.1000–P.1099
Communications involving vehicles	P.1100–P.1199
Models and tools for quality assessment of streamed media	P.1200–P.1299
Telemeeting assessment	P.1300–P.1399
Statistical analysis, evaluation and reporting guidelines of quality measurements	P.1400–P.1499
Methods for objective and subjective assessment of quality of services other than speech and video	P.1500–P.1599

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

Recommendation ITU-T P.863

Perceptual objective listening quality prediction

Amendment 1

Revised Appendix III – Prediction of acoustically recorded narrowband speech

Summary

Amendment 1 to Recommendation ITU-T P.863 replaces Appendix III, which gives advice on how ITU-T P.863 can be used for the prediction of listening quality of acoustically recorded speech data in a narrowband context.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.863	2011-01-13	12	11.1002/1000/11009
1.1	ITU-T P.863 (2011) Amd. 1	2011-11-09	12	11.1002/1000/11463
2.0	ITU-T P.863	2014-09-11	12	11.1002/1000/12174
3.0	ITU-T P.863	2018-03-16	12	11.1002/1000/13570
3.1	ITU-T P.863 (2018) Amd. 1	2020-04-24	12	11.1002/1000/14283

* 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 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/.

Recommendation ITU-T P.863

Perceptual objective listening quality prediction

Amendment 1

Revised Appendix III – Prediction of acoustically recorded narrowband speech

Replace Appendix III with the following:

Appendix III

Prediction of acoustically recorded narrowband speech

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

III.1 Background

Recommendation ITU-T P.863 is specified for the prediction of listening quality of acoustically recorded speech data in a fullband (FB) or super-wideband context only. That means the reference signal in this context is always a fullband or super-wideband speech signal and ITU-T P.863 is used in fullband operational mode.

This appendix advises how ITU-T P.863 can be used for the prediction of listening quality of acoustically recorded speech data in a narrowband context. Narrowband context means that the reference signal is narrowband. The prediction is using a narrowband scale and ITU-T P.863 predicts as a listener in a pure narrowband listening test.

Therefore, no modifications to ITU-T P.863 are required.

III.2 Requirements for acoustically recorded speech data to be assessed by ITU-T P.863

Besides the common rules for speech signals and acoustical recordings as described in ITU-T P.863, the prediction of listening quality of acoustically recorded speech in a narrowband context is restricted to the following items.

- Recordings which are close to the ear by, e.g., using handsets and headphones.
- Low level variation compared to a nominal level in recording and presentation level. The nominal level is 79 dB(A) sound pressure level (SPL) for monotic recording/presentation and 73 dB(A) SPL for diotic recording/presentation.
- The reference signal to ITU-T P.863 must be flat filtered. No IRSsend characteristic must be applied to the reference signal.
- It is recommended to apply the DC-removal filter as described in Annex C of [ITU-T P.501] to any speech signal used before applying the speech signal to ITU-T P.863.

ITU-T P.863 is not recommended for loudspeaker recordings or other recordings with considerable lower levels than the nominal level. ITU-T P.863 is applied to one ear signal only; binaural effects are not taken into account.

III.3 Pre-processing of speech and use of ITU-T P.863

It is recommended to reduce the sampling frequency of the flat reference signal and the test signal to 8 kHz to ensure narrowband audio and resample the signal afterwards to 48 kHz as required for FB mode. In addition, the digital level of both signals should be −26 dB OVL SPL according to [ITU-T P.56] independent of whether the signal was recorded monotically or diotically. These steps have to be performed in a pre-processing step and are not an integral part of ITU-T P.863. Ideally, the reference signal is directly obtained from a flat super-wideband or fullband signal by re-sampling and level readjustment.

Even though in this application reference signals with narrowband spectrum are used, ITU-T P.863 itself has to be used in fullband operational mode. In this mode the ITU-T P.863 internal IRSrcv filter characteristic is not used; instead a flat input filter is applied which is required for acoustically recorded data.

III.4 Interpretation of results

The result of ITU-T P.863 are mean opinion score (MOS) predictions without further mapping. The predicted MOS values are directly predicted on a one to five point scale. On average across various experiments, an excellent approximation of subjective results was achieved, without any noticeable systematic bias or scale deviation.

The MOS-LQO can be interpreted as a prediction of listening quality as it would be perceived in a narrowband listening only test with monotic or diotic presentation on the nominal level.

III.5 Example results

Three narrowband experiments with acoustically recorded speech material were evaluated. The three experiments were provided by Deutsche Telekom. The experiments were all conducted using German speech samples.

The experiments PAAM_1 and PAAM_2 have been conducted in the course of the PAAM project in 2002/2003. They consist of recordings using an ITU-T P.79 type 3.4 ear coupler. In both tests two types of randomly chosen off-the-shelf handset-phones and a headphone have been used as devices in receiving direction, the test conditions included applying a set of typical codecs (PAAM_1) and high-/low-pass conditions (PAAM_2) in front of the device. The purpose of the experiments was not the evaluation and comparison of different handsets rather the evaluation of a reproducible rank-order of the test conditions independent from the used handset when acoustically recorded. Both experiments are based on simulated transmission, only the receiving and playing handsets were physical devices.

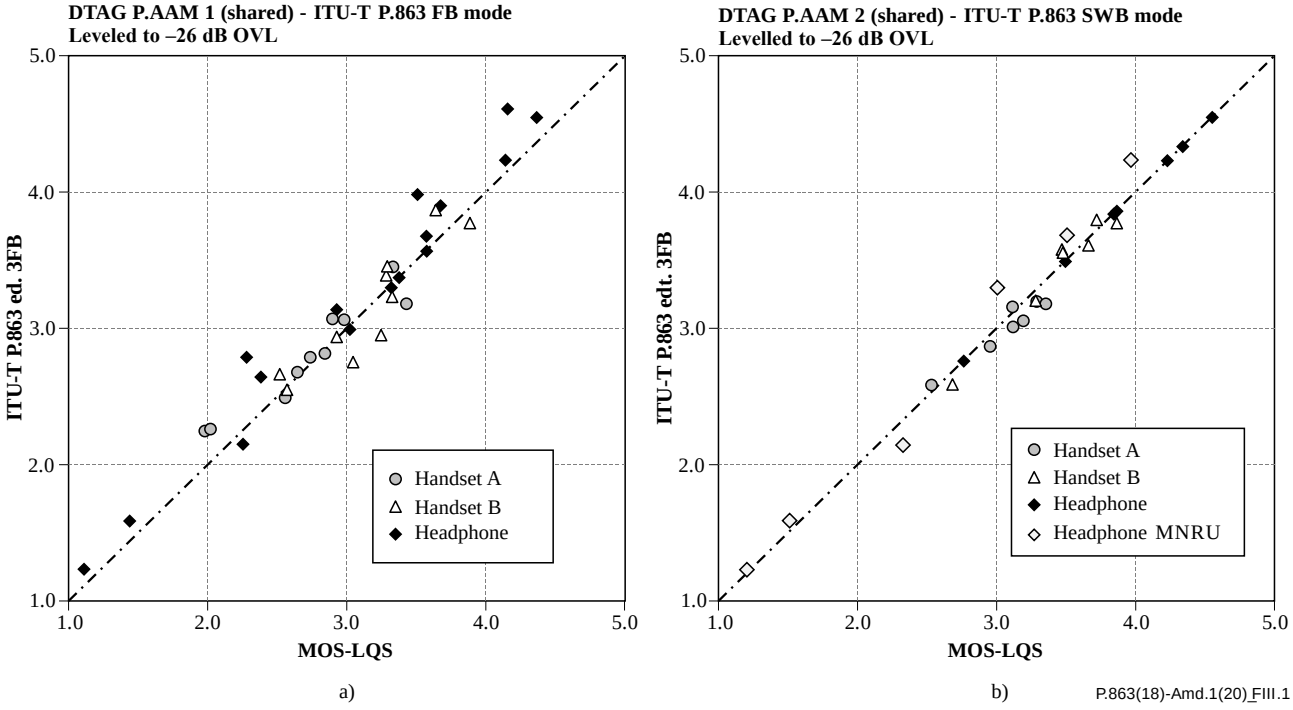
The third experiment (NB_1, 2001) is based mostly on recordings using a good standard headset (headset 1) that was used as a listening device over a variety of real VoIP connections on a laptop-PC. In contrast a few conditions have been produced by using a low-cost headset (headset 2) and a PC loudspeaker/desktop microphone. At sending side, an artificial mouth was used along with a headset microphone, wireless handset devices and ordinarily shaped handsets.

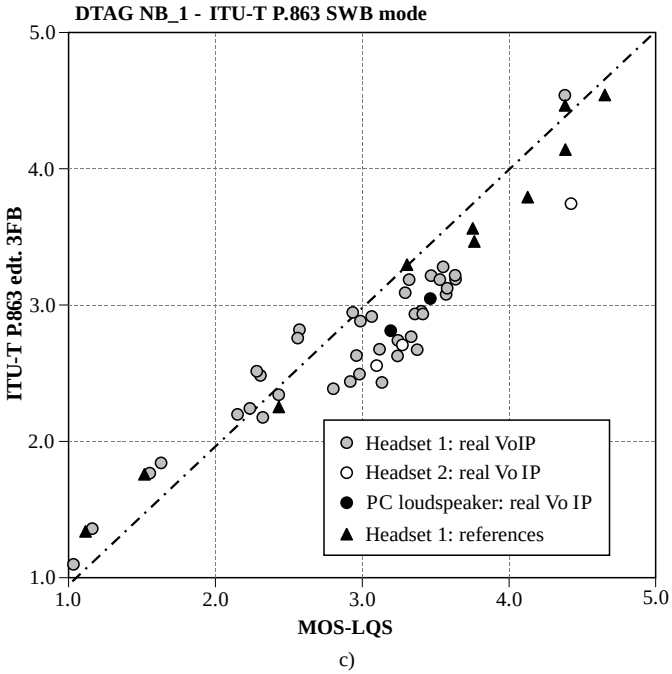
All three experiments were conducted according to [ITU-T P.800] with a naïve listening panel.

	DTAG PAAM_1	**DTAG PAAM_2**	**DTAG NB_1**
Pearson correlation	0.97	0.98	0.95
rmse (raw)	0.21	0.15	0.36
rmse* (raw)	0.09	0.04	0.21
rmse* (1^{st} order mapping)	0.07	0.01	0.12

The epsilon-insensitive rmse* figures have been calculated by the assumption of a ci95 of 0.2, because the actual ci95 for these experiments are not available for analysis. Please consider that the given performance figures above and the scatterplots below are obtained 'per condition' that means, the subjective MOS and the predictions of four different samples forming a test condition are averaged before statistical evaluation.

The experiments are predicted with good accuracy. However, the experiments reflecting transmission technologies as used in 2002/2003 include only a little variance of different acoustical devices but they reflect well the use of an ITU-T P.79 type 3.4 ear coupler.





SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
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	**Telephone transmission quality, telephone installations, local line networks**
Series Q	Switching and signalling, and associated measurements and tests
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, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2020