International Telecommunication Union

ITU-T
TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.863
Amendment 1
(11/2011)

SERIES P: TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Methods for objective and subjective assessment of speech quality

Perceptual objective listening quality assessment

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

Recommendation ITU-T P.863 (2011) – Amendment 1

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

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

Recommendation ITU-T P.863

Perceptual objective listening quality assessment

Amendment 1

New Appendix III – Prediction of acoustically recorded narrowband speech

Summary

Amendment 1 presents a new Appendix III to Recommendation ITU-T P.863 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. 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.

History

Edition	Recommendation	Approval	Study Group
1.0	ITU-T P.863	2011-01-13	12
1.1	ITU-T P.863 (2011) Amd. 1	2011-11-09	12

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/.

Table of Contents

Recommendation ITU-T P.863

Perceptual objective listening quality assessment

Amendment 1

New 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 super-wideband context only. That means the reference signal in this context is always a super-wideband speech signal and ITU-T P.863 is used in super-wideband operational mode.

This Appendix III to ITU-T P.863 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 that are close to the ear as, e.g., using handsets and headphones.

• Low variation compared to a nominal level in recording and presentation level. The nominal level stands for 79 dB(A) SPL in monotic recording/presentation and 73 dB(A) SPL in 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. In addition, the digital level of both signals should be to –26 dB OVL SPL according to [ITU-T P.56] independent of whether the signal was recorded monotically or diotically. Both steps have to be done in a pre-processing step and are not an integral part of ITU-T P.863. At best, the reference signal is directly gained from a flat super-wideband signal by down-sampling and level readjustment.

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

III.4 Interpretation of results

The outcome of ITU-T P.863 requires no further mapping. The predicted mean opinion score (MOS) values are directly given on a one to five point scale. Experiments on test data have not shown a systematic bias or different interpretation of the scale. On average across the experiments, a good approximation of results was reached.

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.
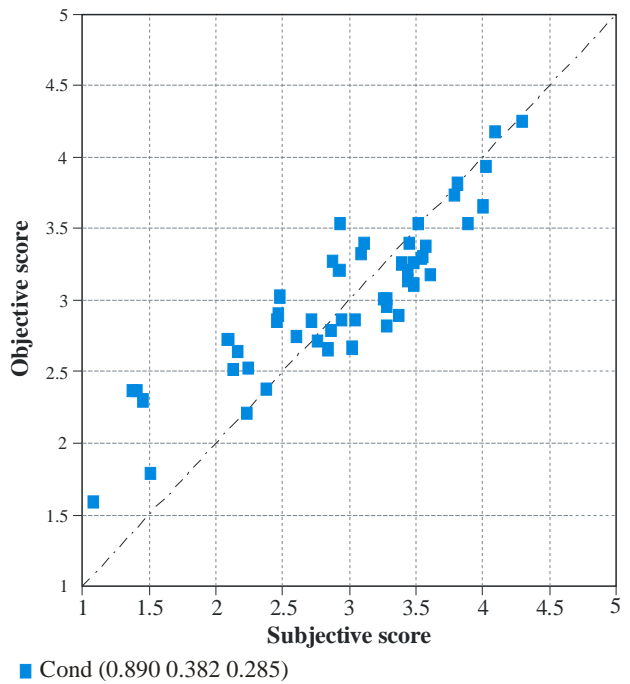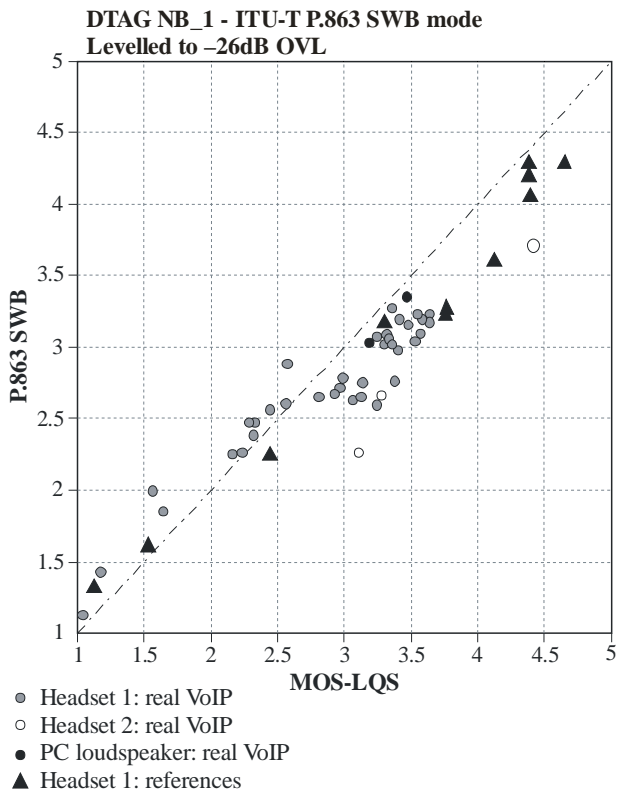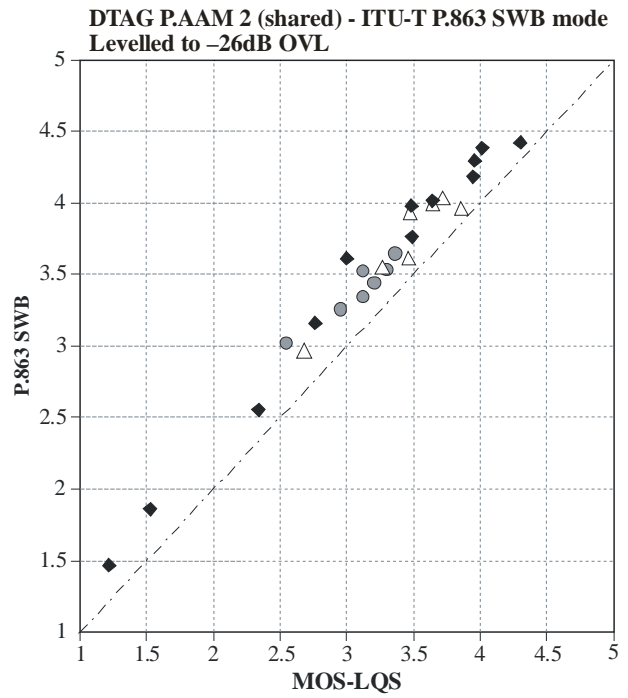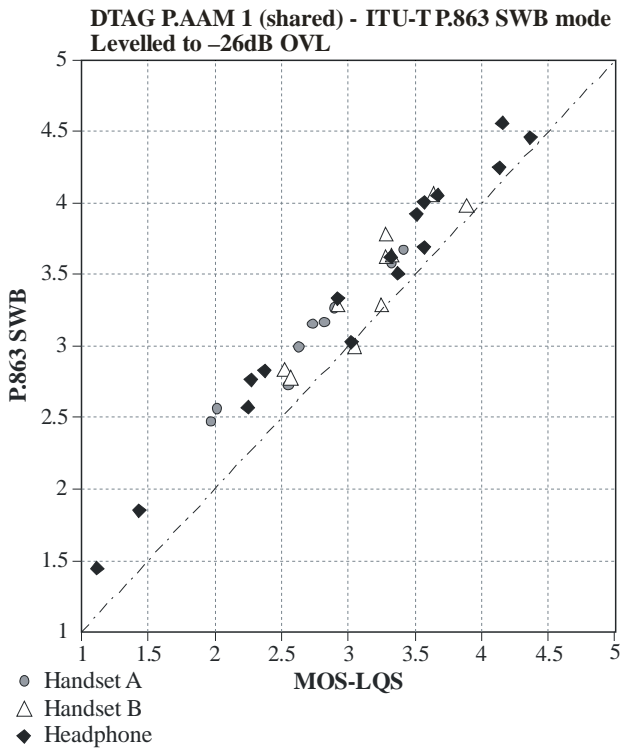
III.5 Example results

Four narrowband experiments with acoustically recorded speech material were evaluated. Three experiments were provided by Deutsche Telekom in German, and one experiment was provided by Netscout in English.

	DTAG PAAM_1	DTAG PAAM_2	DTAG NB_1	Netscout NB
Pearson Correlation	0.98	0.99	0.96	0.89
rmse*	0.02	0.01	0.12	0.23

The experiments are predicted with good accuracy. Each experiment contained a variety of different handsets and headphones at the recording side. The transmission channels were simulated codec conditions and real live channels (mostly GSM and DECT). At the sending side either a real device with acoustical insertion was used or an electrical input to, e.g., an ISDN line was applied.

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

DTAG P.AAM 1 (shared) - ITU-T P.863 SWB mode
Levelled to –26dB OVL

DTAG P.AAM 2 (shared) - ITU-T P.863 SWB mode
Levelled to –26dB OVL

● Handset A
△ Handset B
◆ Headphone

DTAG NB_1 - ITU-T P.863 SWB mode
Levelled to –26dB OVL

● Headset 1: real VoIP
○ Headset 2: real VoIP
● PC loudspeaker: real VoIP
▲ Headset 1: references

■ Cond (0.890 0.382 0.285)

P.863-Amd.1(11)_FIII.1

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	**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