INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.862.1

(11/2003)

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

Methods for objective and subjective assessment of quality

Mapping function for transforming P.862 raw result scores to MOS-LQO

ITU-T Recommendation P.862.1

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

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

ITU-T Recommendation P.862.1

Mapping function for transforming P.862 raw result scores to MOS-LQO

Summary

ITU-T Rec. P.862 provides raw scores in the range –0.5 to 4.5. It is desired to provide a MOS-LQO (P.800.1) score from P.862 to allow a linear comparison with MOS. This Recommendation presents the mapping function and its performance for a single mapping from raw P.862 scores to the MOS-LQO (P.800.1).

This will allow MOS-LQO scores from ITU-T Rec. P.862 to be comparable independent of the implementation of ITU-T Rec. P.862. The given function for transformation presented in this Recommendation has been optimized on a large corpus of subjective data representing different applications and languages.

Source

ITU-T Recommendation P.862.1 was approved by ITU-T Study Group 12 (2001-2004) under the ITU-T Recommendation A.8 procedure on 13 November 2003.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. 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, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

ITU-T Recommendation P.862.1

Mapping function for transforming P.862 raw result scores to MOS-LQO

1 Introduction

ITU-T Rec. P.862 provides raw scores in the range –0.5 to 4.5. It is desired to provide a MOS-LQO (P.800.1) score from P.862 to allow a linear comparison with MOS.

The aim of this separate Recommendation is to provide a single mapping from raw P.862 scores to MOS-LQO. This will allow MOS-LQO scores from ITU-T Rec. P.862 to be comparable independent of the implementation of ITU-T Rec. P.862.

2 Description of the training databases and conditions

The presented mapping function has been trained on a large corpus of test samples that covers VoIP, wireless, fixed and clean conditions/applications. Due to the importance of the applications, different weightings of the conditions/applications have been used (Table 1). Both simulated and field-collected samples and also background conditions have been considered. The field databases exhibit a high percentage of values in the lower end of the MOS scale. The reliability of the calibration for the lower end of the MOS scale is therefore ensured. The simulated databases, especially the clean conditions, guarantee the reliability at the upper end of the MOS scale.

Table 1/P.862.1

Database type	Weight % (No. of files per database type/ Total No. of files %)	Database characteristics	
		Simulated	Field
VoIP	19%	64.3%	35.7%
Wireless	43%	41.7%	58.3%
Fixed	30%	15%	85%
Clean	10%	100%	N/A

The normalization of the experiments has been ensured by the introduction of the certain MNRU conditions in the training process.

2.1 Subjective content

The accuracy of the calibration process is strongly dependent on the MOS panel statistics. The available training databases have been obtained from a large set of experiments. Different training databases are therefore characterized by different individual MOS standard deviation and confidence intervals. A number of at least 24 votes per condition/per file have been used.

2.2 Speech material and test procedure

The speech material is represented by sentence-pairs spoken by both female and male genders. In addition, nine languages are considered (British English, American English, Swedish, French, Italian, German, Finnish, Dutch and Japanese). It should be noted that in the training databases 85% of the samples represented the western languages and 15% of the samples represented the Japanese language.

All the test databases are based on ACR listening quality subjective tests.

3 The mapping function and its performance

The mapping function is given by equation 1.

$$y = 0.999 + \frac{4.999 - 0.999}{1 + e^{-1.4945*x+4.6607}} \quad (1)$$

The graph of the function is presented in Figure 1. Users may approximate this curve using other means (for example a lookup table) but this should be within the following constraints:

– The mapped MOS-LQO score should be within ±0.01 absolute of the curve defined in this Recommendation.

– This constraint should be held for all points, no more than 0.01 apart, over the whole raw P.862 range –0.5 to 4.5.

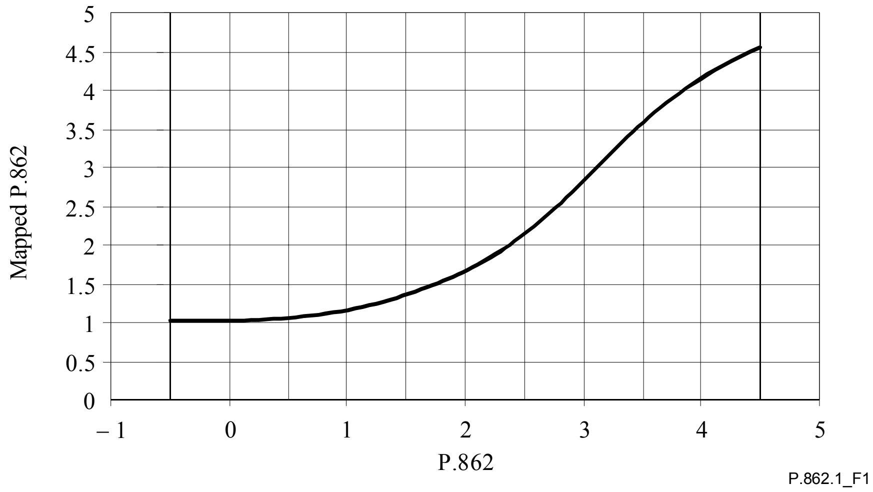

Figure 1/P.862.1 – P.862 Algorithm's mapping function

This constraint will help to ensure that MOS-LQO scores from P.862 will be comparable for all implementation of ITU-T Rec. P.862.

The inverse function that allows the transformation from the mapped P.862 scores to the raw P.862 scores is given by equation 2.

$$x = \frac{4.6607 - \ln\left(\frac{4.999 - y}{y - 0.999}\right)}{1.4945} \quad (2)$$

Three statistical measures have been used to analyse the performance of this mapping function: the Pearson correlation coefficient, R, the prediction error, Ep, the mean residual error, Em. In addition, the distribution of the mean residual error has been determined.

The performances of this mapping function are shown in Tables 2 and 3. The performances have been analysed overall corpus of training samples.

Table 2/P.862.1

Application	Metric	P.862	Mapped P.862
Overall	R	0.876	0.879
	CI 95%-lower limit	0.855	0.86
	Ep	0.492	0.441
	CI 95%-upper limit	0.501	0.449
	Em	–0.121	0.031

Table 3/P.862.1

MOS bins	<0.25	<0.5	<0.75	<1	<1.25	<1.5	<1.75	<2
P.862 scores (%)	36.1	66.63	87.44	96.95	99.56	99.96	99.96	100
Mapped P.862 scores (%)	41.92	72.64	91.22	98.4	99.64	99.88	99.96	100

4 Conclusions and comments

The P.862 algorithm's mapping function presented in this Recommendation has been optimized on a large corpus of subjective data representing different applications and languages.

The overall performance shown in Tables 2 and 3 shows that the presented mapping function performs better than the original P.862 and it is continuous within the whole raw P.862 scale, respectively –0.5 to 4.5. In addition, the mapping ensures a domain rescaling from –0.5 ... 4.5 to 1.02 ... 4.56.

The presented function is, therefore, recommended to be applied for all types of applications in order to get the subjective estimate MOS-LQO from P.862.

It should be noted that the presented function has some practical limitations:

i) The presented mapping function has been optimized on databases that originate from all types of applications. Other mapping functions optimized only for a specific application or language could perform better on that specific application or language than the presented function.

ii) Although the training databases contain a large percentage of samples with scores in the low MOS region, there is a lack of samples within the raw P.862 score range –0.5 to 1. Within this range the mapped P.862 function interpolates and therefore determines a prediction error Ep and a mean residual error Em that could be slightly higher than the one presented in Table 2. It is expected however that the error lies within the 95% CI of the Ep, respectively Em.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
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	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
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 and open system communications
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, 2004