

TTC標準
Standard

JJ-201.11

IP 携帯電話の通話品質評価法

A Method for Speech Quality Assessment
of IP Mobile Telephony

第3版

2025年5月22日制定

一般社団法人
情報通信技術委員会

THE TELECOMMUNICATION TECHNOLOGY COMMITTEE



本書は、一般社団法人情報通信技術委員会が著作権を保有しています。
内容の一部又は全部を一般社団法人情報通信技術委員会の許諾を得ることなく複製、転載、改変、転用及びネットワーク上での送信、配布を行うことを禁止します。

目次

目次	3
<参考>	4
1. 用語と定義	7
2. 関連する勧告・規定など	8
3. 規定範囲	11
4. 対象ネットワークと前提条件	12
4.1 レファレンス接続系	12
4.2 接続パターン	12
5. E-model 拡張	13
5.1 Wideband E-model の概要	13
5.2 Fullband E-model の概要	15
6. R 値の評価方法	17
6.1 R 値パラメータの分類と要規定項目	17
6.2 各パラメータの評価方法	20
6.2.1 音質パラメータ評価法 (leWB, Bpl, Ppl, le,eff,WB)	20
6.2.2 エコー評価法 (TELR)	26
6.2.3 遅延評価法 (T, Ta, Tr)	26
6.3 パラメータの測定条件	27
6.3.1 測定日、測定時間帯	28
6.3.2 評価サンプル時間	28
6.3.3 評価周期	29
7. R 値の解釈ガイドライン	30
8. R 値を補完するパラメータとその評価方法	32
8.1 遅延	32
8.2 エコー	32
8.3 音質	32
8.3.1 評価方法	33
8.3.2 R 値に対応した所要受聴 MOS 値	34
付録 I (参考) 欧米と日本の MOS 値の違い	34
付録 II (参考) POLQA の概要とアプリケーションガイドの要点	36
II.1 POLQA の概要	36
II.2 POLQA アプリケーションガイドの要点	36
II.2.1 一般的事項	36
II.2.2 参照信号	39
II.2.3 劣化信号	40
II.2.4 評価結果の分析	40
付録 III (参考) レファレンス条件を基準にした評価値の変換	41
付録 IV (参考) 相互接続環境における通話品質評価に関する留意事項	43
IV.1 相互接続点が単一の場合	43
IV.2 相互接続点が複数存在する場合	43
IV.3 品質評価時の留意事項	44
参考文献	45

< 参考 >

1. 国際勧告などとの関連

2. 改版の履歴

版数	制定日	改版内容
第1版	2015年11月12日	制定
第2版	2020年8月27日	・超広帯域／フルバンド帯域のIP電話の通話品質評価に関する記述の追加 ・ITU-T勧告G.107.1及びP.863改版に伴う記述の修正
第3版	2025年5月22日	・PESQ (ITU-T 勧告 P.862)、Wideband PESQ (ITU-T 勧告 P.862.2) およびこれらに関連するITU-T 勧告 P.862.1, P.862.3 に関する記述の削除、ITU-T 勧告 G.107 の改訂内容反映、新規に勧告化されたITU-T 勧告 P.833.2, 834.2 に関する記述の追加

3. 工業所有権

本標準に関わる「工業所有権等の実施の権利に係る確認書」の提出状況は、TTC ホームページで御覧になれます。

4. 参照している勧告、標準など

[1] ITU-T G.107.1 (06/2019)

Wideband E-model

ITU-T G.107.1 (2019) Corrigendum 1 (01/2020)

[2] ITU-T G.107 (06/2015)

The E-model, a computational model for use in transmission planning

JT-G.107 E-model 伝送計画のための計算モデル, 2006年6月

[3] ITU-T P.76 (11/1988)

Determination of loudness ratings; fundamental principles

[4] ITU-T P.79 (11/2007)

Calculation of loudness ratings for telephone sets

[5] ITU-T G.113 (09/2024)

Transmission impairments due to speech processing

JT-G.113 音声信号処理による伝送劣化, 2003年4月

- [6] ITU-T P.833.1 (04/2009)
Methodology for the derivation of equipment impairment factors from subjective listening-only tests for wideband speech codecs
- [7] ITU-T P.834.1 (06/2015)
Extension of the methodology for the derivation of equipment impairment factors from instrumental models for wideband speech codecs
- [8] ITU-T P.863 (03/2018)
Perceptual objective listening quality prediction
- [9] ITU-T P.863.1 (06/2019)
Application guide for Recommendation ITU-T P.863
- [10] ITU-T P.501 (05/2020)
Test signals for use in telephony and other speech-based applications
ITU-T P.501 Amendment 1 (06/2022)
New Appendix II: ITU-T P.863 results on ITU-T P.501 and ITU-T P.565 speech samples
- [11] ITU-T G.122 (03/1993)
Influence of national systems on stability and talker echo in international connections
- [12] ITU-T P.310 (06/2009)
Transmission characteristics for narrow-band digital handset and handset telephones
- [13] ETSI TS 101 329-5 (11/2000)
End-to-end quality of service in TIPHON systems; Part5: Quality of service (QoS) measurement methodologies
- [14] ETSI TR 101 329-6 (02/2002)
End-to-end quality of service in TIPHON systems; Part 6: Actual measurements of network and terminal characteristics and performance parameters in TIPHON networks and their influence on voice quality
- [15] ETSI TS 101 329-2 (01/2002)
End-to-end quality of service in TIPHON systems; Part2: Definition of speech quality of service (QoS) classes
- [16] ITU-T E.500 (11/1998)
Traffic intensity measurement principles
- [17] ETSI EG 201 769 (10/2000)
QoS parameter definitions and measurements; Parameters for voice telephony service required under the ONP voice telephony directive
- [18] ITU-T P.800 (08/1996)
Methods for subjective determination of transmission quality

[19] ITU-T G.114 (05/2003)

One-way transmission time

[20] ITU-T G.131 (11/2003)

Talker echo and control

[21] ITU-T G.165 (03/1993)

Echo cancellers

JT-G.165 エコーキャンセラ, 2003年4月

[22] ITU-T G.168 (04/2015)

Digital network echo cancellers

[23] ITU-T P.830 (02/1996)

Subjective performance assessment of telephone-band and wideband digital codecs

[24] ITU-T P.810 (03/2023)

Modulated noise reference unit (MNRU)

[25] ITU-T G.191 (05/2024)

Software tools for speech and audio coding standardization

[26] ITU-T P.56 (12/2011)

Objective measurement of active speech level

[27] ITU-T P.50 (09/1999)

Artificial voices

[28] ITU-T G.107.2 (03/2023)

Fullband E-model

[29] ITU-T P.833.2 (05/2024)

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

[30] ITU-T P.834.2 (09/2024)

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

5. 標準作成部門

網管理専門委員会

1. 用語と定義

IP 電話

音声信号を IP パケットに変換して伝送する VoIP(Voice over IP)技術による電話サービスの総称。

IP 携帯電話

無線通信を利用し、電話機の持ち運びが可能な携帯電話サービスのうち、音声信号を IP パケットに変換して伝送する VoIP(Voice over IP)技術を利用した IP 電話サービスを IP 携帯電話サービスと呼称する。代表的な例として VoLTE サービスが存在する。

VoLTE

音声信号を IP パケットに変換して伝送する VoLTE(Voice over LTE)技術による IP 携帯電話サービスの総称。

IP 携帯電話端末

IP ネットワークと接続可能な携帯電話端末。代表的な例として VoLTE サービスを利用可能なスマートフォン等が存在する。

ACR (Absolute Category Rating)法

絶対範疇尺度法。主観評価試験において評価者に対して「非常に良い」「普通」「非常に悪い」などの絶対的なオピニオン判断を求める試験法。

MOS (Mean Opinion Score)

予め通話品質に関する評価語に点数を割り当て、複数の評価者が受聴あるいは会話試験に基づいてオピニオン判断した結果の平均値を求めたもの。受聴試験により定義される受聴 MOS や、会話試験により定義される会話 MOS などがある。

TCLw (Weighted Terminal Coupling Loss)

端末の送話-受話部分の音響的、電氣的、振動的結合による損失量を周波数重み付けした値 (ITU-T 勧告 P.30, P.310 参照)。

エコーキャンセラ

ITU-T 勧告 G.165 (2.1 項), G.168 (3.8 項) に示されるような、エコー信号を減少させる仕組みの一種。

エコーリタンロス

受信端から端末に入力された信号が、エコーとなって送信端から出力される時の、入力信号とエコー信号とのレベル差 (減衰量: 単位 dB)。

メディアゲートウェイ (MG)

あるネットワークから提供されたメディアを、他のネットワークから要求されるフォーマットに変換する装置 (ゲートウェイ装置)。

揺らぎ吸収バッファ

音声パケットの到着間隔と音声の再生タイミングの差 (揺らぎ) を吸収するために IP 電話装置に実装されているバッファ。

2. 関連する勧告・規定など

[1] ITU-T G.107.1 (06/2019)

Wideband E-model

ITU-T G.107.1 (2019) Corrigendum 1 (01/2020)

[2] ITU-T G.107 (06/2015)

The E-model, a computational model for use in transmission planning

JT-G.107 E-model 伝送計画のための計算モデル, 2006年6月

[3] ITU-T P.76 (11/1988)

Determination of loudness ratings; fundamental principles

[4] ITU-T P.79 (11/2007)

Calculation of loudness ratings for telephone sets

[5] ITU-T G.113 (09/2024)

Transmission impairments due to speech processing

I

JT-G.113 音声信号処理による伝送劣化, 2003年4月

[6] ITU-T P.833.1 (04/2009)

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

[7] ITU-T P.834.1 (06/2015)

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

[8] ITU-T P.863 (03/2018)

Perceptual objective listening quality prediction

[9] ITU-T P.863.1 (06/2019)

Application guide for Recommendation ITU-T P.863

[10] ITU-T P.501 (05/2020)

Test signals for use in telephony and other speech-based applications

ITU-T P.501 Amendment 1 (06/2022)

New Appendix II: ITU-T P.863 results on ITU-T P.501 and ITU-T P.565 speech samples

[11] ITU-T G.122 (03/1993)

Influence of national systems on stability and talker echo in international connections

[12] ITU-T P.310 (06/2009)

Transmission characteristics for narrow-band digital handset and handset telephones

[13] ETSI TS 101 329-5 (11/2000)

End-to-end quality of service in TIPHON systems; Part5: Quality of service (QoS) measurement methodologies

[14] ETSI TR 101 329-6 (02/2002)

End-to-end quality of service in TIPHON systems; Part 6: Actual measurements of network and terminal characteristics and performance parameters in TIPHON networks and their influence on voice quality

[15] ETSI TS 101 329-2 (01/2002)

End-to-end quality of service in TIPHON systems; Part2: Definition of speech Quality of Service (QoS) classes

[16] ITU-T E.500 (11/1998)

Traffic intensity measurement principles

[17] ETSI EG 201 769 (10/2000)

QoS parameter definitions and measurements; Parameters for voice telephony service required under the ONP voice telephony directive

[18] ITU-T P.800 (08/1996)

Methods for subjective determination of transmission quality

[19] ITU-T G.114 (05/2003)

One-way transmission time

[20] ITU-T G.131 (11/2003)

Talker echo and its control

[21] ITU-T G.165 (03/1993)

Echo cancellers

JT-G.165 エコーキャンセラ, 2003年4月

[22] ITU-T G.168 (04/2015)

Digital network echo cancellers

[23] ITU-T P.830 (02/1996)

Subjective performance assessment of telephone-band and wideband digital codecs

[24] ITU-T P.810 (03/2023)

Modulated noise reference unit (MNRU)

[25] ITU-T G.191 (05/2024)

Software tools for speech and audio coding standardization

[26] ITU-T P.56 (12/2011)

Objective measurement of active speech level

[27] ITU-T P.50 (09/1999)

Artificial voices

[28] ITU-T G.107.2 (06/2019)

Fullband E-model

[29] ITU-T P.833.2 (05/2024)

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

[30] ITU-T P.834.2 (09/2024)

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

3. 規定範囲

本標準では、VoLTE に代表される IP 携帯電話サービスの総合品質指標として、ITU-T 勧告 G.107.1[1], G.107.2[31]に準拠した R 値を用いることを前提とし、この算出のためにサービス提供事業者が考慮すべき品質パラメータを特定すると共に、その評価法を規定する。品質パラメータや評価法に関しては、該当する ITU-T 勧告に準拠することを基本とし、国際標準化との整合を図ることとする。

本標準では IP 携帯電話サービスの品質に関する評価を対象とし、既存電話サービスは評価対象としていない。着目する品質要因は通話品質のみであり、サービス機能等の評価はもとより、接続品質や安定品質の評価は本標準の規定範囲外である。IP 携帯電話サービスの品質評価においては、サービス提供前の品質把握に加えて、サービス運用中の定常的な品質管理が重要であり、本標準はこれらの目的に供するものである。本標準が対象とする IP 携帯電話サービスの端末形態は「ハンドセット」であり、音声信号の周波数帯域は広帯域（50-7000 Hz）、超広帯域（50-14000Hz）、フルバンド（20-20000Hz）である¹。

本標準では、R 値による品質表示を、日本の国内事情を鑑みて適切に解釈するためのガイドラインを示す。さらに、ITU-T 勧告 G.107.1[1], G.107.2[28]が考慮していない品質要因がある点、R 値と主観品質の対応関係に関する検証が必ずしも十分でない点を鑑み、R 値を補完する品質パラメータ及びその評価方法も規定する。

¹ 本標準で前提とする ITU-T 勧告 G.107.1 Wideband E-model[1]および ITU-T 勧告 G.107.2 Fullband E-model[28]は、それぞれ広帯域(50-7000 Hz)ハンドセット通話およびフルバンド(20-20000Hz)ハンドセット通話を前提としたオピニオンモデルである。

4. 対象ネットワークと前提条件

4.1 レファレンス接続系

IP 携帯電話サービスの品質評価にあたっては、評価の公平性を確保することが必要であるため、標準系、限界系と呼ぶレファレンス接続系を設定することとする。IP 携帯電話サービスを提供する IP 網において、以下の2種類の系を設定することが考えられる。

標準系：距離、ルータ段数、通信速度、トラフィック量などを考慮し、遅延やパケットロスが平均的な2点間

限界系：距離（例えば最遠区間、サービスエリアの対角区間など）、ルータ段数、通信速度、トラフィック量などを考慮し、遅延やパケットロスが限界値に近い2点間

上記、標準系/限界系に対して IP 網内の2端末間について品質を評価することとする。また、選定した対地が妥当である根拠を示すことが必要である。サービス運用中に、大幅なネットワークの再編、輻輳制御の変更、コーデックの変更、等を実施する場合は、上記観点でレファレンス接続系を再定義する必要がある。

他社 IP 網との相互接続時においても同様に、端末と相互接続点間の標準系や限界系を考慮すること等で標準系/限界系を設定することとする。（相互接続環境におけるレファレンス接続系の設定方法に関する留意事項を付録IVに示す。）

なお、端末の品質を規定するためのレファレンスとなる接続系は、上記の接続系が事業者依存で異なるため特に規定しない。

4.2 接続パターン

端末とネットワークの接続パターンについては、複数の接続パターンが考えられるが、本標準では、既存電話サービスとの相互接続パターンは除き、IP 網で接続された IP 携帯電話端末間の品質を規定対象とする。ここでは、IP 携帯電話端末としてインタフェース(I/F)が IP である端末（スマートフォン等）を想定する。

以下の接続パターンの品質評価にあたっては両端ユーザの評価が必要であり、端末機種が異なる場合も想定されるため、両方向の品質を評価する。

IP 携帯電話端末 — IP 網 — IP 携帯電話端末

IP 網における事業者間の相互接続を考慮すると、上記の形態については、

IP 携帯電話端末 — 自 IP 網 — 相互接続先 IP 網 — IP 携帯電話端末

となり、相互接続点ではルータによる IP 網接続や、メディアゲートウェイのバックトゥバック接続等となる。

5. E-model 拡張

5.1 Wideband E-model の概要

Wideband E-model は ITU-T 勧告 G.107.1[1]で規定された計算モデルであり、狭帯域(300-3400Hz)ハンドセット通話を評価対象とする ITU-T 勧告 G.107 E-model[2]を拡張して、広帯域(50-7000Hz)ハンドセット通話を評価可能とした計算モデルである。Wideband E-model は、ハンドセットを使用する電話サービスの会話品質に影響を与える様々なパラメータの複合効果を推定するため、伝送計画ツールとして有用である。この計算モデルの出力は R 値であり、これは通話品質評価値の一つである会話 MOS 値に対応づけられる。従って、対象とするネットワークによりサービスを提供した場合の通話品質に対するユーザの意見（体感）をサービス提供に先立って評価することができる。

Wideband E-model では、通話品質に影響を与える品質要因（下記参照）が、心理尺度上で相加されることを仮定しており、R 値はこれらの品質要因に対する評価値に基づいて以下のように定義される。

$$R = RoWB - IsWB - IdWB - Ie-effWB + A \quad (5-1)$$

ここで、各項目は次の意味をもつ。

RoWB (Basic signal-to-noise ratio) : 回線雑音、送/受話室内騒音、加入者線雑音による主観品質劣化

IsWB (Simultaneous impairment factor) : OLR(ラウドネス)、側音、量子化歪による主観品質劣化

IdWB (Delay impairment factor) : 送話者エコー、受話者エコー、絶対遅延による主観品質劣化

Ie-effWB (Equipment impairment factor) : 低ビットレート符号化、パケット/セル損失などによる主観品質劣化

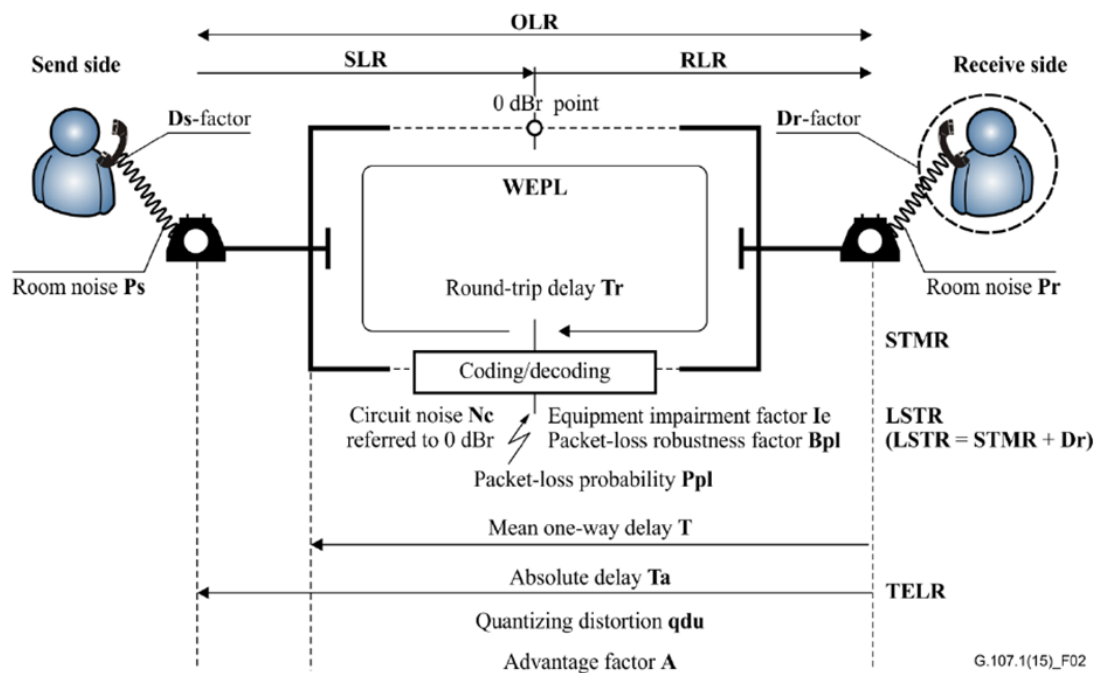
A (Advantage factor) : モバイル通信などの利便性が主観品質（満足度）に与える影響を補完

上記項目を計算するには次のパラメータを使用する。

略号	名称	概要	単位	範囲とデフォルト (カッコ内)
SLR	Send Loudness Rating	電話機を含めた送話ラウドネスを表す。値が小さいほど音量は大きい。	dB	TBD (+8)
RLR	Receive Loudness Rating	電話機を含めた受話ラウドネスを表す。値が小さいほど音量は大きい。	dB	TBD (+2)
STMR	Sidetone Masking Rating	電話機の送話側音量を表す。値が小さいほど側音は大きい。	dB	10 ... 20 (15)
LSTR	Listener Sidetone Rating	電話機の受話側音量を表す。値が小さいほど側音は大きい。	dB	13 ... 23 (18)
Ds	D-Value of Telephone, Send Side	送話側電話機の D ファクタ。受話側側音と送話側側音に対する感度差。		TBD (3)
Dr	D-Value of Telephone Receive Side	受話側電話機の D ファクタ。受話側側音と送話側側音に対する感度差。		TBD (3)
TELR	Talker Echo Loudness Rating	送話者エコー経路のラウドネスを表す。	dB	5 ... 65 (65)
WEPL	Weighted Echo Path Loss	受話者エコー経路のラウドネスを	dB	5 ... 110 (110)

			表す。		
T	Mean one-way Delay of the Echo Path		エコー経路の平均片道遅延	msec	0 ... 500 (0)
Tr	Round Trip Delay of the Echo Path		4W ループ区間の往復伝送遅延	msec	0 ... 1000 (0)
Ta	Absolute Delay in echo-free Connections		エンド・ツー・エンド片道遅延	msec	0 ... 500 (0)
IeWB	Equipment Impairment Factor		低ビットレートコーデックによる符号化歪主観品質劣化。		0 ... 56 (0)
Bpl	Packet-loss Robustness Factor		コーデックの packets 損失耐性をあらわす係数		4.3 ... 7.3 (4.3)
Ppl	Random Packet-loss Probability		ランダムパケット損失率	%	0 ... 20 (0)
Nc	Circuit Noise referred to 0 dBr-point		回線雑音量	dBm0p	TBD (-70)
Nfor	Noise Floor at the Receive Side		加入者線への誘導雑音量	dBmp	TBD (-96)
Ps	Room Noise at the Send Side		送話側の室内騒音量	dB(A)	TBD (35)
Pr	Room Noise at the Receive Side		受話側の室内騒音量	dB(A)	TBD (35)
A	Advantage Factor		利便性などによるユーザ評価向上を見込む補正項。広帯域通話での適用効果は未検討。		0 ... 20 (0)

上記パラメータを図で示すと以下の通り。(G.107.1 より転載)



5.2 Fullband E-model の概要

Fullband E-model は ITU-T G.107.2[28]で規定された計算モデルであり、狭帯域(300-3400Hz)音声通話を評価対象とする ITU-T 勧告 G.107 E-model を拡張して、超広帯域(50-14000Hz)及びフルバンド帯域(20-20000Hz)音声通話を評価可能とした計算モデルである。Fullband E-model は、超広帯域あるいはフルバンド帯域の電話サービスの会話品質に影響を与える様々なパラメータの複合効果を推定するため、伝送計画ツールとして有用である。この計算モデルの出力値は R 値である、これは通話品質評価値の一つである会話 MOS 値に対応づけられる。従って、対象とするネットワークによりサービスを提供した場合の通話品質に対するユーザの意見（体感）をサービス提供に先立って評価することができる。

Fullband E-model では、通話品質に影響を与える品質要因（下記参照）が、心理尺度上で相加されることを仮定しており、R 値はこれらの品質要因に対する評価値に基づいて以下のように定義される。

$$R = R_o, FB - I_s, FB - I_d, FB - I_e, eff, FB + A \quad (5-2)$$

ここで、各項目は次の意味をもつ。

R_{o,FB} (Basic signal-to-noise ratio) : 回線雑音、送/受話室内騒音、加入者線雑音による主観品質劣化

I_{s,FB} (Simultaneous impairment factor) : OLR(ラウドネス)、側音、量子化歪による主観品質劣化

I_{d,FB} (Delay impairment factor) : 送話者エコー、受話者エコー、絶対遅延による主観品質劣化

I_{e,eff,FB} (Equipment impairment factor) : 低ビットレート符号化、パケット/セル損失などによる主観品質劣化

A (Advantage factor) : モバイル通信などの利便性が主観品質（満足度）に与える影響を補完

上記項目を計算するには次のパラメータを使用する。

略号	名称	概要	単位	範囲とデフォルト (カッコ内)
SLR	Send Loudness Rating	電話機を含めた送話ラウドネスを表す。値が小さいほど音量は大きい。	dB	0 ... +18 (+8)
RLR	Receive Loudness Rating	電話機を含めた受話ラウドネスを表す。値が小さいほど音量は大きい。	dB	-5 ... :14 (+2)
STMR	Sidetone Masking Rating	電話機の送話側音量を表す。値が小さいほど側音は大きい。	dB	10 ... 20 (15)
LSTR	Listener Sidetone Rating	電話機の受話側音量を表す。値が小さいほど側音は大きい。	dB	13 ... 23 (18)
D _s	D-Value of Telephone, Send Side	送話側電話機の D ファクタ。受話側音と送話側音に対する感度差。		-3 ... +3 (3)
D _r	D-Value of Telephone Receive Side	受話側電話機の D ファクタ。受話側音と送話側音に対する感度差。		-3 ... +3 (3)
N _c	Circuit Noise referred to 0 dBr-point	回線雑音量	dBm _{0p}	TBD (-96)

Nfor	Noise Floor at the Receive Side	加入者線への誘導雑音量	dBmp	TBD (-96)
Ie,FB	Equipment Impairment Factor	低ビットレートコーデックによる符号化歪主観品質劣化。		0 ... 120 (0)
Bpl	Packet-loss Robustness Factor	コーデックの packets 損失耐性をあらわす係数		7.4 ... 18 (4.3)
Ppl	Random Packet-loss Probability	ランダム packets 損失率	%	0 ... 20 (0)
BurstR	Burst ratio	packets 損失パタンのバースト性		1 ... 8 (1)
Brf	Burstiness robustness factor	バースト損失に対するロバスト性		1 ... 20 (1)
Ta	Absolute Delay in echo-free Connections	エンド・ツー・エンド片道遅延	msec	0 ... 1700 (0)
sT	Delay sensitivity	遅延感度		0.4 ... 1 (1)
mT	Minimum perceivable delay	最小知覚遅延	ms	20 ... 200 (100)
A	Advantage Factor	利便性などによるユーザ評価向上を見込む補正項。広帯域通話での適用効果は未検討。		0 ... 20 (0)

6. R 値の評価方法

6.1 R 値パラメータの分類と要規定項目

Wideband E-model の入力パラメータおよび中間パラメータは以下の通り分類できる：

- A) アナログ伝送における通話品質パラメータなど、現在は影響がないと解釈
 - $N_c = -70$ dBm0p, $N_{for} = -96$ dBmp, WEPL=110 dB
- B) 適用指針が不明確であり、現時点では適用が不適切
 - $A = 0$
- C) 環境要因であり、制御不可能であるため、特定の環境を想定
 - $P_s = P_r = 35$ dB(A)
- D) 端末の設計パラメータであり、標準的特性を想定²
 - SLR=8 dB, RLR=2 dB, STMR=15 dB, LSTR=18 dB, $D_r = D_s = 3$, TELR (端末エコー) =65 dB
- E) IP 携帯電話サービスの設計パラメータであり、R 値導出の際に評価が必要
 - 音質： $I_e, W_B, B_{pl}, P_{pl}, I_{e,eff}, W_B$
 - エコー： TELR(既存電話網エコー)
 - 遅延： T, T_a, T_r

Fullband E-model の入力パラメータおよび中間パラメータは以下の通り分類できる：

- A) アナログ伝送における通話品質パラメータなど、現在は影響がないと解釈
 - $N_c = -96$ dBm0p, $N_{fo} = -96$ dBmp, $R_o, F_B = 148$, $I_s, F_B = 0$
- B) 適用指針が不明確であり、現時点では適用が不適切
 - $A = 0$, $B_{urstR} = 1$, $B_{rf} = 1$, $sT = 1$, $mT = 100$ ms
- C) 環境要因であり、制御不可能であるため、特定の環境を想定
 - なし
- D) 端末の設計パラメータであり、標準的特性を想定エラー! ブックマークが定義されていません。
 - SLR=8 dB, RLR=2 dB, STMR=15 dB, LSTR=18 dB, $D_r = D_s = 3$
- E) IP 携帯電話サービスの設計パラメータであり、R 値導出の際に評価が必要
 - 音質： $I_e, F_B, B_{pl}, P_{pl}, I_{e,eff}, F_B$
 - 遅延： T_a

項目 A から C については、全ての評価において共通的に用いるパラメータ値 (ITU-T 勧告 G.107.1[1], ITU-T 勧告 G.107.2[28]におけるデフォルト値) を示している。

項目 D は、標準的なハンドセット端末の特性を表すパラメータであり、端末機器 (スマートフォン等の IP 携帯電話端末) を提供する際には上記のような特性 (ITU-T 勧告 G.107.1[1], ITU-T 勧告 G.107.2[28]におけるデフォルト値) を実現することが望まれる。音量設計は既存電話網及び IP 携帯電話網との相互接続における過大もしくは過小音量を回避するためにも重要なパラメータであり³、これらのパラメータの評価法は ITU-T 勧告 P.76[3]及び P.79[4]に規定されている。

² 本標準ではハンドセットタイプの IP 携帯電話端末を想定しており、ハンズフリーで利用した場合やラウドスピーカ・マイクフォン等を接続した場合の品質は評価対象外である。

³ ネットワーク区間における音量損失/利得はないものと想定している。

エコー品質パラメータである TELR は、IP 携帯電話端末において生じるエコーを表現する場合と、既存電話網との相互接続時に既存電話網の加入者交換機において生じるエコーを表現する場合の 2 通りがある。前者については項目 D に含まれるが、後者についてはネットワークの設計パラメータであり、項目 E に含まれる。ただし、本標準では、既存電話網との相互接続パターンは規定対象としないため、理想的な値として項目 D で想定する TELR=65 dB を適用する。項目 D で想定している TELR=65 dB は、端末において、エコーキャンセラなどのエコー制御機構を備えていなければ実現困難な値である点に注意が必要である。また、Fullband E-model においては送受話エコーに関する劣化についてはカバーできていない点について注意が必要である。

項目 E は、本標準で評価法を規定するパラメータである。



図 6-1 / JJ-201.11 <パラメータの設定例 (IP 携帯電話端末ユーザ同士の通話の場合)>

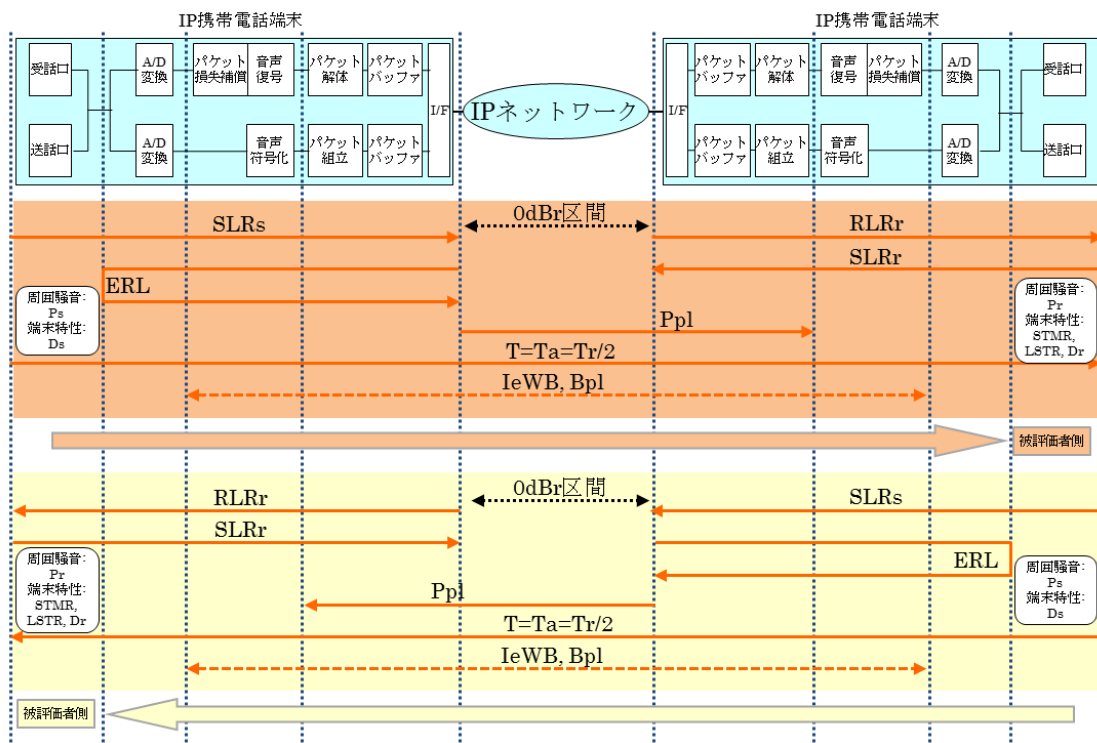


図 6-2 / JJ-201.11 <主要な R 値パラメータの規定点(IP 携帯電話端末ユーザ同士の通話の場合)>

6.2 各パラメータの評価方法

図6-1に、IP携帯電話端末ユーザ同士の通話を例に、パラメータの設定の考え方を示す。また、この時の主要なパラメータ規定点を図6-2に示す。以下、項目Eのパラメータ評価法について述べる。

6.2.1 音質パラメータ評価法 (IeWB, Bpl, Ppl, Ie,eff,WB) ⁴

音質を支配する要因には、音声符号化による符号化歪、ネットワークにおけるパケット損失、パケット到着揺らぎ吸収バッファにおけるパケット損失が挙げられる。Wideband E-modelにおいては、符号化歪は IeWB 値によって、パケット損失は当該コーデックのパケット損失耐性(Bpl 値)とパケット損失率(Ppl 値)によって表現され、これらのパラメータを統合して音質に関わる主観品質劣化量(Ie,eff,WB 値)が求められる。

音質に関わる主観品質劣化量(Ie-effWB 値)を決定する方法には以下の3通りが ITU-T において標準化されている：

- A) ITU-T 勧告 G.113 Appendix IV[5]に示される、符号化方式に対応して予め主観評価試験により求められた Ie,WB 値及び Bpl 値を用い、ITU-T 勧告 G.107.1 7.5 節[1]で規定される計算式で Ie,eff,WB 値を求める方法。⁵
- B) ITU-T 勧告 P.833.1[6]に規定される方法に準じて、主観評価試験から Ie,WB 値を決定する方法を用いて、パケット損失を含む試験条件下で求めた Ie,WB 値を Ie,eff,WB 値とみなす方法。⁶
- C) ITU-T 勧告 P.834.1[7]に規定される方法に準じて、客観評価試験(例えば、ITU-T 勧告 P.863[8])から Ie,WB 値を決定する方法を用いて、パケット損失を含む試験条件下で求めた Ie,WB 値を Ie,eff,WB 値とみなす方法。⁷

プランニングツールである Wideband E-model による R 値の計算は、ITU-T 勧告 G.113[5]で規定される Ie,WB 値及び Bpl 値を活用する方法 A が基本である。しかし、ITU-T 勧告 G.113[5]では、VoLTE に代表される IP 携帯電話で利用される様々な符号化方式に対し、Ie,WB 値及び Bpl 値が十分に定義されていないのが現状である。また、IP 携帯電話端末でのパケット損失率(Ppl 値)の測定が困難である等の課題もあり、方法 A により音質の主観品質劣化量(Ie,eff,WB 値)を求めることは困難である。

方法 B、C は、ITU-T 勧告 G.113[5]で Ie,WB 値及び Bpl 値が規定されていない符号化方式に対して、方法 A を利用するためにデータを充実する目的で勧告化されたものであり、主観/客観評価に基づいて当該符号化方式の Ie,WB 値を求める方法である。また、パケット損失を想定した条件下で、主観/客観評価に基づいて求めた当該符号化方式の Ie,WB 値は Ie,eff,WB 値とみなすとされている。従って、当該条件下で取得したデータでは、パケット損失の影響の重複計算を避けるため、ITU-T 勧告 G.107.1 7.5 節[1]で規定される Ie,eff,WB 値を求める計算式を利用しない点に注意が必要である⁸。

⁴ 本節において Fullband E-model の場合は、Ie,FB, Ie,eff,FB 及び Ie,eff,FB,def と読み替える。

⁵ Fullband E-model の場合は、ITU-T 勧告 G.113 Appendix V に示される Ie,FB 値及び Bpl 値を用い、ITU-T 勧告 G.107.2 7.1 節[28]で規定される計算式で Ie,eff,FB 値を求める。

⁶ Fullband E-model の場合は、ITU-T 勧告 P.833.1[6]ではなく ITU-T 勧告 P.833.2[29]を用いて Ie,eff,FB 値を求める。

⁷ Fullband E-model の場合は、ITU-T 勧告 P.834.1[7]ではなく ITU-T 勧告 P.834.2[30]を用いて Ie,eff,FB 値を求める。

⁸ 方法 C で用いる ITU-T 勧告 P.834.1[7]は、パケット損失条件下での適用は未検証とされてきたが、2015 年の勧告改訂にて、パケット損失を適切に評価可能な客観評価技術を用いることで、パケット損失による音質劣化量の計算も可能となった。

方法 B は、 $I_{e,eff,WB}$ 値を求めるために主観評価試験を実施する必要がある。適切な主観評価値を求めるためには、専用試験設備や主観評価試験ノウハウ、時間、コスト等を要するため、容易に $I_{e,eff,WB}$ 値を求めることはできない。

方法 C は、広帯域(50-7000 Hz)及びフルバンド帯域(20-20000Hz)までの音声符号化やパケット損失によって生じる音質の劣化を、適切に評価可能な ITU-T 勧告 P.863 POLQA[8]等の客観評価技術が存在することから、方法 B よりも比較的容易に $I_{e,eff,WB}$ 値を求めることができる。

従って、本標準では、方法 C (勧告 P.834.1[7]の客観評価活用方式)を基本とし、 $I_{e,eff,WB}$ 値を導出することとする。ただし、方法 A (勧告 G.113[5]のデータベース方式)に必要なデータが全て揃っている場合は、方法 A により $I_{e,eff,WB}$ 値を導出することとする。

次に、方法 C を用いて客観評価試験から $I_{e,eff,WB}$ 値を求める手順について述べる。

まず、ITU-T 勧告 P.834.1[7]に従って、ITU-T 勧告 P.863 POLQA[8]等の客観評価技術による音質評価試験の結果を、 $I_{e,eff,WB}$ 値に変換するための計算モデルを作成する。また、当該計算モデルは、客観評価試験で用いる音声サンプルのセットに対して 1 つ作成する。客観評価試験に用いる音声サンプルを変更しない限り、一度作成した計算モデルを、再度、作成し直す必要はない。例えば、客観評価試験に POLQA を用いる場合、POLQA のアプリケーションガイドラインである ITU-T 勧告 P.863.1[9]にて、最低 8 種類 (男女各 2 話者以上) の音声サンプルを用いて評価することが規定されているため、評価に利用する最低 8 種類の音声サンプルのセットに対して、評価結果を $I_{e,eff,WB}$ 値に変換するための計算モデルを 1 つ作成することになる。

当該計算モデルを用いることで、評価対象の客観評価結果から音質に関わる主観品質劣化量 ($I_{e,eff,WB}$ 値)を求めることができる。

音質の客観評価結果を主観品質劣化量 ($I_{e,eff,WB}$ 値)に変換する計算モデルは、以下の手順で構築する：

- 1) 音質の客観評価試験で用いる音声サンプルのセットを用意し、表 6-1 に示す ITU-T 勧告 P.834.1[7]で規定される 12 種類のリファレンス条件で音声符号化処理を実施する。なお、超広帯域・フルバンド音声を対象とする場合は、表 6-2 の 14 種類のリファレンス条件を用いる。以下、超広帯域・フルバンド音声を対象とする場合は、Wideband E-model を Fullband E-model と読み替え、ITU-T 勧告 P.834.1[7]を ITU-T 勧告 P.834.2[30]と読み替え、12 種類を 14 種類と読み替え、 $I_{e,eff,WB}$ については $I_{e,eff,FB}$ と読み替える。
- 2) 符号化処理された音声サンプルに対して客観評価試験を実施し、12 種類のリファレンス条件毎に評価結果を求める。
- 3) ITU-T 勧告 G.107.1 Annex A⁹[1]及び ITU-T 勧告 G.107 Appendix I[2]、ITU-T 勧告 P.834.1[7]に従って、12 種類のリファレンス条件毎に、客観評価結果を Wideband E-model で扱う R 値のスケール (R スケール) に変換する。ただし、評価結果のうち 4.5 を上回る結果 X が含まれる場合には、すべての評価結果に対して、1~X の値が 1~4.5 の範囲となるような線形補正を行ったうえ

⁹ 超広帯域・フルバンド音声を含む場合は ITU-T 勧告 G.107.2 Annex A に従って変換する。

で、R スケール変換を行う。

- 4) リファレンス条件 No.1(Clean)の客観評価結果の R スケール変換値を基準値として、リファレンス条件毎に音質の主観品質劣化量を表す $I_{e,eff,WB}$ 変換値(K)を求める。 $I_{e,eff,WB}$ 変換値(K)は、リファレンス条件 No.1(Clean)の客観評価結果の R スケール変換値から各条件の客観評価結果の R スケール変換値を減じることで求める。従って、リファレンス条件 No.1(Clean)の $I_{e,eff,WB}$ 変換値(K)は常に 0 となる。
- 5) 12 種類のリファレンス条件毎に対となる客観評価結果から求めた $I_{e,eff,WB}$ 変換値(K)と表 6-1 に示す $I_{e,eff,WB}$ デフォルト値($I_{e,eff,WB,def}$ 値)¹⁰のデータセットを用いて、線形補間により回帰直線を次式のように求め、係数 a と b の値を決定する。

$$K = a \cdot I_{e,eff,WB,def} + b \quad (6-1)$$

図 6-3 は、ITU-T 勧告 P.834.1[7]に記載される線形補間の一例であり、ITU-T 勧告 P.863 POQLA による客観評価結果を用いて求めた $I_{e-effWB}$ 変換値(K)と $I_{e-effWB}$ デフォルト値 ($I_{e-effWB,def}$ 値)の関係及びその回帰直線を示している。

- 6) 評価対象の音質に対する客観評価結果を主観品質劣化量($I_{e-effWB}$ 値)に変換する計算モデル(数式)は、手順 5 で決定した係数 a 及び b の値と回帰直線を用いて、次式のように定義する。

$$I_{e,eff,WB} = \frac{K - b}{a} \quad (6-2)$$

手順 2 から 4 に則って、手順 1 で用意した音声サンプルを用いて評価対象の音質に対する客観評価試験を実施し、評価結果を R スケールに変換して $I_{e,eff,WB}$ 変換値(K)を求め、式 6-2 に代入することで、評価対象の音質劣化量($I_{e,eff,WB}$ 値)を求めることができる。

表 6-1 / JJ-201.11 <音質に関わる主観品質劣化量のリファレンス条件(ITU-T 勧告 P.834.1)>

No.	略称	音声符号化方式	参照勧告	符号化レート (kbit/s)	$I_{e,eff,WB,def}$ 値 (デフォルト値) ¹¹
1	Clean	linear PCM, 16 bits	—	—	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

¹⁰ 表 6-1 に示す $I_{e,eff,WB,def}$ 値は、ITU-T 勧告 P.834.1[7]では、音声符号化方式による音質の主観品質劣化量を表す尺度として $I_{e,WB,def}$ 値と表記される。2015 年の勧告改訂にて、パケット損失による音質劣化を定量化可能となり、パケット損失が発生する条件下で求めた $I_{e,WB}$ 値は $I_{e,eff,WB}$ 値とみなすとされた。従って、本標準では、パケット損失を含む条件下での利用を想定するため、 $I_{e,WB,def}$ 値を $I_{e,eff,WB,def}$ 値と表記する。

¹¹ ITU-T 勧告 P.834.1[7]では、 $I_{e,WB,def}$ 値と表記され、音声符号化方式による音質の主観品質劣化量を表す尺度として表現されている。2015 年の勧告改訂において、パケット損失による音質劣化の定量化が可能となり、本勧告を用いて音質劣化量を求める際は、 $I_{e,WB}$ 値を $I_{e,eff,WB}$ 値として扱うとされるため、本標準では、 $I_{e,eff,WB,def}$ 値と表記する。

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

表 6 - 2 / JJ-201.11 <SWB 音声を含む場合の音質に関わる主観品質劣化量のリファレンス条件>

No.	略称	音声符号化方式	参照勧告	符号化レート (kbit/s)	Ie,eff,FB,def 値 (デフォルト値)
1	SWB Clean	linear PCM,16 bits	—	—	0
2	EVS@48	ACELP/MDCT	3GPP TS 26.445	48	10.2
3	EVS@32	ACELP/MDCT	3GPP TS 26.445	32	8.7
4	EVS@24.4	ACELP/MDCT	3GPP TS 26.445	24.4	7.2
5	EVS@16.4	ACELP/MDCT	3GPP TS 26.445	16.4	10.8
6	EVS@13.2	ACELP/MDCT	3GPP TS 26.445	13.2	17.1
7	EVS@9.6	ACELP/MDCT	3GPP TS 26.445	9.6	22.7
8	G.722.2@23.85	CELP	ITU-T G.722.2	23.85	29
9	G.722.2@12.65	CELP	ITU-T G.722.2	12.65	39
10	G.722.2@8.85	CELP	ITU-T G.722.2	8.85	60
11	G.722@64	ADPCM	ITU-T G.722	64	24
12	G.722@48	ADPCM	ITU-T G.722	48	60
13	G.729.1@32	CELP with TDBWE and TDAC	ITU-T G.729.1	32	26
14	G.729.1@24	CELP with TDBWE and TDAC	ITU-T G.729.1	24	35

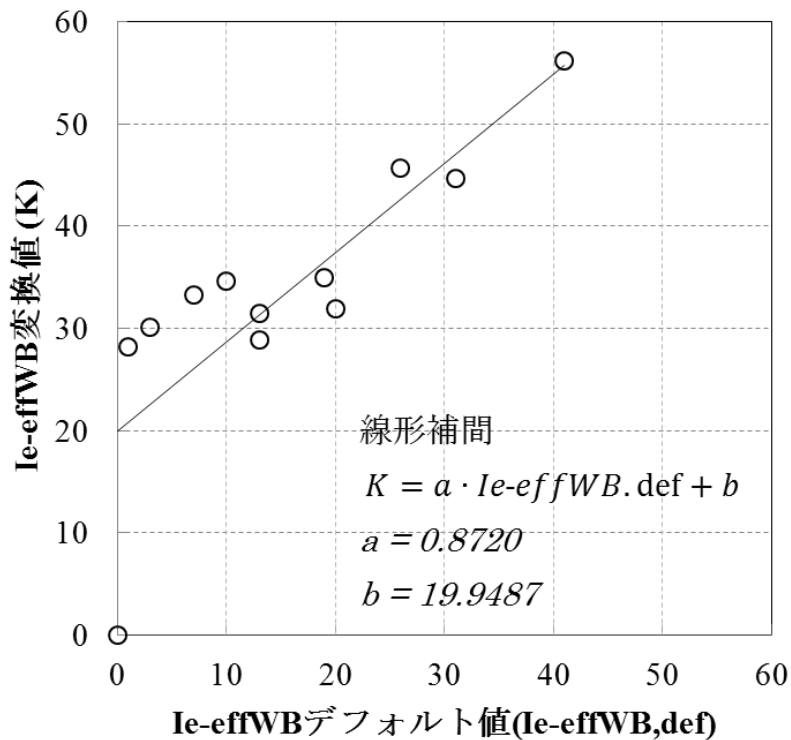


図 6-3 / JJ-201.11 <客観評価結果による音質の主観品質劣化量 (Ie-effWB 値) 計算モデル作成のための線形近似直線 (ITU-T P.834.1)>

手順 1 から手順 6 の実施例として、ITU-T 勧告 P.863 POLQA[8]を用いた実施手順を説明する。

手順 1 では、POLQA により音質の客観評価を実施するため、ITU-T 勧告 P.863.1[9]に従って、最低でも男女各 2 名が発話した 2 つ (合計 8 つ) の音声サンプルを用意する。例えば、ITU-T 勧告 P.501[10]のデータベースで提供される 2 文章 8 秒で構成される音声サンプル等を用いる。音声サンプルの詳細は、P.834.1 Annex A[7]や ITU-T 勧告 P.501[10]を参照されたい。

次に、用意した全ての音声サンプルに対して表 6-1 に示す 12 種類のリファレンス条件で音声符号化処理を実施する。音声符号化処理の詳細は、ITU-T 勧告 P.834.1 6.3 節及び Annex A[7]を参照されたい。

手順 2 では、手順 1 の音声符号化処理の前及び後の音声サンプルを用いて、POLQA による音質の客観評価試験を実施する。表 6-1 に示す 12 種類のリファレンス条件毎に、用意した音声サンプル (最低 8 つ) に対する POLQA による客観評価結果(MOS-LQO)の平均値を求める。POLQA による客観評価の実施手順等の詳細は、付録II及び ITU-T 勧告 P.863[8]、ITU-T 勧告 P.863.1[9]を参照されたい。

手順 3 では、表 6-1 に示す 12 種類のリファレンス条件毎に求めた 8 つ以上の音声サンプルに対する POLQA による客観評価結果 (MOS-LQO) の平均値を、R スケールに変換する。スケール変換は、ITU-T 勧告 G.107.1 Annex A[1]及び ITU-T 勧告 P.834.1 7 章[7]に記載される、狭帯域(300-3400Hz)通話を評価対象として ITU-T 勧告 G.107 E-model[2]により計算される R_{NB} 値と、広帯域(50-7000Hz)通話を評価対象として ITU-T

勧告 G.107.1 Wideband E-model[1]により計算される R 値の関係式¹²：

$$R = 1.29 \cdot R_{NB} \quad (6-3)$$

及び、ITU-T 勧告 G.107 Appendix I[2]に記載される、狭帯域(300-3400Hz)通話を評価対象とした会話 MOS 値 (MOS_{CQO}) と R_{NB} 値との関係式：

$$R_{NB} = \frac{20}{3} \left(8 - \sqrt{226} \cos \left(h + \frac{\pi}{3} \right) \right) \quad (6-4)$$

$$h = \frac{1}{3} \arctan 2 \left(18566 - 6750 \cdot MOS_{CQO}, 15 \sqrt{-903522 + 1113960 \cdot MOS_{CQO} - 202500 \cdot MOS_{CQO}^2} \right) \quad (6-5)$$

$$\arctan 2(x, y) = \begin{cases} \arctan \left(\frac{y}{x} \right) & : x \geq 0 \\ \pi - \arctan \left(\frac{y}{-x} \right) & : x < 0 \end{cases} \quad (6-6)$$

を用いて、POLQA による客観評価結果 (MOS-LQO) の平均値を、次式の MOS_{LQO} に代入して R スケールに変換する¹²：

$$R = 1.29 \cdot R_{NB} = 8.6 \left(8 - \sqrt{226} \cos \left(h + \frac{\pi}{3} \right) \right) \quad (6-7)$$

$$h = \frac{1}{3} \arctan 2 \left(18566 - 6750 \cdot MOS_{LQO}, 15 \sqrt{-903522 + 1113960 \cdot MOS_{LQO} - 202500 \cdot MOS_{LQO}^2} \right) \quad (6-8)$$

以上の手順により、リファレンス条件毎に POLQA による客観評価結果 (MOS-LQO) の平均値を R スケールに変換する。

手順 4 では、手順 3 で表 6-1 に示す 12 種類のリファレンス条件毎に求めた R スケール変換値のうち、リファレンス条件 No.1(Clean)の値を基準値として、リファレンス条件毎の音質劣化量を表す I_{e,eff,WB} 変換値(K)を求める。表 6-1 に示されるリファレンス条件 No. i の POLQA による客観評価結果 (8 つ以上の音声サンプルに対する MOS-LQO の平均値) の R スケール変換値を R_i、I_{e,eff,WB} 変換値を K_i として、次式によりリファレンス条件毎の I_{e,eff,WB} 変換値(K_i)を求める：

$$K_i = R_1 - R_i \quad (6-9)$$

また、リファレンス条件 No.1(Clean)は、式 6-9 の i に 1 を代入するため、K₁ が常に 0 となる。

手順 5 では、表 6-1 に示すリファレンス条件毎に対となる I_{e,eff,WB} デフォルト値(I_{e,eff,WB,def} 値)と手順 4 で求めた I_{e,eff,WB} 変換値(K)のデータセットを用いて、図 6-3 に示す線形補間により、I_{e,eff,WB} デ

¹² Fullband E-model の場合は、(6-3)及び(6-7)の関係式中の定数「1.29」を「1.48」に置き換える。

フォルト値(Ie,eff,WB,def 値)から Ie,eff,WB 変換値(K)を求める回帰直線である式 6-1 を導出し、係数 a 及び b の値を求める。

手順 6 では、手順 5 で決定した係数 a 及び b の値を定数として式 6-2 に代入した数式を、音質の客観評価結果を主観品質劣化量 (Ie,eff,WB 値) に変換する計算モデルとして定義する。

当該計算モデルは、手順 1 で用意した音声サンプルのセット (合計 8 つ以上) に対して一意に決定するため、評価対象の POLQA による客観評価を実施する際に、手順 1 で用意した音声サンプルのセットを用いることが、Ie,eff,WB 値の計算に当該計算モデルを用いるための適用条件となる。

以上の手順で構築した音質の客観評価結果を主観品質劣化量 (Ie,eff,WB 値) に変換する計算モデルにより、評価対象の主観品質劣化量 (Ie,eff,WB 値) を計算することができる。

6.2.2 エコー評価法 (TELR)

エコー経路のラウドネス (音量) を表す TELR は、以下の式で定義される。

$$TELR = SLR_r + ERL + RLR_r \quad (6-10)$$

ここで、SLR_r 及び RLR_r は、受話側 (被評価側) の SLR 及び RLR を表す。ERL(Echo Return Loss)は ITU-T 勧告 G.122[11]に規定されるエコーロスの意味する¹³。

本標準では、6.1 節に記した通り、IP 携帯電話端末において生じるエコーについては、理想的な値 (TELR=65 dB) を想定する。

6.2.3 遅延評価法 (T, Ta, Tr)

遅延時間の測定に際しては、揺らぎ吸収バッファ長及び音声パケット長をサービス提供時の設定にあわせることが重要である。Ta は送信側送話口から受信側受話口までの音声伝送遅延時間である。厳密には、IP 電話においては、T は送信側送話口と遠端の 2W/4W 変換点間で、Tr は近端と遠端の 2W/4W 変換点間で定義されるが、IP 携帯電話端末では、2W/4W 変換点がないことから、両パラメータを Ta の測定結果から以下の式で定義することとする。

$$T = \frac{T_r}{2} = T_a \quad (6-11)$$

Ta は、エンド・ツー・エンドの片道遅延時間を測定した結果で決定する、あるいはエンド・ツー・エンドの往復遅延時間を測定した結果(Tr)の 1/2 により決定することとする。

端末までエンド・ツー・エンドでサービスを提供する場合には、利用する IP 携帯電話端末を用いて Ta を

¹³ 端末機器の受話口と送話口の音響結合などによりエコーが生じる。これによるエコー量は TCLw(weighted Terminal Coupling Loss)により定量化されるが (ITU-T 勧告 P.310[14]参照)、IP 携帯電話品質評価においていかに TCLw を考慮するかについては今後の検討課題とする。

実測する。測定に用いる信号については ETSI TS 101 329-5 Annex B[13]、また、遅延測定例については ETSI TR 101 329-6 5.4 節[14]を参照されたい。

6.3 パラメータの測定条件

本節では、6.1 節 E 項に示した R 値パラメータのうち時間的及び空間的な変動要素のある音質とエンド・ツー・エンド遅延に関する測定条件を規定する。IP 網においては、一定の品質を保つことが困難であり、品質の変動が大きいため、平均特性のみで評価することは適切ではない。従って、ここでは ETSI TS 101 329-2 4 章[15]の規定に従い、『95%確率で品質を保証するための考え方』も合わせて示す。

図 6-4 にエンド・ツー・エンド遅延の測定サンプルを例示する。測定値は、端末間距離やルータ段数等による空間的変動要素と、測定日時による時間的変動要素により、同図のサンプル空間上に分布する。95%確率で品質を保証するためには、例示した全サンプル空間に渡る測定を実施することが本来望ましいが、測定のための稼働などから現実的ではない。特に、空間的変動要素（図中の横軸）全般に渡りサンプルを測定することは困難である。

そこで、空間的変動要素に対しては、標準系、限界系のサンプル測定とし、時間的変動要素の観点から平均特性及び 95%値を規定することとする。サービスの全体的な品質傾向を把握する目的では、例えば、標準系の平均特性を評価することも有効であるが、ユーザ保護の観点からは、限界系における 95%値を把握することが望ましい。

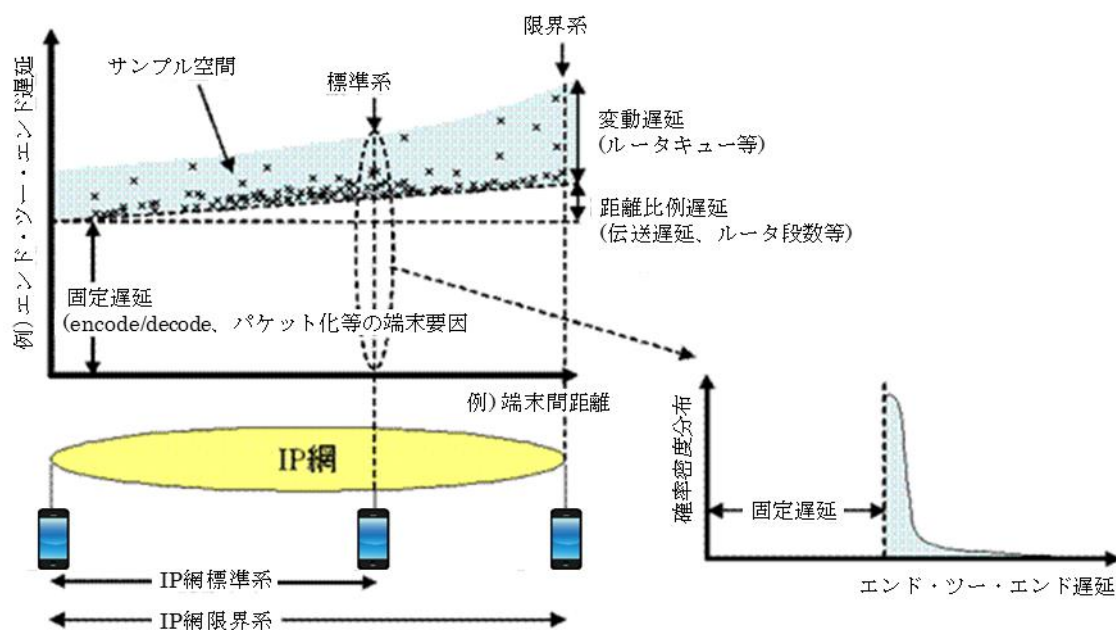


図 6-4 / JJ-201.11 <エンド・ツー・エンド遅延の測定サンプル>

6.3.1 測定日、測定時間帯

例えば、電話サービスの通話トラヒックは、以下の主たる変動要素が考えられる。

- 1) 評価日による違い（例えば、休日と平日）
- 2) 1日における通話時間帯による違い
- 3) 住宅用サービスまたは事務用サービスの違い

これらの通話トラヒック変動要素により、IP 網における音声信号転送用パケットトラヒックが変動し、パケット転送遅延、パケット転送遅延揺らぎ（ジッタ）、さらにはパケット損失といった、IP 網における各種伝送パラメータ量の変動を引き起こす。これらのパラメータ量の変動は、いずれも IP 携帯電話サービスの音声品質に少なからず影響を与える。これらの要因を考慮し、測定を実施する評価日と評価時間帯を選定する。

評価日の選定（ITU-T 勧告 E.500 5.2.3 項[16]など参照）：

電話や IP 網のトラヒックの変動を考慮して評価日を選定する。

時間帯の選定：

トラヒック状況の時間変動を考慮し、24 時間に渡った代表値を測定する。例えば、

- ・ 24 時間全ての測定
- ・ トラヒック量が大小などの時間帯について 1-2 時間程度サンプル測定
当該時間帯の選定根拠（例えば、24 時間のトラヒック変動特性など）が必要。ITU-T 勧告 E.500[18]、ETSI EG 201 769[17]などを参照して測定条件を明示。

上記の選定例を踏まえ、適切な評価日、評価時間帯を選定し、その妥当性を明示することとする。なお、サービス運用中の評価に関しては、特に再評価を行う周期は定めないが、定期的にトラヒック特性を分析し、上記観点で評価日、評価時間帯を適宜見直す必要がある。

6.3.2 評価サンプル時間

評価サンプル時間は、評価パラメータの特性、電話の保留時間、監視周期などを考慮する必要がある。IP 携帯電話を考慮した評価サンプル時間は、おおむね 1-10 分程度であることが望ましい。用いた評価サンプル時間（x min 測定、x sec 測定 など）については明示する。

・音質

本標準では、6.2.1 項に記したように、音質の主観品質劣化量は、客観評価を用いて求めることを基本としている。例えば、ITU-T 勧告 P.863 POLQA[8]を用いる場合、ITU-T 勧告 P.863.1[9]にて、最低でも 8 種類（男女各 2 話者以上）の音声サンプルで評価した結果の平均値を客観評価結果とすることが規定されている。また、音声サンプルの一例として ITU-T 勧告 P.501[10]で提供されている音声サンプルの時間長が 8 秒であることから、評価サンプル時間は、以下のようになる。

$$\begin{aligned} \text{評価サンプル時間} &= \text{音声サンプルの時間長} * \text{音声サンプル数} \\ &= 8 * 8 \end{aligned}$$

・エンド・ツー・エンド遅延

評価サンプル時間内におけるエンド・ツー・エンド遅延とは、評価時間内にサンプル測定した遅延の平均値である。例えば、以下の方法などがある。

- ・ 前述の客観評価を用いた音質測定の際に、音声サンプル毎にエンド・ツー・エンド遅延を測定し、その平均値を測定値とする方法
- ・ 評価サンプル時間内に、サンプル的にエンド・ツー・エンド遅延測定を複数回実施し、その平均

値を測定値とする方法

上記を踏まえ、適切な評価サンプル時間を選定し、その妥当性を明示することが必要である。

6.3.3 評価周期

評価周期、評価サンプル数は、それぞれの変動パラメータの特性を反映したものである必要がある。

- ・ 評価周期

例えば定間隔、ランダム間隔など適切な周期を選定

- ・ 評価サンプル数

評価値の平均や分散、評価値に求める精度などから、評価すべきサンプル数を決定する（ETSI EG 201 769 Annex D[17]など参照）。

- ・ 評価指標

平均値、95%値などの評価指標を選定

（評価サンプルから95%値を求める方法は、ETSI EG 201 769 Annex B[17]など参照）

上記を踏まえ、適切な測定周期、サンプル数を選定し、その妥当性を明示することが必要である。

7. R 値の解釈ガイドライン

通話品質評価の基本は、人間が評価対象系を介して会話した際に感じる品質（主観品質）であり、その評価法は ITU-T 勧告 P.800 Annex A[18]に規定される会話オピニオン評価である。会話オピニオン評価によって得られる評価値は MOS 値（特に、受聴 MOS 値と区別するために会話 MOS 値と呼ぶことがある）であり、1-5 の範囲で定義される。

一方、Wideband E-model が与える R 値は、0-129 の範囲で与えられ、ITU-T 勧告 G.107.1 Annex A[1]では R 値と会話 MOS 値の対応関係を提示している¹⁴。しかし、一般に欧米と日本の MOS 値は直接比較できないことが知られており¹⁵、上記対応関係を用いて得られる MOS 値を日本国内における評価として捉えるのは問題である。本標準では、欧米の MOS 値(MOS_w)を日本の MOS 値(MOS_j)に整合させるために、以下の変換を施すこととする¹⁶。

$$MOS_j = 0.8681 \cdot MOS_w + 0.0271 \quad (7-1)$$

一般に、MOS_j=3.5 は「90%の人が（日常電話を利用する立場から）当該品質は普通以上である」と判断、MOS_j=3.1 は「80%の人が当該品質は普通以上である」と判断、MOS_j=2.3 は「40%の人が当該品質は普通以上である」と判断する品質であると言われている[101]（オピニオン評価の評価語については ITU-T 勧告 P.800[18]参照）。

次に、ITU-T 勧告 G.107.1 Wideband E-model[1]により計算される R 値と、ITU-T 勧告 G.107 E-model[2]により計算される R_{NB} 値¹⁷の解釈の違いについて説明する。

広帯域(50-7000Hz)通話を評価対象とする R 値は 0-129 の範囲で与えられ、狭帯域(300-3400Hz)通話を評価対象とする R_{NB} 値は 0-100 の範囲で与えられており、同一の尺度として扱うことはできない。2 つの尺度が異なる値を示す一例として、狭帯域符号化方式である G.711PLC と G.729A を対象に、Wideband E-model と E-model で計算した R 値と R_{NB} 値の関係を図 7-1 に示す。

図中のプロットは、ITU-T 勧告 G.113[5]で定義される対象符号化方式の I_e 値及び I_{eWB} 値、B_{pl} 値を用いて、パケット損失率(P_{pl} 値)を 0%、3%、10%に設定した各条件に対して、遅延(T_a 値)を 0ms から 500ms の範囲で 50ms 刻みに変動させた条件毎に Wideband E-model と E-model で計算した R 値と R_{NB} 値を示している。

¹⁴ ITU-T 勧告 G.107.1 Wideband E-model [1]はプランニングツールであり、厳密な意味での MOS 値を推定可能なわけではない。Wideband E-model と実測 MOS 値の対応関係に関する問題点は、十分理解する必要がある。

¹⁵ 同一評価条件でも、言語の違いに起因する音声特徴量の違いが低ビットレート符号化等による音質劣化に影響を与えること、オピニオン評価における評価語の意味が言語によって微妙に異なること、国民性の違いにより、各評価語への投票傾向が異なることなどが原因として考えられる。

¹⁶ 本変換式の根拠データを付録 I に示す。

¹⁷ ITU-T 勧告 G.107 では R 値と表記されるが、本標準では、ITU-T 勧告 G.107.1 で計算される R 値と区別するために R_{NB} 値と表記する。

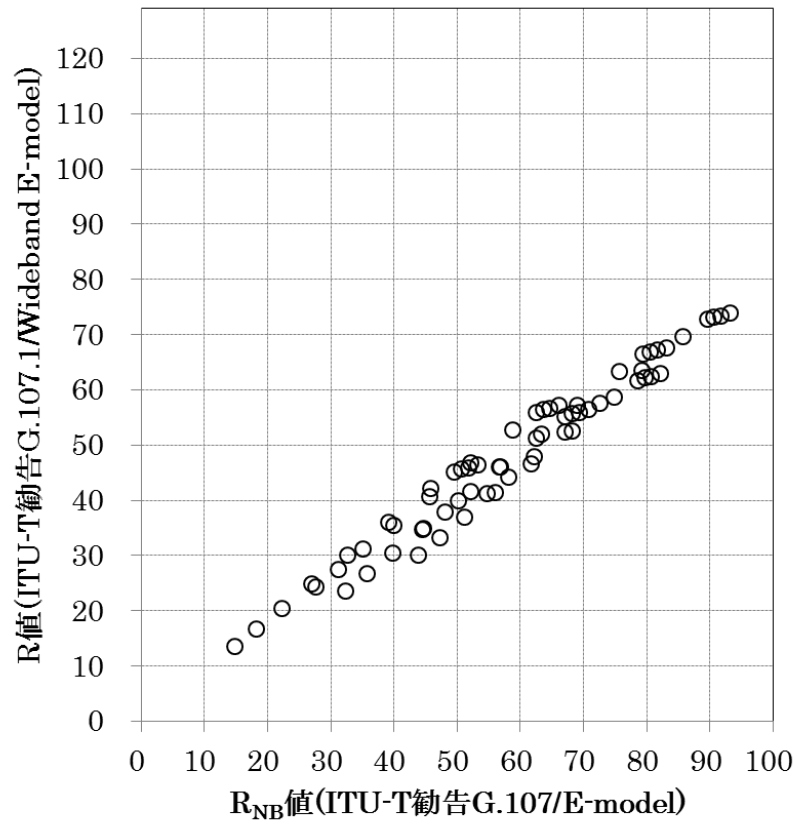


図7-1 / JJ-201.11 <R 値と R_{NB} 値の関係の一例>

8. R 値を補完するパラメータとその評価方法

ITU-T 勧告 G.107.1[1]は、プランニングツールとして開発されたモデルであり、実系を介して収録した音声サンプルに対して Wideband E-model では適切に評価できない品質要因も存在する。

例えば、ITU-T 勧告 G.107.1[1]は音質に関わる品質劣化を、コーデックとそのパケット損失耐性、及びエンド・ツー・エンドのパケット損失率で規定することを基本としており、同じコーデックを用いた場合には常に同じ評価値となる。しかし、一般に IP 携帯電話の品質には機器の実装に依存したばらつきがあり、評価の厳密性・公平性という観点からはこの点を考慮する必要がある。さらに、ITU-T 勧告 G.107.1[1]による主観品質推定精度に関しての検証は必ずしも十分ではなく、特に、音質・遅延・エコーなどの劣化の相加則についての定量的な検討例は少ない。従って、算出された R 値に基づいて IP 携帯電話サービスを評価した結果が必ずしもユーザが享受する品質と対応しない恐れもある。

上記のような点を考慮すると、Wideband E-model による総合通話品質指標 (R 値) を補完する意味で、個別の品質要因に対する評価もあわせて行うことが望ましい。従って、本章では、考慮すべき主な品質要因である「遅延」「エコー」「音質」に対する評価の必要性について記述する。通話品質は端末を含めたエンド・ツー・エンドで評価することが基本であり、特定の端末を想定したサービス提供にあたっては端末特性まで含めた評価が重要である。

8.1 遅延

遅延は本標準 6.2.2 項に示す方法で測定でき、得られた結果が主観品質に与える影響については ITU-T 勧告 G.114[19]に示されている。

8.2 エコー

エコーによる品質劣化は、エコー経路の音量 (TELR) 及びエコー経路の遅延時間 ($2 * T$) に依存する。つまり、同じ TELR であっても、 T が十分短ければエコーは電話機側音と同様に感じられ妨害感は少ないが、 T が長くなるに従いエコーとして知覚されるようになり妨害感が増す。

TELR の算出に必要なエコーリタンロス(ERL)の評価法は ITU-T 勧告 G.122[11]に規定されているが、IP インタフェースを有する端末のエコーリタンロスの具体的な測定法は ITU-T において検討中である。

TELR と片道伝送遅延時間(T)をパラメータとした主観評価試験結果に基づいて、 T に対応した所要 TELR (平均値) を求めた例が ITU-T 勧告 G.131(Figure 1/G.131)[20]に示されている¹⁸。ITU-T は当該勧告の“許容 (Acceptable)特性¹⁹”の適用を推奨している。つまり、TELR 値が、片道伝送遅延時間(T)を上述の“許容特性”に照らして導出される所要 TELR 値以上であることが望ましい。所要 TELR 値を満足できない場合には、エコーキャンセラを適用するなどの対応が必要である²⁰。

8.3 音質

方式の実装依存性やパケット損失パタンの影響を含めた実サービス品質をチェックする観点からは、実系

¹⁸ ITU-T 勧告 G.131 では、TELR の標準偏差を電話交換機の測定データから決定している。今後は、IP 携帯電話端末における TELR の標準偏差について実態値を把握した上で、適切な値の設定が必要である。

¹⁹ 許容特性は、TELR の変動を考慮して、オピニオン評価において「非常に悪い(1)」と評価するユーザの確率を 1%に抑えるために必要な特性である。

²⁰ エコーキャンセラの要求条件については ITU-T 勧告 G.165[21], G.168[22]に規定されている。

に対する主観／客観評価試験のいずれかを実施し、受聴 MOS 値（音声を受聴したときの音質にのみ着目した評価を受聴オピニオン試験と呼び、これにより得られる MOS 値を、会話 MOS 値に対して受聴 MOS 値と呼ぶ）を求めることが望ましい。

8.3.1 評価方法

以下の評価の規定点は、6 章に示した **IeWB** 規定点に準ずる。但し、測定上の都合により、これ以外の規定点で測定する場合、規定点の違いが評価値に影響を与えないよう配慮が必要である。特に、ネットワーク区間のみで品質を測定する際には、音声コーデックや揺らぎ吸収バッファのモデル化など、端末における品質劣化が適切に評価値に反映されるように配慮する必要がある。

8.3.1.1 主観評価試験による受聴 MOS の測定

主観評価試験により受聴 MOS 値を測定する具体的な方法として、ITU-T 勧告 P.800 Annex B[18]に規定される ACR 法を用いることができる。

主観評価値である受聴 MOS 値は、試験の枠組み（試験に用いる音声サンプルの品質バランスなど）の影響を受けるため、同一の評価条件であっても、異なる試験の枠組みで評価された受聴 MOS 値を直接比較することは必ずしも妥当でない。

上記オピニオン評価の問題点を回避するため、主観評価試験によって得られる受聴 MOS 値を以下の手順で正規化することができる²¹：

- (1) 主観評価試験には、評価対象とする音声処理条件に加え、ITU-T 勧告 P.810[24]で規定されるレファレンス条件（MNRU 条件）を含める²²。MNRU の Q 値は 0-40 dB の範囲で最低 5 段階とし、極力等間隔に設定する。（この主観評価試験により得られる受聴 MOS 値を MOS_{tmp} と定義する）
- (2) MNRU 条件に対する主観評価結果（ MOS_{tmp} ）を Sigmoid 関数で近似し、以下の関係式を得る。

$$MOS_{tmp} = f1(Q) \quad (8-1)$$

- (3) MNRU 条件の共通的な評価特性として以下の式を用い、これにより得られる MOS_{ref} と MOS_{tmp} の関係を定式化する。

$$MOS_{ref} = \frac{0.834 - 4.460}{1 + \exp\left(\frac{Q-16.57}{6.043}\right)} + 4.460 \quad (8-2)$$

$$MOS_{ref} = f2(MOS_{tmp}) \quad (8-3)$$

- (4) 上記のように得られた関係式に基づいて、当該試験により得られた MOS を変換する。

$$MOS = f2(MOS_{tmp}) \quad (8-4)$$

²¹ 主観評価値の普遍性を確保する手法としては ITU-T 勧告 P.830[23]に記述されている「等価 Q 値換算法」がある。本標準に示した MOS 値変換法は、本質的に等価 Q 値換算法と同じであり、Q 値を付録 III に示す Q vs. MOS 特性により MOS 軸上で表現したものである。

²² ITU-T 勧告 P.810[24]に準拠した MNRU 条件を実現するソフトウェアは ITU-T 勧告 G.191[25]により提供される。

8.3.1.2 客観評価試験による受聴 MOS の推定

8.3.1.1 項に示した音質に関する主観評価値（受聴 MOS 値）は、主観評価試験を行うことなく物理測定から推定する方法（客観評価法）を用いて求めることもできる。客観評価法の一例としては、ITU-T 勧告 P.863 POLQA[8]が国際標準化されている²³。

POLQA による音質（受聴 MOS）推定精度は、ITU-T SG12 において数多くの主観品質データベースを対象として検証されており、符号化歪やパケット損失による劣化など、IP 携帯電話における音質劣化要因の評価に適用可能である。POLQA の適用に際しては、アプリケーションガイドである P.863.1[9]に記述されている内容を十分理解する必要がある。なお、POLQA の概要と P.863.1 のポイントを付録Ⅱにまとめるので参照されたい。

POLQA の評価においては、ITU-T 勧告 P.863.1[9]に記述されている通り、少なくとも男女各 2 名の発声した 2 つ（計 8 つ）の音声サンプルを用いることとし、評価結果はこれらサンプルに対する評価値の平均値で定義する。つまり、POLQA で出力される評価値は MOS-LQO のため、音声サンプルに対する MOS-LQO の平均値が評価結果となる。特にパケット損失のように時間離散的に発生する劣化の評価においては、同一条件を複数回繰り返して評価し（つまり、上記 8 音声サンプルを複数回用いる）、それらの平均によって評価結果を表現することが望ましい。

8.3.2 R 値に対応した所要受聴 MOS 値

R 値に対応する会話 MOS 値は、遅延・エコー・音量などの会話要因を含めた評価値であり、前項に規定する音質のみに着目した受聴 MOS 値とは定義が異なる。しかし、所要の会話 MOS 値を実現するにはそれと同等以上の受聴 MOS 値を実現することが必要条件であるとの想定から、代表的な R 値に対応する受聴 MOS 値を表 8-1 に示す。

表 8-1 / JJ.201-11 <代表的な R 値に対応した所要受聴 MOS 値>

(A) 所要 R 値	(B) A に対する欧米の会話 MOS 値	(C) B に対する日本の会話 MOS 値	(D) 所要受聴 MOS 値 ^(注1)
100	3.9	3.4	3.4
80	3.2	2.8	2.8
65	2.6	2.3	2.3

(注1) 本表(D)の「所要受聴 MOS 値」とは「所要会話 MOS 値を実現するには、それと同等以上の受聴 MOS 値が必要である」との想定から決定した値であり、主観評価試験に基づいて決定した値ではない。

付録Ⅰ （参考）欧米と日本の MOS 値の違い

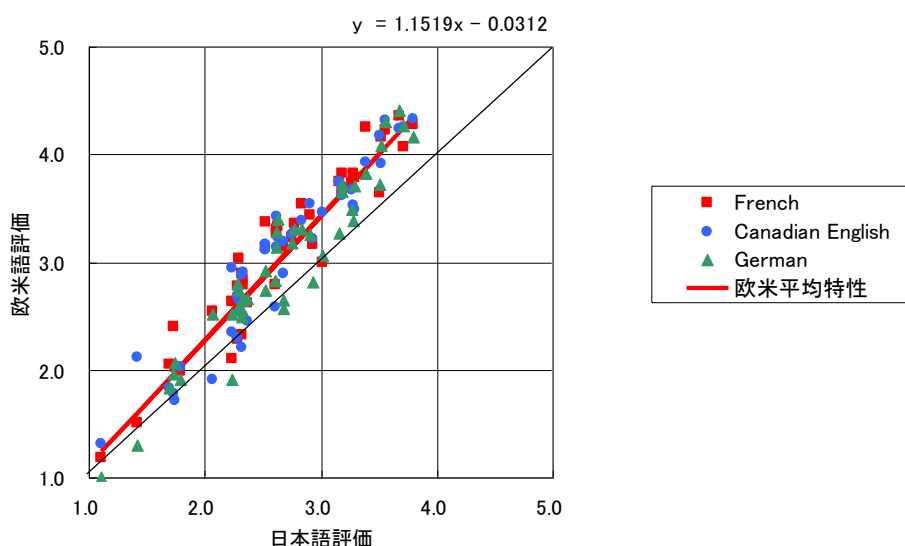
（本付録は参考資料であり、仕様ではない。）

²³ POLQA には、2011 年に Amendment 1 を含めて標準化されたエディション 1.1 (V1.1)、更なる推定誤差の改善がなされ 2014 年 9 月に標準化されたエディション 2.0 (V2.0)、2018 年 3 月に標準化されたエディション 3.0 (V3.0) が存在する。

本付録は、欧米と日本における MOS 値の違いを、ITU-T における国際主観品質評価試験結果に基づいて分析し、両者の関係を定式化する。具体的には、8kbit/s において音声を符号化するコーデックに関する国際標準である ITU-T 勧告 G.729 の標準化の過程において行われた特性規定フェーズ試験(Characterization Phase Tests)の評価結果[102][103]を用いた。当該試験は 3 つの試験から構成されており、本付録では試験 1（音声符号化方式のタンデム接続品質の評価）におけるデータを用いた²⁴。

同一符号化条件に対する欧米の MOS 値と日本の MOS 値の関係をプロットした図を付図 I-1 に示す。これによると、欧米の MOS 値(MOS_w)は日本の MOS 値(MOS_j)より高くなることが多く、平均的には両者の関係は以下の式で近似される。

$$MOS_j = 0.8681 \cdot MOS_w + 0.0271 \quad (\text{付 I -1})$$



付図 I-1 / JJ-201.11 < 欧米と日本の MOS 値の違い >

²⁴ 試験 2 は CCR 法 (ITU-T 勧告 P.800[18]参照) による評価結果であり、評価結果は一般的な MOS 値ではない。また、試験 3 は入力音声として背景騒音重畳音声を用いており、本標準における評価条件とは前提が異なる。

付録Ⅱ（参考） POLQA の概要とアプリケーションガイドの要点

（本付録は参考資料であり、仕様ではない。）

Ⅱ.1 POLQA の概要

POLQA は超広帯域音声及び電話帯域音声の受聴品質を推定する 2 つのモードを有する客観品質評価技術である。対象とする品質要因は、音声符号化歪、遅延の変動による時間構造歪、パケット／セル損失歪などである。詳細は次節(1)項及び ITU-T 勧告 P.863[8]を参照されたい。

POLQA は、参照信号と劣化信号の 2 つの音声サンプルを比較し、人間の聴覚の周波数分解能に対応したバークスペクトル領域での歪を算出し、これを主観的な音量感度に対応するラウドネス領域において表現する方式である。

Ⅱ.2 POLQA アプリケーションガイドの要点

本節では POLQA のアプリケーションガイドである勧告 P.863.1[9]の要点を記述する。詳細については当該勧告を参照されたい。

Ⅱ.2.1 一般的事項

(1)評価対象とする品質要因

POLQA が評価可能な評価項目、評価不可能な評価項目、評価の妥当性が未確認である評価項目を付表Ⅱ-1 から 3 に示す。

付表Ⅱ-1 / JJ-201.11

<POLQA を適用可能な評価項目>

品質要因
コーデックへの音声入力レベル
伝送路エラー
パケット損失及びパケット損失隠蔽処理(PLC)
ビットレート（当該Codecが複数のビットレートモードを有する場合）
コーデックのタンデム接続
送話側周囲騒音（注参照）
受聴試験における遅延変動
短区間の時間構造歪
長区間の時間構造歪
超広帯域モードにおける53～78 dB(A)の受聴レベル
パケット損失及びPCMタイプ符号化のパケット損失隠蔽処理(PLC)
時間的クリッピング及び振幅クリッピング
帯域幅制限やスペクトル整形を含む線形歪（平坦でない周波数特性）
周波数特性
符号化技術
ITU-T G.711, ITU-T G.711 PLC, ITU-T G.711.1
ITU-T G.718, ITU-T G.719, ITU-T G.722, ITU-T G.722.1, ITU-T G.723.1, ITU-T G.726, ITU-T G.728, ITU-

GSM-FR, GSM-HR, global system for mobile communications (GSM), enhanced full rate codec (EFR)
AMR-NB, AMR-WB (ITU-T G.722.2), AMR-WB+
PDC-FR, PDC-HR
Enhanced variable rate codec (EVRC) (ANSI/TIA-127-A), EVRC-B (TIA-718-B)
Skype (SILK V3, iLBC, iSAC and ITU-T G.729)
Speex, QCELP (TIA-EIA-IS-733), iLBC, CVSD (64 kbit/s, "Bluetooth")
MPEG-1 audio layer 3(MP3), advanced audio coding (AAC), AAC-LD
EVS, OPUS

アプリケーション
コーデック評価
端末試験、送受話方向における音響パスおよびトランスデューサの影響
帯域幅拡張
デジタルあるいはアナログ接続による実網評価
擬似網やプロトタイプ網の評価
UMTS, CDMA, GSM, TETRA, WB-DECT, VoIP, POTS, PSTN, Video Telephony, Bluetooth
音声区間検出、自動利得制御
音声強調装置、雑音抑圧
不連続送信、快適雑音挿入
注：周囲騒音のある条件で評価を行う場合、POLQAの参照信号として周囲騒音を含まないCleanな音声信号を、劣化信号として周囲騒音を含む音声信号を用いる。

付表III-2/JJ-201.11 <POLQAによる評価が適切でない評価項目>

品質要因
会話試験における遅延劣化
送話者エコー
側音
受話側の周囲雑音
アプリケーション
インサービス非侵入型評価
双方向通話品質

付表III-3/JJ-201.11 <POLQAの適用の妥当性が未確認である評価項目>

品質要因
話者依存性
複数話者による同時発話
符号器と復号器のビットレートミスマッチ
網情報信号に対する符号化性能
擬似音声信号に対する符号化性能
楽音に対する符号化性能
受話者エコー
符号化技術
4 kbps以下の符号化技術

(2)適用シナリオ

POLQAは、実験室環境における評価や計算機シミュレーションによる評価に加え、実網での評価にも適用可能である。但し、実網評価では同一条件での複数測定が困難である場合があり、その際には品質推定結果の信頼性が低下する点に注意が必要である。

II.2.2 参照信号

P.863.1[9]の内容にもとづく参照信号条件を、POLQA の参照信号条件とする。条件を満たす参照信号として、ITU-T 勧告 P.501[10]のデータベースで提供される音声信号や多言語音声データベース[104]に収録されている日本語文章音声等が存在する。

(1) 信号時間長

音声サンプルの継続時間長は 12 sec 以下とする。長時間のレファレンスファイルを使用する場合には、発話区間が 3 sec 以上かつ 6 sec 以下となるように複数の音声サンプルに区切り、それぞれの POLQA による評価値(MOS-LQO)を算出する。

(2) 発話区間

発話区間は 3 sec 以上かつ 6 sec 以下とし、2 つ以上の短文章を含み、かつ、発話区間のインターバルは 500ms 以上とする。

(3) 音声実効レベル

参照信号をデジタル信号として記録する際には、クリッピング等の劣化を避けるため、ITU-T 勧告 P.56[26]に規定される音声区間の信号レベルを-26 dBov とする必要がある。この時、雑音レベルは-75 dBov(A)未満とする。

(4) 擬似音声の適用

ITU-T 勧告 P.50[27]に規定されるような擬似音声信号の利用は、現時点で推奨されていない。

(5) 音声サンプル録音の要求条件

音声サンプルの録音環境は、背景騒音が 30 dB SPL(A)以下で、200 Hz を超える信号の反響時間が 300 msec 以下でなければならない。また、音声サンプル収録では、マイクと音源の距離が 10 cm 程度とする。超広帯域音声モードで評価する場合の参照音声信号は、16bit Linear PCM, 48 kHz サンプリングとし、50-14000 Hz のバンドパスフィルタリングを行う。また、電話帯域音声モードで評価する場合の参照音声信号は、16bit Linear PCM, 8 kHz サンプリングとし、100-3800 Hz のバンドパスフィルタリングを行う。

(6) 話者と文章のバリエーション

評価は最低でも男女各 2 名が発話した 2 つの音声サンプル (計 8 サンプル) を用いることが推奨されている。POLQA による評価値(MOS-LQO)は、音声サンプルの話者や短文章の違いに依存して変化するため、少ない音声サンプルではこの点が出力値に影響することに注意する必要がある。特に低レート符号化方式の評価においては、選択した話者や短文章により出力値が大きく異なることが確認されている。

(7) 先頭及び末尾の無音区間長

音声サンプルの先頭及び末尾には最低 0.25 sec の無音区間が必要である。

以下、測定ノイズを抑制するために考慮すべき点を記述する。

a) インピーダンス不整合

インピーダンスの不整合は減衰の要因となりスコアの悪化に繋がる。日本の電話機のインピーダンスは 600Ω であるため、測定端末のアナログポートに接続する測定機は 600Ω のものを使用する。

b) アース、ノイズ

測定時のノイズ発生の可能性を抑えるため、測定端末の電源ケーブルはグラウンドピンをもつことが望ましい。電源ケーブルが2ピンであっても、本体のグラウンド端子を測定機のグラウンドと接続することでノイズが低減できる場合がある。また、AC 電源は適切な接地がなされている必要があり、モーター等ノイズ発生源となるような機器が接続されていない系統のものを使用する。

II.2.3 劣化信号

参照信号と同様に、P.863.1[9]の内容にもとづく劣化信号条件を、POLQA の劣化信号条件とする。

(1) 音声実効レベル

クリッピングや雑音の影響を取り除くため、受聴音声レベルは -26 dBov となるように調整する。この -26 dBov は 79 dB SPL の音響レベルに相当するものとしている。

II.2.4 評価結果の分析

評価結果は異なる音声サンプル（最低でも男女各2名が発話した2つの音声サンプル（計8サンプル））に対する評価結果の平均値で示す。つまり、POLQA で出力される評価値は推定受聴 MOS 値である MOS-LQO のため、音声サンプルに対する MOS-LQO の平均値が、POLQA による評価結果となる。

付録Ⅲ (参考) レファレンス条件を基準にした評価値の変換

(本付録は参考資料であり、仕様ではない。)

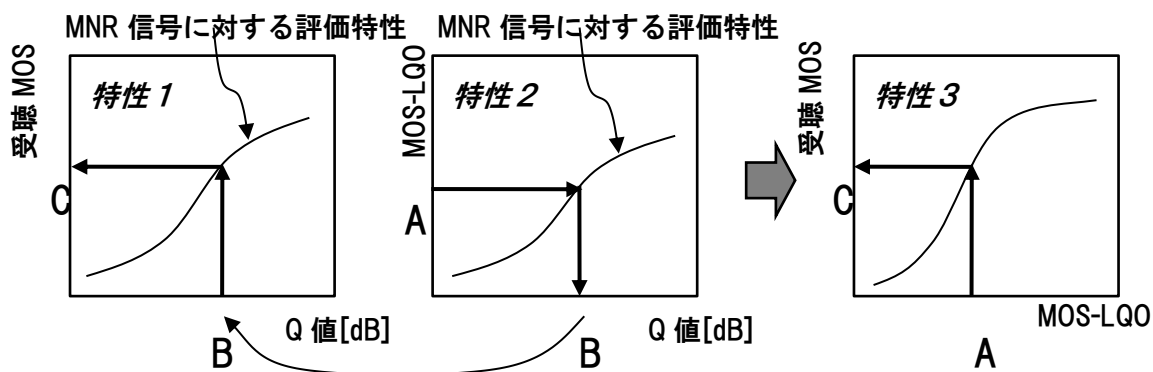
絶対評価値である平均オピニオン評点(MOS: Mean Opinion Score)は、主観評価試験において評価される音声サンプルセットの品質バランス、評価者の国民性、評価者集団の違い、評価者へのインストラクションの与え方等、「試験の枠組み」の影響を受けやすく、異なる試験結果を直接比較することは困難である場合が多い。

このような主観評価の問題点を解決する方法として、評価対象となる音声の品質を、これと MOS 値が等しくなるレファレンス信号の物理量で表す方法が用いられている (ITU-T 勧告 P.830[23]参照)。このときにレファレンス信号として用いられるのが、音声に振幅相関雑音を付加した MNR 信号 (ITU-T 勧告 P.810[24]参照) であり、MNR 信号の品質は S/N 比 (これを Q 値と呼ぶ) により制御される。具体的には、各主観評価試験において得られる Q 値と MOS 値の関係に基づいて、評価対象音声に対する MOS 値と同じ MOS 値を与える MNR 信号の Q 値を求め (この値を「等価 Q 値」と呼ぶ)、これをその音声の評価値とする。この方法によれば、試験の枠組みの違いにより同一の評価対象音声に対する MOS 値が変動する場合にも、評価対象音声とレファレンス信号の品質の相対関係は保存されることが期待できることから、試験の枠組みの影響を回避した評価が可能となる[105]。

客観評価の場合にも、評価対象となる符号化音声と客観評価値が等しくなる MNR 信号の Q 値を求めることにより等価 Q 値を定義できる。しかし、等価 Q 値による品質表現は R 値による品質表現と直接対応付けができないため、ここでは等価 Q 値の考え方をベースとしつつ、MOS 値による品質表現を試みる。

具体的には、「Q 値と受聴 MOS 値の関係【特性 1】」を主観評価試験により定式化すると共に、「Q 値と客観評価値(MOS-LQO)の関係【特性 2】」を客観評価試験により定式化する。これらの関係から「客観評価値(MOS-LQO)と受聴 MOS 値の関係【特性 3】」を決定する。

この関係を付図Ⅲ-1 に示す。例えば、評価対象音声に対する MOS-LQO 値が A であるとき、客観評価から定義される等価 Q 値は B[dB]である。これに対応する受聴 MOS 値は C である。これらの関係を予め求めておけば、等価 Q 値を介さずに、MOS-LQO 値から直接受聴 MOS 値を決定することができる。



付図Ⅲ-1 / JJ-201.11 < MOS-LQO 値と受聴 MOS 値の関係の定式化 >

本方法は、本質的には等価 Q 値により品質を表現することと等価であるが、その評価値が MOS 尺度で得られる点異なる。

このようにレファレンス信号に対する評価特性により評価値を正規化するためには、レファレンス信号に対する客観／主観評価特性の対応関係と評価対象信号に対するそれとが一致していることが前提となる。低ビットレート符号化による劣化やパケット損失による劣化は、雑音付加音声である MNR 信号の劣化とは聴感上の性質が異なるため、客観／主観評価特性の対応関係が異なることが危惧されるが、文献[106]における試験結果は、POLQA の前身の客観品質評価法である PESQ がこれらを統一的に評価可能であることを示している。

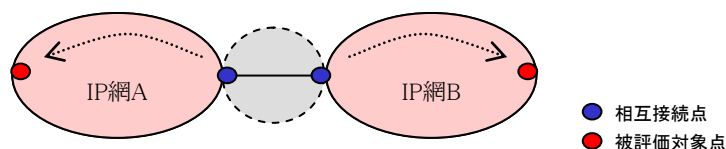
本手法で与えられる受聴 MOS 値は、一定の音声レベルで音声を受聴した場合の品質に関する指標であり、音量・遅延・エコー等の要因を含めた総合評価値としての会話 MOS とは異なる。

付録Ⅳ（参考）相互接続環境における通話品質評価に関する留意事項

（本付録は参考資料であり、仕様ではない。）

本付録は、異なる IP 網を相互接続する環境において、通話品質を評価する際のレファレンス接続系設定の考え方及び品質評価時の留意事項を示す。

Ⅳ.1 相互接続点が単一の場合



付図Ⅳ－１／JJ-201.11 <相互接続点が単一の IP 網の接続モデル>

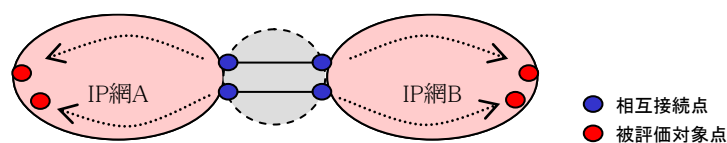
① 被評価対象点の設定

IP 網間の相互接続点を考慮して、相互接続環境下におけるレファレンス接続系（標準系／限界系）を定める。具体的には、各々の相互接続点から各々の IP 網に対して、距離（相互接続点からの最遠区間など）、ルータ段数、回線速度、トラフィック量などを考慮し、遅延やパケットロスが限界値に近い地点を限界系の被評価対象点に設定する。また、標準系の被評価対象点も同様な考え方で、遅延やパケットロスが平均的な地点に設定する。

② レファレンス接続系の設定

各々の IP 網に対して、上記①で定めた限界系及び標準系の被評価対象点の2点間をそれぞれ相互接続環境におけるレファレンス接続系（限界系及び標準系）とする。

Ⅳ.2 相互接続点が複数存在する場合



付図Ⅳ－２／JJ-201.11 <相互接続点が複数存在する IP 網の接続モデル>

① 被評価対象点の設定

相互接続点が単一の場合と同じ方法を用いて、相互接続点毎に限界系及び標準系の被評価対象点を設定する。

② レファレンス接続系の設定

各々の IP 網に対して、上記①で定めた限界系及び標準系の被評価対象点の全ての組み合わせ（但し、限界系の被評価対象点と標準系の被評価対象点の組み合わせは除く）をレファレンス接続系（限界系及び標準系）とする。

IV.3 品質評価時の留意事項

- ① 品質評価項目は、基本的に本標準に規定された内容に準拠すること。
- ② 相互接続点間及び相互接続対象 IP 網内の時間的変動要素を考慮し（トラフィック変動特性などの差異を踏まえ、適切な評価日、評価時間を設定する等）、95%値を算出できるよう十分なサンプル測定を行うこと。
- ③ 相互接続点にメディアゲートウェイのバックトゥバック接続が入る場合は、コーデック変換による影響を考慮すること。
- ④ 評価を行う際の基準となる端末は、各々の IP 網であらかじめ設定し、評価条件を明示すること。

参考文献

- [101] 浅谷編著, 「通信ネットワークの品質設計」, 電子情報通信学会発行, 1993.
- [102] ITU-T SG12 SQ-46.95R3, "Subjective Test Plan for Characterization of an 8 kbit/s Speech Codec," Sep. 1995.
- [103] ITU-T SG15 TD-66, "Results and Preliminary Analyses of Experiments to Characterize the Subjective Performance of Proposed Rec. G.729," Nov. 1995.
- [104] "Multi-Lingual Speech Database for Telephonometry 1994," NTTアドバンステクノロジー株式会社, 1994年6月
- [105] ITU-T Contribution COM12-74, "Review of Validation Tests for Objective Speech Quality Measures (NTT)," May 1996.
- [106] A. Takahashi, "Performance Evaluation of Objective Quality Measure, ITU-T Recommendation P.862 PESQ," Proc. of CQR2002, May 2002.