

# 技術調査資料

## － PESQ/POLQA差分調査 －

第1.0版

2025年7月1日



一般社団法人 情報通信ネットワーク産業協会

ユーザネットワークシステム委員会

電話・TA WG

## 0. 改訂履歴

| 版数  | 改訂日付     | 改定箇所 | 改定内容 | 備考 |
|-----|----------|------|------|----|
| 1.0 | 2025.7.1 | —    | 初版制定 |    |
|     |          |      |      |    |

# 1. 本資料について

昨今、音質基準に使用されていたPESQが、POLQAでの測定に見直されている。NTT東西がメタル回線の代わりに携帯電話回線を活用した「ワイヤレス固定電話」サービスの提供を2024年4月1日に始めたが、音質基準はPOLQAを指定している。

ITU-Tにて元勧告のP.862シリーズ（PESQ関連の勧告）の廃棄が合意され、2024年1月に標準化勧告から外され、POLQAが標準化されている。TTC標準化（JJ-201.01/JJ-201.11からもPESQの記載を削除）も進んでいる。

本資料では、PESQとPOLQAの差分をまとめ、POLQAの重要性について考察する。

## 2. POLQAとPESQの概要

**POLQA (Perceptual Objective Listening Quality Analysis) : ITU-T P.863**  
**PESQ (Perceptual Evaluation of Speech Quality) : ITU-T P.862**

POLQA及びPESQは、参照信号と劣化信号を比較し、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。

### ■ POLQA

従来の電話帯域音声域（300Hz～3.4kHz）と 4G/LTE、VoIP等で用いられる超広帯域音声（50Hz～14kHz）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。

PESQで測定不可だった超広帯域/モバイル端末向けコーデックの評価が可能。

### ■ PESQ

電話帯域音声とアプリケーションに応じて50Hz～7kHz の帯域幅の受聴品質を推定する客観品質評価技術である。

対象とする品質要因は、音声符号歪、パケット/セル損失歪などである。

### 3. POLQAとPESQの差分

| 項目             | PESQ                                       | POLQA                                             | 備考                                                           |
|----------------|--------------------------------------------|---------------------------------------------------|--------------------------------------------------------------|
| 音声帯域           | △狭帯域対応：50Hz-3.4KHz<br>△広帯域対応：50Hz-7KHz     | ○超広帯域対応：<br>50Hz-14KHz                            |                                                              |
| MOS（主観測定）との相関性 | △低い<br>相関性を高めるためP.862.1 と P.862.2 が後から規格化※ | ○高い<br>MOS-LQO（リスニング品質）としてMOS 値をスコアリングする様に設計されている | ※低スコア（3以下）になると、MOS よりも高めのスコアを出す傾向があり                         |
| 音声品質改善機能※との親和性 | △スコアが低く出る傾向                                | ○正しいスコアを出す様に設計                                    | ※コーデックの無音圧縮機能が動作、VoIP の遅延変動、電話音声の周波数変動、ノイズサプレッサ、AGC、エコーキャンセラ |
| 広帯域コーデックとの親和性  | △MOS より低い値が出る                              | ○MOS との相関性が高い正確な値が出る                              | SLIK ,AMR-WB, EVRC-WB, iLBC, AMR+, AAC, iSAC等                |

詳細は別紙 

## 4. POLQAを使用するメリット

- ◆対応音声帯域  
狭帯域～超広帯域対応：50Hz-14KHz  
PESQでは実施できない7kHz以上の音声評価が可能
- ◆MOS（主観測定）との相関性  
高スコア（5）から低スコア（1）までMOS との相関性が高い
- ◆パケット損失およびCELPタイプ符号化のパケット損失隠蔽処理  
G711に対しても評価
- ◆音声品質改善機能（ノイズサプレッサ、A G C、エコーキャンセラ）の親和性  
聴取品質への影響を正しく反映
- ◆評価適切な項目がPESQより増加、不適切な項目がPESQから減少  
適切な項目3項目増、不適切な項目1項目減、妥当性未確認5項目減



## 5. 考察（まとめ）

音声品質を評価するには、POLQA を使用することが最良の選択であると判断する。

- ・ MOSとの相関性が高くなる様に設計されており、主観測定との差が出にくい。
- ・ 様々な音声品質改善機能（ノイズサプレッサ、A G C、エコーキャンセラ）による聴取品質への影響を正しく反映可能。
- ・ 50Hz-14KHzまでの広い帯域をカバーしており、タイムワーピングコーデックに対して正しいスコアが算出可能



## <参考資料>

# 参考 1. 評価が適切な項目 (1/3)



○ : 適切    – : 不適切

| 項目                                  | POLQA | PESQ | 項目の内容(意味合い)                                                                                                                                                                                |
|-------------------------------------|-------|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Codecへの音声入力レベル                      | ○     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途                                                                                                                                                                 |
| 伝送路エラー                              | ○     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途                                                                                                                                                                 |
| パケット損失およびCELPタイプ符号化のパケット損失隠蔽処理(PLC) | ○     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>G.729、G.723.1、AMR等のCodecにてパケット損失 (0%,3%,5%)を与えてMOS評価し、PESQ,ITU-T P.863スコアとの相関を見る                                                                             |
| 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)  | ○     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>ITU-T P.863の要求仕様に含まれる要素と用途<br>G.723.1、G.726等                                                                                                                 |
| Codecのタンデム接続                        | ○     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>異なるCodecの組み合わせで評価する                                                                                                                                          |
| 送話側周囲騒音                             | ○     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>【P.862より】ノイズ録音は、修正IRS送信特性と似た適切なフィルタを通過させ（これは、ハンドセット・フィルタにより大きく減衰される車のノイズのような低周波信号にとって特に重要である）、次に試験用の希望のレベルにレベル整合される。                                         |
| 受聴試験における遅延変動                        | ○     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>音声信号の遅延が時間とともにどのように変動するかを評価する項目。<br>具体的には、音声信号が送信されてから受信されるまでの遅延が一定でない場合、その変動が音声品質にどのような影響を与えるかを測定する。<br>遅延変動が大きいと、音声途切れたり、エコーが発生したりすることがあり、これが通話品質の低下につながる。 |

# 参考 1. 評価が適切な項目 (2/3)



○ : 適切    - : 不適切

| 項目                      | POLQA | PESQ | 項目の内容(意味合い)                                                                                                                                                                                                                                                                                                                                                                                            |
|-------------------------|-------|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 短区間の時間構造歪               | ○     | ○    | <p>ITU-T P.863の要求仕様に含まれる要素と用途<br/> 【P.863 App.II「アルゴリズムのトレーニングと検証の為に実施される超広帯域コンテキストでの主観テストの説明」<br/> II.3 Design rules of test conditions for full-scale mandatory testsより】At least 15% variable delay (VoIP, video-telephony)/time warping</p> <p>音声信号の短い時間区間における時間的な歪みを評価する。<br/> 具体的には、音声信号が短い時間内でどのように変化するかを測定し、その変化が音声品質に与える影響を評価する。<br/> 例えば、音声信号が急激に変化する場合、その変化がリスナーにとってどの程度不自然に感じられるかを評価する。</p> |
| 長区間の時間構造歪               | ○     | ○    | <p>ITU-T P.863の要求仕様に含まれる要素と用途<br/> 音声信号の長い時間区間における時間的な歪みを評価する。<br/> 音声信号が長い時間内でどのように変化するかを測定し、その変化が音声品質に与える影響を評価する。<br/> 例えば、音声信号が長時間にわたって一定の遅延を持つ場合、その遅延がリスナーにとってどの程度不自然に感じられるかを評価する。</p> <p>評価対象の時間スケール：短区間の時間構造歪はミリ秒単位の短い時間スケールで評価するのに対し、長区間の時間構造歪は秒単位の長い時間スケールで評価する。<br/> 影響の範囲：短区間の歪みは音声の瞬間的な品質に影響を与えるのに対し、長区間の歪みは全体的な音声品質や会話の流れに影響を与える。</p>                                                      |
| 時間的クリッピング<br>及び振幅クリッピング | ○     | -    | <p>ITU-T P.863の要求仕様に含まれる要素と用途<br/> .POLQAでは、音声の途切れ（時間的クリッピング）と歪み（振幅クリッピング）をより正確に検出できる。</p> <p>時間的クリッピングは、音声信号の一部が時間的に切り取られる現象を指す。これは、ネットワークの遅延やバッファリングの問題などによって発生することがある。時間的クリッピングが発生すると、音声途切れ途切れになり、リスナーにとって非常に不自然に感じられる。</p> <p>振幅クリッピングは、音声信号の振幅が一定の閾値を超えた場合に、その超えた部分が切り取られる現象を指す。これは、音声信号が過大入力された場合や、音声処理装置の設定が不適切な場合に発生する。振幅クリッピングが発生すると、音声歪んで聞こえ、リスナーにとって不快な音質になる。</p>                           |

# 参考 1. 評価が適切な項目 (3/3)



○：適切   －：不適切

| 項目                               | POLQA | PESQ | 項目の内容(意味合い)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|----------------------------------|-------|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 帯域幅制限やスペクトラム整形を含む線形歪(平坦でない周波数特性) | ○     | -    | <p>ITU-T P.863の要求仕様に含まれる要素と用途<br/>           帯域制限したテスト信号による評価<br/>           試験条件は以下の通り。</p> <ul style="list-style-type: none"> <li>・超広帯域（50 Hzから14,000 Hz）の音声信号のうち、10%以上が帯域制限されている</li> <li>・広帯域（300 Hzから3,400 Hz）の音声信号のうち、30%以上が帯域制限されている</li> <li>・狭帯域（300 Hzから3,400 Hz）の音声信号のうち、30%以上が帯域制限されている</li> </ul> <p>帯域幅制限は、音声信号の周波数帯域が制限される現象を指す。これは、音声コーデックや伝送路の特性によって発生することがある。帯域幅が制限されると、特定の周波数成分が失われ、音声の明瞭度や自然さが低下する。</p> <p>スペクトラム整形は、音声信号の周波数スペクトラムが特定の形状に整えられる現象を指す。これは、音声処理装置やコーデックの設計によって意図的に行われることがある。スペクトラム整形が適切でない場合、音声の自然さや品質が損なわれることがある。</p> <p>線形歪は、音声信号の周波数特性が平坦でない場合に発生する歪みを指す。これは、音声信号が伝送される過程で特定の周波数成分が強調されたり減衰されたりすることによって発生する。線形歪が大きいと、音声のバランスが崩れ、リスナーにとって不自然に感じられることがある。</p> <p>【補足】<br/>           スペクトラム：フーリエ変換を用いて音声波形を周波数成分に分解したもの。音声信号が占める周波数帯域のこと。<br/>           線形：入力信号と出力信号の間における振幅と位相（信号の時間的な位置）の比例関係を指す。</p> |
| 周波数特性                            | ○     | -    | <p>ITU-T P.863の要求仕様に含まれる要素と用途<br/>           POLQAでは、音声の品質（特定の周波数が強調されたり減衰されたりしていないか）をより正確に検出できる。</p> <p>音声信号の周波数応答がどのように変化するかを評価するもの。</p> <p>周波数特性は、音声信号が異なる周波数でどのように伝送されるかを示す。理想的には、音声信号は全ての周波数で均等に伝送されるべきだが、実際には特定の周波数が強調されたり減衰されたりすることがある。これが「平坦でない周波数特性」と呼ばれる現象。</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

## 参考 2. 評価が適切でない項目



○ : 適切でない    - : 適切

| 項目               | POLQA | PESQ | 項目の内容(意味合い)                                                                                   |
|------------------|-------|------|-----------------------------------------------------------------------------------------------|
| 受聴音量レベル          | -     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>ITU-T P.863ではフルバンドに対して53～78dB(A)(SPL)で設計されているので、これ以外のレベルでの精度を評価 |
| 音量損失             | -     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途                                                                    |
| 会話試験における<br>遅延劣化 | ○     | ○    | ITU-T P.863はこれらの変数とともに使用することを意図していない                                                          |
| 送話者エコー           | ○     | ○    | ITU-T P.863はこれらの変数とともに使用することを意図していない                                                          |
| 側音               | ○     | ○    | ITU-T P.863はこれらの変数とともに使用することを意図していない<br>側音レベルを変えて評価                                           |
| 受話側の周囲雑音         | ○     | -    | ITU-T P.863はこれらの変数とともに使用することを意図していない<br>周囲騒音レベルを変えて評価                                         |

## 参考3. 妥当性未確認な項目（1/2）



○：未確認－：確認済

| 項目                                          | POLQA | PESQ | 項目の内容(意味合い)                                                                                                                                                                                                                                              |
|---------------------------------------------|-------|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PCMタイプCodecの<br>パケット損失および<br>パケット損失隠蔽<br>処理 | -     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>G.711 App.IについてNo.3と同様に評価する<br>PESQは話頭の時間的クリッピング（特に単語の欠落などが生じた場合）に対して主観評価より過敏になる傾向がある。逆に、話中や話尾の時間的クリッピングに対しては主観評価より感度が悪い。                                                                                                  |
| 時間的クリッピング                                   | -     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>POLQAでは、評価が適切な項目となり、音声の途切れをより正確に検出できる。<br>【P.863 App.II「アルゴリズムのトレーニングと検証の為に実施される超広帯域コンテキストでの主観テストの説明」II.2 Anchor conditionsより】<br>試験条件は以下の通り。<br>・パケット損失率が2%と20%の2種類がある<br>・1つのパケットサイズは20ms<br>・パケット隠蔽技術を使用しない             |
| 振幅クリッピング                                    | -     | ○    | ITU-T P.863の要求仕様に含まれる要素と用途<br>POLQAでは、評価が適切な項目となり、音声の歪みをより正確に検出できる。<br>【P.863 App.II「アルゴリズムのトレーニングと検証の為に実施される超広帯域コンテキストでの主観テストの説明」II.3 Design rules of test conditions for full-scale mandatory testsより】<br>試験条件は以下の通り。<br>・音声信号の約5%がクリップされている（切り取られている） |
| 話者依存性                                       | ○     | ○    | ITU-T P.863が現在検証されていない要素、技術、およびアプリケーション<br>テスト信号として男女数名を用意する？ P.501                                                                                                                                                                                      |
| 複数話者による同時<br>発話                             | ○     | ○    | ITU-T P.863が現在検証されていない要素、技術、およびアプリケーション                                                                                                                                                                                                                  |
| 符号器と複合器の<br>ビットレートミス<br>マッチ                 | ○     | ○    | ITU-T P.863が現在検証されていない要素、技術、およびアプリケーション                                                                                                                                                                                                                  |

## 参考 3. 妥当性未確認な項目 (2/2)



○ : 未確認 – : 確認済

| 項目                    | POLQA | PESQ | 項目の内容(意味合い)                                                                                        |
|-----------------------|-------|------|----------------------------------------------------------------------------------------------------|
| 網情報信号に対する<br>符合化性能    | ○     | ○    | ITU-T P.863が現在検証されていない要素、技術、およびアプリケーション<br>IRやCARを指すのか？                                            |
| 擬似音声信号に対す<br>る符号化性能   | ○     | ○    | ITU-T P.863が現在検証されていない要素、技術、およびアプリケーション<br>トーカーを指すのか？ITU-T P.50のような音声をさすのか？                        |
| 楽音に対する符号化<br>性能       | ○     | ○    | ITU-T P.863が現在検証されていない要素、技術、およびアプリケーション<br>Gシリーズのコーデックは音声を対象としているため、CELPに近いが、mp3,AAC等のコーデックを用いた場合  |
| 受話者エコー                | ○     | ○    | ITU-T P.863が現在検証されていない要素、技術、およびアプリケーション                                                            |
| エコーキャンセラに<br>よる品質劣化   | -     | ○    | ラインエコーキャンセラのon/off比較<br>TTC標準「IP電話の通話品質評価法」の「付録Ⅱ（参考）エコーリタンロスの測定法」はエコーキャンセラによる影<br>響も加味した評価法となっている。 |
| 雑音抑制アルゴリズム<br>による品質劣化 | -     | ○    | ノイズキャンセラのon/off比較<br>TTC標準「IP電話の通話品質評価法」のPOLQAにおいて評価が適切な項目の「アプリケーション」に「雑音抑制」が<br>ある。               |

## 参考 4 . PESQ詳細説明

PSQMと別の品質推定手法であるPAMS (Perceptual Analysis Measurement System) を統合した PESQ (Perceptual Evaluation of Speech Quality) がITU-T勧告P.862として2001年に標準化。PESQは、電話などでの音声の受聴品質の客観評価を行うためのアルゴリズムである。原音となる参照信号と評価系からの出力である被試験信号との比較によって評価を行い、評価の適用対象は、音声符号化歪、伝送遅延の変動、パケットまたはセルの損失、時間クリッピング、伝送チャンネルのエラー等である。PESQ による評価結果は、絶対範疇尺度法による受聴品質を推定するスコアである。

また、PESQ は以下の2段階の処理により評価値を推定する。

- 知覚モデリング : リファレンス/評価対象音声を人間の聴覚心理モデルに基き周波数領域でパラメータ化
- 認知モデリング : パラメータ化した値から雑音や歪みなどの妨害値を計算し MOS 値にマッピング

ITU-T P.862は、300-3400Hz の電話帯域の音声信号の評価を対象とする。

同様のアルゴリズムを用い 7kHz の広帯域音声を対象とする勧告としてITU-T P.862.2 がある。

## 参考 5. POLQAとPESQの差分(1/2)

### ■ 音声帯域

PESQ P.862, P.862.1 狭帯域対応：50Hz-3.8KHz

PESQ P.862.2 広帯域対応：50Hz-7KHz

POLQA P.863 狭帯域～超広帯域対応：50Hz-14KHz

POLQA には狭帯域モードと超広帯域モードの二つのモードがある。

狭帯域モードはPESQ との互換性を保つために用意されたものである。

超広帯域モードはPOLQA のみに対応しており、超広帯域でのスコアの最大値は5 となる。

広帯域、狭帯域ではそれに準じてスコアが低く出る様に設計されている。

広帯域測定はPOLQA では超広帯域モードを使用して行う。

### ■ MOS（主観測定）との相関性

PESQ P.862 は主観測定であるMOS との相関性が欠けており、相関性を高めるためP.862.1 とP.862.2 の二つの規格が後から規格化された。

一方POLQA は初めからMOS-LQO（リスニング品質）としてMOS 値をスコアリングする様に設計されている。

PESQ ：低スコア（3 以下）になると、主観測定のMOS よりも高めのスコアを出す傾向があり、相関性が低い。

POLQA ：高スコア（5）から低スコア（1）までMOS との相関性が高くなる



### ■ 音声品質改善機能（ノイズサプレッサ、AGC、エコーキャンセラ）との親和性

PESQ : GSM/WCDMA で使われているAMR-WB コーデックの無音圧縮機能が動作している場合や、VoIP の遅延変動、電話音声の周波数変動、ノイズサプレッサ、AGC、エコーキャンセラなどが働いている場合、スコアが低く出る傾向がある。

POLQA : 上記の様な音質改善機能が動作している環境でもより正しいスコアを出す様に設計されている。

### ■ 広帯域コーデックとの親和性

PESQ : タイムワーピング機能を持つ最新のコーデックを使った環境ではスコアが主観測定 MOS より低い値が出てくる。（タイムワーピング：周波数成分を変えずに圧縮する技術）SLIK ,AMR-WB, EVRC-WB, iLBC, AMR+, AAC, iSACなどの広帯域コーデックを使った環境では、PESQ は正しいスコアを出せない。

音声品質改善機能による聴取品質への影響を正しく反映できず、また最新のコーデックによる違いを正確にスコアリングできない。

POLQA : 上記の様なコーデックに対してMOS との相関性が高い正確なスコアが出る。特にVoice Over LTE で使用するコーデックを比較試験する場合に、PESQ では様々なこのため最新のコーデックによる音声品質を評価するにはPOLQA を使うことが最良の選択であると言える。

## ユーザーネットワークシステム委員会 電話・TA WG

|     |      |                    |
|-----|------|--------------------|
| 主査  | 葉迫悠介 | 沖電気工業(株)           |
| 委員  | 亦野康弘 | 岩崎通信機(株)           |
| //  | 塩谷洋翔 | 岩崎通信機(株)           |
| //  | 岡崎浩幸 | (株)ナカヨ             |
| //  | 畠山永吉 | (株)ナカヨ             |
| //  | 木下博之 | N E Cプラットフォームズ(株)  |
| //  | 横山達彦 | N E Cプラットフォームズ(株)  |
| //  | 鈴木政光 | 沖電気工業(株)           |
| //  | 星川拓哉 | 富士通(株)             |
| 事務局 | 笹野 潤 | (一社)情報通信ネットワーク産業協会 |
| //  | 谷 敏樹 | (一社)情報通信ネットワーク産業協会 |

技術調査資料「PESQ/POLQA差分調査」第1.0版

発行日 第1.0版：2025年7月1日

発行人 ユーザネットワークシステム委員会 電話・TA WG

発行元 一般社団法人 情報通信ネットワーク産業協会（CIAJ）

〒103-0026 東京都中央区日本橋兜町2 1 番 7 号

HF日本橋兜町ビルディング6階

URL： <https://www.ciaj.or.jp/>

TEL：03-5962-3452

FAX：03-5962-3455

本「技術調査資料「PESQ/POLQA差分調査」第1.0版」に関し、  
全部又は一部を無断で転載を行うことを禁ずる。

