



- 2700 Series
- APx585 family
- APx525 family
- APx515

Perceptual Audio Testing

By David Mathew

はじめに

本テクニカルノートでは、Audio Precision の APx500 シリーズのアナライザで使用可能な Perceptual Audio 測定の概要と、APx500 ソフトウェアで使用方法について説明します。豊富な追加情報が ITU-T（国際電気通信連合の電気通信標準化部門）および OPTICOM GmbH（本書末尾の参考文献セクションを参照）から入手できます。POLQA と PESQ の要約された OPTICOM 社のマニュアルは、本テクニカルノートに含まれており、Audio Precision Web サイトから入手できます。

Perceptual audio 測定

知覚的な測定（Perceptual Methods）は、データ圧縮コーデックを使用する通信チャネルおよびネットワークにおけるオーディオ品質の評価や、従来の客観的なオーディオ測定が有効ではない場合の評価を可能にするために開発されました。プライマリモデルは、オーディオ品質の主観的な意見を提供する人間の聴取者を使用し、Mean Opinion Score（MOS）に集約されます。

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

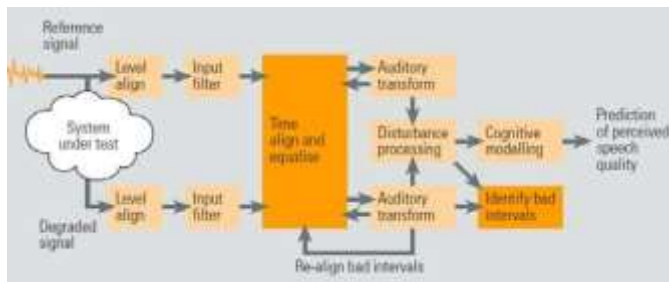
ITU およびその他の団体は、人の聴取者によって報告されたものと同等の MOS 結果と、原因分析のための他の結果を提供することを目的として、機械試験を使用して通信チャネルのスピーチ品質を知覚的に評価するための手法と推奨を開発することに長年協力してきました。これらのモデルには、PSQM、PESQ、POLQA などの技術が含まれています。関連する方法である PEAQ は、音楽などのフルレンジのオーディオ信号の伝送チャネルを評価するように設計されています。

Audio Precision では、最新の 2 つの推奨事項 POLQA と PESQ を使用して、APx500 シリーズアナライザで知覚的な音声評価を提供します。PESQ は 2001 年に ITU-T 勧告として承認され、現在も広く使用されています。2011 年に ITU-T によって承認された POLQA は、PESQ の後継機種であり、数多くの改良点があります。

POLQA と PESQ を検討すべきユーザー

Audio Precision ユーザーは、Bluetooth チャネルのオーディオを特徴付けるための POLQA または PESQ の必須ツールを見つけることができます。特に、HFP および HSP Bluetooth プロファイルで使用される CVSD コーデックおよび mSBC コーデックによって生成される劣化音声信号は、電話ネットワーク、VoIP（Voice over Internet Protocol）接続などを通じた損失の多いコーデックに依存する音声と同様に、知覚オーディオテスト方法を使用して最も評価されるアプリケーションです。

PESQ ITU-T P.862 (2001)

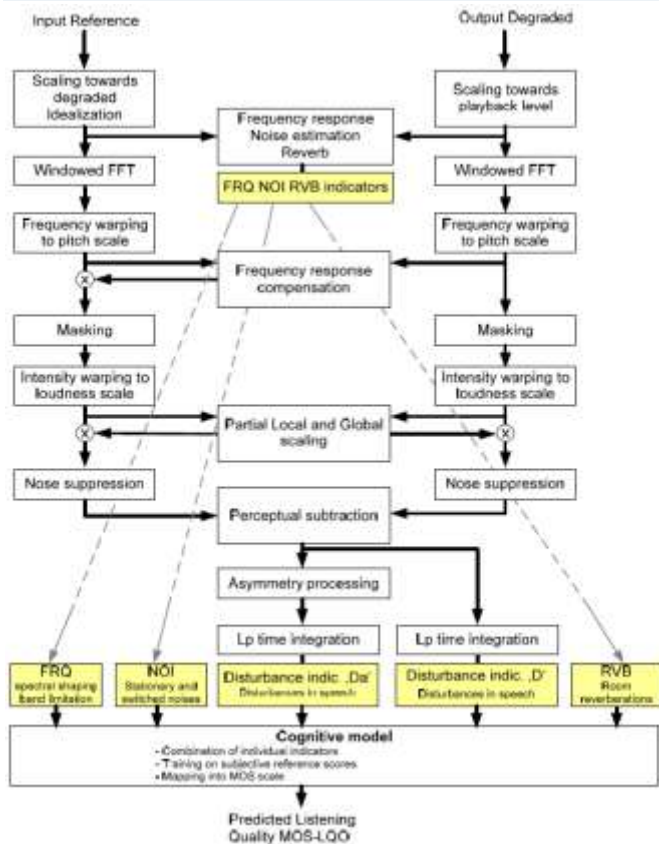


PESQ block diagram, courtesy of OPTICOM GmbH.

PESQ（音声品質の知覚評価）アルゴリズムは、ITU-T 勧告 P.862（2001）で定義されています。PESQ は、MOS の結果と原因分析に有用な様々な結果を提供します。

PESQ は、300～3400Hz の帯域幅を有する狭帯域モードと、アプリケーションに応じて 50～7000Hz の帯域幅を有する広帯域モードを提供します。リファレンス信号と劣化信号は処理されて比較され、MOS 結果が生成されます。

POLQA ITU-T P.863 (2011)



POLQA block diagram, courtesy of OPTICOM GmbH.

POLQA（Perceptual Objective Listening Quality Analysis）アルゴリズムは、ITU-T 勧告 P.863（2011）で定義されています。POLQA は PESQ（ITU-T Rec. P.862）の後継機種です。POLQA は MOS の結果と原因分析に有用な様々な結果を提供します。

POLQA は P.862 / PESQ モデルの弱点を回避し、より高い帯域幅のオーディオ信号を処理するように拡張されています。また、多くの遅延を伴う信号や時間依存する信号を取り扱えるように改善されています。

P.862 と同じく、POLQA は、共通電話帯域（300～3400Hz）での測定をサポートしますが、さらに広帯域および超広帯域音声信号（50～14000Hz）で HD 音声の評価するための第 2 の動作モードを備えています。

POLQA はマウスシミュレータとイヤースミュレータを備えた人工頭部によって音響的に記録された音声信号の評価もターゲットにしています。

POLQA と PESQ の比較

- ITU-T による評価で、狭帯域と広帯域の両方のモードで POLQA が PESQ よりもパフォーマンスが優れることが示されています。
- POLQA は、より正確に（主観的なリスニングと比較して）レベルの差を評価します。
- POLQA は、最新の HD 音声コーデックとネットワークにスーパーワイドバンドモードに対応します。
- POLQA では、PESQ がサポートしていなかった、人間の知覚に沿った線形周波数ひずみの重み付けができます。
- PESQ はコンバータ間のクロック差に非常に影響を受けやすく、POLQA はクロック差やその他の時間スケリング効果を補償します。
- POLQA は、音響測定を完全にサポートしており、頭部および胴部シミュレータ、マウスシミュレータ、イヤースミュレータによるハンドセットとトランスデューサのテストをサポートしています。

音声サンプル

POLQA と PESQ は、人間の音声サンプルをソース信号として使用します。サンプルはシステムまたは試験対象（DUT）を介して再生され、DUT の出力は POLQA または PESQ アルゴリズムを使用して元のスピーチサンプルと比較されます。スピーチサンプルは正しく形成され、男性と女性の両方の声を含む多数の異なるサンプルがアルゴリズムによってテストされ、平均化されなければなりません。

Audio Precision では、APx POLQA および PESQ ソリューションのセットアップとデモンストレーションの例として、いくつかの音声サンプルのライブラリを提供しています。実際のテストでは、ITU-T、IEEE または他の認定された標準機関から音声サンプルファイルを取得するか、ご使用の言語とアプリケーションに最適化された音声サンプルの独自のライブラリを作成することを推奨します。本ドキュメントの末尾に、知覚的な音声測定のための音声サンプルの作成に関する情報へのリンクや、参考文献の音声サンプルファイルへのリンクがあります。

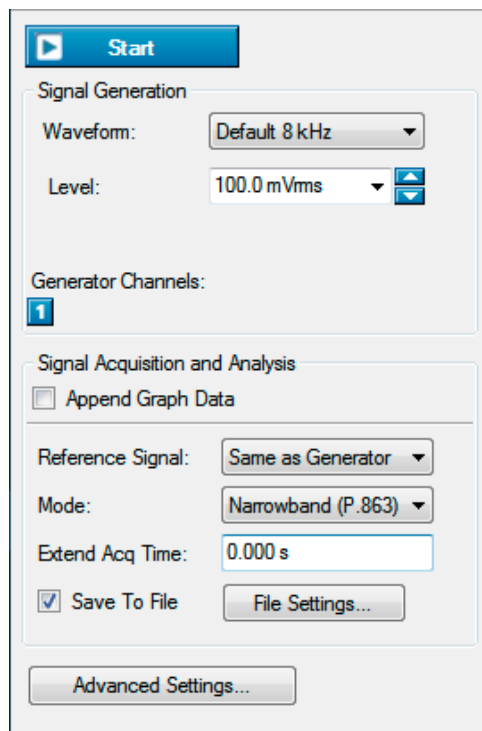
APx での使用

APx POLQA と PESQ の実装は似ており、各方法には一連の取得で平均した結果を提供するものと、単一の取得で結果を提供するものの 2 つの APx 測定値があります。

注：POLQA または PESQ アルゴリズムのすべての結果が APx ユーザーインターフェイス上に表示されるわけではありません。

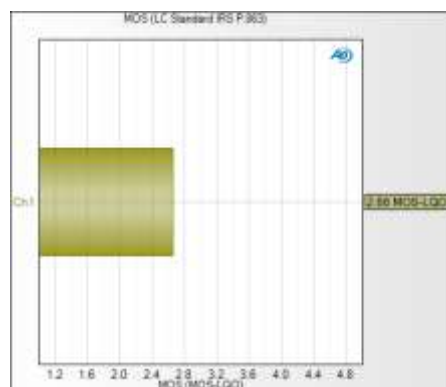
POLQA と PESQ は、APx アナライザのオプションの測定技術であり、動作するにはそれぞれソフトウェアキーが必要です。すべての POLQA および PESQ 測定は、1 つまたは複数の適切に形成された音声サンプルを必要とします。歪みのない音声サンプルは、参照ファイルとしての APx に指定され、リファレンス波形として表示することができます。これらは、システムまたは被試験デバイス（DUT）を通じた際に劣化し、APx 上で Acquisition Waveform として表示されます。

POLQA measurement (one acquisition)



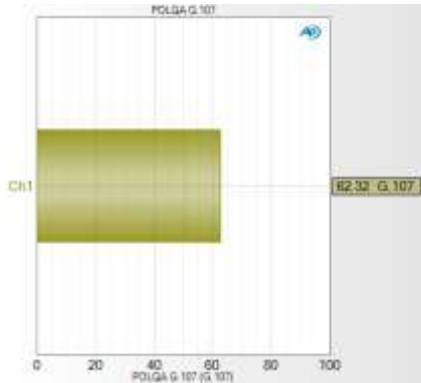
一回の測定サイクルから結果を表示します。これらの結果は、セットアップおよび原因分析に有用です。

MOS



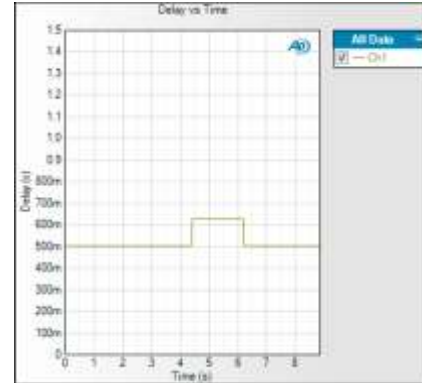
アルゴリズムから MOS スコアを表示します。Mode 設定に応じて、LC 標準 IRS P.863 または LC SWIDE H P.863 にすることができます。

POLQA G.107



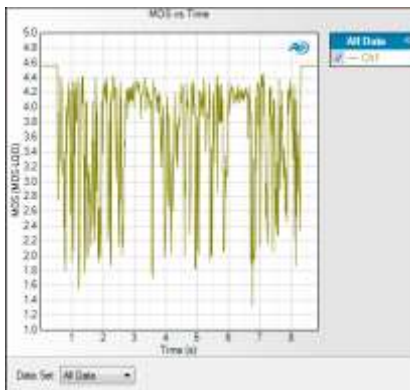
G.107（ITU-T トランスミッション定格モデル）に準拠した、MOS 変換前の POLQA の生データを表示します。Narrowband Mode でのみ使用可能です。

Delay vs. Time



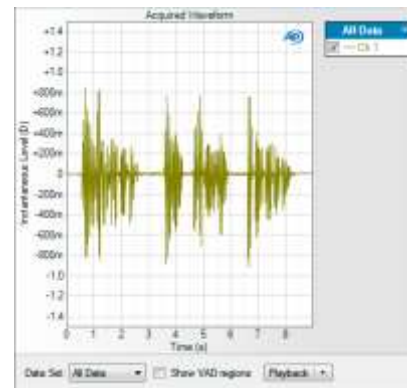
測定中のソース信号と取得信号間の遅延の遷移を表示します。原因分析に有用です。

MOS vs. Time



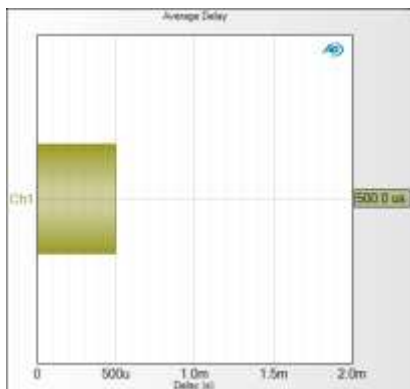
測定中の MOS の遷移を表示します

Acquired Waveform



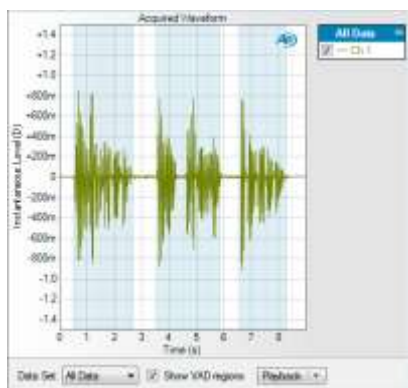
取得した（劣化した）音声サンプル波形を示します。グラフの下の **Playback** ボタンで、取得した波形を PC で再生できます。**Save To File** が設定されている場合、取得した信号は WAV ファイルとしてディスクに保存されます。原因分析に有用です。

Average Delay



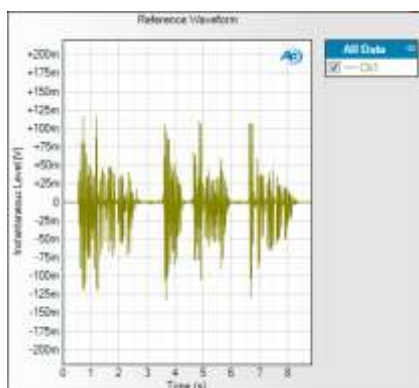
各チャンネルについて測定全体の平均信号遅延を表示します。原因分析に有用です。

Acquired Waveform VAD



Show VAD regions が設定されている場合、アルゴリズムの Voice Activity Detector が音声を検出し分析を行った領域が強調表示されます。

Reference Waveform



リファレンス音声サンプル波形を示します。この結果は、Generator レベルの設定に基づいて、APx の出力における実際の値を表すようにスケールされます。原因分析に有用です。

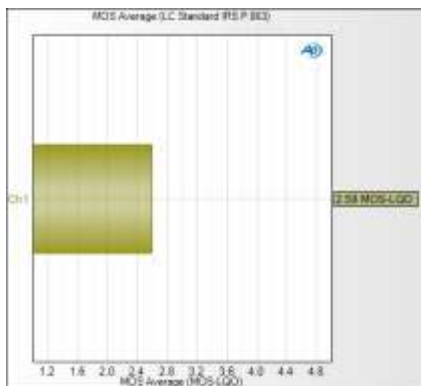
POLQA (Averaged) measurement

	File Name	Hz
1	Female1_1st_Set_16k .wav	16k
2	Female1_2nd_Set_16k .wav	16k
3	Female1_3rd_Set_16k .wav	16k
4	Male1_1st_Set_16k .wav	16k
5	Male1_2nd_Set_16k .wav	16k
6	Male1_3rd_Set_16k .wav	16k

ITU-T P.863 では、実際の通信システムをテストするための POLQA MOS の結果が、内容が変化する音声サンプルを使用して、男性と女性の音声混在する多数の取得の平均であることを推奨しています。

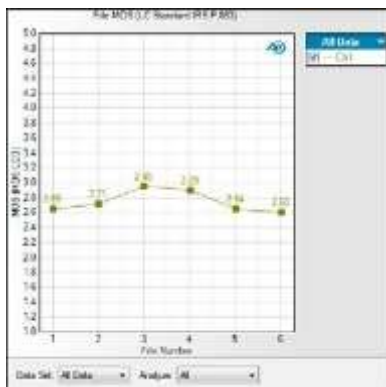
Reference Files リストには、最大 64 個のファイルをロードできます。**Start** をクリックすると、最初のファイルがテスト中のシステムから再生され、POLQA アルゴリズムで取得され、処理され、MOS 結果が提供されます。これはリスト上の全てのファイルで繰り返され、MOS 結果の実行平均が結果に表示されます。

POLQA (Averaged): MOS Average



この測定結果は P.863 勧告を満たしており、実際の通信システムをテストする際に最適の APx POLQA 結果です。モード設定に応じて、LC 標準 IRS P.863 または LC SWIDE H P.863 にすることができます。

POLQA (Averaged): File MOS



各音声サンプルの個々の MOS 結果を Reference Files リストに表示します。この結果は原因分析に役立ちます。

Other POLQA settings

Mode

POLQA 測定は、以下の 3 つのモードのいずれかに設定できます。

- Narrowband (P.863)
- Wideband (P.863)
- SuperWideband (P.863)

Extend Acquisition Time

想定よりも MOS の結果が悪い場合は、取得した波形を確認して下さい。DUT による遅延により切り捨てられている場合があります。この場合、波形全体が取得されるように測定時間を延長することができます。

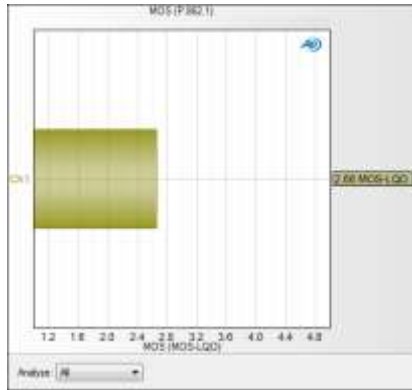
Reference Signal Selection

オープンループテストでは、リファレンス信号を指定する必要があります。クローズドループテストでは、リファレンスファイルは通常はジェネレータ出力になります。ただし、EQ または他の方法で変更されたファイルを使う場合は、ここでは EQ のような代替ファイルを指定することができます。

PESQ measurement (one acquisition)

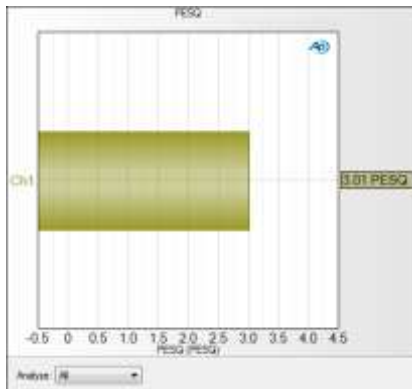
一回の測定サイクルから結果を表示します。これらの結果は、セットアップおよび原因分析に有用です。

MOS



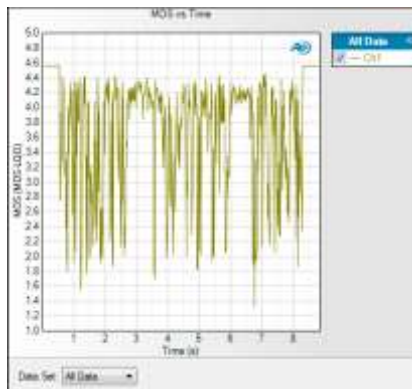
アルゴリズムから MOS スコアを表示します。**Mode** 設定に応じて、P.862.1 または P.862.2 にすることができます。

PESQ



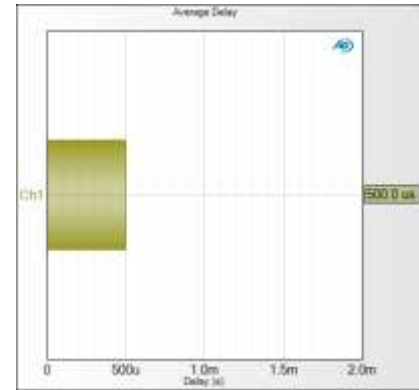
MOS 変換前の POLQA の生データを表示します。**Narrowband Mode** でのみ使用可能です。

MOS vs. Time



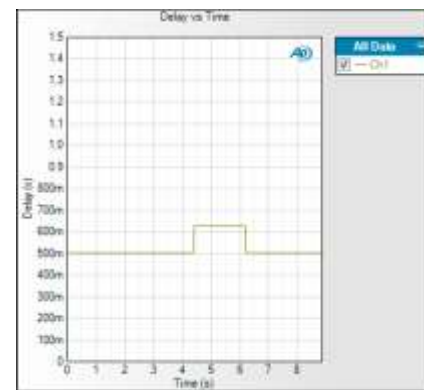
測定中の MOS の遷移を表示します

Average Delay



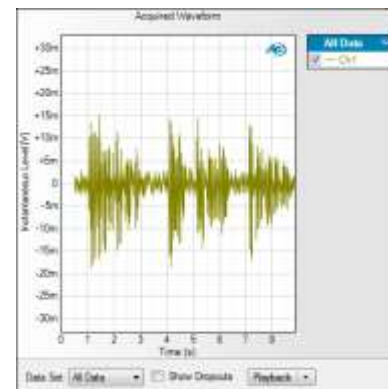
各チャンネルについて測定全体の平均信号遅延を表示します。原因分析に有用です。

Delay vs. Time



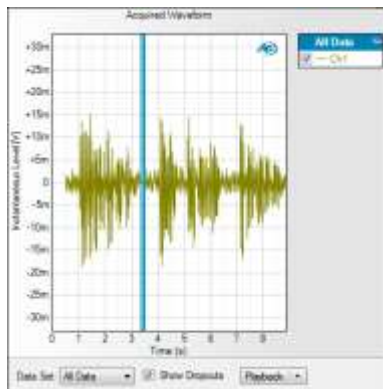
測定中のソース信号と取得信号間の遅延の遷移を表示します。原因分析に有用です。

Acquired Waveform



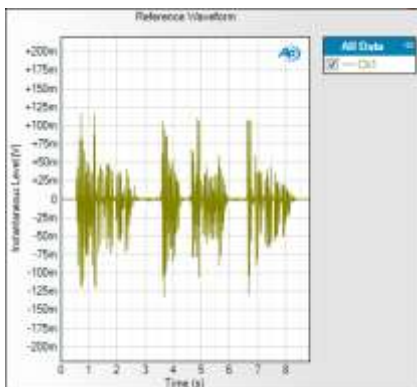
取得した（劣化した）音声サンプル波形を示します。グラフの下の **Playback** ボタンで、取得した波形を PC で再生できます。**Save To File** が設定されている場合、取得した信号は WAV ファイルとしてディスクに保存されます。原因分析に有用です。

Acquired Waveform Dropout Display



Show Dropouts を設定すると、ドロップアウトが強調表示されます。ドロップアウトは、リファレンス波形には信号があり、取得された波形には信号がない領域です。

Reference Waveform



リファレンス音声サンプル波形を示します。この結果は、Generator レベルの設定に基づいて、APx の出力における実際の値を表すようにスケールされます。原因分析に有用です。

PESQ (Averaged)

	File Name	Hz
1	Female2_1st_Set_16k .wav	16k
2	Female2_2nd_Set_16k .wav	16k
3	Female2_3rd_Set_16k .wav	16k
4	Female2_4th_Set_16k .wav	16k
5	Male1_3rd_Set_16k .wav	16k
6	Male2_3rd_Set_16k .wav	16k

ITU-T P.862 では、実際の通信システムをテストするための PESQ MOS の結果が、内容が変化する音声サンプルを使用して、男性と女性の音声混在する多数の取得の平均であることを推奨しています。

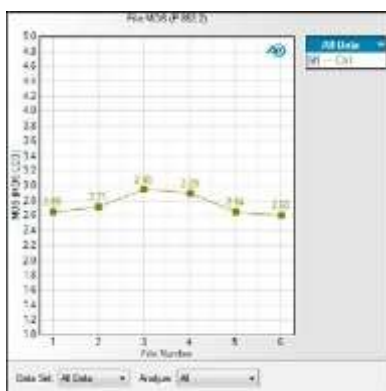
Reference Files リストには、最大 64 個のファイルをロードできます。**Start** をクリックすると、最初のファイルがテスト中のシステムから再生され、PESQ アルゴリズムで取得され、処理され、MOS 結果が提供されます。これはリスト上の全てのファイルで繰り返され、MOS 結果の実行平均が結果に表示されます。

PESQ (Averaged): MOS Average



この測定結果は、P.862 勧告に準じており、実際の通信システムをテストする際に最適の APx PESQ 結果です。PESQ では分析を **All**、**Active Speech** または **Silence** で行うことができます。

PESQ (Averaged): File MOS



各音声サンプルの個々の MOS 結果を Reference Files リストに表示します。この結果は原因分析に役立ちます。

その他の PESQ の設定

Mode

- Narrowband (P.862)
- Wideband (P.862)

Extend Acquisition Time

想定よりも MOS の結果が悪い場合は、取得した波形を確認して下さい。DUT による遅延により切り捨てられている場合があります。この場合、波形全体が取得されるように測定時間を延長することができます。

Reference Signal Selection

オープンループテストでは、リファレンス信号を指定する必要があります。クローズドループテストでは、リファレンスファイルは通常はジェネレータ出力になります。ただし、EQ または他の方法で変更されたファイルを使う場合は、ここでは EQ のような代替ファイルを指定することができます。

Analyze

PESQ では、Signal, Silence, 信号全体(All)のアクティブな音声部分を分析するセレクトがあります

ライセンスについて

PESQ は、OPTICOM GmbH, Psytechnics Limited により、TNO Telecom、British Telecommunications の PSQM の後継製品として開発され、2001 年に ITU-T が勧告 P.862 として標準化しました。

POLQA は PESQ のアップグレード版として OPTICOM、SwissQual、TNO によって開発され、ITU-T によって 2011 年に勧告 P.863 として標準化されました。

これらの測定に使用されるアルゴリズムは著作権によって保護されています。Audio Precision の POLQA および PESQ 測定の各処理は、これらの技術の唯一のライセンスエージェントとして機能する OPTICOM のライセンスを受けています。APx アナライザで POLQA または PESQ を使用するための iButton キーは、Audio Precision から購入できます。

POLQA チャンネル数について

POLQA は OPTICOM 社からチャンネルごとにライセンスされています。Audio Precision POLQA-2 iButton キーは、オーディオチャンネル 1 と 2 で POLQA を有効にします。チャンネルを追加する場合は別途購入して下さい。

PESQ iButton キーは、接続された APx アナライザのすべてのチャンネルを有効にします。

References

PESQ:

<http://www.pesq.org/>

POLQA:

<http://www.opticom.de/technology/polqa.php>

OPTICOM:

<http://www.opticom.de/index.php>

ITU-T

PESQ and POLQA standards at ITU-T:

<http://www.itu.int/en/ITU-T/publications/Pages/recs.aspx>

Navigate to P series, then to P.862 (PESQ) or P.863 (POLQA). Most documents and evaluation packages are available to the public. At least one document requires ITU membership and an ITU-T “TIES” account.

Speech Samples

ITU-T

<http://www.itu.int/en/ITU-T/publications/Pages/recs.aspx>

Information on creating speech samples for perceptual audio measurements can be found in the ITU-T Recommendation P.830, section 7.2. Navigate to P series, then to P.830.

Additionally, speech sample files are available at

<http://www.itu.int/net/itu-t/sigdb/genaudio/AudioForm-g.aspx?val=10000501>

IEEE

Recommended Practice for Speech Quality Measurements

<http://ieeexplore.ieee.org/xpl/tocresult.jsp?reload=true&isnumber=26069>

Navigate to document download for *IEEE Recommended Practice for Speech Quality Measurements*. This document points to the Harvard Sentences, 720 phonetically balanced phrases for English speakers. These and other speech samples can be found at the Open Speech Repository.

Open Speech Repository

http://www.voiptroubleshooter.com/open_speech/index.html