psytechnics
behind every word

PESQ Tools

White Paper

Prepared by:
Psytechnics Limited

23 Museum Street
Ipswich, Suffolk
United Kingdom
IP1 1HN

t: +44 (0) 1473 261 800
f: +44 (0) 1473 261 880
e: info@psytechnics.com

August 2003

Table of Contents

What you should do now

If you are a voice network operator, voice network designer or a voice device designer then it is time to think again about how to optimize your hardware for speech quality.

Q: You've got a score for the quality of voice that your device or network can deliver, but how can you influence this quality?

A: Use PESQ Tools to present the information to help you manage and develop your network or device.

PESQ Tools enables voice quality assessment to be taken into the domain of network and device diagnostics.

Overview

PESQ Tools is designed to give essential diagnostic information alongside the overall PESQ quality score.

The main users of PESQ Tools are likely to be:

Operators
Need to know the quality of voice being delivered by their networks.

Designers
Need to determine the quality of voice that their networks or devices will deliver.

It is essential that users in both these categories have access to the voice quality measure. However, in order to know what to do with this measure, PESQ Tools provides further information which will help both to pinpoint problems and aid development that will generate real improvements in voice quality.

PESQ Tools is a powerful set of analysis tools that allow the user to asses and diagnose a range of network and device problems related to voice. PESQ Tools works by communicating directly with the core of PESQ in order to create and provide the desired diagnostic information.

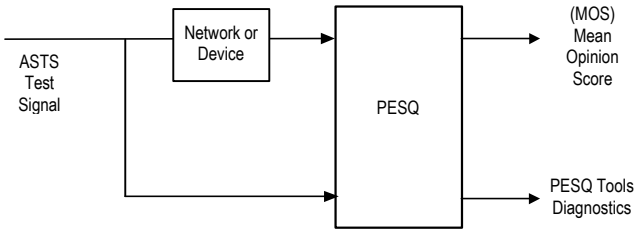

Figure 1 – An example of voice quality monitoring for a chosen network or device

Benefits of using PESQ Tools

In 2000 Psytechnics co-developed PESQ as the world standard for voice quality assessment. PESQ provides an accurate voice quality score for networks, systems and devices. This score is based on a subjective scale and can be mapped onto the results from thousands of subjective mean opinion score (MOS) results worldwide. If you simply need a voice quality score then PESQ is the ideal tool.

While PESQ gives an accurate rating of voice quality, the extra diagnostic information available from PESQ Tools can be valuable to network operators and designers as well as device designers. PESQ Tools provides detailed information, based on real voice passing through a network, to help pinpoint problems and optimize configurations. PESQ Tools is the ideal step up from standard PESQ.

Applications for PESQ Tools

For **Analog Networks**, PESQ Tools allows issues with noise, levels and filtering to be identified quickly and efficiently. These can have a strong effect on voice transmission quality. Once fixes and modifications have been implemented, PESQ Tools enables operators and engineers to quickly verify that they work.

Wireless Networks present many unique challenges, but still require tight control of the factors affecting speech quality. PESQ Tools enables operators and engineers to asses and analyze the time varying parameters of voice quality within their networks. PESQ Tools allows issues such as voice quality at cell boundaries and during handovers to be analyzed as well as the effect of network and handset VQE.

VoIP success relies heavily on the ability to control network and gateway/IP phone properties that affect voice quality. PESQ Tools offers the ability to identify issues such as jitter arising from variations in delay over time, packet losses that have an impact on the speech quality, and VAD and comfort noise issues. This allows users to realize the advantages of VoIP.

While **Codec and Algorithm** designers and developers hope to improve the performance of their devices and systems, they need the relevant information to do this. PESQ Tools provides this information to allow a development of their products targeted on voice quality. Information available from PESQ Tools can be used in conjunction with the frame-by-frame PESQ score to identify and solve problem conditions. Once designs or modifications have been implemented, PESQ Tools allows engineers to quickly verify their work.

Features of PESQ Tools

The advanced diagnostics and analysis features of PESQ tools are outlined below.

It is important to note that certain features are listed under specific headings, but that these features may also be useful in other circumstances.

Advanced analysis of speech quality

PESQ score frame-by-frame
Frame-by-frame score provides a rough guide to the location and relative magnitude of distortions.

PESQ-LQ
PESQ-LQ is a mapping that presents the PESQ score on the same scale as subjective MOS.

E-Model Ie
PESQ-Ie is a mapping that transforms the PESQ score to an Ie value suitable for use with the ITU/ETSI E-model.

Delay statistics
Frame-by-frame delay from PESQ is the best way of tracking how delay varies during the signal due to effects in the system under test.

Error indicators
PESQ computes two disturbance parameters that describe the amount and distribution of audible errors:

- Symmetric disturbance
- Asymmetric disturbance

A more detailed view of distortions is given by the error surface, which shows how audible disturbances are in time and frequency.

The following are some examples of errors that these measures can help to diagnose.

- Front-end clipping
- Muting
- Packet loss
- Addition of background noise
- Coding distortion
- Bit or frame errors

Linear spectrum analysis

Bark and linear frequency spectra
These provide views of the frequency content of the speech and noise in the two signals, for example showing the spectrum of additive noise, VAD comfort noise, or noise reduction.

Impulse response estimate
PESQ Tools makes it possible to measure the impulse response of analog or acoustic connections, or echo and sidetone paths.

Transfer function estimation
Four different transfer function estimates are provided. These show the effect of filtering in the terminals or network, which can be used to diagnose problems in analog connections, terminals and echo cancellers.

Speech analysis diagnostics

Spectrogram and LPC spectrogram
Frame-by-frame signal spectra are calculated on a linear frequency scale. These measures can be used to compare the spectrum of different signals and compare speech and noise, for example illustrating comfort noise mis-match.

Pitch identification
This provides an estimate of the vocal pitch during voiced sections of the two signals, showing the impact on speech of distortions such as robotisation due to very low bit-rate coding or packet loss.

Signal excitation
Excitation approximates the speech before the signal spectrum is modified by the effects of the vocal tract and lip radiation, and is used in most speech codecs. This output gives a view of processing of codecs on the signal excitation.

Voiced / unvoiced decision
Vowel-like speech sound such as “ah” are voiced, while noise-like sounds such as “sh” are unvoiced. This PESQ Tools diagnostic shows the effect of the device on the voicing of speech.

PESQ Tools also provides a number of extra diagnostics specifically for the speech:

- Frequency of first four formants, f1 – f4, in Hertz
- Frame-by-frame power envelope
- Probability of speech.

Analysis of level variations

- Active Speech Level
- Noise Level
- Speech and Noise Gain
- SNR and SNR change

These features show the speech and noise levels in the signals, which vary, for example due to gain in the system, noise addition, or signal processing. SNR is the ratio between the speech and noise levels.

Utterance Level and ALC gain
Levels calculated for each speech utterance show the effect of time-varying gain, such as automatic level control or dynamic noise reduction.

DC offset
This diagnoses offset problems in analog-to-digital conversion which can cause a drop in speech quality due to quantization or clipping.

Testing of VoIP

Frame-by-frame delay
The variation of delay during a measurement shows how jitter in the network and the processing of the jitter buffer affect the end-to-end audio delay.

Comfort noise and VAD analysis
PESQ Tools provides information on processing due to voice activity detectors (VADs), including the front and back-end clipping that they can cause, the hangover time, and comfort noise, allowing all of these behaviors to be analyzed.

Temporal clipping and muting
Several sets of clipping statistics are provided by PESQ Tools for identifying events of clipping/muting due to VADs and/or packet loss, and their averages across the whole signal.

Glossary

Active Speech Level	A measure of the absolute level of voice within a signal.
ALC	Automatic Level Control adaptively varies the amount of gain in a system to compensate for connections that are too quiet or too loud.
Bark	A perceptual frequency scale.
Comfort noise	A noise signal added by echo cancellers or VADs to make their operation less noticeable.
Condition	The particular network connection under test, including all processing such as codecs and errors due to packet loss or mobile distortions.
Degraded	Processed version of the reference speech file that is measured at the output of the system under test.
Error surface	Time-frequency image of distortion sign and amplitude. Positive error means noise that is added; negative error is due to signal components that have been lost.
ETSI	European Telecommunications Standards Institute
Gain	The amount of amplification or attenuation of the speech signal.
ITU	International Telecommunications Union
LPC	Linear Predictive Coding technique used in voice coding systems.
MOS	Mean Opinion Score, the average of votes for a group of subjects for the given network condition.
Noise Gain	The amount of amplification applied to background noise in a signal.
Noise Level	The absolute level of background noise in a signal.
PESQ	Perceptual Evaluation of Speech Quality, ITU-T recommendation P.862 - the standard measure of voice quality.
PESQ score	The standard quality measure returned by P.862 PESQ.
PESQ-Ie	E-model impairment factor Ie calculated from PESQ score (P.834).
PESQ-LQ	PESQ Listening Quality score is on the MOS scale (1-5) and is calculated from PESQ score.
Pitch	The instantaneous vocal pitch (fundamental frequency only) in voiced speech.
Reference	'Clean' original speech file injected into the system under test.
Spectrogram	Time-frequency image of the signal loudness..
Utterance	A section of speech, typically a continuous sentence.
Utterance Level	The signal level (loudness) of an utterance.
VAD	Voice Activity Detector, makes decisions on when there is voice and when there is silence, usually so that transmission is stopped during silent periods.
Voicing	Voiced parts of speech-like vowels are strongly pitched while unvoiced parts are noise-like.
VoIP	Voice over Internet Protocol allows voice to be sent over packet networks.
VQE	Voice Quality Enhancement devices reduce noise and echo in networks and/or terminals.

psytechnics
behind every word

PRO_PESQ00_XXX_ES0058_01