OPTICOM

PESQ – Perceptual Evaluation of Speech Quality

OPTICOM's PESQ (Perceptual Evaluation of Speech Quality) provides accurate and repeatable estimates of speech quality degradation occurring through e.g. a telephony network. It compares the audio signal input to a network with the corresponding (degraded) audio signal output from the network. P.862 is based on the earlier P.861 (PSQM) technology developed by KPN Research extended to include a time alignment model.

Being already widely used in active testing in both lab environments and in network test equipment, OPTICOM's PESQ has achieved a great reputation for its ease of use and its reliability of its results while lowering the costs for the measurements.

OPTICOM's PESQ is fully compliant to ITU-T P.862.x covering the applicability to narrow band as well as wide band audio signals.

OPTICOM who is the leading provider of signal based perceptual measurement technologies offers algorithms for voice, audio and video quality measurements.

OPTICOM Product Line:	
Voice/Audio Quality	
PESQ	ITU-T P.862
3SQM	ITU-T P.563
PSQM	ITU-T P.861
ECHO	
PEAQ	ITU-T BS.1387
Video Quality	
PEVQ	
Network Quality	
VQmon	

Principle

The degraded signal and the reference signal are individually level aligned and filtered with the transfer characteristics of a receiving device (handset). Basically two different filter functions – a narrow band (IRS) filter and a wide band filter – are available and may be chosen depending on the preferred application. After that the two signals are time aligned in order to compensate for small time shifts that can occur in e.g. Voice over IP networks due to delay and jitter changes and coding.
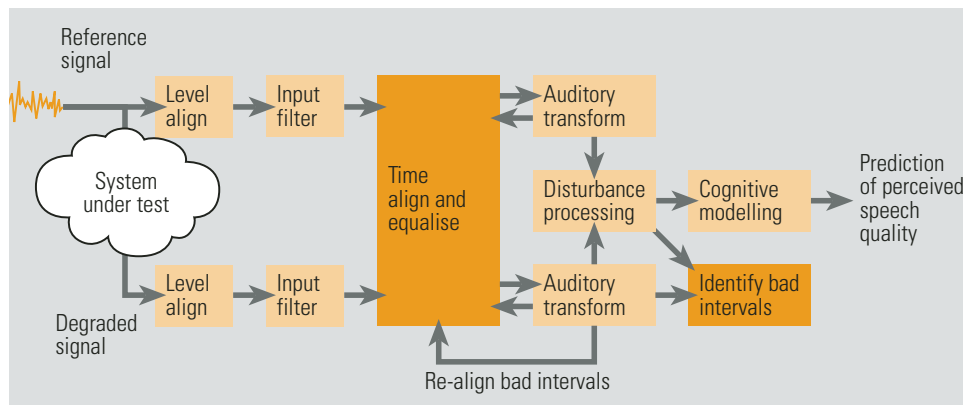
In order to account for the distortions that are actually perceived by a human listener the model transforms the two aligned and filtered signals from the time-amplitude domain into a frequency-loudness domain (auditory transform).

By subtracting the two signal representations an estimate of the audible differences is derived. The audible differences are accumulated over time while they are weighted differently depending on whether a distortion was added to the signal or if parts of the signal were missing after the transmission (cognitive model).

Finally after the analysis a single Mean Opinion Score (MOS) is generated. The MOS is commonly used to describe the voice quality on a scale from 1 (bad quality) to 5 (excellent quality).

Key Features:

- Perceptual analysis of degradations in voice signals
- Well established, repeatable results
- Output score correlates very well with subjective MOS
- Automatic time alignment capability
- Available as DLL/Library and Source

The Perceptual Quality Experts.

One of the major advantages of PESQ over other measurement approaches is that it does not need to make any assumptions on the network under test or the distortion types that may occur in a specific telephony test scenario. The only prerequisite employed within the model is the scientific knowledge on how human speech is produced and how humans perceive it.

PESQ already includes the effects of both packet level impairments (loss, jitter) and signal related impairments such as noise, clipping and distortions caused by coding processes. The impairment model, based on this knowledge is as generic as it can be and therefore independent from current and future telephony applications and networks.

Specifications

PESQ – Perceptual Evaluation of Speech Quality

Functionality

- ITU-T P.862.x Perceptual Evaluation of Speech Quality, with mapping to MOS scale
- Narrowband and Wideband operation

Input

- 16 bit linear audio sampled at 8kHz and 16kHz
- Input voice files 6 to 20 seconds in length

Output

- VAD measurements (front end clipping, hold over time)
- Drop out measurement
- Delay measurement and statistics (min, max, average, delay vs. time, histogram)
- MOS vs. time
- MOS for speech and silent parts separately
- Attenuation
- Speech level (total, active speech, silence, reference, test separately)
- Loudness (reference, test)
- AGC measurement
- R Factor (LQ)

Complexity

- P.862 is processed in about 10-20 times faster than real-time on a Pentium III 733MHz (44-125MFLOPs)
- Data memory: 5.2MB for file pair with 8s length
- Code memory: App. 200-400kb (depending on used features and optimizations)
- Floating point library required

Platforms

- Windows
- Linux
- Solaris

The Perceptual Quality Experts.