**P.563 – Single Sided Speech Quality Measure**

OPTICOM’s single-sided speech quality measure provides estimates of the speech quality degradation caused for example by telephony networks. Therefore the measurement model analyses the degraded voice signal output from a device or a network under test. The measure is based on models of the human vocal tract and the human perception of abnormalities in a voice signal.

OPTICOM’s single-sided speech quality measure includes and is fully compliant to ITU-T Recommendation P.563.

**Principle**

The evaluation of the voice quality with P.563 requires several steps.

Before the voice signal quality is assessed, it needs to be preprocessed. The important steps in the pre-process are: IRS receive filtering, speech level adjustment and separation in voice and non-voice parts via a Voice Activity Detection (VAD).

In the second step distortion and speech parameters are extracted from the voice signal parts. Three main functional analysis blocks that correspond to the main groups of distortion types of P.563 help classifying the signal and its distortions. The main distortion types are defined as: unnaturalness of speech, basic speech quality, robotic voice, unnatural voice like beeps, strong additive noise, background noise, low segmental SNR, interruptions, mutes and clipping.

Finally after the analysis a dominant distortion class is determined and 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).

One of the major advantages of P.563 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 it is perceived by humans. P.563 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 distortion model, based on this knowledge is as generic as it can be and therefore independent from current and future telephony applications and networks.

**Key Features:**

- Perceptual analysis of degradations in voice signals
- Non-intrusive, single ended analysis method
- Output score correlates well with subjective MOS
- Available as DLL/Library or Source

---

**Voice Quality Testing**

OPTICOM Product Line:

<table>
<thead>
<tr>
<th>Voice/Audio Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>PESQ ITU-T P.862</td>
</tr>
<tr>
<td>P.563 ITU-T P.563</td>
</tr>
<tr>
<td>PSQM ITU-T P.861</td>
</tr>
<tr>
<td>ECHO</td>
</tr>
<tr>
<td>PEAV ITU-T BS.1387</td>
</tr>
</tbody>
</table>

**Video Quality**

- PEVD

**Network Quality**

- VQmon

---

**Detected distortion**

- Domination
- Mapping to final quality estimate

**Voice Signal**

- Pre-process
- Unnatural Speech
- Noise Analysis
- Intermittent, Mutes...

**Single sided speech quality measure**

(based on P.563)

**Detection of dominant distortion**

**Mapping to final quality estimate**

**MOS-LQ0**
Specifications

P.563 – Single Sided Speech Quality Measure

Functionality

- ITU-T P.563 Perceptual Estimation of Speech Quality, with mapping to MOS scale

Input

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

Output

Descriptors for mutes and interruptions in the signal:
Duration of muted signal parts, Indicator for interruptions, Counter for sharp declines, Indicator for silenced signal parts.

Results of the noise analysis:
Estimated background noise level, Estimated segmental signal to noise ratio, Mean spectral level deviation and level range, Speech level, Noise level, Signal to noise ratio, Descriptor for variations of high frequencies, Spectral clarity indicator, Percentage of samples that were classified as local background noise, Mean energy of frames that were classified as local background noise, Indicator for the number of occurrences of severe local background noise, Percentage of samples that were classified as global background noise.

Descriptors for unnatural voice, e.g. by describing the vocal tract:
Indicator for validity of estimated vocal tract, Maximum section size of the first VTP tube, Averaged section size of the last vocal tract tube, Average of maximum VTP section sizes, Average of VTP back section sizes, Ratio of total voiced signal length over the total length of speech sections, Average of peak to mean difference of cross power of 2 consecutive voiced frames, Average of peak to mean ratio of cross power of 2 consecutive voiced frames, Number of repeated frames occurrences.

Average of energy sum of all repeated frames, Number of detected tones/beeps, Average of energy sum of all frames containing tones/beeps, Average sum of samples that contain beeps, Indicator for the amount of highly periodic signals, Absolute standard deviation of the cepstrum, Skewness of the cepstrum, Kurtois of the cepstrum, Kurtois of the LPCs, Skewness of the LPCs, Absolute Skewness of the LPCs.

Basic voice descriptors:
Speech level in dBov, Average of the pitch frequency, Level variations between sentences.

Basic speech quality metrics:
Aggregation of the asymmetric frame disturbances, Output of the psychoacoustic model.

Distortion classes found in the speech signal:
Distortion due to a high background noise level, muted (all zeros) parts or interruptions or sharp declines, Noise related to the signal envelope, Strongly reverberant voice signal due to e.g. band limitation in GSM networks, Basic voice quality indicator.

The finally estimated MOS-LQO value:
The MOS-LQO value according to the P.900 standard within a range from 1 (bad) to 5 (excellent), R-Factor (LQ) according to standard G.107 which is derived from final MOS.

Complexity

- Estimated MFLOPS required for real-time operation: 175.262
- Floating point library required
- Program Memory: 200...490Kb (depending on used features and optimizations)
- Data Memory: 4...5Mb per 8s speech file. The total amount is dependent on the voice file length

Platforms

- Windows
- Linux

About OPTICOM

With PSQM, PESQ and PEAQ, OPTICOM GmbH, the pioneer in perceptual quality testing has been providing four international world-class standards for voice and audio quality measurement since its foundation in 1995. With their new single-sided speech quality measure based on ITU-T P.563, a joint development with partners, the perceptual experts from Germany now presented their latest ITU standard. At its 10th anniversary, the presentation of the new PEVQ™ video measure leverages the company’s huge experience towards the multimedia testing domain.

Recognized an industry reference, OPTICOM’s OPERA voice/audio quality test tools are available to users worldwide. And while specialized on OEM customers in particular, the directory of OEM licensees today reads like the ‘Who-is-Who’ of the Telecoms industry. OPTICOM is a privately held company located in Erlangen, Germany.

Sales Contacts:

OPTICOM GmbH
Nägelsbachstraße 38
91052 Erlangen, GERMANY
Phone: +49-9131/53020-0
Fax: +49-9131/53020-20
info@opticom.de
www.opticom.de

North America:
Telchemy Inc.
Phone: +1-770-614-6944

JDSU - Acterna U.S.
Phone: +1-301 353 1560 2850

Europe, Latin America, Middle East & Africa, Asia Pacific, CIS Countries:

JDSU - Acterna Germany GmbH
Phone: +49-7121 86 2222

Through our distributor network, we are represented in more than 80 countries. To find your local sales office, please contact info@opticom.de