Perceptual wideband speech and audio quality measurement

Dr Antony Rix
Psytechnics Limited
Agenda

Background

Perceptual models
- BS.1387 PEAQ
- P.862 PESQ
- Scope
- Extension to wideband

Performance of wideband PESQ
- Results for speech
- Results for audio
- Next steps – discussion

AMR-WB case study
Psytechnics background

- Solutions for measuring/monitoring speech, audio, video quality
- Extensive subjective testing background
- Main products are objective quality models (software)
  - Intrusive (P.862 PESQ, …) – for testing
  - Non-intrusive (P.VTQ/psyVoIP, P.563 SEAM/NiQA, P.562 CCI) – for monitoring
- Experience in wideband in both subjective testing and objective models (PAMS, PESQ).
BS.1387 PEAQ

- High-quality audio model for small impairments
- Comparable with BS.1116 subjective tests

- General audio model, not designed or optimised for “wideband speech”
- Mobile/IP multimedia is at edge of or outside scope
- Some issues with accuracy (see BS.1387 for results).

- Not currently applicable to 16kHz wideband speech
P.862 PESQ

- Speech quality model for telephony applications
- Comparable with P.800 subjective tests
- Assumes listening through narrowband IRS handset
- Was not extensively tested on perceptual waveform codecs (e.g. MP3, AAC) or with non-speech signals

- Not currently applicable to 16kHz wideband speech or audio
P.862 PESQ – scope

Reference signal

Level align

Input filter

Auditory transform

Time align and equalise

System under test

Degraded signal

Level align

Input filter

Auditory transform

Disturbance processing

Cognitive modelling

Identify bad intervals

Prediction of perceived speech quality

Re-align bad intervals

Input filter

Level align

Auditory transform
P.862 PESQ – scope

Scope assumes narrowband telephone handset listening, and speech signals
Extending PESQ for wideband speech & audio

Modification proposed in COM12-D7:

Input filter replaced by 100Hz high-pass with 9dB additional gain. No other changes (e.g. same psychoacoustic model).
Use of WPESQ

• Select wideband mode whenever headphone listening is used
• Also operates at 8kHz sampling rate (same filter frequency response)
• Be careful about mixing narrowband and wideband PESQ – binaural headphone listening is more sensitive, so the results are different
• Reference signal should normally be full bandwidth
WPESQ results – speech

Eurescom P905 exp1
Multiple audio bandwidths

Eurescom P905 exp2a
8kHz conditions only

ETSI wideband workshop, 8-9 June 2004
Copyright (c) Psytechnics Limited, 2004.
WPESQ results – speech

Eurescom P905 exp2b
16kHz conditions only

BT AES experiment
Multiple audio bandwidths
WPESQ results – NTT

• Morioka & Takahashi have published an independent evaluation of wideband PESQ
  – Wideband results: 91.2% correlation
  – Main issue is slight offset between G.722.1 and other conditions – will be investigated further
  – Problem with analysis – used narrow-band PESQ for 8kHz (wideband headphone) conditions although WPESQ should be used for this.
  – This caused offset between 8kHz and 16kHz conditions
    • Wideband PESQ is more critical than narrowband
  – 8kHz and overall results not included here.
WPESQ results – audio

- New subjective test by Psytechnics using:
  - 8 audio signals representative of PC and mobile multimedia (advertisement, movies, news documentary, pop music, speech, sports), of duration 8-12sec
  - 20 conditions
  - Range of codecs (AAC, AMR, G.711, G.722, and direct)
  - Range of bandwidths (8, 11.025, 12, 16kHz sample rates)
  - Presented to subjects and model at 16kHz, mono
  - Wideband binaural free field equalised headphones at 76dB SPL
  - Bit-rates from 4.75-256kbit/s
WPESQ results – audio

PESQ performance, $R=0.952$ RMSE=0.278
WPESQ results – overall

<table>
<thead>
<tr>
<th>Test</th>
<th>R %</th>
</tr>
</thead>
<tbody>
<tr>
<td>P905 exp 1 (speech)</td>
<td>95.2</td>
</tr>
<tr>
<td>P905 exp 2a (speech)</td>
<td>98.1</td>
</tr>
<tr>
<td>P905 exp 2b (speech)</td>
<td>97.7</td>
</tr>
<tr>
<td>AES107 (speech)</td>
<td>94.9</td>
</tr>
<tr>
<td>NTT wideband results (speech)</td>
<td>91.2</td>
</tr>
<tr>
<td>Psytechnics multimedia (16kHz mono audio)</td>
<td>95.2</td>
</tr>
<tr>
<td>Overall mean</td>
<td>95.4</td>
</tr>
</tbody>
</table>
WPESQ discussion

- WPESQ shows excellent correlation with MOS, comparing favourably with narrowband PESQ.
- Explore issues identified in P905 exp1 and NTT test:
  - Bandwidth and context effect
  - G.722.1 codec
- Can be used for both wideband speech and 16kHz mono audio – e.g. mobile multimedia applications
- Mapping between WPESQ and subjective MOS is required (like P.862.1 MOS-LQO).
Case study – Validation of AMR-WB (G.722.2) floating-point codec

- Fixed-point AMR-WB codec had been approved; needed to validate non-bit-exact floating-point version
- Used WPESQ to compare speech quality of codecs over 1280 test cases.
  - Identified bug in fixed-point codec mode-switching
  - Showed bug was corrected in floating-point and modified fixed-point codecs
  - Found no significant difference in quality between (corrected) fixed-point and floating-point codecs.
  - Took just 2 days of processing and analysis.
Conclusions

• BS.1387 PEAQ and P.862 PESQ not originally designed for wideband speech quality measurement
• By changing PESQ to use an appropriate input filter, WPESQ is able to make accurate quality measurements of wideband speech and 16kHz audio
• WPESQ allows interesting new applications in wideband speech and 16kHz audio quality testing, such as codec development, multimedia quality
• Some issues with subjective tests remain to be explored and further testing is desirable.
References


Rix, A. W. and Hollier, M. P. *Perceptual speech quality assessment from narrowband telephony to wideband audio*. 107th AES Convention, New York, preprint 5018, September 1999.


Eurescom P905. *AQUAVIT - Assessment of Quality for Audio-Visual signals over Internet and UMTS*

Rix, A. W. et al. *Proposed modification to draft P.862 to allow PESQ to be used for quality assessment of wideband speech*. ITU-T COM12-D007, Feb 2001.


Barrett, P. A. and Rix, A. W. *Verification of floating-point implementation of AMR-WB using Wideband-PESQ*. 3GPP Tdoc S4 (02)0049r1 and S4 (02)0124, Feb 2002.
Dr Antony Rix
Psytechnics Limited
antony.rix@psytechnics.com
www.psytechnics.com