



Report Qvoice™ PESQ™ verification

Overview

Psytechnics conducted a detailed set of tests to verify Ascom's DSP implementation of PESQ in Qvoice.

PESQ (Perceptual Evaluation of Speech Quality) is the international standard for testing the speech quality of telecommunications networks, ITU-T Recommendation P.862, which Psytechnics co-developed.

This document describes:

- the need for conformance testing
- the data sets and methods that were used to conduct these tests
- the version of Qvoice that was tested
- other assumptions that were made
- the results of these tests.

It is concluded that the version of Qvoice that was tested conforms to P.862.

Conformance testing

PESQ (ITU-T P.862) is a software algorithm that is implemented in floating-point arithmetic. In general, floating-point arithmetic differs slightly between different compilers and hardware platforms. This can lead to differences in the results of software such as PESQ when it is ported to different platforms.

These differences are normally so small that the user should not notice them. The purpose of conformance testing is to make sure that this is the case.

P.862 sets out a set of tests to verify whether an implementation of PESQ conforms to P.862. In addition to the fixed data specified in P.862, more than 7,000 additional cases were also tested. These include conditions that are representative of almost all major network technologies that are available today. Conditions that probe the limits of the PESQ algorithm were also included, to check for problems or bugs in an implementation.

Data sets

Qvoice PESQ was tested on the following four data sets.

Test 1

P.862 Annex A conformance test 1b.

- 1,736 file pairs in five different languages.
- Data from ITU-T Supplement 23 to the P series recommendations.
- Data re-sampled to 8kHz sampling rate according to P.862 Annex A, using ITU Software Tools Library 2000 release 3.

Test 2

P.862 Annex A conformance test 2b.

- 40 file pairs in two different languages.
- Data supplied at 8kHz sampling rate with P.862 for testing variable delay conditions.

Test 3

P.862 Annex A conformance test 3, with Psytechnics speech database.

- This test is open-ended and allows users to supply their own test data.
- Approx. 7,000 file pairs in five different languages.
- Data from the Psytechnics speech database covering a very wide range of telephone network conditions.
- Where appropriate, data re-sampled to 8kHz sampling rate according to P.862 Annex A, using ITU Software Tools Library 2000 release 3.

Test 4

P.862 Annex A conformance test 3, with Psytechnics mobile network test scenarios database.

- 100 file pairs made up using speech-like test signals
- Data designed by Psytechnics to model drive-testing of networks using 2G and 3G GSM in a wide range of conditions.

Notes

Note that Qvoice PESQ operates at 8kHz sampling rate, so tests 1a and 2a of P.862 Annex A (January 2003) do not apply as these are at 16kHz sampling rate.

The conformance testing procedure was according to P.862 Annex A, January 2003. At the time of writing, this revision to P.862 had been approved by ITU-T Study Group 12 subject to review under AAP, which will not be completed until May 2003.

The January 2003 revision to P.862 extends the conformance tests that had been available in the previous revision (February 2001) of P.862. Only the conformance tests changed between these two revisions of P.862.

The main difference between these two revisions is that conformance test 1, using Supplement 23, was only available at 16kHz sampling rate in the February 2001 revision of P.862, and therefore did not apply to Qvoice PESQ.

The conclusions of the evaluation are exactly the same whether the February 2001 or January 2003 revisions of the P.862 conformance testing procedure are used.

Method	<p>The conformance testing process consists of the following steps:</p> <ul style="list-style-type: none"> • Process the data through a reference implementation of P.862 • Process the data through the PESQ implementation under test • Compare the PESQ scores • Test whether the absolute differences in PESQ score between the two implementations are within the limits set by the relevant conformance test in P.862. These limits are as follows:
Test 1	Difference may exceed 0.05 in not more than 2 file pairs (approx. 0.1% of cases). Difference may not exceed 0.1 in any case.
Test 2	Difference may exceed 0.05 in not more than 1 file pair (2.5% of cases). Difference may not exceed 0.5 in any case.
Test 3	Difference may exceed 0.05 in not more than 0.5% of cases (advisory); difference may exceed 0.05 in not more than 5% of cases (mandatory).
Test 4	As for test 3.

Product tested

The implementation of PESQ in Qvoice that was tested was prepared for this evaluation by Ascom and provided to Psytechnics as a DSP card and the software to drive it. The file versions were as follows:

FPGA	5.0.0
dspqv3.out	1.5

Important notes	This section sets out some important notes and assumptions that were made in this evaluation.
<i>Test data</i>	A large and wide-ranging dataset was used, in particular for tests 3 and 4. However, due to the nature of the numerical differences that are being tested, other data may give different results.
<i>PESQ version</i>	The conclusions of this report apply only to the PESQ implementation that was tested. We are satisfied from the processing load on the host PC that PESQ was executing on the DSP.
<i>Test signals</i>	<p>Tests 1–3 were based on natural speech material of between 6 and 10sec duration.</p> <p>Test 4 was based on concatenated speech material of a type suited for drive-testing. The test signal duration was 5sec, with 80% speech activity.</p> <p>The implementation of PESQ that was tested allowed files to be a maximum of 79978 samples (9.997sec) in duration for .wav files, or 80000 samples (10.000sec) for headerless files. A normal duration of signals for drive-testing is 5–8sec, so this is unlikely to be a problem for users.</p>
<i>Additional processing in Qvoice</i>	The Qvoice measurement system includes some other processing, such as transcoding to G.711 (for example, as part of the ISDN interface) and the electrical interface to the test mobile. These components should have minimal effect on the PESQ score, but were not included in this study.

Results This section sets out the results of this study.

- Test 1**
- One file pair (approx. 0.05% of cases) had a difference of 0.06, out of a total of 1736 file pairs. The other pairs all had difference ≤ 0.05 .
 - This passes the requirement for accuracy: *Difference may exceed 0.05 in not more than 2 file pairs (approx. 0.1% of cases). Difference may not exceed 0.1 in any case.*
- Test 2**
- All 40 file pairs had a difference ≤ 0.05 .
 - This passes the requirement for accuracy: *Difference may exceed 0.05 in not more than 1 file pair (2.5% of cases). Difference may not exceed 0.5 in any case.*
- Test 3**
- 6 file pairs (approx. 0.08% of cases) had a difference > 0.05 , out of a total of 6,949 file pairs. The other pairs all had difference ≤ 0.05 .
 - This passes the tighter (advisory) requirement for accuracy: *Difference may exceed 0.05 in not more than 0.5% of cases (advisory); difference may exceed 0.05 in not more than 5% of cases (mandatory).*
- Test 4**
- All 100 file pairs had a difference ≤ 0.05 .
 - This passes the tighter (advisory) requirement for accuracy: *Difference may exceed 0.05 in not more than 0.5% of cases (advisory); difference may exceed 0.05 in not more than 5% of cases (mandatory).*

Conclusions

This study has found that the implementation of PESQ that was tested conforms to ITU-T Recommendation P.862.

Contact details

For further information on this report, please contact:

Tim Beck

Senior Account Manager

Psytechnics Limited

Fraser House, 23 Museum Street, Ipswich IP1 1HN, UK

Tel. +44 (0) 1473 261 805

Mobile (Cellphone) +44 (0) 7734 558508

Switchboard +44 (0) 1473 261800

Fax +44 (0) 1473 261 880

E-mail tim.beck@psytechnics.com



Fraser House, 23 Museum Street, Ipswich, Suffolk, IP1 1HN, United Kingdom
Tel: 44 (0)1473 261800 Fax: 44 (0)1473 261880 Email: info@psytechnics.com
Web: www.psytechnics.com