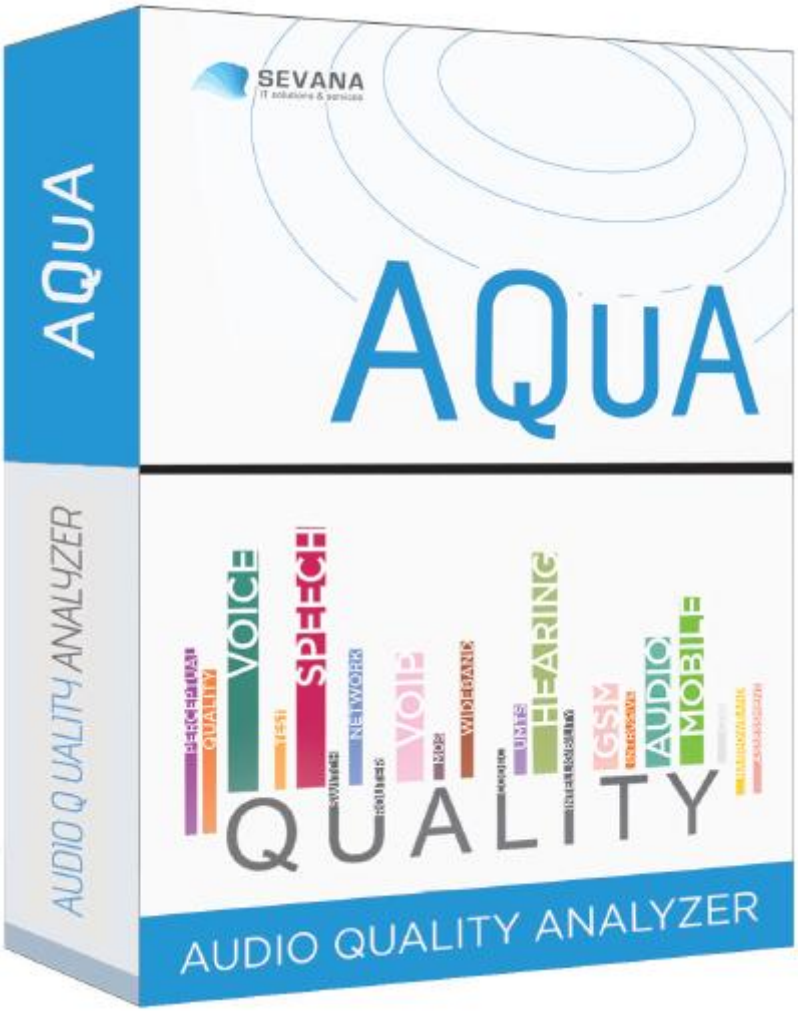
Sevana AQuA - Audio Quality Analyzer 6.x User Manual



Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

Contents

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Introduction

AQuA is a simple but powerful tool to provide perceptual voice quality testing and audio file comparison in terms of audio quality. This is the easiest way to compare two audio files and test voice quality loss between original and

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

degraded files. Besides this a most demanded functionality of the software can also test audio codecs and generate audio signals for voice quality testing.

AQuA gives a unique opportunity to design your own voice quality testing solution not being dependent on particular hardware and software. It is available as a library for Windows and Linux, portable to Java and mobile devices.

Functionality

Requirements

AQuA is capable to work with audio files represented in .wav or .pcm formats. Audio files should have the following parameters depending on the version one uses

AQuA Voice: 8kHz, 16 bit, Mono.

AQuA WB: 8kHz and up to 192kHz, 32 bit, Mono, Stereo

Generate test signals

Aqua allows generating test signals with the following parameters:

- Choosing voice type for synthesized sound: male or female
- Allows to define name of the synthesized sound or to run generation of full speech sounds distribution
- Allows choosing duration for synthesized sound equal in amount of samples
- Signals are built according to internal speech model
- Allows test sound signals generated as short, normal and long

Test voice codecs

- Allows output of codec speed performance indicator
- Allows testing any codec library(Windows only)

Compare wav files

- Allows intrusive testing of original wav file against degraded wav file
- Allows testing using internally generated audio test files
- Allows direct wav files comparison quality wise
- Allows measuring voice quality for any language

Testing parameters

AQuA supports the following parameters for voice quality testing:

- Choosing type of quality measurement: overall quality loss or voice naturalness
- Choosing filenames for original, test and generated audio files
- Define type of weight coefficients: uniform, linear or logarithmic
- Allows energy normalization
- Allows setting envelope smoothing level from 1 to 10

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

- Allows choosing source for original sound (external file or generated internally)
- Contains audio synchronization and voice activity detection
- Provides reasons for voice quality loss (quite unique feature on the market)
    - Duration distortion
    - Changes in signal spectrum
    - Distortion detection in low, medium and high frequency bands
- Provides voice quality feedback in:
    - Percentage of similarity
    - MOS-like value
    - PESQ-like value
- Enabling advanced psycho-acoustic model:
    - Psycho-acoustic filter
    - Normalization to loudness level at 1kHz
    - Spectrums transform into detectable range of loudness
- Audio synchronization – trimming silence in the beginning and end of the test file
- Adjusting ratio between calculation performance and quality score forecast accuracy

Scientific Background

The human ear is a non-linear system, which produces an effect named masking. Masking occurs on hearing a message against a noisy background or masking sounds.

As result of the research of the harmonic signal masking by narrow-band noise Zwiker has determined that the entire spectrum of audible frequencies could be divided into frequency groups or bands, recognizable by the human ear. Before Zwiker, Fletcher, who had named the selected frequency groups as critical bands of hearing, had drawn a similar conclusion.

Critical bands determined by Fletcher and Zwiker differ since the former has defined bands by means of masking with noise and the latter – from the relations of perceived loudness.

Sapozhkov has determined a critical band as “a band of frequency speech range, perceptible as a single whole”. In his earlier researches he even suggested that sound signals in a band could be substituted by an equivalent tone signal, but experiments did not confirm this assumption. Critical bands determined by Sapozhkov differ from those determined by Fletcher and Zwiker since Sapozhkov proceeded from the properties of speech signal.

Pokrovskij has also determined critical bands on the basis of speech signal properties. According to his definition the bands provide equal probability of finding formants in them.

The value of spectrum energy in bands can be used for different purposes; one of which is the sound signal quality estimation. However, using only one author’s critical bands (for example, Zwiker’s critical bands are used in prototype) does not allow getting an estimation objective enough, since they show only one of the aspects of perception or speech production. AQuA can determine energy in various critical bands as well as in logarithmic and resonator bands, that allows taking into consideration more properties of hearing and speech processing.

Taking into account that the bands determined by Pokrovskij and Sapozhkov are better for speech signal and not for sound signal, in general allows increasing the accuracy of estimation depending on its purpose.

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

AQuA utilized research results of the above mentioned scientists implementing different algorithms in one software solution. AQuA also has several advantages compared to other existing voice quality measurement software:

Besides critical bands new AquA implements a more advanced psycho-acoustic model, which consists of three layers:

- psy-filtering
- level normalization
- transform into detectable range

Psycho-acoustic model is based on dependencies obtained during experiments. The most complex phase is psy-filtering represented at Fig. 1.

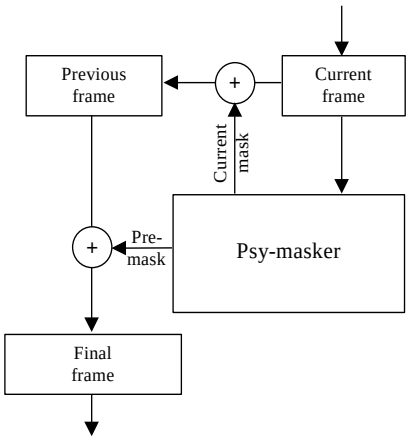


Fig. 1. General scheme of psy-filtering

Masking procedure includes the following sequence of actions:

1. hearing threshold processing
2. fluid level masking
3. spectrum separation into tones and noises
4. creating masks from tone components
5. creating masks from noise components
6. joining tone and noise mask components
7. joining current mask with post-mask
8. preparing post-mask for the next frame
9. creating mask for the previous frame

Hearing threshold corresponds to ear sensitivity towards intensity of sound energy, and minimal sound pressure that produces feeling of hearing is called hearing threshold. Threshold level depends on type of sound fluctuations and measuring conditions. One of possible options to detect hearing threshold (implemented in AQuA 5.x) is standardized in ISU/R-226.

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

Psycho-acoustic model implemented in AQuA 5.3 introduces the so-called range of detectable loudness, which is minimal change of signal amplitude detectable by a human ear. It's a well-known fact that depending on signal loudness level and frequency human perception varies from 2 up to 40%.

AQuA algorithms have certain advantages:

- it is universal since it allows measuring signals quality from various sources and processed in different ways;
- one can optimize quality estimation depending on the purpose:
  - for speed (for example, it is possible to receive rough estimation quickly);
  - by signal type (using different bands for speech signals and sound signals in general);
- resulting estimations correlate well with that of MOS;
- quality estimations received for speech signals can be translated in values of various scores of intelligibility.

AQuA Command Line parameters

AQuA Usage:

AquA-XX <license file> [options]

Print sounds' names: -h sndn

sndn - prints list of sounds names;

There are 54 sounds in the database

```
------------------------------------------------------------------------
Num   Name      Type                    Num   Name      Type
------------------------------------------------------------------------
000    a0 <<--- Vocal                   001    a1 <<--- Vocal
002    a2 <<--- Vocal                   003    a4 <<--- Vocal
004    e0 <<--- Vocal                   005    e1 <<--- Vocal
006    i0 <<--- Vocal                   007    i1 <<--- Vocal
008    i4 <<--- Vocal                   009    o0 <<--- Vocal
010    o1 <<--- Vocal                   011    o4 <<--- Vocal
012    u0 <<--- Vocal                   013    u1 <<--- Vocal
014    u4 <<--- Vocal                   015    y0 <<--- Vocal
016    y1 <<--- Vocal                   017     l <<--- Sonor
018    l' <<--- Sonor                   019     m <<--- Sonor
020    m' <<--- Sonor                   021     n <<--- Sonor
```


Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

```
022    n' <<--- Sonor                  023     j <<--- Sonor
024     v <<--- Noised                 025    v' <<--- Noised
026    zh <<--- Noised                 027     z <<--- Noised
028    z' <<--- Noised                 029     r <<--- Noised
030    r' <<--- Noised                 031     b <<--- Voiced Explosiv
032    b' <<--- Voiced Explosiv        033     g <<--- Voiced Explosiv
034    g' <<--- Voiced Explosiv        035     d <<--- Voiced Explosiv
036    d' <<--- Voiced Explosiv        037     f <<--- UnVoiced
038    f' <<--- UnVoiced               039     h <<--- UnVoiced
040    h' <<--- UnVoiced               041     s <<--- UnVoiced
042    s' <<--- UnVoiced               043    sh <<--- UnVoiced
044   sch <<--- UnVoiced               045     k <<--- Occlusive
046    k' <<--- Occlusive              047     p <<--- Occlusive
048    p' <<--- Occlusive              049     t <<--- Occlusive
050    t' <<--- Occlusive              051     c <<--- Occlusive
052    ch <<--- Occlusive              053     _ <<--- Silencer
-------------------------------------------------------------------------
```

Print samples of program usage: -h exam

exam - prints samples of program usage;

In order to test voice quality between original and reference files use the following set of parameters:

aqua-v.exe tst.lic -mode files -src file ORIGINAL_FILE -tstf REFERENCE_FILE

To test voice codec provided as a DLL library use the following set of parameters:

aqua-v.exe tst.lic -mode codec -clibf <DLL_LIBRARY_NAME> -src file <TEST_AUDIO_FILE>

e.g.

aqua-v.exe tst.lic -mode codec -clibf GSM610.dll -src file short.wav

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

Define program mode: -mode <mod>

Defines AQuA mode of operation. The following modes are available:

<mod> :

codec - codec testing mode;

files - audio file comparison mode;

folder - folder comparison mode;

generate - test signals generation mode

For example: *-mode codec*

aqua-v.exe tst.lic -mode files -src file ORIGINAL_FILE -tstf REFERENCE_FILE

aqua-v.exe tst.lic -mode codec -clibf <DLL_LIBRARY_NAME> -src file <TEST_AUDIO_FILE>

aqua-v.exe tst.lic -mode codec -clibf GSM610.dll -src file short.pcm

Command line argument: -clibf <file>

- codec library file name;

Use initial sound file as source or internal signal generator: -src file <fname> | gen <mode> | folder <fname>>

Determines source of initial sound: <file> - external sound file, <gen> - internal signal generator or <folder> external sound files from a folder.

In <file> mode one should specify name of audio file.

The signal generator mode may have one of the following option: short, normal or long.

In folder mode user specifies a path to audio files and file extensions (.wav or .pcm).

Set type of weight coefficients: -ct <ctype>

uniform, linear, logarithmic or htrsdelta;

Set name of the file under test: -tstf <fname>

<folder> mode sets name of folder, with the files to test.

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Set name of report file for folder processing mode: -frep <fname>

Set name of the file generated by the speech model: -dst <fname>

Generate full speech sounds distribution or synthesized sound: -sn <all | <sname>>

runs generation of full speech sounds distribution or defines name of the synthesized sound

Examples:

aqua-v.exe tst.lic -mode generate -sn all -dst SPEECH_MODEL_FILE

Here are options for generating **speech model** audio signal:
aqua-v.exe tst.lic -mode generate -sn all -dst generated_01.pcm

instead of "all" parameter one can specify **separate sound**s from the table of sounds you can see in the manual.

aqua-v.exe tst.lic -mode generate -sn a0 -voit female -slen 8000 -dst generated_02.pcm

aqua-v.exe tst.lic -mode generate -sn i0 -voit male -slen 8000 -dst generated_03.pcm

For separate sounds one can also set type of voice "-voit male/female" and duration of the sound to be generated "-slen 8000"

Set voice type: -voit <female | male>

- sets voice type for synthesized sound;

Set duration: -slen <num>

sets duration for synthesized sound equal to <num> samples

Set quality loss or naturallness: -qt <quality | naturalness>

- sets type of quality measurement **overall quality loss** or **voice naturalness**

Enable indication of codec speed performance: -power <on | off>

enables output of codec speed performance indicator;

Enable energy normalization -enorm <on | off>

- enables energy normalization;

Set number of link points: -npnt <num | auto>

sets number of link points (num = 1..10);

auto - enables detection of optimal amount of linking points (recommended);

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

Set precision of spectral analysis: -acr <num | auto>

sets spectral analysis precision. num = 8..16,
auto - enables automated analysis precision detection according to sampling frequency.

Set envelope smoothing level: -miter <num>

smoothing level is in the range of (1..10);

Turn on "waiting for key press" after showing voice quality output: -gch

turns on waiting for a key press after output of voice quality

Print reason of quality loss: -fau <fname>

prints reasons for quality loss to the file specified;

Set voice quality output type: -ratem <% | m | p>

%: voice / audio quality in percentage,

m: MOS score prediction (objective score of P.800 MOS prediction)

Set spectral analysis precision: -acr <num | auto>

sets spectral analysis precision. num = 8..16, auto - enables automated analysis precision detection according to sampling frequency

Set delta correction mode: -decor <on | off>

enable/disable delta correction;

Set spectrums integrating mode: -emode <normal | log | 10log>

Sets one of the integration modes: normal - linear, [10]log – logarithmic.

Set signsl type: -mprio <on | off>

sets signal type: on - music, off - voice

Set initial delay: -tdel <num>

sets delay in samples <num> from the beginning of test file. In order to obtain correct number of samples for certain period in milliseconds please use this formula: <num> = (delay (ms) * sampling frequency (Hz)) / 1000, and vice versa: delay (ms) = <num> * 1000 / Sampling frequency (ms).

Enable perception correction: -spfrcor <on | off>

turns on/off perception correction. This option introduces additional coefficients to specific frequencies is preferred for VoIP or G.729 signal only (8kHz only).

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Enable processing speech related frequency bands only -voip <on | off>

turns on/off processing of only speech related and specific frequency bands. In particular this parameter forces AQuA to consider signals only in the range between 300Hz and 3.4kHz (telephone frequency band). When the option is turned on differences in signals spectrum outside of the range above is not considered. This option is recommended for VoIP, mobile, PSTN and converged networks transmitting telephone-like speech signals.

Set psychoacoustics: -psyf <on | off>

sets psycho-acoustic filter on/off

Set psychoacoustics: -psyn <on | off>

sets psycho-acoustic normalyzer on/off

Set level gradation: -grad <on | off>

allows / forbids amplitude gradation

AQuA performance calculation: -tmc <on | off>

allows / forbids quality score calculation time measurement

Set average levels correction: -avlp <on | off>

enables / disables average levels correction

Smart energy normalization: -smtnrm <on | off>

enables /disables smart energy normalization. Performs energy normalization according to energy levels in integral spectrums of the most significant frequency band.

Export spectral pairs into CSV file: -specp <num> <fname>

exports specified amount (<num>) of spectral pairs into the file specified (<fname>). <num> parameter may be equal to 8,16 or 32. This is important for visualizing differences in original and degraded signals' spectrums.

Set program performance speed: -fst <num>

sets program performance speed. Increasing the speed decreases score accuracy. <num> should be in the range between 0.0 (slow) and up to 1.0 (fast).

Set silence trimming: -trim[-src|-tst] <a | r> <level>

sets silence trimming type: absolute (a) (should be below average signal level), or relative (r) threshold (should be below SNR level), the <level> parameter is set in dB and varies from 0.0 up to 120.0.

- -trim-src run trimming for source file only;
- -trim-tst run trimming for depredated file only;
- -trim run trimming for both files (synchronizes both audio files in time domain).

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

AQuA Command Line usage

Most of our customers represent the following business segments:

- VoIP service providers
- Mobile service providers
- PSTN service providers
- Sattelite service providers
- Audio and web conferencing providers
- Radio communications
- Unified communications
- Solution providers for telecom

AQuA helps telecom business to solve a wide range of tasks:

- test conference bridges quality when dialing from different locations
- monitor quality on a conference bridge
- monitor quality to certain destinations
- monitor quality at different terminations by end-to-end testing with termination's echo server
- test quality in converged networks (f.e. Mobile-VoIP)
- device testing in noisy environment
- audio improvement algorithms development

In all cases AQuA is the means end-to-end intrusive (active) testing, which involves a reference audio file compared to the one passed through a network, device or any other environment that may introduce degradation (f.e. a voice codec).

In order to show how AQuA performs perceptual voice quality assessment we are going to use WAV files one can download from Microtronix web site (http://www.microtronix.ca/pesq.html). However, one can use any audio files within AQuA Wideband or those that are recorded at 8Khz sampling and are 16 bit mono (in case of AQuA Voice).

Compare two audio files and learn about reasons for voice quality loss

To compare two audio files in AQuA Command Line version when one is interested to get extensive feedback from the software we suggest to invoke AQuA in the following manner:

aqua-wb.exe tst.lic -mode files -src file Or272.wav -tstf Dg002.wav -acr auto -npnt auto -miter 1 -ratem %m -fau log.txt

As result you will received the following output:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

Evaluation

File Quality is

Percent value 39.31

MOS value 2.23

Thus one can see that file comparison gives only 39.31% of similarity what corresponds to 2.23 MOS. By the way, this is an example of when AQuA does detect voice quality loss and PESQ does not (please read more details about this test case on Microtronix page).

After test was executed log.txt file contains quantitative reasons for voice quality loss:

Source SNR : 63.160200.

Degradated SNR : 71.187480.

Duration distortion.

Audio stretching corresponds to 1.41 percent.

Delay of audio signal activity.

Signal delayed by 100.000000 ms.

Audio signal activity mistiming (unsynchronization) is 1.25 percent.

Corrupted signal spectrum.

Overall spectral energy distortion approaches 62.18 %

Vibration along the whole spectrum [-19.73, 42.45] %

Significant distortion in low frequencies band.

Energy distortion approaches 32.27 %

Spectrum vibration in low frequency band [-16.91, 15.36] %

Significant distortion in medium frequencies band.

Energy distortion approaches 27.10 %

Amplification approaches 24.29 %

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

Test two audio files and receive audio quality score

In case we like to simply compare two audio files and get feedback on how similar the quality of the one under test is towards the reference audio we suggest invoking AQuA in the following manner:

aqua-wb.exe tst.lic -mode files -src file Or272.wav -tstf Dg001.wav -acr auto -npnt auto -miter 1 -ratem %m

Result will be:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Copyright (c) 2012 by Sevana Oy, Finland. All rights reserved.

Evaluation

File Quality is

Percent value 92.08

MOS value 4.89

or invoking it for the other degraded file:

aqua-wb.exe tst.lic -mode files -src file Or272.wav -tstf Dg002.wav -acr auto -npnt auto -miter 1 -ratem %m

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Copyright (c) 2012 by Sevana Oy, Finland. All rights reserved.

Evaluation

File Quality is

Percent value 39.31

MOS value 2.23

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Adapting AQuA to actual environment

AQuA parameters have pre-set values by default, however, in some cases it is required to adapt the algorithm to actual environment, which is network, device, or specific codec. Majority of our customers don't require adjusting AQuA parameters, but in some cases software tuning makes test results more consistent. There is no common case when it's 100% required, but some of our customers mentioned that when doing tests in mobile networks, or VoIP-mobile this tuning gives better scores.

In case your tests show unexpected results means that AQuA engine or VAD may need tuning. We suggest to start with these parameters first:

- -npnt
  - This parameter sets the amount of linking points required to catch different "holes" inside the signal. By default the value is 5.

- -miter
  - Sets amount of voice activity detector frames that are used during smoothing. By default it's 5. This is required to smooth the detector's vibration.

For example:

aqua-wb.exe tst.lic -mode files -src file Or272.wav -tstf Dg002.wav -acr auto -npnt 1 -miter 5 -ratem %m

Result is:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Copyright (c) 2012 by Sevana Oy, Finland. All rights reserved.

Evaluation

File Quality is

Percent value 39.73

MOS value 2.25

or invoking it for the other degraded file:

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

aqua-wb.exe tst.lic -mode files -src file Or272.wav -tstf Dg001.wav -acr auto -npnt 1 -miter 5 -ratem %m -fau log.txt

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Evaluation

File Quality is

Percent value 90.22

MOS value 4.82

In fact this result is much closer to what one would hear, however, the file was degraded. One can find the reasons for voice quality loss in the log.txt file, e.g.:

Source SNR : 63.160200.

Degraded SNR : 60.847082.

Duration distortion.

Audio stretching corresponds to 14.15 percent.

Advancing of audio signal activity.

Signal advances the original by -400.000000 ms.

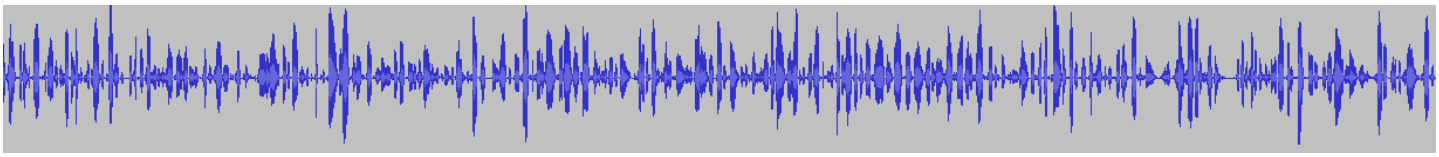
Audio signal activity mistiming (unsynchronization) is 1.34 percent.

Synchronizing reference and test files using AQuA 6.x

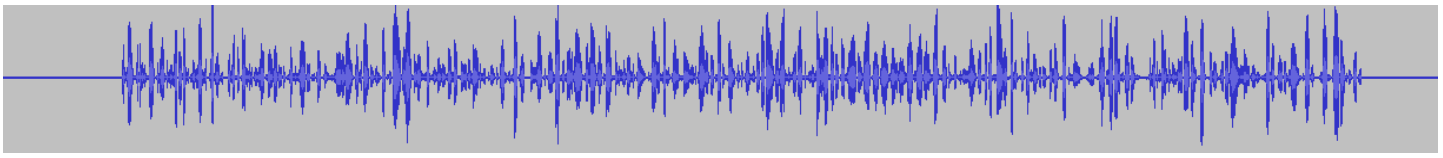
In many cases when monitoring voice quality in real life one receives degraded file from the network containing pauses (silence) before and/ or after the actual audio. Let's consider an example received from one of our customers while doing voice quality monitoring in a mobile network. Initial audio is a male voice pronouncing a phrase in English language with the following wave form:

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

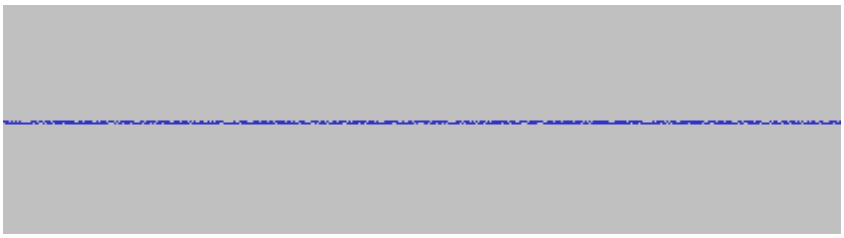
Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178



This audio is sent over a mobile network and then recorded back, but due to delays before the call is established and after hang-up degraded file has delays in the beginning and end of the audio:



Furthermore, if one zooms into the “silence” he will realize that it contains noise:



According to AQuA algorithms introduction of silence or noise into audio signal leads to quality degradation, and taking into account that establishing a test call as well as then detecting disconnect tone may take even a couple of seconds, this may significantly decrease the final quality score.

In order to trim irrelevant parts of the test signal in the beginning and end of the degraded file one just needs to invoke AQuA with a -trim parameter:

aqua-wb.exe tst.lic -mode files -src file male.wav -tstf male_5s_delay_5s_end_-36db_whitenoise.wav -acr auto -npnt auto -miter 1 -trim r 5 -ratem %m -fau log.txt

AQuA output will be:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Evaluation

File Quality is

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Percent value 74.59

MOS value 3.98

or one can use another option as described above:

aqua-wb.exe tst.lic -mode files -src file male.wav -tstf male_5s_delay_5s_end_-36db_whitenoise.wav -acr auto -npnt auto -miter 1 -trim a 45 -ratem %m -fau log.txt

AQuA output will be:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Copyright (c) 2012 by Sevana Oy, Finland. All rights reserved.

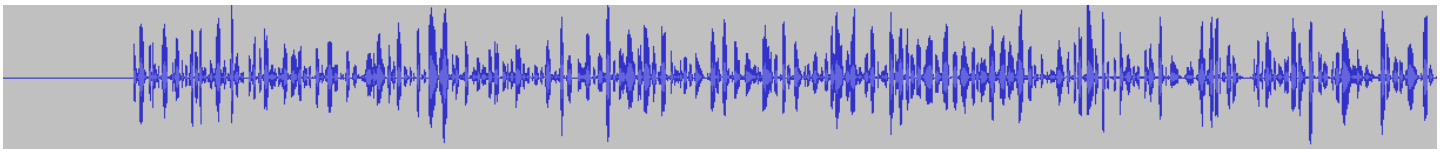
Evaluation

File Quality is

Percent value 75.28

MOS value 4.02

However, in order to be absolutely sure that the trimming works properly let's test it with an artificially created file containing silence:



aqua-wb.exe tst.lic -mode files -src file male.wav -tstf male_5s_delay_5s_beginning.wav -acr auto -npnt auto -miter 1 -trim a 45 -ratem %m -fau log.txt

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Copyright (c) 2012 by Sevana Oy, Finland. All rights reserved.

Evaluation

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

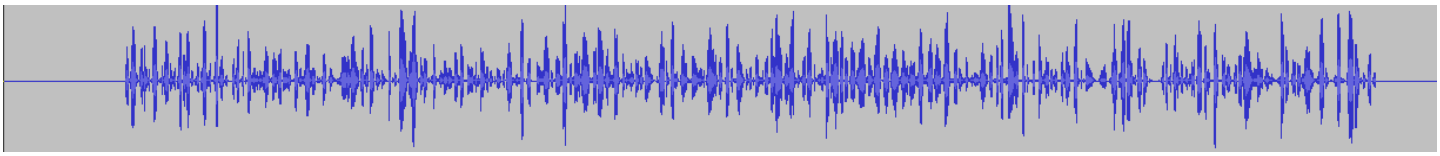
Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

File Quality is

Percent value 100.00

MOS value 5.00

and another file with silence in the beginning and the end of the file:



aqua-wb.exe tst.lic -mode files -src file male.wav -tstf male_5s_delay_5s_end.wav -acr auto -npnt auto -miter 1 -trim a 45 -ratem %m -fau log.txt

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Evaluation

File Quality is

Percent value 100.00

MOS value 5.00

Analysis of possible reasons for voice and audio quality loss

Besides audio quality score AQuA gives a possibility to determine and analyze possible reasons that caused audio signal degradation. Software automatically prepares analysis results that are stored in a log file.

Additional audio quality metrics returned by the system may not look trivial to understand and this chapter is devoted to the main principles of how these metrics are built and how one can interpret them.

AQuA returns additional metrics only in the case when they are out of range for their “typical values” (exception Signal/Noise Ratio (SNR) that is always present in the report). In case the metrics are within the range the system returns “Cannot determine the major reason for audio quality loss”.

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

Signal/Noise Ratio (SNR)

These metrics represent SNR both in the original and degraded files

Source SNR : XX.XX.

Degradated SNR : XX.XX.

These metrics show the signal/noise ratio of the original and degraded signals. Typically signal quality gets lower when SNR value decreases (or significantly increases) in the test audio.

Duration distortion

This metric represents continuity of compared audio files. Ideally amount of audio data in the original signal and file under test should be the same. During audio processing or transfer over communication channels audio fragments may be lost as well as inserted into the audio. If such audio degradation took place then value of this metric is lower than 100. The bigger the difference the stronger the degradation, however, this metric does not consider possible starting pauses.

When the value is less than 100% this means that audio data was lost and analysis result will be:

Audio shrinking corresponds to XX.XX percent.

where XX.XX corresponds to deviation from 100%.

When the actual value is more than 100% this means that data was inserted and analysis result will be:

Audio stretching corresponds to XX.XX percent.

where XX.XX corresponds to deviation from 100%.

Tolerance range for this value is set to 100% $\pm$ 1%.

Delay/Advancing of audio signal activity

This metric represents signal shift in test file compared to the original and determines how much active level of the test signal delays/advances active level of the reference (original) signal. When it is delayed analysis returns the following

Signal delayed by XX.XX ms.

where XX.XX is delay time in milliseconds. Correspondently, when the signal advances the original the return string is

Signal advances the original by -XX.XX ms.

where XX.XX is advancing time.

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Tolerance range for this value is interval of ±50 ms.

Audio signal activity mistiming

This metric represents unsynchronization of active levels in reference and test signals. Original (reference) audio signal and test signal are merged to determine characteristics of audio activity, and when the characteristics of audio activity do not match system increases unsynchronization counter. After processing the final unsynchronization value is presented as percentage of cases when unsynchronization was detected.

If the metric value is not zero analysis result represents it as:

Audio signal activity mistiming (unsynchronization) is XX.XX percent.

where XX.XX is percentage of unsynchronization. The value is not considered if it is less than 1%.

Corrupted signal spectrum

This represents a set of metrics reflecting differences in integral energy spectrums of the original signal and audio under test. If overall spectrums difference is more than 15% than analysis returns the following string:

Corrupted signal spectrum.

If difference in spectrums is multidirectional (goes both into positive and negative zones) analysis returns the following string:

Vibration along the whole spectrum [-XX.XX, YY.YY] %

where XX.XX and YY.YY are deviations to negative and positive zones correspondently. Tolerance range of the deviation is ±5%.

If spectrum distortions are unidirectional (only negative or only positive) analysis returns this string:

Amplification approaches YY.YY %

When distortions are positive, or

Attenuation approaches XX.XX %

when distortions are negative.

Other metrics returned by analysis correspond to distortions occurred in different frequency groups. Analysis of different frequency bands performs in a similar manner to spectrum analysis. When talking about frequency bands in question we consider:

Low frequencies – below 1000 Hz

Medium frequencies – from 1000 Hz to 3000 Hz

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

High frequencies are those that are greater than 3000 Hz

When analyzing frequency bands we use different tolerance range for different bands. Distortion in low frequencies is considered when they are greater than 5%, in medium frequencies – 10% and in high frequencies – 30%.

Multidirectional spectrum changes (vibration) are considered when they are greater than 2.5% in low frequencies, 7% in medium frequencies and 15% in high frequencies.

Unidirectional distortions (no matter positive or negative) are considered when they are greater than 5% in low frequencies, 10% in medium frequencies and 25% in high frequencies.

Visualizing signals spectrum for analysis

AQuA 5.x has a special parameter to store pairs of spectrum energy in critical bands of original and degraded audio to a .csv file:

aqua-wb tst.lic -mode files -src file male.wav -tstf male_5s_delay_5s_end.wav -npnt auto -miter 1 -ratem %p -fau report.txt -tmc on -gch -psyn off -psyf off -smtnrm on -enorm on -grad on -specp 32 spect.csv

This command produces the following output:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Copyright (c) 2012 by Sevana Oy, Finland. All rights reserved.

Evaluation

File Quality is

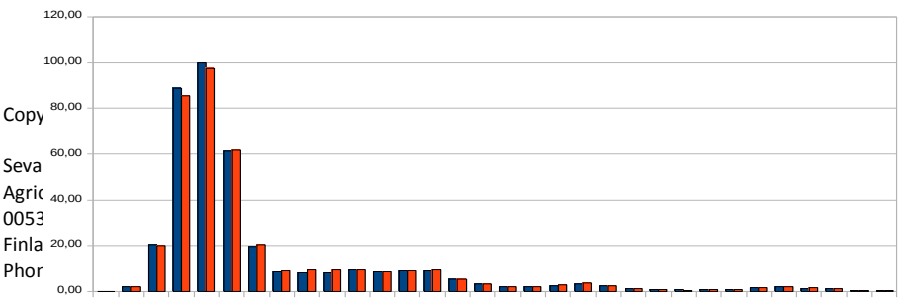
Percent value 91.37

PESQ value 3.30

Calculating time 1.424000 sec.

Press any key to continue....

File spect.csv contains 32 pairs related to spectrum energies of both files, so after importing the file into electronic spreadsheet we can plot a diagram visualizing differences in signals' spectrum:



As one can see the difference is not big and the reasons for received MOS score are stored in report.txt:

Source SNR : 73.214584.

Degradated SNR : 73.277180.

Duration distortion.

Audio stretching corresponds to 9.81 percent.

Delay of audio signal activity.

Signal delayed by 4990.000000 ms.

Audio signal activity mistiming (unsynchronization) is 1.14 percent.

As one can see the main reasons are mistiming and delay, which we have not removed, and if we remove it as described in previous chapter:

aqua-wb tst.lic -mode files -src file male.wav -tstf male_5s_delay_5s_end.wav -npnt auto -miter 1 -ratem %p -fau report.txt -tmc on -gch -psyn off -psyf off -smtnrm on -enorm on -grad on -trim r 5 -specp 32 spect.csv

we receive result showing that the files are of identical quality:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

Copyright (c) 2012 by Sevana Oy, Finland. All rights reserved.

Evaluation

File Quality is

Percent value 99.94

PESQ value 4.49

Calculating time 1.386000 sec.

Press any key to continue....

AQuA Online

AQuA Online is an olnine service that we implemented to show our customers and partners how simply one can create an advanced audio quality analysis service. The online script uses the same information presented in this

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

manual to provide end users with full analysis of audio quality and recommendations on what we consider important to improve in order to achieve higher perceptual quality. Access to online demo or service is available on demand.

Let's use the same voice files and test Or272.wav against Dg001.wav, here is output of our online service based on the latest AQuA version:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

```
***     Percent value   93.50 ***
***     MOS value       4.93 ***
```

Source SNR : 63.160200.
Degradated SNR : 60.457320.
Duration distortion.
Analysis shows that during audio processing or transfer over communication channels audio fragments might have been lost from the original signal. Tolerance range for this value is set to 1%
Audio shrinking corresponds to 1.33 percent.

Low frequencies - below 1000 Hz
Medium frequencies - from 1000 Hz to 3000 Hz
High frequencies are those that are greater than 3000 Hz

Distortions in low frequencies are considered when greater than 5%, in medium frequencies - 10% and in high frequencies - 30%.
Multidirectional spectrum changes (vibration) is considered when greater than 2.5% in low frequencies, 7% in medium frequencies and 15% in high frequencies.
Unidirectional distortions (no matter positive or negative) are considered when greater than 5% in low frequencies, 10% in medium frequencies and 25% in high frequencies.

And now let's analyze Or272.wav against Dg002.wav, which according to the information from Microtronix web site was equalized such that there is far less low frequency and high frequency energy when compared to the original file. Our online service gives the following:

Sevana Audio Quality Analyzer - AQuA-Wideband v.6.12.3.28.

```
***     Percent value   28.77 ***
***     MOS value       1.68 ***
```

Source SNR : 63.160200.
Degradated SNR : 69.740421.
Duration distortion.
Analysis shows that during audio processing or transfer over communication channels audio fragments might have been lost from the original signal. Tolerance range for this value is set to 1%
Audio shrinking corresponds to 1.05 percent.

Characteristics of audio activity between the origianl and degraded file don't match. We recommend

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

unsynchrnization value below 2%
Audio signal activity mistiming (unsynchronization) is 5.18 percent.

Overall difference between signals' spectrums exceeds 15%!
Corrupted signal spectrum.
Overall spectral energy distortion approaches 70.55 %
Difference in signals' spectrums is both in positive and getative zones. Deviation range should not be out of -5% to +5% range.
Vibration along the whole spectrum [-16.16, 54.39] %

Significant distortion in low frequencies band.
Energy distortion approaches 26.51 %
Spectrum vibration in low frequency band [-14.39, 12.12] %

Significant distortion in medium frequencies band.
Energy distortion approaches 36.06 %
Spectrum distortions are in positive zone.
Amplification approaches 34.29 %

Low frequencies - below 1000 Hz
Medium frequencies - from 1000 Hz to 3000 Hz
High frequencies are those that are greater than 3000 Hz

Distortions in low frequencies are considered when greater than 5%, in medium frequencies - 10% and in high frequencies - 30%.
Multidirectional spectrum changes (vibration) is considered when greater than 2.5% in low frequencies, 7% in medium frequencies and 15% in high frequencies.
Unidirectional distortions (no matter positive or negative) are considered when greater than 5% in low frequencies, 10% in medium frequencies and 25% in high frequencies.

As one can see AQuA did detect the equalization providing quite detailed information on what happened to the test audio. The fact that the quality scores in AQuA 6.x and AQuA 5.x may differ is result of our software improvement based on customer feedback and research made during the past years, and we believe that AQuA 6.x can better objectively predict the subjective mean opinion scores as in a P.800 listening setup.

AQuA Benefits

Among AQuA benefits one will definitely appreciate that:

- AQuA is available for Windows, Linux and MAC OS operating systems
- AQuA is availbale for both for 32 and 64 bit systems
- AQuA is easy to deploy and use for software products development
- AQuA provides perceptual estimation of audio quality and can be utilized in VoIP, PSTN, ISDN, GSM, CDMA, LTE/4G, sattelite and radio networks and combinations of those

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

SEVANA
IT solutions & services

AQuA Error Messages

The following error messages may appear during AQuA run time:

When writing spectrum pairs into a file

(Please note if error occurs when writing file with spectrum pairs then error message can be found in the output file)

Wrong number of spectrum pairs!

Source file is too short!

Test file is too short!

Wrong numbers of spectrum pairs! Please check AQuA Manual.

An error occurred while calculating spectrum pairs!

When reading command line parameters the following errors may occur:

Parameters listing error.

Mode error (-mode).

Data source error (-src).

Sound data generator error (-src).

Weights of sound bands error (-ct).

Synthesized voice type error (-voit).

Incorrect sound duration to synthesize (-slen).

Unknown type of quality estimation defined (-qt).

Incorrect numbers of links points (-npnt).

Incorrect envelope smoothing level (-miter).

Incorrect out quality mode (-ratem).

Incorrect accuracy value (-acr).

Test file name is missing (-tstf).

Speech model file name is missing (-dst).

Name of synthethized sound is missing (-sn).

Codec performance measurement is not set (-power on|off).

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Unknown parameter in codec performance measurement option (-power).

Parameter is missing in energy normalization management option (-enorm).

Unknown parameter in in energy normalization management option (-enorm).

File name to store reasons for audio quality loss is not specified.

Parameter is missing in delta-correction management option (-decor).

Unknown parameter in delta-correction management option (-decor).

Parameter is missing in integration mode option (-emode).

Unknown parameter in integration mode option (-emode).

Parameter is missing in signal type selection option (-mprio).

Unknown argument in signal type selection option (-mprio).

Value of initial delay is missing (-tdel).

Value of initial delay is negative or integer value set is incorrect (-tdel).

Parameter is missing in option (-spfrcor).

Unknown parameter in option (-spfrcor).

Parameter is missing in psycho-acoustic model option (-psyf).

Unknown parameter in option setting psycho-acoustic filter on (-psyf).

Parameter in time measurement option is missing (-tmc).

Unknown parameter in time measurement option (-tmc).

Parameter in smart energy normalization option is missing (-smtnrm).

Unknown parameter in smart energy normalization option (-smtnrm).

Parameter is missing in option (-avlp).

Unknown parameter in option (-avlp).

Parameter in psy-normalizer management option is missing (-psyn).

Unknown parameter in psy-normalizer management option (-psyn).

Incomplete parameters in spectrum pairs output option (-specp).

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178

Incorrect amount of spectrum pairs (-specp).

Parameter is missing in filter management option (-voip).

Unknown parameter in filter management option (-voip).

Incompatible set of parameters of psy-normalization and amplitude ranges.

Parameter is missing in amplitude ranges management option (-grad) .

Unknown parameter in amplitude ranges management option (-grad).

Parameter is missing in program performance speed option (-fst).

Program performance parameter is out of range (-fst).

Parameter(s) is missing in silence trimming option (-trim).

Unknown type of silence level detection (-trim).

Incorrect silence level (-trim).

Source audio files folder or extension are not defined (-src).

Data source type is not defined (-tstf).

Test files folder is not defined (-tstf).

Report file name is missing (-frep).

When calculating quality score the following errors may occur:

Error opening source file!

Error opening file under test (degraded)!

Error: files have different sampling frequencies!

Error: sampling frequency is not supported.

Error: files have different channels!

Error: sampling frequency (in source file) is not supported.

Error: sampling frequency (in degralated file) is not supported.

Sevana Oy
Agricolankatu 11
00530 Helsinki
Finland
Phone: +358 9 2316 4165

Sevana Oü
Rohtlaane 12
76911 Huuru kula
Estonia (Harjumaa)
Phone: +372 53485178