Intrusive Monitoring of Speech Quality

Pero Juric
Director System Engineering
SwissQual AG
CH – 4528 Zuchwil
http://www.swissqual.com Email: pero.juric@swissqual.com

Summary

This document presents a new approach to assess the end-to-end speech quality by using both noisy and clean speech test signals. Two new models Squad-LQ and Squad-NS are presented. Squad-LQ is an algorithm for objective assessment of one-way speech quality and is able to quantify the causes for speech quality degradation. Squad-NS is an objective model for rating of Noise Reduction by using of noisy speech reference signals. The model is able to determine Quality Index for Noise Reduction as a function of several analytic parameters. By combining the results of the two models a comprehensive objective view of the telecommunication network under noisy- and clean conditions is possible.

After its transmission through a telephone channel a speech signal usually experiences variety of degradations caused by filters, low bit-rate speech codecs, compression circuits, voice activity detectors, comfort noise generators, echo cancellation, noise reduction, adaptive level control, packet-switched transmission, etc. Due to the effects of one or more of the network devices mentioned above, a speech signal sent through such a system may suffer from a variety of degradations, including the following: temporal clipping, amplitude clipping, spectral degradation, fixed and/or variable delay, speech transcoding, adaptive level, high background noise, bad comfort noise insertion, channel interferences or presence of signaling tones. Further an end-to-end speech quality can be improved if noise reduction devices are activated.

Measuring of only MOS (Mean Opinion Score), is not sufficient to find out the causes of degradation in a network. VoIP network operators, for example, are very interested in detecting calls degraded by temporal clipping and signal interruptions typically caused by packet loss. In the case of end-to-end measurement, especially in internetworking situations, such an additional expertise of degradation causes is the only help to the users when trouble-shooting.

Figure 1 shows the range of employment for both models. **Squad-LQ** is comparable with current ITU-T's standard P.862 with additional wideband capability, whereas **Squad-NS** is designed for Noise Reduction testing.

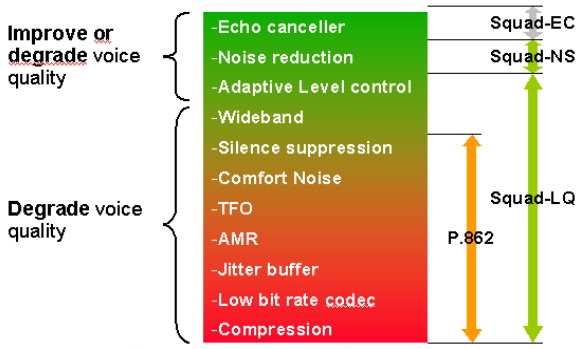

Figure 1: *Squad-LQ rates speech degradations whereby Squad-NS assesses noise reduction devices. Squad algorithm family covers the whole range of speech quality impairer and enhancer.*

Squad-LQ as shown in **Figure 2** consists out of two main modules. The module on the left side predicts listening only speech quality on ACR scale (see ITU P.800) and the module on the right side is designed for expertise of degradation causes. The evaluation of cause values can be separated into four groups:

1) Time domain
- Circuit noise estimation
- Speech level measurement
- Mean value of the signal power
- Gain estimation
- Temporal clipping detection
- Calculation of a delay distribution
- Signal energy envelopes

2) Frequency domain
- Frequency shift detection

- Transfer characteristics

3) Impulsive noise detection
4) Analysis of quality distribution

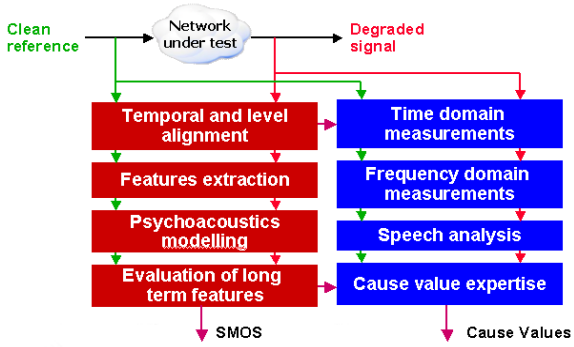

Figure 2: *Left part of Squd-LQ is psichoacoustics modelling of human hearing mechanism. SMOS is predicted speech quality value. Right part analyses the causes of speech quality degradations.*

The final output of the Analysis of Cause Values module is a list of one or more of the following cause values:
"MOS-drops", "Circuit noise", "Variable delay", "Front end clipping", "Wrong level", "High DC-Offset" , "Impulse noise", "Frequency shift" and "Non-Linear filtering".
The list of new parameters can be easily expanded if new requirements in speech quality measurement have to be met.

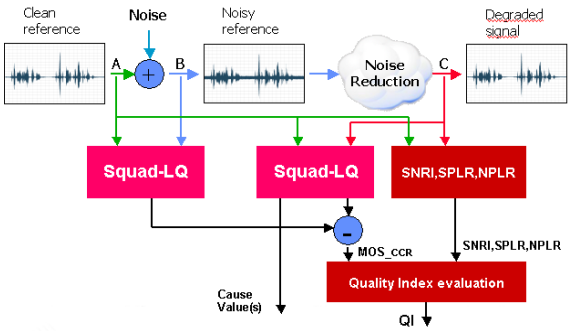

Figure 3: *Squad-NS uses Squad-LQ for its calculation. MOS_ccr is combined with other objective numbers in order to obtain Quality Index for Noise Reduction*

A model for Noise Reduction testing is shown in **Figure 3**. Squad-LQ is executed two times, first time for reference alignment, and second time for particular Noise Reduction measurement. First run has to be executed only ones per reference. The difference MOS_ccr is equal zero if there were no degradation or enhancement in voice channel. The value is scaled according to a CCR scale (ITU P.800). Signal Noise Ratio Improvement (SNRI), Noise Power Level Reduction (NPLR) and Speech Power Level Reduction (SPLR) are further parameter needed for Quality Index (QI) calculation. QI is mapped onto scale 1 to 5 where 3 stays for god Noise Reduction. The range bellow 3 is reserved for bad Noise Reduction (les then 2 means there is no noise reduction).

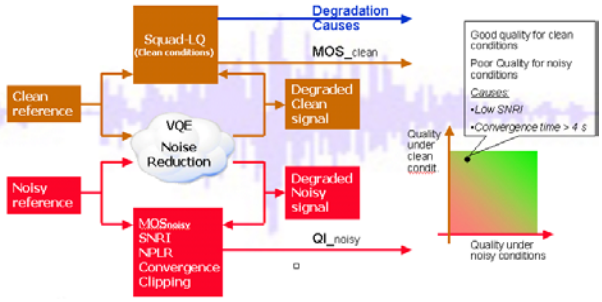

Figure 4: *By combining of both models Squad-LQ and Squad-NS it is possible to assess a speech quality for noisy- and clean conditions with additional cause value expertise.*

The combined method as shown in **Figure 4** is a cascaded test, which includes speech quality measurement with clean- and noisy- reference signal. The clean speech quality is done with Squad-LQ and evaluated by means of sending the clean speech reference through the transmission channel with the result expressed in MOS units scaled in range 1 to 5. The noisy speech quality is calculated as a combination of four parameters: NPLR, SNRI, SPLR and noisy speech quality. Both quality values are then presented on 2-D plane showing the qualities for both noisy- and clean conditions. Additionally the cause values are calculated for each measurement.

Conclusion

*We propose to add a new speech quality class called "Better Then Reference" within standardisation bodies (ETSI, ITU,...) in the tables including definition of the speech quality classes. With **Squad** measurement methods developed by SwissQual it is possible to measure the new quality class.*

Considering the increased use of mobile-to-mobile calls and the increased percentage of mobile calls being made in noisy environments, it is of high importance to conduct test reflecting such situations and subscriber behaviour. Furthermore, with the increased use of network elements such as Voice Quality Enhancers (VQE), the use of noisy speech samples for testing is the only way to verify the effect of such equipment.