# QUALITY ASPECTS OF DIGITAL HEARING AID ALGORITHMS

*T. Rohdenburg, V. Hohmann, R. Huber, B. Kollmeier*

Thomas.Rohdenburg@uni-oldenburg.de
University of Oldenburg, Medical Physics Group, D-26111 Oldenburg, Germany

## Abstract

For several modern signal-processing schemes for hearing aids (e.g. single- or multi-channel noise reduction schemes, dynamic compression, dereverberation) technical measures are used during the development process for performance prediction that are either inappropriate or can only approximately predict the success for hearing impaired (and also normal hearing) people. In this context, not only the speech intelligibility, but also the subjective signal quality is of importance. In this study we present and compare subjective and objective assessment methods. Particularly, the objective quality measure PEMO-Q that is based on modern auditory models is evaluated for different hearing aid algorithms. This approach appears to be significantly superior to the classic SNR-based approaches in some quality aspects.

## 1. ALGORITHMS AND SIGNALS

In this exemplary study, STSA algorithms according to Ephraim and Malah's weighting rules [1] where employed as single channel state-of-the-art algorithms. These algorithms are characterized by a strong reduction of noise while introducing only little of the well known *musical tones* or *musical noise* that result from subtracting an average noise spectrum from a non-stationary frame-based spectral estimate. A detailed description of the involved filter parameters can be found in Cappé [2]. The most important parameters are two signal-to-noise ratio (SNR) estimates: An instantaneously estimated (a posteriori) SNR and an a priori SNR estimate that is calculated by a recursive smoothing of preceding a posteriori values. The considered algorithms need a reliable noise power estimation. Here, the minimum statistics method (MinStat) by Martin [3] and a Voice Activity Detection (VAD) algorithm by Marzinzik [4] are used.
The speech signals used here were taken from the Oldenburg Logatome Speech Corpus (OLLO) [5] and consisted of six sentences spoken by German male and female speakers. The noise signals were speech-shaped noise, cafeteria noise, icra7 noise (speech like modulated noise) and white gaussian noise. All signals had an approximate duration of 20 seconds and a sampling rate of 16 kHz. In the simulation system the signals were mixed at a SNR of 0 dB and 5 dB. The calculation of the time-variant filter was made on this mixture while the filtering process was also done on the separate speech and noise signals for subsequent quality assessment and the calculation of the SNRE and other quality measures.

## 2. OBJECTIVE MEASURES

Several objective measures were assessed in this study. The quality prediction method PEMO-Q is based on a psychoacoustically validated quantitative model of the "effective" peripheral auditory processing by Dau et al. [6].
The perceptual similarity measure (PSM), obtained from PEMO-Q, is a correlation measure between two so called *internal representations* of acoustic stimuli, i.e., the output of the modeled peripheral auditory system. PSM serves to predict the perceived similarity between two given signals, generally a reference signal and a test signal whose quality is to be measured.
As another psychoacoustical measure, the ITU standard measure PESQ is used. PESQ stands for "Perceptual Evaluation of Speech Quality" and is an

enhanced perceptual quality measurement for voice quality in telecommunications [7].
PSM and PESQ were varied with different options and reference signals (e.g. SNR_PSM, PSM_b, ΔPESQ,), for details see [8].
Besides the above perceptual measures also more "technically based" quality measures where incorporated (see [8] for more details). These were signal-to-noise ratio enhancement (SNRE), coherence, a critical bandwidth weighted SNRE (freq. wt. SNRE), and the quality evaluation measures defined by Hansen and Pellom [9]: segmental SNR, Log-Area Ratio (LAR), Log-Likelihood Ratio (LLR) and the Itakura-Saito Distance (ISD). The last three measures are based on a LPC-Model.

## 3. EXPERIMENTS AND RESULTS

The recursive smoothing parameter $\tau$ of the STSA algorithm was varied in the range from $0 - 800$ms to cover a broad range of noise reduction and signal quality. All signals were processed with the noise estimators Minstat and VAD, respectively. For each setting, the above mentioned quality measures were calculated for a number of speech signals mixed with different types of noise. Subjective listening tests were done according to ITU-T Recommendation P.835 [10] which describes a methology for evaluating the subjective quality of speech in noise and is particularly appropriate for the evaluation of noise suppression algorithms. The data from subjective listening tests was then correlated with the quality predictions of the objective measures (see table 1 for the overall quality prediction). The best correlated measures (see Fig. 1) were chosen to predict the subjective rating curves. The objective curves have been linearly fitted to match the scaling of the subjective curves. As for the background noise rating, the highest correlations are gained by the SNRE. This means that SNRE is a good measure to rate the amount of noise reduction by an algorithm, independent of the speech signal quality. In terms of speech-signal rating the perceptual measures with the noisy reference (SNR_PSM, SNR_PESQ) have the highest correlations with subjective data. The best correlation in terms of overall quality rating show the perceptual measures with clean speech reference (PSM_b and PESQ).

| Correlation with overall quality rating | Cafeteria noise | White noise | Speech-shaped noise | ICRA7 noise | Overall-Correlation |
|---|---|---|---|---|---|
| SNRE | 0.35 | 0.66 | 0.41 | 0.29 | 0.35 |
| Coherence | 0.83 | 0.88 | 0.93 | 0.93 | 0.65 |
| seg. SNRE | 0.74 | 0.89 | 0.89 | 0.71 | 0.53 |
| freq. wt. SNRE | -0.17 | 0.40 | -0.05 | -0.11 | 0.00 |
| mean LAR | -0.46 | -0.90 | -0.45 | -0.28 | -0.07 |
| mean LLR | -0.75 | -0.94 | -0.61 | -0.68 | -0.43 |
| mean ISD | -0.24 | -0.79 | -0.04 | -0.16 | -0.34 |
| PSM | 0.82 | **0.92** | 0.93 | 0.87 | 0.70 |
| PSM_b | 0.85 | 0.91 | **0.94** | 0.93 | 0.76 |
| PESQ | 0.85 | **0.92** | **0.94** | **0.94** | **0.81** |
| ΔPSM | 0.71 | 0.69 | 0.58 | 0.54 | 0.39 |
| ΔPESQ | **0.86** | **0.92** | 0.48 | 0.65 | 0.47 |
| SNR_PSM | 0.12 | -0.04 | 0.09 | 0.01 | 0.04 |
| SNR_PESQ | 0.09 | -0.36 | 0.05 | 0.07 | 0.00 |

Table 1: Correlation between objective and subjective measures for the overall signal quality. (from [8])

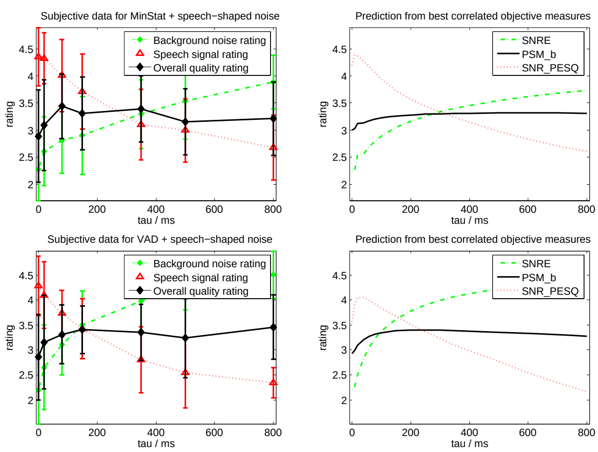


Figure 1: Subjective (left panel) and objective (right panel) data for speech-shaped noise and two algorithms (from [8])

## 4. REFERENCES

[1] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean–square error short–time spectral amplitude estimator," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 32, no. 6, pp. 1109–1121, 1984. 1

[2] O. Cappe, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE Trans. on Speech and Audio Processing*, vol. 2, no. 2, pp. 345–349, 1994. 1

[3] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Trans. on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, 2001. 1

[4] M. Marzinzik, *Noise Reduction Schemes for Digital Hearing Aids and their Use for Hearing Impaired*, Ph.D. thesis, University of Oldenburg, Nov. 2000. 1

[5] T. Wesker, B. Meyer, K. Wagener, B. Kollmeier, and A. Mertins, "Oldenburg logatome speech corpus (ollo) for speech recognition experiments with humans and machines," in *Interspeech*, 2005. 1

[6] T. Dau, D. Püschel, and A. Kohlrausch, "A quantitative model of the effective signal processing in the auditory system i," *Journal of the Acoustical Society of America (JASA)*, vol. 99, no. 6, 1996. 2

[7] ITU-T, "Perceptual evaluation of speech quality (pesq), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," Series P: Telephone Transmission Quality Recommendation P.862, International Telecommunications Union, Feb 2001. 2

[8] T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Objective perceptual quality measures for the evaluation of noise reduction schemes," in *9th International Workshop on Acoustic Echo and Noise Control*, C.P. Janse, M. Moonen, and P.C.W. Sommen, Eds., Eindhoven, 2005, pp. 169–172. 2, 1, 1

[9] J.H.L Hansen and Bryan Pellom, "An effective quality evaluation protocol for speech enhancement algorithms," in *Proc. Int. Conf. on Spoken Language Processing (ICSLP)*, Sydney, Australia, Dec. 1998. 2

[10] ITU-T, "Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm," Series P: Telephone Transmission Quality Recommendation P.835, ITU, Nov. 2003. 3