Robustness analysis for multi-channel hearing aid algorithms with binaural output by means of objective perceptual quality measures

Thomas Rohdenburg, Volker Hohmann, Birger Kollmeier

Medizinische Physik, Universität Oldenburg, 26111 Oldenburg

Introduction

According to the ITU-T P.835 recommendation, subjective quality evaluation of noise reduction schemes involves (i) the perceived quality of the speech signal, (ii) the quality of the background signal and (iii) the overall quality. In [7] it has been shown that these subjective measures are predictable by objective measures in the case of monaural noise reduction schemes. In this study we extend the quality prediction to the case of multi-channel algorithms. These microphone array based algorithms have other influences on signal quality than single channel envelope filters as they exploit the spatial configuration of interfering signals and therefore in general lead to less signal distortion. For hearing aid applications, data from literature suggest that it is important that the beamformer preserves the binaural information so that the listener can make use of the effect of spatial unmasking. In order to generate a binaural output [6] was adopted.

Signal model and algorithms

The signals were generated using two 3-channel hearing aid headsets mounted on a dummy head. 6-channel HRTFs in an anechoic room and real-world environmental noise in a cafeteria have been recorded. The input signal was composed from two directional signals filtered with HRTFs (target and interferer from 30° and -135° azimuth, respectively) and mixed with the recorded cafeteria noise to generate a near-to realistic scenario. The multi-channel algorithms used here are fixed superdirective beamformers that are designed by the well-known constraint Minimum Variance Distortionless Response (MVDR) solution [2]. This solution allows to include different assumptions on the wave propagation of the target signal and the characteristics of the noise field as described by its cross power spectral density matrix. Three different beamformers were designed with the assumptions about wave propagation (i) in free-field (aka far-field assumption) (ff), (ii) in a simple spherical head model according to [3] (hm) and (iii) with measured 6-channel HRTFs in an anechoic room (hr). These beamformers had monaural outputs that were enhanced by a binaural post-filter according to [6]. The processing block-diagram is shown in Figure 1.

Signal independent quality measures

The beam-pattern is a well-established measure to evaluate the signal independent directional response of a beamformer. It is computed as the response of the array to a wavefront coming from a specific angle at a specific frequency [2]. In general, beam-patterns are only evaluated for far-field propagation. When beamformer coefficients that are designed for far-field are used in a near-field environment with head influences, the constraint of distortionless response may not be fulfilled and the far-field beampattern does not reflect the measured directional response. Therefore, in the near-field or if head-shadow and diffraction effects play a role, these effects also have to be incorporated in the beampattern calculation. Figure 2 shows the beampattern for farfield, beamformer coefficients steered to 30°, (a) evaluated in farfield and (b) evaluated in the nearfield (HRTF). As the beamformer should be designed for the head-mounted array, beampattern (b) shows the more realistic behavior. It can be seen that the target-signal will be distorted and the lateral noise reduction is poor, which is in line with the signal dependent perfomance measures (see below). Also, for other perfomance measures like the *directivity index* the head-shadow and diffraction effects need to be incorporated.

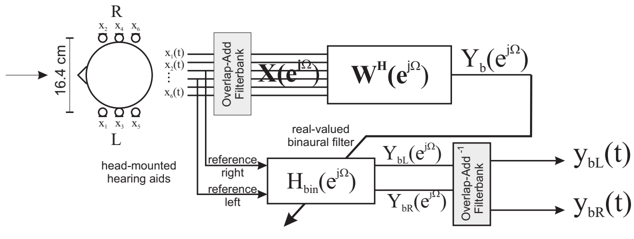


Figure 1: Signal model and beamformer setup

Signal dependent quality measures

SNRE

The SNR-Enhancement (SNRE) is the difference of the signal-to-noise ratio (SNR) at the output of the beamformer and a reference input-SNR, both measured in dB. For a comparison of multi-channel algorithms the choice of the reference is crucial. Here, the SNRE to different references (left, right, source, best microphone) are evaluated for a comparison with the perceptual measures (see below).

PSM

The quality measure PSM from PEMO-Q [4] estimates the perceptual similarity between the processed signal and the clean speech source signal. For monaural noise reduction schemes this measure has shown a high correlation with subjective overall quality ratings according to [5, 7]. Here, the PSM is measured between the clean speech source (before HRTF filtering) and the beamformer output (monaural) or the output of the binaural post-filter, respectively.

SRT

The speech reception threshold (SRT) is defined as the signal-to-noise ratio (SNR) at 50% speech intelligibility. In [1] a binaural model of speech intelligibility based on the equalization-cancelation (EC) processing by Durlach had been defined which is able to predict the SRT with high accuracy. For the objective quality assessment of binaural signals, we define a deduced measure here, namely the SRT Gain. The SRT Gain is calculated iteratively by reducing the SNR of the beamformer input signal until the predicted SRT has the same value as the original unprocessed reference signals. Thus, the SRT Gain is the amount of SNR reduction achieved by the algorithm as estimated by intelligibility estimates including spatial unmasking.

Results

The results in table 1 show that the beamformers with binaural outputs (D,E,F) in general have a higher SNRE. Although for the monaural A) a SNRE source of 5.8 dB was measured, the SNRE compared to the best microphone is almost zero. The same effect can be derived from the SRT Gain, it says that the SRT of the output is worse than the SRT of the unprocessed reference signal. This implies that this algorithm is not helpful to the listener, although the SNR is enhanced by 5.8 dB. Similar effects can also be seen for the other algorithms. The binaural algorithm F) has only a 1.5 dB higher mean SNRE than its monaural counterpart C), but the binaural output leads to an SRT Gain that is 4.3 dB higher. This means that the binaural algorithm can deal with an input signal that is 4.3 dB lower to gain the same speech intelligibility as the monaural algorithm.

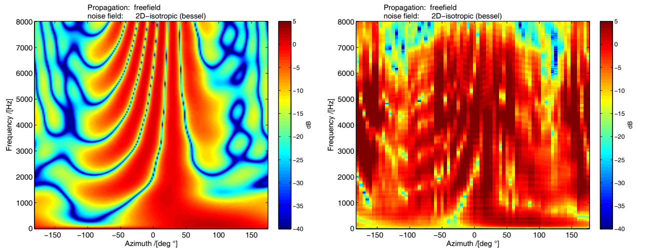
	Algorithm	SNRE Ref L	SNRE Ref R	SNRE source	SNRE best Mic	ΔPSM Ref L	ΔPSM Ref R	SRT Gain
A	mon-rec-ff-2d-fixed	1.6 dB	9.2 dB	5.8 dB	0.4 dB	0.06	0.30	-0.3 dB
B	mon-rec-hm-2d-fixed	2.3 dB	9.8 dB	6.4 dB	1.1 dB	0.12	0.36	1.7 dB
C	mon-rec-hr-hrtf-fixed	4.7 dB	12.2 dB	8.8 dB	3.5 dB	0.20	0.43	3.9 dB
D	bin-rec-ff-2d-fixed	5.3 dB	8.5 dB	7.3 dB	4.1 dB	0.20	0.23	4.5 dB
E	bin-rec-hm-2d-fixed	6.4 dB	7.9 dB	7.5 dB	5.2 dB	0.25	0.29	7.2 dB
F	bin-rec-hr-hrtf-fixed	9.0 dB	11.0 dB	10.3 dB	7.8 dB	0.29	0.40	8.2 dB

Table 1: Performance results for 3 beamformer designs with monaural and binaural outputs

Figures 3 (a-b) show preliminary robustness results for the three beamformers with binaural outputs. The performance measures are plotted over the steering mismatch. The results show that for all quality measures, the free-field and the head-model beamformers do not reach the optimal value at a steering mismatch of 0°. This is because the head diffraction is not (or not sufficiently) incorporated in the coefficients which leads to a steering to higher angles. On the other hand, the gradient of the performance curves is slightly steeper for the hrtf-beamformer which points out that it is more sensitive to steering errors.

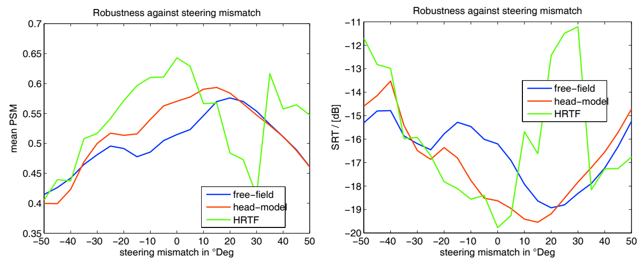
Outlook

Preliminary results have shown the importance of the incorporation of head-shadow and diffraction influences in



(a) (b)

Figure 2: Beampatterns for far-field beamformer coefficients steered to 30° and used (a) in far-field and (b) in near-field environment



(a) PSM (b) SRT

Figure 3: Robustness against steering mismatch

both the beamformer designs and the performance measures. Furthermore, the performance measures showed a significantly higher quality if the beamformer was extended by a binaural post-filter. The new binaural quality measure showed encouraging results and is an important step towards a robustness testbench for multichannel hearing aid algorithms with binaural outputs.

Work supported by the EC (DIRAC project IST-027787) and BMBF

References

[1] Beutelmann, R. ; Brand, T. : Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. In: *Journal of the Acoustical Society of America* 120 (2006), Nr. 1, S. 331–342

[2] Kap. 2 In: Bitzer, J. ; Simmer, K. U.: *Superdirective Microphone Arrays*. Springer, 2001, S. 19–38

[3] Brown, P. C. ; Duda, R. O.: A Structural Model For Binaural Sound Synthesis. In: *IEEE Trans. on Speech and Audio Processing* 6 (1998), Sep, Nr. 5, S. 476–488

[4] Huber, R. ; Kollmeier, B. : PEMO-Q - A new Method for Objective Audio Quality Assessment using a Model of Auditory Perception. In: *IEEE Trans. on Audio, Speech and Language Processing* (2006). – Special Issue on Objective Quality Assessment of Speech and Audio

[5] ITU-T: Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm / ITU. 2003 (Recommendation P.835). – Series P: Telephone Transmission Quality

[6] Lotter, T. ; Sauert, B. ; Vary, P. : A Stereo Input-Output Superdirective Beamformer for Dual Channel Noise Reduction. In: *Interspeech*. Lisbon, Portugal, sep 2005, S. 2285–2288

[7] Rohdenburg, T. ; Hohmann, V. ; Kollmeier, B. : Objective Perceptual Quality Measures for the Evaluation of Noise Reduction Schemes. In: Janse, C. (Hrsg.) ; Moonen, M. (Hrsg.) ; Sommen, P. (Hrsg.): *9th International Workshop on Acoustic Echo and Noise Control*. Eindhoven, 2005, S. 169–172