Implementation and evaluation of an experimental hearing aid dynamic range compressor

Giso Grimm[1,2], Tobias Herzke[2], Stephan Ewert[1], Volker Hohmann[1,2]
[1] *Medizinische Physik and Cluster of Excellence Hearing4all, Universität Oldenburg, Germany*
[2] *HörTech gGmbH, Oldenburg, Germany*
g.grimm@uni-oldenburg.de

Introduction

Hearing loss causes elevated thresholds of hearing in quiet, manifesting as reduced audibility of sounds at low and moderate levels. Moreover, hearing loss typically causes steeper than normal loudness growth (loudness recruitment), such that a reduced dynamic range is usable for the hearing-impaired listener, with sounds rising from just audible above the elevated threshold to uncomfortably loud at similar levels as for normal-hearing listeners. Dynamic range compression is the most important hearing-aid strategy for compensating the recruitment problem and to aid audibility of soft sounds while avoiding amplification at high levels.

For the development of hearing aid algorithms software-based research platforms are commonly used and offer versatility and lately sufficient performance for real-time applications [6, 4]. As a framework for hearing aid algorithms such as single-channel noise reduction, beamforming, and feedback cancelation, multiband dynamic compression is required to offer the key function of a research hearing aid. Here an experimental multiband dynamic range compressor is suggested which aims to represent a large class of commonly used hearing aid dynamic range compressors. The technical function of the compressor is described and the compressor is evaluated using instrumental measures as well as speech intelligibility measurements with hearing impaired listeners.

Algorithm

The compression algorithm proposed here was implemented in the HörTech Master Hearing Aid platform [7] as sketched in Figure 1. It processes audio signal in blocks of configurable duration, to enable real-time processing. It can either process a time domain signal or a short time Fourier transform (STFT) signal computed by an overlap-add procedure [1]. The signal is split into frequency bands by a filterbank. The filterbank used here is computed in the STFT domain and allows for overlapping frequency bands by applying, for each filter band k, frequency depended weights $W(\nu, k)$ to each STFT bin ν of the input spectrum $X(\nu)$ [7]. Effectively, each filter is a linear phase FIR filter.

The short time level in each filter band is computed from the signal's intensity. For the STFT domain, the negative frequency bins are not stored, but need to be taken into account when computing the band's intensity. A first

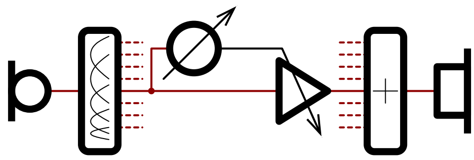

Figure 1: Schematic flow chart of the experimental dynamic range compressor: the microphone signal is analyzed by a filterbank. In each band, the input level is computed and determines the insertion gain to apply. The signal from all bands is then summed to a broad-band signal and acoustically presented to the user.

order recursive attack-release filter $\langle\cdot\rangle_{\tau_a,\tau_r}$ is applied to the logarithmic short time level; the filtered level $L_{\text{in}}(k)$ is the input into the gain table:

$$L_{\text{in}}(k) = \left\langle 10\log_{10}\left(\sum_{\nu=0}^{N} |X(\nu)W(\nu,k)C(\nu)|^2\right)\right\rangle_{\tau_a,\tau_r} \quad (1)$$

with $C(\nu) = 1$ for $\nu = 0$ and the Nyquist bin, otherwise $C(\nu) = 2$. A gain table contains insertion gains in each band for different input levels. Input level granularity is configurable. For input levels between table entries, gains are interpolated on a log-log scale. The gains are applied to the overlapping filters in the spectral domain before re-synthesis. For input levels outside the range covered by the gain table, the gain is computed by extrapolation. When dynamic compression is applied to STFT spectra, then by definition the band-specific gains apply to the complete duration of the STFT frame. Gain jumps between consecutive blocks are smoothed by the overlapping windows of the overlap-add procedure. When dynamic compression is applied to the time signal, then the filtered level and the gain are updated for every sample.

In this study, signals at 16kHz sampling rate were processed in the STFT domain. Processing frames are 64 samples in duration with Hann windows and 50% overlap (hop size: 32 samples). The 64-sample frames were zero-padded to 128 samples before fast Fourier transform (FFT). The filterbank analyzed the signal in 8 rectangular, non-overlapping, $\frac{3}{4}$ octave wide bands. 20ms attack and 100ms release time constants were used to compute the input levels. The gain table granularity was 1dB. The gain table was calculated according to the fitting rule [5] as described in the following section. In the subjective evaluation of the proposed algorithm, a commercial behind-the-ear (BTE) reference hearing aid (Siemens Motion 501DP) was used for comparison.

Gain prescription

The gains used in the proposed dynamic compressor aim to restore the function of the damaged outer hair cells (OHC) in the inner ear according to [5]. The basilar membrane input-output characteristics in humans has been measured for normal-hearing and hearing-impaired listeners [11, 13, 14, 12]. It is assumed that at low levels, i.e., below 25 dB SPL, the OHC amplify the signal by approximately 40 dB [12]. This amplification decreases with increasing level. At levels above about 87 dB the gain caused by the function of the OHCs is assumed to be zero dB. In hearing impaired ears, the maximum gain (at low levels) is assumed to be reduced. However, the compression rate ($\approx 1:4$) and the level at which the basilar membrane response equals the passive response remain unaffected.

To find the required gain for the dynamic compressor, the input level at the ear drum has to be found which creates the same basilar membrane response as normal hearing listeners have for the non-amplified input, i.e., on a basilar membrane input-output plot, the prescribed gain would be the level dependent horizontal difference between the healthy and the impaired basilar membrane input-output function. [12] measured the basilar membrane IO function only for a filter center frequency of 4 kHz. The filter effect of the outer and middle ear probably leads to different knee points in the input-output function. In this study the knee points are shifted according to the 80 phone loudness-contour to compensate for the outer and middle ear filter. Furthermore, the OHC gain is assumed to be frequency dependent resulting in reduced gains for lower and higher frequencies, as described in the gain prescription by [5] which was used here to precribe gains based on the audiogram and categorical loudness scaling.

The commercial hearing aid used in the subjective evaluation was fitted using the NAL-NL2 gain prescription rule [9].

Instrumental evaluation

In the instrumental evaluation, the experimental dynamic range compressor was fitted to a hearing loss representative for the subject group used in the subjective evaluation. The hearing threshold with loudness scaling data is shown in Figure 2.

Percentile gain analysis

A percentile gain analysis [8] was applied to estimate the effective gain. The input and output signal of the algorithms were time aligned and analyzed in third-octave filter bands. The filterbank output was cut into overlapping windows of 125 ms duration with 50% overlap. The percentile input level was calculated for the 30%, 50%, 65%, 95% and 99% percentile. The resulting input and output percentile levels are shown in the top panel of Fig. 3. In the next step all input windows have been selected whose level did not differ from a given percentile level by more than 3 dB. For those windows the applied gain was measured by dividing the output RMS by the input RMS level; the gain was averaged over this selection in the dB domain (Fig. 3, bottom panel).

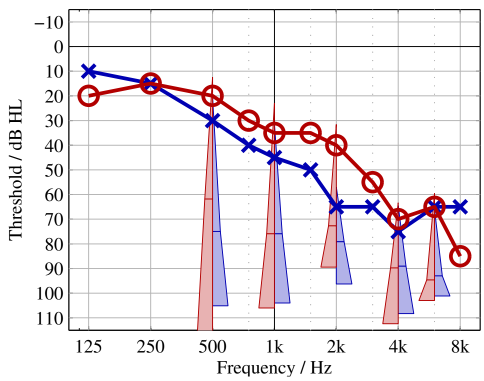

Figure 2: Underlying hearing loss used in the instrumental evaluation of the compressor. The wedges indicate categorical loudness scaling data used for the gain prescription according to [5].

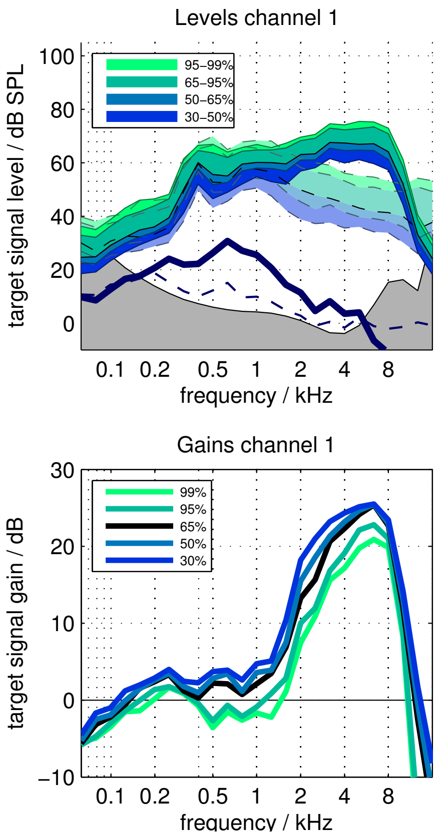

Figure 3: Percentile level and gain analysis using the international speech test signal [8]. Top panel: Target signal level of the processed (solid lines) and unprocessed signal. Bottom panel: Percentile gains.

The analysis of the percentile gains shows that with the proposed dynamic compressor, the prescribed gains are reached.

SNR and speech intelligibility prediction

Dynamic compression schemes do not only compress the signal dynamics, but also the signal to noise ratio (SNR): At positive SNR, the signal peaks typically belong to the target signal and are thus compressed, whereas at negative SNR the signal peaks belong to the noise signal. The compression of the SNR depends on the signal statistics of target and noise signal, and on the properties of the dynamic compressor.

In order to explore the SNR compression for the suggested dynamic compressor, the output SNR was recorded as a function of input SNR and for two different noise types, a fluctuating speech shaped noise, ICRA5-250 [3, 16], and a stationary speech shaped noise. To measure the output SNR, the mixed input signal was processed by the compressor, and the applied time- and frequency-dependent gains were recorded and applied to the target and noise signal separately in a parallel filterbank. The segmental output SNR [10] was calculated in 125-ms time segments, and limited to the range $[-20, 35]$ dB. The dB values were averaged over time. The difference to a linear amplification of the same signal is shown in the upper panel of Fig 4.

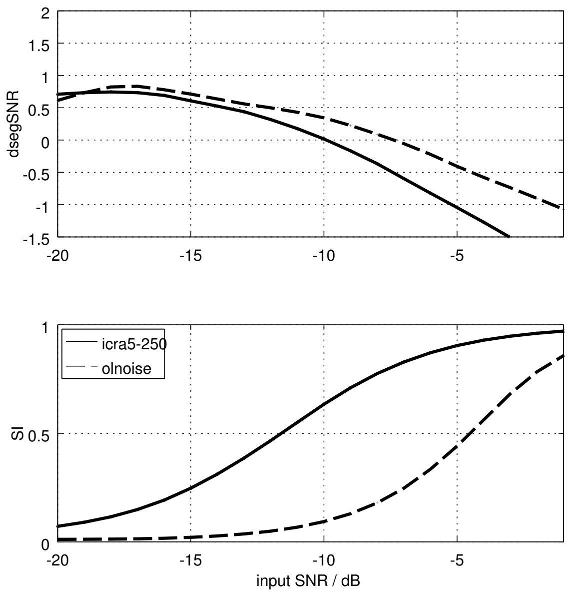

Figure 4: Segmental SNR [10] improvement (upper panel) and speech intelligibility [2] improvement (lower panel) as a function of input SNR, for stationary (olnoise) and fluctuating (icra5250) speech shaped noise.

For prediction of speech intelligibility, a auditory preprocessing based intelligibility model [2] was applied. The mixed output and the unprocessed clean speech signal are the inputs to the model.

Subjective evaluation

Speech reception thresholds (SRT) were measured using the OLSA sentence test [15] in stationary (olnoise) and fluctuating (ICRA5-250) speech shaped noise. 20 elderly hearing impaired listeners (average age 73.5 years, standard deviation 7.8 years) participated. All listeners were experienced hearing aid users.

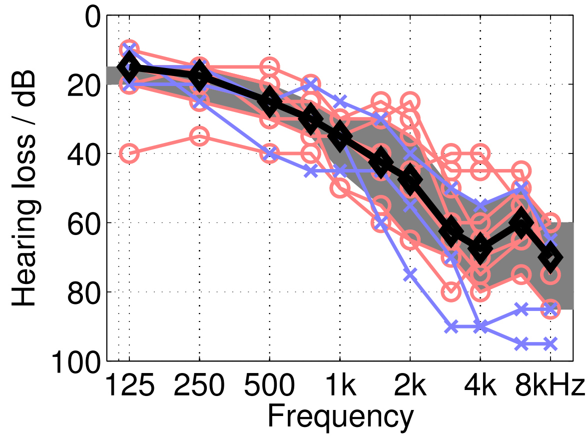

Figure 5: Audiograms of the test subjects, with average (black line) and inter-quartile range (shaded area). The better ear was used in the speech test.

The test was presented monaurally via headphones; the better ear was chosen for presentation. The individual hearing thresholds with the average and inter-quartile ranges is shown in Fig. 5 for the proposed experimental compressor ('prop.') and for the commercial hearing aid ('comm'). It is obvious that the SRTs are comparable for both dynamic compressors for all different test conditions. No statistically significant differences of the SRTs for the proposed experimental compressor and the commercial hearing aid were found.

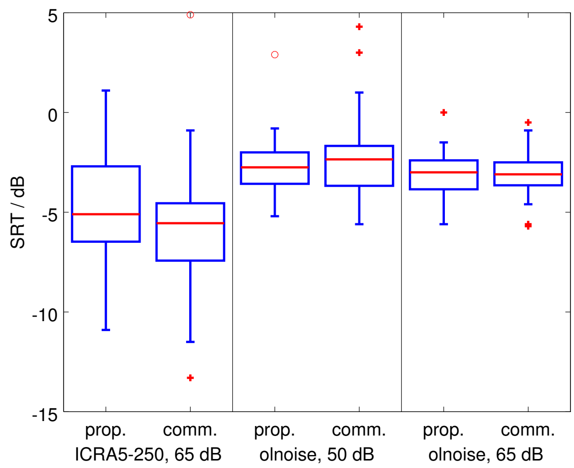

Figure 6: Speech reception thresholds of the OLSA sentence test, in fluctuating noise (left panel), stationary noise at 50 dB (center panel) and at 65 dB SPL (right panel).

Conclusions

An experimental dynamic range compressor was analyzed using instrumental measures, and was compared to a commercially available hearing aid in terms of speech intelligibility. The intrumental measures show that the desired algorithm behavior was achieved: Segmental

SNR analysis revealed an SNR benefit at negative SNRs. The percentile gain analysis (ISMADHA method) indivates that the prescribed gains were reached.

The subjective results show that the proposed algorithm yields comparable performance to a commercial hearing aid algorithm: No significant difference in speech reception thresholds between the commercial device and the proposed experimental algorithm were found in fluctuating and stationary noise.

The proposed dynamic compressor is thus suited as a framework for hearing aid algorithms such as single-channel noise reduction, or beamforming, providing multiband dynamic compression as key function of a research hearing aid.

Aknowledgments

This study was funded by BMBF 13EZ1127D.

References

[1] J. B. Allen. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Trans. Acoust., Speech, Signal Processing*, 25(3):235–238, June 1977.

[2] C. Christiansen, M. S. Pedersen, and T. Dau. Prediction of speech intelligibility based on an auditory preprocessing model. *Speech Communication*, 52(7-8):678 – 692, 2010.

[3] W. A. Dreschler, H. Verschuure, C. Ludvigsen, and S. Westermann. Icra noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. *International Journal of Audiology*, 40(3):148–157, 2001.

[4] S. Ernst, G. Grimm, and B. Kollmeier. Evaluation of binaurally-synchronized dynamic-range compression algorithms for hearing aids. In *Proceedings of Meetings on Acoustics*, volume 19, page 050085. Acoustical Society of America, 2013.

[5] S. Ewert and G. Grimm. Model-based hearing aid gain prescription rule. In *Speech perception and auditory disorders, Int. Symposium on Audiological and Auditory Research (ISAAR)*, Nyborg, Denmark, August 2011.

[6] S. D. Ewert, S. Kortlang, and V. Hohmann. A model-based hearing aid: Psychoacoustics, models and algorithms. In *Proceedings of Meetings on Acoustics*, volume 19, page 050187. Acoustical Society of America, 2013.

[7] G. Grimm, T. Herzke, D. Berg, and V. Hohmann. The Master Hearing Aid – a PC-based platform for algorithm development and evaluation. *Acta Acustica united with Acustica*, 92:618–628, 2006.

[8] I. Holube, S. Fredelake, M. Vlaming, and B. Kollmeier. Development and analysis of an international speech test signal (ists). *International journal of audiology*, 49(12):891–903, 2010.

[9] G. Keidser, H. Dillon, M. Flax, T. Ching, and S. Brewer. The nal-nl2 prescription procedure. *Audiology Research*, 1(1):24, 2011.

[10] P. Mermelstein. Evaluation of a segmental snr measure as an indicator of the quality of adpcm coded speech. *The Journal of the Acoustical Society of America*, 66(6):1664–1667, 1979.

[11] A. J. Oxenham and C. J. Plack. A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. 101(6):3666–3675, 1997.

[12] C. J. Plack, V. Drga, and E. A. Lopez-Poveda. Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss. *J. Acoust. Soc. Am.*, 115(4):1684–1695, April 2004.

[13] C. J. Plack and A. J. Oxenham. Basilar-membrane nonlinearity and the growth of forward masking. 103(3):1598–1608, 1998.

[14] C. J. Plack and A. J. Oxenham. Basilar-membrane nonlinearity estimated by pulsation threshold. 107(1):501–507, 2000.

[15] K. Wagener and T. Brand. Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters. *International Journal of Audiology*, 44(3):144–156, 2005.

[16] K. Wagener, T. Brand, and B. Kollmeier. The role of silent intervals for sentence intelligibility in fluctuating noise in hearing- impaired listeners. *International Journal of Audiology*, 45(1):26–33, 2006.