Compensation of hearing deficiencies in the inner ear

Volker Hohmann

Medizinische Physik, Universität Oldenburg

Keywords

auditory models, peripheral compression, dynamic compression, cochlear amplifier, cochlear hearing loss

Introduction

Hearing impairment of cochlear origin is generally thought to be mainly a consequence of a loss of outer haircells (see Moore (1995) for an overview). It causes a frequency-specific elevation of the hearing threshold level and a reduced dynamic range between threshold level and uncomfortable loudness level (‘Recruitment’-phenomenon). In order to compensate for Recruitment, single- or multi-frequency-band amplification and dynamic compression schemes are commonly used in hearing instruments. Established methods are Automatic Volume Control (AVC) and Automatic Gain Control (AGC). Whereas AVC basically aims at maintaining the long-term signal level at a comfortable loudness level, AGC involves shorter control time-constants that allow for changing the relative loudness levels of speech segments, i.e., increasing the audibility of weak components like unvoiced consonants relative to the high-level vowels. The prevailing view is that time-constants close to or shorter than the phoneme rate deteriorate hearing in many everyday listening conditions, because fast amplitude compression introduces nonlinear distortions, which are particularly prominent in case the AGC is performed independently in a large number of frequency bands (see Hansen (2002) and Herzke and Hohmann (2005), for an overview). In contrast to the detrimental effects found in fast multi-band AGC systems, the healthy auditory system provides a large number of overlapping frequency bands and almost instantaneous amplitude compression due to the function of the outer haircells (see, e.g., Moore (1995)). This suggests that fast or even instantaneous compression should in principle be possible without degrading speech intelligibility.

This contribution sheds some light on the possible reasons for the discrepancy between the physiological finding of the presence of instantaneous compression in the healthy auditory system and the audiological finding of the detrimental effect of fast multi-band compression in cases of cochlear hearing loss. In particular, a model of the instantaneous spectro-temporal compressive auditory processing is introduced, which might help to improve compression schemes in hearing instruments in the future.

Model of instantaneous spectro-temporal compressive processing

The cochlea acts as a non-linear spectro-temporal analyzer that codes the frequency-distribution of the incoming acoustic energy (frequency-to-place transformation). The nonlinearity in the cochlear response to sound is mediated by the outer haircells, which provide a compressive saturating gain characteristics, i.e., a high gain at low input levels that decreases with level (‘cochlear amplifier’). The exact way of how gain and compression is realized and how it is controlled is investigated by assessing the basilar membrane (BM) response to sounds in mammals using Mößbauer techniques (e.g., Robles et al. (1986)) or laser interferometry (e.g., Cooper and Rhode (1992)). These techniques revealed a compressive response to tones at a fixed place along the BM only when stimulated with the characteristic (or: best) frequency of that place (‘on-frequency tone’). The compression was found to be almost instantaneous (see van der Heijden (2005), for a detailed view) and is associated with a high gain of about 40-50 dB at low levels. However, gain and compression decrease with increasing deviation of the stimulus frequency from the best frequency towards low frequencies, and gain decreases with increasing deviation towards high frequencies, yielding a sharp frequency tuning of the place-specific response (e.g., Russell and Nilsen (1997)). This means that the response to ‘off-frequency tones’ linearizes and that the frequency resolution of the cochlear amplifier is higher than that of the passive cochlear resonant response. Corresponding results have been found using auditory nerve recordings (e.g., Yates (1990)) and can also be found in psychophysical experiments with humans (e.g., frequency tuning (Patterson, 1976) and on-frequency compression (Plack et al., 2004)). Suppression is another important consequence of the cochlear nonlinearity. In two-tone experiments, an off-frequency tone presented simultaneously with an on-frequency tone suppresses the response to the on-frequency tone at its best place, with the amount of suppression depending on frequency separation and tone levels (e.g., Cooper (1996)). Suppression can also be measured using psychophysical methods in humans, e.g., pulsation thresholds (Duifhuis, 1980) and nonlinear growth of simultaneous masking (Oxenham and Plack, 1998). The reduction of cochlear gain and compression associated with a loss of outer haircells (cochlear hearing loss) can be measured both physiologically and psychophysically. In humans, masking data from on- and off-frequency forward masking (Nelson et al., 2001) provide evidence for a loss of gain and compression in the individual case (Plack et al., 2004).

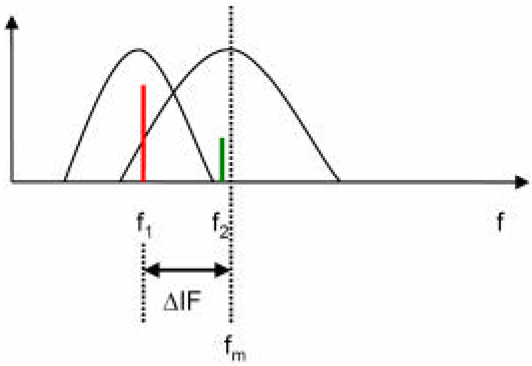

Figure 1: Sketch of the frequency response of an auditory filter with center frequency f_m (solid line, right filter) and spectrum of an on-frequency component with frequency f_2 (green bar) and an off-frequency component with frequency f_1 (red bar). On- and off-frequency components can be distinguished by measuring the instantaneous frequency at the output of the auditory filter and calculating its deviation DIF from the center frequency f_m. Measuring the level of a component at the output of the filter does not allow for this distinction. Note that the instantaneous frequency can be measured within a time period corresponding to only a few waveforms of the narrowband signal at the filter output.

In summary, the physiological and psychoacoustical findings reveal that the cochlear amplifier does not amplify and compress all spectral components of the input signal, but provides amplification and compression only for selected spectral components. Selection is determined by (i) the distinction between on- and off-frequency components and (ii) suppression effects. Time-constants of the selection process seem to be not more than a few waveforms of the respective on-frequency tone at each BM-position. In order to simulate the selection process, Hohmann and Kollmeier (2007) proposed to measure the instantaneous frequency in each frequency band of an auditory filterbank. The deviation of the inst. frequency from the bands respective center frequency is then used as a control signal for the amount of compression and gain: The higher the inst frequency deviation, the less gain and compression is applied. In this way, a direct measure of the relative prominence of off-frequency components relative to the on-frequency components in each filter-band is revealed with a very high time and frequency-resolution. Figure 1 shows the idea of using the deviation in instant. frequency as an indicator for the presence of off-frequency components. Figure 2 shows a block diagram of the nonlinear auditory filterbank that implements a compression stage in each filter-band that is steered by the deviation in inst. frequency. For model details see Hohmann and Kollmeier (2007).

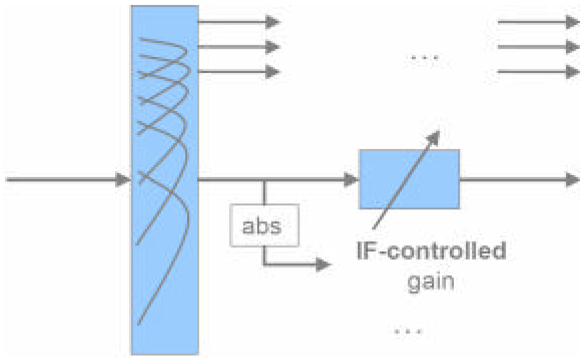

Figure 2: Block diagram of the auditory filterbank that simulates nonlinear spectro-temporal processing in the healthy cochlea (Hohmann and Kollmeier, 2007). At the output of a linear filterbank (complex-valued Gammatone filterbank (Hohmann, 2002)) that simulates the passive resonant properties of the basilar membrane, a compression stage provides instantaneous gain and compression in each filter-band (simulation of the outer haircell function). For this, the level is measured ('abs'-block), from which a gain is derived and applied to the signal. The amount of gain and compression is steered by the deviation of the inst. frequency from the respective bands center frequency ('IF-controlled gain').

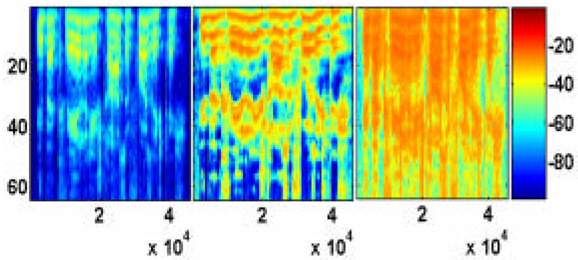

Figure 3: Excitation pattern, i.e., response of the nonlinear auditory filterbank of Figure 2 to a speech signal (one sentence). Plotted is the magnitude of the response in dB (color-coded) re. full scale as a function of time in samples and frequency (number of channel from low (top row) to high frequencies (bottom row). Left panel: no gain and compression (passive filterbank); right panel: full gain and compression in all filter-bands; middle panel: gain and compression with instantaneous frequency control.

Results

Figure 3 shows the excitation pattern, i.e., the magnitude in dB of the response of the model as a function of frequency and time for a speech signal (one sentence). The left panel shows the response of the passive model, i.e., without any gain and compression. It shows the response of a passive basilar membrane that is purely determined by the resonant properties. The right panel shows the output with full gain and compression in all filter-bands at all times. As expected, the model smears out the spectro-temporal pattern of the signal and destroys the spectro-temporal contrasts. On the other hand, the model with instantaneous frequency control (middle panel) provides gain for the prominent peaks in the spectro-temporal pattern only and thus sharpens the pattern. At a certain instant of time, only a few of the auditory filters get gain and thus compression. Most of the bands do not receive full gain and compression because the energy found in these bands is more determined by off-frequency components. The selection process is very fast, i.e., gain and compression is switched on within a few waveforms of the signal after on-frequency components become prominent in a certain frequency band.

Discussion

Given that cochlear hearing loss results in a linearization of the basilar membrane response due to the (partial) loss of the cochlear amplifier, a model-based approach to compensating for this type of loss should provide gain and very fast (even instantaneous) compression in many frequency bands. However, detrimental effects of fast sub-band compression in terms of perceived speech quality and speech intelligibility, and a strong interdependence between number of frequency bands, amount of compression and control time constants of the AGC have been found in many studies (e.g., Hansen (2002)). As outlined above, one possible reason for this finding might be that the spectro-temporal characteristics of the nonlinear response of the healthy cochlea is not properly accounted for in the compressor design. In particular, current multi-band fast compression systems provide independent compression in all frequency bands and thus neither simulate the linear response to off-frequency signals nor the suppression effects found in the healthy cochlea. Because both effects lead to a linearization of the response to wideband signals and to a spectral sharpening of the response to complex stimuli like speech, they might be required for successfully applying instantaneous compressive processing in frequency sub-bands. The aim of the work presented in this contribution was therefore to put forward the model-based approach to the rehabilitation of cochlear hearing loss. A signal processing technique was developed that allows for a simulation of a realistic cochlear response pattern including the most fundamental effects of cochlear nonlinearity. Based on this nonlinear auditory filterbank, an instantaneous compression scheme can easily be developed. Further work will reveal its applicability to hearing-aid processing.

Acknowledgment

Work supported by DFG and BMBF. Thanks to Birger Kollmeier and the members of the Medical Physics group for continuous support and discussions. Special thanks to T. Bisitz for drawing Figures 1 and 2.

References

Moore BCJ (1995) Perceptual Consequences of Cochlear Damage. Oxford University Press, Oxford.

Hansen M. (2002) Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. Ear Hear, 23, 369-80.

Herzke, T., Hohmann, V. (2005) Effects of instantaneous multi-band dynamic compression on speech intelligibility. EURASIP JASP 2005(18): 3034-3043.

Robles, L., Ruggero, M. A., and Rich, N. C. (1986) Basilar membrane mechanics at the base of the chinchilla cochlea I. Input-output functions, tuning curves, and phase responses. J. Acoust. Soc. Am. 80, 1364–1374.

Cooper, N.P., Rhode, W.S. (1992) Basilar membrane mechanics in the hook region of cat and guinea-pig cochleae: sharp tuning and nonlinearity in the absence of baseline position shifts. Hear Res. 63(1-2):163-90.

van der Heijden, M. (2005) Cochlear gain control. J. Acoust. Soc. Am. 117, 1223.

Russell, I.J., Nilsen, K.E. (1997) The location of the cochlear amplifier: Spatial representation of a single tone on the guinea pig basilar membrane. Proc. Natl. Acad. Sci. USA Vol. 94, pp. 2660-2664.

Yates, G.K. (1990) Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions. Hear Res. 50(1-2):145-62.

Patterson, R. (1976) Auditory filter shapes derived with noise stimuli. J. Acoust. Soc. Am. 59, 640.

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

Cooper, N. P. (1996) Two-tone suppression in cochlear mechanics. J. Acoust. Soc. Am. 99, 3087–3098.

Duifhuis, H. (1980) Level effects in psychophysical two-tone suppression. J. Acoust. Soc. Am. 67, 914-927.

Oxenham, A.J. and Plack, C.J. (1998) Suppression and the upward spread of masking. J. Acoust. Soc. Am. 104, 3500-3510.

Nelson, D. A., Schroder, A. C., and Wojtczak, M. (2001) A new procedure for measuring peripheral compression in normal-hearing and hearing impaired listeners.’ J. Acoust. Soc. Am. 110, 2045–2064.

Hohmann, V., Kollmeier, B. (2007) A nonlinear auditory filterbank controlled by sub-band instantaneous frequency estimates. In: Hearing – From Sensory Processing to Perception – Proc. Int. Symp. on Hearing - ISH 2006 (Kollmeier et al., Eds.), Springer.

Hohmann, V. (2002) Frequency analysis and synthesis using a Gammatone filterbank. Acustica / acta acustica 88(3): 433-442.