AN INTEGRATED APPROACH TO HEARING AID ALGORITHM DESIGN FOR ENHANCEMENT OF AUDIBILITY, INTELLIGIBILITY AND COMFORT

Bert de Vries and Rob A.J. de Vries

{bdevries,rdevries}@gnresound.com
GN ReSound, Dept. of Research and Core Technology
Horsten 1, 5612 AX Eindhoven, the Netherlands

ABSTRACT

Over the past decade, signal processing for industrial hearing aids has evolved from analog electronics for amplitude compression to sophisticated digital systems that integrate multiple adaptive algorithms for beamforming, feedback cancellation, amplitude compression, noise reduction and signal quality enhancement. Commonly, these modules are designed independently and consequently the opportunities for undesired interactive effects are abound. In this paper we present an approach to integrate adaptive modules for amplification, amplitude compression, noise reduction and signal quality (distortion) control.

1. INTRODUCTION

Over the past few years, much research and development has been spent on developing 'feature-rich' hearing aids. As a result, in today's hearing aid algorithms, multiple adaptive modules interact in unpredictable and sometimes undesirable ways. In this paper we present an approach to integrate gain function modules for loudness restoration, noise reduction and distortion control.
In the next section, we analyze the hearing aid signal processing goals. The signal processing modules are developed in Section 3.

2. THE HEARING AID SIGNAL PROCESSING GOAL

We focus on the most common (and interesting) hearing aid (HA) patient who suffers from Sensorineural Hearing Loss (SNHL). The hearing problem for a SNHL patient extends beyond conductive loss (due to mechanical malfunctioning in the peripheral ear) to include neural pathology in (but not limited to) the cochlea.

2.1. The Audibility Problem

Conductive losses (a.k.a. audiometric hearing loss), such as caused by defects to the eardrum or mechanical transducers (incus, malleus, stapes), are commonly measured as a function of frequency by a hearing threshold test (the audiogram, see Figure 1). Our audible range of input power level is frequency dependent and lower bounded by the *Hearing Threshold* (HT) and limited from above by an *Uncomfortable Level* (UCL), where audibility transgresses into pain sensation. As we grow older, our hearing threshold increases in particular for the higher frequencies, while the maximal level of comfortable input power level is much less affected. The compensating signal processing remedy is straightforward: we need amplification, but generally more for soft (and high-frequency) signals (to make them louder than the HT) than for loud (and low-frequency) signals (to keep them below the UCL). In principle, these two signal processing filters, frequency-dependent *amplification* and *compression*, resolve the audibility problem for the hearing impaired patient. In order to estimate the optimal amount of amplification and compression for a patient, it is a reasonable objective to restore (normal hearing) loudness perception, although this is a point of discussion in the audiological community.

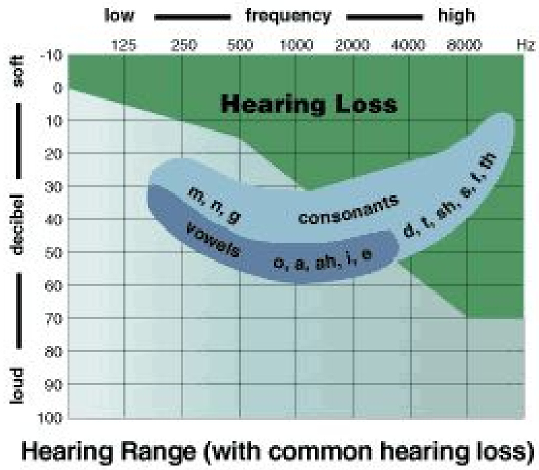

Figure 1: *The audiogram measures hearing loss as a function of frequency.* (source: www.phonak.co.nz).

2.2. The Intelligibility Problem

The primary task of a hearing aid is to make the signal audible. After all, great noise suppression and low signal distortion is of no use if the signal is not made audible to the patient. That said, even if the signal is (made) audible, the SNHL patient will perform worse on average than a normal hearing person on speech-in-noise intelligibility tests. Psycho-acoustic research has revealed that intelligibility scores, i.e., the recognition rate of words, is almost completely determined by the signal-to-noise ratio's in frequency bands of 'critical' bandwidth [1]. Effectively, due to cochlear damage, the SNHL patient needs a higher input signal-to-noise ratio (SNR) than a normal hearing person in order to recognize the same fraction of words in a sentence, even if the input signal is completely audible. The amount of necessary SNR improvement for a patient in order to resemble the scores for the average NH population has been named *SNR Loss* by Killion, in analogy to the term hearing loss for the audibility problem [2]. The signal processing strategy to combat SNR loss is obviously to improve the signal-to-noise ratio of the received signal, either by reducing the interfering noise or by speech (or music) enhancement.

2.3. Distortion

Unfortunately, signal processing agents for amplification, compression and noise reduction may negatively affect the quality of the signal. For instance, too much amplification leads to whistling (feedback) of the instrument. Too fast compression leads to an annoying distortion often referred to as noise modulation. In practice, the targeted compression gain is therefore smoothed by low-pass filters that are parameterized by attack and release times. In general, there is an engineering trade-off between restoring loudness and intelligibility versus 'tolerable perceptual signal distortion' for the patient. Nobody but the patient himself knows what is meant by 'tolerable distortion'. Based on the foregoing discussion, we can now state the hearing aids signal processing goal as follows:

The hearing aids signal processing goal is to maximize loudness restoration and intelligibility while keeping the signal distortion perceptually acceptable.

3. SIGNAL PROCESSING MODULES

3.1. Notation and Definitions

We assume that the hearing aid receives a signal $x(t) = s(t) + n(t)$, where $s(t)$ is a meaningful signal such as speech or music, $n(t)$ a noise signal and t a (discrete) time index. The received signal $x(t)$, which we will call *xignal* to distinguish it from signal $s(t)$, gets analyzed

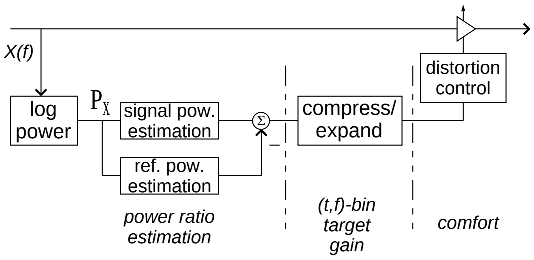

Figure 2: *Integrated gain module for audibility, intelligibility and comfort enhancement.*

every few milliseconds into a time-frequency representation $X(t, f)$. The indices t and f are often dropped from equations if their values are the same for all terms. The specifics of the frequency analysis is not of interest in this paper, but an approximate critical band resolution over the frequency range is assumed.
The power spectral density (PSD) is defined as $P_{ab}(t, f) \equiv \mathcal{E}[A(t, f)B^*(t, f)]$ and for P_{aa} we shortly write P_a. If the signal $s(t)$ and noise $n(t)$ are uncorrelated, then $P_{sn}(t, f)$ equals zero and consequently $P_x = P_s + P_n$. The script-letters $\mathcal{G} \equiv 20 \log(G/G_{ref})$ and $\mathcal{P} \equiv 10 \log(P/P_{ref})$ are used to indicate that we are working in the log-domain. This paper is about designing an appropriate linear time-varying filter $G(t, f) = G(\hat{P}_x(t, f))$ that processes $X(t, f)$ to $Y(t, f) = G(t, f)X(t, f)$. The hearing aid output $y(t)$ is obtained by frequency synthesis over $Y(t, f)$.

3.2. Design Approach

We will now develop a gain agent $G(t, f)$ that integrates compressive amplification and noise reduction with explicit facilities for signal distortion control. The system is split into three subsystems (fig. 2). The first subsystem estimates power ratios for each (t, f)-bin. In the second subsystem, the 'optimal' integrated gain for the (t, f)-bin is computed. In the last subsystem, signal distortion is controlled by smoothing target gains over time and neighboring bins.

3.3. Loudness Restoration

The hearing range for the normal hearing person is bounded from below by the (frequency-dependent) hearing threshold $\mathcal{P}_T^n(f)$ and from above by the uncomfortable level or pain threshold $\mathcal{P}_U(f)$. The hearing impaired subject suffers from an increased hearing threshold $\mathcal{P}_T^i(f) \geq \mathcal{P}_T^n(f)$. For computational simplicity, we assume that the uncomfortable level is the same for both impaired and normal hearing subjects. If we assume that all powers and gains are in the log-domain, then the power of the processed signal is given by $\mathcal{G} + \mathcal{P}_x$. For linear loudness restoration, we apply a gain $\mathcal{G}_l$ such that the relative log-power above threshold is the same for impaired and normal listeners, i.e.,

$$\frac{(\mathcal{G}_l + \mathcal{P}_x) - \mathcal{P}_T^i}{\mathcal{P}_U - \mathcal{P}_T^i} = \frac{\mathcal{P}_x - \mathcal{P}_T^n}{\mathcal{P}_U - \mathcal{P}_T^n} \tag{1}$$

Solving for $\mathcal{G}_l$ leads to

$$\mathcal{G}_l(t,f) = \beta(f)[\mathcal{P}_u - \mathcal{P}_x] \tag{2}$$

where we introduced the *compression slope* $\beta = \frac{\mathcal{P}_T^i - \mathcal{P}_T^n}{\mathcal{P}_U - \mathcal{P}_T^n}$. It follows from $\mathcal{P}_T^n \leq \mathcal{P}_T^i \leq \mathcal{P}_U$ that $0 \leq \beta \leq 1$. The first term in (2), $G_a = \beta \mathcal{P}_U$, implies a fixed linear amplification and the second term $G_c = -\beta \mathcal{P}_x$ compresses the xignal. Thus, loudness restoration leads to a *compressive amplification* filter. In the linear domain, eqn. 2 corresponds to $G_l = (\frac{P_U}{P_x})^{\beta/2} = (XUR)^{-\beta/2}$, where $XUR = P_x/P_U$ is the xignal-to-UCL power ratio. While alternative loudness correction criteria are plentiful, the general scheme remains estimation of a xignal-to-'reference' ratio followed by a compressive gain function.

3.4. Noise Reduction

Single-microphone based noise reduction algorithms in hearing aids are usually based on the Wiener filter. Consider the signal model $X = S + N$ where S and N are uncorrelated (i.e., $P_x = P_s + P_n$) and the filter $Y = GX = GS + GN$. In a hearing aid, traffic noise should still sound like traffic noise after processing, so we are only interested in partial noise reduction. Define the target signal $Z = S + \gamma N$, where $0 < \gamma < 1$. The difference between the target and output signal is the residual $R = Z - Y = (1 - G)S + (\gamma - G)N$ with PSD $P_r = (1-G)^2 P_s + (\gamma - G)^2 P_n$, since $P_{sn} = 0$. The gain that minimizes P_r is found by solving

$$\partial P_r / \partial G = 0 \tag{3}$$

which leads to the (Wiener) noise reduction filter

$$G_n = \frac{P_s + \gamma P_n}{P_s + P_n} = \frac{SNR + \gamma}{SNR + 1}. \tag{4}$$

Since $SNR \geq 0$, we find $G_n \geq \gamma$ and recognize γ as the minimal gain. For growing signal-to-noise ratio, the gain G_n increases. In general, noise reduction involves estimation of a signal-to-noise ratio followed by an expansive gain function.

3.5. Integrating Loudness Restoration and Noise Reduction

The goal of the noise reduction agent G_n is to improve the SNR of the xignal. It is pertinently *not* the aim of the

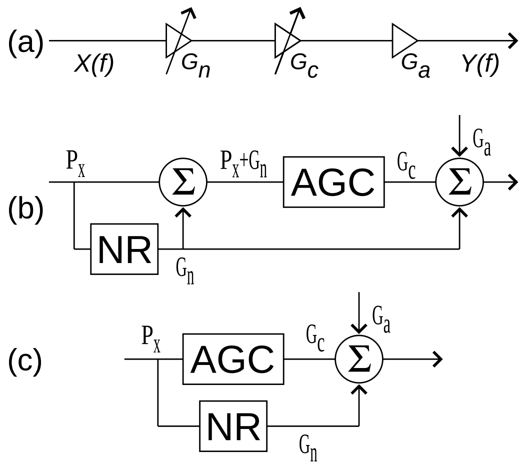

Figure 3: *Combining noise reduction and loudness restoration agents. See text for comments. (All gains and powers in log-domain.)*

noise reduction (NR) agent to affect the loudness. However, without corrective action, the NR gains in eqn. 4 do affect loudness. Similarly, the loudness restoration gains G_c and G_a are meant to affect the perceived loudness only. The inevitable site-effect on SNR is not a design goal and therefore essentially a processing artifact.

How do we integrate a noise reduction (NR) and loudness restoration (LR) gain agent into a single consistent structure? The LR criterion of eqn. 1 and the NR criterion eqn. 3 together specify two equations for one variable (the gain in band f). Thus, we cannot satisfy both criteria in the same frequency band. In order to find a solution for an integrated gain, we need to relax the rules a bit. Next, we discuss two techniques for combining NR and LR gains.

3.5.1. No Compromises on Loudness Restoration

It seems appropriate to worry more about correct loudness restoration than an accurate noise reduction gain, since the cleanest signal is worthless if it is not made audible. Thus, we choose to restore loudness after noise reduction processing, as is displayed in fig. 3a. This structure shows that the compressor technically works on the (NR-processed) signal $G_n(f)X(f)$ rather than $X(f)$. Correct gain agent integration therefore requires that the power estimate $\mathcal{P}_x + \mathcal{G}_n$ is made available to the compressor, as shown in fig. 3b. Note that the compressor cannot take responsibility for accurate loudness restoration anymore, if the compressor would precede or be placed parallel to the NR agent (as in fig. 3c).

3.5.2. *Relaxing Frequency Resolution*

In the previous section we removed the constraint that both rules need to be solved simultaneously: the noise reduction agent is applied straight to the received signal that has not been compensated for loudness (yet).

There are alternative ways to relax the prescriptions in order to find an integrated solution. We will now sketch an approach based on relaxing the frequency resolution of the solution. Although we wish to restore the perceived loudness, it may not be necessary that this is accomplished by restoring the specific loudness in each frequency band. Let us relax the loudness restoration criterion by requiring that the (combined) loudness in two neighboring bands (say, indexed by f and $f+1$) is restored by criterion eqn. 1. If we assume that the total power in bands f and $f+1$ is equal to the sum of the powers in the individual bands, then it is possible to write this criterion in terms of the original power and gain estimates for bands f and $f+1$. The loudness restoration criterion over the combined bands f and $f+1$ can then be written as

$$\frac{\log \frac{(G^2(f)+G^2(f+1))(P_x(f)+P_x(f+1))}{P_T^i(f)+P_T^i(f+1)}}{\log \frac{P_U(f)+P_U(f+1)}{P_T^i(f)+P_T^i(f+1)}} = \frac{\log \frac{P_x(f)+P_x(f+1)}{P_T^n(f)+P_T^n(f+1)}}{\log \frac{P_U(f)+P_U(f+1)}{P_T^n(f)+P_T^n(f+1)}} \tag{5}$$

This is an equation with two unknowns: $G(f)$ and $G(f+1)$. Similarly, the noise reduction (Wiener) criterion is relaxed to apply for the combined powers over two neighboring frequency bands and rewrite the equation in terms of the individual powers (and gains) for bands f and $f+1$. This provides a second equation for the variables $G(f)$ and $G(f+1)$, which both can now be solved for exactly. This procedure provides a method for simultaneously complying with both the loudness restoration and noise reduction gain rules on a half-resolution frequency grid. Apart from a loss of frequency resolution (which may not be any problem if we have enough frequency bands), a downside is that we need to solve two complicated equations for each frame. The upside is of course that accurate loudness restoration and noise reduction gains are obtained.

3.6. Distortion Control

For each frame and each frequency bin, the compression and NR agents produce gains $G_c(t,f)$ and $G_n(t,f)$, respectively. These gains may be regarded as *target gains* for the (t,f)-bin. The criteria for noise reduction and loudness restoration (eqs. 1 and 3) are based on power estimates for the (t,f)-bin only. The time- and frequency-varying behavior of the gains causes signal distortion.

Consider the system $Y(t) = G(t)X(t)$. There would be no signal distortion only if $G(t)$ were a constant. In that case, $G(t) = G(t-1)$, and we would have the (distortionless) output $G(t-1)X(t)$. Now subtract the distortionless output from the actual output to get the signal distortion $Y_d(t) = [G(t) - G(t-1)]X(t)$ that is a caused by the time-varying nature of G.

Spectral distortion (colorization) occurs if the gains vary over the frequency bands. We need a lowpass filter to soften the fluctuations of $G(t,f)$ if the distortion in $y(t)$ is perceptually intolerable. Such can be achieved by a spectro-temporal leaky integrator,

$$\overline{G}(t,f) = G(t,f) - \lambda[\partial_t G(t,f)] - \mu[\partial_f G(t,f)] \tag{6}$$

where $\overline{G}$ is the smoothed filter output and we defined the temporal *proposed gain fluctuation* $\partial_t G(t) \equiv G(t) - \overline{G}(t-1)$ and spectral proposed gain fluctuation as

$$\partial_f G(f) \equiv (-G(f-1) + 2G(f) - G(f+1))/3.$$

Note that eqn. 6 scales the temporal proposed gain fluctuation by means of the (temporal) distortion scaling factor $0 \le \lambda < 1$ according to

$$\underbrace{\overline{G}(t) - \overline{G}(t-1)}_{actual\,fluctuation} = (1-\lambda)\underbrace{(G(t) - \overline{G}(t-1))}_{proposed\,fluctuation}. \tag{7}$$

A similar argument holds for the spectral smoothing dimension. For $\lambda = \mu = 0$, the target gain passes unchanged through the filter; for $\lambda = \mu = 1$, maximal smoothing is obtained.

4. DISCUSSION

We have introduced a gain agent with parameter vector $\Omega = \{\beta, \gamma, \lambda, \mu\}$ for control of (the amount of) loudness restoration, noise reduction, temporal and spectral signal distortion, respectively. Two methods for integrating loudness restoration and noise reduction were presented. In practice, finding satisfying values for Ω is a very challenging task.

Acknowledgements. We wish to express our appreciation to our colleagues at the GN ReSound Algorithm R&D group for stimulating discussions.

5. REFERENCES

[1] J.B. Allen, "How do humans process and recognize speech," *IEEE Trans. on Speech and Signal Proc.*, vol. 2, no. 4, pp. 567–576, 1994.

[2] M.C. Killion, "Snr loss: I hear what people say, but i can't understand them," *Hearing Review*, vol. 4, no. 12, pp. 8–14, 1997.