EVALUATION OF TWO SPEECH AND NOISE ESTIMATION METHODS FOR THE ASSESSMENT OF NONLINEAR HEARING AIDS

Nicolas Ellaham†, Christian Giguère‡‡, Wail Gueaieb†‡

†School of Information Technology and Engineering, University of Ottawa, 800 King Edward Ave., Ottawa, ON K1N6N5
‡School of Rehabilitation Sciences, University of Ottawa, 451 Smith Rd., Ottawa, ON, K1H8M5

ABSTRACT

In this paper, we present a comparative evaluation of two speech and noise estimation methods commonly used with nonlinear hearing aids: the coherence function used for spectral estimation, and a noise phase-inversion scheme used to perform signal separation. The speech and noise spectra estimated at the output with both methods are compared for normal-hearing subjects and for hearing-impaired subjects using linear and nonlinear hearing aid processing. The spectra are found to be similar except for very low SNR conditions. However, Speech Intelligibility Index (ANSI S3.5-1997 [1]) estimates are relatively unaffected by differences in the spectra obtained with each method.

Index Terms— Hearing aids, nonlinear distortion, adaptive systems, speech processing, spectral analysis

1. INTRODUCTION

Hearing aids aim to improve the ability to understand speech in the presence of background noise. However, quantifying the benefits of a hearing aid system is a difficult challenge. While the signal-to-noise ratio (SNR) provides a rough indication, more sophisticated measures, such as the Speech Intelligibility Index (SII) (ANSI S3.5-1997 [1]) for instance, have become standard methods for evaluating these benefits. The SII accurately predicts the intelligibility of speech under linear processing in the presence of additive noise.

Hearing aids often contain nonlinear signal processing circuitry like the ones found in noise reduction algorithms (NRA) and automatic gain control (AGC). Such hearing aids have signal-adaptive characteristics that introduce nonlinear distortions at the output. Coherence measurements using broadband noise signals have become standard procedures to evaluate such systems (ANSI S3.42-1992 [2]). Kates and Arehart [3] derived an objective measure to predict speech intelligibility using the coherence function to estimate the speech and noise spectra at the output of the hearing aid. These estimates are used to calculate a signal-to-distortion ratio (SDR), which is incorporated into the standard SII computation using a three-level partitioning of the input and output signals. Their predictions using this method were consistent with intelligibility scores of normal-hearing and hearing-impaired subjects for conditions of additive noise, peak-clipping and center-clipping distortions.

In a study evaluating SNR improvements due to NRA and compression (AGC) in hearing aids, Hagerman and Olofsson [4] proposed a simple method to obtain estimates of the speech and noise signals at the hearing aid output. Their approach is based on performing two measurements using simultaneous presentation of speech and noise, one of them with the noise phase reversed. These estimates are used to compute the hearing aid’s equivalent speech and noise transfer functions, from which a frequency-weighted SNR is derived. Changes in the SNR thus obtained were found to correlate well with subjective intelligibility scores when using fast-acting compression hearing aids [5].

In this study, we compare the signal-separation potential of the coherence function and the method proposed in [4] under the same nonlinear hearing aid processing conditions for a wide range of SNRs. The goal is to identify a candidate for a future offline binaural objective measure of speech intelligibility. Normal-hearing subjects are considered, along with hearing-impaired subjects using three types of hearing aids: linear, with compression, and with NRA. The speech and noise spectra obtained with both methods are presented, and the impact of differences in spectral estimations on speech intelligibility is investigated using the SII.

2. METHODS

2.1. Material

The test materials used in the simulations presented in this paper are sentences taken from the Hearing-in-Noise-Test (HINT) [6]. The digitized sentences (re-sampled at 20 kHz) are arranged into 25 lists of 10 sentences. The HINT material also includes a noise signal with the same long-term spectrum as the average spectrum for each list. A software (MATLAB) simulator is used to implement hearing aid processing for impaired subjects. The software implements functions such as Volume Control, Spectral tilt, the NAL-RP gain rule [7], a single-channel compression amplifier, a wavelet-based noise reduction algorithm [8], and a hard-
Table 1: Hearing thresholds (dB HL) of hearing impaired subjects.

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>250</th>
<th>500</th>
<th>1k</th>
<th>2k</th>
<th>3k</th>
<th>4k</th>
<th>6k</th>
<th>8k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild-moderate</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>40</td>
<td>35</td>
<td>45</td>
<td>40</td>
<td>55</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>50</td>
<td>55</td>
<td>65</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>Severe</td>
<td>80</td>
<td>75</td>
<td>60</td>
<td>55</td>
<td>65</td>
<td>80</td>
<td>75</td>
<td>70</td>
</tr>
</tbody>
</table>

clipping output limiter. Audiograms of three hearing-impaired subjects were taken from a local database to perform the simulations (cf. Table 1).

2.2. Coherence method

The coherence function, also referred to as the magnitude-squared coherence (MSC), is a function valued between 0 and 1. Provided two stationary random processes \( x(n) \) and \( y(n) \), the MSC is computed according to equation (1):

\[
|\gamma(f)|^2 = \frac{|P_{xy}(f)|^2}{P_{xx}(f) \cdot P_{yy}(f)} \tag{1}
\]

where \( P_{xy}(f) \) is the cross-spectral density between \( x(n) \) and \( y(n) \), and \( P_{xx}(f) \) and \( P_{yy}(f) \) are the respective auto-spectral densities computed using Welch’s averaged periodogram method [9]. In this work, a 706-point Hamming window (approximately 35 ms) with 50% overlap is used to compute the spectra.

The MSC represents the fraction of \( y(n) \) that is linearly dependent on \( x(n) \), and is decreased by nonlinear distortions and additive noise. The complementary fraction \( 1 - |\gamma(f)|^2 \) represents the proportion of \( y(n) \) that is unrelated to \( x(n) \). With \( x(n) \) denoting the pure speech, and \( y(n) \) the distorted signal (with additive noise and hearing aid processing where applicable), estimates of the speech and noise power spectra are obtained using equations (2) and (3) respectively:

\[
\hat{P}_x(f) = |\gamma(f)|^2 \cdot P_{yy}(f) \tag{2}
\]
\[
\hat{N}(f) = \left[1 - |\gamma(f)|^2\right] \cdot P_{yy}(f) \tag{3}
\]

2.3. Phase-inversion method

In [4], signal separation is performed by presenting the speech and noise mixture to the hearing aid twice; the second time with the phase of the noise inverted. The input-output pairs for each presentation \((a_{in}/a_{out} \text{ and } b_{in}/b_{out})\) are summarized in equations (4) and (5):

\[
a_{in}(n) = u(n) + v(n) \quad a_{out}(n) = u'(n) + v'(n) + e_1(n) \tag{4}
\]
\[
b_{in}(n) = u(n) - v(n) \quad b_{out}(n) = u'(n) - v'(n) + e_2(n) \tag{5}
\]

where \( u(n) \) and \( v(n) \) denote the input speech and noise signals respectively, \( u'(n) \) and \( v'(n) \) the output speech and noise signals respectively, and \( e_1(t) \) and \( e_2(t) \) denote error signals resulting from nonlinear distortions introduced by the device. Assuming the error signals are sufficiently small, we can recover estimates of the output speech and noise signals using equations (6) and (7) respectively:

\[
c(n) = a_{out}(n) + b_{out}(n) = 2u'(n) + e_1(n) + e_2(n) \tag{6}
\]
\[
d(n) = a_{out}(n) - b_{out}(n) = 2v'(n) + e_1(n) - e_2(n) \tag{7}
\]

The speech and noise spectra are obtained from these estimates using Welch’s method, with the same windowing parameters used for the coherence method to simplify the comparison. The assumption that the error signals are small for nonlinearities typically used in hearing aids has been validated in [4] using a procedure for measuring nonlinear distortion based on the Hilbert transform, described in [10].

2.4. SII prediction

The SII is essentially calculated as a weighted average of the amount of information available to the listener over a chosen number of frequency bands. In this work, we use the third-octave band computational procedure to calculate the SII based on the speech and noise spectra obtained with each method. The hearing thresholds are interpolated over the third-octave band (logarithmic) frequency scale. The band audiability function for “short passages” is used in our computations (Table B.2 in ANSI S3.5-1997 [1]). The ANSI standard provides a detailed description of the calculation.

3. RESULTS

In this section, we present the estimated speech and noise spectra for normal-hearing subjects (Figure 1) and for one hearing-impaired subject using three types of hearing aids (Figure 2). The first sentence of the HINT test and an equal-length segment of the HINT speech-shaped noise are used for these experiments [6]. Figure 3 shows the average SII as a function of SNR, computed over the 10 sentences that make up the first list of the HINT test. For each hearing aid, the plots in Figure 3 show the results obtained for all three hearing-impaired subjects (cf. Table 1). In all experiments, the overall speech level is maintained at 65 dB SPL. The noise level is varied to achieve different SNR conditions.

3.1. Normal-hearing subjects

Figure 1 shows the speech and noise spectra recovered using the coherence and phase-inversion methods for normal hearing subjects at different SNRs. Without hearing-aid processing, the phase-inversion method reduces to the theoretical case, providing a basis for evaluating the coherence method. The noise spectra estimated with both methods (bottom panel) overlap perfectly for all SNR levels. For the speech spectra (top panel), however, the coherence method deviates from the theoretical estimates in low SNR conditions. For SNRs of -5 dB and -10 dB, small deviations are observed at high frequencies where the speech level is too low compared to the corresponding noise spectrum level.
At -30 dB SNR, the speech spectrum estimated using the coherence method is up to about 10 dB above the theoretical estimate at low frequencies, and 25 dB at high frequencies. Nevertheless, the effect of these deviations on speech intelligibility is negligible: the noise level is too high for the speech content to contribute to intelligibility. This is verified on the plot of SII vs. SNR for normal hearing subjects (top-left panel of Figure 3), which shows very similar SII predictions for both methods.

3.2. Hearing-impaired subjects

In Figure 2, we show the speech and noise spectra estimated at an SNR of 5 dB for a hearing-impaired subject using three types of hearing aids: linear, with compression, and with noise reduction (the NAL-RP gain-frequency rule is used in all three types). Results are presented for a subject with moderately-severe hearing loss. Similar plots were obtained for the other subjects in Table 1. Settings for inter-syllabic compression were used for the AGC hearing aid (threshold = 50 dB, ratio = 2:1, attack time = 5 ms, release time = 50 ms). For the hearing aid with noise reduction, the NRA strength was set to 50%. The results for the linear and compression hearing aids show a significant overlap between the spectra estimated using the coherence and phase-inversion method. The maximum deviation between the two methods (observed at the lower frequencies) falls within 2 dB. For the hearing aid with noise reduction, larger differences up to 5 dB are observed at high frequencies. However, the effect of these differences on the SII is small (less than 0.03 in Figure 3).

Figure 3 reveals a few additional observations. First, the benefit of compression compared to a linear hearing aid is shown by an improvement in speech intelligibility which increases with the severity of the hearing loss. Second, the SII plots for noise reduction seem to be similar to the linear case over the entire range of SNRs tested. However, a closer look at the figure reveals that: for very low SNRs, where the noise level is so elevated that it completely masks the speech, removing some of the noise does not affect speech intelligibility; on the other end of the scale (large SNR) the speech component is so dominant such that removing some of the noise has negligible effect on intelligibility. The range of SNR from -10 to 10 dB is most important, since any noise reduction in this range would affect speech intelligibility. In fact, a clear improvement in SII due to noise reduction can be observed over this range. For example, at 0 dB SNR an SII benefit of 0.1 is observed for moderately-severe hearing loss. This is equivalent to a 3.3 dB SNR improvement.
4. CONCLUSION

In this work, we compared the signal-separation potential of the coherence and phase-inversion methods under linear and nonlinear processing conditions commonly used in hearing aids. In the context of the SII, our results show the two methods to be alike. This study could be extended by testing the two methods under additional processing conditions: higher presentation levels that would result in peak-clipping, AGC with different compression ratios or time constants, and more aggressive noise reduction.

The aim of this work is not to design a new version of the monaural SII per se, but to identify a suitable signal-separation method that can be integrated in a future offline binaural objective measure to predict speech intelligibility. Such a measure requires access to the time-domain version of the separated speech and noise signals at the two hearing aids. The phase-inversion method requires two passes of the signals through the hearing aid, and is thus suitable for offline estimations. The present study shows that it could be an attractive solution for the purpose set forth.

5. ACKNOWLEDGEMENT

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Graduate Scholarship in Science and Technology (OGSST).

6. REFERENCES