Computers in Biology and Medicine 46 (2014) 29–41

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

A complementary low-cost method for broadband noise reduction in hearing aids for medium to high SNR levels

Márcio Holsbach Costa *

Department of Electrical Engineering, Federal University of Santa Catarina, 88040-900 Florianópolis, SC, Brazil

ARTICLE INFO

Article history:
Received 13 June 2013
Accepted 18 December 2013

Keywords:
Hearing-aids
Noise reduction
Adaptive filter
Speech processing
Noise cancelling

ABSTRACT

This work presents a complementary broadband noise reduction scheme for hearing aid applications. It is designed to attenuate uncorrelated and small-correlation-length acoustic noise with controlled speech distortion. Noisy speech signals are pre-processed by the proposed strategy before being subjected to an existing narrowband noise reduction system. The clean speech signal is estimated by a convex combination of the unprocessed speech signal and the output of a linear predictor. The convex combination coefficient is adjusted to provide noise suppression while avoiding significant unvoiced utterance distortions. The proposed method is optimized to minimize speech mean-square prediction-error. A low-cost adaptive implementation is proposed and compared to the conventional adaptive linear predictor showing an improved performance, as predicted by theory. Four different objective quality measures and subjective assessment performed by normal hearing volunteers indicate that the combined use of the proposed technique with a narrowband noise reduction system consistently improves speech quality for a range of signal to noise ratios. Low-cost digital hearing aids that make use of the conventional adaptive predictor for broadband noise reduction can be easily modified to incorporate the new proposal with a minimum amount of extra computational resources.

1. Introduction

Hearing-aids are essential devices for the social integration of people that suffer from hearing limitations or neurosensorial losses. These conditions affect about 9% of the world's population [1,2].

Hearing-aids are very complex systems consisting of several processing units that perform tasks such as adaptive directionality, noise reduction, dynamic compression and feedback cancellation. These elements interact to improve intelligibility and provide better acoustic comfort for the user. Due to their small size and power consumption requirements, the availability of computational resources for each subsystem is restricted. As a result, each technique should be designed as sparingly as possible.

Despite the great advancements in this area, two main causes of sensorial discomfort – noise amplification and reverberation – still persist [3]. One of the major complaints of hearing aid users is poor speech intelligibility due to background noise. Many studies have demonstrated that hearing impaired people need an SNR-50[1] from 10 to 30 dB higher than that required for the non-impaired [4].

The basic function of a noise reduction system is to lessen the user's perception of environmental acoustic noise, minimizing distortion and masking effects[2]. With noise reduction, acoustic comfort increases while fatigue decreases, which in turn increases the equipment's acceptability[3] [5,6].

Although multi-microphone hearing-aids may have many advantages over single microphone devices[4], some commercial gadgets are still equipped with the latter [7,8].

The most common single microphone noise reduction approaches are [9]: (a) subspace decomposition [10], (b) statistical and parametric modelling [11–13], and (c) Wiener filtering. The first two approaches, although feasible, incur considerable computational complexity (even when look-up tables are used to alleviate the effort) and large time delay due to signal processing [14]. Wiener filtering has also been applied to hearing-aids; however, it tends to generate unpleasant

* Corresponding author. Tel.: +55 48 3721 9506; fax: +55 48 3721 9280.
E-mail address: costa@eel.ufsc.br

[1] SNR-50 is the signal to noise ratio needed for the comprehension of 50% of the speech in a conversation.

[2] Distortion happens when the hearing impaired perceives sound but its intelligibility is compromised due to the lack of high frequency information. The problem is worsened by the existence of background noises, resulting in the complaint that "it is possible to hear, but not understand, speech" [2]. Moreover, masking decreases speech redundancy; as a result, small amounts of noise can lead to significant intelligibility degradation [50].

[3] Although noise reduction systems can provide a substantial improvement in speech quality (acoustical comfort), the same effect does not necessarily occur upon intelligibility. Intelligibility improvements can be obtained by speech enhancement systems [16,51], whose aim is not to reduce noise, but rather highlight the contrast between vowels and consonants [1].

[4] Multi-microphone techniques take advantage of the spatial separation among acoustic sources [1], but they require a considerable distance among microphones. This condition makes them inappropriate for ear-canal devices [52].

0010-4825/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiomed.2013.12.009

acoustic artefacts called "musical noise" [3]. Without exception, all techniques present a trade-off between noise reduction and speech distortion [9,15,16].

The limited computational resources available in commercial devices greatly limit the development of new techniques for hearing-aid improvement. In recent years, manufacturers have provided specific hardware to lessen this problem. An example of a very successful commercial architecture for the (software/hardware) implementation of a noise reduction system can be found in Ref. [17]. In this architecture, the digitized acoustic signal is split into different frequency channels; each of them is then subjected to independent attenuation factors before signal reconstruction. Signals from high signal to noise ratio (SNR) channels are preserved, while low SNR channel signals are attenuated. This approach, introduced in Ref. [18], works well only for narrowband noise. For broadband noise, all channels tend to suffer approximately the same attenuation, maintaining the same global SNR.

In Ref. [19], an extensive analysis of background noise databases showed that not all daily-life noises can be referred to as narrowband (low-frequency) background noises. As a result, the authors of Ref. [19] suggest that hearing aids should not only be fine-tuned to the individual audiogram but also to environmental conditions. In fact, in Ref. [19] it was shown that noise in environments such as industry and nature preponderantly present flat spectrum without temporal modulations. In work conditions (industry), hearing-aid users cannot reduce volume due to the possibility of warning sounds [20]. Consequently, many hearing-impaired workers are often forced to endure a certain degree of discomfort. In addition, [21] stated that the most difficult listening situations that commonly face persons with hearing loss feature broadband competition.

Some attempts to overcome the broadband noise problem in hearing-aids can be found in Refs. [22,23]. This work will focus on broadband acoustic noise characterized by small correlation-length[5] (such as low-pass noise sampled at near-Nyquist frequency rates[6]).

The conventional linear adaptive predictor (CLAP) [24–26] is a low-complexity solution that performs well in reducing broadband noise. However, it also tends to cancel uncorrelated speech components, which constitute about 20 to 25% of natural speech in the English language [27]. As a result, it produces muffled speech sounds and musical noise which can severely affect both speech intelligibility and naturalness. Some works have recently addressed the design of practical low distortion broadband noise cancellers based on the CLAP structure. In Ref. [28], a weighted sum of contaminated signal and CLAP output was proposed. This approach aims to enhance the quasi-stationary components of speech (voiced sounds), improving intelligibility and, secondarily, SNR. However, many intelligibility problems can be attributed to poor comprehension of unvoiced sounds. In Ref. [29], a first attempt was made to control the CLAP attenuation of uncorrelated speech components. However, the authors were not successful in accurately determining the optimal control parameter due to the use of very restrictive theoretical assumptions. In Ref. [30], CLAP output and error signals were linearly combined using attenuation factors directly related to instantaneous SNR. This approach provides poor results when unvoiced speech and uncorrelated noise occur simultaneously. Hence, low-cost reduction of uncorrelated noise remains an open issue of great interest for hearing-aids designers.

This work proposes a complementary low-cost technique for broadband noise reduction in hearing-aids for pre-processing of noisy-speech signals before narrowband noise reduction. Clean speech is estimated using a convex combination of the original contaminated signal and the output of a linear predictor. The convex combination weight factor establishes a trade-off between uncorrelated noise reduction and unvoiced speech distortion. An adaptive version of the algorithm is proposed. The proposed system is fitted to hearing-aid applications due to three main requirements: (a) availability of a narrowband noise reduction system to alleviate acoustical discomfort due to CLAP coefficient fluctuations and to reduce narrowband noise; (b) small signal processing time delay (since 6 to 8 ms delays can be undesirably perceived by users, while over 10 ms can be considered annoying [1]); and (c) low extra computational cost (in addition to the processing load previously existent in the hearing aids processing system). The problem is mathematically described in Section 2. Section 3 briefly reviews the prediction of a signal immersed in noise. Section 4 presents the proposed method and its optimization strategy. Section 5 presents a simple adaptive implementation of the proposed technique. Section 6 shows simulation results using synthetic and real speech signals. The results obtained corroborate the theoretical derivations and illustrate the performance of the proposed method. Final conclusions are presented in Section 7[7]. The proposed method would be particularly useful for commercial devices in which it would be used together with an existing narrowband noise reduction system [17,18,31]. Throughout this text, bold uppercase and lowercase letters represent matrices and vectors, respectively, while italics represent scalars.

2. Problem description

The sampled acoustic signal at time instant n is modelled as the sum of a speech signal $x(n)$ and noise $\eta(n)$, resulting in

$$y(n) = x(n) + \eta(n). \tag{1}$$

here, noise $\eta(n)$ is assumed stationary, independent of $x(n)$, zero-mean with power σ_η^2 and with a small correlation-length so that $|E\{\eta(n)\eta(n-k)\}| \leq \varepsilon$ for $k \geq K$, where ε is a very small constant and K is a finite integer smaller than the speech correlation-length. Noise is white in the particular case of $K=1$ and $\varepsilon=0$.

Speech signal $x(n)$ has zero mean with power σ_x^2 and is modelled by an autoregressive process with a small correlation-length for unvoiced utterances or a large correlation-length for voiced utterances. The model coefficients are assumed constant in a given time window (about 20 ms).

The mean-square prediction-error (MSPE), resulting from predicting the clean speech $x(n)$ by the unprocessed (contaminated) speech $y(n)$, is given by

$$J_{US} = E\left\{[x(n) - y(n)]^2\right\} = \sigma_\eta^2, \tag{2}$$

where $E\{\cdot\}$ denotes statistical expectation.

3. Prediction of a signal immersed in noise

The Wiener filter is a widespread noise reduction technique that presents a known trade-off between speech distortion and noise reduction. Since noise and speech usually share the same frequency range, noise statistics are usually estimated during voice pauses. In practical applications, large time periods between estimations can lead to substantial degradation of the noise reduction process. However, assuming noise has a small correlation-length compared to the speech, a prediction approach can be used to continuously compute pseudo-optimum noise reduction filters without significant performance loss.

[5] The correlation length of a random signal $x(n)$, with exponential decaying autocorrelation function, is defined as $L_X = \sum_{l=0}^{\infty} E\{x(n)x(n-l)\}/E\{x^2(n)\}$ [32]. It measures the signal memory.

[6] This condition can be found in low-cost or severely limited computational systems, as is the case of hearing aid devices.

[7] Preliminary results were published in [53].