In-Ear Microphone Hybrid Speech Enhancement

Nils Westerlund
Nils.Westerlund@bth.se

Mattias Dahl
Mattias.Dahl@bth.se

Ingvar Claesson
Ingvar.Claesson@bth.se

Department of Telecommunications and Signal Processing
Blekinge Institute of Technology
Ronneby, Sweden

ABSTRACT
This paper presents a novel speech enhancement approach for performing noise reduction in severely disturbed environments. A small microphone for communication purposes is placed inside the external auditory canal to pick up the speech signal originating from the speech production organ. The speech enhancement is achieved by using three different noise reduction methods: High frequencies are attenuated by passive absorbers, low frequency components are attenuated by employing active noise control and finally a broadband noise reduction is achieved by using spectral subtraction.

KEY WORDS
Speech and Acoustic Processing, Active Noise Control, Spectral Subtraction, Ear-Mic.

1 INTRODUCTION

Many occupations of today requires the usage of personal preservative equipment such as a pair of ear-muffs to damp high sound pressure levels. At the same time, there is often a need to be able to communicate via some communication equipment. The primary concern with noise is not only the potential risk of damage to the hearing. Another effect caused by noise is the masking effect it has on speech, degrading speech intelligibility.

A conventionally placed communication microphone in front of the mouth is exposed to the surrounding noise. The sound originating from the vocal cords also propagates to the External Auditory Canal (EAC) via the bone structure in the head. This means that it is possible to pick up the speech inside a protected ear, instead of in front of the mouth.

However, in a personal communication situation, such as inter-com, it is desirable to obtain both protection for the near-end user and enhanced speech quality for the far-end user in noisy environments with high sound pressure levels. To achieve that, this paper presents an approach where three noise reduction methods are combined and applied to the speech signal inside the EAC:

- **Passive absorbers** are usually used to attenuate high frequency noise. One problem is that passive attenuation of low frequencies implies heavy and bulky absorbtion materials. A typical application where passive absorbers are used to attenuate noise, is a pair of ear-muffs [1].
- **Active Noise Control (ANC)** is an attractive complement to passive reduction of unwanted low frequency noise. ANC in headsets dates back to the 1950's, and a good active headset will effectively combine low frequency active attenuation with high frequency passive attenuation to provide high attenuation of the exterior noise at a wider frequency range [2]. ANC can be both feedforward [3], feedback [4] or a combination of both [5].
- **Spectral subtraction** is a non-linear, yet simple way of reducing unwanted broadband noise acoustically added to a speech signal [6]. The method estimates the magnitude frequency spectrum of the underlying clean speech by subtracting an estimate of the noise magnitude spectrum from the noisy speech magnitude spectrum. The noise estimate is obtained during non-speech activity. A number of enhancements of the original algorithm has been proposed during the years [7], [8], [9].

In a noisy environment, the combination of an in-ear microphone with both passive ear-muffs and ANC protects the near-end user from high sound pressure levels. Still, the speech signal picked up inside the ear will be contaminated by some reduced noise. A post processing spectral subtraction can enhance the speech quality and intelligibility before communicating to the far-end user. However, this does not directly protect the near-end user from harmful noise. Nevertheless, the near-end user can indeed benefit by the spectral subtraction since some individual speech monitoring feedback is necessary when using effective ear-muffs in an intercom system. Spectral subtraction significantly reduces the noise feedback in such a system.

Finally, in the proposed system the ANC also helps in equalizing the distortion of the speech signal caused by the bone conduction, i.e. the mouth-to-ear channel. This channel represents a fairly simple low pass filtering, which is to some extent equalized by the low frequency suppression of the ANC [10].

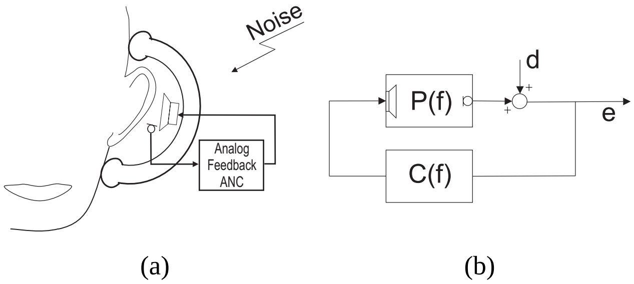

Figure 1. (a) Feedback ANC in an ear-muff. (b) The corresponding block scheme where $P(f)$ represents the channel transfer function, $C(f)$ represents the control transfer function, d is the primary noise and e is the resulting noise.

2 THE NOISE REDUCTION METHODS

In this section, a brief survey of the three noise reduction methods will be given.

In a pair of *passive ear-muffs*, a tight fit and good seal around the ear is crucial for the performance. To achieve this, the cushions needs to be soft and flexible. However, these cushions allow the ear-muff shell to vibrate when exposed to external sound and this vibration is perceived by the user. It has been shown that the passive attenuation of a pair of ear-muffs, behaves as a second order mechanical system [1], i.e the passive frequency response, $G_P(f)$, can be written as

$$G_P(f) = \frac{K_v}{K_v + K_c + j2\pi fR - (2\pi f)^2 M} \qquad (1)$$

where M is the shell mass, R is the cushion damping and K_v and K_c are the stiffness of the air inside the shell volume and the cushion stiffness, respectively.

Imperfect seal around the ear degrades low frequency attenuation and attenuation of higher frequencies are determined by the high frequency dynamics of the shell [2].

Feedback ANC implemented in a pair of ear-muffs, employs an internal error microphone placed inside the shell. The input signal from this microphone is phase shifted, weighted and fed back to a loudspeaker also placed inside the shell, in order to create a noise cancelling sound field, see Fig. 1a. In Fig. 1b, the corresponding feedback control system block scheme is shown, with $P(f)$ representing the channel, i.e. the response from the loudspeaker input to the microphone output, and $C(f)$ representing the analogue controller. The noise inside the ear-muff shell is denoted by d whereas e represents the error signal. The active response, $G_A(f)$, of the system to the right in Fig. 1 can then be written as [2]

$$G_A(f) = \frac{e}{d} = \frac{1}{1 - C(f)P(f)} \qquad (2)$$

The *spectral subtraction* algorithm makes some assumption regarding signal characteristics [6]. The noise is

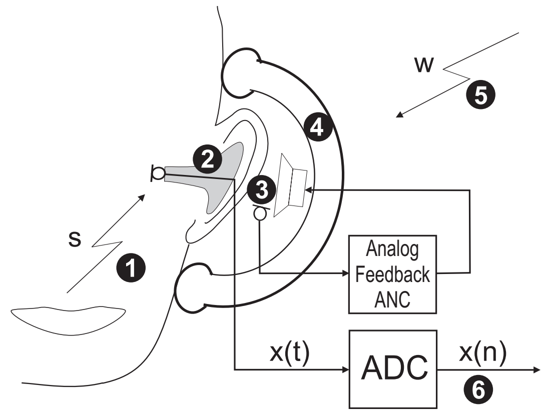

Figure 2. Measurement setup and signal paths. (1) Bone conducted speech signal. (2) Ear-mic inserted in custom-made acrylic earplug. (3) ANC reference microphone and loudspeaker. (4) Passive ear-muff. (5) Surrounding noise. (6) Digital output speech signal from ear-mic.

assumed to be additive and uncorrelated to the speech signal. A slow varying noise environment is acceptable as long as there is enough time to calculate a new estimate of the noise magnitude spectra and apply this estimate before the noise characteristics have changed. Furthermore, the speech is assumed to be short-time stationary.

Often the conventional spectral subtraction equation is written as

$$\hat{S}_N(f) = H_N(f) X_N(f) \qquad (3)$$

where $\hat{S}_N(f)$ is the N-point estimate of the clean speech magnitude spectra, $X_N(f)$ is the N-point noisy speech magnitude spectra and

$$H_N(f) = \left[1 - k \cdot \frac{|W_N(f)|^a}{|X_N(f)|^a} \right]^{\frac{1}{a}} \qquad (4)$$

where $W_N(f)$ is the N-point estimate of the background noise magnitude spectra calculated during non-speech activity. The function $H_N(f)$ is denoted the weighting function. The parameter a decides whether to use a power spectral subtraction ($a = 2$) or a magnitude spectral subtraction ($a = 1$ or some other preferred value). The parameter k adjusts the noise reduction. A larger k reduces the noise level more than a smaller k does, but the resulting speech will be more distorted.

3 METHOD

The measurement setup and signal paths are illustrated in Fig. 2. The speech signal, s, originating from the human speech organ propagates in the actual system mainly through the skull, causing the interior wall of the EAC to vibrate. The acoustic path, i.e. the air borne speech sound transmission, is assumed to be negligible due to the attenuation by the ear-muffs. The high frequency part of the

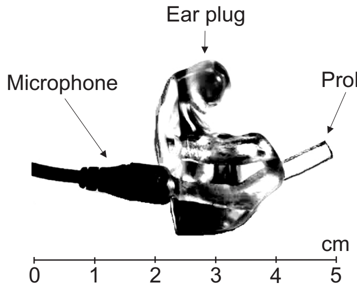

Figure 3. Custom-made acrylic ear-plug with microphone and microphone probe attached.

surrounding noise w is attenuated by the passive absorbers in the ear-muffs but the speech signal inside the EAC is still contaminated by the low frequency components of the noise that passive absorbers are unable to attenuate efficiently. However, the ANC frequency response, $G_A(f)$, attenuates those remaining noise components. The low frequencies of the speech signal inside the EAC are also attenuated by $G_A(f)$. The resulting speech signal, $x(t)$, is picked up by the ear-mic and properly sampled and bandlimited by the analog-to-digital converter (ADC) forming the digital output signal $x(n)$. Further noise reduction is obtained by applying the spectral subtraction weighting function, $H_N(f)$, to the signal.

4 EVALUATION AND RESULTS

A narrow canal was drilled in a custom-made, acrylic ear-plug and a thin microphone probe was inserted into this canal. The effects of using microphone probes are thoroughly investigated in [11]. A small-size microphone was attached to the other end of the probe, see Fig. 3. The ear-plug was inserted into the EAC of the test person and the ANC equipped ear-muffs were then fitted onto the head.

Disturbing noise is usually low pass, i.e. most of the noise energy is located at lower frequencies. Hence, two different representative noise environments were used: Helicopter rotor blade noise (exterior noise) and noise from a helicopter cockpit (interior noise). The interior and exterior helicopter noise are dominated by tonal components in the lower frequency range ($<$ 500 Hz).

The effect of the ANC in the ear-muffs becomes apparent when investigating the power spectral densities (PSD) of the signals inside the EAC. These PSDs, both for exterior and interior helicpter noise, are shown in Fig. 4. The ANC in the ear-muffs also increases the SNR. This is illustrated in Fig. 6.

Another positive side effect, is that the indirect high pass filtering that the ANC represents, results in increased speech quality when combined with an ear-mic. In addition, this indirect high pass filtering is practically delayless, an important property in communication equipment of today.

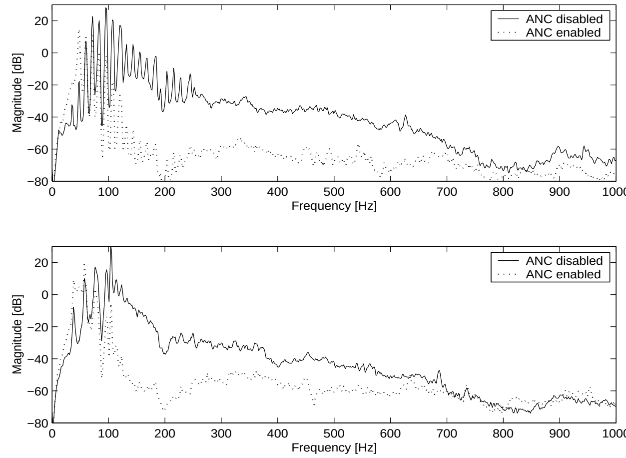

Figure 4. (Upper plot) The PSD of the signal inside the EAC with surrounding exterior helicopter noise. (Lower plot) The PSD of the signal inside the EAC with surrounding interior helicopter noise.

However, some background low frequency noise is still present in the speech signal inside the EAC. The spectral subtraction algorithm reduces this noise significantly and in addition, the two systems, i.e. the ANC and the spectral subtraction, overlaps in the frequency domain. The spectral subtraction algorithm described in section 2 were used with $a = 1$ and $k = 2$. The length of the FFTs were 128 and a hanning window was used to reduce the effects due to framing of the signal. Furthermore, the weighting function $H_N(f)$ was exponentially averaged to mitigate the effects of spectral subtraction musical tones caused by rapid fluctuations of the FFT bins.

The speech signal inside the EAC with interior helicopter noise is shown in Fig. 5. The noise suppressing effect due to the spectral subtraction is clear and a subjective listening test reveals that the speech distortion caused by the noise reduction is acceptable.

5 CONCLUSIONS

In a severely disturbed environment, with high surrounding noise sound pressure levels, an ear-mic combined with a pair of ear-muffs equipped with ANC is superior to a conventionally placed microphone in front of the mouth. Not only is the microphone protected from the surrounding noise both by passive absorbers and by the ANC, but also protected against mechanical damage.

Furthermore, a spectral subtraction algorithm, even in its simplest form, significantly reduces the remaining background noise level.

Altogether, the hybrid speech enhancement system combining an ear-mic, passive and active noise reduction and a spectral subtraction algorithm, offers several advantages in an environment severely disturbed by noise and the method is widely applicable in numerous situations.

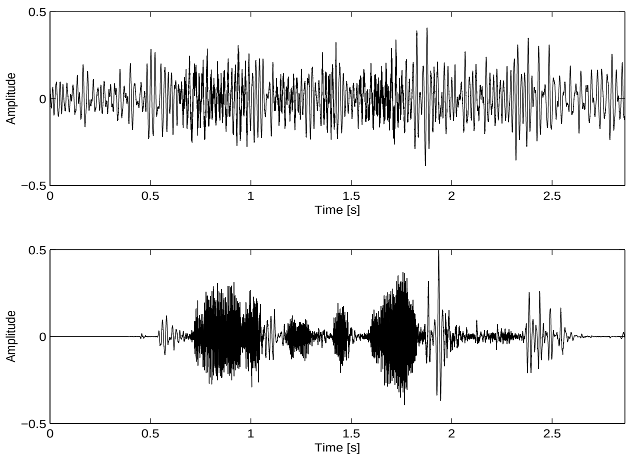

Figure 5. (Upper plot) In-ear speech signal before spectral subtraction. (Lower plot) In-ear speech signal after spectral subtraction.

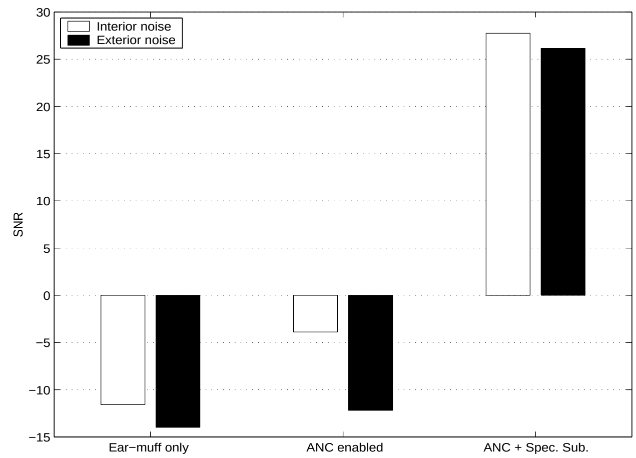

Figure 6. The SNR improvement for both interior and exterior helicopter noise.

References

[1] E. A. G. Shaw and G. J. Thiessen, "Acoustics of circumaural earphones," *Journal of the Acoustical Society of America*, vol. 34, no. 9, pp. 1233–1246, 1962.

[2] B. Rafaely, "Active noise reducing headset — an overview," in *Proc. of Internoise*, August 2001.

[3] S. M. Kuo and D. R. Morgan, *Active noise control systems*, John Wiley & Sons, Inc., New York, 1st edition, 1996.

[4] B. Rafaely, J. Garcia-Bonito, and S. J. Elliot, "Feedback control of sound in headrests," in *Proc. of Active '97*, 1997, pp. 445–456.

[5] M. Winberg, S. Johansson, T. Lagö, and I. Claesson, "A new passive/active hybrid headset for a helicopter application," *International Journal of Acoustics and Vibration*, vol. 4, no. 2, pp. 51–58, 1999.

[6] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust. Speech and Sig. Proc.*, vol. ASSP-27, pp. 113–120, April 1979.

[7] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoust. Speech and Sig. Proc.*, vol. ASSP-32, pp. 1109–1121, December 1984.

[8] O. Cappé, "Elimination of the musical noise phenomenon with the ephraim and malah noise suppressor," *IEEE Trans. Speech and Audio Proc.*, vol. 2, no. 2, pp. 345–349, 1994.

[9] H. Gustafsson, S. Nordholm, and I. Claesson, "Spectral subtraction using reduced delay convolution and adaptive averaging," *IEEE Trans. Speech and Audio Proc.*, vol. 9, no. 8, 2001.

[10] N. Westerlund, M. Dahl, and I. Claesson, "In-ear microphone equalization exploiting an active noise control," in *Proc. of Internoise*, August 2001, Invited paper.

[11] P. A. Hellström and A. Axelsson, "Miniature microphone probe tube measurements in the external auditory canal," *Journal of the Acoustical Society of America*, vol. 93, no. 2, pp. 907–919, 1993.