PREDICTION-ERROR-METHOD-BASED ADAPTIVE FEEDBACK CANCELLATION IN HEARING AIDS USING PITCH ESTIMATION

Kim Ngo[1], Toon van Waterschoot[1], Mads Græsbøll Christensen[2], Marc Moonen[1], Søren Holdt Jensen[3] and Jan Wouters[4]

[1]Katholieke Universiteit Leuven, ESAT-SCD, Leuven, Belgium
[2]Aalborg University, Dept. of Media Technology, Aalborg, Denmark
[3]Aalborg University, Dept. Electronic Systems, Aalborg, Denmark
[4]Katholieke Universiteit Leuven, ExpORL, O. & N2, Leuven, Belgium

ABSTRACT

Acoustic feedback is a well-known problem in hearing aids, which is caused by the undesired acoustic coupling between the loudspeaker and the microphone. The goal of adaptive feedback cancellation (AFC) is to adaptively model the feedback path and estimate the feedback signal, which is then subtracted from the microphone signal. The main problem in identifying the feedback path model is the correlation between the near-end signal and the loudspeaker signal, which is caused by the closed signal loop. In this paper, a novel prediction-error-method (PEM)-based AFC is presented using a harmonic sinusoidal near-end signal model. Furthermore, the prediction error filter (PEF) is designed to incorporate a variable order and a variable amplitude next to a variable pitch. Simulation results for a hearing aid scenario indicate an improvement up to 6dB in maximum stable gain increase and up to 8dB improvement in terms of misadjustment.

1. INTRODUCTION

Acoustic feedback is a well-known problem in hearing aids, which is caused by the undesired acoustic coupling between the loudspeaker and the microphone. Acoustic feedback limits the maximum amplification that can be used in a hearing aid if howling, due to instability, is to be avoided. In many cases this maximum amplification is too small to compensate for the hearing loss, which makes feedback cancellation algorithms an important component in hearing aids [1] [2].

The goal of adaptive feedback cancellation (AFC) is to adaptively model the feedback path and estimate the feedback signal, which is then subtracted from the microphone signal. The main problem in identifying the feedback path model is the correlation between the near-end signal and the loudspeaker signal, which is caused by the closed signal loop. This correlation problem causes standard adaptive filtering algorithms to converge to a biased solution. The challenge is therefore to reduce the correlation between the near-end signal and the loudspeaker signal. Typically, there exist two approaches to this decorrelation [3], i.e., decorrelation in the closed signal loop and decorrelation in the adaptive filtering circuit. Recently proposed methods for decorrelation in the closed signal loop consist in the insertion of all-pass filters [4] in the forward path of the hearing aid or in clipping [5] of the feedback signal arriving at the microphone. Alternatively, an unbiased identification of the feedback path model can be achieved by applying decorrelation in the adaptive filtering circuit, i.e., by first prefiltering the loudspeaker and microphone signals with the inverse near-end signal model before feeding these signals to the adaptive filtering algorithm [6], [7]. The near-end signal model and the feedback path model can be jointly estimated using the prediction-error-method (PEM). For near-end speech signals, a linear prediction (LP) model is commonly used [6]. For audio signals a pole-zero LP, warped LP or a pitch prediction model cascaded with a LP model have been proposed [7].

In [8], different frequency estimation techniques, namely pitch estimation methods [9] and constrained pole-zero LP (CPZLP) [10], were compared in PEM-based AFC, where simulation results showed an improvement in AFC performance, when pitch estimation were used. The main difference is that pitch estimation relies on harmonicity, i.e., the sinusoidal frequencies are assumed to be integer multiples of a fundamental frequency, whereas CPZLP estimates the sinusoidal frequencies independently.

In [8], only the pitch were used in the prediction error filter (PEF) design leading to infinite suppression of the sinusoids. This may not be the optimal solution since speech generally can be considered voiced or unvoiced, resulting in different amplitudes and number of harmonics. In this paper, a novel PEM-based AFC approach is presented using a harmonic sinusoidal near-end signal model. Furthermore, the PEF is designed to incorporate a variable order and a variable amplitude next to a variable pitch. Simulation results for a hearing aid scenario indicate a significant improvement in terms of misadjustment and maximum stable gain increase, compared to previous work on PEM-based AFC using CPZLP.

The paper is organized as follows. Section 2 reviews the PEM-based AFC concept. Section 3 describes the near-end signal model that is used. Section 4 explains the PEF design. In Section 5, simulation results are presented. The work is summarized in Section 6.

2. ADAPTIVE FEEDBACK CANCELLATION (AFC)

The AFC scheme is shown in Fig.1. The microphone signal

This research work was carried out at the ESAT Laboratory of Katholieke Universiteit Leuven, in the frame of the EST-SIGNAL Marie-Curie Fellowship program (http://est-signal.i3s.unice.fr) under contract No. MEST-CT-2005-021175, the Concerted Research Action GOA-MaNet, Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011, and the K.U.Leuven Research Council CoE EF/05/006 Optimization in Engineering (OPTEC). The scientific responsibility is assumed by its authors.

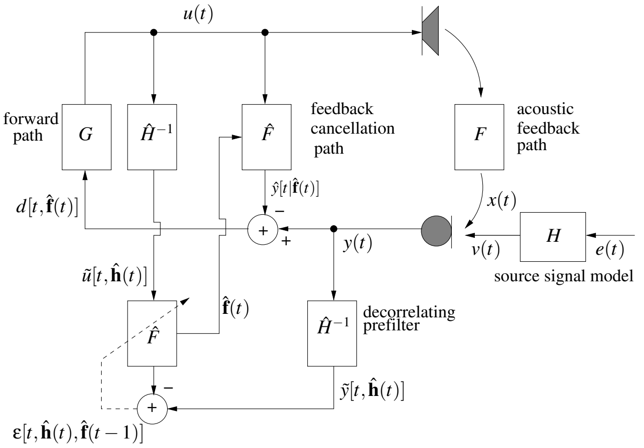


Figure 1: AFC with decorrelating prefilters in the adaptive filtering circuit.

is given by

$$y(t) = v(t) + x(t) = v(t) + F(q,t)u(t) \tag{1}$$

where q denotes the time shift operator and t is the discrete time variable. $F(q,t)$ is the feedback path between the loudspeaker and the microphone, $v(t)$ is the near-end signal, $x(t)$ is the feedback signal. The forward path $G(q,t)$ maps the microphone signal $y(t)$, possibly after AFC, to the loudspeaker signal $u(t)$. The aim of the AFC is to place an estimated finite impulse response (FIR) adaptive filter $\hat{F}(q,t)$ in parallel with the feedback path, having the loudspeaker signal as input and the microphone signal as the desired output. The feedback canceller $\hat{F}(q,t)$ produces an estimate of the feedback signal $x(t)$ which is then subtracted from the microphone signal $y(t)$. The feedback-compensated signal is given by

$$d(t) = v(t) + [F(q,t) - \hat{F}(q,t)]u(t). \tag{2}$$

The main problem in identifying the feedback path model is the correlation between the near-end signal and the loudspeaker signal, due to the forward path $G(q,t)$, which causes standard adaptive filtering algorithms to converge to a biased solution. This means that the adaptive filter does not only predict and cancel the feedback component in the microphone signal, but also part of the near-end signal, which results in a distorted feedback-compensated signal $d(t)$.
An unbiased identification of the feedback path model can be achieved by applying decorrelation in the adaptive filtering circuit, i.e., by first prefiltering the loudspeaker and microphone signals with the inverse near-end signal model $\hat{H}^{-1}(q,t)$ (see Fig.1) before feeding these signals to the adaptive filtering algorithm. The near-end signal model and the feedback path model can be jointly estimated using the PEM [3], [6], [7]. The PEM delivers an unbiased estimate of the feedback path coefficient vector $\mathbf{f}(t) = [f_0(t) \quad f_1(t) \quad \dots \quad f_{n_F}(t)]$, by minimization of the prediction error criterion

$$\min_{\mathbf{f}(t)} \sum_{k=1}^{t} \varepsilon^2(k) \tag{3}$$

if the prediction error is calculated as

$$\varepsilon(t) = H^{-1}(q,t)\,[y(t) - F(q,t)u(t)] \tag{4}$$

where $H(q,t)$ is a linear model for the source signal $v(t)$, i.e.,

$$v(t) = H(q,t)r(t) \tag{5}$$

where $r(t)$ is an uncorrelated signal.

3. NEAR-END SIGNAL MODEL

In this paper, the goal is to use a harmonic sinusoidal near-end signal model instead of a LP model in PEM-based AFC, such that the sinusoids are assumed to have frequencies that are integer multiples of a fundamental frequency ω_0, i.e., $\omega_n = n\omega_0$. This follows naturally from voiced speech being quasi-periodic.

3.1 Harmonic sinusoidal near-end signal model

The near-end signal $v(t)$ is assumed to consist of a sum of real harmonically related sinusoids and additive noise,

$$v(t) = \sum_{n=1}^{P} a_n \cos(n\omega_0 t + \phi_n) + r(t), \tag{6}$$

where $\omega_0 \in [0,\pi]$ is the fundamental frequency, a_n the amplitude, and $\phi_n \in [0,2\pi)$ the phase of the nth sinusoid, and $r(t)$ the noise which is assumed to be autoregressive (AR), i.e., $r(t) = \frac{1}{C(q,t)}e(n)$, with

$$C(q,t) = 1 + \sum_{i=1}^{n_C} c^{(i)}(t)q^{-i}. \tag{7}$$

We should stress that, in constrast to the approach in [8], none of the signal model parameters $(P, a_n, \omega_0, \phi_n, C(q,t))$ are assumed to be known but are estimated, as explained next.

3.2 Near-end signal model parameter estimation

The pitch estimation technique used here is based on optimal filtering (optfilt) of the feedback-compensated signal $d(t)$, which ideally corresponds to the near-end signal $v(t)$. The idea behind pitch estimation based on filtering is to find a set of filters that pass power undistorted at the harmonic frequencies $n\omega_0$, while minimizing the power at all other frequencies. This filter design problem can be stated mathematically as [9]

$$\min_{\mathbf{h}} \mathbf{h}^H \mathbf{R} \mathbf{h} \quad \text{s.t.} \quad \mathbf{h}^H \mathbf{z}(n\omega_0) = 1, \quad \text{for} \quad n = 1, \dots, P, \tag{8}$$

where $\mathbf{h}$ is the length-N impulse response of the filter, $\mathbf{z}(\omega) = [e^{-j\omega_0} \; \dots \; e^{-j\omega(M-1)}]$ and $\mathbf{R}$ is the covariance matrix defined as

$$\mathbf{R} = \mathrm{E}\{\tilde{\mathbf{d}}(t)\tilde{\mathbf{d}}^H(t)\}, \tag{9}$$

where $(\cdot)^H$ denotes Hermitian transpose and $\tilde{\mathbf{d}}(t)$ is a vector containing M consecutive samples of the analytical counterpart of the feedback-compensated signal $d(t)$ [9]. Using the Lagrange multiplier method, the optimal filters can be shown to be

$$\mathbf{h} = \mathbf{R}^{-1}\mathbf{Z}\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\mathbf{1} \tag{10}$$

with $\mathbf{1} = [1 \quad \dots \quad 1]^T$ and $\mathbf{Z} = [\mathbf{z}(\omega_0) \quad \dots \quad \mathbf{z}(P\omega_0)]$ the Vandermonde matrix containing the sinusoids. This filter is signal adaptive and depends on the unknown fundamental frequency. Intuitively, one can obtain a fundamental frequency estimate by filtering the signal using the optimal filters for various fundamental frequencies and then picking the one for which the output power is maximized, i.e.,

$$\begin{aligned} \hat{\omega}_0 &= \arg\max_{\omega_0} \mathrm{E}\{|\mathbf{h}^H \tilde{\mathbf{d}}(t)|^2\} \\ &= \arg\max_{\omega_0} \mathbf{1}^H \left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1} \mathbf{1}. \end{aligned} \tag{11}$$

This method has demonstrated to have a number of desirable features, namely excellent statistical performance and robustness against periodic interference [9]. Once ω_0 is known, the amplitude of the sinusoids can be estimated using a least squares approach:

$$\hat{\mathbf{a}} = \left(\mathbf{Z}^H \mathbf{Z}\right)^{-1} \mathbf{Z}^H \mathbf{d} \tag{12}$$

with $\hat{\mathbf{a}} = [\hat{a}_1 \quad \dots \quad \hat{a}_P]$. Finally, the number of harmonics P can be determined by using a maximum a posteriori (MAP) criterion [9] [11],

$$\hat{P} = \arg\min_{P} M \log \hat{\sigma}_P^2 + P \log M + \frac{3}{2} \log M \tag{13}$$

where the first term is a log-likelihood term which comprises a noise variance estimate that depends on the candidate model order, the second term is the penalty associated with the amplitude and phase, while the third term is due to the fundamental frequency. The last model parameters that need to be estimated are the AR parameters of the noise component $r(t)$, which is straightforward using LP of the output signal of the first PEF $H_1^{-1}(q,t)$, see Section 4.

4. PREDICTION ERROR FILTER DESIGN

4.1 PEF for sinusoidal components

It is well known that a sum of P sinusoids can be described exactly using an all-pole model of order $2P$, with mirror symmetric LP coefficients. However, it has been shown that the all-pole model is not exact when noise is added, and in this case a pole-zero model of order $2P$ should be used [12]. Still, by constraining the poles and zeros to lie on common radial lines in the z-plane, the number of unknown parameters in the pole-zero model can be limited to P and the LP parameters can be uniquely related to the unknown frequencies [10]. The PEF for the sinusoidal components can hence be written as a cascade of second-order sections:

$$H_1^{-1}(q,t) = \prod_{n=1}^{P} \frac{1 - 2\nu_n \cos n\omega_0 z^{-1} + \nu_n^2 z^{-2}}{1 - 2\rho_n \cos n\omega_0 z^{-1} + \rho_n^2 z^{-2}} \tag{14}$$

where the poles and zeros are on the same radial lines, with the poles positioned between the zeros and the unit circle, i.e., $0 \ll \rho_n < \nu_n \leq 1$.

In [8], the PEF in (14) was designed with the zero radii fixed to $\nu_n = 1$ and the pole radii fixed to $\rho_n = 0.95$, and the order fixed to $P = 15$. For an example speech frame, with a spectrum shown in Fig.2, this design leads to the PEF response shown in Fig.3 (when CPZLP is used for frequency

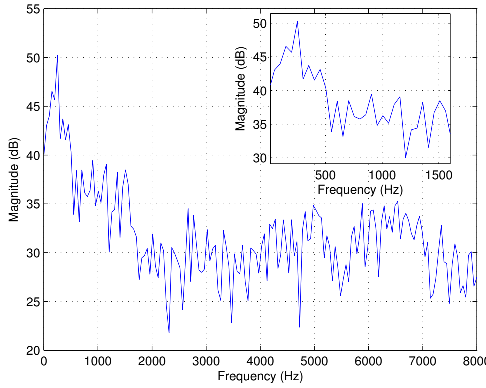


Figure 2: Speech spectrum used to estimate the PEF.

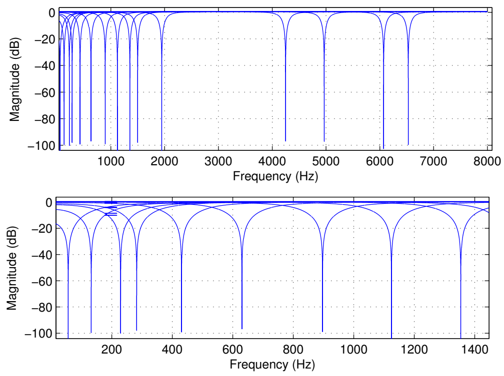


Figure 3: PEF using CPZLP for frequency estimation, (top) notch filters up to 8000Hz and (bottom) notch filters up to 1400Hz.

estimation) and Fig.4(top, when pitch estimation is used). A first observation is that the PEF applies equal (infinity) suppression for all frequencies when all the zeros are placed on the unit circle. The PEF using pitch estimation in Fig.4(top) shows that the PEF has a more dense structure in the low frequency region when harmonicity is assumed.

Pitch and variable order estimates are straightforward to include in the PEF, by setting $\omega_0 = \hat{\omega}_0$ and $P = \hat{P}$. From the design of the PEF it is clear that the zero radius determines the notch depth, i.e., the inverse of estimated amplitudes. Including the amplitude in the PEF then follows from the design rule in [13], i.e.,

$$\nu_n = \max\left(\rho_n, 1 - \frac{1-\rho_n}{\hat{a}_n}\right). \tag{15}$$

Including the amplitude estimated in (12) and the pitch estimated in (11) the PEF is shown in Fig.4(bottom) which shows a more signal dependent behaviour, when comparing to the corresponding speech spectrum shown in Fig.2.

In previous work [8], besides assuming infinite notch depth, the model order is also assumed to be equal for every speech frame. This may not be the optimal solution since speech generally can be considered voiced or unvoiced, resulting in different amplitudes and number of harmonics. A histogram

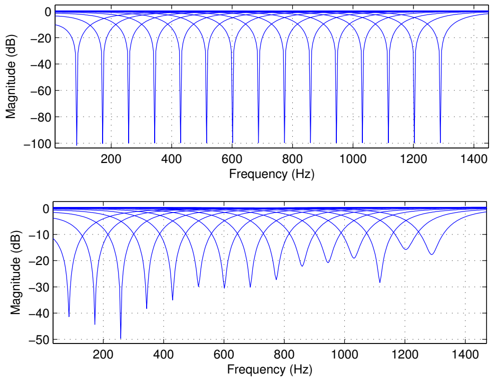


Figure 4: PEF using optimal filtering based pitch estimation (top) without amplitude and (bottom) with amplitude.

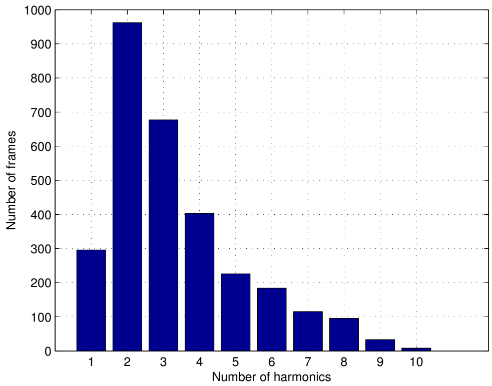


Figure 5: Number of harmonics $\hat{P}$ for each frame.

of the estimated number of harmonics using (13) for the speech signal used in the evaluation in Section 5 is shown for different frames in Fig.5. This indeed suggests, that the harmonic sinusoidal near-end signal model order varies across different frames and that the fixed model order of $P = 15$ used in [8] indeed is too high compared to the estimated model order $\hat{P}$. The prediction error using a cascaded near-end signal model can then be written as

$$\varepsilon(t) = H_2^{-1}(q,t)H_1^{-1}(q,t)\left[y(t) - F(q,t)u(t)\right] \quad (16)$$

with the PEF for the noise component $r(t)$ defined as $H_2^{-1}(q,t) = \hat{C}(q,t)$.

5. EVALUATION

In this section, simulation results are presented in which different PEF designs are compared in PEM-based AFC with cascaded near-end signal models in a hearing aid setup. The harmonic sinusoidal near-end signal model order is evaluated for different fixed orders, i.e., P = 15, 10, 5 and compared with a variable order. The effect of including a variable amplitude in the PEF is also illustrated. The near-end noise model order is fixed to $n_C = 30$. Both near-end signal models are estimated using 50% overlapping data windows of length $M = 320$ samples. The optimal filtering length is set to $N = \frac{M}{4}$. The NLMS adaptive filter length is set equal to the acoustic feedback path length, i.e., n_F = 200 (measured hearing aid feedback path). The near-end signal is a 30 s male speech signal sampled at f_s= 16 kHz. The forward path gain $K(t)$ is set 3 dB below the maximum stable gain (MSG) without feedback cancellation.

To assess the performance of the AFC algorithm the following measures are used. The achievable amplification before instability occurs is measured by the MSG, which is defined as

$$\text{MSG}(t) = -20\log_{10}\left[\max_{\omega\in\mathscr{P}} |J(\omega,t)[F(\omega,t) - \hat{F}(\omega,t)]|\right] \quad (17)$$

where $J(q,t) = \frac{G(q,t)}{K(t)}$ denotes the forward path transfer function without the amplification gain $K(t)$, and $\mathscr{P}$ denotes the set of frequencies at with the feedback signal $x(t)$ is in phase with the near-end signal $v(t)$. The misadjustment between the estimated feedback path $\hat{\mathbf{f}}(t)$ and the true feedback path $\mathbf{f}$ represents the accuracy of the feedback path estimation and is defined as

$$\text{MA}_F = 20\log_{10}\frac{||\hat{\mathbf{f}}(t) - \mathbf{f}||_2}{||\mathbf{f}||_2}. \quad (18)$$

5.1 Simulation results

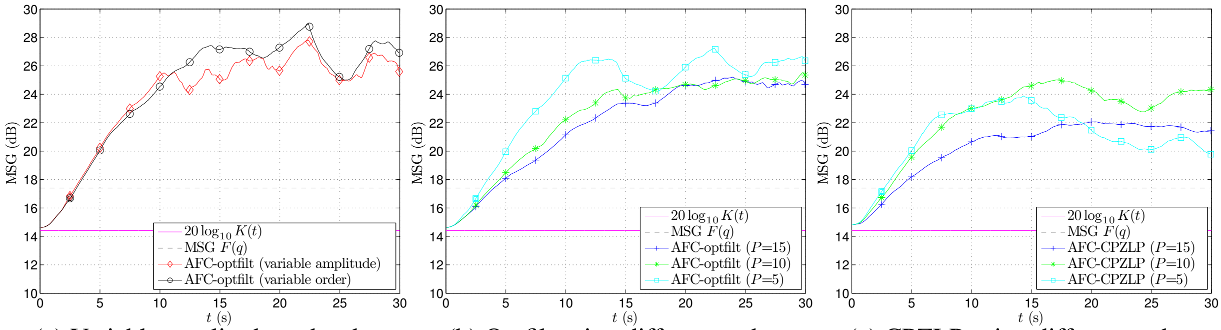
The instantaneous value of the MSG(t) is shown in Fig.6 and the corresponding misadjustment is shown in Fig.7. The MSG(t) curves have been smoothed with a one-pole low-pass filter to improve the clarity of the figures. The instantaneous value of the forward path gain $20\log_{10}K(t)$ and the MSG without acoustic feedback control (MSG $F(q)$) are also shown.

In general the MSG is higher for AFC-optfilt compared to AFC-CPZLP and the corresponding misadjustment is also lower for AFC-optfilt. For the AFC-CPZLP a fixed order of 20 seems to be the best choice whereas for AFC-optfilt a fixed order of 10 gives the best result. The fact that AFC-optfilt can achieve a better performance than AFC-CPZLP with a lower order can be explained by using Fig.3 and 4. The structure of the PEF is more dense towards lower frequencies when the pitch estimation method is used and it is therefore anticipated that the PEF using CPZLP does not sufficiently suppress the tonal components when a lower order is used. Furthermore it is also clear that a fixed order of 30 is too high, which can be seen in Fig.5, especially when the PEF applies infinite suppression at the sinusoidal frequencies.

The MSG performance of the AFC-optfilt when variable order and variable amplitude is used is shown in Fig.6(a). Using a variable order and a variable amplitude (with order 30) almost results in the same AFC performance, with a small advantage too the variable order performance. The performance when both variable order and variable amplitude are included is not shown since the performance is similar when only a variable amplitude is used. This probably happens because at very low amplitude the PEF results in a pole-zero cancellation and no suppression is applied.

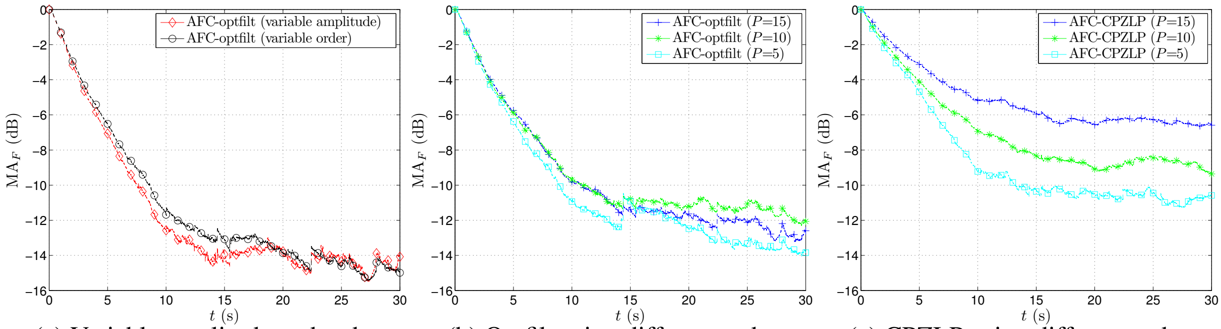
6. CONCLUSION

In this paper, a PEM-based AFC approach is introduced that uses a harmonic sinusoidal near-end signal model based on



(a) Variable amplitude and order (b) Optfilt using different order (c) CPZLP using different order

Figure 6: Instantaneous MSG vs. time for simulations with speech for PEM-based AFC in hearing aids.



(a) Variable amplitude and order (b) Optfilt using different order (c) CPZLP using different order

Figure 7: Misadjustment between the estimated feedback path $\hat{\mathbf{f}}(t)$ and the true feedback path $\mathbf{f}$.

pitch estimation. The proposed PEF design results in increased performance in terms of MSG and misadjustment compared to using a non-harmonic near-end signal model. In the pitch estimation the sinusoids are assumed to have frequencies that are integer multiples of a fundamental frequency, which results in a more dense PEF at lower frequencies, and therefore a lower order can be used. Furthermore, it is shown that the PEM-based AFC performance can be further improved by including a variable amplitude and a variable order in the PEF next to a variable pitch.

REFERENCES

[1] Harvey Dillon, *Hearing Aids*, Boomerang Press, Thieme, 2001.

[2] Mark Kahrs and Karlheinz Brandenburg, Eds., *Applications of digital signal processing to audio and acoustics*, Kluwer Academic Publishers, Norwell, MA, USA, 1998.

[3] T. van Waterschoot and M. Moonen, "50 years of acoustic feedback control: state of the art and future challenges," *Proc. IEEE*, submitted for publication, Feb. 2009, ESAT-SISTA Technical Report TR 08-13, Katholieke Universiteit Leuven, Belgium.

[4] C. Boukis, D. P. Mandic, and A. G. Constantinides, "Toward bias minimization in acoustic feedback cancellation systems," *J. Acoust. Soc. Am.*, vol. 121, no. 3, pp. 1529–1537, Mar. 2007.

[5] D. J. Freed, "Adaptive feedback cancellation in hearing aids with clipping in the feedback path," *J. Acoust. Soc. Am.*, vol. 123, no. 3, pp. 1618–1626, Mar. 2008.

[6] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.

[7] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio applications," *Signal Processing*, vol. 89, no. 11, pp. 2185–2201, Nov. 2009.

[8] K. Ngo, T. van Waterschoot, M. G. Christensen, M. Moonen, S. H. Jensen, and J. Wouters, "Adaptive feedback cancellation in hearing aids using a sinusoidal near-end signal model," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Dallas, Texas, USA, Mar. 2010.

[9] M. G. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, Morgan & Claypool, 2009.

[10] T. van Waterschoot and M. Moonen, "Constrained pole-zero linear prediction: an efficient and near-optimal method for multi-tone frequency estimation," in *Proc. 16th European Signal Process. Conf. (EUSIPCO '08)*, Lausanne, Switzerland, Aug. 2008.

[11] P. M. Djuric, "Asymptotic map criteria for model selection," *IEEE Trans. Signal. Process.*, vol. 46, no. 7, pp. 2726–2735, Oct. 1998.

[12] Y. T. Chan, J. M M. Lavoie, and J. B. Plant, "A parameter estimation approach to estimation of frequencies of sinusoids," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 29, no. 2, pp. 214–219, Apr. 1981.

[13] T. van Waterschoot and M. Moonen, "A pole-zero placement technique for designing second-order IIR parametric equalizer filters," *IEEE Trans. Audio Speech Lang. Process.*, vol. 15, no. 8, pp. 2561–2565, Nov. 2007.