ADAPTIVE FEEDBACK CANCELLATION FOR AUDIO SIGNALS USING A WARPED ALL-POLE NEAR-END SIGNAL MODEL

Toon van Waterschoot and Marc Moonen

Katholieke Universiteit Leuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
{toon.vanwaterschoot,marc.moonen}@esat.kuleuven.be
http://homes.esat.kuleuven.be/∼tvanwate

ABSTRACT

Sound amplification systems having a closed signal loop often suffer from acoustic feedback, which limits the achievable amount of amplification and severely affects sound quality. A promising solution to the feedback problem consists in predicting the feedback signal using an adaptive filter, however, a bias is then introduced due to signal correlation. In speech applications, a prediction-error-method-based approach to adaptive feedback cancellation has proven to be capable of providing sufficient decorrelation without sacrificing speech quality. This approach, which is based on estimating an all-pole near-end signal model, appears to be unappropriate for musical audio signals because of their large degree of tonality. We propose a novel prediction-error-method-based adaptive feedback cancellation algorithm that features a frequency-warped all-pole near-end signal model, which is better suited for tonal audio signals. Simulation results show a doubling of the convergence speed, with only a relatively small increase in computational complexity.

Index Terms— Acoustic feedback, adaptive filters, audio signal processing, warped linear prediction, prediction error identification

1. INTRODUCTION

Acoustic feedback is a physical phenomenon arising in several speech and audio applications, which may severely degrade sound quality and may even cause damage to human hearing and to loudspeaker components. When a sound signal is picked up by a microphone and is played back after amplification in the same acoustic environment, then a closed signal loop is created, which may give rise to system instability. The existence of an acoustic feedback path limits a sound system's performance in two ways. First of all, there is an upper limit to the amount of amplification that can be applied if the system is required to remain stable, which is called the maximum stable gain (MSG). Second, the sound quality is affected by occasional howling when the MSG is exceeded, or, even when the system is operating below the MSG, by ringing and excessive reverberation.

Apart from manual feedback control, the two most promising solutions to the acoustic feedback problem are notch-filter-based feedback suppression (NFS) and adaptive feedback cancellation (AFC). Notch-filter-based solutions aim at detecting ringing or howling sounds in the microphone signal spectrum, and subsequently placing suitably designed notch filters in the closed signal loop to stabilize the system and remove the narrowband interferences. The main advantages of the notch-filter-based approach are its relatively small computational cost and its robustness when operating in an unstable sound system. A large shortcoming, however, is the reactive nature of NFS, i.e., feedback suppression is only achieved after the system has become (nearly) unstable, hence signal distortion can never be avoided.

Adaptive feedback cancellation (AFC) is a more recent approach to acoustic feedback control, which is based on the concept of acoustic echo cancellation (AEC). If the loudspeaker signal is fed to an adaptive filter, having the microphone signal as its desired signal, then the feedback component in the microphone signal may be predicted by the adaptive filter. However, in contrast to the AEC situation, in AFC the loudspeaker signal is correlated with the near-end signal, which causes standard adaptive filtering algorithms to converge to a biased solution [1]. A variety of decorrelation techniques has been suggested for bias removal, which can be divided in two categories: either the decorrelation is performed in the closed signal loop, or in the adaptive filtering circuit. In this paper, we focus on prediction-error-method-based AFC algorithms [1],[2], which belong to the latter category. In these algorithms, decorrelation is achieved by prefiltering the adaptive filter's input and desired signal with an inverse model of the near-end signal, which is also adaptively estimated. In contrast to the NFS approach, an AFC system performs proactive feedback control, i.e., if the acoustic feedback path characteristics can be effectively tracked by the adaptive filter, then the sound system is guaranteed to remain stable, and moreover ringing and reverberation effects are also avoided. The main drawback of the AFC approach is its high computational cost, especially in room acoustic applications.

The majority of research efforts in acoustic feedback control has dealt with speech applications. In this paper, we explicitly focus on feedback control in audio applications involving musical signals, e.g., public address (PA) systems in concert venues, or hearing aids (HA) operating in a musical environment. When dealing with audio instead of speech applications, three major issues should be taken into account [3]. First of all, whereas in speech applications intelligibility is of prime interest, sound quality becomes much more important for audio applications. Second, a higher sampling frequency is required in audio than in speech applications. Finally, most audio signals exhibit a much higher degree of tonality than speech signals,

Toon van Waterschoot is a Research Assistant with the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of K.U.Leuven Research Council: CoE EF/05/006 Optimization in Engineering, the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P6/04 ("Dynamical systems, control and optimization", 2007-2011), the Concerted Research Action GOA-AMBioRICS, and IWT project 040803: "SMS4PA-II: Sound Management System for Public Address Systems". The scientific responsibility is assumed by its authors.

while many feedback control techniques are not designed to work with tonal signals.

The aim of this paper is to propose a modification to an existing prediction-error-method-based AFC algorithm, known as PEM-AFROW [2],[4], such that it becomes capable of dealing with tonal audio signals too. The PEM-AFROW algorithm is based on the prediction error method for system identification [5], and performs decorrelation by prefiltering the adaptive filter's input and desired signal with the inverse of a time-varying all-pole model of the near-end signal, which is estimated by batch (frame-based) linear prediction (LP) of the feedback-compensated signal. The algorithm has inherited its name from the fact that no modeling approximations are introduced such that the prefiltering operation only performs row operations in the loudspeaker signal data matrix. The all-pole model is appropriate for efficiently modeling speech signals, but is usually not a suited model for tonal audio signals. It is well known that linear prediction of audio signals should rather be performed using a frequency-warped all-pole model [6],[7], resulting in a so-called warped linear prediction (WLP). In this paper, we illustrate how the WLP concept can be applied to AFC in a prediction error identification framework. It should be emphasized the WLP model is not the only alternative to the conventional all-pole model for tonal audio signals. In [8], we have investigated several alternative LP models for audio signals. A derivation and evaluation of AFC algorithms for audio applications based on pitch prediction all-pole models, downsampled all-pole models, and pole-zero models can be found in [3].

This paper is organized as follows. In Section 2, the acoustic feedback problem and the AFC concept are briefly explained. In Section 3, we introduce a prediction error identification criterion that features a warped all-pole near-end signal model. From this criterion we derive the new warped PEM-AFROW (WPEM-AFROW) algorithm, and we evaluate its computational complexity as a function of the algorithm parameters. In Section 4, we describe simulation results that illustrate the performance of the proposed algorithm in terms of convergence speed for a real audio signal. Finally, Section 5 concludes the paper.

2. ACOUSTIC FEEDBACK AND AFC CONCEPT

The acoustic feedback problem is depicted in Fig. 1 for a setup with one microphone and one loudspeaker. In this setup, we refer to the source signal $v(t)$ as the near-end signal, and to the loudspeaker signal $u(t)$ as the far-end signal (adopting the terminology from acoustic echo cancellation). The acoustic feedback path $F\{\cdot\}$ is assumed to be a linear and (slowly) time-varying system of finite order n_F,

$$F(q,t) = f_0(t) + f_1(t)q^{-1} + \ldots + f_{n_F}(t)q^{-n_F} \quad (1)$$

where q denotes the time shift operator, i.e., $q^{-k}u(t) = u(t-k)$. The electro-acoustic forward path $G\{\cdot\}$ is defined as the cascade of the characteristics of the microphone, the A/D-converter, the amplifier, the D/A-converter, the loudspeaker, and any signal processing device that is inserted in the signal loop, such as an equalizer, a compressor, etc. The forward path mapping will typically be nonlinear for large signal amplitudes, due to amplifier or loudspeaker saturation, or because of compression. In the closed-loop system analysis, however, it is usually assumed that the forward path mapping $G(q,t)$ is linear, time-varying, and possibly of infinite order ($n_G \to \infty$). Note that the forward path is assumed to contain (at least) a unit delay, i.e., $g_0(t) \equiv 0$, for avoiding an algebraic loop.

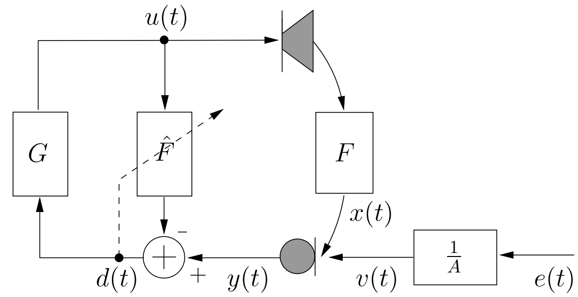

Fig. 1. Adaptive feedback cancellation with near-end signal model.

The far-end signal and the near-end signal are related by the closed-loop transfer function as follows:

$$u(t) = \frac{G(q,t)}{1 - G(q,t)F(q,t)}v(t) \quad (2)$$

According to Nyquist's stability criterion, the closed-loop system will become unstable if there exists a radial frequency ω for which

$$\begin{cases} |G(e^{j\omega},t)F(e^{j\omega},t)| \geq 1 & (3) \\ \angle G(e^{j\omega},t)F(e^{j\omega},t) = n2\pi, \quad n \in \mathbb{Z} & (4) \end{cases}$$

where the short-time frequency responses $G(e^{j\omega},t)$ and $F(e^{j\omega},t)$ of the forward and feedback path, respectively, are obtained using the short-time Fourier transform (STFT).

The AFC concept consists in placing an FIR adaptive filter $\hat{F}(q,t)$ in parallel with the acoustic feedback path, i.e., having the far-end signal as its input and the microphone signal as its desired signal, see Fig. 1. The feedback signal $x(t)$ is then predicted by the adaptive filter output signal $\hat{y}[t,\hat{\mathbf{f}}(t)] = \hat{F}(q,t)u(t)$, which is subtracted from the microphone signal to deliver the feedback-compensated signal $d[t,\hat{\mathbf{f}}(t)] = y(t) - \hat{y}[t,\hat{\mathbf{f}}(t)]$, with

$$\hat{\mathbf{f}}(t) \triangleq \left[\hat{f}_0(t) \quad \ldots \quad \hat{f}_{n_F}(t)\right]^T. \quad (5)$$

Throughout this paper, we will assume that the acoustic feedback path model order n_F is known and equals the adaptive filter order. The closed-loop transfer function of the system with AFC is given by

$$u(t) = \frac{G(q,t)}{1 - G(q,t)[F(q,t) - \hat{F}(q,t)]}v(t). \quad (6)$$

It can be understood that, as the mismatch between $\hat{F}(q,t)$ and $F(q,t)$ decreases, the forward path gain can be increased and the feedback-compensated signal $d[t,\hat{\mathbf{f}}(t)]$ approaches the near-end signal $v(t)$, which can be related to increasing sound quality.

3. WARPED PEM-AFROW ALGORITHM

The estimation of the adaptive filter coefficients in $\hat{\mathbf{f}}(t)$ should be approached from a closed-loop system identification point of view. It is well known that if the near-end signal $v(t)$ is a correlated sequence, such as speech or music, then the standard Wiener or least-squares (LS) estimate converges to a biased solution [1]. The so-called direct method is the only closed-loop system identification method which does not require the insertion of an additional noise signal in the closed signal loop [1], and is hence preferred in audio applications. An unbiased feedback path estimate can be obtained with the direct method only when a model of the near-end signal is taken into account in the identification. Using an autoregressive near-end signal model, the data model can be written as

$$y(t) = F(q,t)u(t) + \frac{1}{A(q,t)}e(t) \tag{7}$$

with $e(t)$ an uncorrelated sequence such as Gaussian white noise or a Dirac impulse. However, because of the nonstationarity of speech and music, the near-end signal model $A^{-1}(q,t)$ is time-varying and should be estimated concurrently with the acoustic feedback path $F(q,t)$. This is possible by applying a prediction error system identification method [5], as shown in [1],[2].

In speech AFC applications, the near-end signal model $A^{-1}(q,t)$ is a conventional all-pole model [1],[2], i.e.,

$$A(q,t) = 1 + a_1(t)q^{-1} + \ldots + a_{n_A}(t)q^{-n_A}. \tag{8}$$

If the near-end signal is a tonal audio signal, then a warped all-pole model may be more appropriate than a conventional all-pole model [6]-[8]. A warped all-pole model is obtained by replacing the unit delay operator q^{-1} in the conventional all-pole model (8) by a first order all-pass filter $D(q,\lambda)$, i.e.,

$$A(q,t,\lambda) = 1 + a_1(t)D(q,\lambda) + \ldots + a_{n_A}(t)D^{n_A}(q,\lambda) \tag{9}$$

with the warping parameter $\lambda \in (-1,1)$, and

$$D(q,\lambda) = \frac{q^{-1} - \lambda}{1 - \lambda q^{-1}}. \tag{10}$$

Concurrent estimation of $F(q,t)$ and $A(q,t,\lambda)$ is possible by minimizing the sum of squared prediction errors

$$\min_{\hat{\mathbf{f}}(t),\hat{\mathbf{a}}(t)} \frac{1}{2N} \sum_{k=1}^{t} \zeta^{-1}(k,t)\left\{\hat{A}(q,t,\lambda)[y(k) - \hat{F}(q,t)u(k)]\right\}^2 \tag{11}$$

where $\zeta^{-1}(k,t)$ is a weighting factor for discounting old data and compensating for power variations in the near-end excitation signal $e(t)$, N denotes the effective window length after data weighting, and

$$\hat{\mathbf{a}}(t) \triangleq \begin{bmatrix} \hat{a}_1(t) & \ldots & \hat{a}_{n_A}(t) \end{bmatrix}^T. \tag{12}$$

Following the approach in [1], [2], the minimization of (11) can be decoupled w.r.t. the parameters in $\hat{\mathbf{f}}(t)$ and $\hat{\mathbf{a}}(t)$, such that different data windows and estimation schemes can be used for both. We will adopt the PEM-AFROW estimation schemes [2], in which the near-end signal model parameters in $\hat{\mathbf{a}}(t)$ are estimated by linear prediction of the feedback-compensated signal $d[t,\hat{\mathbf{f}}(t-1)] = y(t) - \hat{F}(q,t-1)u(t)$ and the feedback path coefficients in $\hat{\mathbf{f}}(t)$ are estimated using a normalized least mean squares (NLMS) adaptive algorithm with input signal $D_0^{-1}(q,\lambda)\hat{A}(q,t,\lambda)u(t)$ and desired signal $D_0^{-1}(q,\lambda)\hat{A}(q,t,\lambda)y(t)$. Note that the prefiltering operation for calculating the NLMS input and desired signal includes an additional whitening filter

$$D_0^{-1}(q,\lambda) = \frac{1 - \lambda q^{-1}}{\sqrt{1-\lambda^2}} \tag{13}$$

which is necessary for compensating the undesired low-pass ($\lambda > 0$) or high-pass ($\lambda < 0$) behavior of the inverse warped near-end signal model $A(q,t,\lambda)$ [6],[7].

The warped PEM-AFROW (WPEM-AFROW) algorithm is displayed in Table 1. Since the near-end signal model is estimated in a frame-based manner, with frames of length M and hop size P (i.e., overlap $M-P$), steps 1) to 4) are performed only every P iterations. The WLP algorithm in step 2) consists in calculating the warped autocorrelation and subsequently applying the Levinson-Durbin algorithm [7]. Two different estimates of the prediction error variance, $\sigma_A^2(t)$ and $\sigma_\varepsilon^2(t)$, are averaged in step 6) to obtain $\sigma^2(t)$ which is used as a time-varying regularization parameter (next to the fixed regularization parameter α) in the NLMS weight update.

The computational complexity of the eight steps in the WPEM-AFROW algorithm is shown in Table 2, in terms of the average number of multiplications per iteration (i.e., averaged over P iterations). The complexity measures for the PEM-AFROW and NLMS algorithms are also given. It should be emphasized that M and P are typically much larger than n_A, while in room acoustic applications n_F is of the same order of magnitude as M and P, see Section 4 for some realistic values.

Table 2. Complexity comparison: average number of multiplications per iteration

	WPEM-AFROW	PEM-AFROW	NLMS
1)	$\frac{M(n_F+1)}{P}$	$\frac{M(n_F+1)}{P}$	
2)	$\frac{M(2n_A+1)+n_A(n_A+4)}{P}$	$\frac{M(n_A+1)+n_A(n_A+4)}{P}$	
3)	$2(n_A+2)$		
4)	$\frac{n_A(2M+n_F)}{P}$	$\frac{n_A(2M+n_F)}{P}$	
5)	n_F+1	n_F+1	n_F+1
6)	4	4	
7)	$2(n_F+2)$	$2(n_F+2)$	$2(n_F+2)$
8)	n_F+1	n_F+1	n_F+1

4. SIMULATION RESULTS

Matlab computer simulations at $f_s = 44100$ Hz were performed to compare the misadjustment,

$$\text{misadjustment (dB)} = 20\log_{10}\frac{\|\hat{\mathbf{f}}(t) - \mathbf{f}\|_2}{\|\mathbf{f}\|_2} \tag{14}$$

of the WPEM-AFROW, PEM-AFROW, and NLMS algorithms. The acoustic feedback path impulse response $\mathbf{f}$ of length $n_F + 1 = 4410$ (100 ms) was measured in a rectangular room of about $5 \times 3 \times 3$ m. The forward path is the cascade of a time delay d, a broadband gain K, and a saturation function. The gain is chosen such that the closed-loop system operates at a gain margin of at least 3 dB, and the delay is chosen equal to the frame hop size, $d = P = 1024$. A frame length of $M = 8192$ (185.8 ms), which is substantially larger than the typical PEM-AFROW frame length in speech applications (ca. 20 ms), was found to be the best choice for the current simulation. The near-end signal is a 60 s excerpt from the Partita No. 2 in D minor (Allemande) for solo violin by J. S. Bach. The near-end signal model order was chosen as $n_A = 14$ in both the WPEM-AFROW and PEM-AFROW algorithms. The warping parameter in the WPEM-AFROW algorithm was set to the Bark optimal value $\lambda = 0.75641$, using the approximation in [9]. Finally, $\alpha = 10^{-6}$ and $\lambda_\varepsilon = 1 - 1/M$.

The convergence curves are shown in Fig. 2. The NLMS step size was tuned individually for each algorithm, such that the steady-state misadjustment would be approximately equal for all

Table 1. WPEM-AFROW algorithm, $j = t \bmod P = 0$ (top), $j = t \bmod P \in [0, P-1]$ (bottom)

1)	$\mathbf{d}[t, \hat{\mathbf{f}}(t-1)] = \begin{bmatrix} y(t+P-M) \\ \vdots \\ y(t+P-1) \end{bmatrix} - \begin{bmatrix} u(t+P-M) & \dots & u(t+P-M-n_F) \\ \vdots & \ddots & \vdots \\ u(t+P-1) & \dots & u(t+P-1-n_F) \end{bmatrix} \hat{\mathbf{f}}(t-1)$
2)	$\{\hat{\mathbf{a}}(t), \sigma_A^2(t)\} = \text{wlp}(\mathbf{d}[t, \hat{\mathbf{f}}(t-1)])$
3)	$\begin{cases} \bar{u}(k,\kappa) = D_0^{-1}(q,\lambda)D^\kappa(q,\lambda)u(k), & k \in [t, t+P-1],\ \kappa \in [0, n_A] \\ \bar{y}(k,\kappa) = D_0^{-1}(q,\lambda)D^\kappa(q,\lambda)y(k), & k \in [t, t+P-1],\ \kappa \in [0, n_A] \end{cases}$
4)	$\begin{cases} \tilde{u}[k, \hat{\mathbf{a}}(t)] = \bar{u}(k,0) + \begin{bmatrix} \bar{u}(k,1) & \dots & \bar{u}(k,n_A) \end{bmatrix} \hat{\mathbf{a}}(t), & k \in [t+P-M-n_F, t+P-1] \\ \tilde{y}[k, \hat{\mathbf{a}}(t)] = \bar{y}(k,0) + \begin{bmatrix} \bar{y}(k,1) & \dots & \bar{y}(k,n_A) \end{bmatrix} \hat{\mathbf{a}}(t), & k \in [t+P-M, t+P-1] \end{cases}$
5)	$\begin{cases} \tilde{\mathbf{u}}[t, \hat{\mathbf{a}}(t-j)] = \begin{bmatrix} \tilde{u}[t, \hat{\mathbf{a}}(t-j)] & \dots & \tilde{u}[t-n_F, \hat{\mathbf{a}}(t-j)] \end{bmatrix}^T \\ \varepsilon[t, \hat{\mathbf{a}}(t-j), \hat{\mathbf{f}}(t-1)] = \tilde{y}[t, \hat{\mathbf{a}}(t-j)] - \tilde{\mathbf{u}}^T[t, \hat{\mathbf{a}}(t-j)]\hat{\mathbf{f}}(t-1) \end{cases}$
6)	$\begin{cases} \sigma_\varepsilon^2(t) = \lambda_\varepsilon \sigma_\varepsilon^2(t) + (1-\lambda_\varepsilon)\varepsilon^2[t, \hat{\mathbf{a}}(t-j), \hat{\mathbf{f}}(t-1)] \\ \sigma^2(t) = [\sigma_A^2(t-j) + \sigma_\varepsilon^2(t)]/2 \end{cases}$
7)	$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mu \frac{\tilde{\mathbf{u}}[t, \hat{\mathbf{a}}(t-j)]\varepsilon[t, \hat{\mathbf{a}}(t-j), \hat{\mathbf{f}}(t-1)]}{\tilde{\mathbf{u}}^T[t, \hat{\mathbf{a}}(t-j)]\tilde{\mathbf{u}}[t, \hat{\mathbf{a}}(t-j)] + \sigma^2(t) + \alpha}$
8)	$d[t, \hat{\mathbf{f}}(t)] = y(t) - \begin{bmatrix} u(t) & \dots & u(t-n_F) \end{bmatrix} \hat{\mathbf{f}}(t)$

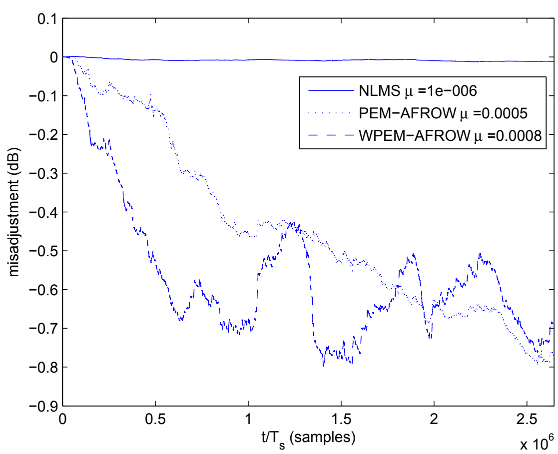

Fig. 2. Misadjustment convergence curves of computer simulations.

algorithms. It can be seen that the WPEM-AFROW algorithm converges roughly twice as fast as the PEM-AFROW algorithm, while the NLMS algorithm converges extremely slowly.

5. CONCLUSION

In this paper, we have proposed a novel AFC algorithm which is obtained by replacing the all-pole near-end signal model in the PEM-AFROW algorithm by a warped all-pole model. If the warping parameter is chosen such that the warping operation approximates the Bark scale, then the warped all-pole model is more appropriate for predicting audio signals. As a consequence, the novel WPEM-AFROW algorithm is well-suited for AFC with audio signals. Simulation results show that the WPEM-AFROW misadjustment converges roughly twice as fast as the PEM-AFROW misadjustment, while the increase in computational complexity is relatively small.

6. REFERENCES

[1] 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.

[2] G. Rombouts, T. van Waterschoot, K. Struyve, and M. Moonen, "Acoustic feedback suppression for long acoustic paths using a nonstationary source model," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3426–3434, Sept. 2006.

[3] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio applications," Oct. 2007, ESAT-SISTA Technical Report TR 07-30, Katholieke Universiteit Leuven, Belgium.

[4] G. Rombouts, T. van Waterschoot, and M. Moonen, "Robust and efficient implementation of the PEM-AFROW algorithm," *J. Audio Eng. Soc.*, vol. 55, no. 11, pp. 955–966, Nov. 2007.

[5] L. Ljung and T. Söderström, *Theory and practice of recursive identification*, MIT Press, Cambridge, Massachusetts, 1986.

[6] H. W. Strube, "Linear prediction on a warped frequency scale," *J. Acoust. Soc. Amer.*, vol. 68, no. 4, pp. 1071–1076, Oct. 1980.

[7] A. Härmä and U. K. Laine, "A comparison of warped and conventional linear predictive coding," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 5, pp. 579–588, July 2001.

[8] T. van Waterschoot and M. Moonen, "Comparison of linear prediction models for audio signals," *EURASIP J. Audio, Speech, Music Process.*, submitted for publication, Sep. 2007.

[9] J. O. Smith and J. S. Abel, "Bark and ERB bilinear transforms," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 6, pp. 697–708, Nov. 1999.