Acoustic Echo Cancellation in the Presence of Continuous Double-Talk

Toon van Waterschoot*, Geert Rombouts*, Kris Struyve† and Marc Moonen*

*Katholieke Universiteit Leuven,
ESAT-SCD,
Kasteelpark Arenberg 10,
B-3001 Leuven, Belgium,
http://www.esat.kuleuven.ac.be/scd
toon.vanwaterschoot@esat.kuleuven.ac.be

†Televic N.V.,
L. Bekaertlaan 1,
B-8870 Izegem, Belgium,
http://www.televic.com/
k.struyve@televic.com

Abstract

The use of a double-talk detector in acoustic echo cancellation cannot improve the adaptive algorithm's performance if near-end noise is continuously present. We propose a new way of dealing with such a continuous double-talk situation, which may occur e.g. in an automatic gain adjustment application. If the microphone and loudspeaker signals are prefiltered with the inverse near-end signal model, a minimum variance room impulse response estimate can be obtained. However, the near-end signal model is unknown and time-varying and has to be estimated concurrently with the room impulse response. We apply three different prediction error identification algorithms to this problem, that were originally developed for adaptive feedback cancellation. Simulation results indicate that only the prediction error method based adaptive filtering algorithm applying row operations (PEM-AFROW) outperforms standard RLS or NLMS adaptive algorithms.

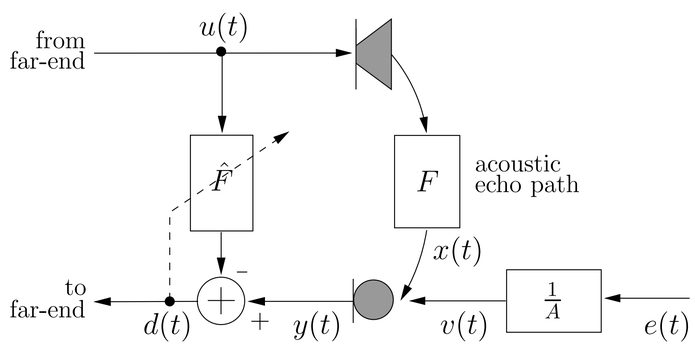
1 Introduction

Consider the acoustic echo cancellation (AEC) problem

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

in which an echo-compensated signal based on an estimate of the room impulse response $\hat{F}(q,t)$ is sent to the far-end side:

$$d(t) = y(t) - \hat{F}(q,t)u(t). \tag{2}$$



We focus on the continuous double-talk situation (simultaneous far-end and near-end activity 100 % of the time) which may occur in

- a noisy teleconferencing scenario,
- an automatic gain adjustment application,
- an acoustic feedback scenario.

The use of a double-talk detector is irrelevant in this scenario since it would freeze the adaptation continuously.

Our aim is to develop a double-talk robust adaptive filtering algoritm for estimating the room impulse response (RIR) $F(q,t)$. This will be accomplished by including a model of the near-end signal:

$$y(t) = F(q,t)u(t) + \frac{1}{A(q,t)}e(t). \tag{3}$$

The RIR model $F(q,t)$ and the inverse near-end signal model $A(q,t)$ have to be identified concurrently. We propose three algoritms for concurrent identifcation based on prediction error (PE) identification theory.

Concatenating the parameters vectors of $F(q,t)$ and $A(q,t)$,

$$\mathbf{f}(t) \triangleq \left[f_0(t)\ f_1(t)\ \dots\ f_{n_F}(t)\right]^T, \tag{4}$$
$$\mathbf{a}(t) \triangleq \left[a_1(t)\ a_2(t)\ \dots\ a_{n_A}(t)\right]^T. \tag{5}$$

leads to a vector containing all to-be-identified parameters:

$$\boldsymbol{\theta}(t) \triangleq \begin{bmatrix}\mathbf{f}(t)\\ \mathbf{a}(t)\end{bmatrix} \tag{6}$$

The prediction error criterion for estimating $\boldsymbol{\theta}(t)$ may be defined as

$$V_{PE}(t,\boldsymbol{\theta}(t)) = \frac{1}{2}\sum_{k=1}^{t}\frac{\lambda_{\boldsymbol{\theta}}^{t-k}}{\hat{\sigma}_k^2}\big(A(q,t)[y(k) - F(q,t)u(k)]\big)^2 \tag{7}$$

with:

- exponential weighting with forgetting factor $\lambda_{\boldsymbol{\theta}}$ to allow tracking of a time-varying RIR
- weighting with the inverse variance of the near-end excitation signal $e(t)$, estimated as

$$\hat{\sigma}_k^2 = \lambda_\sigma\hat{\sigma}_{k-1}^2 + (1-\lambda_\sigma)\varepsilon^2(k,\hat{\boldsymbol{\theta}}(k-1)), \tag{8}$$

using the prediction error

$$\varepsilon(t,\boldsymbol{\theta}(t)) = A(q,t)[y(t) - F(q,t)u(t)]. \tag{9}$$

2 Recursive PE identification

2.1 Two-channel adaptive filtering (2ch-AF)

By modelling the convolution of the inverse near-end signal model and the RIR as one FIR filter,

$$B(q,t) \triangleq -A(q,t)F(q,t), \tag{10}$$

the prediction error criterion can be linearized as follows:

$$V_{2ch}(t,\boldsymbol{\theta_{2ch}}(t)) = \frac{1}{2}\sum_{k=1}^{t}\frac{\lambda_{\boldsymbol{\theta}}^{t-k}}{\hat{\sigma}_k^2}\big(A(q,t)y(k) + B(q,t)u(k)\big)^2 \tag{11}$$
$$= \frac{1}{2}\sum_{k=1}^{t}\frac{\lambda_{\boldsymbol{\theta}}^{t-k}}{\hat{\sigma}_k^2}\big(y(k) + \boldsymbol{\phi_{2ch}}^T(k)\boldsymbol{\theta_{2ch}}(t)\big)^2, \tag{12}$$

with

$$\boldsymbol{\theta_{2ch}}(t) \triangleq [b_0(t)\ \dots\ b_{n_B}(t)\quad a_1(t)\ \dots\ a_{n_A}(t)]^T, \tag{13}$$
$$\boldsymbol{\phi_{2ch}}(k) \triangleq [u(k)\ \dots\ u(k-n_B)\quad y(k-1)\ \dots\ y(k-n_A)]^T \tag{14}$$

An ordinary RLS algorithm can be applied to estimate $\boldsymbol{\theta_{2ch}}(t)$:

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{2ch}}(t) = \hat{\boldsymbol{\theta}}_{\boldsymbol{2ch}}(t-1) - \frac{1}{\hat{\sigma}_t^2}\mathbf{R}^{-1}(t)\boldsymbol{\phi_{2ch}}(t)\varepsilon(t,\hat{\boldsymbol{\theta}}_{\boldsymbol{2ch}}(t-1)),$$
$$\mathbf{R}(t) = \lambda_{\boldsymbol{\theta}}\mathbf{R}(t-1) + \frac{1}{\hat{\sigma}_t^2}\boldsymbol{\phi_{2ch}}(t)\boldsymbol{\phi_{2ch}}^T(t), \tag{15}$$
$$\varepsilon(t,\hat{\boldsymbol{\theta}}_{\boldsymbol{2ch}}(t-1)) = y(t) + \boldsymbol{\phi_{2ch}}^T(t)\hat{\boldsymbol{\theta}}_{\boldsymbol{2ch}}(t-1). \tag{16}$$

Pro: → convex cost function (no local minima)
Contra: → estimation of $F(q,t)$ and $A(q,t)$ is performed on the same data window,
→ deconvolution needed to obtain $F(q,t)$ from $B(q,t)$ (complexity)

2.2 Prediction error method based adaptive filtering (PEM-AF)

Recursive prediction error identification of $F(q,t)$ and $A(q,t)$ from the non-linear PE criterion leads to a recursion for parameter vector $\hat{\boldsymbol{\theta}}(t)$:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{1}{\hat{\sigma}_t^2}\mathbf{R}^{-1}(t)\boldsymbol{\psi}(t,\hat{\boldsymbol{\theta}}(t-1))\varepsilon(t,\hat{\boldsymbol{\theta}}(t-1)). \tag{17}$$

The gradient vector $\boldsymbol{\psi}(t,\boldsymbol{\theta}(t))$ is defined as

$$\boldsymbol{\psi}(t,\boldsymbol{\theta}(t)) = \begin{bmatrix}\boldsymbol{\psi_f}(t,\mathbf{a}(t))\\ \boldsymbol{\psi_a}(t,\mathbf{f}(t))\end{bmatrix} \triangleq -\begin{bmatrix}\frac{\partial}{\partial\mathbf{f}(t)}\varepsilon(t,\boldsymbol{\theta}(t))\\ \frac{\partial}{\partial\mathbf{a}(t)}\varepsilon(t,\boldsymbol{\theta}(t))\end{bmatrix}. \tag{18}$$

Decoupling of the data windows on which $F(q,t)$ and $A(q,t)$ are estimated can be achieved by block-diagonalizing $\mathbf{R}(t)$ as

$$\mathbf{R}(t) = \begin{bmatrix}\mathbf{R_f}(t) & \mathbf{0}_{(n_F+1)\times n_A}\\ \mathbf{0}_{n_A\times(n_F+1)} & \mathbf{R_a}(t)\end{bmatrix}. \tag{19}$$

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \frac{1}{\hat{\sigma}_t^2}\mathbf{R_f}^{-1}(t)\boldsymbol{\psi_f}(t)\varepsilon(t), \tag{20}$$
$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \frac{1}{\hat{\sigma}_t^2}\mathbf{R_a}^{-1}(t)\boldsymbol{\psi_a}(t)\varepsilon(t), \tag{21}$$
$$\mathbf{R_f}(t) = \lambda_{\mathbf{f}}\mathbf{R_f}(t-1) + \frac{1}{\hat{\sigma}_t^2}\boldsymbol{\psi_f}(t)\boldsymbol{\psi_f}^T(t), \tag{22}$$
$$\mathbf{R_a}(t) = \lambda_{\mathbf{a}}\mathbf{R_a}(t-1) + \frac{1}{\hat{\sigma}_t^2}\boldsymbol{\psi_a}(t)\boldsymbol{\psi_a}^T(t). \tag{23}$$
$$\varepsilon(t) = \hat{A}(q,t-1)\xi(t), \tag{24}$$
$$\boldsymbol{\psi_f}(t) = [\hat{A}(q,t-1)u(t)\dots\hat{A}(q,t-n_F-1)u(t-n_F)]^T, \tag{25}$$
$$\boldsymbol{\psi_a}(t) = \left[\xi(t-1)\ \dots\ \xi(t-n_A)\right]^T, \tag{26}$$
$$\xi(t) \triangleq y(t) - \hat{F}(q,t-1)u(t). \tag{27}$$

Pro: → decoupling of the data windows for estimation of $F(q,t)$ and $A(q,t)$
Contra: → non-convex cost function (local minima),
→ approximate prediction error and gradient vectors, assuming that $A(q,t)$ remains stationary during n_F samples

2.3 Prediction error method based adaptive filtering algorithm applying row operations (PEM-AFROW)

A two-stage prediction error identification algorithm:

Stage 1: linear prediction of the near-end signal model $\{\mathbf{a_i}, \sigma_i^2\}$ on a frame of length M of the echo-compensated signal vector

$$\mathbf{d_i} = \mathbf{y_i} - \mathbf{U_i}\hat{\mathbf{f}}((i-1)M), \tag{28}$$

using a previous estimate for the RIR $F(q,t)$.

Stage 2: exponentially windowed RLS algorithm to estimate $F(q,t)$ using the prefiltered data

$$y_{\hat{\mathbf{a}}_\mathbf{j}}(k) \triangleq \left[y(k)\ \dots\ y(k-n_A)\right]\begin{bmatrix}1\\ \hat{\mathbf{a}}_\mathbf{j}\end{bmatrix}, \tag{29}$$
$$u_{\hat{\mathbf{a}}_\mathbf{j}}(k) \triangleq \left[u(k)\ \dots\ u(k-n_A)\right]\begin{bmatrix}1\\ \hat{\mathbf{a}}_\mathbf{j}\end{bmatrix}, \tag{30}$$

and the exact gradient vector

$$\boldsymbol{\psi_f}(t,\hat{\mathbf{a}}_\mathbf{i}) \triangleq \left[u_{\hat{\mathbf{a}}_\mathbf{i}}(t)\ \dots\ u_{\hat{\mathbf{a}}_\mathbf{i}}(t-n_F)\right]^T. \tag{31}$$

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \frac{1}{\hat{\sigma}_i^2}\mathbf{R}^{-1}(t)\boldsymbol{\psi_f}(t,\hat{\mathbf{a}}_\mathbf{i})\varepsilon(t,\hat{\mathbf{f}}(t-1),\hat{\mathbf{a}}_\mathbf{i}) \tag{32}$$
$$\mathbf{R}(t) = \lambda_{\mathbf{f}}\mathbf{R}(t-1) + \frac{1}{\hat{\sigma}_i^2}\boldsymbol{\psi_f}(t,\hat{\mathbf{a}}_\mathbf{i})\boldsymbol{\psi_f}^T(t,\hat{\mathbf{a}}_\mathbf{i}) \tag{33}$$
$$\varepsilon(t,\hat{\mathbf{f}}(t-1),\hat{\mathbf{a}}_\mathbf{i}) = y_{\hat{\mathbf{a}}_\mathbf{i}}(t) - \boldsymbol{\psi_f}^T(t,\hat{\mathbf{a}}_\mathbf{i})\hat{\mathbf{f}}(t-1) \tag{34}$$

Pro: → decoupling of the data windows for estimation of $F(q,t)$ and $A(q,t)$,
→ exact calculation of prediction error and gradient vectors (only row operations)
Contra: → non-convex cost function (local minima)

3 Simulation results

