Digitale Audiotechnik 2

Acoustic Echo Cancellation

Control Strategies

Martin Hagmüller
hagmueller@tugraz.at

Signal Processing and Speech Communication Laboratory
Graz University of Technology – Austria
Introduction
Review - AEC System

\[ x(n) \rightarrow h(n) \rightarrow y(n) \]

Echo-kompensations-filter

\[ d(n) \rightarrow s(n) \]

\[ e(n) \rightarrow b(n) \]

\[ x(n) \]

\[ x(n) \]

\[ h(n) \]

\[ d(n) \]

\[ s(n) \]

\[ b(n) \]
Control Issues

- Fast initial convergence ↔ Final missadjustment
- Convergence ↔ Divergence
- Double talk
NLMS Convergence
Least Mean Squares (LMS) Algorithm

- Minimize error iteratively at every time instance $n$
  \[ \min \{e^2[n]\} \]

- Update equation:
  \[ \hat{h}_i[n + 1] = \hat{h}_i[n] + \mu (e[n]x[n - i]) \]

- Vector notation:
  \[ \hat{h}[n + 1] = \hat{h}[n] + \mu e[n]x[n] \]

- Normalization & Regularization
  \[ \hat{h}[n + 1] = \hat{h}[n] + \mu \frac{e[n]x[n]}{\|x[n]\|^2 + \Delta} \]
System Distance

- **Definition:**
  - $||h_\Delta [n]||^2 = ||h[n] - \hat{h}[n]||^2$

- **White excitation signal $x[n]$, no local noise $n[n]$**
  - $||h_\Delta [n]||^2 = \frac{E\{e[n]^2\}}{E\{x[n]^2\}}$

- **White excitation signal $x[n]$, local noise $n[n]$**
  - lower bound: $\lim_{n \to 0} E \{||h_\Delta [n]||^2\} |_{\Delta=0} \approx \frac{\mu}{2-\mu} \frac{\sigma_n^2}{\sigma_x^2}$
  - Depends on
    - SNR
    - Stepsize $\mu$

- **Equivalent relations for regularization parameter $\Delta$**
Convergence Speed

- Initial convergence:
  - $10 \log_{10} \left( 1 - \frac{\mu(2-\mu)}{N} \right)$ dB per iteration
  - Depends on:
    - stepsize $\mu$
    - Filter length $N$

- Equivalent relations for regularization parameter $\Delta$
Convergence vs. Final Missadjustment (NLMS)

\[ \hat{h}[n + 1] = \hat{h}[n] + \mu \frac{e[n]x[n]}{\|x[n]\|^2} + \Delta \]

- Fast initial convergence →
  - large stepsize \( \mu \)
  - small regularization parameter \( \Delta \)

- Small final missadjustment (\( \|h_\Delta[n]\|^2 \rightarrow 0 \)) →
  - small stepsize \( \mu \)
  - large regularization parameter \( \Delta \)

- Conflicting requirements →
  time-varying stepsize \( \mu[n] \) & regularization parameter \( \Delta[n] \).
Variable Stepsize

\[ h[n+1] = h[n] - \mu \frac{e[n]x[n]}{\|x[n]\|^2} + \Delta \]

- **Optimal stepsizes:**
  - \( \mu_{opt} \approx \frac{E\{|e_u[n]|^2\}}{E\{|e[n]|^2\}} \), where \( e_u[n] \) … undistorted error
  - \( e[n] = e_u[n] + n[n] \)

- **Implementation:**
  - \( \hat{\mu}_{opt} = \frac{|e_u[n]|}{|e[n]|} \)
  - \( |e^2[n]| = \gamma |e^2[n-1]| + (1 - \gamma) e^2[n] \)
  - \( |e_u^2[n]| = \gamma |e_u^2[n-1]| + (1 - \gamma) (e[n] - n[n])^2 \)

- **Problem:** \( e_u \) not directly accessible
Variable Stepsize - Example

Parameters:
- \( x[n], n[n] \) ... white noise
- SNR = 30dB
- \( N = 1000 \)
- \( \gamma = 0.997 \)
Detection & Estimation
Short-term Power Estimation

- First order IIR smoothing:

\[
\|x^2[n]\| = \gamma x^2[n-1] + (1 - \gamma)x^2[n]
\]

\[
\gamma[n] = \begin{cases} 
\gamma_r, & \text{if } x^2[n] \geq x^2[n-1] \\
\gamma_f, & \text{otherwise}
\end{cases}
\]
Short-term Power Estimation

- Local Background Noise Power

\[
|\hat{b}^2[n]| = \min\{|y^2[n]|, |\hat{b}^2[n-1]|\}(1 + \epsilon)
\]

\[
B_\Delta(\epsilon) = f_s 20 \log_{10}(1 + \epsilon)
\]
Short-term Power Estimation Example

-1 0 0.5

-5 -10 -15 -20 -25 -30

200 400 600 800 1000 1200 1400 1600 1800

var_y power_y bg noise_y
System Distance Estimation

- Undisturbed error signal:
  \[ E\{e_u^2[n]\} = E\{x^2[n]\} E\{\|h_\Delta[n]\|^2\} \]

- Echo coupling:
  \[ \beta[n] = E\{\|h_\Delta[n]\|^2\} = E\{\|h[n] - \hat{h}[n]\|^2\} \]

- Pseudo-optimal stepsize:
  \[ \mu[n] = \frac{E\{e_u^2[n]\}}{E\{e^2[n]\}} = \frac{E\{x^2[n]\} \beta[n]}{E\{e^2[n]\}} \]
Echo Coupling
Coupling Factor Estimation

- $\beta_p$ based on power estimation:
  \[
  \beta_p[n] = \frac{|e^2_u[n]|}{|x^2[n]|} \begin{cases} 
    \gamma \beta_p[n-1] + (1 - \gamma) \frac{|e^2[n]|}{|x^2[n]|}, & \text{if remote single talk} \\
    \beta_p[n-1], & \text{otherwise}
  \end{cases}
  \]

- Remote single talk detection necessary

- Power comparison:
  \[
  \beta_L[n] = \begin{cases} 
    \gamma \beta_L[n-1] + (1 - \gamma) \frac{|y^2[n]|}{|x^2[n]|}, & \text{if large remote exitation} \\
    \beta_L[n-1], & \text{otherwise}
  \end{cases}
  \]

- Thresholding:
  \[
  |y^2[n]| \geq K_L |x^2[n]| \beta_L[n]
  \]
System Example - Echo Coupling

Diagram showing the process of echo cancellation with blocks for single-talk detection, coupling factor estimation, and adaptation and step-size control. The diagram includes signals like $x(n)$, $e(n)$, $e^2(n)$, and $h(n)$. The overall system includes feedback and convolution processes.
System Example - Single Talk Detection
Double Talk
Convergence vs. Divergence

Expected system distance

Convergence time
Divergence time

Iterations

dB

Silence
Far end single talk
Double talk

Introduction
NLMS Convergence
Detection & Estimation

Double Talk
- Convergence vs. Divergence
- Intro
- Geigel DTD
- X-Corr DTD
- Norm. X-Corr DTD
- VIRE

Signal Processing and Speech Communication Laboratory

Graz, SS 2010
Martin Hagmüller

DAT 2 – AEC - p. 22/27
Introduction - DTD

- Detection statistic $\xi$ using signals $x, y, e, \hat{h}$
- Compare $\xi$ to threshold $T$: Double-talk if $\xi < T$
- Adaptation of $\hat{h}$ is disabled (for minimum time $t_{hold}$)
- Optimum behaviour:
  - $\xi \geq T$, if $n[n] = 0$, i.e. no double-talk present
  - $\xi < T$, if $n[n] \neq 0$, i.e. double-talk present
  - $\xi$ insensitive to echo path variation
- Performance evaluation
  - Probability of false alarm
  - Probability of detection
  - Probability of Miss
Introduction

NLMS Convergence

Detection & Estimation

Double Talk

- Convergence vs. Divergence
- Intro
- Geigel DTD
- X-Corr DTD
- Norm. X-Corr DTD
- VIRE

Geigel DTD

- Most basic algorithm
- Computationally very simple
- Little memory needed

\[ \xi[n] = \frac{\max\{|x[n], \ldots, x[n - N_g + 1]|}{|y[n]|} < T \]

\( N_g \) ... Memory, \( T \) ... Threshold
Introduction

NLMS Convergence

Detection & Estimation

Double Talk

■ Convergence vs. Divergence
■ Intro
■ Geigel DTD
■ X-Corr DTD
■ Norm. X-Corr DTD
■ VIRE

Cross-Correlation DTD

Cross-correlation between \( x[n] \) and \( y[n] \)

\[
c_{xy} = \frac{E\{x[n]y[n]\}}{\sqrt{E\{x^2[n]\}E\{y^2[n]\}}} = \frac{r_{xy}}{\sigma_x \sigma_y} = [c_{xy,0}, c_{xy,1}, \ldots, c_{xy,N-1}]
\]

Decision variable:

\[
\xi = \|c_{xy}\|_\infty < T
\]

Problem: Statistics of excitation signal
Normalized Cross-Correlation DTD

- Accounts for statistics of excitation signal

\[ \xi = \sqrt{\frac{\sigma_d^2}{\sigma_y^2}} = \sqrt{\frac{r_{xy}^T R_{xx}^{-1} r_{xy}}{\sigma_y^2}} \]

\[ \xi = \sqrt{\frac{\mathbf{h}^T R_{xx} \mathbf{h}}{\mathbf{h}^T R_{xx} \mathbf{h} + \sigma_n}} \]

\[ c_{xy} = (\sigma_y^2 R_{xx})^{-1/2} r_{xy} \]

- Equivalent to adaptive pre-whitening filter
- Simplified normalized cross-correlation:

\[ \xi = \sqrt{\frac{\sigma_d^2[n]}{\sigma_y^2[n]}} \]
Variance Impulse Response (VIRE)

- Based on variance of adaptive filter
- Near end speech → corrupting adaption → fluctuation of filter coefficients

\[ \xi[n] = \gamma \xi[n - 1] + (1 - \gamma) (\kappa - \overline{\kappa})^2 \]

\[ \overline{\kappa} = \gamma \overline{\kappa}[n - 1] + (1 - \gamma) \kappa \]

\[ \kappa = \max (\hat{h}_0, \hat{h}_1, \ldots, \hat{h}_{N-1}) \]

- Computationally simple
- Sensitive to echo path variations