DSP Challenges in Digital Hearing Aids

Marc Moonen
KU Leuven, E.E.Dept/SCD
marc.moonen@esat.kuleuven.be

EDERC 2012 - Amsterdam
September 14, 2012
Acknowledgements

Joint work with


- Simon Doclo, Ann Spriet, Tim Van den Bogaert, Thomas Klasen, Sylwester Szczepaniak, Bram Cornelis, Kim Ngo, Alexander Bertrand, Romain Serizel, Joe Szurley…
Outline

- **Part 1: Long Introduction…**
  - Hearing instruments
  - Hearing impairment
  - Signal processing challenges
  - Binaural hearing
  - Bilateral & binaural hearing aids

- **Part 2: Noise reduction**
  - Monaural noise reduction
  - **Binaural noise reduction**
  - Perceptual evaluation
  - Future: multi-node noise reduction / sensor networks

- Conclusions
Hearing Instruments

Hearing Aids (HAs)

• Audio input/audio output (‘microphone-processing-loudspeaker’)
• ‘Amplifier’, but so much more than an amplifier!!
• History:
  • Horns/trumpets/…
  • ‘Desktop’ HAs (1900)
  • Wearable HAs (1930)
  • Digital HAs (1980)
• State-of-the-art:
  • MHz’s clock speed
  • Millions of arithmetic operations/sec, …
  • Multiple microphones
Cochlear Implants (CIs)
- Audio input/electrode stimulation output
- Stimulation strategy (cfr infra) preprocessing similar to HAs
- History:
  - Volta’s experiment…
  - First implants (1960)
  - Commercial CIs (1970-1980)
  - Digital CIs (1980)
- State-of-the-art:
  - MHz’s clock speed, Mops/sec, …
  - Multiple microphones

Other: Bone anchored HAs, middle ear implants, …
Hearing Instruments

Cochlear Implants (CIs)

Tonotopy of inner ear:
spatial arrangement of where sounds of different frequency are processed

Electrical stimulation
for high frequency

Electrical stimulation
For low frequency

Intra-cochlear electrode
Hearing Instruments

Cochlear Implants (CIs)

External Processor:
- Digital/analog-conversion (cfr infra)
- Digital processing & filterbank
- Etc..
Hearing Instruments

Cochlear Implants (CIs)

Coil
- Inductive/magnetic coupling

Implant
- Electrode array

PS: number of CI-implantees worldwide approx. 200,000
PS: 1 CI is approx. 25kEURO, plus surgery, revalidation,..
PS: 3 companies (Cochlear LtD, Med-El, Advanced Bionics)
Hearing Impairment

Hearing loss types:
- conductive
- sensorineural
- mixed

Typical causes:
- aging
- exposure to loud sounds

one in six adult Europeans (and still increasing…)

[Source: Lapperre]
Hearing Impairment

Dynamic range and audibility

Normal hearing subjects

Hearing impaired subjects

Level

100dB

0dB
Hearing Impairment

Dynamic range and audibility

- Dynamic range compression (DRC) (rather than `amplification`)

![Diagram showing dynamic range compression](image)

Design: multiband DRC, attack time, release time, …
Hearing Impairment

Audibility vs speech intelligibility

- Audibility does not imply intelligibility
- Hearing impaired subjects need 5..10dB larger signal-to-noise ratio (SNR) for speech understanding in noisy environments

- Need for noise reduction (≈speech enhancement) algorithms:
  - State-of-the-art: monaural 2-microphone adaptive noise reduction
  - Near future: binaural noise reduction (see below)
  - Not-so-near future: multi-node noise reduction (see below)
DSP Challenges

**HA technology requirements**

- Small form factor (cfr. user acceptance)
- Low power: 1…5mW (cfr. battery lifetime ≈ 1 week)
- Low processing delay: 10msec (cfr. synchronization with lip reading)

**DSP in HA’s**

- Dynamic range compression (cfr supra)
- Dereverberation: undo filtering (‘echo-ing’) by room acoustics
- Feedback cancellation
- Noise reduction
DSP Challenges: Feedback Cancellation

- **Problem statement**: Loudspeaker signal is fed back into microphone, then amplified and played back again.
- Closed loop system may become unstable (howling).
- Similar to feedback problem in public address systems (for the musicians amongst you).
**DSP Challenges: Feedback Cancellation**

- **Adaptive filtering approach:** Identify **mathematical model** for feedback path, then reconstruct feedback signal and subtract from microphone signal.
- Model: $x(t) = a.u(t) + b.u(t-T) + c.u(t-2T) + d.u(t-3T) + \ldots$
- Hundreds of ‘coefficients’ (a,b,..) & time-varying
- Least-squares estimation of a,b,\

**PS:** similar to echo cancellation in GSM handsets, Skype,…
• **Adaptive filtering approach:** Identify mathematical model for feedback path, then reconstruct feedback signal and subtract from microphone signal

• Need additional functionality to remove ‘bias’ in estimation…
Multimicrophone ‘beamforming’

HA/CI typically has 2 microphones, e.g. ‘directional’ front microphone ‘omnidirectional’ back microphone
DSP Challenges: Noise Reduction

Multimicrophone ‘beamforming’

“filter-and-sum” the microphone signals

constructive interference for target speech direction
destructive interference for other directions
Multimicrophone ‘beamforming’

- Each microphone signal $m(t)$ is filtered:
  
  $\text{filtered\_signal} = a.m(t)+b.m(t-T)+c.m(t-2T)+d.m(t-3T)+\ldots$

- Filtered microphone signals are summed
- Hundreds/thousands of coefficients ($a,b,\ldots$) & time-varying
- Compute optimal coefficients: least squares estimation, matrix decompositions, etc..
Binaural Hearing

**Binaural auditory cues**

- ITD (interaural time difference)
- ILD (interaural level difference)

Binaural cues (ITD: f < 1500Hz, ILD: f > 2000Hz) used for

- Sound localization
- Noise reduction

= `Binaural unmasking` (‘cocktail party’ effect) 0-5dB
Bilateral & Binaural Hearing Aids

Bilateral hearing aids

- Two hearing aids (L+R) but no coordination/cooperation
- Perceptual evaluation [Van den Bogaert et al 2006]

- Localization performance

NH=normal hearing subjects
NO=hearing impaired subjects with HA switched off
O=hearing impaired subjects with non-adaptive bilateral HA
A=hearing impaired subjects with adaptive bilateral HA

Bilateral HA operation distorts binaural cues !!!!
Binaural hearing aids

- Two hearing aids (L&R) with wireless link & cooperation
- Opportunities:
  - More signals (e.g. 2*2 microphones)
  - Better sensor spacing (17cm i.o. 1cm)
- Challenges:
  - Improved localization through cue preservation
  - Improved noise reduction + benefit from binaural unmasking
- Constraints: power/bandwidth/delay of wireless link
  - ..10kBit/s: coordinate program settings, parameters,…
  - ..300kBits/s: exchange 1 or more (compressed) audio signals
Outline

- **Part 1: Long Introduction…**
  - Hearing instruments
  - Hearing impairment
  - Signal processing challenges
  - Binaural hearing
  - Bilateral & binaural hearing aids

- **Part 2: Noise reduction**
  - Monaural noise reduction
  - **Binaural noise reduction**
  - Perceptual evaluation
  - Future: multi-node noise reduction (sensor networks)

- Conclusions
Monaural noise reduction (1/7)

Single-microphone noise reduction:

- Data model:

\[ y[k] = s[k] + n[k] \]

- `Spectral subtraction`:

\[ \hat{S}_i(\omega) = G_i(\omega)Y_i(\omega) \]

\( i \) is frame index, \( \omega \) is frequency, \( 0 < G_i < 1 \) is gain function, e.g.

\[ G_i(\omega) = \frac{\text{SNR}_i(\omega)}{1 + \text{SNR}_i(\omega)} \]
Monaural noise reduction (2/7)

Single-microphone noise reduction:

- Need a VAD (voice activity detection) to distinguish between speech+noise and noise-only frames
- Many ingeneous (not-so-simple) gain functions have been developed

- Conceptually very simple
- Computationally very cheap
- Can reduce full-band SNR, but cannot reduce $\text{SNR}(\omega)$
- Cannot improve speech intelligibility (much)
- Can improve listening comfort
Multi-microphone noise reduction

- Data model:

\[
Y(\omega) = S(\omega) + N(\omega) = d(\omega) \cdot S(\omega) + N(\omega)
\]

\[
\begin{bmatrix}
Y_1(\omega) \\
Y_2(\omega) \\
\vdots \\
Y_M(\omega)
\end{bmatrix} = \begin{bmatrix}
H_1(\omega) \\
H_2(\omega) \\
\vdots \\
H_M(\omega)
\end{bmatrix} \cdot S(\omega) + \begin{bmatrix}
N_1(\omega) \\
N_2(\omega) \\
\vdots \\
N_M(\omega)
\end{bmatrix}
\]

- \(M\) is number of microphones (\(M=2,3,..\))
- \(Hi(\omega)\) is acoustic transfer function from speaker to \(i\)-th microphone
- \(d(\omega)\) is speech `steering vector`
- \(N(\omega)\) is noise vector

\[
E\{N(\omega) \cdot N^H(\omega)\} = \Phi_{NN}(\omega)
\]
Monaural noise reduction (4/7)

Multi-microphone noise reduction

- Will use `filter-and-sum’ estimates

\[
\hat{S}_1(\omega) = F^H(\omega) . Y(\omega) = \begin{bmatrix} F_1^*(\omega) & F_2^*(\omega) & \ldots & F_M^*(\omega) \end{bmatrix} \begin{bmatrix} Y_1(\omega) \\ Y_2(\omega) \\ \vdots \\ Y_M(\omega) \end{bmatrix}
\]

reference microphone (arbitrary)

\[
S_1(\omega) = H_1(\omega) . S(\omega)
\]

- Consider different solutions
  - Fixed beamforming
  - Adaptive beamforming
  - Multi-channel Wiener filtering (MWF)
Multi-microphone noise reduction

\[ \hat{S}_1(\omega) = F^H(\omega).Y(\omega) \]

- **Fixed beamforming**: `superdirective beamforming`
  \[ F(\omega) = \alpha.\Phi^{-1}_{NN}(\omega).d(\omega) \]
  - Need a priori info: noise covariance + speech steering vector
- **Adaptive beamforming**: `generalized sidelobe canceller`
  \[ F(\omega) = \beta.\Phi^{-1}_{NN}(\omega).d(\omega) \]
  - Need a priori info: speech steering vector
  - Need VAD (voice activity detection)
  - Adaptively estimates noise covariance matrix
Monaural noise reduction (6/7)

Multi-microphone noise reduction

- Multi-channel Wiener Filtering (MWF)

\[
\min_{F(\omega)} E\{ |S_1(\omega) - F^H(\omega).Y(\omega)|^2 \} \quad \text{`MMSE'}
\]

\[
F(\omega) = \left( \frac{E\{Y(\omega).Y^H(\omega)\}}{E\{Y(\omega).S_1^*(\omega)\}} \right)^{-1} E\{Y(\omega).S_1^*(\omega)\}.
\]

\[
= \ldots \quad (\text{with } E\{S(\omega).N_1^*(\omega)\} = 0)
\]

\[
= E\{Y(\omega).Y^H(\omega)\}^{-1} (E\{Y(\omega).Y_1^*(\omega)\} - E\{N(\omega).N_1^*(\omega)\})
\]

- Compute during speech+noise frames
- Compute during noise-only frames
Monaural noise reduction (7/7)

Multi-microphone noise reduction

- **Multi-channel Wiener Filtering** (MWF)

\[
F(\omega) = E\{Y(\omega)Y^H(\omega)\}^{-1}\left(E\{Y(\omega)Y^*(\omega)\} - E\{N(\omega)N^*(\omega)\}\right)
\]

- All quantities can be computed, no a priori info
- Need VAD (voice activity detection)
- MWF-solution can be shown to provide a super-directive beamformer cascaded with a (spectral subtraction type) single-channel postfilter
- Can also use alternative criteria, e.g.

\[
\min_{F(\omega)} E\left\{ \left| (S_1(\omega) + \eta N_1(\omega)) - F^H(\omega)Y(\omega) \right|^2 \right\}
\]

To better preserve noise binaural cues (see next slide)
Binaural Noise Reduction

`Full blown’ binaural MWF

• In monaural MWF, number of microphones (M) is mostly 2 (sometimes 3)
• In binaural MWF, an MWF can be operated at each side, with number of microphones M=2*2, i.e. each HA can use its own microphone signals plus 2 contralateral signals.
• Speech binaural cues (between L&R reference microphones) are well preserved (cfr MMSE cost function)
• Noise binaural cues can also be preserved, at the expense of a reduced noise reduction (see previous slide)
**Binaural Noise Reduction**

**Binaural MWF under bandwidth constraints**
- Use only 1 contra-lateral microphone signal
  - Contra-lateral reference microphone signal
  - Contra-lateral fixed beamformer
  - Contra-lateral `optimal’ beamformer: db-MWF
    (=distributed MWF, converges to `full blown MWF solution as if all microphone signals were available at each side!)
    [Doclo et al 2009] [Bertrand et al 2010]
- Transmit **compressed** microphone signals
- Theoretical analysis based on rate-distortion theory
  [Roy & Vetterli 2008] [Srinivasan 2009] [Doclo et al 2010]
- Impact in practice to be analysed
Perceptual Evaluation: Set-up

Test subject with 2 HA’s 8 loudspeakers
Microphone signals are processed by 2 real-time processing units
Output signals are routed back to HA loudspeakers

\[ RT_{60} = 0.62s \]
Perceptual Evaluation: S0-N45

Speech:
Dutch sentences (VU)

Noise:
multitalker babble (Auditec)

$RT_{60} = 0.62$ s
Perceptual Evaluation: S0-N45

S0N45 - Improvement vs. "Unprocessed"

BILATERAL

BINAURAL

SRT improvement [dB]
Perceptual Evaluation: S0-N90/180/270

Speech:
Dutch sentences (VU)

Noise:
3 multitalker babble (Auditec)

$RT_{60} = 0.62s$
Perceptual Evaluation: S0-N90/180/270

S0N90/180/270 - Improvement vs. "Unprocessed"

SRT improvement [dB]

-2
-1
0
1
2
3
4
5
6

Bil. MWF - P. VAD
Bil. MWF - R. VAD
Bil. Fixed BF
Bin. MWF - P. VAD
Bin. MWF - R. VAD

NH
HI

Perceptual Evaluation: S0-N90/180/270
What the future should bring…

Multi-node noise reduction - sensor networks

Initial work: `DANSE’ algorithm = generalized db-MWF

[Bertrand et al 2010]
Conclusions

• Have demonstrated relevance of DSP in HA’s

• Current topic: Binaural noise reduction
  • ≠ Duplicating monaural noise reduction
  • Binaural cue preservation!
  • Initial algorithms & evaluation
  • Study/evaluate latency, coding, .. in wireless link
  • Room for improvement

• Future is challenging…
Questions..