Distributed Signal Processing for Binaural Hearing Aids

Olivier Roy

LCAV - I&C - EPFL

Joint work with Martin Vetterli

July 24, 2008

Outline

Motivations (1/5)

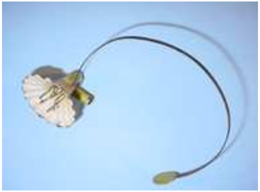
Hearing aids through the ages



Source: `http://www.hearingaidmuseum.com`

Motivations (2/5)

Deafness in disguise



Source: `http://beckerexhibits.wustl.edu/did`

Motivations (3/5)

State-of-the-art technology

- Types: BTE, ITE, ITC, CIC
- Analog vs. digital
- 2-3 (omni)directional microphones, 1 loudspeaker

Motivations (4/5)

Ultimate goal: improve speech intelligibility

- Spectral shaping
- Beamforming
- Assistive listening devices

(a) (b)

Figure: Assistive listening devices. (a) Remote microphone. (b) Binaural hearing aids.

Motivations (5/5)

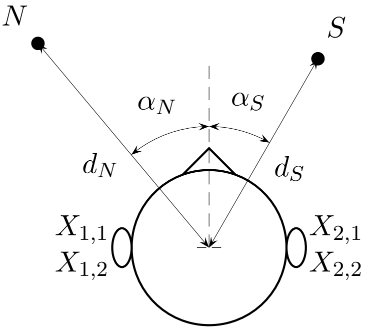
Wireless collaboration

- Analog vs. digital
- Transmission method (e.g. Bluetooth)
- Limited communication bitrate: coding issues

$\boxed{\text{Gain Rate Trade-off}}$

Information-theoretic Analysis (1/9)

Recording setup

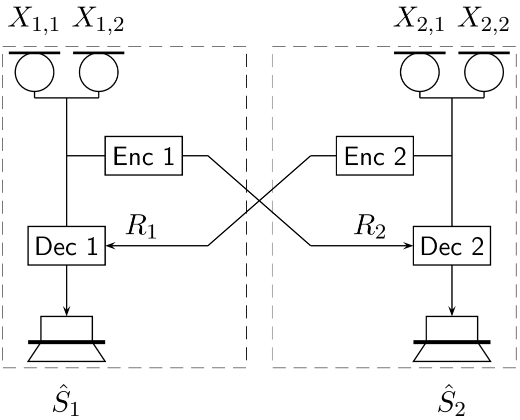


Recorded signals ($m = 1, 2$)

$$
\begin{aligned}
X_{1,m}[n] &= X^s_{1,m}[n] + X^n_{1,m}[n]\,, \\
X_{2,m}[n] &= X^s_{2,m}[n] + X^n_{2,m}[n]\,.
\end{aligned}
$$

Information-theoretic Analysis (2/9)

Wireless collaboration



Distortion criterion $d(S, \hat{S})$ (e.g. MSE, perceptual, etc.)

Information-theoretic Analysis (3/9)

Source coding in a nutshell

$$S \longrightarrow \boxed{\text{Enc}} \xrightarrow{R} \boxed{\text{Dec}} \longrightarrow \hat{S}$$

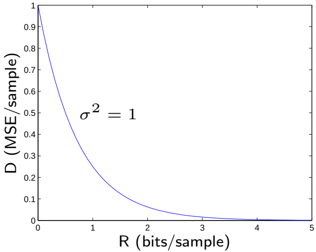
- Given: a source (signal) S and a distortion criterion $d(S, \hat{S})$
- Question: for a given rate R, what is the minimum achievable distortion?
- Answer: the rate distortion function
- Assumption: unbounded coding delay and complexity

Information-theoretic Analysis (4/9)

Example: the Gaussian case

- We observe $X_1, X_2, \ldots$ where $X_k \sim \mathcal{N}(0, \sigma^2)$ i.i.d.
- Rate distortion function given by

$$D(R) = \sigma^2 2^{-2R} \quad \text{(MSE/sample)}$$



- simple 1-bit quantization $\approx 0.36\sigma^2$, optimal $= 0.25\sigma^2$

Information-theoretic Analysis (5/9)

Variations on a theme

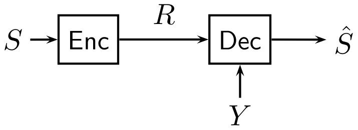
- Remote source coding

$$S \rightsquigarrow X \rightarrow \boxed{\text{Enc}} \xrightarrow{\quad R \quad} \boxed{\text{Dec}} \longrightarrow \hat{S}$$

Information-theoretic Analysis (5/9)

Variations on a theme

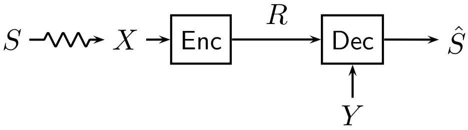
- Source coding with side information at the decoder



Information-theoretic Analysis (5/9)

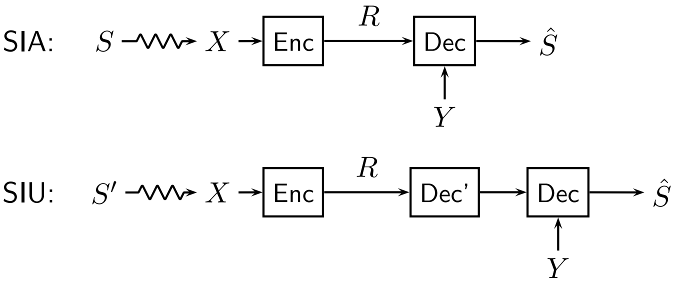
Variations on a theme

- Remote source coding with side information at the decoder



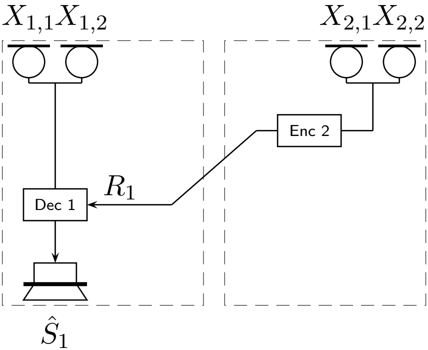
Variations on a theme

- Remote source coding with side information at the decoder



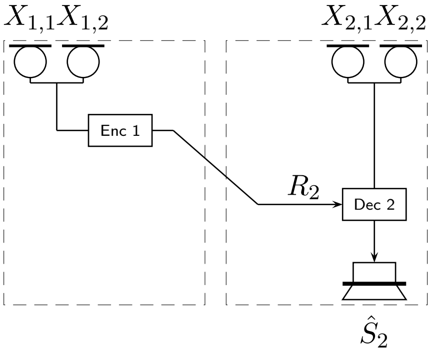
Information-theoretic Analysis (6/9)

What about collaborating hearing aids? Monaural perspective



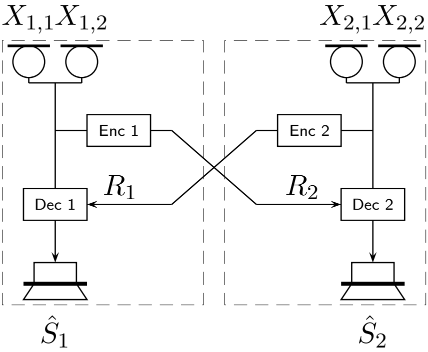
Information-theoretic Analysis (6/9)

What about collaborating hearing aids? Monaural perspective



Information-theoretic Analysis (6/9)

What about collaborating hearing aids? Binaural perspective



Information-theoretic Analysis (7/9)

Results:

- Mean-square optimal gain rate trade offs

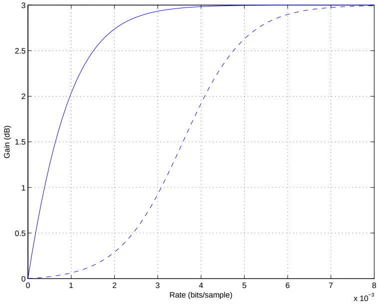


Figure: Examples of gain rate trade offs (SIA vs. SIU)

Information-theoretic Analysis (8/9)

- Mean-square optimal rate allocation

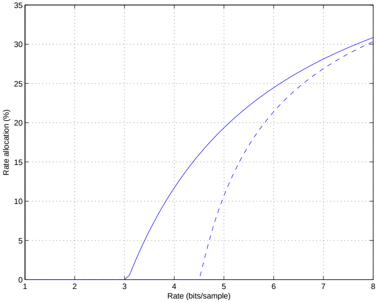


Figure: Examples of rate allocations (SIA vs. SIU)

Information-theoretic Analysis (9/9)

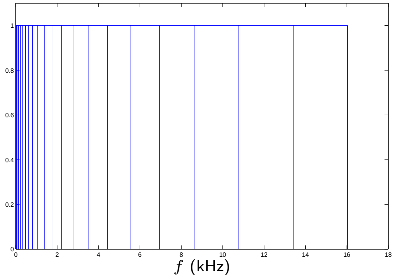
Usefulness of information-theoretic analysis

- Provides upper bounds to gains achieved by practical systems
- Suggests optimal coding architectures
  - Multichannel Wiener filtering
  - Scalar distributed source coding
- Correlation induced by recording setup can be used
  - A priori vs. learned

Example: Distributed Coding of Binaural Cues (1/2)

Binaural Cues

- Scene analysis
  - Classification
  - Source localization
  - Voice activity detection
- Time-frequency representation, one value per critical band $\mathcal{B}_l$



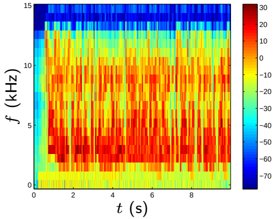
Example: Distributed Coding of Binaural Cues (2/2)

- Inter channel level difference (ICLD)

$$\Delta p[l] = p_1[l] - p_2[l]\,,$$

where

$$p_m[l] = 10\log_{10}\left(\frac{1}{|\mathcal{B}_l|}\sum_{k\in\mathcal{B}_l}|X_m[k]|^2\right) \quad \text{for } m = 1, 2.$$



Example: Distributed Coding of Binaural Cues (2/2)

Centralized coding

$$\Delta p[l] \in \left[\, \Delta p_{min}[l] \,, \Delta p_{max}[l] \,\right]$$

$\Longrightarrow$ scalar quantizer with range $\Delta p_{max}[l] - \Delta p_{min}[l]$

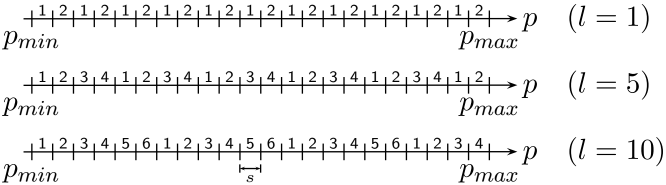
Example: Distributed Coding of Binaural Cues (2/2)

Distributed coding

- Scalar quantization of $p_1[l]$ and $p_2[l]$

$$
\begin{aligned}
i_1[l] - i_2[l] \quad &\in \quad \left\{\Delta i_{min}[l], \ldots, \Delta i_{max}[l]\right\} \\
&= \quad \left\{ \left\lfloor \frac{\Delta p_{min}[l]}{s} \right\rfloor, \ldots, \left\lceil \frac{\Delta p_{max}[l]}{s} \right\rceil \right\}
\end{aligned}
$$

- Modulo coding approach = index reuse



Example: Distributed Coding of Binaural Cues (2/2)

Centralized vs. distributed coding

- Same coding efficiency
- Distributed scheme takes head shadowing into account (i.e., a priori correlation)
- Assumption must be verified!!

Application: distributed spatial audio coding

- original & reconstruction (KEMAR)
- original & reconstruction (BRIR, $T_{60} \approx 600$ ms)

Conclusions

- Binaural noise reduction as a distributed source coding problem
- Information-theoretic analysis
- Distributed coding of binaural cues
- Take home message: correlation/structure that is known a priori is most relevant for distributed source coding

Conclusions

- Binaural noise reduction as a distributed source coding problem
- Information-theoretic analysis
- Distributed coding of binaural cues
- Take home message: correlation/structure that is known a priori is most relevant for distributed source coding

Thanks for your attention!!