KU LEUVEN



Embedded optimization algorithms for multi-microphone dereverberation

Toon van Waterschoot, Bruno Defraene,
Moritz Diehl, and Marc Moonen

STADIUS Center for Dynamical Systems,
Signal Processing, and Data Analytics

ESAT – Department of Electrical Engineering

KU Leuven, Belgium

EUSIPCO 2013, Marrakech, Morocco

1. Introduction
2. Problem statement
3. Embedded optimization algorithms
4. Evaluation
5. Conclusion

1 Introduction

2 Problem statement

3 Embedded optimization algorithms

4 Evaluation

5 Conclusion

- Overview of dereverberation methods
  - beamforming approach
  - speech enhancement approach
  - blind system identification and inversion approach
- Blind system identification and inversion
  - two-step procedure:
    - blind room impulse responses (RIRs) identification
    - inverse filter design using identifed RIRs
  - inversion step is often ill-posed (high sensitivity to RIR estimation errors, near-common RIR zeros, ...)
  - procedure needs to be regularly repeated in time-varying scenarios

- Multi-microphone dereverberation by embedded optimization
  - direct estimation of source signal of interest (jointly with auxiliary variables such as RIRs)
  - "intermediate" inverse filter design step is skipped, thus avoiding ill-posedness
  - more flexibility than "traditional" recursive implementations of closed-form estimators
  - prior knowledge on RIRs and source signal can be elegantly incorporated

1 Introduction

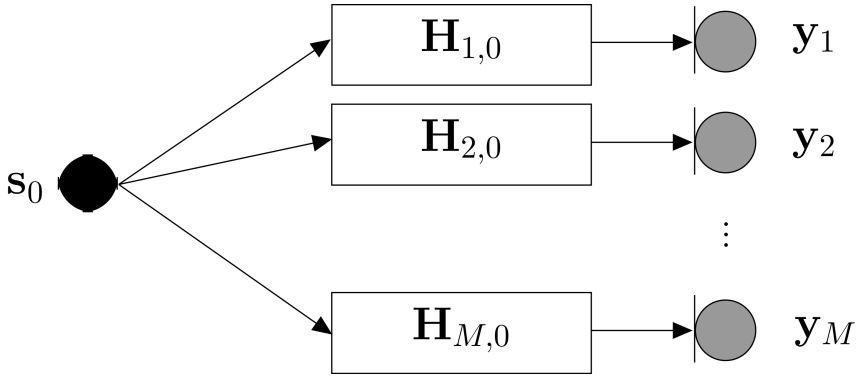
2 Problem statement

3 Embedded optimization algorithms

4 Evaluation

5 Conclusion

- Multi-microphone sound acquisition



- Assumptions
  - batch operation ($t = 1, \ldots, N$), zero initial source signal conditions
  - time-invariant RIRs, with fixed and known length $L \leq N$
  - no measurement noise

Problem (Multi-microphone dereverberation)

Given a length-MN vector $\mathbf{y}$ of microphone signals,
find the best possible estimate of the length-N source signal vector $\mathbf{s}_0$.

- True system

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_M \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{1,0} \\ \vdots \\ \mathbf{H}_{M,0} \end{bmatrix} \mathbf{s}_0 = \mathbf{H}_0 \mathbf{s}_0 \tag{1}$$

- Parameter vectors to be estimated

$$\begin{aligned} \text{RIRs:} \quad & \mathbf{h} = \begin{bmatrix} \mathbf{h}_1^T & \dots & \mathbf{h}_M^T \end{bmatrix}^T \\ \text{source signal vector:} \quad & \mathbf{s} = \begin{bmatrix} s(1) & \dots & s(N) \end{bmatrix}^T \\ \text{estimation error signal vector:} \quad & \mathbf{e} = \begin{bmatrix} \mathbf{e}_1^T & \dots & \mathbf{e}_M^T \end{bmatrix}^T \end{aligned}$$

- Data model (two equivalent formulations)

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e} \tag{2}$$

$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e} \tag{3}$$

1 Introduction

2 Problem statement

3 Embedded optimization algorithms

4 Evaluation

5 Conclusion

- Rationale
  - joint estimation of unknown parameter vectors $\mathbf{h}$ and $\mathbf{s}$
  - incorporate prior knowledge through regularization
  - adopt sequential minimization approach (block coordinate descent, BCD) to deal with nonlinearity
  - solve linearized subproblems using (convex) numerical optimization tools
- Proposed problem formulations
  - NLS: nonlinear least squares
  - ℓ_2-RNLS: ℓ_2-regularized NLS, exploiting prior knowledge on $\mathbf{h}$
  - ℓ_1/ℓ_2-RNLS: ℓ_1/ℓ_2-regularized NLS, exploiting prior knowledge on $\mathbf{s}$ and $\mathbf{h}$

Problem (NLS)

$$\min_{\mathbf{h},\mathbf{s},\mathbf{e}} \quad \|\mathbf{e}\|_2^2 \tag{4}$$

$$s.\ t. \quad \mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e} \tag{5}$$

$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e}. \tag{6}$$

- BCD Solution strategy
  - minimize (4),(5) w.r.t. $\{\mathbf{s}, \mathbf{e}\}$ for a fixed value of $\mathbf{h} = \hat{\mathbf{h}}$
  - minimize (4),(6) w.r.t. $\{\mathbf{h}, \mathbf{e}\}$ for a fixed value of $\mathbf{s} = \hat{\mathbf{s}}$
  - repeat this procedure for a number of iterations
- Properties
  - multiple local solutions to NLS problem exist
  - initialization is crucial for BCD algorithm to converge to global solution

Problem (ℓ_2-RNLS)

$$\min_{\mathbf{h},\mathbf{s},\mathbf{e}} \quad \|\mathbf{e}\|_2^2 + \lambda_1 \|\mathbf{h}\|_{\mathbf{W}}^2 \tag{7}$$

$$s.\ t. \quad \mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e} \tag{8}$$

$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e} \tag{9}$$

- Properties
  - ℓ_2-norm regularization results in smoothing effect, facilitating convergence to meaningful local solution
  - regularization matrix $\mathbf{W}$ allows to incorporate prior knowledge on $\mathbf{h}$, e.g., using Polack's model for late reverberation (with fixed decay α)

$$\mathbf{W} = \mathbf{I}_M \otimes \text{diag}\left\{1, e^{2\alpha}, \ldots, e^{2(L-1)\alpha}\right\}. \tag{10}$$

Problem (ℓ_1/ℓ_2-RNLS)

$$\min_{\mathbf{h},\mathbf{s},\mathbf{e}} \quad \|\mathbf{e}\|_2^2 + \lambda_1 \|\mathbf{h}\|_{\mathbf{W}}^2 + \lambda_2 \|\mathbf{D}\mathbf{s}\|_1 \tag{11}$$

$$s.\ t. \quad \mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e} \tag{12}$$

$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e}. \tag{13}$$

- Properties
  - ℓ_1-norm regularization results in additional smoothing effect
  - $\|\mathbf{D}\mathbf{s}\|_1$ promotes source signal sparsity in basis defined by $\mathbf{D}$
  - sparsity in spectral basis (DFT, DCT, ...) is meaningful for speech/audio
  - minimization w.r.t. $\{\mathbf{s}, \mathbf{e}\}$ does not yield closed-form solution, but allows efficient numerical solution using convex optimization tools (e.g., interior point method, method of multipliers, ...)

1 Introduction

2 Problem statement

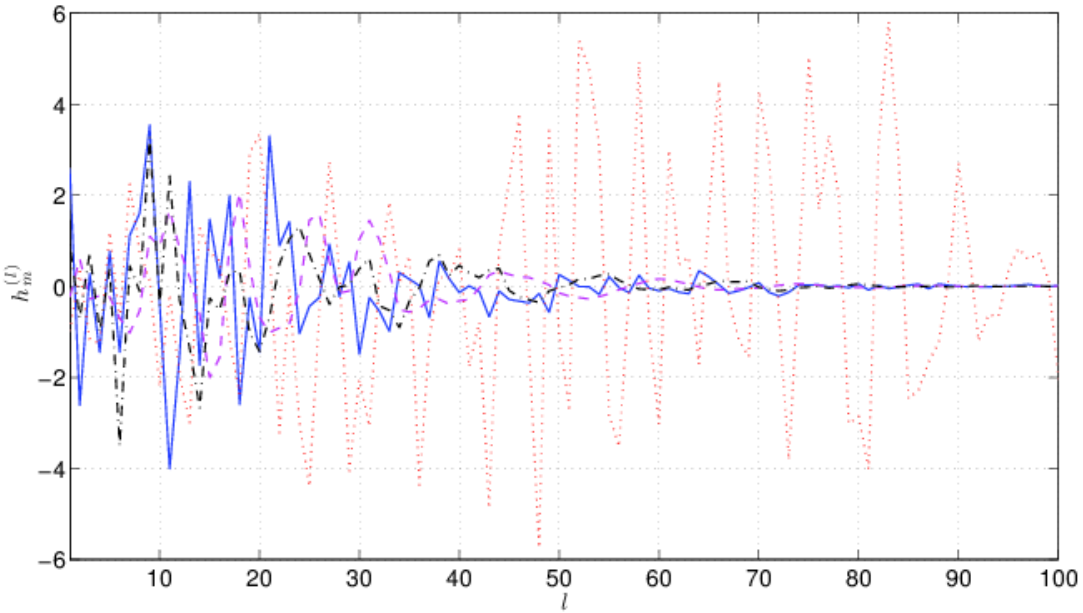
3 Embedded optimization algorithms

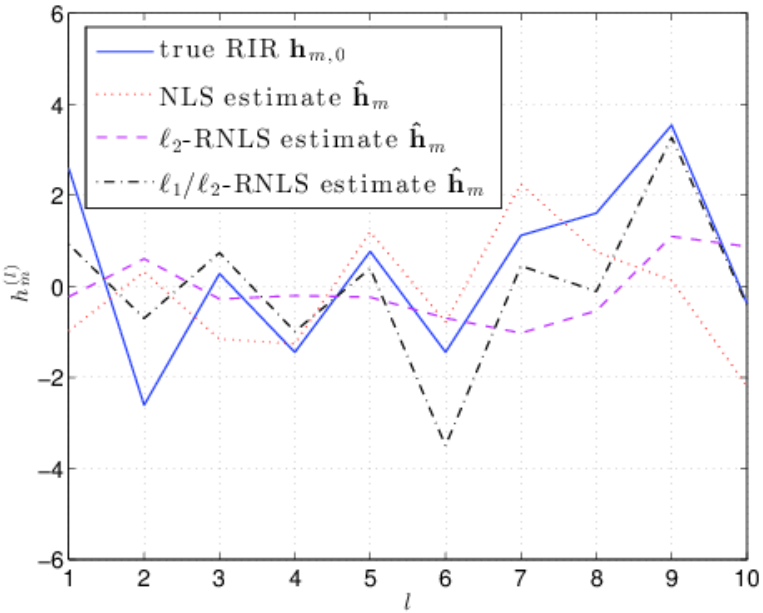
4 Evaluation

5 Conclusion

- Acoustic scenario
  - source signal = quasi-stationary voiced speech segment
  - source signal length $N = 1024$ at $f_s = 8$ kHz
  - $M = 5$ microphones
  - RIRs = GWN shaped by Polack's model envelope ($\alpha = 0.025$)
  - RIRs length $L = 100$
- Algorithm parameters
  - regularization parameters $\lambda_1 = \lambda_2 = 0.1$
  - $\mathbf{D}$ = DCT matrix
  - random (GWN) initialization of RIRs parameter vector estimate
  - fixed number of $k_{\max} = 10$ iterations

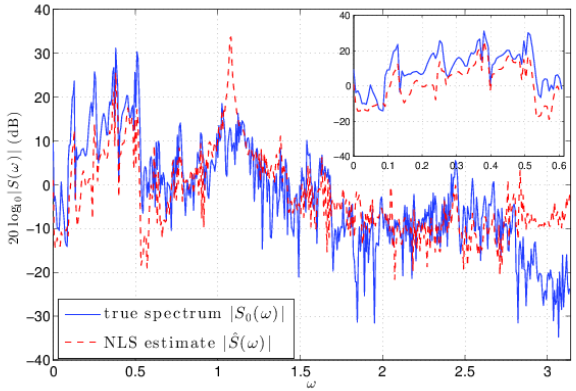
- Estimated RIR for microphone $m = 2$

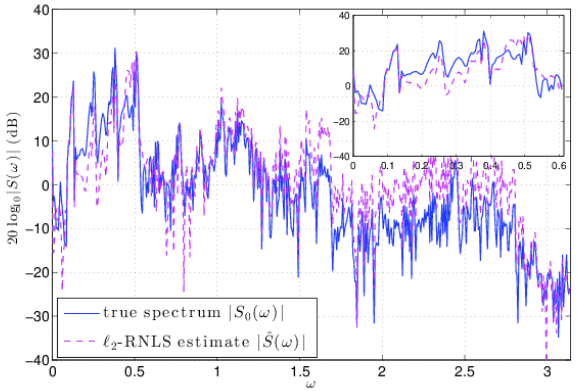


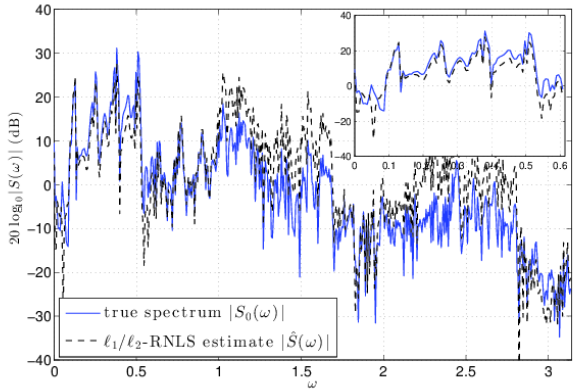


- Observations
  - NLS estimate corresponds to erroneous local solution
  - ℓ_2-norm regularization improves overall estimation performance
    = improved RIR envelope
  - ℓ_1-norm regularization improves local estimation performance
    = improved early RIR coefficients

- Estimated source signal magnitude spectrum







- Observations
  - NLS estimate corresponds to inaccurate local solution
  - ℓ_2-norm regularization improves overall estimation performance
    = improved source signal spectrum envelope
  - ℓ_1-norm regularization improves local estimation performance
    = improved quasi-harmonic source signal components

1 Introduction

2 Problem statement

3 Embedded optimization algorithms

4 Evaluation

5 Conclusion

- Multi-microphone dereverberation by embedded optimization
  - allows more general problem formulations ($\neq$ closed-form estimators)
  - achieves direct source signal estimation ($\neq$ inverse filter design)
  - facilitates use of prior knowledge
  - yields promising simulation results
- Future research challenges
  - batch $\rightarrow$ online (frame-based) processing
  - synthetic $\rightarrow$ realistic RIRs
  - zero $\rightarrow$ non-zero measurement noise
  - general-purpose $\rightarrow$ fast dedicated numerical optimization tools
  - include perceptual criteria in problem formulation
  - evaluate resulting dereverberation performance