Noise Reduction in Robot Audition

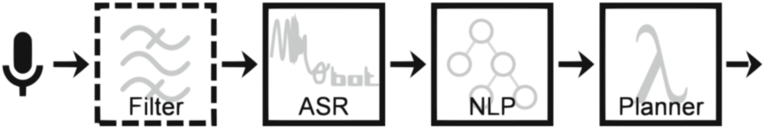
Tanja Flemming

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics
Technical Aspects of Multimodal Systems

16. December 2019

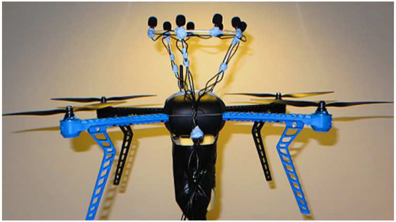
Outline

What is Robot Audition?

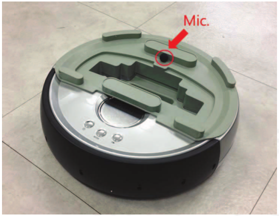


[MHP19]

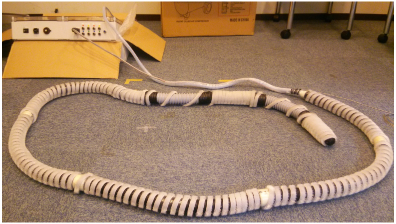
Use Cases



[WC16]



[PRS+14]



[MYM+18]

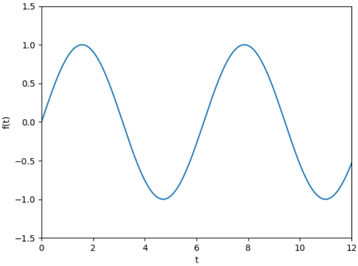
[NAO]

Main Challenges in Robot Audition

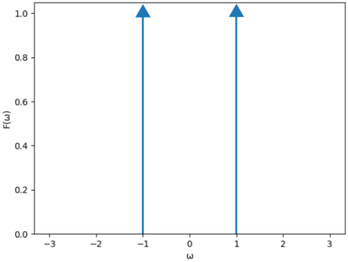
- Real-time processing
- Robustness against noise
  - Background noise
  - Reverberation
  - Ego noise

1D signal characterization:

- Amplitude
- Phase
- Frequency

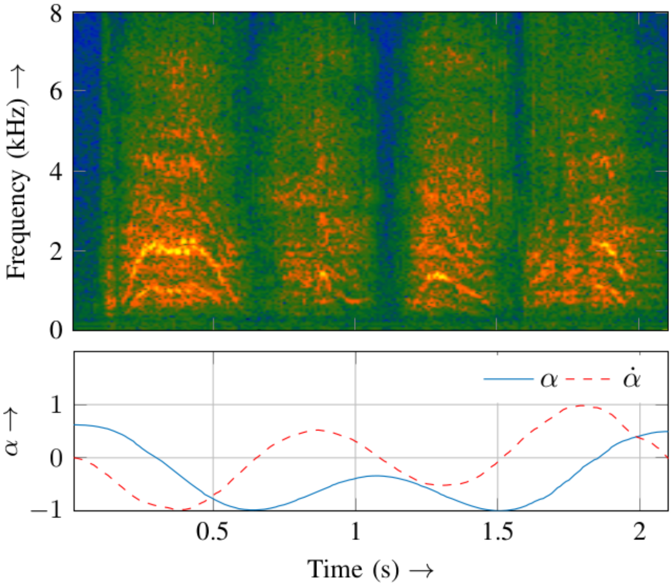


spatial



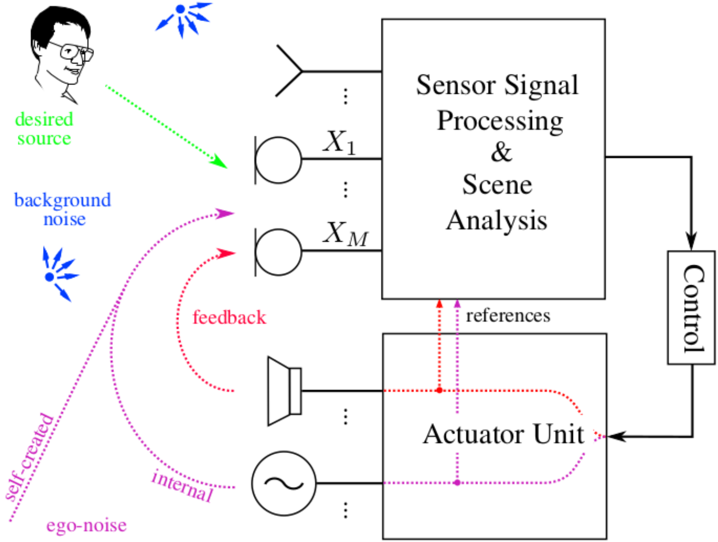
spectral

Local Analysis



[SDK16]

Typical Sound Field



[SLK18]

Presented Approaches

- Dictionary based:

  Ego-Noise Reduction Using a Motor Data-Guided Multichannel Dictionary
  Alexander Schmidt[1], Antoine Deleforge[2] and Walter Kellermann[1]
  2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)

- Matrix Factorization:

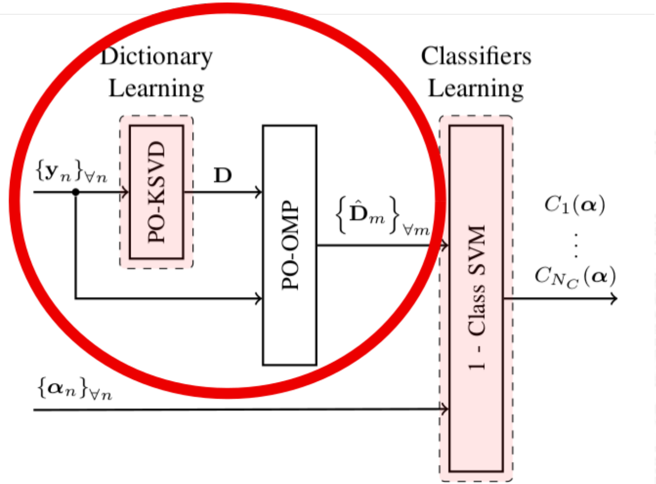
  Multichannel Nonnegative Matrix Factorization for Ego-Noise Suppression
  Thomas Haubner[1], Alexander Schmidt[1] and Walter Kellermann[1]
  2018, Speech Communication; 13th ITG-Symposium

[1]Friedrich-Alexander University, Erlangen-Nürnberg
[2]INRIA center of Rennes, France

Strategy - Dictionary based Approach

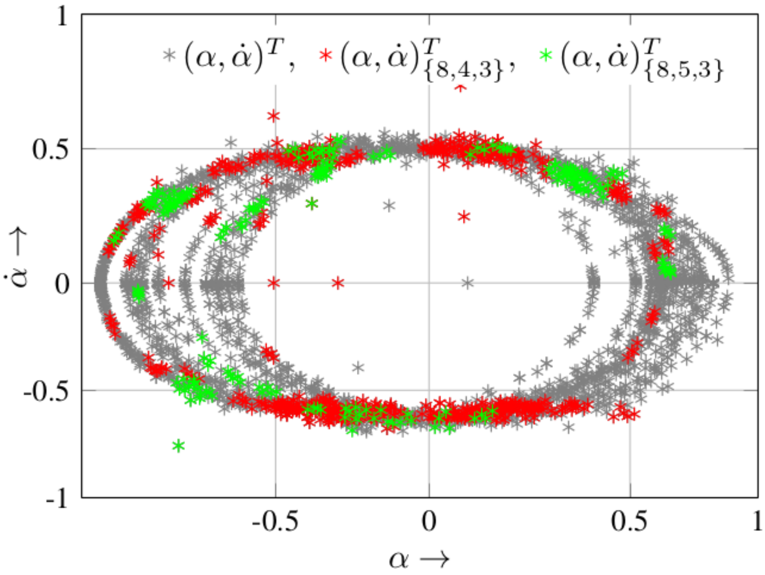
- Capture characteristics of ego noise
- Save prototype signals (atoms) in dictionaries
- Associate motor data to atoms
- Noise removal by subtracting atoms

Approach: Starting Point



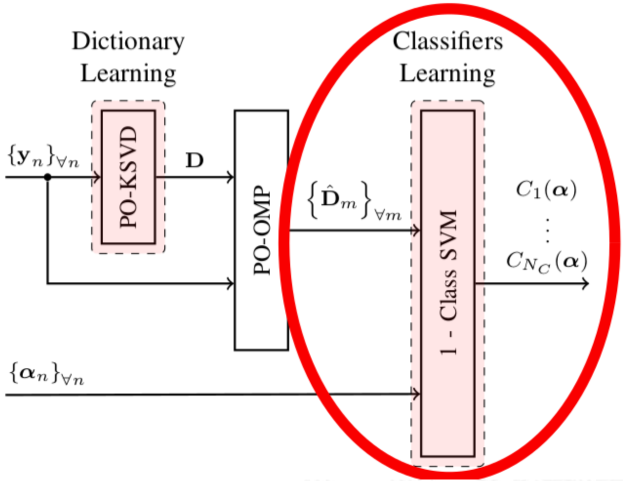
Adapted from [SDK16]

Motor Data - Atom Association



[SDK16]

Adaption



Adapted from [SDK16]

Presented Approaches

- Dictionary based:

  Ego-Noise Reduction Using a Motor Data-Guided Multichannel Dictionary
  Alexander Schmidt[1], Antoine Deleforge[2] and Walter Kellermann[1]
  2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)

- Matrix Factorization:

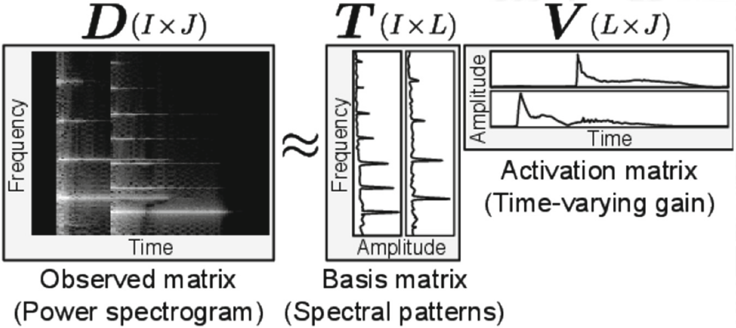
  Multichannel Nonnegative Matrix Factorization for Ego-Noise Suppression
  Thomas Haubner[1], Alexander Schmidt[1] and Walter Kellermann[1]
  2018, Speech Communication; 13th ITG-Symposium

---

[1]Friedrich-Alexander University, Erlangen-Nürnberg
[2]INRIA center of Rennes, France

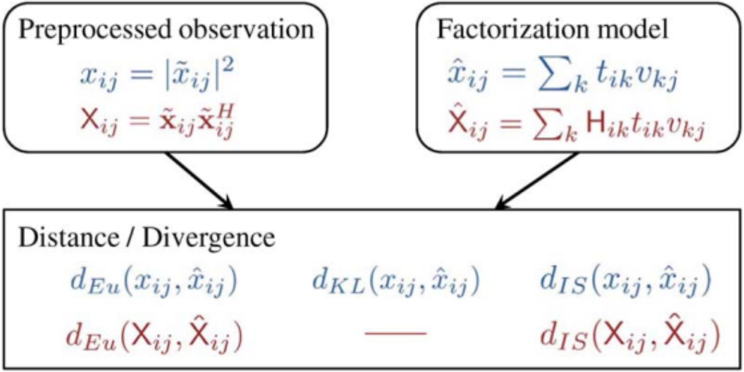
Strategy - Matrix Factorization Approach (MNMF)

- Goal: Separate target source from noise
- Approximate signal with basis and activation matrices
- Minimize difference between original and approximated signal
- Assign bases to noise or speech
- Reconstruct speech signal



[MIM+17]

General MNMF Approach

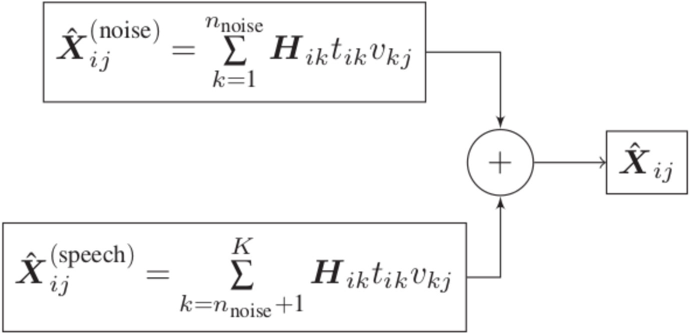


[SKAU13]

blue: single-channel NMF
red: multichannel NMF

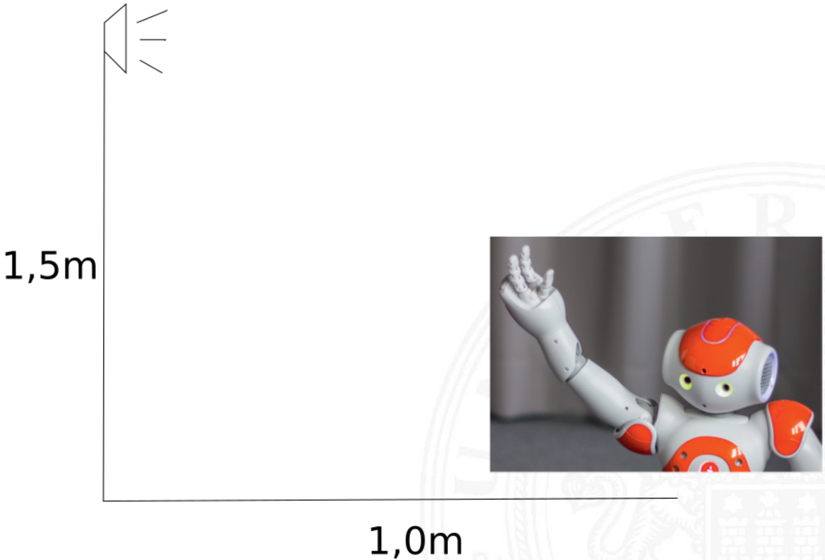
Adaption

1. Learn ego noise model
2. On input signal:
   2.1 Add bases and transfer matrices to model
   2.2 Minimize difference to real signal
   2.3 Assign bases to noise resp. speech
   2.4 Reconstruct speech signal



[HSK18]

Evaluation Experiment



Adapted from [HSK18]

Results

Dictonary based

	SIR [dB]	SDR [dB]	RR[%]
Classifier	14.71	2.64	73.0
PO-OMP	14.46	2.57	71.8
NMF	2.51	0.8	45.2
Unprocessed	-5.48	-8.15	36.1

[SDK16]

Matrix Factorization

	SDR in dB	SIR in dB	SAR in dB
Unprocessed	-2.30	-2.26	258.18
SNMF	6.77	12.59	9.31
ILRMA	8.44	10.96	12.85
Proposed	10.76	23.70	13.37

[HSK18]

SIR: Signal-to-Inference-Ratio

SDR: Signal-to-Distortion-Ratio

Conclusion

Dictionary based

- Good noise suppression
- Fast execution on input signal
- Complex training is needed

Matrix Factorization

- Stronger noise suppression
- Minimization for every incoming signal required
- Complex training is needed

Thank you for your attention.
Do you have any questions?

References

[HSK18] T. Haubner, A. Schmidt, and W. Kellermann, *Multichannel nonnegative matrix factorization for ego-noise suppression*, Speech Communication; 13th ITG-Symposium, Oct 2018, pp. 1–5.

[MHP19] Mauricio Matamoros, Karin Harbusch, and Dietrich Paulus, *From commands to goal-based dialogs: A roadmap to achieve natural language interaction in robocup@home*, RoboCup 2018: Robot World Cup XXII (Cham) (Dirk Holz, Katie Genter, Maarouf Saad, and Oskar von Stryk, eds.), Springer International Publishing, 2019, pp. 217–229.

References (cont.)

[MIM+17] Narumi Mae, Masaru Ishimura, Shoji Makino, Daichi Kitamura, Nobutaka Ono, Takeshi Yamada, and Hiroshi Saruwatari, *Ego noise reduction for hose-shaped rescue robot combining independent low-rank matrix analysis and multichannel noise cancellation*, Latent Variable Analysis and Signal Separation (Cham) (Petr Tichavský, Massoud Babaie-Zadeh, Olivier J.J. Michel, and Nadège Thirion-Moreau, eds.), Springer International Publishing, 2017, pp. 141–151.

References (cont.)

[MYM+18] Narumi Mae, Koei Yamaoka, Y Mitsui, Mitsuo Matsumoto, Shoji Makino, Daichi Kitamura, Nobutaka Ono, T Yamada, and Hiroshi Saruwatari, *Ego noise reduction and sound localization adapted to human ears using hose-shaped rescue robot*, Proc. International Workshop on Nonlinear Circuits, Communications and Signal Processing, 2018, pp. 371–374.

[NAO] *Picture of NAO robot from SoftBank Robotics*, Accessed: 19.12.2019.

References (cont.)

[PRS+14] S. Park, J. Rho, M. Shin, D. K. Han, and H. Ko, *Acoustic feature extraction for robust event recognition on cleaning robot platform*, 2014 IEEE International Conference on Consumer Electronics (ICCE), Jan 2014, pp. 145–146.

[SDK16] A. Schmidt, A. Deleforge, and W. Kellermann, *Ego-noise reduction using a motor data-guided multichannel dictionary*, 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Oct 2016, pp. 1281–1286.

References (cont.)

[SKAU13] H. Sawada, H. Kameoka, S. Araki, and N. Ueda, *Multichannel extensions of non-negative matrix factorization with complex-valued data*, IEEE Transactions on Audio, Speech, and Language Processing **21** (2013), no. 5, 971–982.

[SLK18] A. Schmidt, H. W. Löllmann, and W. Kellermann, *A novel ego-noise suppression algorithm for acoustic signal enhancement in autonomous systems*, 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2018, pp. 6583–6587.

References (cont.)

[WC16] L. Wang and A. Cavallaro, *Ear in the sky: Ego-noise reduction for auditory micro aerial vehicles*, 2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Aug 2016, pp. 152–158.