Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

SP Signal Processing

Prof. Dr.-Ing. Timo Gerkmann

# Signal Processing Group

Signal Processing Group
Department of Informatics
Universität Hamburg

2017

# Zur Person

- 2005, Siemens Corporate Research, Princeton, USA
- 2005-2010, Promotion, IKA, Ruhr-Universität Bochum
- 2010-2011, Postdoktorand, KTH Royal Institute of Technology, Stockholm
- 2011-2015, Juniorprofessor of Speech Signal Processing, Cluster of Excellence Hearing4all, Universität Oldenburg
- 2015-2016, Principal Scientist, Technicolor R&I, Hannover
- seit 2016, Professor of Signal Processing, Universität Hamburg

## Was wir machen

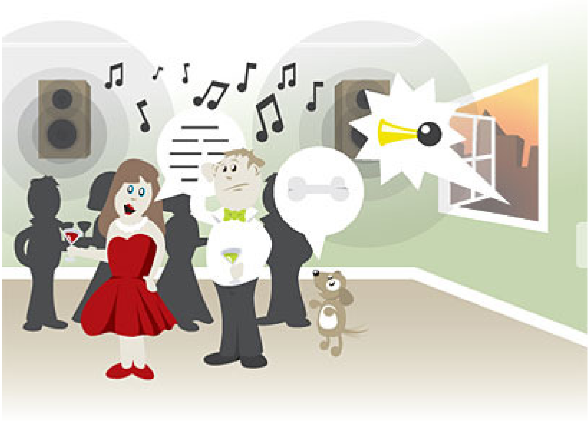
- Signalverarbeitung mit Schwerpunkt auf Audio und Sprachsignale
- Moderne probabilistische Verfahren / Machine Learning
- Beispiele
  - Verbesserung gestörter Signale
  - Quellentrennung
  - Robuste Spracherkennung

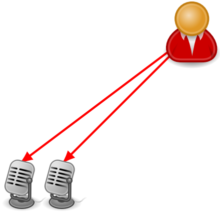
# Signal acquisition in noisy environments



- Speech communication disturbed by external noise sources

➜ Make information more easily accessible by humans and machines

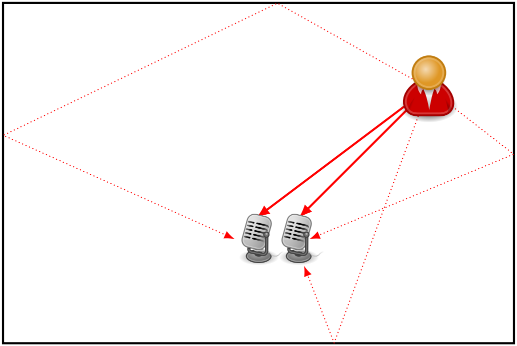




- Signal model: $y_m(n) = s_m(n)$

Goal: Extract desired source $s(n)$ from recorded mixture $y_m(n)$

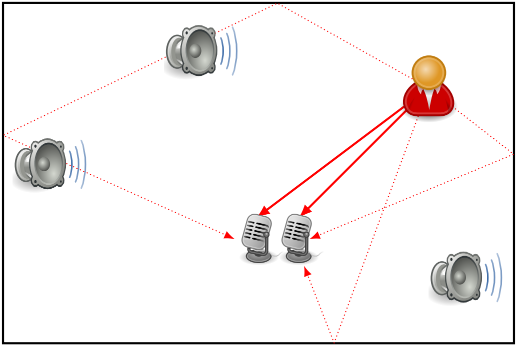
# Signal Model



- Signal model: $y_m(n) = s(n) * h_m$
- Conversation disturbed by
  - reflections from the walls

Goal: Extract desired source $s(n)$ from recorded mixture $y_m(n)$

# Signal Model



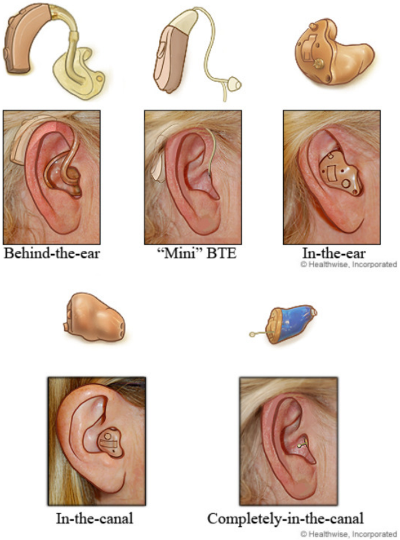
- Signal model: $y_m(n) = s(n) * h_m + \sum_{i=1}^{I} v_{i,m}(n)$
- Conversation disturbed by
  - reflections from the walls
  - additive noise

Goal: Extract desired source $s(n)$ from recorded mixture $y_m(n)$

# Signal Processing in smart phones


- speech coding
- noise reduction
- speech recognition
  - speech control
  - virtual assistent

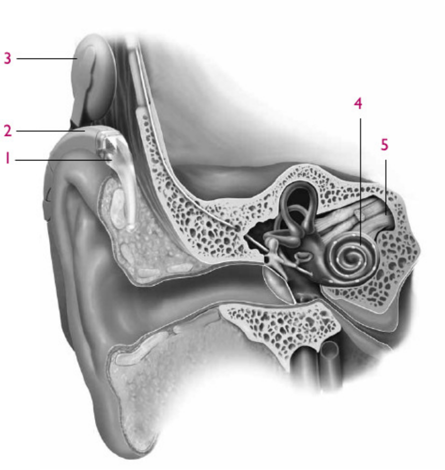
# Signal processing for hearing devices



# Signal processing for hearing devices


Quelle: designaffairs.com

## Cochlear Implants




1 Sounds are picked up by the microphone.

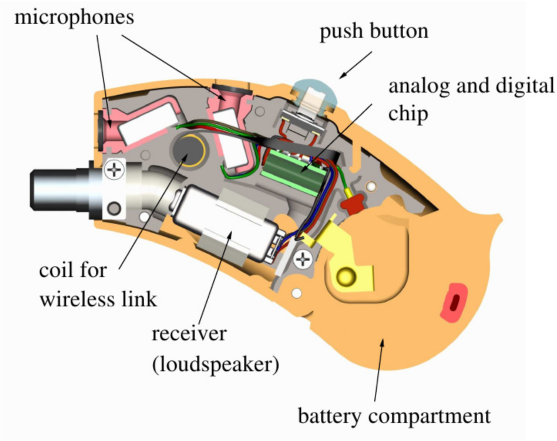
2 The signal is then "coded" (turned into a special pattern of electrical pulses).

3 These pulses are sent to the coil and are then transmitted across the skin to the implant.

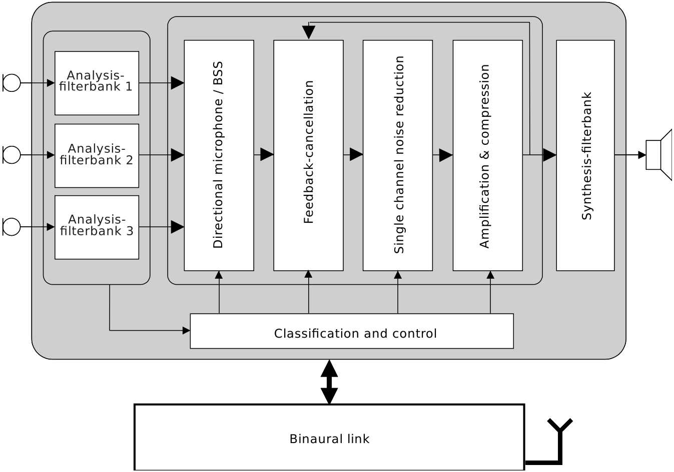
4 The implant sends a pattern of electrical pulses to the electrodes in the cochlea.

5 The auditory nerve picks up these electrical pulses and sends them to the brain. The brain recognizes these signals as sound.

Quelle: Handbook for Educators, Med-El




Quelle: Siemens Audiologische Technik




Hamacher et al. in: Martin et al. (eds.), Wiley 2008

# Signal processing for hearing devices

Quelle: `http://www.nuheara.com/`

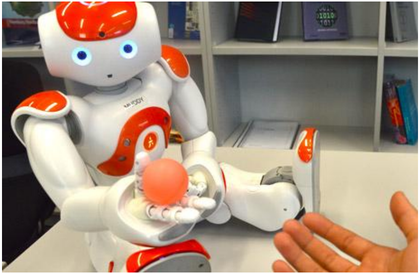
Wireless earbuds for assisted listening

- no prescription needed ▶ much faster time to market
- computations can done on smartphone / cloud

Typical Algorithms/Functionality

- Music Streaming
- Blended Audio Worlds
- Noise Cancellation
- Advanced Speech Amplification (like a hearing aid)

# Human-machine interfaces



- robust speech recognition required

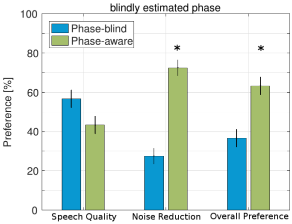
Dr.-Ing. Martin Krawczyk-Becker

- Single microphone speech enhancement
- Short-time Fourier coefficients
  $\rightarrow$ complex-valued
  - Traditionally only the amplitude is modified
  - The spectral phase is neglected
- Contributions:
  1. Novel method for estimating spectral phase
  2. Use spectral phase as additional information for further improvements

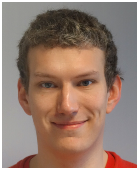
# Phase-Aware Speech Enhancement

Dr.-Ing. Martin Krawczyk-Becker

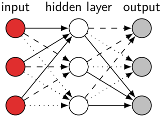
- Single microphone speech enhancement
- Short-time Fourier coefficients
  $\rightarrow$ complex-valued
  - Traditionally only the amplitude is modified
  - The spectral phase is neglected
- Contributions:
  1. Novel method for estimating spectral phase
  2. Use spectral phase as additional information for further improvements
- Results:
  - Listening tests confirmed improved performance
  - Specifically in challenging acoustic scenarios



Robert Rehr, M.Sc.



- Generic approaches:
  - Completely blind estimation of the speech signal
  - ⊕ Robust also in unseen scenarios
- Machine-learning-based:
  - Exploit information available from training (DNNs, NMF, GMM-HMM,...)
  - ⊕ Performance gains when testing and training data match




→ Exploit synergies between generic and machine-learning-based approaches for speech enhancement

## Lasse Vetter, M.Sc.

- Background in
  - automatic mixing and mastering
  - compression and amplification
- Interpretation of DNNs from a DSP perspective
- Machine Listening