UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

DoA-Based Microphone Array Position Self-Calibration Using Circular Statistics

Florian Jacob, Joerg Schmalenstroeer, Reinhold Haeb-Umbach

Department of Communications Engineering - University of Paderborn

May 29, 2013

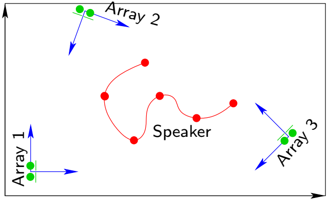
Table of contents

Motivation

Problem statement

- Application areas of spatially distributed microphone arrays:
  - Speaker localization & tracking
  - Signal enhancement (e.g. beamforming)

⇒ Geometry calibration required:
  - Position & orientation of microphone arrays



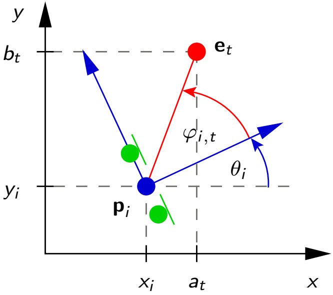
Task

- Automatic calibration procedure
- Input: Reverberant speech signal
- Direction-of-arrival measurements per array (avoid inter array clock sync.)

Overview

System overview

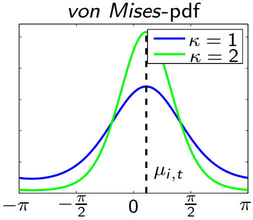
- Unknown parameters:
  - Speaker position $\mathbf{e}_t = [a_t, b_t]$
  - Array position $\mathbf{p}_i = [x_i, y_i]$
  - Array orientation θ_i
- Measured data:
  - Direction of arrival $\varphi_{i,t}$



Observation model of a microphone pair

- Measurement: $\varphi_{i,t} := \mu_{i,t} + \text{Noise}$
- Observation modeled by *von Mises*-pdf:

$$p(\varphi_{i,t}; \mu_{i,t}, \kappa_{i,t}) = \frac{\exp\left(\kappa_{i,t}\cos(\varphi_{i,t} - \mu_{i,t})\right)}{2\pi I_0(\kappa_{i,t})}$$



Maximum-Likelihood Geometry Calibration

Loglikelihood function

- Moving speaker $\Rightarrow$ Multiple measurements to a speaker position unavailable
- Independent observations $\varphi_{i,t}$ from I microphone pairs and T speaker positions:

$$\ell(\boldsymbol{\mu}) := \sum_{t=1}^{T}\sum_{i=1}^{I} \kappa_{i,t} \cos(\varphi_{i,t} - \mu_{i,t})$$

- $\#\text{Unknowns} = 2 \cdot I \cdot T$
- $\#\text{Observation} = I \cdot T$

$\Rightarrow \#\text{Unknowns} > \#\text{Observation}$

How to overcome these drawbacks?

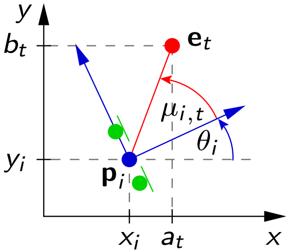
- Incorporate geometric relation:

$$\mu_{i,t} := \mu(\mathbf{p}_i, \theta_i, \mathbf{e}_t) = \operatorname{atan}\left(\frac{b_t - y_i}{a_t - x_i}\right) - \theta_i$$

- Approximate $\kappa_{i,t}$:

$$\kappa_{i,t} \propto d_{i,t} = |\mathbf{e}_t - \mathbf{p}_i|$$

$\Rightarrow \#\text{Unknown}=3 \cdot I + 2 \cdot T$



Maximum-Likelihood Geometry Calibration

Optimization problem

- Resulting equation:

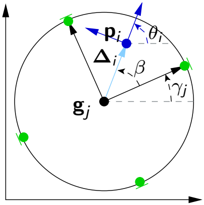
$$\langle \mathbf{p}^*_{2:I}, \theta^*_{2:I}, \mathbf{e}^*_{1:T} \rangle = \underset{\mathbf{p}_{2:I}, \theta_{2:I}, \mathbf{e}_{1:T}}{\operatorname{argmin}} \left\{ \sum_{t=1}^{T} \sum_{i=1}^{I} d_{i,t} \left[1 - \cos\left(\varphi_{i,t} - \mu(\mathbf{p}_i, \theta_i, \mathbf{e}_t) \right) \right] \right\}$$

- Iterative optimization using Newton root finding algorithm
- Recovers array positions, orientations and speaker positions
- Result has arbitrary scaling

Absolute calibration using circular arrays

How to recover the scaling?

- Literature: Knowledge of distance between two sensor nodes
- Proposed: Obtain scaling information from radius of circular microphone arrays



Calibration using circular arrays

- Express position of microphone pair relative to center of circular array:

$$\mathbf{p}_i = \mathbf{g}_j + \boldsymbol{\Delta}_i \text{ and } \theta_i = \gamma_j + \beta_i$$

- Determine position and orientation of circular arrays instead of microphone pairs

$$\langle \mathbf{g}^*_{2:J}, \gamma^*_{2:J}, \mathbf{e}^*_{1:T} \rangle = \underset{\mathbf{g}_{2:J}, \gamma_{2:J}, \mathbf{e}_{1:T}}{\operatorname{argmin}} \left\{ \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{i=1}^{I} d_{i,t} \left[1 - \cos\left(\varphi_{i,t} - \mu(\mathbf{g}_j, \gamma_j, \boldsymbol{\Delta}_i, \mathbf{e}_t) \right) \right] \right\}$$

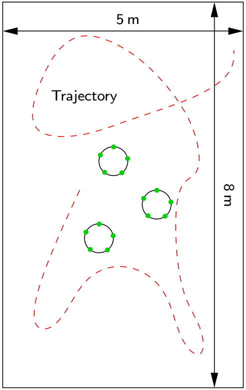
- Apply same optimization as before

Simulation Framework

Scenario

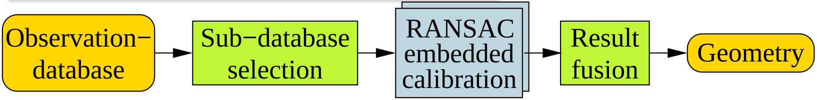
- 3 circular arrays with 5 microphones
- 5 array arrangements, approx. distance of 1 m
- 6 min long speaker trajectory
- Reverberation times form 0 ms to 500 ms
- DoA estimates obtained by adaptive beam-forming, on blocklength of 128 Samples @ 16 kHz

sample sensor configuration



Random **Sa**mple **C**onsenus Framework

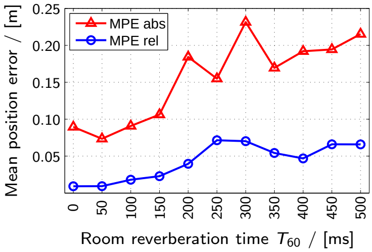
- Outlier rejection
- Handle large data sets

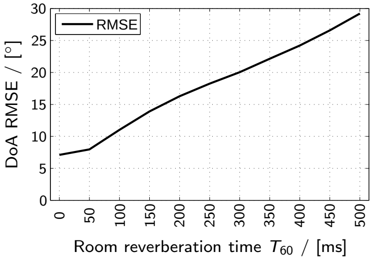


Simulation Results

Error measure

- **M**ean **p**osition **e**rror: Average distance between real and estimated sensor position
- Apply rigid body transformation to match real and estimated positions
- Distinguish between 2 error types:
  - Geometric + Scaling errors (MPE abs): Rotation + Translation
  - Geometric errors (MPE rel): Rotation + Translation + Scaling



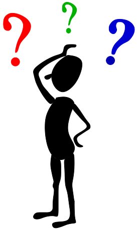


Conclusions

Summary

- Geometry calibration framework base on reverberated speech input
- Circular arrays enable absolute geometry calibration
- ML-Estimator in case of *von Mises* distributed observations
- RANSAC-Framework: Outlier rejection and handling of large data sets
- Calibration Error:
  - Relative positioning error $< 10\,\text{cm}$
  - Absolute positioning error $< 25\,\text{cm}$

  } Even in case of high reverberation



Thank you for your attention!

Questions ?

Florian Jacob

University of Paderborn
Department of Communications Engineering

jacobn@nt.uni-paderborn.de
nt.uni-paderborn.de

This work has been supported by Deutsche Forschungsgemeinschaft (DFG) under contract no. Ha3455/7-1.