3D Single Source Localization Based on Euclidean Distance Matrices

Klaus Brümann, Simon Doclo

Dept. of Medical Physics and Acoustics and Cluster of Excellence Hearing4all, University of Oldenburg, Germany

Introduction

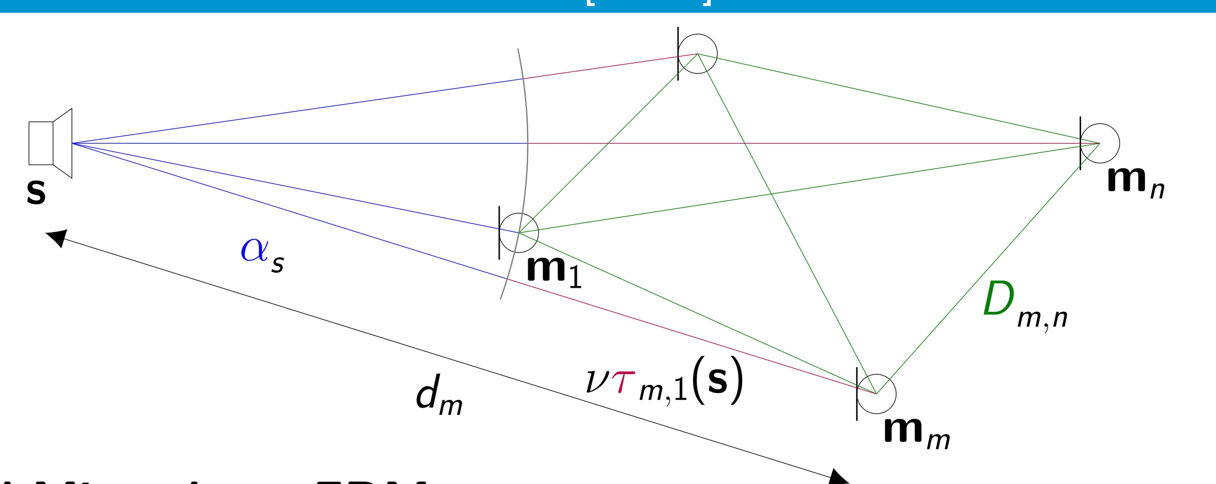
3D source localization approach based on steered response power (SRP) requires optimization of **three continuous position variables**

MAIN IDEAS

- 3D source localization method based on Euclidean distance matrices (EDMs) and estimated time-differences of arrival (TDOAs), which depends on a **single continuous distance variable**
- Consider **multiple candidate TDOA estimates** per microphone pair and select the best combination to improve localization performance in reverberant environments

EDM-Based Source Localization

Properties of Euclidean Distance Matrices [1, 2, 3]



Source and Microphone EDM:

$$\bar{\mathbf{D}} = \begin{bmatrix} \mathbf{D} & \mathbf{d} \\ \mathbf{d}^{\mathrm{T}} & 0 \end{bmatrix}$$

- $\mathbf{D} = [D_{i,j}^2]$: inter-microphone EDM (known distances)
- $\mathbf{d} = [d_1^2, \ldots, d_M^2]^{\mathrm{T}}$: vector of (unknown) squared distances between source and microphones

Gram Matrix:

$$\mathbf{G} = -\frac{1}{2}(\mathbf{I} - \mathbf{1}\mathbf{e}^{\mathrm{T}})\bar{\mathbf{D}}(\mathbf{I} - \mathbf{e}\mathbf{1}^{\mathrm{T}}),$$

For 3D scenarios, the rank of the Gram matrix is **at most 3**

Reconstruction of relative microphones and source positions matrix:

$$\mathbf{P}_{\text{rel}} = \left[\operatorname{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}\right) \mid \mathbf{0}_{3\times((M+1)-3)}\right]\mathbf{U}^{\mathrm{T}},$$

with λ_i and $\mathbf{U}$ eigenvalues and eigenvectors of Gram matrix $\mathbf{G}$.

$\mathbf{P}_{\text{rel}}$ is related to absolute microphones and source positions matrix $\mathbf{P} = [\mathbf{M}|\,\mathbf{s}]$ via arbitrary translation/rotation/reflection

EDM-Based Cost Function

Idea 1: Decompose distance between source and m-th microphone as

$$d_m = \alpha_s + \nu\,\tau_{m,1}(\mathbf{s})$$

Assuming (for now) that TDOAs $\tau_{m,1}(\mathbf{s})$ are available: write distance d_m as function of unknown variable α

$$d_m(\alpha) = \alpha + \nu\,\tau_{m,1}(\mathbf{s})$$

Formulate **cost function** in single variable α using all but 3 largest eigenvalues $\lambda_i(\alpha)$ of Gram matrix $\mathbf{G}(\alpha)$ and minimize to determine α_s

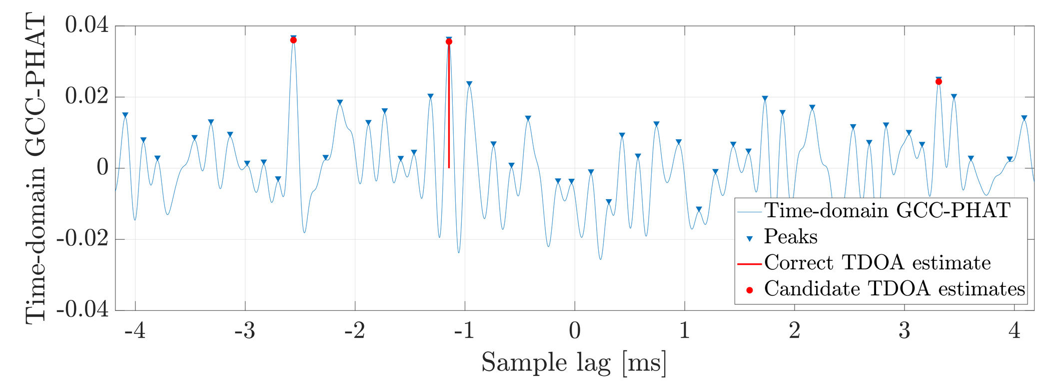
$$J(\alpha) = \sum_{i=3+1}^{M+1} |\lambda_i(\alpha)| \qquad \Rightarrow \alpha_s = \underset{\alpha}{\operatorname{argmin}}\; J(\alpha)$$

TDOA Selection

Idea 2: Consider multiple candidate TDOA estimates per microphone pair

$$d_m(\alpha, \hat{\tau}_{m,1}^{c_m}) = \alpha + \nu\,\hat{\tau}_{m,1}^{c_m}, \quad c_m = 1, \ldots, C$$

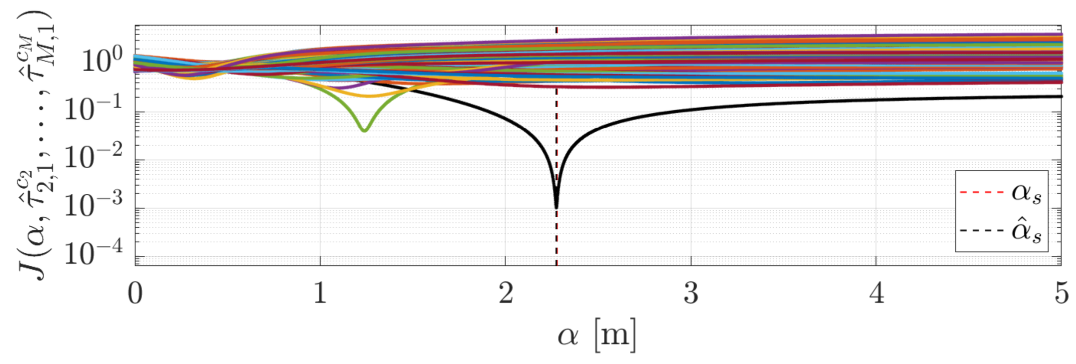
where TDOAs are estimated using GCC-PHAT [4]



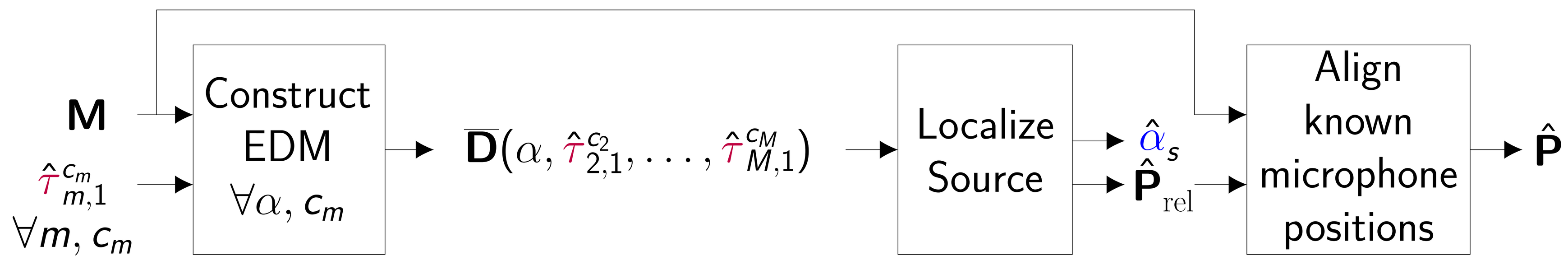
Formulate cost function in single variable using candidate TDOA estimates

$$\hat{\alpha}_s = \underset{\alpha, c_2, \ldots, c_M}{\operatorname{argmin}}\; J(\alpha, \hat{\tau}_{2,1}^{c_2}, \ldots, \hat{\tau}_{M,1}^{c_M})$$

The optimal variable α_s is estimated via a (discretized) search of the continuous parameter α and $M-1$ discrete parameters $c_2, \ldots, c_M$, e.g.,



Overview of Proposed Method:



Experimental Evaluation

Framework and Acoustical Parameters

- Acoustic scenarios simulated with RIR generator [5]
- $6 \times 6 \times 2.4$ m room with equally reflective walls and DRR $= 0$ dB (avg. in each mic.)
- Microphones randomly positioned within cube with cube length 2 m
- Four different distances between source and centroid of the microphones were simulated: $\alpha_c \in \{0.5, 1, 2, 3\}$ m
- 100 acoustic scenarios for each distance α_c (random 5 s speech signal, array location & geometry, and speech source location), with babble noise at SNR $= 5$ dB
- Sampling frequency 16 kHz, 512 sample (32 ms) frame length, 50% overlap between frames, 1024 sample FFT-length

Experiment (Results Below):

- Analysis of localization error $\varepsilon_s = \|\mathbf{s} - \hat{\mathbf{s}}\|$
- Comparison of proposed EDM-based approach with up to 3 candidate TDOA estimates per microphone pair with SRP-PHAT-based method (similar to [6])

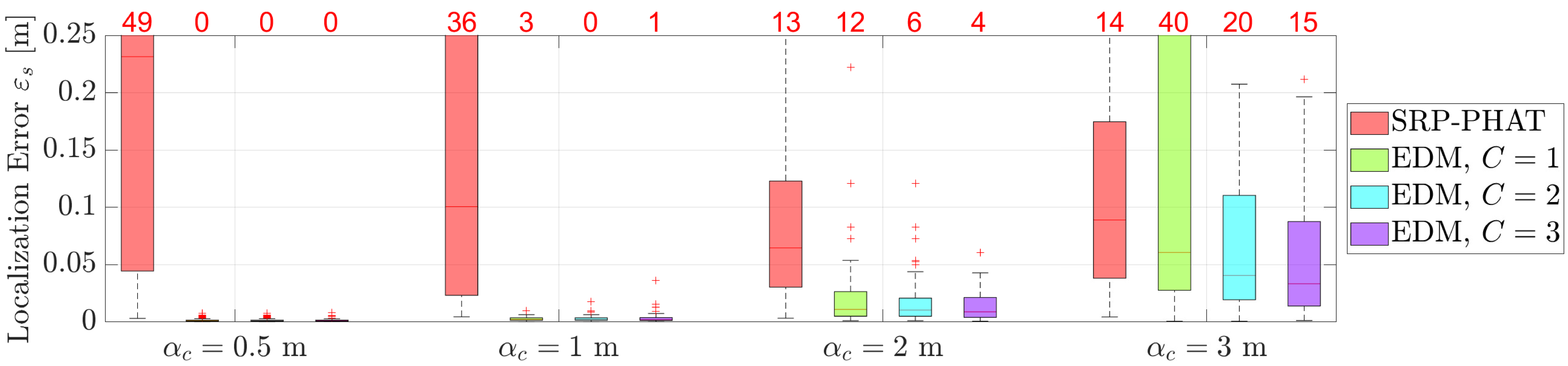
Conclusions and Outlook

Conclusions

- ✓ Proposed EDM-based method results in lower median error for all distances
- ✓ Proposed candidate TDOA estimate selection further reduces localization error

Outlook

- EDM-based DOA estimation
- EDM-based 3D multi source localization



Box plots of the localization errors ε_s (over 100 scenarios) for the SRP-PHAT method and the EDM-based method (with different numbers of candidate TDOA estimates C per microphone pair), for different distances α_c between the source and the centroid of the distributed microphones. The number of results outside of the plotted range are denoted by red numbers at the top

References

[1] W. S. Torgerson, "Multidimensional scaling: I. theory and method," *Psychometrika*, vol. 17, no. 4, pp. 401–419, 1952.

[2] J. C. Gower, "Euclidean distance geometry," *Math. Sci*, vol. 7, no. 1, pp. 1–14, 1982.

[3] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, "Euclidean distance matrices: essential theory, algorithms, and applications," *IEEE Signal Processing Magazine*, vol. 32, no. 6, pp. 12–30, 2015.

[4] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. on Audio, Speech, Language Processing*, vol. 24, no. 4, pp. 320–327, 1976.

[5] E. A. P. Habets, *RIR-Generator*, available at *https://github.com/ehabets/RIR-Generator*.

[6] H. Do and H. F. Silverman, "A fast microphone array SRP-PHAT source location implementation using coarse-to-fine region contraction (CFRC)," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, 2007, pp. 295–298.

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2177/1 - Project ID 390895286 and Project ID 352015383 - SFB 1330 B2.