Acoustic Source Localization With Distributed Asynchronous Microphone Networks

A. Canclini, F. Antonacci, A. Sarti, *Member, IEEE*, and S. Tubaro, *Member, IEEE*

Presented by Sarit Ben-Sultan

Outline

- Introduction
- Article algorithm
- Simulation
- Algorithm drawback
  - sensitivity to reverberations
- Suggested extension
- Further work

Introduction

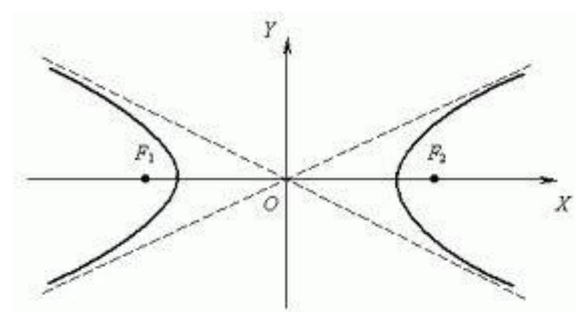
- Source localization:
  - Finding the spatial location of an acoustic source, usually a speaker.
- Application:
  - Automated camera steering towards the speaker in video conferences.

Article algorithm

- The method is based on data recorded from distributed independent (unsynchronized) acoustic sensors.
- Each sensor is made of two or more synchronized microphones.

Article algorithm

- Algorithm stages:
  - Time Differences of Arrival (TDOAs)
  - Geometric constraints – hyperbola
  - Combining the constraints from different sensors



Article algorithm

- Algorithm parameters:
  - M sensors, m=1,2,…,M
  - L microphones in sensor, l=1,2,…,L
  - Microphones location $\boldsymbol{x}_l^m = (x_l^m, y_l^m)^T$
  - Source location $\boldsymbol{x}_s = (x_s, y_s)^T$
  - Sample frequency fs

Article algorithm

- TDOAs calculation using GCC-PHAT

$$\hat{n}_{lj}^{m} = \arg\max_{n} \Gamma_{lj}^{m}(n) \quad \text{s.t.} \quad n \in [-n_{\max}, n_{\max}]$$

  - GCC-PHAT improve robustness against reverberation in comparison with CC.
  - Knapp and Carter (1976)

→ $\hat{\tau}_{lj}^{m} = (\hat{n}_{lj}^{m})/(f_s)$

Article algorithm

- Geometric constraints
  - Let us look at 2 microphone of sensor m in locations $(x_l^m, y_l^m)^T$ and $(x_j^m, y_j^m)^T$
  - The distance to source difference between the microphones signals is $T_{lj}^m = c\hat{\tau}_{lj}^m$
  - That constrains the source to lie on the hyperbola:

(1) $$\sqrt{(x_l^m - x)^2 + (y_l^m - y)^2} - \sqrt{(x_j^m - x)^2 + (y_j^m - y)^2} = T_{lj}^m$$

Article algorithm

- Geometric constraints
  - Wanting to write the equation in the form

$$a_{lj}^{m}x^2 + b_{lj}^{m}xy + c_{lj}^{m}y^2 + d_{lj}^{m}x + e_{lj}^{m}y + f_{lj}^{m} = 0$$

  - We calculate the coefficients according to equation (1).

Article algorithm

- Combining the constraints
  - We define

$$\varepsilon_{lj}^{m}(\boldsymbol{x}) = [a_{lj}^{m}x^2 + b_{lj}^{m}xy + c_{lj}^{m}y^2 + d_{lj}^{m}x + e_{lj}^{m}y + f_{lj}^{m}]$$

  - We stack all the $\varepsilon_{lj}^{m}(\boldsymbol{x})$ in a column vector $\varepsilon(\boldsymbol{x})$

And get the minimization problem:

(2) $$\hat{\boldsymbol{x}}_s = \arg\min_x J(\boldsymbol{x}), \quad J(\boldsymbol{x}) = \varepsilon(\boldsymbol{x})^T \varepsilon(\boldsymbol{x})$$

Article algorithm

- Combining the constraints
  - To solve (2) we adapt an iterative technique using Taylor series expansion

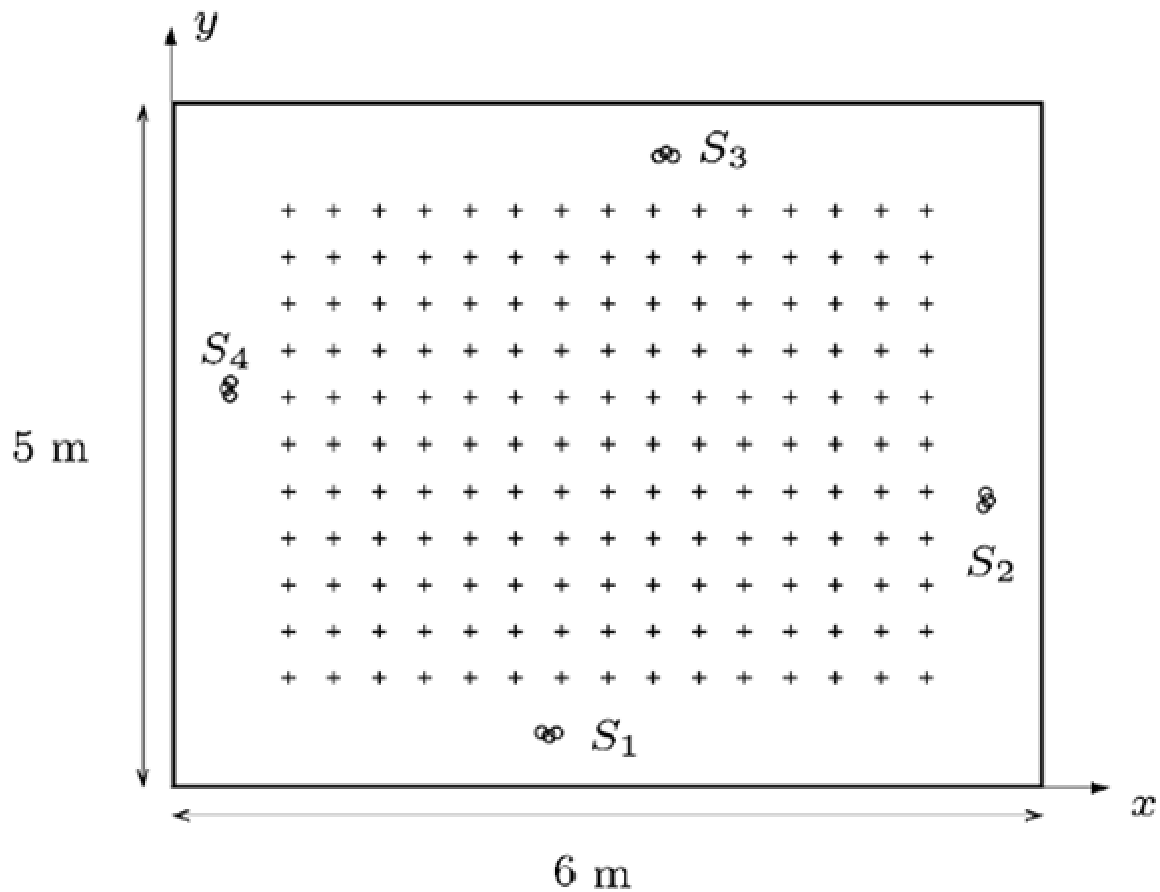
$$\varepsilon(\boldsymbol{x}) \simeq \varepsilon(\boldsymbol{x}_{s,0}) + \nabla\varepsilon\,\Big|_{\boldsymbol{x}_{s,0}} \cdot (\boldsymbol{x} - \boldsymbol{x}_{s,0})$$

$$\nabla\varepsilon = \begin{bmatrix} \frac{\partial\varepsilon^{1}_{1,2}}{\partial x} & \frac{\partial\varepsilon^{1}_{1,2}}{\partial y} \\ \vdots & \vdots \\ \frac{\partial\varepsilon^{M}_{L-1,L}}{\partial x} & \frac{\partial\varepsilon^{M}_{L-1,L}}{\partial y} \end{bmatrix}$$

$$\hat{\boldsymbol{x}}_{s,i+1} = \hat{\boldsymbol{x}}_{s,i} - \nabla\varepsilon^{\dagger}\,\Big|_{\hat{\boldsymbol{x}}_{s,i}} \varepsilon(\boldsymbol{x}_{s,i})$$

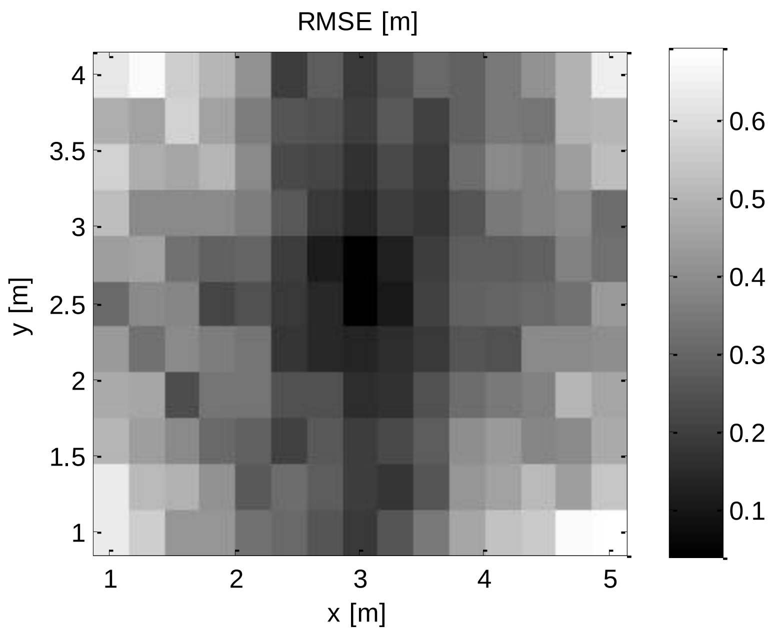
Simulation

- M=4, L=3, room size=5mx6m



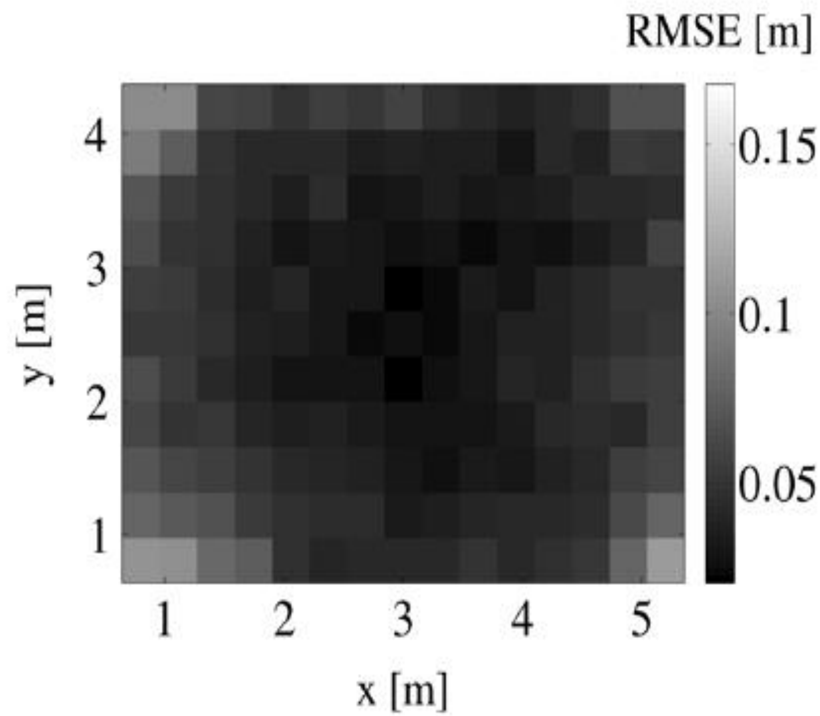
Simulation - moderately reverberant room

My simulation



T60=0.4 sec
Time it takes the sound to decay by 60dB

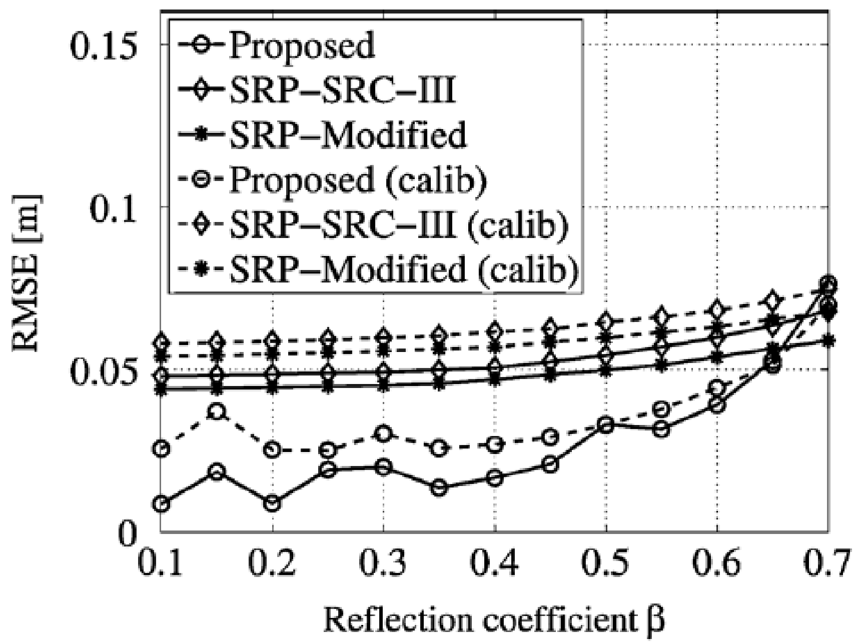
Article results



$\beta = 0.5$
Reflection coefficient

Algorithm drawback

- Sensitivity to reverberations



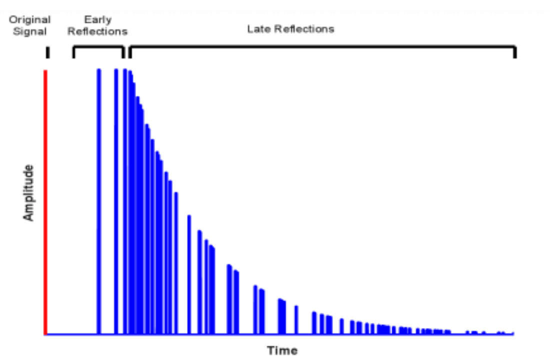
As β grows the RMSE grows

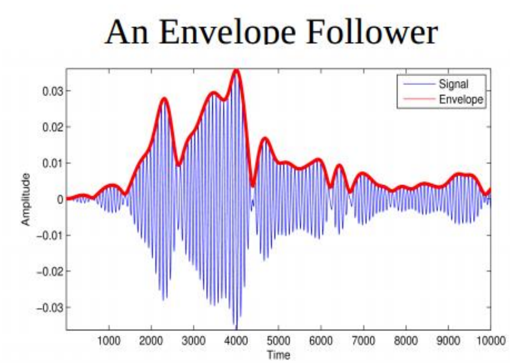
Algorithm drawback

- For source location at (x,y)=(1,4)
  - T60=0[sec] – RMSE=0.096[m]
  - T60=0.4[sec] – RMSE=0.626[m]
  - T60=1[sec] – RMSE=0.8528[m]

Suggested extension

- Dereverberation before using the algorithm
  - A Simple Blind Dereverberation Approach
    - Focus only on late reflections
    - Use signal envelopes
    - Envelope of late reflections has an exponential decay form: $\exp(-1/\tau)^n = a^n$

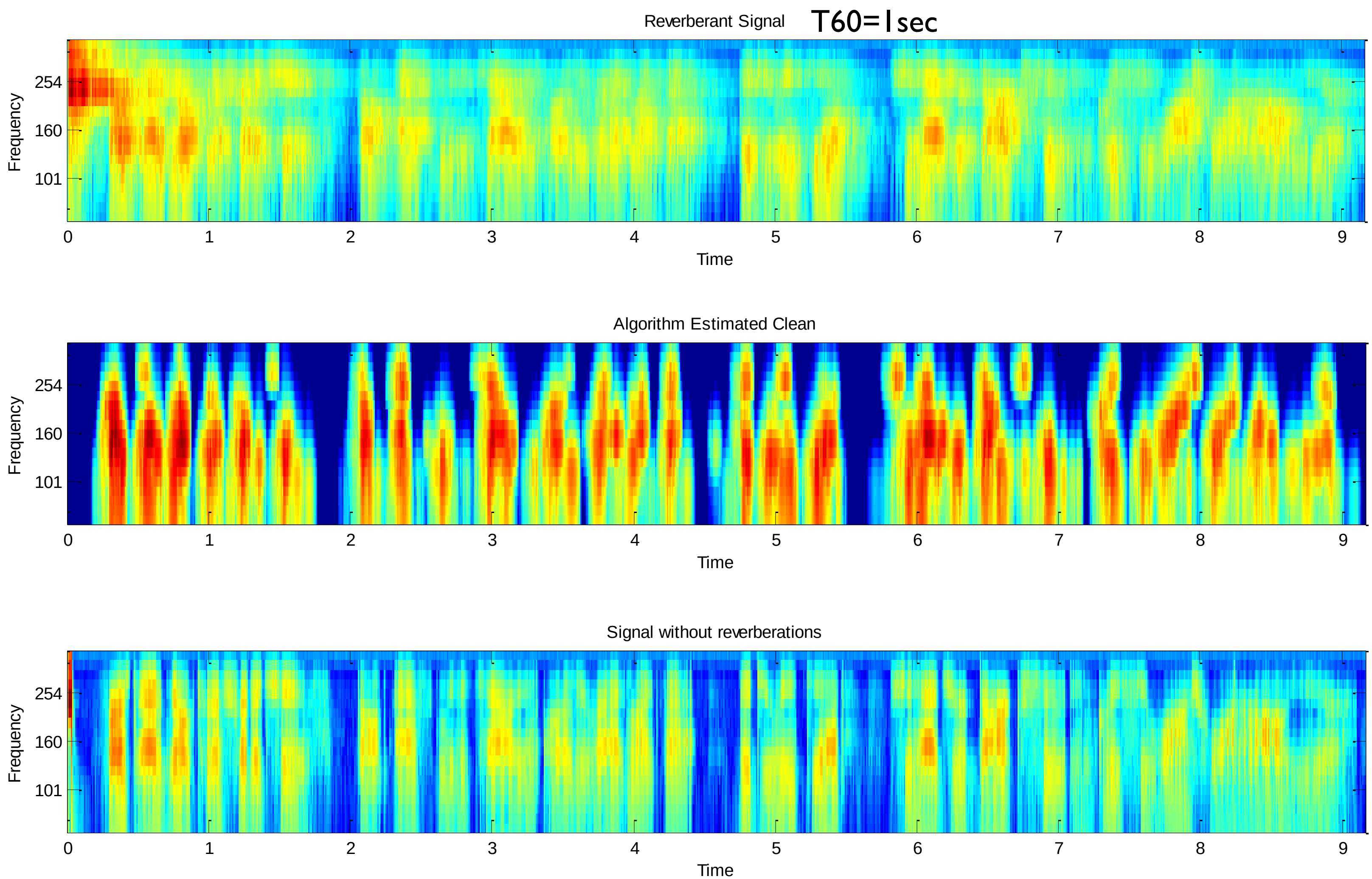




Suggested extension

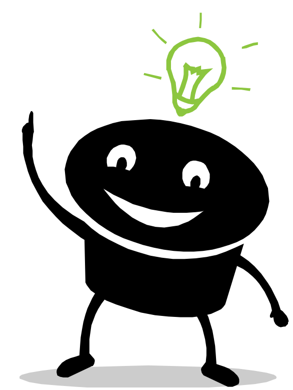
- A Simple Blind Dereverberation Approach (cont.)
  - Basic idea:
    - Estimate a from the envelope of the reverberant signal
    - Use estimate to find and remove regions of the reverberant signal that are mostly reverberant decay

Suggested extension



Further work

- Using a more accurate dereverberation technique:
  - LIME method:
    - M. Delcroix, T. Hikichi, and M. Miyoshi, "Precise dereverberation using multichannel linear prediction," *IEEE Trans. Audio, Speech and Language Process.*, vol. 15, no. 2, pp. 430-440, Feb. 2007.




Thank You!