Using Global Optimization Methods for Acoustic Source Localization

Malgoezar, Anwar; Snellen, Mirjam; Simons, Dick; Sijtsma, Pieter

Publication date
2016

Document Version
Final published version

Published in
Proceedings of the 23rd International Congress on Sound and Vibration

23rd International Congress on Sound & Vibration
Athens, Greece 10-14 July 2016
ICSV23

USING GLOBAL OPTIMIZATION METHODS FOR ACOUSTIC SOURCE LOCALIZATION

Anwar Malgoezar, Mirjam Snellen and Dick Simons

Delft University of Technology, Faculty of Aerospace Engineering, Kluyverweg 1, 2629 HS, Delft, The Netherlands
email: a.m.n.malgoezar@tudelft.nl

Pieter Sijtsma

PSA3 Advanced AeroAcoustics, Prinses Margrietlaan 13, 8091 AV, Wezep, The Netherlands

Conventional beamforming is a common method to localize sound sources with a microphone array. The method, which is based on the delay-and-sum beamforming, provides an estimate value for the source strength at a given spatial position. It suffers from low spatial resolution at low frequencies, high side lobe levels and requires the user to initialize a two dimensional scan area at a certain distance from the array which can trouble source identification. In this work we use the global optimization method Differential Evolution to efficiently search for source locations. The source locations maximizes the agreement between the modelled signal and measurement. This method also allows for inclusion of more unknowns, such as environmental parameters or a search in three dimensions. Using simulated data, results show that the acoustic source can be identified very accurately with good spatial resolution.

1. Introduction

Beamforming is a widely applied method for imaging noise sources [1,2]. The method is based on the phase differences between the microphones and their known locations. The general approach is to define a number of scan points and estimate the source amplitude for each point. These amplitudes are often depicted in a so-called source map, i.e. an image where high levels indicate the presence of a source.

In general, side-lobes and potentially grating lobes are visible in the source plots as well, indicating high levels at locations where actually no sound source is present. This approach of calculating the levels for all scan points can be considered as an exhaustive search for those locations with high levels. A major drawback of such an exhaustive search is that it restricts the problem to a limited number of unknowns. Typically, beamforming is applied for source determination in two dimensions, often using a scan plane parallel to the array at a fixed distance.

In this work we use optimization methods to overcome this drawback with the possibility to extend this search to more unknowns, for example, the speed of sound. The presence of sidelobes, however, might result in the optimization often converging to these "local optima" while missing the global optimum. In contrast to local search methods, global optimization methods have the capability to escape local optima. This ability is essential for the application considered with sidelobes present. We introduce the use of global optimization method Differential Evolution [3,4] to locate the source positions. The method will be applied to a case of benchmark data for array methods [5]. A case with a single acoustic source is considered.

2. Signal model

To derive the signal model a wavefield is considered generated by L acoustic monopoles where each monopole is located at $\mathbf{x}_{S,l}$ and $l=1,\ldots,L$. Let the microphone position m be given as $\mathbf{x}_m$, where $m=1,\ldots,M$ and M the total number of microphones in the array. The $M\times 1$ array output vector $\mathbf{y}$ for frequency ω is [6]

$$\mathbf{y}(\omega)=\sum_{l=1}^{L}\mathbf{a}_l(\mathbf{x}_{S,l},\omega)s_l(\omega), \tag{1}$$

where $\mathbf{a}_l=\left[a_{l,1},\cdots,a_{l,M}\right]^T$ is the steering vector, s_l the acoustic waveform of source l and $[\cdot]^T$ denotes the transpose of the vector. The element of $\mathbf{a}_l$ for a microphone m is given by

$$a_{l,m}=\frac{r_{l,0}}{r_{l,m}}e^{-j\omega\left(r_{l,m}-r_{l,0}\right)/c}, \tag{2}$$

with c the speed of sound, $r_{l,m}=\left|\mathbf{x}_{S,l}-\mathbf{x}_m\right|$ and $r_{l,0}=\left|\mathbf{x}_{S,l}-\mathbf{x}_0\right|$ the distance between the source and microphone and distance between the source and the array centre location, respectively. Eq.(1) can also be written as

$$\mathbf{y}(\omega)=\mathbf{A}(\omega)\mathbf{s}(\omega), \tag{3}$$

with $\mathbf{A}(\omega)=[\mathbf{a}_1(\mathbf{x}_{S,1},\omega),\ldots,\mathbf{a}_L(\mathbf{x}_{S,L},\omega)]$ the $M\times L$ steering matrix and $\mathbf{s}(\omega)=[s_1(\omega),\ldots,s_L(\omega)]^T$ the signal waveforms. The cross spectral matrix (CSM) of the received microphone signals is then given as

$$\mathbf{C}=\mathbf{y}(\omega)\mathbf{y}^H(\omega), \tag{4}$$

where $(\cdot)^H$ denotes the conjugate transpose of the argument. Using both Eq.(3) and Eq.(4), $\mathbf{C}$ can be written as

$$\mathbf{C}=\mathbf{APA}^H, \tag{5}$$

where

$$\mathbf{P}=\mathbf{s}(\omega)\mathbf{s}^H(\omega). \tag{6}$$

For uncorrelated sources $\mathbf{P}$ is a diagonal matrix where each element presents the power of a source.

In general the covariance matrix is obtained either from actual measurements or synthesized from model predictions. For this research the data considered is synthetic data and is obtained from benchmark cases that were generated in the framework of the workshop Benchmarking Array Analysis Methods in Dallas 2015 [5].

The approach taken in delay-and-sum beamforming is to treat each grid point as a potential source location and estimate the source strength at the grid point as

$$\tilde{y}(\mathbf{x}',\omega)=\frac{1}{M^2}\mathbf{a}^H(\mathbf{x}',\omega)\mathbf{C}(\omega)\mathbf{a}(\mathbf{x}',\omega), \tag{7}$$

where $\tilde{y}$ is the beamformer output, i.e, the estimate for the source strength at grid point $\mathbf{x}'$. Using Eq.(7) is known as conventional beamforming.

For the work considered in the current contribution a different approach is taken. Here, Eq.(3) and alternatively Eq.(5) are used as the forward model. An objective function, sometimes denoted as energy function, is defined that provides a measure for the difference between the measured CSM and that predicted from Eq.(3) or Eq.(5) given a set of values for the unknown parameters. The energy function used in this work is defined as [7]

$$E_{csm}(\mathbf{g}) = \sum \left\{ \left[\mathrm{real}(\mathbf{C}) - \mathrm{real}(\mathbf{C_g}) \right]^2 + \left[\mathrm{imag}(\mathbf{C}) - \mathrm{imag}(\mathbf{C_g}) \right]^2 \right\} \tag{8}$$

where $\mathbf{C}$ is the cross spectral matrix as provided by the workshop, $\mathbf{C_g}$ is the modelled covariance matrix corresponding to parameter vector $\mathbf{g}$ containing the trial values for the unknown parameters. For example, in the case of one source it could have the form of $\mathbf{g} = \mathbf{g}(\mathbf{x}_{S,1}, s_1)$ which would be 4 parameters considering only the spatial position and amplitude of the source. The summation for Eq.(8) is over all *M*x*M* elements of the matrices containing the differences between $\mathbf{C}$ and $\mathbf{C_g}$. The covariance matrices are for specific frequency ω.

3. Differential Evolution

Differential evolution (DE) is a method that optimizes a problem by iteratively trying to improve candidate solutions with regard to a given measure of quality [4]. The subsequent iterations are denoted as generations. DE makes use of a population of candidate solutions per generation and creates new candidate solutions by combining existing ones, and then keeping improved candidate solutions. For creating the new candidate solutions for the next generation, promising solutions of the current population are selected. Still, to allow for escaping local optima, also less good solutions have a probability of being selected for creating new candidate solutions. This probability decreases for subsequent generations.

DE starts with an initial population of randomly chosen parameter value combinations. The population consists of q members, each containing trial values for the unknown parameters. At each generation, a partner population is created from the population members $\mathbf{g}_{k,u}$ as

$$\mathbf{h}_{k,u} = \mathbf{g}_{k,r_1} + F(\mathbf{g}_{k,r_2} - \mathbf{g}_{k,r_3}), \tag{9}$$

with $u, r_1, r_2, r_3 \in \{1, 2, \ldots, q\}$, integer and mutually exclusive and F a scalar multiplication factor between 0 and 1. The values for r_1, r_2, r_3 are chosen at random. A higher value of F indicates an increased difference between original parameter values $\mathbf{g}_{k,u}$ and those contained in the partner population $\mathbf{h}_{k,u}$.

The next step is to calculate its descendant $\mathbf{d}_{k,u}$ by applying crossover to $\mathbf{g}_{k,u}$ and $\mathbf{h}_{k,u}$ with a probability p_C. For each parameter v of $\mathbf{d}_{k,u}$ we get

$$d_{k,uv} = \begin{cases} g_{k,uv} & \text{if } rand[0,1]_v \le p_c \\ h_{k,uv} & \text{if } rand[0,1]_v > p_c \end{cases}, \tag{10}$$

with $rand[0,1]_v$ the v-th evaluation of the uniform distribution with values between 0 and 1. Setting the value of p_C high means that more values are replaced by the partner population, while a low value of p_C results in generations that differ only slightly regardless of the value of F.

To create the new generation $k+1$ from the previous generation k, the member $\mathbf{g}_{k,u}$ is replaced by $\mathbf{d}_{k,u}$ only if it yields a smaller value for the energy function E as

$$\mathbf{g}_{k+1,u} = \begin{cases} \mathbf{d}_{k,u} & \text{if } E(\mathbf{d}_{k,u}) < E(\mathbf{g}_{k,u}) \\ \mathbf{g}_{k,u} & \text{if } E(\mathbf{d}_{k,u}) \geq E(\mathbf{g}_{k,u}) \end{cases}. \tag{11}$$

Doing this for all members u in the population we obtain the next generation $k+1$. This process is repeated for N_G generations.

The performance of global optimization methods, i.e. their success in locating the global optimum in an efficient way, is dependent on a number of so-called setting parameters. For DE these are

- Population size q,
- Multiplication factor F,
- Crossover probability p_C,
- Number of generations N_G.

These settings must be set beforehand to suitable values, and can be problem specific, to maximize the probability to locate the global optimum. In this work the best values for the parameters were found to be q = 128, F = 0.35, p_c = 0.75 and N_G = 600.

4. Results

For the test case a single monopole source is considered, located at $\mathbf{x}_S$ = (0.3 m, 0.4 m, 1.0 m) with source amplitude of 1 Pa. The array consists of 48 microphones. Figure 1 shows the array geometry. The data provided is simulated data and consists of the cross spectral matrix of the microphone measurements at the frequencies 500 Hz to 6000 Hz in 500 Hz steps.

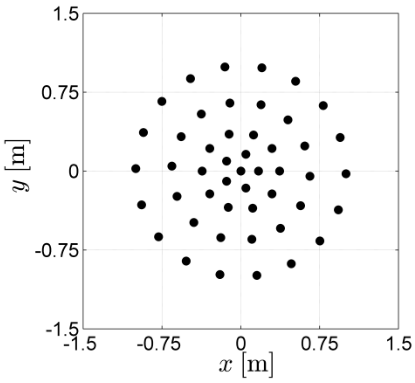

Figure 1: Array geometry for test case 1.

As a reference we start with applying conventional beamforming for which we define a scan plane at the source location parallel to the array. The plane was set at z = 1.0 m. Beamforming is performed for 500 and 5000 Hz. It can be seen that a 500 Hz signal has a wider main lobe than 5000 Hz, but many more side lobes can be seen at 5000 Hz in the 54 to 72 dB range.

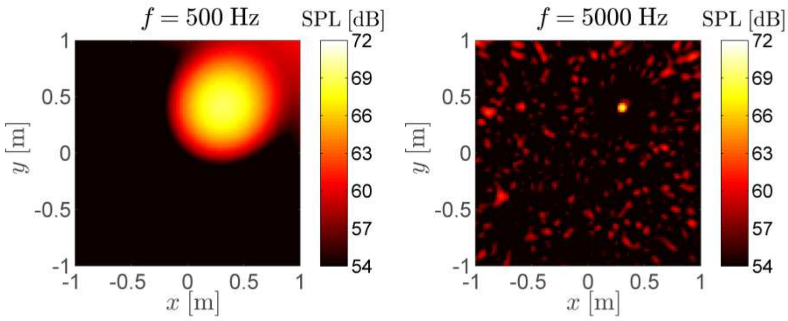

Figure 2: Beamforming for 500 and 5000 Hz.

Figure 3 shows the results for the same frequencies by using the proposed inversion method, employing the energy function. To obtain this result the settings of DE were set to q = 128, N_G = 600, p_C = 0.75 and F = 0.35. The number of independent runs was set to 50. For both frequencies it can be seen that the source position is retrieved correctly, since the values for x, y and z that correspond to the lowest energy values are in agreement with the true source position. The value for the amplitude is obtained correctly at 1 Pa. To show the rate of convergence, **Figure 4** is given, where the energy is given for the number of generations. For most runs convergence to the correct position is achieved well within the 600 generations. For 5000 Hz, three runs are seen not to have reached zero energy value at 600 generations. This corresponds to runs being stuck in a local optima. Having additional generations for higher frequencies can solve this.

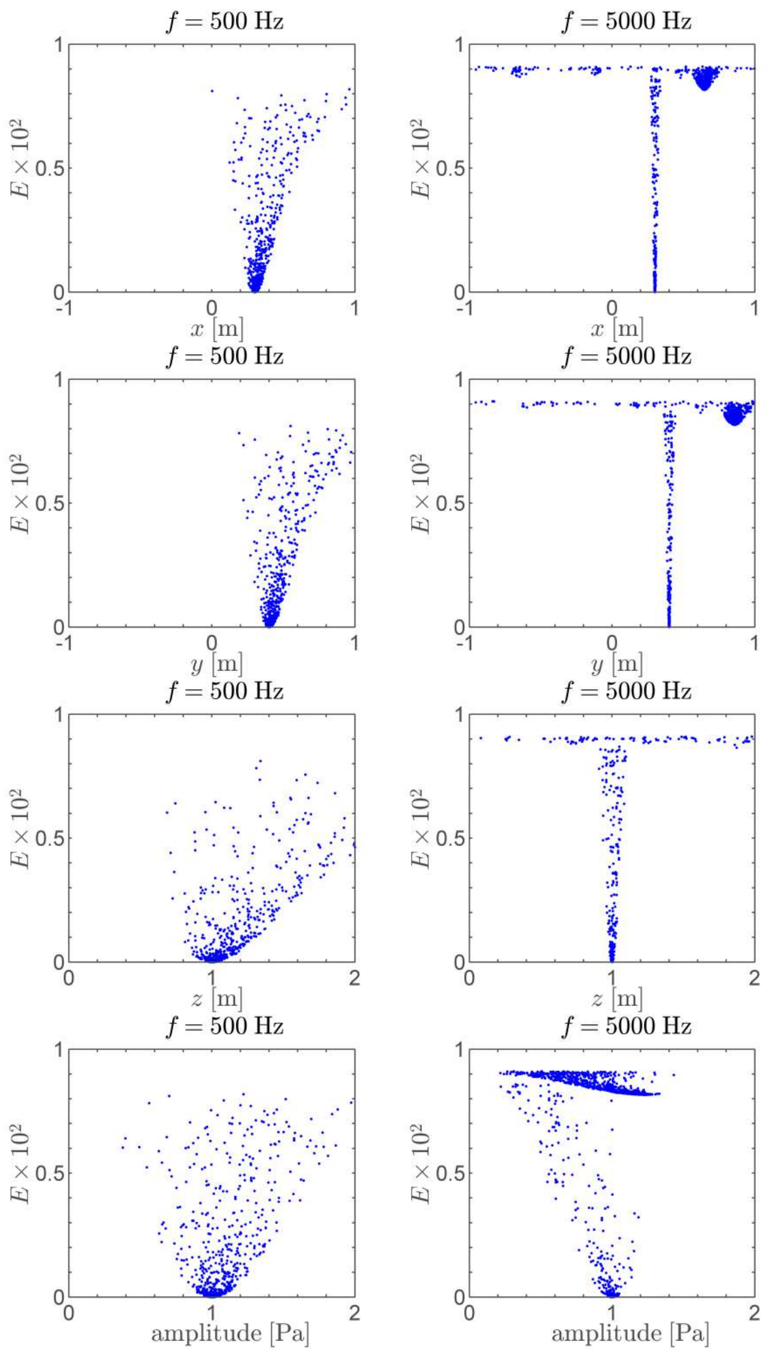

Figure 3: Inversion for 500 and 5000 Hz for spatial position and amplitude of the source.

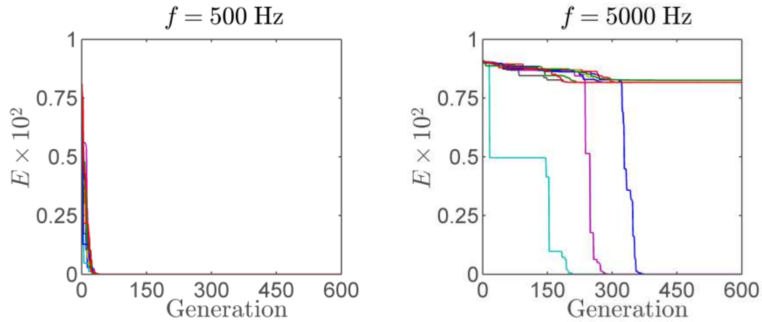

Figure 4: Energy as function of the generation for 500 and 5000 Hz.

5. Summary and conclusion

In this work an alternative method is presented to localize the acoustic sources using a microphone array. Use is made of the global optimization method Differential Evolution (DE). For this purpose an energy functions is formulated which uses the cross spectral matrix (CSM) of the measurement and signal model.

A simulated single monopole case was used to assess the performance of localization. The optimization method showed that it could easily identify the source for both 500 and 5000 Hz. Both spatial position and amplitude of the source corresponded to the lowest energy level.

This work shows that using global optimization methods can prove very useful for acoustic source localization. Source localization is easily extended in the third dimension. Additional advantage is the possibility to include even more unknowns, such as environmental parameters. The energy function can be adapted to the situation at hand. It can account for multiple sound sources, reflections or refractions of the sound, thus reflecting the actual measurement environment to a large extent.

REFERENCES

1 Van Veen, B. D. and Buckley, K. M. Beamforming: A Versatile Approach to Spatial Filtering, *IEEE ASSP Magazine*, **5** (2), 4-24, (1988).

2 Sijtsma, P. Experimental techniques for identification and characterisation of noise sources, *Netherlands Aerospace Center NLR*, NLR-TP-2004-165, (2014).

3 Snellen, M. and Simons, D. G. An Assessment of the Performance of Global Optimization Methods for GEO-Acoustic Inversion, *Journal of Computational Acoustics*, **16** (2), 199-223, (2008).

4 Storn, N. and Price, K. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, *Journal of global optimization*, **11** (4), 341–359, (1997).

5 Benchmarking Array Analysis Methods: http://www.b-tu.de/fg-akustik/lehre/aktuelles/arraybenchmark

6 Yardibi, T., Li, J., Stoica, P. and Cattafesta III, L.N. Sparsity constrained deconvolution approaches for acoustic source, *J. Acoust. Soc. Am.*, **123** (5), (2008).

7 Blacodon, D. and Élias, G., Level estimation of extended acoustic sources using an array of microphones, AIAA paper 2003-3199.