Article Identifier
DOI: 10.22032/dbt.59141
URN: urn:nbn:de:gbv:ilm1-2023isc-033:3

60th ILMENAU SCIENTIFIC COLLOQUIUM
Technische Universität Ilmenau, 4 – 8 September 2023
URN: urn:nbn:de:gbv:ilm1-2023isc:1

COMPARATIVE STUDY OF A BIOINSPIRED SOUND SOURCE LOCALIZATION ALGORITHM AND A STANDARD BEAMFORMER

Philipp Jünemann[1,2], Malte Mechtenberg[1,2], Axel Schneider[1,2], Joachim Waßmuth[1,2]

[1] Biomechatronics and Embedded Systems Group, Faculty of Engineering and Mathematics, University of Applied Sciences and Arts (HSBI), Bielefeld, Germany
[2] Institute of System Dynamics and Mechatronics, University of Applied Sciences and Arts, Bielefeld, Germany

ABSTRACT

This work presents a study that compares the angular accuracy of a binaural bioinspired method with classical delay-and-sum beamforming using linear acoustic arrays with an increasing number of sensors. In addition, this study also focuses on the frequency band dependent angular accuracy of the bioinspired method, as previous work has optimized it only for localizing narrowband sound sources. Therefore, a comparison of the localization accuracy for narrowband and broadband excitatory sounds is presented. This enables conclusions about the potential to reduce resources in terms of sensor array spatial requirements and the number of sensors if a bioinspired system is used.

Index Terms - Direction-of-arrival estimation, sound source localization, two-microphone array, multidimensional signal processing, beamforming, bioinspired coupling system, *Ormia ochracea*

1. INTRODUCTION

Animal species have evolved sensor modalities of different quality and range to perceive their environment. Near range sensing, e.g., touch based on specialized hair cells in insects [1], is at one end of this arbitrary range scale, far range sensing, e.g., active sound production in bats for echolocation [2], is at the other end. Within the scale are many different examples such as weakly electric fish that actively generate and sense electric fields to detect nearby objects [3]. This work focusses on the example of directional hearing / sound source localization based on bioinspired acoustic signal processing. The parasitoid fly *Ormia ochracea*, as the biological example, possesses a mechanically coupled auditory system that, despite limiting physiological properties, enables it to localize its cricket host based on interaural time differences of the sensed chirping sound as emitted by the cricket [4]. A model-based replication of the biological system has already shown that a change in mechanical parameterization leads to a change in directional hearing with respect to the most sensitive angle of incidence and frequency [5,6,7]. Compared to conventional beamforming methods, the bioinspired method [6] represents an optimization in terms of its angular accuracy when localizing mono-frequency sounds with two acoustic sensors. However, the bioinspired method must be extended for the localization of broadband sound sources. For this purpose, the optimization of the parameters of the coupled auditory system is adapted.

Miles et al. replicated the coupled auditory system of the fly *Ormia ochracea* with an idealized mechanical analogue that consists of an intertympanal bridge with two rigid bars [4]. These are connected by a coupling spring and a damper. At the extreme end of each bar there is also a spring and a damper. The replication includes the corresponding spring, damper and mass constants (k, k_c, c, c_c, m). From the idealized mechanical analog, a system of second-order differential equations can be derived, from which, using the Laplace transform, transfer functions can be obtained for a system with two inputs and two outputs. The block diagram of this system with its transfer functions $H_1(s)$ and $H_2(s)$ is illustrated in figure 1. The two outputs are labeled as $Y_1(s)$ and $Y_2(s)$ and the two inputs of the coupling system are connected to microphones m_1 and m_2, which are used to measure sound. The illustration shows the incidence of a spherical sound wave emitted by a sound source at a distance r_0 from the center point between the microphones. The microphones are at a distance d from each other. The radial distances r_1 and r_2 from the sound source to the microphones are depending on the angle of incidence Θ_{in} and can be determined with equations (1) and (2) [6]:

$$r_1(\Theta_{in}, r_0) = \sqrt{\left(\frac{d}{2}\right)^2 + r_0^2 - d\, r_0 \sin(\Theta_{in})} \tag{1}$$

$$r_2(\Theta_{in}, r_0) = \sqrt{\left(\frac{d}{2}\right)^2 + r_0^2 + d\, r_0\, sin(\Theta_{in})} \tag{2}$$

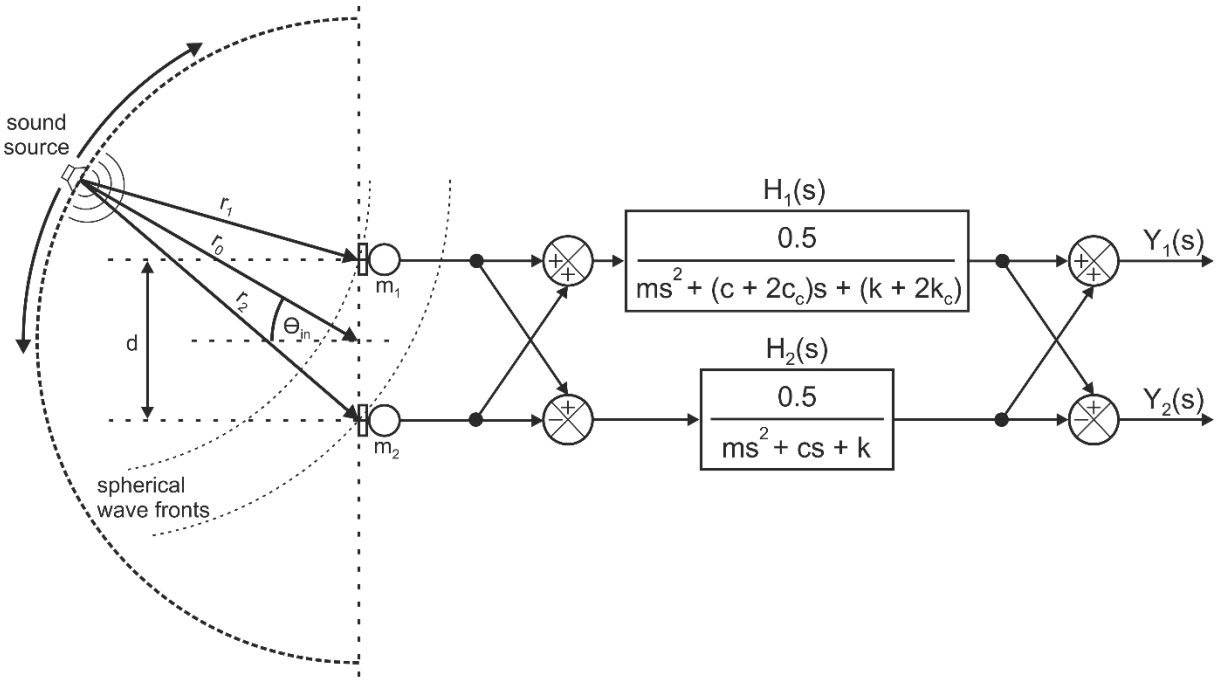

Figure 1. Spherical wave excitation on a linear array with two microphones m_1 and m_2 placed at a distance d. The sound source is located on a semicircle with radius r_0 and its center is placed in the middle between the microphones. The angle of incidence Θ_{in} is measured from the normal of the microphone plane at its center point. The radial distances r_1 and r_2 are measured from the sound source to the microphones. The microphones are connected directly to the inputs of the coupling system, which is shown as a block diagram on the right. Transfer functions $H_1(s)$ and $H_2(s)$ are represented in the Laplace domain. $Y_1(s)$ and $Y_2(s)$ are the outputs of the coupling system [6].

The auditory coupling system of *Ormia ochracea* can be characterized as a two-dimensional filtering system, which has a filtering effect over the spatial frequency and the temporal frequency. The angle of incidence Θ_{in} of the incoming sound wave is the spatial frequency and the propagating frequency f of the sound wave is the temporal frequency. In [6] it was shown that the coupling system can be parameterized for a specific most sensitive angle of incidence, which means that it responds with the largest amplitude difference between its two outputs $Y_1(s)$ and $Y_2(s)$ to a sound wave coming from this angle of incidence. The same applies to the most sensitive frequency, i.e., the temporal frequency to which the coupling system responds with the largest amplitude difference. The parameterization of the coupling constants (k, k_c, c, c_c, m) changes the sensitivity of the coupling system, i.e., the filter characteristic changes over the spatial and the temporal frequency. In the approach for parameter optimization presented in [6], the most sensitive frequency was chosen to be $f = 1000$ Hz and the most sensitive angle of incidence was varied from 0° to 90°. As a result, the localization of sound sources was only possible for mono-frequency signals and the two-dimensional filter effect was limited to the spatial frequency respectively the angle of incidence of the sound wave.

2. METHOD

For a comparison with a conventional delay-and-sum beamformer, broadband sound sources should also be localizable. Therefore, the parameter optimization approach from [6] was extended to consider a wider interval of temporal frequencies. To do this, the equation was considered that must be satisfied to parameterize the coupling system sensitively for a specific angle of incidence of a spherical sound wave as shown in figure 1:

$$\Psi(\omega, \Theta_{in}, r_0) \;=\; \chi(\omega, \vec{\varphi}) \text{ with } \vec{\varphi} = [k, k, k_c, c, c_c, m]^{\mathsf{T}} \tag{3}$$

$$\Leftrightarrow \frac{r_1^2 - 2r_1 r_2 \cos\big(\omega(\tau_1 - \tau_2)\big) + r_2^2}{r_1^2 + 2r_1 r_2 \cos\big(\omega(\tau_1 - \tau_2)\big) + r_2^2} = \frac{(c\omega)^2 + (k - m\omega^2)^2}{(k + 2k_c - m\omega^2)^2 + \big((c + 2c_c) \cdot \omega\big)^2} \tag{4}$$

Here $\omega = 2\pi f$ is the angular frequency depending on the temporal frequency f. $\tau_1 = \frac{r_1}{v_c}$ and $\tau_2 = \frac{r_2}{v_c}$ are the time delays associated with the radial distances r_1 and r_2 and with the propagation speed of the sound $v_c = 343\,\frac{\mathrm{m}}{\mathrm{s}}$.

In figure 2 the left side of the equation (4) is visualized as a surface plot for a fixed microphone distance of $d = 0.06$ m and a fixed radial distance of $r_0 = 1.0$ m. The temporal frequency f is varied from 0 Hz to 2000 Hz and the angle of incidence Θ_{in} from 0° to 90°. For small $\Theta_{in} \leq 50°$ the surface is flat over the whole interval of f and its maximum value is below 4. Since the right-hand side of equation (4) also depends on ω, a $\vec{\varphi}$ is sought for each angle of incidence Θ_{in} that satisfies equation (4) for all ω and for all f from 0 Hz to 2000 Hz, respectively.

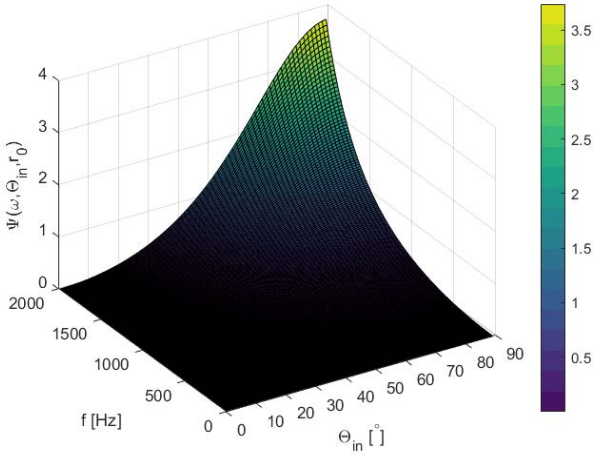

Figure 2. Surface plot of the left side of equation (2) with a fixed microphone distance $d = 0.06$ m, a fixed radial distance $r_0 = 1.0$ m. The temporal frequency range is varied from 0 Hz to 2000 Hz and the angle of incidence Θ_{in} is varied in a range from 0° to 90°.

In [6], a set of coupling parameters was optimized for each angle of incidence Θ_{in}, but this was designed only for a frequency of $f = 1000$ Hz. This was now used to determine $\chi(\omega, \vec{\varphi})$ for each Θ_{in} and over all f. In figure 3, the error function

$$err(\omega, \Theta_{in}) = \Psi(\omega, \Theta_{in}, r_0) - \chi(\omega, \vec{\varphi}) \tag{5}$$

associated to equation (3) is shown. For small angles of incidence, the error over the entire frequency interval is already small, whereas it increases for larger angles of incidence and higher frequencies. This shows that part of the parameter sets can already be used for different frequencies between 0 Hz and 2000 Hz. The remaining deviations can be further reduced by a renewed parameter optimization. For the doubled frequency interval from 2000 Hz to 4000 Hz, a similar surface is obtained when the microphone distance is halved to $d = 0.03$ m.

Based on this, for each possible angle of incidence $\Theta_{\mathrm{in}} \leq 50°$ in steps of 0.25° another optimization was performed with the corresponding parameter set $\vec{\varphi}$. For this, a gradient descent algorithm (lsqnonlin, Matlab R2022b (The Mathworks Inc., Natick, MA, USA)) was initialized using Trust-region-reflective as optimization algorithm. Then for each angle of incidence equation (5) was set as error function with a minimum step size of $1 \cdot 10^{-10}$ and a maximum of 400 iterations. For the optimization over $\omega = 2\pi \cdot [1\,\mathrm{Hz}, 2\,\mathrm{Hz}, \ldots, 2000\,\mathrm{Hz}]$ a microphone distance of $d = 0.06$ m was set. A lower boundary of $\mathrm{lb} = 0.5 \cdot \vec{\varphi}_{\mathrm{init}}$ and an upper boundary of $\mathrm{ub} = 2.0 \cdot \vec{\varphi}_{init}$ were set, with $\vec{\varphi}_{init}$ from [6] chosen as the initial parameter set for the optimization.

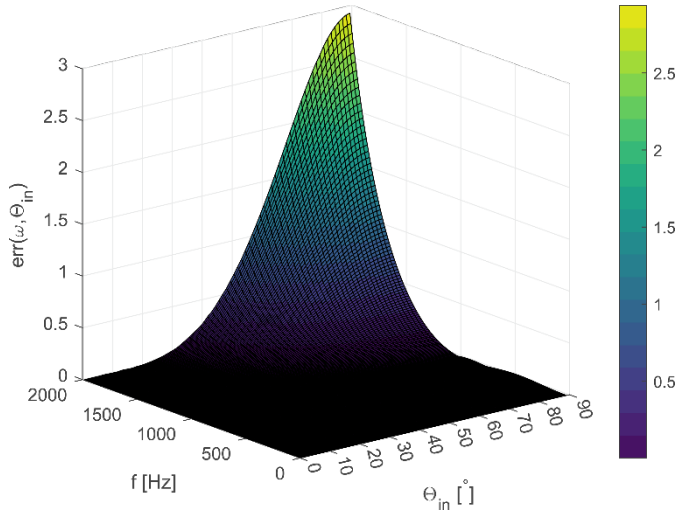

Figure 3. Surface plot of the error function $err(\omega, \Theta_{in})$ from equation (5). A fixed microphone distance $d = 0.06$ m and a fixed radial distance $r_0 = 1.0$ m is chosen. *The temporal frequency range is varied from 0 Hz to 2000 Hz and the angle of incidence Θ_{in} is varied in a range from 0° to 90°.*

A second optimization was done for another frequency interval $\omega = 2\pi \cdot [2000\ \text{Hz}, 2001\ \text{Hz}, \ldots, 4000\ \text{Hz}]$. Therefore, a microphone distance of $d = 0.03$ m was chosen as already described above. In contrast to the parameter optimization in [6], no nonlinear equality constraint was implemented here.

3. SIMULATION RESULTS

A simulation model of the coupling system that was implemented in Simulink (Matlab R2022b) was used as in [6]. In addition, a sound source was implemented that emits spherical sound waves so that they strike the center point of the microphones from a given angle of incidence Θ_{in}. Considering equations (1) and (2), the radial distances and the corresponding time delays were calculated to simulate the respective sound pressure signals at the two microphones m_1 and m_2. These were used as the two inputs of the coupling system. The sound waves were configurable with randomly selected frequencies that lie in a specific frequency band. In the experiments, sound waves with 10 and 20 additively superposed frequencies were simulated in chunks of 4096 samples at a sampling frequency of 48 kHz. To obtain a scalar value, the root mean square (RMS) of the output signals Y_1 and Y_2 of the coupling system was calculated and divided. The quotient of the RMS values expressed in decibels resulted in the amplitude difference. For the simulated sound pressure signals of a specific direction of incidence, the parameter sets $\vec{\varphi}$ were tested for all possible angles of incidence Θ_{in}. The ode14x fixed-step solver was used for the simulations.

3.1 Simulations with the bioinspired sound source localization method

As an example, two time-delayed signals are simulated for a microphone distance of $d = 0.06$ m and a radial distance of the sound source to the center point of the microphone array of $r_0 = 1.0$ m. The angles of incidence of the sound wave were chosen as $\Theta_{in} = \{10°, 20°, 30°, 40°\}$ to the center of the microphone array. The signals have 10 randomly selected frequencies between 200 Hz and 2000 Hz. Figure 4 shows the results of the sound source localization. The maximum amplitude difference indicates the calculated direction of incidence of the sound wave. For the smallest angle of incidence ($\Theta_{in} = 10°$) the amplitude difference is highest and decreases for larger angles of incidence.

In addition, a sound signal close to the application was tested in the simulation. This is the operating noise of a washing machine (example sound from Matlab). To create a second time-delayed signal with respect to the specified angle of incidence, the original signal was Fourier transformed, then the phase was shifted, and the signal was inverse Fourier transformed back into the time domain. Both signals were filtered with a bandpass filter. Then they were used as input signals for the coupling system. For the simulation, the procedure described above was carried out and the position of the sound source was localized. In figure 5(a), 4096 samples of each of the time-delayed signals are shown when the washing machine noise comes from an angle of incidence of $\Theta_{in} = 20°$. In figure 5(b) the localization result is visualized. The maximum amplitude difference of 20,88 dB for $\Theta_{in} = 19{,}5°$ shows only a deviation of 0.5° from the simulated angle of incidence.

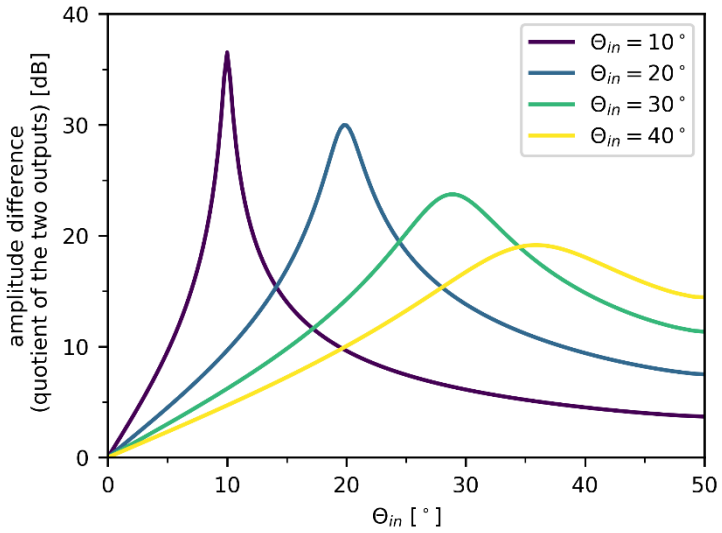

Figure 4. Acoustic point sources were simulated with angles of incidence of $\Theta_{in} = \{10°, 20°, 30°, 40°\}$. *The microphone distance was* $d = 0.06$ *m and the radial distance was* $r_0 = 1.0$ *m. The maximum amplitude difference indicates the angle of incidence calculated for the simulation. Recorded data (4096 samples for each microphone) was fed into the two inputs of the coupling system model, which was parameterized and simulated with the parameters* $\vec{\varphi}$ *optimized for all possible angles of incidence* $\Theta_{in} \leq 50°$, *as described in the section 3.*

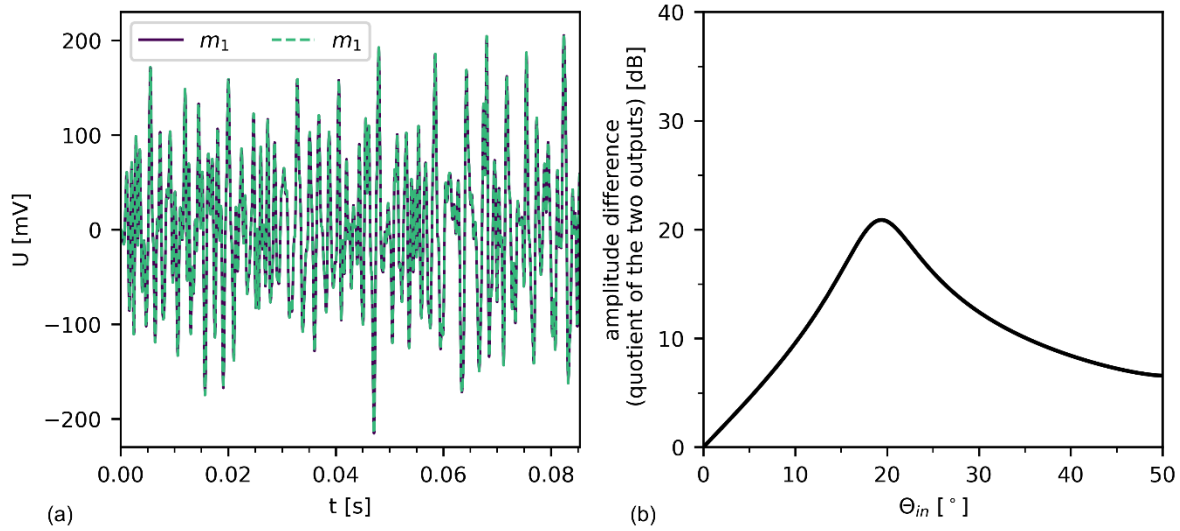

Figure 5. A point source emitting the sound of a running washing machine was simulated. The angle of incidence of the sound source was selected as $\Theta_{in} = 20°$. *The microphone distance was* $d = 0.06$ *m and the radial distance was* $r_0 = 1.0$ *m. The sound source localization method was performed as described in figure 4. In (a) the simulated measurement of the two microphone channels is shown. The sampling frequency was 48 kHz. In (b) the result of sound source localization is shown. The maximum amplitude difference indicates the angle of incidence calculated for the simulation.*

For the evaluation of the bioinspired method, 200 simulations each were performed for randomly selected angles of incidence of the sound wave from five intervals. The intervals were selected from 1° to 10°, 11° to 20°, 21° to 30°, 31° to 40° and 41° to 50°. This was done for signals with 10 and 20 randomly selected frequencies in the range from 200 Hz to 2000 Hz and again in the range from 2000 Hz to 4000 Hz. For the frequency range from 200 Hz to 2000 Hz, parameter sets were optimized using a microphone distance of $d = 0.06$ m (see section 3). Figure 6(a) shows box and whisker plots of the absolute error (AE) between the angle of incidence as simulated and the angle with highest amplitude difference when the excitation frequencies were in the range from 200 Hz to 2000 Hz. The first three intervals show a small AE for simulations with 10 and with 20 frequencies as well. In figure 6(b) the 3dB beam width of each simulation is shown.

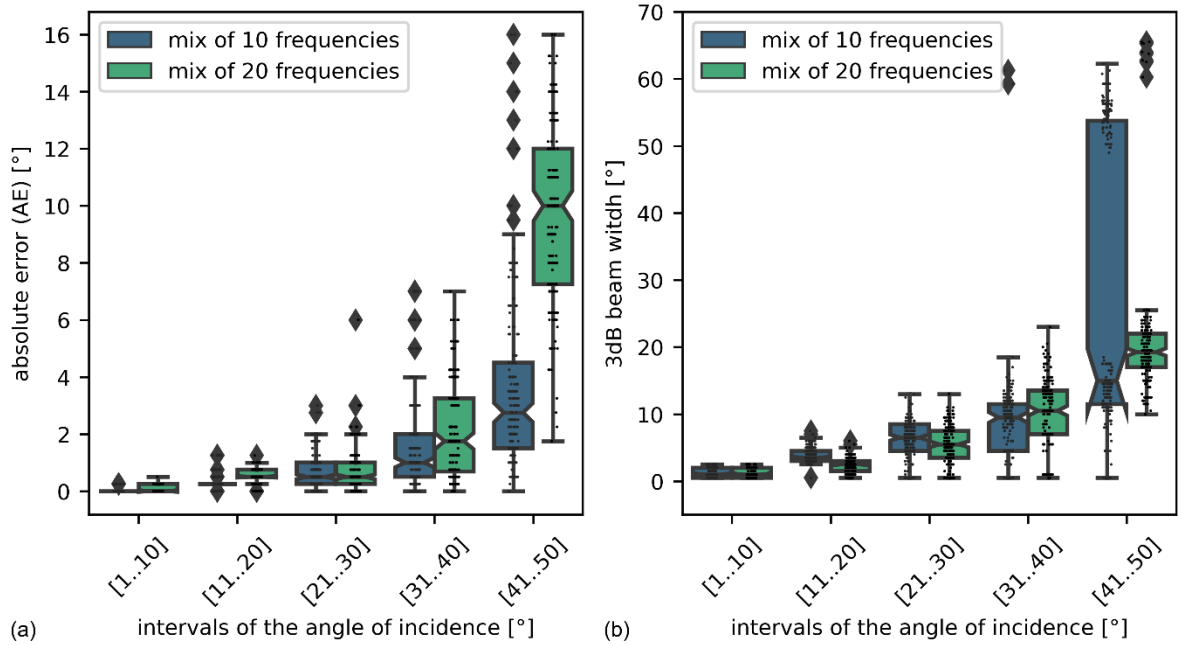

Figure 6. Acoustic point sources were simulated at a distance $r_0 = 1\ m$ from the center of the microphone array and with randomly chosen angles of incidence from the intervals 1° to 10°, 11° to 20°, 21° to 30°, 31° to 40°, and 41° to 50° 200 times each. The emitted sound waves were recorded over 4096 samples at the locations of the two microphones m_1 and m_2 at a distance $d = 0.06\ m$. In (a) the absolute error (AE) between the angle of incidence as simulated and the angle with highest amplitude difference when the excitation frequencies were in the range from 200 Hz to 2000 Hz are shown. In the first run, 10 frequencies were randomly selected from the described frequency interval for each simulation, and 20 frequencies were selected in the second run. In (b), the 3dB beam width was calculated for each simulation and mapped into the intervals of the angles of incidence.

For the second frequency range from 2000 Hz to 4000 Hz, further parameter sets were optimized using a microphone distance of $d = 0.03$ m (see section 3). Figure 7(a) shows the AE that is below 2° for the intervals 1° to 10°, 11° to 20° and 21° to 30°. For the intervals from 31° to 40° and 41° to 50°, the AE remains below that from figure 6(a). As previously described in [6], figure 7(b) shows that the 3dB beam width increases with growing angle of incidence.

3.2 Simulations with a delay-and-sum beamformer

For the comparison of the bioinspired sound source localization method and a classical delay-and-sum beamformer the theoretical description of an acoustic antenna for spherical sound waves from [8] was used and implemented in Matlab R2022b. As with the bioinspired method, the beamformer was set up with a microphone spacing of $d = 0.06$ m and a radial distance of $r_0 = 1.0$ m from the sound source to the center of the microphone array. The number of microphones n_{mic} was varied between 2, 10 and 20. Spherical sound waves were generated, and sound pressure signals were simulated at the microphones as described in section 3.1. In figure 8(a), the beampattern of the three beamformers is shown for an angle of incidence of $\Theta_{\text{in}} = 0°$ and for a frequency of $f = 1000$ Hz of the sound wave. For a beamformer with only two microphones, no 3dB beam width can be specified at all. For 10 microphones the 3dB beam width is 29.5° and for 20 microphones 16.5°. Even for 20 microphones, the delay-and-sum beamformer still has a larger 3dB beam width than the bioinspired method.

For comparison of the two methods, the same experiments as described in section 3.1 were performed for the delay-and-sum beamformer with 200 simulations each and for randomly selected angles of incidence from the five intervals from 1° to 10°, 11° to 20°, 21° to 30°, 31° to 40° and 41° to 50°. This was again performed for signals with 10 and 20 randomly selected frequencies in the range from 200 Hz to 2000 Hz. The box and whisker plots in figure 8(b) show that, unlike the bioinspired method, the delay-and-sum beamformer has a constant 3dB beam width over all intervals of the angles of incidence and regardless of the number of frequencies. For each interval, the median of the calculated 3dB beam width is between 15° and 18°. However, these are still higher than the medians of the bioinspired method, which uses only two microphones.

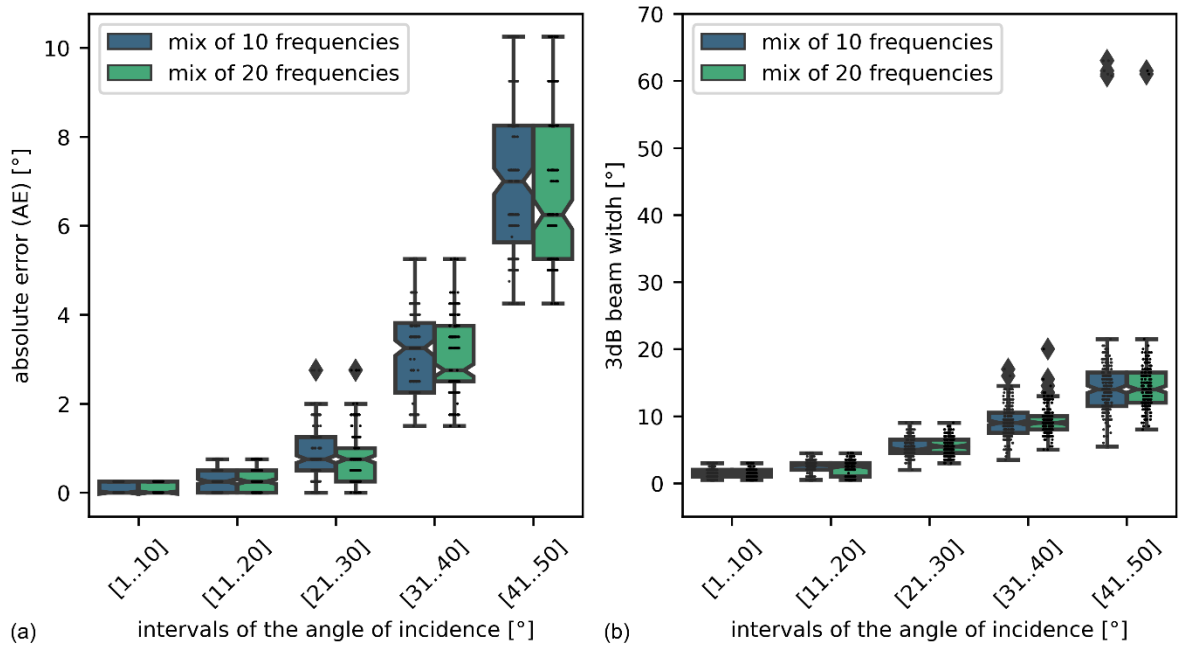

Figure 7. Acoustic point sources were simulated in the same way as in figure 6. The locations of the two microphones m_1 and m_2 were at a distance $d = 0.03$ m. In (a) the absolute error (AE) between the angle of incidence as simulated and the angle with highest amplitude difference when the excitation frequencies were in the range from 2000 Hz to 4000 Hz are shown. In the first run, 10 frequencies were randomly selected from the described frequency interval for each simulation, and 20 frequencies were selected in the second run. In (b), the 3dB beam width was calculated for each simulation and mapped into the intervals of the angles of incidence.

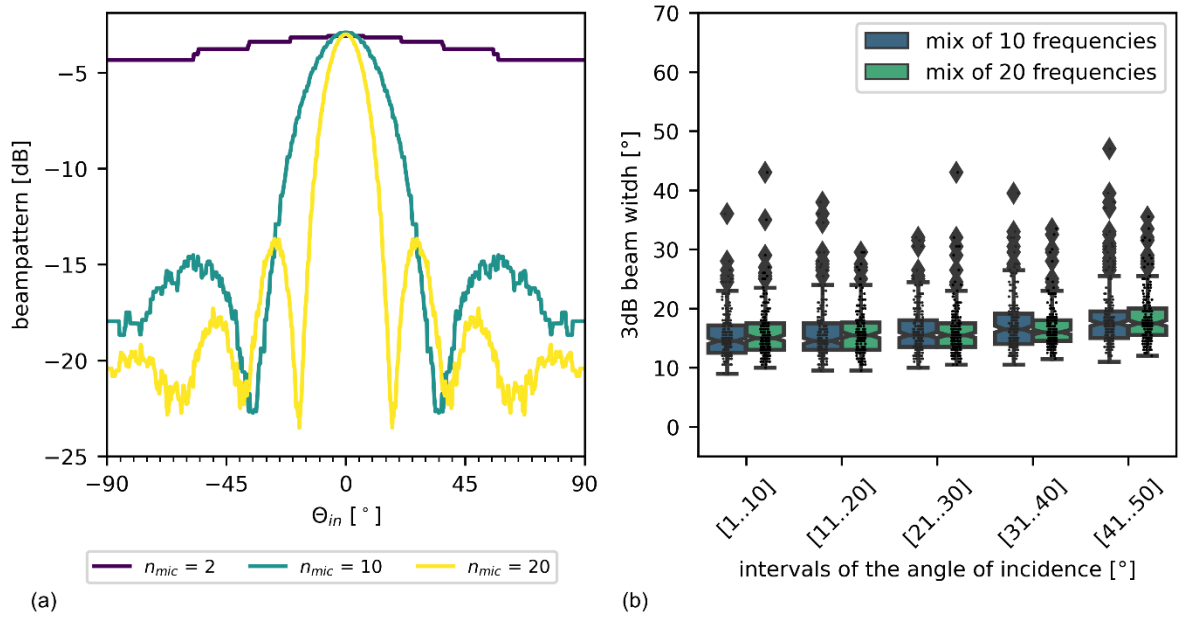

Figure 8. Delay-and-sum beamformers for spherical sound waves were simulated with 2, 10 and 20 microphones and a distance of $d = 0.06$ m between microphones [7]. (a) shows the beampattern for each of the three beamformers. The sound source was placed with an angle of incidence of $\Theta_{in} = 10°$ and a radial distance of $r\_0 = 1.0$ m. In (b), acoustic point sources were simulated at a distance $r_0 = 1$ m from the center of the linear microphone array with 20 microphones ($n_{mic} = 20$). The angles of incidence were randomly selected from the intervals 1° to 10°, 11° to 20°, 21° to 30°, 31° to 40°, and 41° to 50° 200 times each. In the first run, 10 frequencies were randomly selected from the frequency band between 200 Hz and 2000 Hz and additively superimposed for each simulation; in the second run, 20 frequencies were selected. The 3dB beam width was calculated for each simulation and mapped into the intervals of the angles of incidence.

4. DISCUSSION

In this work, a bioinspired method for localizing multifrequency sound sources is presented and evaluated. This is an extension of the method presented in [6], so that multifrequency signals in different frequency bands can be localized. Sound source localization with frequencies in the range from 200 Hz to 2000 Hz can be performed with two microphones at a distance of $d = 0.06$ m. For the frequency range from 2000 Hz to 4000 Hz a microphone distance of $d = 0.03$ m is used. The extension of the method for localization of multifrequency sound sources is a valuable addition for comparison with conventional delay-and-sum beamformers. These are characterized by a wide frequency range for localization, but optimally require a large number of microphones for this purpose. The method presented here uses only two microphones and still has a narrower 3dB beam width, as the simulation results have shown (see section 4). The accuracy of the sound source localization is also confirmed in the simulation results. For angles of incidence smaller than 30° the absolute error (AE) between the angle of incidence as simulated and the angle with highest amplitude difference is in the range of plus/minus 2°. This also applies to a different number of additively overlapping frequencies in the excitation sound signal. The method also determines the correct angle of incidence for realistic noise sources, shown here for the simulated noise of a washing machine (see section 4). Moreover, by finer parameter optimization as described in section 3, the AE can be further reduced. In contrast, the implemented delay-and-sum beamformer showed that at least 20 linearly arranged microphones at a distance of $d = 0.06$ m were required to achieve a result with a similarly narrow 3dB beam width. However, the 3db beam width remained constant over all intervals of the angles of incidence.

For the evaluation of the bioinspired method, the front-back ambiguity must be considered. In contrast, a positive or negative angle of incidence can be distinguished based on a positive or negative amplitude difference. The separability and localization of multiple sound sources is also not possible by the presented extension. In future work, the method must be extended for the localization of multiple sound sources in 2D and 3D space. For the 3D space the use of a 2D microphone array and multiple pairs of microphones is necessary. The use of several pairs of microphones at different distances should make it possible to adjust different frequency bands so that wide-band speech sources can also be localized.

ACKNOWLEDGEMENTS

This work has been supported by the Bundesministerium für Bildung und Forschung (BMBF, Federal Ministry of Education and Research) - Reference No. 01IS18009B and by a Career@Bi grant of the University of Applied Sciences and Arts, Bielefeld, Germany.

REFERENCES

[1] S. Zill, J. Schmitz, and A. Büschges, Load Sensing and Control of Posture and locomotion. Arthropod Struct. Dev., vol. 33, no. 3, pp. 273–86, 2004, DOI: https://doi.org/10.1016/j.asd.2004.05.005
[2] G. Jones, Echolocation. Current Biology, vol. 15, no. 13, pp. R484-8, 2005, DOI: https://doi.org/10.1016/j.cub.2005.06.051
[3] A.B. Sichert, R. Bamler, and J.L. van Hemmen, Hydrodynamic Object Recognition: When Multipoles Count. Phys. Rev. Lett., vol. 102, 2009, DOI: https://doi.org/10.1103/PhysRevLett.102.058104
[4] R.N. Miles, D. Robert, and R.R. Hoy, Mechanically coupled ears for directional hearing in the parasitoid fly *Ormia ochracea*. The Journal of the Acoustical Society of America, vol. 98, no. 6, pp. 3059-70, 1995, DOI: https://doi.org/10.1121/1.413830
[5] Y. Hou, J. Su, K. Xia, and X. Ji, Analysis on performance of biomimetic coupled processing system. 2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN), 2017, DOI: https://doi.org/10.1109/ICCSN.2017.8230233
[6] P. Jünemann, A. Schneider, and J. Waßmuth, Direction-of-arrival estimation for acoustic signals based on direction-dependent parameter tuning of a bioinspired binaural coupling system, Bioinspiration & Biomimetics, IOP Publishing, 2023
[7] M. Akcakaya, and A. Nehorai, Biologically inspired coupled antenna beampattern design, Bioinspiration & Biomimetics, IOP Publishing, 2010
[8] M. Möser, Messtechnik der Akustik, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

CONTACTS

P. Jünemann, M.Sc.
email: philipp.juenemann@hsbi.de
ORCID: https://orcid.org/0000-0002-5890-928X

M. Mechtenberg, M.Sc.
email: malte.mechtenberg@hsbi.de
ORCID: https://orcid.org/0000-0002-8958-0931

Prof. Dr. rer. nat. A. Schneider
email: axel.schneider@hsbi.de
ORCID: https://orcid.org/0000-0002-6632-3473

Prof. Dr.-Ing. J. Waßmuth
email: joachim.wassmuth@hsbi.de
ORCID: https://orcid.org/0009-0006-1863-1275