Applied Acoustics 231 (2025) 110560

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Bearing estimation using foghorn sounds

Ajinkya Gorad [a,*], Zheng Zhao [b], José M. Vallet García [c], Ville Lehtola [c,d], Toni Hammarberg [c], Henrik Ramm-Schmidt [e], Saiful Islam [c], Sarang Thombre [c], Simo Särkkä [a]

[a] Aalto University, Otakaari 3, Espoo, 02150, Finland
[b] Linköping University, Linköping, 58183, Sweden
[c] Finnish Geospatial Research Institute, Vuorimiehentie 5, Espoo, 02150, Finland
[d] University of Twente, 7500 AE, Enschede, PO Box 217, Netherlands
[e] Fleetrange Oy, Ylätie 6 B, Espoo, 02300, Finland

ARTICLE INFO

Keywords:
Autonomous ship
Sound based localization
Maritime awareness
Bearing estimation

ABSTRACT

Situational awareness is critical for safe navigation and autonomous ship operations. To improve safety and automate navigation, multi-sensor systems could involve sound-based direction-finding of foghorns. In this paper, we study the performance of a three-microphone array in bearing estimation when a distant foghorn is audible. We adopt an approach which utilizes Hilbert transforms to obtain phase delay from filtered sounds, and get bearing estimate by minimization of an energy function. We compare this approach with a cross correlation method and GPS-based ground-truth bearing. We present results from simulated tests, laboratory tests with synthetic foghorn sounds, and field tests done on a cruise vessel in the Gulf of Finland. The results suggest that vessel bearing estimation is possible through foghorn sounds received by the microphone array. Best accuracy of 7° is achieved in bearing estimation using Hilbert transform approach with target-to-receiver distance of 3.2 km, and processing time of 19 ms for 1 second of sound information. We provide recommendations on how the sound could play an important role in multi-sensor systems. It can be concluded that sound sensory information is useful in autonomous navigation.

1. Introduction

Autonomous navigation is essential for future of transportation. When navigating on a water body, awareness of the surroundings is a crucial factor for implementing safe navigation techniques. Typical methods for that purpose make use of sensing devices such as visible light and infrared cameras, lidar, radar, sonar, and Global Navigation Satellite System (GNSS) receivers [33,51]. Visible light cameras can aid in maritime awareness of small objects such as buoys and nearby ships [19]. However, such cameras cannot work in the dark. The resolution of infrared cameras [28] is not yet sufficient for maritime use. Although Automatic Identification System (AIS) and radars provide vessel traffic information required for navigation, the information provided by them can be unreliable: GNSS receiver may be jammed [31] or small targets in radar view can be undetected or be unrecognizable, which can cause a significant danger of collision. Therefore, it is important to augment these systems with alternative, independent measurements systems. The goal of this work is to detect foghorn sounds from another vessel and determine its bearing relative to the measuring-vessel. Our aim is to evaluate the feasibility of using sound-based information for this purpose and, in particular, using sounds originating from the foghorn of a remote vessel.

Multiple sensors can be fused to provide detailed information for navigation: a brief review on sensors and information fusion using Artificial Intelligence (AI) methods is presented in Thombre et al. [45]. Future maritime transportation could fuse sound-based measurements with available sensors such as AIS, radar, and imaging methods in order to utilize the sound sensory information for localization. Although wide literature exists on underwater sound-based localization [42] and the technology for microphone array based localization is well known [8,9,22], sound-based techniques are not well studied for applications on the sea. For this purpose, our aim is to study the possibility of equipping maritime vessels with microphones arrays in order to estimate the relative direction or bearing of another vessel from foghorn sounds.

Fog signals are a protocol invented to provide warnings for ships in scenarios where optical signal from a lighthouse is occluded by a thick

* Corresponding author.
E-mail address: ajinkya.gorad@aalto.fi (A. Gorad).

https://doi.org/10.1016/j.apacoust.2025.110560
Received 31 October 2023; Received in revised form 13 January 2025; Accepted 24 January 2025

Available online 27 January 2025

fog [38,1]. Such signals were first produced with cannons or fog guns installed on ship, and were developed in Europe during the 18th century, but were then replaced by foghorns after the arrival of steam engines. After the invention of radar and AIS, the role of the foghorn was reduced to a backup safety device. The sophisticated multi-radar systems that modern ships are equipped with, can detect other vessels even in fog. AIS signals that communicate vessel information like positions are not largely affected by the weather. However, radar systems have blind spots at a close range and sometimes may fail to detect small vessels at normal operating range. Risks become even higher when a small vessel lacks an operational AIS transceiver. A foghorn may help in these situations. Therefore, by existing regulations, ships need to have a "means of making sound signals", which is an requirement according to Safety of Life at Sea (SOLAS) regulations, and sound signals must be utilized in "restricted visibility" (COLREG rule 35) [16]. Foghorn can thus be seen as a backup safety device.

We argue that the future advanced sensor fusion methods will not diminish the role of the foghorn but, instead, make it even more useful as a part of a multi-sensor system. First, having a foghorn integrated to an autonomous system makes the new system compatible with the established code of conduct at sea, as not all vessels have state-of-the-art sensor systems. Second, having an autonomous sound reception system for foghorn signals enables the system to digitally integrate these signals with ones from other sensors enhancing the situational awareness in terms of safety, which is the primary driving factor within the e-Navigation strategy of International Maritime Organization (IMO MSC 85/26/Add.1 Annex 20) [21]. However, this integration, although based on new methods, faces well-known challenges. Hubbard [20] writes that there is a great variance in the audibility of the sound from a foghorn. In some scenarios an aerial echo can be heard and in some the sound can be lost in so-called *silent zones*, as observed by a captain: "*In moving directly away from a fog signal station, the sound is audible for a distance of a mile, is then lost for about the same distance after which it is again distinctly heard for a long time*" [20]. Obviously, these phenomena constitute a real danger for a mariner developing a feeling of confidence that the foghorn signal alone could be used to guarantee safety. In the light of modern technology, aerial echoes and silent zones represent weaknesses of a single sensor which may be compensated by the use of multi-sensor systems and sensor fusion techniques.

In this article, we describe an algorithm to estimate the bearing of a target vessel from a received foghorn sound. The algorithm is validated using computer simulations, lab tests, and field tests. Simulations are used to find the performance of the algorithm incorporating various foghorn sound intensities and sampled background noise from the sea. Laboratory tests are conducted to estimate the short-range performance of artificially synthesized foghorn sounds generated through speakers. We validate the algorithm with field tests conducted on-board the Ro-pax (Roll-On-Roll-Off-Passenger-ship) ferry MS Megastar sailing between Helsinki and Tallinn in the Gulf of Finland. The on-board system includes an array of three microphones, related recording equipment, and suggested processing computational methods.

The article is organized as follows. In Section 2, we review the previous work and general background of methods used in localizing acoustic targets. The proposed computational methods for sound detection and localization using a three-microphone array are presented in Section 3. Related equipment, experiments and results in simulations, laboratory tests, and field tests are presented in Section 4. Future work related discussion based on results is provided in Section 5. Limitations and challenges are discussed in Section 6, and finally the article is concluded in Section 7.

2. Background

Using microphones for foghorn direction estimation is almost a-century-old problem [47]. Most ships are already required to have foghorn sound reception systems on-board, for example, SOLAS [40] states that vessels of over 500 GT (Gross Tonnage) and with a completely enclosed bridge must be equipped with a sound reception system. These commercial systems offer a directional hint on the sound origin, typically restricted to 8 directions: front, starboard-front, starboard, starboard-aft, aft, port-aft, port, and port-front.

Localization using sound is challenging in windy conditions and is even more challenging in the sea-environment conditions [20,50,25,43]. Attenuation of foghorn sounds takes place due to weather conditions [50], which degrades the signal quality and the phase information. Under normal conditions the rough sea surface caused by wind-driven waves induces random scattering of the reflected sound [5,46,7]. However, low frequency sounds from, for example, wind farms can still propagate far over ten kilometers depending on the temperature of the water and the air, the wind, and the presence of atmospheric turbulence [4]. This is because at short ranges, up to a few kilometers, the sound attenuation experiences spherical $\sim 1/r^2$ loss, but at long ranges, it becomes cylindrical, $\sim 1/r$ loss [4]. In our work, the field experiments are done within the short-range regime, assuming that the multi-path effects from the sea surface and atmospheric effects can largely be neglected as in Bolin et al. [5]. However, wind induced noise cannot be neglected in the recorded data.

Audio signals could play an active role in navigation. A natural demonstration of active sound-based navigation is the echo-localization used by bats. Such active strategies have been implemented in robots using two ear-like microphones with an acoustic transmitter and a bat-motivated algorithm called BatSLAM [41,12]. In the sea, *odontocetes* or toothed whales have similar bio-sonars, but they use them underwater [e.g., 2]. Echo-localization has a transmitter which transmits a specific signal repeatedly, and the echo signal is detected using microphones or hydrophones [10]. Echo-localization method can obtain the distance to target, which cannot be easily obtained from sound emanating objects using a passive approach.

Foghorn-based sound direction estimation can use a planar microphone array and perform either a time delay, phase delay, or amplitude based estimation [13]. Wave properties of sound, like propagation phase delay, can be used to estimate the direction of incidence using multiple microphones. Multiple studies have been conducted for underwater localization of submarines using sound signals [see, e.g., 3]. However, few studies exist for sound-based localization of maritime objects above the sea surface [11,23,6]. Work by Choi and Kim [11] provides a numerical simulation on the use a symmetric deployment with 4 equidistant microphones to find the direction of the transmitting ship using sound propagation model in 3D, which uses boundary element method. A more recent work utilizing wave property of sound using a Luneberg's metalens made of acrylic pipes and using 8 microphones is studied in Kim et al. [23]. Practical experiments with ship foghorn sound utilizing a 24-microphone array were reported in Brix [6]. They studied the methods of time-difference-of-arrival (TDoA) estimated using generalized cross correlation (GCC) and EB-ESPRIT involving cylinder harmonics which were introduced in [44]. Typical reported results in such experiments use conventional sound sources in indoor spaces, and include steered beamforming [52,9,48], TDoA [36].

Beamforming methods have also been tested for foghorn sound direction estimation using quad-microphone array in an anechoic chamber in Vilhjálmsson and Engmarksgaard [49]. Valin et al. [48] used steered beamforming with particle filter to simultaneously localize and track the azimuth of the sound source. Several beamforming based acoustic imaging methods are discussed in Merino-Martínez et al. [32]. Work by DiBiase [14] describes speaker localization methods like Generalized Cross Correlation Phase Transform (GCC-PHAT) and Steered Power Response Phase transform (SRP-PHAT). GCC-PHAT$-\rho\gamma$ with guided spectral-temporal (ST) position method is used in Li et al. [26]. Hioka and Hamada [18] proposes an algorithm which uses an equilateral arrangement of omni-directional microphones to track a speaker. Grondin and Michaud [17] proposed SRP-PHAT-HSDA which utilized TDoA and Kalman filter to track the direction in 3D space. However, such algo-

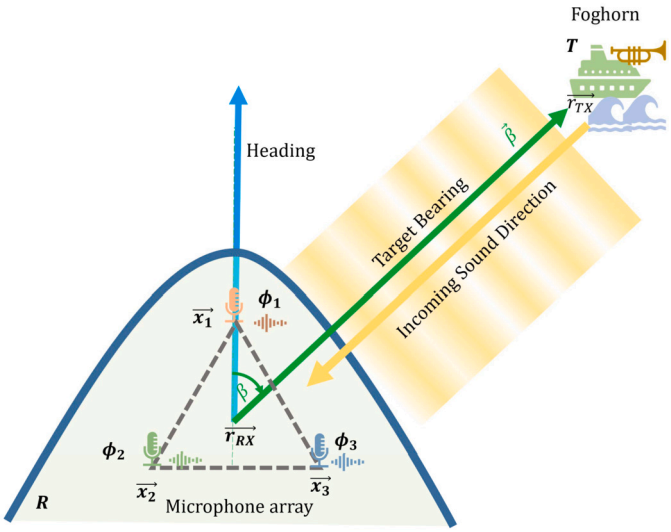


Fig. 1. Ferry foghorn sound bearing estimation setup. A target vessel at $\vec{r}_{TX}$ makes foghorn sound which is received by the microphone array on the receiving vessel at $\vec{r}_{RX}$. In the figure, ϕ_1, ϕ_2 and ϕ_3 are the incident phases of the dominant foghorn frequency harmonic on the corresponding microphones. The use of a sound direction estimation algorithm determines the bearing of the emanating sound source.

rithms are adapted to indoor conditions, and an outdoor field sound localization is even more challenging.

In this work, we restrict our goal of positioning to finding only the angular localization of target vessel, that is, to target bearing estimation using a single, passive microphone array. Methods such as bearing-only tracking [see, e.g., 27,39] which utilize multiple microphone arrays placed at sufficient distances can provide a full localization result. In addition, active echo-localization based approaches could also provide distance over the air. However, using multiple arrays or an active sensor are not considered in this work.

3. Methodology

The microphone assembly and the algorithm for the direction estimation of foghorn sounds is discussed in this section. The geometry of the direction (relative bearing) estimation problem is shown in Fig. 1 and it can be mathematically described as follows. Assume a maritime scene of sea surface with a transmitting vessel TX at $\vec{r}_{\mathrm{TX}}$ and a receiving vessel RX at $\vec{r}_{\mathrm{RX}}$. The target ship direction unit vector is $\vec{\beta}$ which is at angle β measured clockwise from the receiving ship's forward direction (Fig. 1). A sound wave travels from TX to RX at speed of c in time Δt. Therefore, the position of the transmitter vessel is formulated in ideal scenario as:

$$\vec{r}_{\mathrm{TX}} = \vec{\beta}\, c\, \Delta t + \vec{r}_{\mathrm{RX}}. \tag{1}$$

We use microphone array consisting of three microphones. Let $\vec{x}_1$, $\vec{x}_2$, and $\vec{x}_3$ in $\mathcal{R}^2$ be the microphone positions in an equilateral arrangement relative receiving vessel's position $\vec{r}_{\mathrm{RX}}$ in the front, on the left, and on the right side respectively. Side length of this triangle is D. Three microphones are positioned on these vertices installed in a plane parallel to the sea surface. These microphones record incident sound pressure signals (in Pascals) $u_i(t)$ for microphones $i = 1, 2, 3$.

Goal of the algorithm is to estimate the direction of the target vessel TX from its foghorn sound with respect to the receiving vessel RX by extracting the phase information in received sound. Fig. 2 presents the flow diagram of the algorithm. For each microphone received signal, a rectangular window of time-length T_{win} is applied on the signal. Following that, a second order infinite impulse response (IIR) resonator filter with suitable frequency band for the dominant harmonic is applied. Thereafter, instead of an conventional correlation based approach to calculate delays [11,24], we use Hilbert transform [34,29] on the filtered sound pressure signals which allows straightforward instantaneous frequency calculations [29], as well as envelope calculations [11]. After computing the phase and frequency, an optimization approach is then used on the computed phase information with the microphone array geometry to obtain the sound direction. Resulting direction is then converted to a real-world bearing value in degrees with respect to the heading of the receiving vessel.

The algorithm assumes that the ship foghorn sounds originate in the same plane as the microphone array. The effect of sea-surface reflections was not as dominant as the effect of reflections from the metallic structure of ship, which were compensated using a simple sound absorber behind the microphones for certain foghorn events. However, our model does not account for the multipath from the sea. During our experiments, ferry engine, wind, and other sounds apart from foghorn represented the background noise for our model. There are no long-distance echoes on the open sea. The foghorn sounds are generated while target vessel is in motion and therefore, a Doppler effect is observed in the received signal. Our strategy is based on tracking the small changes in sound source frequency in a narrow frequency window evident from Hilbert transform operation while phase differences are independently observed. It is also valid for locating non-moving point targets. The ego-motion of the observer is captured with on-board GNSS receiver.

The fundamental ship foghorn frequency f_0 as well as the dominant harmonic n are found from the spectrum of recorded foghorn sound pressure signals $u_i, i = 1, 2, 3$. Foghorn sound frequency-window range for the narrow band resonator is taken such that the dominant harmonic is included in the frequency-window. Window is around the n-th harmonic with bandwidth b_{W}. Window is also chosen such that its corresponding wavelength is below $D/2$, where D is the inter-microphone distance, depending on the setup. A rectangular time window of T_{win} is applied on the received microphone channel signals. In this time window, signal u_i is filtered using H_f, a second order resonator [`iirpeak` in Matlab, 35] to produce v_i with filter center frequency $f \approx nf_0$ and filter bandwidth b_{W}. It removes the outband noise and focuses only on the frequency components present in that band, mainly a single dominant harmonic from the foghorn sound that can be used to estimate the direction. A wider bandwidth b_{W} allows for a larger range of unknown Doppler shifted components to be filtered. In our algorithm, the choice of the harmonic is crucial depending upon the microphone configuration and its intensity received by the microphones. The selected frequency for phase detection must satisfy the frequency criteria to such that the wavelength at the dominant peak is at least twice the microphone array distances in equilateral configuration to avoid phase difference ambiguities in the received signals [15], that is:

$$f < \frac{c}{2D}. \tag{2}$$

Let us now take a closer look at the algorithm outlined in Fig. 2. First, sound data is passed through a resonator filter (IIR filter) to extract the foghorn frequency component. After the filter, as in Choi and Kim [11], the Hilbert transform $\mathcal{H}$ of the windowed signal is computed for each microphone signal $u_i (i = 1, 2, 3)$ with which we then form the analytic signal:

$$H_i(t) = v_i(t) + j\mathcal{H}v_i(t). \tag{3}$$

From this analytic signal $H_i(t)$, the total filtered signal energy ϵ_k in the filtered frequency band received by the microphone array is computed as:

$$\epsilon_k = \sum_{i=1}^{3} \int_{t=(k-1)T_{\mathrm{win}}}^{t=kT_{\mathrm{win}}} H_i(t) H_i^*(t) dt, \tag{4}$$

where k is the index for the time-window. From that, we can compute the instantaneous power P_k of the signal received by microphone array:

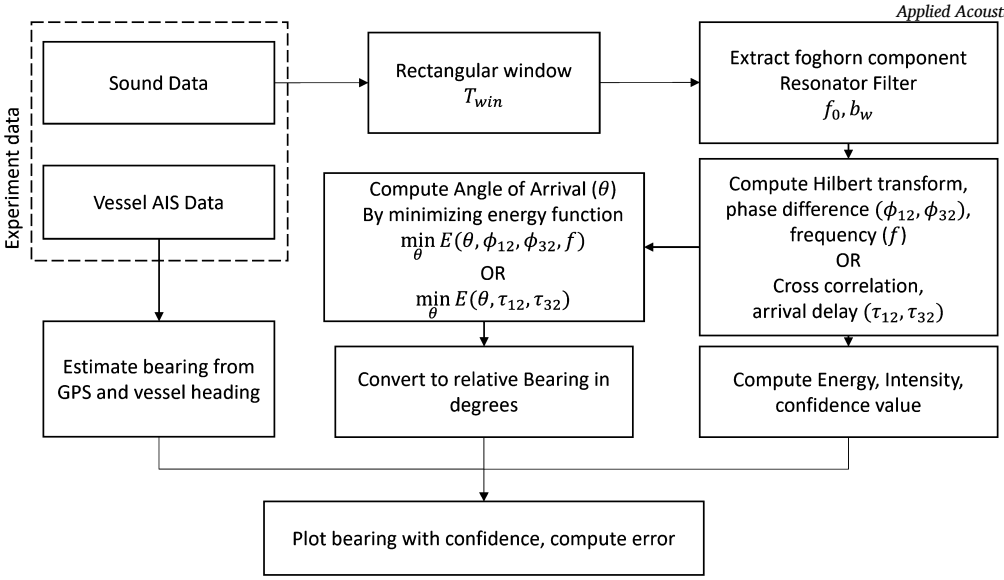


Fig. 2. Sound processing method. Microphone array sound input is fed to the algorithm. Bearing and confidence value are the outcome of the algorithm, and is compared with the true heading from GPS. Error is calculated from AIS calculated bearing of the target ship and sound direction estimated bearing. Results are plotted and error is computed.

$$P_k = \frac{\varepsilon_k}{T_{\text{win}}}, \tag{5}$$

and convert it to decibels:

$$P_{k,dB} = 10 \log_{10} P_k. \tag{6}$$

We then use the following smoothing filter on the power values (in decibels):

$$I_k = \alpha I_{k-1} + (1-\alpha) P_{k,dB}, \tag{7}$$

where

$$\alpha = 1 - \frac{T_{\text{win}}}{\tau_s}. \tag{8}$$

Using a smooth log-intensity signal provides a clean foghorn event detection window for signals with low SNR (signal-to-noise ratio). From this processed signal we devised a heuristic confidence measure ρ_k is obtained using a threshold I_θ of 33 dB using a piecewise linear function:

$$\rho_k = \begin{cases} 0, & I_k \le I_\theta, \\ 0.1\,(I_k - I_\theta), & I_\theta \le I_k \le 10 + I_\theta. \\ 1, & I_k \ge 10 + I_\theta. \end{cases} \tag{9}$$

The rationale behind using this as confidence window is that the stronger components will have better phase detection than the weak ones, this way it heuristically allows to gauge the quality of the bearing obtained. In our work, envelope-like quantity in Equation (7) is used in Equation (9) to get the confidence of the foghorn bearing estimation. However, in Choi and Kim [11], delay estimates are computed from the correlations on the envelope signals, which can be computed as the magnitude of the analytic signal in Equation (3).

In our algorithm, phase differences $\phi_{12} = \phi_1 - \phi_2$ and $\phi_{32} = \phi_3 - \phi_2$ are obtained from the phase information available in the analytic signals $H_1(t)$, $H_2(t)$, and $H_3(t)$, and their complex conjugates as:

$$\phi_{ij}(t) = \angle(H_i(t)\ H_j^*(t)). \tag{10}$$

ϕ_{12}, ϕ_{32} corresponds to the phases of the front and the right microphones relative to the left microphone, respectively. In addition, instantaneous phase ϕ_i for each microphone is calculated:

$$\phi_i(t) = \angle(H_i(t)), \tag{11}$$

to obtain the average instantaneous frequency $f(t)$:

$$f(t) = \frac{1}{3} \sum_{i=1}^{3} \frac{f_s}{2\pi} \frac{\partial \phi_i(t)}{\partial t}. \tag{12}$$

After that, in each window, quantities $\phi_{12}(t)$, $\phi_{32}(t)$, and $f(t)$ are averaged to provide $\phi_{12,k}$, $\phi_{32,k}$ and f_k for the k-th window.

We also compare Hilbert phase estimation approach for direction estimation to delay estimation using cross correlation method [37]. The delay between two signals $v_i(t)$ and $v_j(t)$ in continuous time can be estimated using the cross-correlation function $R(\tau)$, defined as:

$$R_{ij}^k(\tau) = \int_{t=(k-1)T_{\text{win}}}^{t=kT_{\text{win}}} v_i(t) v_j(t+\tau)\, dt \tag{13}$$

The time delay $\tau_{ij,k}$ is estimated by finding the value of τ that maximizes the cross-correlation:

$$\tau_{ij,k} = \arg\max_\tau R_{ij}^k(\tau). \tag{14}$$

From Hilbert based approach, the estimated phase differences between the microphones and instantaneous frequency incident on the microphone array, a relative bearing estimate $\hat{\beta}_k$ of target ship relative to microphone configuration (that is aligned to the heading of the receiving ship) is calculated using a grid-based minimization of:

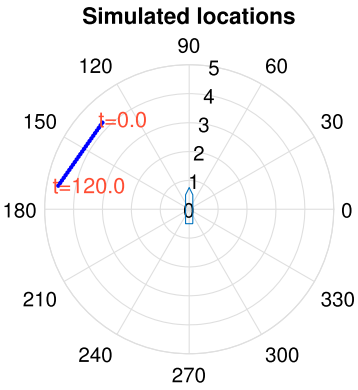
$$\hat{\beta}_k = \min_\beta \left(\left(\frac{\phi_{32,k}}{2\pi f_k} - \frac{D \sin\beta}{c} \right)^2 + \left(\frac{\phi_{12,k}}{2\pi f_k} - \frac{D \cos(\pi/6 - \beta)}{c} \right)^2 \right), \tag{15}$$

and for cross correlation (Xcorr) approach:

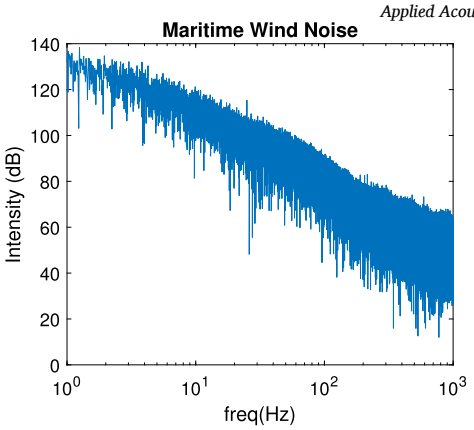
$$\hat{\beta}_k = \min_\beta \left(\left(-\tau_{32,k} - \frac{D \sin\beta}{c} \right)^2 + \left(-\tau_{12,k} - \frac{D \cos(\pi/6 - \beta)}{c} \right)^2 \right), \tag{16}$$

where the minimization is conducted over Angle of View (AoV) of 2π in simulated and laboratory tests, and is π in the field tests. Results of the bearing measurement are obtained as the tuple $(\hat{\beta}_k, \rho_k)$, whose samples arrive at a rate of 10 Hz in the algorithm. Value ρ_k represents the confidence of the occurrence of the foghorn sound, and $\hat{\beta}_k$ provides the estimate for the bearing of that foghorn sound.

In field tests, additional bearing measurements are obtained using AIS Global Positioning System (GPS) receiver measurements from the



(a) Simulated maritime foghorn target vessel TX moving across blue line for sound localization using microphones relative to receiving vessel RX positioned at origin.



(b) Spectrum of recorded maritime wind noise signal used in the simulation.

Fig. 3. Simulation trajectory and wind noise spectrum.

receiving ship itself which is taken as the ground-truth in the measurements. Nearby vessels broadcast AIS messages. Hence, GPS information of both the transmitting vessel and receiving vessel is known. Heading of the ship θ is obtained from AIS as well. The GPS coordinates latitude and longitude of the receiving vessel are obtained as $(\varphi_{\mathrm{RX}}, \lambda_{\mathrm{RX}})$ and the transmitting vessel as $(\varphi_{\mathrm{RX}}, \lambda_{\mathrm{TX}})$. The relative bearing from GPS is estimated using the formula:

$$\beta_{\mathrm{GPS}} = \arctan\left(\frac{\sin(\lambda_{\mathrm{TX}} - \lambda_{\mathrm{RX}})\cos\varphi_{\mathrm{TX}}}{\cos\varphi_{\mathrm{RX}}\sin\varphi_{\mathrm{TX}} - \sin\varphi_{\mathrm{RX}}\cos\varphi_{\mathrm{TX}}\cos(\lambda_{\mathrm{TX}} - \lambda_{\mathrm{RX}})}\right), \tag{17}$$

and relative bearing reference value β_0 related to heading of the source ship is calculated using:

$$\beta_0 = \theta - \beta_{\mathrm{GPS}}. \tag{18}$$

Since the GPS data has typical time-resolution of 1 second, microphone bearing estimation with their confidence values are resampled at 1 Hz, averaging over 10 observations at 10 Hz. The error ϵ between microphone-estimated bearing $\beta_{0,k}$ and GPS-estimated bearing $\hat{\beta}_k$ is computed by taking into account only the errors such that the confidence value ρ_k is above a certain threshold ρ_θ:

$$\epsilon = \sqrt{\frac{\sum_k \mathbb{I}(\rho_k > \rho_\theta)(\hat{\beta}_k - \beta_{0,k})^2}{\sum_k \mathbb{I}(\rho_k > \rho_\theta)}}, \tag{19}$$

where $\mathbb{I}$ is the indicator function.

4. Simulation and experiment results

In this section we present the equipment used in the experiment with the setup itself and the results that are obtained for simulation tests, laboratory tests, and the field tests. Simulations were done to validate the presented algorithm for foghorn sound angular localization which incorporated Doppler effect and radial attenuation of the sound source. Algorithmic performance with source intensity for a given wind noise was analyzed. Laboratory tests were conducted outdoors with foghorn sounds played from a speaker. Field tests were done in the Gulf of Finland on-board MS Megastar by measuring foghorn sounds from the MS Star and MS Finlandia vessels by the microphone array. Processing and reporting of results was done in Matlab.

4.1. Simulation tests

In following, we describe the simulation method to validate the bearing-estimation algorithm presented in Section 3 for foghorn sound bearing estimation problem. We also evaluate the performance on a transmitter-receiver vessel scenario where transmitter vessel is producing the monotonic foghorn sound which is received by the receiving vessel. On the receiving vessel the bearing-estimation algorithm computes the direction of the incoming foghorn sound through error minimization on the measured phases or time delays. Bearing estimation error is obtained for foghorn sound intensities for a given wind noise.

In simulation tests, a single-frequency sinusoid signal having amplitude of pressure p, second harmonic $n = 2$ of the foghorn frequency of f_0 on the transmitting vessel TX is generated as a sinusoid signal with time t:

$$s_{\mathrm{TX}}(t) = p\sin(2n\pi f_0 t). \tag{20}$$

Here, the pressure amplitude p of the sound is calculated from the known source intensity I_0 in dB sound pressure level (dBSPL) as:

$$p = p_0\, 10^{I_0/20}, \tag{21}$$

where reference sound pressure is $p_0 = 20$ μPa. This generated sound source TX moves in a straight line with constant velocity relative to the receiving vessel RX with position as:

$$\vec{r}_{\mathrm{TX}}(t) = \vec{r}_0 + \vec{v}\,t + \vec{r}_{\mathrm{RX}}(t). \tag{22}$$

Microphone has coordinates relative to receiving vessel position r_{RX} as:

$$\vec{x}_1 = (D/2, \sqrt{3}D/2),\ \vec{x}_2 = (0,0),\ \vec{x}_3 = (D,0). \tag{23}$$

Signal received at each microphone is calculated assuming microphones at position $\vec{r}_{\mathrm{RX}} + \vec{x}_i$, $i = 1,2,3$. When sound reaches the i-th microphone on the receiving vessel RX, time required is $\Delta_i(t)$ and is given by:

$$\Delta_i(t) = \frac{|\vec{r}_{\mathrm{TX}} - (\vec{r}_{\mathrm{RX}} + \vec{x}_i)|}{c}. \tag{24}$$

Signal received by individual microphones is then $s_{\mathrm{RXi}}(t)$ which is calculated from signal delay $\Delta_i(t)$ and inverse distance attenuation [4] with added noise as:

$$s_{\mathrm{RXi}}(t) = \frac{s_{\mathrm{TX}}(t + \Delta_i(t))}{c\Delta_i(t)} + \eta_i(t). \tag{25}$$

Where p is the pressure amplitude, c is speed of sound taken to be 343.9 m/s, η_i is the recorded maritime noise signal added to the ith receiving microphone having spectrum shown in Fig. 3b.

4.2. Simulation results

From the aforementioned simulation strategy, received signals are calculated with following simulation parameters. Transmitting vessel TX has speed of $|\vec{v}| = 22.2$ m/s and a heading of $\angle\vec{v} = 235$ deg. Initial

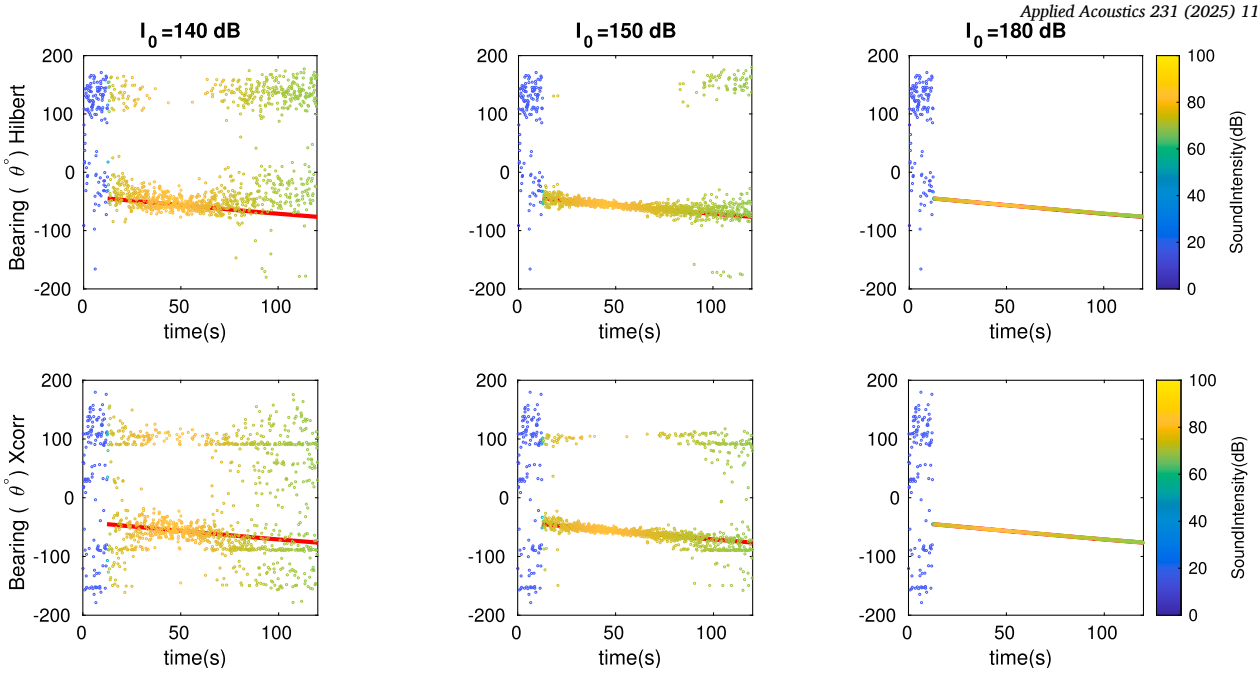


Fig. 4. Simulated trajectory localization results for source intensities of 140, 150 and 180 dB and method of Hilbert transform (top), and cross correlation (bottom). A sufficient threshold is required for algorithm to be more reliable.

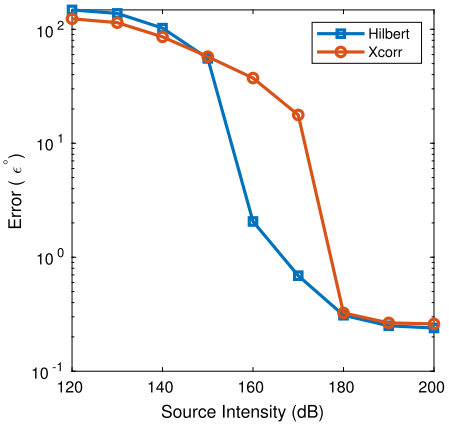


Fig. 5. Angular localization RMS error vs source foghorn signal intensity for Hilbert method of phase estimation, and cross correlation for delay estimation. Result obtained from simulated test on trajectory with various source intensities and added maritime wind noise.

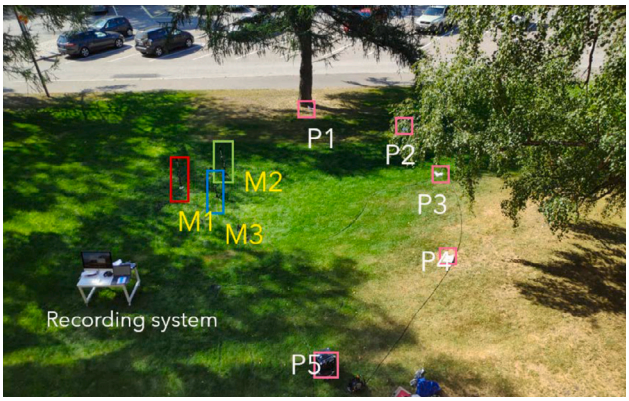
position of the transmitting vessel is $r_0 = (-3\text{ km}, 3\text{ km})$. Microphone separation of $D = 0.5$ m is used on the receiving ship. Rectangular time window of $T_w in = 0.1$ is used. Quantity c is speed of sound taken to be 343.9 m/s. Foghorn monotone sound frequency is $f_0 = 130$ Hz. Receiving vessel is assumed to be static, that is, $\vec{r}_{\text{RX}} = (0, 0)$. Simulated vessel trajectory is presented in Fig. 3a. Intensity spectrum of the recorded heavy wind noise signal $\eta(t)$ taken from the field experiments from 3 microphones is presented in Fig. 3b. Various sound source intensities $I_0 = 140$, 150 and 180 dB were simulated for direction estimation using both the methods of Hilbert transform and cross correlation (Xcorr) and their results are summarized in Fig. 4. For lower foghorn loudness $I_0 = 140$ dB there is a large variation in the obtained bearing estimations, and cross correlation performs better than Hilbert based method, but as the foghorn intensity is increased $I_0 = 150$ dB, the obtained variation decreases and the accuracy increases more for the Hilbert method. Soon cross correlation also performs better at increasing loudness. Both Hilbert and cross correlation based methods reach RMSE of 0.24° and 0.26° respectively at $I_0 = 200$ dB. Error is seen gradually decreasing to a steady value as intensity of foghorn is increased, as shown in Fig. 5. Thus, it suggests there requires a certain threshold of foghorn sound to have a desired bearing accuracy. This trend is consistent with the findings in experimental results of RoboArm 360 system in [30]. Accuracy result is in agreement with the simulation results published in Brix [6]. The simulation result presented in Choi and Kim [11] also shows accuracy $< 1°$, from the boundary element method using sound propagation model in 3D.

4.3. Laboratory tests

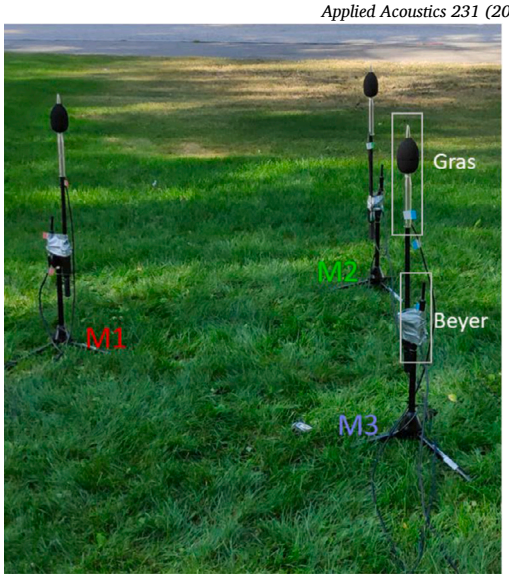
Laboratory tests were performed outside Otakaari 3 building, at Aalto University. Experiment setup is shown in Fig. 6a. The first set of microphones used in the laboratory experiments were *GRAS* 41AC-3 (a set including the membrane and a suitable preamplifier), which are pre-polarized, constant current power (CCP), IEC 61672-1 compliant outdoor measurement microphones for long-term unattended installations. Their main electrical characteristics are a frequency range of 3.15 Hz to 20 kHz (2 dB), a dynamic range of 17 dB(A) to 138 dB and a sensitivity of 50 mV/Pa. They are also equipped with a foam-type wind screen that partially eliminates the effects of air flow, and are overall optimized to reduce the impact on the microphone in the sound reception. In addition to the readiness for use in harsh outdoor environments, having a relatively large dynamic range is also an important feature, since strong wing gusts of the type frequently occurring in sea can easily and recurrently saturate the microphone and preamplifier outputs.

It is then also essential to select a data acquisition (DAQ) with a suitable matching dynamic range. The electrical signals sent by the microphone pre-amplifiers were sampled and recorded by a Dewesoft Sirius Mini data acquisition card (DAC), which can simultaneously capture signals from up to four different inputs. Its dynamic range is larger than 140 dB, with an input range between ± 500 mV to ± 10 V sampled by a 24 bits delta-sigma dual Analog-to-Digital Converter (ADC) for each channel, and capable of up to 200 KHz sampling. This DAC can also synchronize its internal time and sampling triggering directly using an network time protocol (NTP) server. This allowed us to closely match sound samples to signals coming from any other (time-synchronized) sensor such as GPS.

In addition to *GRAS* microphones, Beyerdynamic MM1 measurement microphones were used with Roland OCTA-CAPTURE recorder with



(a) Experiment setup with recording system, Microphone positions M_1, M_2 and M_3 and sound source locations P_1-P_5.



(b) *GRAS* and Beyer microphones placement.

Fig. 6. Laboratory test setup.

Logic Pro X software in Mac system for laboratory tests (Fig. 6a). Both selected *GRAS* and Beyerdynamic microphones are omni-directional. Addition to these, a general loudspeaker Classic Session 30 Mark II Bass combo was used to play artificial foghorn sounds in laboratory tests.

The tests consisted of analyzing the performance of 2 types of microphone arrays in an equilateral triangle arrangement - *GRAS* microphones with sampling rate of 50 KHz and Beyerdynamic MM1 measurement microphones with sampling rate of 44.1 KHz. In an array, microphones are labeled as M_1, M_2, and M_3 are placed in a equilateral triangle arrangement. Microphones were kept 1.5 m apart in an outdoor setup as shown in Fig. 6a. Sound source or the loudspeaker is oriented towards the microphone array from a distance of $\approx$ 8 m. Angle between subsequent positions P_1 to P_5 is kept $\approx$ 30 deg. Two kinds of microphones *GRAS* and Beyer shown in Fig. 6b were used to determine their performance which could be used on the ship. Artificial ship foghorn sounds of 10 seconds are constructed using the first 10 harmonics extracted from a real foghorn sounds having fundamental frequencies of ~90 Hz and ~110 Hz. Finally, 28 sound recordings for the microphone array were created for each location and each foghorn frequency, and 28 sound measurements are recorded as 6, 6, 6, 6, and 4 samples for P_1 to P_5 respectively.

4.4. Laboratory results

Results of the experiments conducted in backyard are presented in this section. We obtained quite satisfactory results with the direction estimation of the sources from positions P_1 to P_5 of sound as shown in Fig. 7. Summarized RMS error for both the types of microphones, both frequencies, and both methods are presented in Table 1. For Hilbert method, RMS error obtained for 90 and 110 Hz for *GRAS* microphones is 5.63° and 10.47° respectively, and for Beyer microphones is 6.84° and 5.23° respectively. For 110 Hz foghorn, error calculation involved only the points from the correct side of the plane (in range $[90°, 270°]$) and off-placed localizations are omitted from the error calculation due to their significant contribution to the error. Although, the deviation in the measurements, that is, error noise deviation is typically quite low, and is <1 deg. We found that direction results are affected by the speaker orientation and 110 Hz foghorn sound-based direction-estimation is affected by the relatively large microphone separation. For 90 Hz, *GRAS* microphones give RMS error of 5.63°, which is slightly lower than Beyer microphone RMS error 6.84°, and measured error noise deviation of 0.7° and 0.96° respectively is obtained for the microphones. For 110 Hz, inclusion of all localized points gives *GRAS* microphones the RMS error of 10.08°, while it is 104.91° for the Beyer microphones. This rather large error contribution is from the mislocated points P1 and P4 via the algorithm as seen in Fig. 7. In all, net RMS error of 7.04° with error noise deviation of 0.91° is obtained for the experiment for Hilbert method. On the other hand, cross correlation method shows lower errors for lower frequency of 90 Hz for both microphones compared to Hilbert method, but on the contrary higher errors for higher frequency of 110 Hz for both methods, and can be inferred from Table 1. Fig. 7 also suggests that error has offset due to possible speaker and microphone-array misalignment, since it has a narrow error deviation of $\approx 1°$ in backyard lab tests. Offset error can be attributed to the source speaker alignment and angle noise deviation provides resolution of the angle measurement, which is close to 1° in these tests. Further, the effect of reflections from the surroundings can be compensated for an even more accurate result.

Table 1
RMS Error in degrees for source direction estimation in 2D using equilateral triangle microphone configuration for Hilbert based and cross correlation method. For 110 Hz, error calculation only takes directions from range $[90°, 270°]$.

Method	Microphone	90 Hz RMSE (°)	110 Hz RMSE (°)
Hilbert	*GRAS*	5.63(0.7)	10.47(0.67)
Hilbert	*Beyer*	6.84(0.96)	5.23(1.33)
Xcorr	*GRAS*	5.61(0.63)	10.49(0.57)
Xcorr	*Beyer*	6.78(0.93)	10.81(1.04)

4.5. Field test

The field test campaign was conducted on-board MS Megastar vessel in the Gulf of Finland, in Baltic Sea. Goal of this campaign was to evaluate the sound bearing estimation accuracy of foghorn sounds from other vessels. Setup consisted of a *GRAS* microphone array, shown in Fig. 8a, which was installed on the deck above the command bridge in front of the vessel with a wall towards the aft, as shown in Fig. 8c. This position was desirable for installing microphones but offered a compromise between minimal wind noise and wider view angle, and was approximately located in the middle. This placement was aimed at reducing the vessel generated noises and the exposure of the microphones to the harsh winds experienced in the upper-most open deck. Microphones were also needed to be placed such that there are no obstructions to the incoming sound waves.

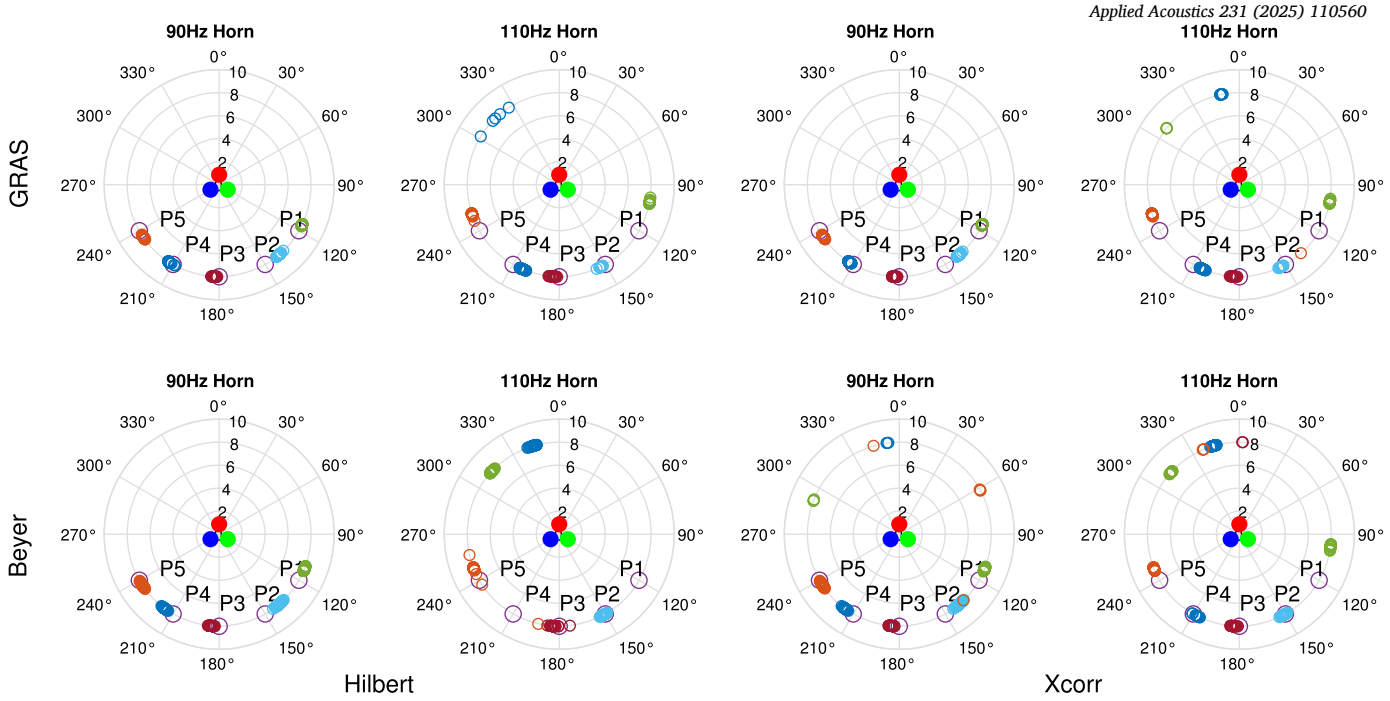
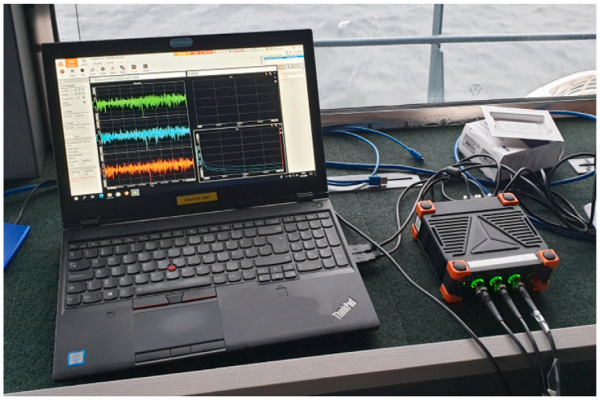


Fig. 7. Direction estimation results for the laboratory experiment for both Hilbert on the left halve and cross correlation method on right halve. Sound direction estimation from *GRAS* and Beyer microphones on simulated foghorn sounds with base frequency of 90 Hz and 110 Hz. Microphone array is placed at the center and loudspeaker points P1, P2, P3, P4, P5 with angular separation of 30 deg. Radial distances are in meters and are for graphing purpose.

(a) Detail of the *GRAS* microphone array configuration, front microphone points to ship heading.



(b) DAC and its associated software during a measurement campaign.

(c) Location of the microphone array, denoted by the red dot. Image copyright/courtesy of Tallink Group.

(d) Improvised sound shield used in selected experiments to alleviate the impact of reflections.

Fig. 8. Microphone setup on MS Megastar.

Table 2
Summary of events and results, net RMS error of * deg. Angular localization within 15° is represented in bold numbers.

Event	Encounter Time (UTC)	Ship	Absorber Used	RMSE (Hilbert) deg	RMSE (Xcorr) deg
1	10-Dec-2020 11:17:53	STAR	Yes	**10.74**	15.18
2	10-Dec-2020 12:17:25	FINLANDIA	Yes	**6.96**	**10.80**
3	10-Dec-2020 14:18:13	STAR	Yes	49.57	36.98
4	10-Dec-2020 15:17:26	FINLANDIA	Yes	**11.61**	18.12
5	10-Dec-2020 17:20:00	STAR	No	15.43	15.10
6	10-Dec-2020 20:14:00	STAR	No	below threshold	below threshold
7	03-Dec-2020 11:19:00	STAR	No	18.95	19.00
8	03-Dec-2020 12:10:30	FINLANDIA	No	24.76	26.59
9	03-Dec-2020 14:19:30	STAR	No	below threshold	below threshold

The inter-microphone distance $D = 50$ cm was decided in this field test which covered typical foghorn sound frequencies in maritime regions. The selected sampling frequency was of 50 KHz. Microphones had full symmetric visibility towards the bow measured clockwise from left to right as -90° to +90°, and were protected of the strongest winds along the bow by the wall at the aft. However, the protecting wall also presented itself as a good planar reflector of sounds, and as a consequence, foghorn sound reflections interfered with the tests on 3rd December. As a result, in order to quickly and informally curb their impact, we improvised the setup for few events foghorn sounds on 10th December with a handmade sound absorbing shield to avoid the reflections from the aft, which is shown in Fig. 8d. This shield was made out of many layers of clothing materials such as blankets and foam-based camping mattresses. It avoided suspected reflections from the wall behind the microphone installation. Fig. 8b shows the DAQ and the laptop running the DAQ software during one of the experiments.

MS Megastar has the pre-installed GNSS unit *DGNSS SAAB R5* with the AIS unit *R5 SUPREME AIS*. In addition to this AIS GPS unit, a manually installed commercial-grade multi-frequency multi-GNSS Topcon unit with measured precision of 0.25-0.5 m is used. This commercial-grade GNSS equipment has more resolution than the conventional pre-installed GSNN hardware. The sampling frequency of GNSS position receiver is 1 Hz.

In the field experiment, Captains of the other ships (MS Star and MS Finlandia) were asked to make the foghorn sounds at suggested times during the encounter between the ships. These foghorn sounds were recorded on-board the MS Megastar ship using the microphone array. Overall 9 sound event recordings were made for this experiment and their results are summarized in Table 2. Setup in events 1 to 4 had an additional cloth-based sound absorber on the back to attenuate any interference of suspected reflections. For ship foghorn, ferry vessel MS Star has foghorn model Kockum Sonics series *ELECTRO-TYFON* Mt 150/130 with fundamental frequency of 130 Hz. Ferry vessel MS Megastar has foghorn Zet-Horn 90AC with fundamental frequency of 90 Hz manufactured by Zöllner Signal GmbH, which is accordingly tested with standard COLREG72 Anne II/3, IEC 60945(2003).

4.6. Field test results

We present the results from the field tests conducted on-board MS Megastar. From the spectrum of the recorded foghorn sounds in the field tests. First harmonic of the foghorn for MS Star ferry was observed to be within 136-138 Hz, and for MS Finlandia was around 162 Hz. Second harmonics were found to be louder than the first harmonic. A rectangular time window of $T_{\text{win}} = 0.1$ s is used on the received microphone channel signals. Resonator filter settings used in the filter are $n = 2$ with the bandwidth $b_{\text{W}} = 12$ Hz, which centered at the observed recorded frequency for each vessel. In the algorithm, the bearing estimation optimization in Equation (15) was done on directions $\beta \in [-\pi/2, \pi/2]$ relative to MS Megastar, with 0° direction being the heading.

Foghorn sounds produced by other vessels are highlighted in a radar plot in Fig. 9, where we combine both AIS GPS measurements with the energy of the filtered received foghorn signal. This plot shows the ship trajectories and foghorn sound levels of MS Finlandia and MS Star relative to MS Megastar. Major portion of foghorn sounds are highlighted in the radar plot. It can be seen that the foghorn sound in event 1 is the loudest. Foghorn sounds particularly within 4 km radius can be heard by the algorithm sound, although, in the events 6 and 9, the sounds are not detected by the algorithm. In total, 10 separate foghorn sounds were audible in the 9 sample recordings. AIS GNSS information is used on experiment on 3rd and 10th December 2020, while Topcon GNSS receiver information is recorded only for experiments on 10th December 2020.

Foghorn sound bearing estimation with time of the ships from MS Megastar are provided in the Fig. 10. Angle estimations are more accurate when an absorber is used behind the microphone array, except in event 3. Error is evaluated as the root mean square (RMS) between the estimated bearing of target ship with respect to MS Megastar to the actual bearing obtained from AIS. Numerical error in sound direction obtained from Topcon GPS as reference was same as the error values obtained from AIS GPS data, as no difference was visible in the numerical values of bearing estimation. A detection is assumed during the time interval when the confidence value is larger than the confidence threshold. In this experiment the confidence threshold is assumed to be 0.5 (50%). Error is computed over the window of successful detection. Related obtained bearing error is presented in Table 2.

Referring to Fig. 9, a foghorn sound event from a target may be detected even up to 5 km away (as in event 8), but not always (as in the event 6 and event 9). Wind noise environment makes it challenging to detect far-away noise sources. Foghorn sound in event 6 was not heard by the algorithm, whereas, foghorn sound in the event 9 is audible to a human, but not to the algorithm. Maritime environment can drastically increase the error, as can be seen from event 3, in Fig. 10, and for RMSE calculations we have omitted events 3,6, and 9, unless explicitly mentioned. A good localization was achieved in the event 2 with the RMS error of 6.96° at distance of approximately 3.2 km. Table 2 summarizes the event details such as encounter time with target transmitting ship, use of the absorber, and the obtained RMS error in bearing estimation of the sound source. Event 1 to 4 involved the use of a makeshift sound absorbing screen whereas event 5 to 9 did not. Average performance in terms of detection success and RMS error appears better when the sound absorbing screen is employed. RMS error over events 1, 2, and 4 is found to be 9.97°, while without the sound absorbing screen is 20.08° for Hilbert approach. On the other hand, cross correlation method has RMSE of 15.00° with absorbing screen for events 1, 2, and 4, and has RMSE of 20.80° without. Average RMS error over all the listed events is 23.80° for Hilbert method, and 21.84° for cross correlation method. However, from individual experiments, Hilbert based approach can provide more accuracy in certain scenarios for direction estimation. Aforementioned algorithm can also perform in real-time, as 1 second of sound data processing takes 19 ms using Hilbert transform, and 30 ms using cross correlation method, on an Intel(R) Core(TM) i9-9920X CPU @ 3.50 GHz with 24 logical cores. Hence, these methods provide feasibility in real-time applications.

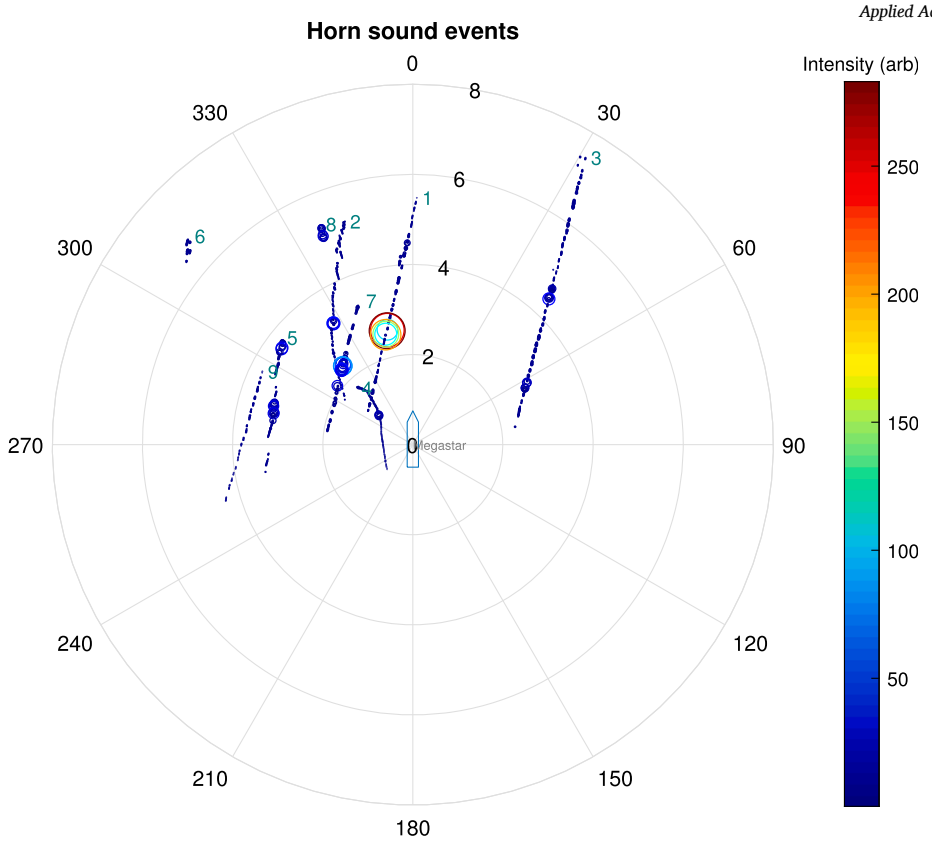


Fig. 9. Field tests foghorn sound events at various locations relative to MS Megastar for results presented in Table 2. Foghorn sound energy signals are fused with target vessel AIS GPS positions and are displayed on a radar chart.

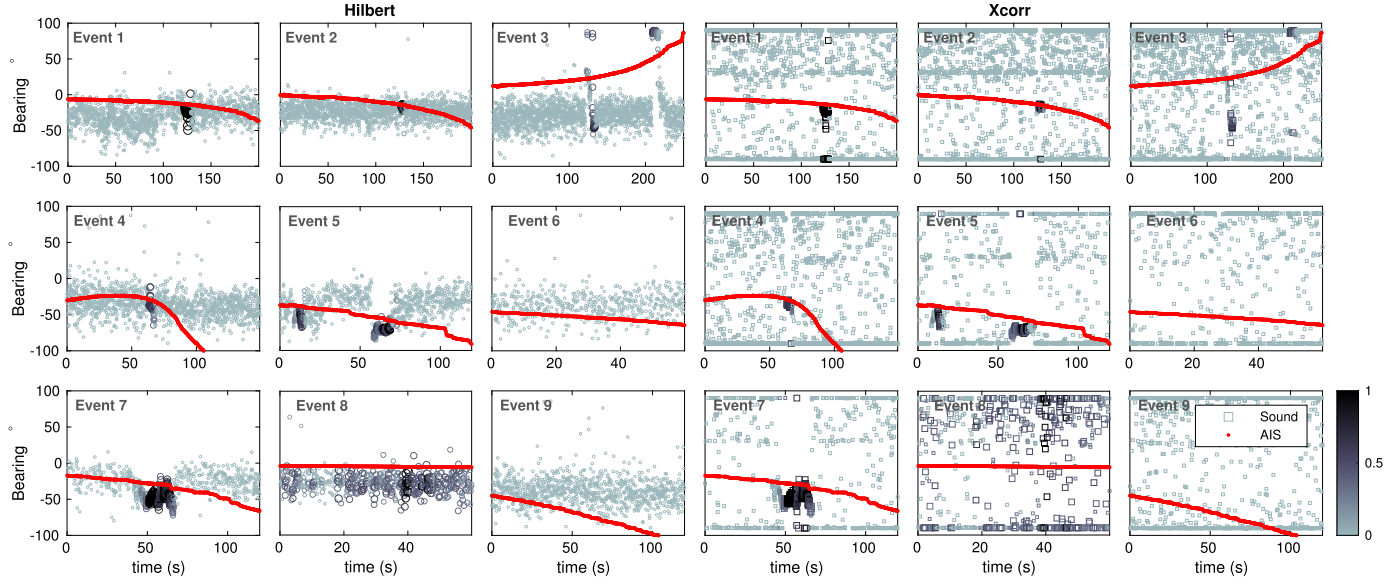


Fig. 10. Bearing results from the proposed algorithm for field tests conducted in the Gulf of Finland, Baltic Sea. Relative bearing estimation of target vessel for events 1 to 9 are shown with start time of encounter in UTC (see also Table 2) for Hilbert phase estimation method in left half and cross correlation time delay method in right half. Red curve indicates the AIS calculated bearing of target ship with respect to MS Megastar, and is taken as the ground truth. X-axis is the time in seconds since the start of event. Colorbar denotes the foghorn event confidence, and is measured in range 0 to 1. Darker regions with confidence higher than 0.5 were used to evaluate the performance of sound-based localization in maritime scenario using Equation (19).

5. Discussion

Use of foghorn is widespread and in addition many ships have a pre-installed digital sound reception system (SOLAS regulation, see Section 2). The question is, how would the current these systems fit into an autonomous perception system. In principle, interesting quantities to be measured are (i) the direction of the incoming sound, (ii) classification of the foghorn type, and (iii) the range to the sound origin, since these lead towards the purpose of (iv) matching the sound observations with visual (and other) observations.

(i) In the field tests, sound direction is estimated with a RMS error of $\sim 24°$ using Hilbert method, and $\sim 22°$ using cross correlation method, and RMSEs per event are listed in Table 2. This is clearly above the precision offered by current sound reception systems, which provide a resolution of 45° (8 directions as in Section 2). However, adverse weather conditions at sea may degrade this precision, as discussed in Section 2. We observe a significant source of error in direction estimation due to reflections from the observing vessel. Unfortunately, the abundant metal surfaces of vessels are extremely good sound reflectors. Therefore, we use a sound-absorbing structure on the background of the microphone array.

(ii) The classification of the foghorn types is plausible in our laboratory tests when differentiating between the foghorn frequencies of 90 and 110 Hz. However, at sea, more information is needed as it is impossible to cover different vessel types with frequency variations only. For this, we recommend future work as follows. Preferably, the sound receiving system should have a dictionary of specifications of the fog signals that can be observed. Also, the fog signals could adopt a form of Morse codes instead of emitting a constant sound so that the emitting vessel's types could be more easily distinguished from each other. Here, maritime regulations on foghorn sounds limit the use of the acoustic localization methods.

(iii) The range to the sound origin requires prior knowledge of the emission amplitude or would need multiple microphone arrays placed at sufficient distances. This is beyond this work. However, for future work, we note that even if the emission amplitude is known, the ranging is still strongly affected by the environmental conditions, especially the wind, as discussed in Section 2.

(iv) When hearing a foghorn sound, the captain of a ship is interested in visually finding the vessel that is the origin of the fog signal. As long as the origin of the sound source is not visually observed, a state of alertness is present in the command bridge due to uncertainty. Mimicking this with an autonomous system has (at least) the following caveat. Fusing the sound signal information with vision observations (as presented in Thombre et al. [45]) requires that the classification of the foghorn and hence the associated vessel to be successful. Otherwise, an autonomous system could only match the incoming fog signal with the first detected vessel in an image in the direction of the incoming sound omitting the chance of another vessel being there, out of sight.

6. Limitations and challenges

Despite the challenging conditions, we obtained satisfactory results from the experiments for sound localization using foghorn sounds. The challenges involved the nature of the sound data received by the microphone array on the MS Megastar, as it had: wind noise, Doppler effect, multi-path from sea-surface and ship itself, ship engine noise and chatter, birds near microphone array, water wave sounds, ferry's own whistle sounds, and microphone's mechanical response to the vessel vibrations. This makes the data rather noisy and filtering it is very challenging. Arising limitation for this work is that only a single bearing can be detected at a time. However, the same signal can be analyzed in parallel for multiple frequencies, depending on computational resources. Regardless, if multiple vessels make sound at the same time, then the direction of source cannot be obtained with the provided algorithm.

In collecting the field test data we also encountered numerous practical challenges due to which data collected was scarce in the experiments. First of all, gathering the data for the experiment was challenging due to cold weather and harsh winds for the microphones and as well as the crew members on the top deck of the ship. Secondly, vessels do not usually sound their foghorn without some specific reason, and hence foghorn events are typically rare at sea. Thirdly, the foghorn sound from nearby vessels was quite low and tended to be drowned out by background noise from wind and the sea waves. Solutions for overcoming these challenges need to be thought before future experiments.

7. Conclusion

In this paper we have presented the results of the experiments conducted to evaluate the feasibility of using foghorn sound for bearing estimation in the maritime scenarios. Based on the results, such microphone array installed on a ferry can be used for obtaining bearings of active foghorns of the other vessels. Best estimated angle-of-arrival RMS error with sound processing was found to be 6.96°. Average RMS error of 9.97° was obtained with the use of sound absorbing plane, and error of 20.08° without sound absorbing plane. We found the Hilbert phase estimation based approach to be more accurate than cross correlation based approach on our data. Although the bearing accuracy is limited, the measurement could still be included as an additional type of measurement in multi-sensor systems. The measurement might also be useful in search-and-rescue operations where sound source in the dark is the only signaling approach. Sound direction can also be used to align other instruments and focus them on the suspected direction so that more precise navigation tools can be aided for collision avoidance.

Algorithm and equipment used in this work has obvious potential improvements left for the future work. However, the improvements must be made with respect to the maritime scenarios which may involve redesigning the microphone array or involve a different approach altogether with sound localization techniques. We think, the presented algorithm is well suited for future developments to planar microphone array foghorn sound bearing estimation.

CRediT authorship contribution statement

Ajinkya Gorad: Data curation, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Zheng Zhao:** Data curation, Investigation, Methodology. **José M. Vallet García:** Data curation, Resources, Writing – review & editing. **Ville Lehtola:** Funding acquisition, Project administration, Supervision, Writing – review & editing. **Toni Hammarberg:** Data curation, Resources, Writing – review & editing. **Henrik Ramm-Schmidt:** Data curation, Project administration. **Saiful Islam:** Data curation. **Sarang Thombre:** Data curation, Funding acquisition, Project administration. **Simo Särkkä:** Data curation, Funding acquisition, Project administration, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We acknowledge the European Space Agency (ESA) Contract No. 4000125887/18/NL/DB, for the 'Artificial Intelligence / Machine Learning Sensor Fusion for Autonomous Vessel Navigation (Maritime AI-NAV)' project funding under NAVISP Element 1 Funding Program. We also thank the support of MS Megastar Tallink captains and crew members for allowing us to do the field experiments aboard the ship. Lastly, we thank the reviewers for supporting comments on the manuscript.

Data availability

Data will be made available on request.

References

[1] Allan JL. The foghorn's lament: the disappearing music of the coast. Hachette UK; 2021.

[2] Arranz P, De Soto NA, Madsen PT, Brito A, Bordes F, Johnson MP. Following a foraging fish-finder: diel habitat use of Blainville's beaked whales revealed by echolocation. PLoS ONE 2011;6:e28353.

[3] Bogue R. Underwater robots: a review of technologies and applications. Ind Robot, Int J 2015.
[4] Bolin K, Almgren M, Ohlsson E, Karasalo I. Long term estimations of low frequency noise levels over water from an off-shore wind farm. J Acoust Soc Am 2014;135:1106–14.
[5] Bolin K, Boué M, Karasalo I. Long range sound propagation over a sea surface. J Acoust Soc Am 2009;126:2191–7.
[6] Brix N. Array-based detection, identification and localization of marine signal horns. Master's thesis. Faculté Polytechnique; 2010.
[7] Chai SR, Zou YF, Li K, Li J, Wei YW, Guo LX, et al. Acoustic scattering from rough bubbly ocean surface based on SSA and EMM. IEEE J Ocean Eng 2024.
[8] Chen JC, Yao K, Hudson RE. Source localization and beamforming. IEEE Signal Process Mag 2002;19:30–9.
[9] Chen JC, Yao K, Hudson RE. Acoustic source localization and beamforming: theory and practice. EURASIP J Adv Signal Process 2003;2003:1–12.
[10] Choi J, Park J, Lee Y, Jung J, Choi HT. Robust directional angle estimation of underwater acoustic sources using a marine vehicle. Sensors 2018;18:3062.
[11] Choi JW, Kim KJ. A new sound reception system using a symmetrical microphone array and its numerical simulation. J Ship Ocean Technol 2004;8:18–25.
[12] Christensen JH, Hornauer S, Stella XY. BatVision: learning to see 3D spatial layout with two ears. In: 2020 IEEE international conference on robotics and automation (ICRA). IEEE; 2020. p. 1581–7.
[13] Cobos M, Antonacci F, Alexandridis A, Mouchtaris A, Lee B. A survey of sound source localization methods in wireless acoustic sensor networks. Wirel Commun Mob Comput 2017;2017.
[14] DiBiase JH. A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays. Ph.D. thesis. Brown University; 2000.
[15] Dmochowski J, Benesty J, Affès S. On spatial aliasing in microphone arrays. IEEE Trans Signal Process 2008;57:1383–95.
[16] Francis CFJ. A review of the current collision regulations to accommodate multiple ship situations and MASS 2019;1357:012011.
[17] Grondin F, Michaud F. Lightweight and optimized sound source localization and tracking methods for open and closed microphone array configurations. Robot Auton Syst 2019;113:63–80.
[18] Hioka T, Hamada T. Speaker direction tracking using microphones located at the vertices of equilateral-triangle. In: 2005 IEEE international symposium on circuits and systems. IEEE; 2005. p. 2847–50.
[19] Huang H, Sun D, Wang R, Zhu C, Liu B. Ship target detection based on improved YOLO network. Math Probl Eng 2020;2020.
[20] Hubbard B. Influence of atmospheric conditions upon the audibility of fog signals. J Acoust Soc Am 1931;3:111–25.
[21] IMO. Strategy for the development and implementation of e-navigation (e-NAV). International Maritime Organisation (IMO); 2006.
[22] Jekateryńczuk G, Piotrowski Z. A survey of sound source localization and detection methods and their applications. Sensors 2023;24:68.
[23] Kim SH, Ahn BW, Park KM, Lim GS, Das MP. Sound reception system by an acoustic meta-lens. J Int Marit Saf Environ Aff Shipp 2021;5:67–72.
[24] Knapp C, Carter G. The generalized correlation method for estimation of time delay. IEEE Trans Acoust Speech Signal Process 1976;24:320–7.
[25] Lebedev AV, Salin BM. Effects of localization of the areas of sound scattering by surface wind waves. Acoust Phys 2004;50:704–16.
[26] Li X, Shen M, Wang W, Liu H. Real-time sound source localization for a mobile robot based on the guided spectral-temporal position method. Int J Adv Robot Syst 2012;9:78.
[27] Li X, Zhao C, Yu J, Wei W. Underwater bearing-only and bearing-Doppler target tracking based on square root unscented Kalman filter. Entropy 2019;21:740.
[28] Li Y, Li Z, Zhu Y, Li B, Xiong W, Huang Y. Thermal infrared small ship detection in sea clutter based on morphological reconstruction and multi-feature analysis. Appl Sci 2019;9:3786.
[29] Luo H, Fang X, Ertas B. Hilbert transform and its engineering applications. AIAA J 2009;47:923–32.
[30] MacDonald JA. An algorithm for the accurate localization of sounds. Technical Report ADA454675. Army Research Laboratory; 2005.
[31] Medina D, Lass C, Marcos EP, Ziebold R, Closas P, García J. On GNSS jamming threat from the maritime navigation perspective. In: 2019 22th international conference on information fusion (FUSION). IEEE; 2019. p. 1–7.
[32] Merino-Martínez R, Sijtsma P, Snellen M, Ahlefeldt T, Antoni J, Bahr CJ, et al. A review of acoustic imaging methods using phased microphone arrays. CEAS Aeronaut J 2019;10:197–230.
[33] Nešković Đ, Marković M, Kašca L, Drajić D. Analysis of maritime positioning techniques for use in maritime internet of things systems. In: 2024 23rd international symposium INFOTEH-JAHORINA (INFOTEH). IEEE; 2024. p. 1–5.
[34] Oppenheim AV, Schafer RW, Buck JR. Discrete-time signal processing. 2nd ed. Prentice Hall, Inc.; 1999.
[35] Orfanidis SJ. Introduction to signal processing. Pearson; 1995.
[36] Paulose S, Sebastian E, Paul B. Acoustic source localization. Int J Adv Res Electric Electron Instrum Eng 2013;2:933–9.
[37] Proakis JG. Digital signal processing: principles algorithms and applications. Pearson Education India; 2001.
[38] Renton A. Lost sounds: the story of coast fog signals; 2001. Dundurn.
[39] Särkkä S, Vehtari A, Lampinen J. Rao–Blackwellized particle filter for multiple target tracking. Inf Fusion 2007;8:2–15.
[40] SOLAS. Rule 35, sound signals in restricted visibility. COLREG; 1972.
[41] Steckel J, Peremans H. BatSLAM: simultaneous localization and mapping using biomimetic sonar. PLoS ONE 2013;8:e54076.
[42] Su X, Ullah I, Liu X, Choi D. A review of underwater localization techniques, algorithms, and challenges. J Sens 2020;2020.
[43] Takahashi T, Matsumoto S, Nakayama M. A method of estimating degree of influence from wind for improving sound source localization in windy place. In: INTER-NOISE and NOISE-CON congress and conference proceedings. Institute of Noise Control Engineering; 2024. p. 6466–73.
[44] Teutsch H, Kellermann W. EB-ESPRIT: 2D localization of multiple wideband acoustic sources using eigen-beams. In: Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP). IEEE; 2005. p. 89–92.
[45] Thombre S, Zhao Z, Ramm-Schmidt H, Vallet García JM, Malkamäki T, Nikolskiy S, et al. Sensors and AI techniques for situational awareness in autonomous ships: a review. IEEE Trans Intell Transp Syst 2020.
[46] Tsukui T, Hirata S, Hachiya H. Investigation of the variability characteristics of sound waves reflected on the sea surface considering surface wavelength. J Mar Acoust Soc Jpn 2023;50:1–14.
[47] Tucker WS. Direction finding by sound. Nature 1936;138:111–8.
[48] Valin JM, Michaud F, Rouat J. Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering. Robot Auton Syst 2007;55:216–28.
[49] Vilhjálmsson BA, Engmarksgaard ME. Acoustic beamforming for identifying the relative direction of foghorns on ships. Technical Report. University of Southern Denmark; 2019.
[50] Wiener FM. Sound propagation over ocean waters in fog. J Acoust Soc Am 1961;33:1200–5.
[51] Wright RG. Intelligent autonomous ship navigation using multi-sensor modalities. TransNav: Int J Mar Navig Saf Sea Transp 2019;13.
[52] Zhuo DB, Cao H. Fast sound source localization based on SRP-PHAT using density peaks clustering. Appl Sci 2021;11:445.