Underwater Acoustic Source Localization Based On A Vector Sensor System: A Field Case Study

Xiaoyu Zhu[1], Hefeng Dong[1], and Guosong Zhang[2]

[1]Department of Electronic Systems, Norwegian University of Science and Technology, 7491 Trondheim, Norway
[2]Institute of Marine Research, Bergen, Norway

Contact author: Xiaoyu Zhu (email: xiaoyu.zhu@ntnu.no).

Abstract: *Underwater acoustic source localization is an important task for many marine applications. Compared to the conventional hydrophone array, an underwater acoustic vector sensor can simultaneously measure the acoustic pressure along with particle motions in three orthogonal directions. An experiment designed for source localization was conducted using a vector sensor in Trondheim Fjord, Norway in 2022. In this paper, the experiment and the analysis results of the collected data are presented. In the experiment, NTNU research vessel R/V Gunnerus was used to tow a broadband transducer that transmitted multitone signals from 400 to 3000 Hz. The data collected by the vector sensor are used for training a self-supervised machine learning model to provide the source range and post-processed to provide the azigram showing the source direction. The results are compared well with the known source location and the direction provided by the GPS system of the vessel, which illustrates the utility of the vector sensor system for source localization.*

Keywords: *Underwater acoustic source localization, vector sensor, machine learning.*

1. INTRODUCTION

Underwater source localization has become a hot topic in underwater acoustics for several decades, which is conventionally conducted by matched filter processing (MFP) based on acoustic pressure data collected by a hydrophone-based sensor system [1]. Compared with a hydrophone-based sensor, an acoustic vector sensor (AVS) can simultaneously measure the scalar acoustic pressure together with acoustic particle velocity at a specific location [2, 3]. Recently, AVS has become more popular in underwater acoustic applications, such as source localization [4] and azimuth estimation [2].

With the development of artificial intelligence, machine learning (ML) has shown superior performance compared to MFP [5], especially under the mismatch condition. Most existing ML-based studies in underwater acoustics are based on supervised learning which requires a vast labeled dataset to train the ML model. Unfortunately, the insufficiency of labels (e.g., source range/azimuth angle) is a common scenario in the field, which incurs performance degradation of the ML model. To deal with the limitation of labels, self-supervised learning (SSL) has been introduced for source localization, which obtains supervisory signals from unlabeled data. The localization performance of the SSL model has been demonstrated to outperform the purely supervised learning model under the scenario with limited labels in the room acoustic environment [6, 7] and ocean environment [8, 9, 10].

In this paper, we report a field experiment designed for source localization, that was conducted in Trondheim Fjord, Norway in 2022. In the experiment, a single AVS system from GeoSepctrum Technology Inc. (GTI) was mounted on the seabed to simultaneously collect the acoustic pressure data and the particle velocities in three orthogonal directions. NTNU research vessel R/V Gunnerus was used to tow a broadband transducer that transmitted multitone signals from 400 to 3000 Hz. Based on the collected data, a self-supervised model [9, 10] and the active intensity method [11, 12] are adopted for source localization and azimuth estimation, respectively. In addition, principle component analysis (PCA) [13] is exploited for better visualization and interpretation of the estimated azigram. In the performance analysis, the estimated source ranges and azimuths are compared with the ground truth provided by the GPS system of the vessel.

The remainder of this paper is organized as follows. Sec. 2 introduces the methods used for source localization and azimuth estimation. In Sec. 3, the details of the field experiment are presented. The comprehensive results analysis is given in Sec. 4. Finally, the conclusions are given in Sec. 5.

2. METHODOLOGY

2.1. SELF-SUPERVISED SOURCE LOCALIZATION

The self-supervised source localization follows a common workflow [10], which includes the training and the testing stages for building the source localizer and estimating the source location in real scenarios, respectively. In the training stage, a self-supervised feature extractor is trained based on an unlabeled dataset. In this paper, contrastive predictive coding (CPC) is adopted as the feature extractor, which consists of a non-linear encoder for compressing the input sequence into the latent space and a recurrent neural network (RNN) for summarizing

part of the latent codes into a context representation. During the training stage, the encoder and RNN are trained to jointly optimize a loss function based on Noise-Contrastive Estimation (NCE) for maximizing the mutual information. More details about the theory of the CPC model can be found in [14, 15].

After training the CPC-based feature extractor, only the encoder with frozen parameters is obtained for extracting high-level features. In this step, the task is to train the source localizer with the purely supervised learning scheme for solving the downstream task, i.e., source localization.

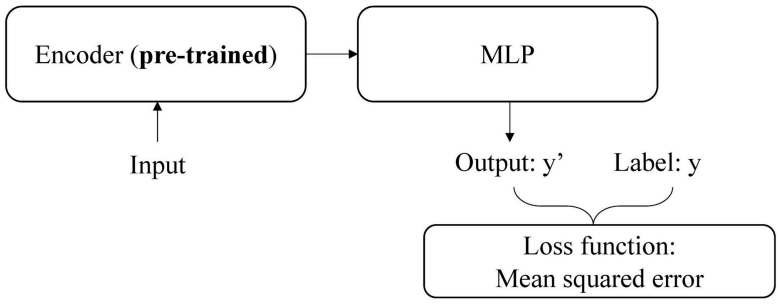

Figure 1: Architecure of source localizer.

The architecture of the source localizer is shown in Fig. 1, which consists of a pre-trained encoder for extracting high-level features from the input data and a multi-layer perceptron (MLP) for solving the downstream task. The MLP consists of several blocks that contain a dense layer, a batch norm layer or a dropout layer, and an activation function. Since source localization can be formulated as a regression problem, mean squared error (MSE) is chosen as the loss function.

2.2. AZIMUTH ESTIMATION

In this paper, the active intensity method is adopted to estimate the azigram [11, 12]. The active intensity I of the acoustic field at one time-frame T and frequency f is formulated as

$$I_k(f,T) = \text{Re}(p(f,T)v_k^*(f,T)) \tag{1}$$

where p and v refer to acoustic pressure and particle velocity, respectively. k indicates either the x or y direction. Re(.) is the real part operator. And $*$ represents a complex conjugate.

The estimated azimuth of the signal at time-frame T and frequency f is given by

$$\theta(f,T) = \tan^{-1}(\frac{I_y(f,T)}{I_x(f,T)}) \tag{2}$$

where the x and y direction point to the north and east, respectively. The estimated azimuth is arranged following the geographic convention, i.e., θ increases clockwise from the x axis.

3. FIELD EXPERIMENT

The field experiment was carried out on 29 September 2022 in Trondheim Fjord, Norway. It was designed to test the AVS system for acoustic source localization. The bathymetric map of

the area is shown in Fig. 2. Point A (red dot) refers to the location of the AVS system. Points M and N (the blue dots) indicate the start and end points of the vessel's track, respectively.

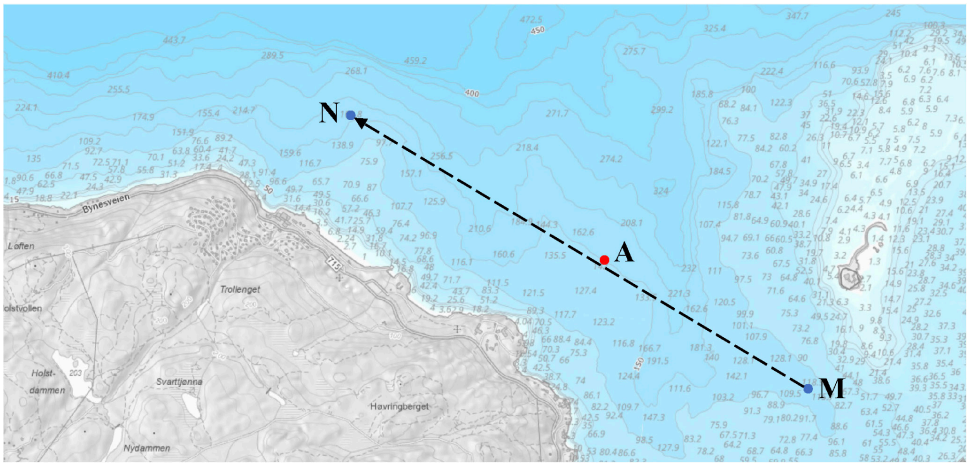

Figure 2: Bathymetry diagram of the experiment site.

During the experiment, a broadband transducer continuously transmitting multitone signals (from 400 to 3000 Hz with a 50 Hz interval) was towed by the NTNU research vessel R/V Gunnerus from point M to N. Along the track, the water depth varies from about 100 to 200 m, which incurs difficulties for the conventional source localization method. The AVS system was deployed on the seabed at 150 m depth, which collected four-channel signals (i.e., acoustic pressure and the particle velocities in three orthogonal directions) with a sampling rate of 8000 Hz. The length of the recorded signal was 2392 seconds. In addition, the GPS data collected by the built-in system of the vessel were converted into the source ranges (horizontal distances between the source and the AVS system) and the azimuth angles related to the north and increasing clockwise. Note that the data have been band-pass filtered between 300-3000 Hz and transformed into spectrograms with a 1-sec time window for further processing.

4. RESULTS ANALYSIS

4.1. AZIMUTH ESTIMATION

Fig. 3(a) provides the normalized spectrogram of the pressure channel, where the horizontal axis and vertical axis refer to the time and frequency axes, respectively. In the figure, the transmitted multitone signals mainly spanning from 500 to 3000 Hz can be clearly revealed. Along the time axis, three main events happened around 12.5 min, 17.5 min, and 38.5 min. According to the ground truth of the source ranges shown in Fig. 3(c), the research vessel was passing the closest point of approaching (CPA) of Point A around 17.5 min. The other events with higher intensity in the spectrogram can be interpreted as the unknown ships passing Point A. For a better description, the events happening around 12.5 min and 38.5 min are named noise events 1 and 2, respectively. The calculated azigram is plotted in Fig. 3(b) which shows that the azimuth of the unknown ships in noise events 1 and 2 changes from around 90 to 300 degrees and 120 to 230 degrees, respectively. However, due to the contamination of these noise events, the azimuth of the research vessel is hard to be recognized directly from the original azigram.

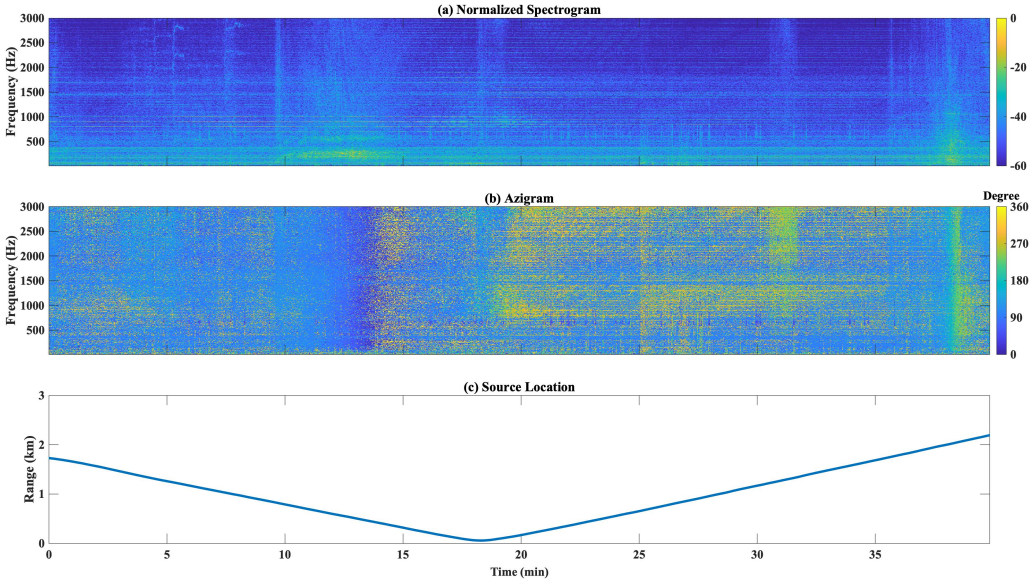

Figure 3: (a) Normalized spectrogram of the pressure channel, (b) Azigram, and (c) Ground truth of source range.

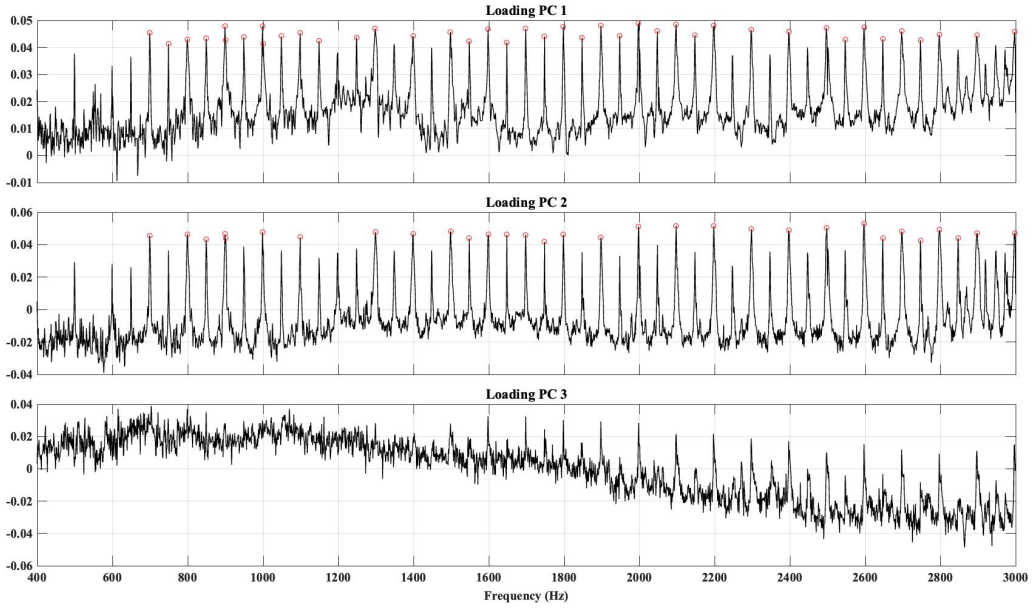

Figure 4: Loadings of PCs 1-3.

The azigram is post-processed by the PCA method for better visualization and interpretation. The loadings of the first three principle components (PCs) are shown in Fig. 4, where the transmitted tones can be clearly viewed in the loadings of PC 1 and PC 2 but not in that of PC 3. This means that the first two PCs mainly express source-related information. In addition, the red circles indicate the significant variables selected according to a preset threshold for representing the source information.

The azigram reconstructed by PCs 1-2 is shown in Fig. 5(a), which is much less noisy compared to the original azigram shown in Fig. 3(b). In Fig. 5(b), the blue and red curves indicate the estimated and ground truth of the source azimuths, respectively. The estimated

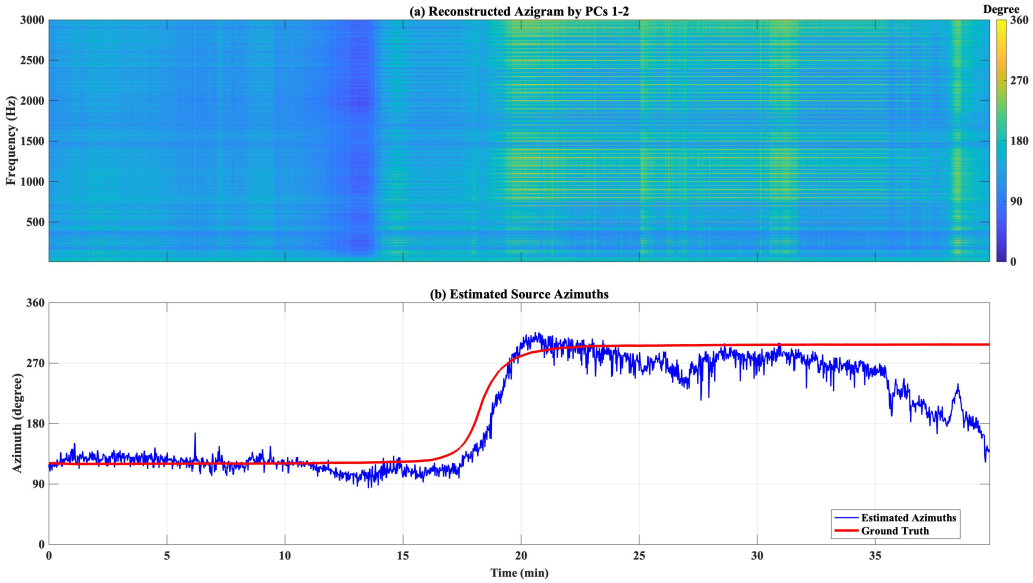

Figure 5: (a) Reconstructed azigram by PC 1 and PC 2, and (b) estimated source azimuths with the ground truth.

azimuth is calculated by the mean azimuth of the selected significant variables (marked by the red circles in Fig. 4). The curve of the estimated azimuth is mainly consistent with that of the ground truth with the exception of noise event 2, which indicates the benefits of the PCA method for better visualization and interpretation of the azigram.

4.2. SOURCE LOCALIZATION

As mentioned in Sec. 2.1, the first step is to train the self-supervised feature extractor based on all collected unlabeled data for extracting the high-level features. After that, to mimic the real scenario with limited labels, only 12.5% labeled data (i.e., 299 samples) are used to train the CPC-based source localizer.

To assess the benefits of the AVS system compared to the single hydrophone for source localization, three CPC-based source localizers with the same architecture are trained based on channel P (only the data from the pressure channel), channel PXY (the data from the pressure channel as well as the x and y channels of particle velocities), and channel PXYZ (the data from all channels), respectively. The performance comparison is shown in Tab. 1, where the metric is root mean squared error (RMSE) in meters.

Channel P	Channel PXY	Channel PXYZ
145.02	142.27	102.78

Table 1: Performance comparison between localizers trained based on different channels.

From Tab. 1, an intuitive phenomenon is revealed, i.e., compared to only using the pressure channel, adding more channels can improve the localization performance. Specifically, adding horizontal channels of particle velocities slightly improves the performance, however, adding channel Z can further boost performance significantly. The localization result by the localizer

trained by the data with all channels is shown in Fig. 6, where the red line and blue dot refer to the true and estimated source locations, respectively.

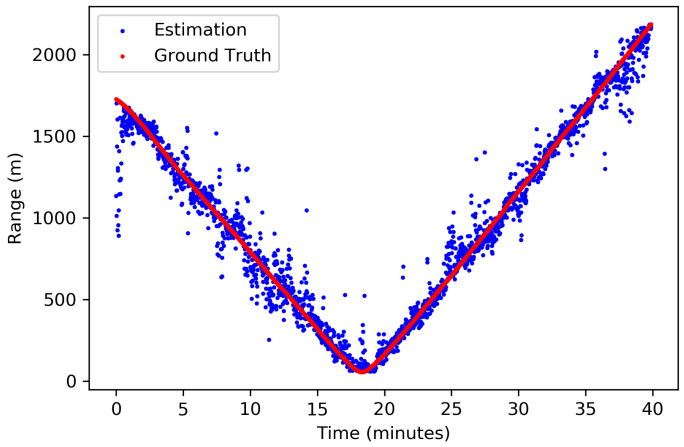

Figure 6: Source localization result by the localizer trained based on the data with all channels.

5. CONCLUSIONS

In this paper, we present a field experiment conducted in Trondheim Fjord, Norway in September 2022. The collected multi-channel data are used for source localization and azimuth estimation by a CPC-based source localizer and the active intensity method, respectively. In addition, the PCA method is exploited as a post-processing for a better interpretation of the calculated azigram. The results compared well with the ground truth provided by the GPS system of the research vessel, which demonstrates the effectiveness of the AVS system for acoustic source localization and azimuth estimation.

6. ACKNOWLEDGEMENTS

Hefeng Dong would like to acknowledge Equinor for providing the data and acknowledge the Norwegian Research Council and the industry partners of the GAMES consortium at NTNU for financial support (grant no. 294404). This work was partially supported by the SFI Centre for Geophysical Forecasting under grant 309960. Xiaoyu Zhu would like to acknowledge the China Scholarship Council (CSC) for the fellowship support (no. 201903170205). The authors acknowledge the scientists, technicians, and captain crew of R/V Gunnerus who participated in the field experiment.

7. REFERENCES

[1] A. B. Baggeroer, W. Kuperman, and H. Schmidt, “Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem,” *The Journal of the Acoustical Society of America*, vol. 83, no. 2, pp. 571–587, 1988.

[2] A. Zhao, L. Ma, J. Hui, C. Zeng, and X. Bi, "Open-lake experimental investigation of azimuth angle estimation using a single acoustic vector sensor," *Journal of Sensors*, vol. 2018, 2018.

[3] P. Stinco, A. Tesei, G. Ferri, S. Biagini, M. Micheli, B. Garau, K. D. LePage, L. Troiano, A. Grati, and P. Guerrini, "Passive acoustic signal processing at low frequency with a 3-d acoustic vector sensor hosted on a buoyancy glider," *IEEE Journal of Oceanic Engineering*, vol. 46, no. 1, pp. 283–293, 2020.

[4] P. Felisberto, O. Rodriguez, P. Santos, E. Ey, and S. M. Jesus, "Experimental results of underwater cooperative source localization using a single acoustic vector sensor," *Sensors*, vol. 13, no. 7, pp. 8856–8878, 2013.

[5] M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot, and C.-A. Deledalle, "Machine learning in acoustics: Theory and applications," *The Journal of the Acoustical Society of America*, vol. 146, no. 5, pp. 3590–3628, 2019.

[6] M. J. Bianco, S. Gannot, and P. Gerstoft, "Semi-supervised source localization with deep generative modeling," in *2020 IEEE 30th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, IEEE, 2020.

[7] M. J. Bianco, S. Gannot, E. Fernandez-Grande, and P. Gerstoft, "Semi-supervised source localization in reverberant environments with deep generative modeling," *IEEE Access*, vol. 9, pp. 84956–84970, 2021.

[8] X. Zhu, H. Dong, P. Salvo Rossi, and M. Landrø, "Feature selection based on principal component regression for underwater source localization by deep learning," *Remote Sensing*, vol. 13, no. 8, p. 1486, 2021.

[9] X. Zhu, H. Dong, P. S. Rossi, and M. Landrø, "Self-supervised underwater source localization based on contrastive predictive coding," in *2021 IEEE Sensors*, pp. 1–4, IEEE, 2021.

[10] X. Zhu, H. Dong, P. S. Rossi, and M. Landrø, "Time-frequency fused underwater acoustic source localization based on contrastive predictive coding," *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13299–13308, 2022.

[11] D. R. Dall'Osto, P. H. Dahl, and J. Woong Choi, "Properties of the acoustic intensity vector field in a shallow water waveguide," *The Journal of the Acoustical Society of America*, vol. 131, no. 3, pp. 2023–2035, 2012.

[12] A. M. Thode, T. Sakai, J. Michalec, S. Rankin, M. S. Soldevilla, B. Martin, and K. H. Kim, "Displaying bioacoustic directional information from sonobuoys using "azigrams"," *The Journal of the Acoustical Society of America*, vol. 146, no. 1, pp. 95–102, 2019.

[13] F. Westad, M. Hersletha, P. Lea, and H. Martens, "Variable selection in pca in sensory descriptive and consumer data," *Food Quality and Preference*, vol. 14, no. 5-6, pp. 463–472, 2003.

[14] A. v. d. Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," *arXiv preprint arXiv:1807.03748*, 2018.

[15] C.-I. Lai, "Contrastive predictive coding based feature for automatic speaker verification," *arXiv preprint arXiv:1904.01575*, 2019.