ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10094978

LEARNING ENVIRONMENTAL STRUCTURE USING ACOUSTIC PROBES WITH A DEEP NEURAL NETWORK

Toros Arikan[⋆]*, Amir Weiss*[⋆]*, Hari Vishnu*[‡]*, Grant Deane*[◦]*, Andrew Singer*[†]*, and Gregory Wornell*[⋆]

[⋆]Dept. of EECS MIT | [‡]ARL NUS | [◦]Scripps Inst. of Oceanography UCSD | [†]Dept. of ECE UIUC

ABSTRACT

Learning the physical environment is an important yet challenging task in reverberant settings such as the underwater and indoor acoustic domains. The locations of reflective boundaries, for example, can be estimated using echoes and leveraged for subsequent, more accurate localization. Current boundary estimation methods are constrained to a regime of high signal strength, or mitigate noise with heuristic (suboptimal) filters. These limitations can lead to fragile estimators that fail under non-ideal conditions. Furthermore, many algorithms in the literature also require a correct assignment of echoes to boundaries, which is combinatorially hard. To evade these limitations, we develop a convolutional neural network method for robust 2D boundary estimation, given known emitter and receiver locations. Our method uses as its input data format transform images, which are the potential boundary locations mapped into curves. We demonstrated in simulations that the proposed neural network method outperforms alternative state-of-the-art algorithms.

Index Terms— Convolutional neural networks, delay estimation, localization, underwater acoustics.

1. INTRODUCTION

Environmental learning in reverberant settings is an important task for difficult domains such as underwater and indoor acoustics [1–3]. Depending on the application, environment learning itself may be the goal, or a step in a processing chain for enhanced localization accuracy. For example, to passively localize an unknown emitter with a collection of receivers, line of sight (LOS) arrivals to the receivers are used for time difference of arrival (TDOA) [4] or time of arrival (TOA) localization [5], [6]. Morevoer, leveraging the non-line of sight (NLOS) arrivals can enhance localization performance [7].

Within the general scope of environment learning, we focus on reflective boundary estimation for shallow-water underwater acoustics, as in Fig. 1. We envision estimating the emitter location using LOS arrivals, and then using it and the NLOS arrivals to accurately estimate the boundaries. While prior knowledge of the sea surface and seafloor is typically available, more accurate estimates of their positions can be required. Over short ranges, these boundaries can be approximated as planar, thus giving rise to mirror images of the true emitter as 'virtual' emitters, as per Snell's Law. Euclidean distance matrices (EDM) [8] or other methods [9] can then be used for boundary estimation through virtual emitter localization. Ocean applications, however, can feature low signal-to-noise ratios (SNRs) [10] and model mismatch due to dynamic environments, which is not addressed by existing methods.

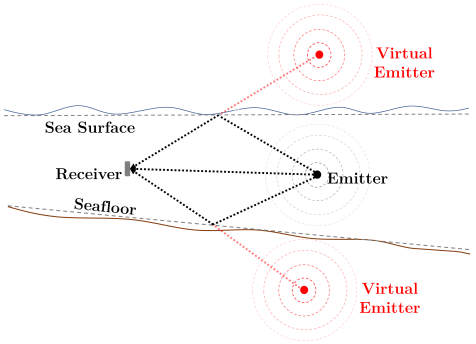


Fig. 1: A general underwater acoustic setting, highlighting the typical NLOS arrivals and corresponding virtual emitters.

An alternative boundary estimation approach leverages the fact that in 2D, a NLOS arrival corresponds to a path distance of d_{NLOS} and yields an ellipse whose foci are the emitter and receiver location, as in Fig. 2. By definition, points on the ellipse have a total distance of d_{NLOS} to the emitter and receiver, and the reflective boundary itself is a tangent to this ellipse. With multiple receivers, multiple ellipses are defined by such NLOS arrivals, and the boundary is their common tangent. Therefore, by fitting common tangents to ellipses, the boundaries can be estimated while avoiding the echo labeling problem in multipath environments, as illustrated in Fig. 3. Assigning ellipses to tangents is combinatorially complex [11] and complicated by missing or spurious arrivals, motivating a solution that bypasses this task.

In light of these challenges, we propose a convolutional neural network-based (CNN) method for boundary estimation. We paramterize the tangents to ellipses by their range ρ and azimuth θ [7], calling this ρ- and θ-space the common tangents to spheroids (COTANS) domain. We then map the

This work was supported, in part, by ONR under Grants N00014-19-1-2661, N00014-19-1-2662 and N00014-19-1-2665, and NSF under Grant CCF-1816209.

Authorized licensed use limited to: MIT Libraries. Downloaded on June 21,2023 at 14:34:32 UTC from IEEE Xplore. Restrictions apply.

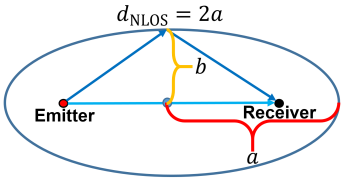


Fig. 2: NLOS arrivals define ellipses with foci at the emitter and receiver positions.

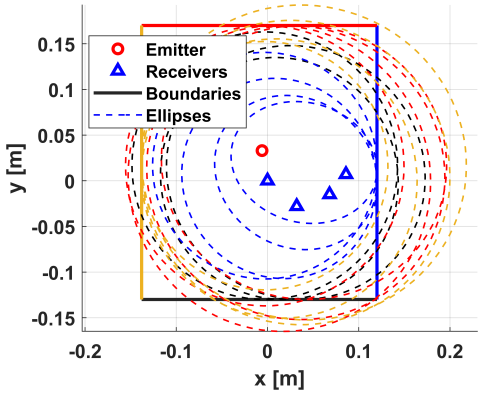


Fig. 3: Illustration of a multipath setting: each NLOS arrival gives rise to an emitter-receiver pair ellipse, as in Fig. 2.

environment and the time-delay estimates to images in the COTANS domain, transforming the data into a more natural input representation for the CNN. The proposed COTANS neural network (NN), termed COTANS-NN, is a modified AlexNet [12] that is trained on synthetic data to estimate boundaries from unlabeled NLOS arrivals. The resulting NN can be used with both simulated and recorded data.

Fitting tangent planes to spheroids for boundary estimation was proposed in [13], [14] as the common tangent (COTA) algorithm. In [15], a Hough transform-inspired methodology was used, with provisions to reject incorrectly-chosen echoes. To avoid conflating the plane-fitting method with the Hough transform, we refer to a COTANS transform and COTANS domain. These methods typically apply a heuristic smoothing filter to COTANS images, followed by the extraction of maxima. This hand-crafted approach is generally suboptimal and setting-specific.

If, however, a NN is trained with a wide range of geometries and realistic estimation errors, it can potentially learn the optimal inference rule, which can be viewed as joint (and implicit) filtering and peak extraction. The resulting NN can also be re-trained for different environments, for applicability to a wide range of settings (e.g., different numbers of boundaries and receivers). Our main contribution is exactly such an architecture, which provides superior boundary estimates—without echo labeling—using COTANS images as its input.

The paper is organized as follows. In Section 2, we formulate the environment and signal models. The COTANS-NN method is detailed in Section 3. Simulation results are presented in Section 4, and final remarks in Section 5.

2. PROBLEM FORMULATION

We model a static 2D environment with N planar boundaries, where N is known. These boundaries are described by the range $\rho \in \mathbb{R}_+$ and azimuth $\theta \in [-\pi, \pi)$ of their normal vector relative to the (arbitrarily-chosen) origin. Thus, the j-th boundary is parametrized as the vector $\boldsymbol{\eta}_j = [\rho_j \ \theta_j]^{\mathrm{T}}$, for all boundaries $j \in \mathcal{S}_N$, where we denote $\mathcal{S}_K \triangleq \{1, \ldots, K\}$ for some $K \in \mathbb{N}$. We assume a single isotropic emitter in the environment at a known location $\mathbf{p}_{\mathrm{e}} = [x_{\mathrm{e}} \ y_{\mathrm{e}}]^{\mathrm{T}}$, and M isotropic receivers at known locations $\mathbf{p}_{\mathrm{r},i} = [x_i \ y_i]^{\mathrm{T}}, i \in \mathcal{S}_M$. The speed of sound, denoted by v_{s}, is assumed to be constant, which is a reasonable simplification at short ranges in a well-mixed shallow-water underwater environment [16].

The received signal at the i-th receiver, $r_i(t) \in \mathbb{R}$, is modeled as the sum of the LOS arrival and single-reflection NLOS arrivals, delayed by their respective TOAs. The LOS TOA, denoted by $\tau_{i,0}$, is given by:

$$\tau_{i,0} = \frac{\|\mathbf{p}_{\mathrm{r},i} - \mathbf{p}_{\mathrm{e}}\|_2}{v_{\mathrm{s}}}, \ \forall i \in \mathcal{S}_M. \tag{1}$$

For the j-th boundary, we obtain the virtual emitter location $\mathbf{p}_j$ by finding the corresponding reflection of $\mathbf{p}_{\mathrm{e}}$ (see Fig. 1). The NLOS TOA to the i-th receiver from the j-th boundary, denoted by $\tau_{i,j}$, is equal to the TOA from the i-th receiver ($\mathbf{p}_{\mathrm{r},i}$) to the corresponding j-th virtual emitter ($\mathbf{p}_j$):

$$\tau_{i,j} = \frac{\|\mathbf{p}_{\mathrm{r},i} - \mathbf{p}_j\|_2}{v_{\mathrm{s}}} \triangleq \frac{d_{i,j}}{v_{\mathrm{s}}}, \ \forall i \in \mathcal{S}_M, \ \forall j \in \mathcal{S}_N. \tag{2}$$

Merging the effects of attenuation and reflection into the equivalent attenuation coefficient $\alpha_{i,j}$, the received signal at the i-th receiver is modeled by:

$$r_i(t) = \sum_{j=0}^{N} \alpha_{i,j} s\left(t - \tau_{i,j}\right) + \xi_i(t), \tag{3}$$

where $s(t)$ is the known emitted signal, and $\xi_i(t)$ is a noise signal that is a realization of a spectrally-flat Gaussian process. Given the energy of a received (and attenuated) pulse as E_r, the SNR is defined as E_r/N_0, where N_0 is the one-sided power spectral density of the noise $\xi_i(t)$. In practice, the environment can be reverberant, and $r_i(t)$ can feature higher-order reflections and noise that may not be Gaussian [17]. We work with a discrete-time sampled version of (3) as $r_i[n] \triangleq \{r_i(t)|_{t=nT_{\mathrm{s}}}\}_{n\in\mathbb{Z}}$, where T_{s} is the sampling period.

The geometric information for boundary estimation consists of the known $\mathbf{p}_{\mathrm{e}}$ and $\{\mathbf{p}_{\mathrm{r},i}\}$, and the unknown $\{\tau_{i,j}\}$. Hereafter, we assume that the NLOS TOAs are estimated using an (at least asymptotically) optimal estimator. For example, these estimates can be obtained by matched-filtering $r_i[n]$ with $s[n] \triangleq s(nT_{\mathrm{s}})$ and picking the TOAs corresponding to the N largest peaks in the result, after removing the LOS arrival, as $\{\hat{\tau}_{i,j}\}$. The distance estimates $\{\hat{d}_{i,j} \triangleq v_{\mathrm{s}}\hat{\tau}_{i,j}\}$ (from (2)) are then used for estimating the boundaries as $\{\hat{\boldsymbol{\eta}}_j\}_{j=1}^{N}$.

At high SNR, the error in $\hat{\tau}_{i,j}$ due to Gaussian noise in

Authorized licensed use limited to: MIT Libraries. Downloaded on June 21,2023 at 14:34:32 UTC from IEEE Xplore. Restrictions apply.

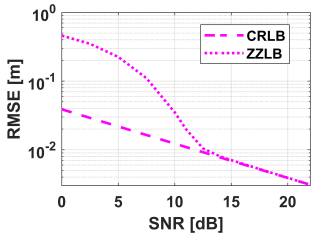


Fig. 4: Bounds on the range estimation root-mean squared error (RMSE) for a Gaussian pulse of 15.4 kHz bandwidth.

(3) is Gaussian itself, with mean squared error (MSE) that asymptotically coincides with the Cramér-Rao lower bound (CRLB) for time-delay estimation [18]. However, there is an SNR threshold below which large 'global errors' occur, leading to a drastic performance reduction as captured by the Ziv-Zakai lower bound (ZZLB) [19], depicted in Fig. 4. In this case, the matched-filtered signal can have larger peaks than the peak of the true arrival; picking such a spurious peak causes $\hat{\tau}_{i,j}$ to be distributed uniformly on the observation time interval. A global error yields an unserviceable time-delay estimate for boundary estimation; thus, a practical estimation method must be robust to large errors in a subset of $\{\hat{d}_{i,j}\}$.

A key motivation for the use of NNs in this context is that, while it is reasonable to assume Gaussian noise in $r_i(t)$ (in (3)), it is certainly not the case for errors in $\hat{\tau}_{i,j}$. The nonlinear time-delay estimation leads to a non-Gaussian estimation error. Hence, it is hard to analytically design a (non-asymptotic) optimal estimator for the boundary locations based on $\{\hat{\tau}_{i,j}\}$, encouraging a data-driven approach instead.

3. COTANS-NN FOR BOUNDARY ESTIMATION

We now discuss how the COTANS transform is used to generate images for a given geometry and the estimated $\{\hat{d}_{i,j}\}$. We then detail the COTANS-NN method for estimating the boundaries $\{\boldsymbol{\eta}_j\}$ from such images.

3.1. Generation of COTANS Images

In 2D, a boundary defined by ρ and θ can be conceptualized as a point (ρ, θ) in a COTANS domain; working out the (ρ, θ) expression of a line is to take its COTANS transform [20]. Discretizing the space $\rho \times \theta$ as a matrix and incrementing this 'accumulator' over every potential (ρ, θ) for NLOS ellipses yields a COTANS-domain image (as in Fig. 5), with maxima at the true boundaries $\{(\rho_j, \theta_j)\}$ in the absence of errors.

For each θ, we obtain ρ for a standard ellipse (Fig. 6(a)) as:

$$\rho(\theta) = \sqrt{a^2 \cos^2\theta + b^2 \sin^2\theta}, \tag{4}$$

where $a = d_{\text{NLOS}}/2$ and $b = \sqrt{d_{\text{NLOS}}^2 - d_{\text{LOS}}^2}/2$, as in Fig. 2. Then, (4) is used to generate the $\{(\rho, \theta)\}$ for a standard ellipse of the same size as the true ellipse (e.g., Fig. 6(b)). The COTANS transform then modifies ρ and θ based on rotations and translations that map one ellipse to the other.

In the absence of time-delay estimation errors, picking the maxima in COTANS images yields the exact and correct

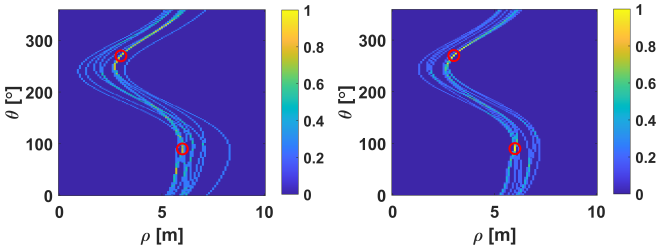


Fig. 5: COTANS images for low SNR with dispersed curves (a) and high SNR (b), with the true boundaries marked in red.

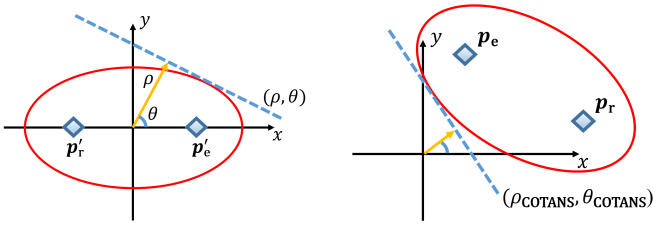


Fig. 6: Taking a tangent's COTANS transform: start from a standard ellipse (a), and modify to match the true setting (b).

(ρ_j, θ_j) for each boundary. Plane-fitting boundary estimation methods currently apply heuristic filters to these images, to locally average COTANS curves that do not exactly intersect. This methodology does not match the underlying statistics of COTANS images, and requires hand-crafting parameters such as filter sizes [7]. Our NN automates these tasks.

3.2. Proposed COTANS-NN Architecture

The COTANS-NN method re-purposes the 8-layer AlexNet architecture [12] by replacing the classification layer with a regression layer [21]. The inputs are COTANS images, which encapsulate the relevant information for boundary estimation. The outputs are boundary parameter estimates $[\hat{\rho}_1 \cdots \hat{\rho}_N \ \hat{\theta}_1 \cdots \hat{\theta}_N]^{\text{T}}$, each scaled to a dimensionless range of $[0,\ 1]$ by dividing each ρ by a $\rho_{\max}$ (10 m in our case), and each θ by 360°. We use the correct $[\rho_1 \cdots \rho_N\ \theta_1 \cdots \theta_N]^{\text{T}}$ for training, using MSE as the cost function. Thus, the output size is $2N$. Training images are generated by simulating scenarios with randomized $\mathbf{p}_{\text{e}}$ and $\{\mathbf{p}_{\text{r},i}\}$, as in Fig. 7.

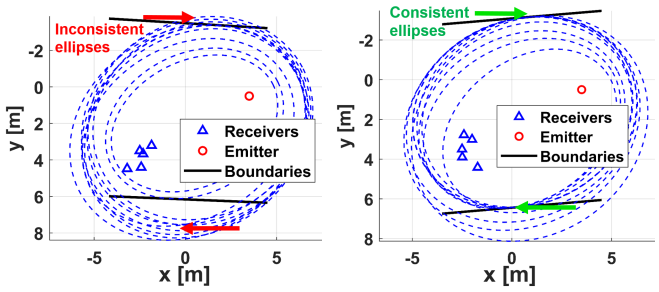
In this framework, the advantages of COTANS-NN relative to analytical methods become clearer. The least-squares (LS) algorithm [22], for example, assumes that the distance estimates are independently corrupted by Gaussian errors:

$$\hat{d}_{i,j} = d_{i,j} + \epsilon_{i,j}, \quad \epsilon_{i,j} \sim \mathcal{N}\left(0,\ \sigma^2\right). \tag{5}$$

Another example is the EDM algorithm [23], which minimizes the following s-stress cost function (sequentially and iteratively) over potential Cartesian coordinates $\widetilde{x}_j$ and $\widetilde{y}_j$ for the virtual emitters [23, Eq. S15]:

$$C\left(\widetilde{x}_j,\ \widetilde{y}_j\right) = \sum_i \left[\left(\widetilde{x}_j - x_i\right)^2 + \left(\widetilde{y}_j - y_i\right)^2 - \hat{d}_{i,j}^2\right]^2, \tag{6}$$

Authorized licensed use limited to: MIT Libraries. Downloaded on June 21,2023 at 14:34:32 UTC from IEEE Xplore. Restrictions apply.



(a) (b)

Fig. 7: Random training settings and their NLOS ellipses: dispersed at 2 dB SNR (a); and consistent at 20 dB SNR (b).

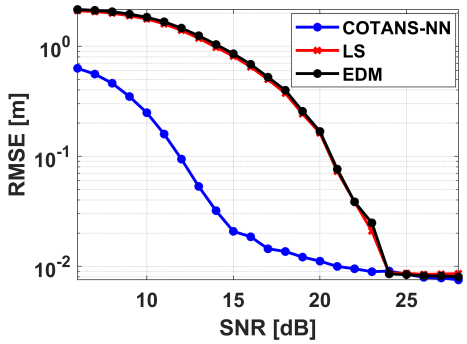


Fig. 8: Average range RMSE vs. SNR of COTANS-NN, LS and EDM. COTANS-NN is uniformly superior.

and is therefore also assuming (even if implicitly) statistically independent Gaussian errors. As mentioned earlier, this assumption is valid only at high SNR; furthermore, the echo labeling problem has to be solved as well to correctly group the $\{\hat{d}_{i,j}\}$ together for each j to achieve a reasonable estimate.

COTANS-NN, by contrast, is data-driven over a synthetic dataset of both small and large errors. Using COTANS images eliminates the need to solve the echo labeling problem altogether: the input to the NN is a transformation of all the unlabeled $\{\hat{d}_{i,j}\}$, rather than a sorted version of them.

4. SIMULATION RESULTS

We now present the performance of COTANS-NN in a simulation setting that is representative of the shallow-water underwater acoustic channel (i.e., with two boundaries), and compare it to the LS and EDM algorithms.

We train COTANS-NN on 14 SNR levels, equally spaced in the 10 to 20 dB SNR range (covering the transition region of global errors and the high-SNR regime), generating 50,000 training, 3,000 validation, and 50,000 test images per SNR. One boundary has its ρ- and θ-parameters as uniformly distributed random variables supported on the intervals $[3\,\text{m},\ 3.5\,\text{m}]$ and $[260^\circ,\ 280^\circ]$, respectively, and similarly for the other boundary on $[6\,\text{m},\ 6.5\,\text{m}]$ and $[80^\circ,\ 100^\circ]$. This yields a scenario where boundaries are known to be roughly at $(3\,\text{m},\ 270^\circ)$ and $(6.25\,\text{m},\ 90^\circ)$. The $\mathbf{p}_\text{e}$ and $\{\mathbf{p}_{\text{r},i}\}$ are drawn from a uniform distribution over 2 meter-wide areas centered on the points $(3.5,\ 0.5)$ m and $(-2.5,\ 3.5)$ m, respectively. Our performance metric is the range RMSE (in m) over all N boundaries and K environment realizations for each SNR S:

$$\rho_\text{RMSE}(S) \triangleq \sqrt{\frac{\sum_{j=1}^{N}\sum_{k=1}^{K}\left(\rho_{j,k}^{(S)} - \hat{\rho}_{j,k}^{(S)}\right)^2}{NK}}. \qquad (7)$$

While we have simulated three-boundary environments as well, and have also estimated θ, the resulting performance curves are qualitatively similar to the performance for ρ presented here. Hence, we only present the range estimation results for two-boundary environments.

We compare COTANS-NN to LS and EDM when these alternative methods are initialized with correct echo labeling. Note that unlike its competitors, COTANS-NN does not require such side-information, or an initialization at all, since it is by nature a non-iterative method. Fig. 8 shows the average range RMSE of the two boundaries vs. the SNR for the three methods. Evidently, COTANS-NN outperforms LS and EDM by a wide margin of up to 9 dB SNR, and maintains its stability despite global errors. When global errors arise, LS and EDM severely deteriorate, demonstrating that their 'small-error' assumption plays a vital role in their proper operation. In tougher error regimes, COTANS-NN is not merely robust but also continues to perform well.

LS and EDM perform very similarly, which arises from how they both minimize the squared error between measured and estimated distances, albeit with different optimization routines and cost functions. Since LS and EDM require echo labeling that in practice can also cause large errors, their performances will be worse than what is reported in Fig. 8. In contrast, COTANS-NN does not require echo labeling, and therefore cannot suffer from such performance deterioration.

5. CONCLUDING REMARKS

We propose the COTANS-NN method for 2D boundary estimation, exploiting the multiscale filtering capabilities of CNNs. Our method leverages a large training set covering SNR regimes with both large and small estimation errors to deliver robust performance, which is superior to the state-of-the-art alternatives that rely on high-SNR assumptions. COTANS-NN avoids the echo labeling and ad-hoc image filtering steps that further degrade the performance of alternative methods. It simplifies the use of domain knowledge to aid the boundary estimation task, and is capable of learning and using the *true* time-delay estimation error statistics.

Future work will focus on applying COTANS-NN to real-world data, with larger parameter spaces and model mismatch, to quantify its robustness. Derivation of theoretical results on boundary estimation will also help assess COTANS-NN's performance. While COTANS-NN was presented in a 2D setting, its operation can be extended to 3D. However, this may already be computationally non-trivial, and some algorithmic modifications could be required. Among other aspects, future work will also address this generalization.

Authorized licensed use limited to: MIT Libraries. Downloaded on June 21,2023 at 14:34:32 UTC from IEEE Xplore. Restrictions apply.

6. REFERENCES

[1] H. Niu, E. Ozanich, and P. Gerstoft, "Ship localization in Santa Barbara Channel using machine learning classifiers," *J. Acoust. Soc. Am.*, vol. 142, no. 5, pp. 455–460, 2017.

[2] H. Niu, Z. Gong, E. Ozanich, P. Gerstoft, H. Wang, and Z. Li, "Deep-learning source localization using multi-frequency magnitude-only data," *J. Acoust. Soc. Am.*, vol. 146, no. 1, pp. 211–222, 2019.

[3] Y. Wu, R. Ayyalasomayajula, M. J. Bianco, D. Bharadia, and P. Gerstoft, "Sslide: Sound source localization for indoors based on deep learning," in *ICASSP 2021-2021 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Jun 2021, pp. 4680–4684.

[4] T. Korhonen, "Acoustic localization using reverberation with virtual microphones," in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2008, pp. 211–223.

[5] F. Ribeiro, C. Zhang, D. A. Florêncio, and D. E. Ba, "Using reverberation to improve range and elevation discrimination for small array sound source localization," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 7, pp. 1781–1792, 2010.

[6] A. Brutti, M. Omologo, and P. Svaizer, "Multiple source localization based on acoustic map de-emphasis," *EURASIP J. Audio, Speech, Music Process.*, pp. 1–17, 2010.

[7] H. Naseri and V. Koivunen, "Cooperative simultaneous localization and mapping by exploiting multipath propagation," *IEEE Signal Process. Mag.*, vol. 65, no. 1, pp. 200–211, 2016.

[8] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, "Euclidean Distance Matrices: Essential theory, algorithms, and applications," *IEEE Signal Process. Mag.*, vol. 32, no. 6, pp. 12–30, Nov 2015.

[9] H. Naseri, M. Costa, and V. Koivunen, "Multipath-aided cooperative network localization using convex optimization," in *48th Asilomar Conf. Signals, Syst. Comput.*, 2014, pp. 1515–1520.

[10] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proc. IEEE*, vol. 97, no. 2, pp. 404–426, 2000.

[11] M. Crocco, A. Trucco, and A. D. Bue, "Uncalibrated 3D room geometry estimation from sound impulse responses," *J. Franklin Institute*, vol. 354, no. 18, pp. 8678–8709, 2017.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in neural inf. process. syst.*, 2012, vol. 25.

[13] F. Antonacci, A. Sarti, and S. Tubaro, "Geometric reconstruction of the environment from its response to multiple acoustic emissions," in *IEEE Int. Conf. on Acoust., Speech, Signal Process.*, 2010, pp. 2822–2825.

[14] F. Antonacci, J. Filos, M. R. Thomas, E. A. Habets, A. Sarti, P. A. Naylor, and S. Tubaro, "Inference of room geometry from acoustic impulse responses," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 10, pp. 2683–2695, 2012.

[15] S. Park and J. Choi, "Iterative echo labeling algorithm with convex hull expansion for room geometry estimation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, no. 3, pp. 1463–1478, 2021.

[16] Y. M. Too, M. Chitre, G. Barbastathis, and V. Pallayil, "Localizing snapping shrimp noise using a small-aperture array," *IEEE J. Oceanic Eng.*, vol. 44, no. 1, pp. 207–219, 2017.

[17] M. A. Chitre, "A high-frequency warm shallow water acoustic communications channel model and measurements," *J. Acoust. Soc. Am.*, vol. 122, no. 5, pp. 2580–2586, 2007.

[18] C. Cook, *Radar signals: An introduction to theory and application*, Elsevier, 2012.

[19] D. Dardari, C. Chong, and M. Win, "Improved lower bounds on time-of-arrival estimation error in realistic UWB channels," in *2006 IEEE Int. Conf. Ultra-Wideband*, 2006, pp. 531–537.

[20] D. Borrmann, J. Elseberg, K. Lingemann, and A. Nüchter, "The 3D Hough transform for plane detection in point clouds: A review and a new accumulator design," *3D Research*, vol. 2, no. 2, pp. 3, 2011.

[21] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in *Advances in neural inf. process. syst.*, 2013, vol. 26.

[22] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1121–113, 2004.

[23] I. Dokmanic, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli, "Acoustic echoes reveal room shape," *Proc. Nat. Acad. Sciences*, vol. 110, no. 30, pp. 12186–12191, 2013.

Authorized licensed use limited to: MIT Libraries. Downloaded on June 21,2023 at 14:34:32 UTC from IEEE Xplore. Restrictions apply.