

# A Review of Sound Source Localization in Robotics

Mofazzal Hossain\*

Department of Electrical Engineering, Cambridge University, Trinity Ln, UK

## Introduction

In the overall scheme of robot audition, sound source localization (SSL) in a robotic platform has been critical. It enables a robot to locate a sound source solely based on sound. It has a significant impact on other robot audition modules, such as source separation, and it improves human-robot interaction by augmenting the robot's perceptual abilities. The main goal of this review is to thoroughly map the reader's current state of the SSL field [1]. The goal of sound source localization (SSL) is to estimate the position of sound sources automatically. This functionality is useful in robotics in a variety of situations, such as locating a human speaker in a waiter-type task, a rescue scenario with no visual contact, or mapping an unknown acoustic environment. Because its estimations are frequently used in subsequent processing stages such as sound source separation, sound source classification, and automatic speech recognition, its performance has a significant impact on the rest of a robot audition system.

SSL in real-world scenarios must account for the possibility of multiple sound sources being active in the environment. As a result, estimating the position of multiple simultaneous sound sources is also required. Furthermore, both the robot and the sound source are refined by traditional techniques such as single direction-of-arrival (DOA) estimation, learning-based approaches (such as neural network and manifold learning), beamforming-based approaches, subspace methods, source clustering through time, and tracking techniques such as Kalman filters and particle filtering, the robotics community has significantly advanced SSL. Several aspects relevant to SSL in robots have become apparent while implementing these techniques onto robotics platforms, including: the number and type of microphones used, the number and mobility of sources, robustness against noise and reverberation, the type of array geometry to be employed, the type of robotic platforms to build upon, and so on [2].

Other mapping procedures exist besides grid-search. Their primary goal is to train the mapping function using recorded data from known sources. As a result, the learned mapping function implicitly encodes the propagation model. This type of mapping procedure is referred to as learning-based mapping in this survey. These are based on various training methodologies such as neural networks, locally-linear regression, and manifold learning, among others. Each mapping procedure is described in greater detail in the relevant branches.

## Description

Based on binaural techniques and multiple-microphone arrays, this work provides an overview of the robot audition field as a whole. The goal of this work is to conduct a literature review on SSL implementations on any type

of robot, such as service, rescue, swarm, industrial, and so on. We also look at efforts that are aimed at being implemented in a robotic platform but are not actually implemented in one. Furthermore, we examine resources for SSL training and evaluation, including some that were not gathered from a robotic standpoint but could be applied to a robotic task. Finally, we incorporate SSL research that uses only one microphone, which, while not applied in a robotic platform, we believe has an interesting potential for SSL robotics [3].

SSL adoption in robotics is still in its early stages. According to our knowledge, it began in 1989 with the robot Squirt, which was the first to have an SSL module. Squirt was a small robot with two competing behaviours: hiding in the dark and finding a sound source. Brook's own research team later investigated the idea of using SSL as a behaviour to drive interaction in a robot, which resulted in an SSL system for the Cog robot. Meanwhile, several Japanese researchers began to investigate the potential of SSL in a robot. Takanashi et al. investigated an anthropomorphic auditory system for a robot (as described by ) in 1993 [4].

During the 2000s, there was a significant shift in the research motivations in the robot audition field, specifically in SSL techniques. The motivation to imitate nature: using only two ears/microphones, cemented binaural audition. On the other hand, there was the incentive to improve performance (discussed in Section), which pushed for the use of more microphones. This paved the way for source localization techniques that employ a large number of sensors (such as MUSIC and beamformers) to perform SSL in a robot. As a result, the facets of the SSL problem were broadened, yielding a diverse range of solutions from the robot audition community [5].

## Conclusion

It is a common term for the feature set formed by the IPD and the ILD working together. This feature set is frequently used in conjunction with learning-based mapping. They are frequently extracted at the start to reduce the effect of reverberation. In practise, it has been demonstrated that temporal smoothing of this feature set makes the resulting mapping more resistant. A SSL mapping procedure should be able to map a given extracted feature to a location. A common approach is to use a propagation model directly, such as the freefield/far-field model or the Woodworth-Schlosberg spherical headmodel, both of which are discussed in Section. However, some features (particularly those used for multiple-source-location estimation) necessitate an exploration or optimization of the SSL solution space. A common approach is to perform a gridsearch, which involves applying a mapping function throughout the SSL space and recording the function output for each tested sound source location. This result in a solution spectrum with peaks (or local maximums) regarded as SSL solutions.

## References

1. Desai, Dhvani and Ninad Mehendale. "A review on sound source localization systems." *Arch Comput Methods Eng* (2022): 1-12.
2. Murray, John C., Harry R. Erwin and Stefan Wermter. "Robotic sound-source localisation architecture using cross-correlation and recurrent neural networks." *Neural Networks* 22 (2009): 173-189.
3. Chen, Guoliang and Yang Xu. "A sound source localization device based on rectangular pyramid structure for mobile robot." *J Sens* 2019 (2019).
4. Yalta, Nelson, Kazuhiro Nakadai and Tetsuya Ogata. "Sound source localization using deep learning models." *J Robot Mechatron* 29 (2017): 37-48.

\*Address for Correspondence: Mofazzal Hossain, Department of Electrical Engineering, Cambridge University, Trinity Ln, UK, E-mail: mofazzalhossain@gmail.com

Copyright: © 2022 Hossain M. This is an open-access article distributed under the terms of the creative commons attribution license which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Date of Submission: 02 June, 2022, Manuscript No: jees-22-80101; Editor Assigned: 04 June, 2022, PreQC No: P-80101; Reviewed: 16 June, 2022, QC No: Q-80101; Revised: 21 June, 2022, Manuscript No: R-80101; Published: 28 June, 2022, DOI: 10.37421/2332-0796.2022.11.27

5. King, E. A., A. Tatoglu, D. Iglesias and A. Matriss. "Audio-visual based non-line-of-sight sound source localization: A feasibility study." *Appl Acoust* 171 (2021): 107674.

**How to cite this article:** Hossain, Mofazzal. "A Review of Sound Source Localization in Robotics." *J Electr Electron Syst* 11 (2022): 27.