Demo: Microstructure-guided Spatial Sensing for Low-power IoT

Nakul Garg‡
nakul22@umd.edu
University of Maryland College Park

Yang Bai‡
yangbai8@umd.edu
University of Maryland College Park

Nirupam Roy
niruroy@umd.edu
University of Maryland College Park

(‡ Co-primary Student Authors)

ABSTRACT

This demonstration presents a working prototype of *Owlet*, an alternative design for spatial sensing of acoustic signals. To overcome the fundamental limitations in form-factor, power consumption, and hardware requirements with array-based techniques, *Owlet* explores wave's interaction with acoustic structures for sensing. By combining passive acoustic microstructures with microphones, we envision achieving the same functionalities as microphone and speaker arrays with less power consumption and in a smaller form factor. Our design uses a 3D-printed metamaterial structure over a microphone to introduce a carefully designed spatial signature to the recorded signal. *Owlet* prototype shows 3.6° median error in Direction-of-Arrival (DoA) estimation and 10 cm median error in source localization while using a $1.5cm \times 1.3cm$ acoustic structure for sensing.

CCS CONCEPTS

• **Computer systems organization → Embedded and cyber-physical systems**; **Sensors and actuators**.

KEYWORDS

Low-power sensing; IoT; Acoustic metamaterial; Spatial sensing

ACM Reference Format:
Nakul Garg, Yang Bai, and Nirupam Roy. 2021. Demo: Microstructure-guided Spatial Sensing for Low-power IoT. In *The 19th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '21), June 24–July 2, 2021, Virtual, WI, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3458864.3466906

1 OVERVIEW

Low-power design and miniaturization are inherent trends in ubiquitous computing. However, this trend is not apparent in acoustic sensing devices, particularly for the solutions that require spatial information on the sound. Traditionally spatial sensing requires a microphone array for spatial sensing, which demands significantly higher power and space. These requirements limit the application scenarios to those that can provide a larger device size and enough power supply. We explore acoustic microstructures to achieve high-resolution spatial sensing in a small form-factor and at a limited power budget.

MobiSys '21, June 24–July 2, 2021, Virtual, WI, USA

ACM ISBN 978-1-4503-8443-8/21/06…$15.00
https://doi.org/10.1145/3458864.3466906

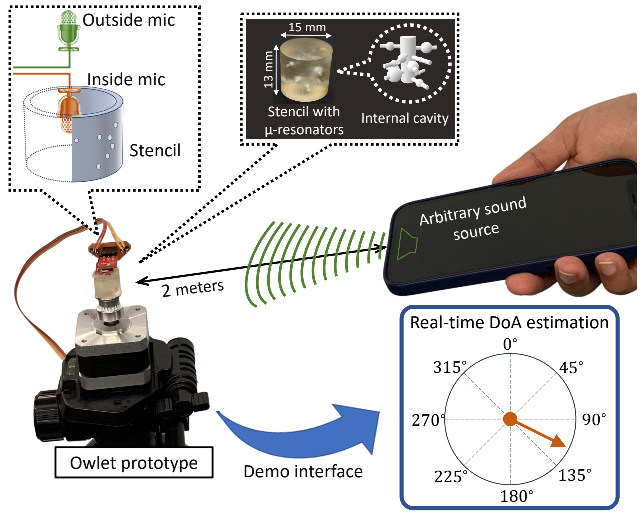

Figure 1: The demonstration of the real-time direction-of-arrival(DoA) estimation of arbitrary sound signals using a microstructure-based acoustic front-end. This technique achieves similar performance of a large linear microphone array, while using orders of magnitude less power and space.

In this demonstration we show the DoA estimation capability of the microstructure-guided sensing prototype, we call *Owlet*. Traditional techniques for obtaining spatial information of sound require multiple streams of simultaneously recorded sound using an array of microphones. Given the conventional DoA estimation algorithms fundamentally depends on this spatial sampling model, the dimension of the array and number of microphones used is crucial for their performance. According to the sampling theorem, a separation equal to the signal's half-wavelength ($\lambda/2$) between microphones in a linear array is considered ideal for DoA estimation. Moreover, the angular resolution (in terms of the inverse of the Half Power Beam-width) of the DoA is proportional to the total length of the array. Therefore, the traditional algorithms require a large hardware setup to achieve fine-grained resolution for DoA estimation. Moreover, arrays require multiple microphones sampled at the same time, which increases power consumption and hardware complexity proportional to the number of microphones. As shown in Figure 1, in this demo we seek to develop an alternative method for spatial signal processing. We break away from the spatio-temporal sampling model and explore the interaction of the waves with structures for a low-power, low-complexity, and miniaturized solution.

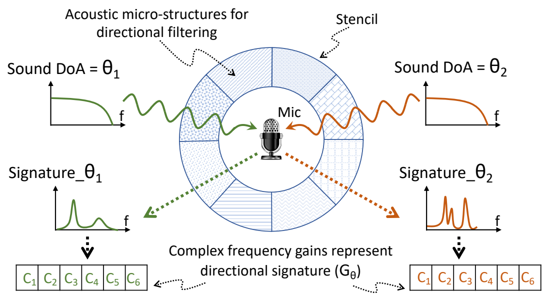

Figure 2: The concept of passive directional filtering using a stencil of acoustic microstructure. The stencil embeds a signature response to the recorded sound unique to its direction of arrival (DoA). The spectrum of complex gains represents the signature for further computation.

2 INTUITIONS AND SYSTEM DESIGN

Fundamentally, we aim to design a predictable environment around the microphone such that the recorded signal contains a unique 'direction-specific' channel impulse response. This impulse response can be extracted from the microphone recording and can serve as a signature of the sound's angle of arrival. While a regular room environment or larger objects near a microphone are known to create a diverse multipath effect to add reasonable direction-specific responses to the signal, we envision achieving a compact form factor for the system by combining the concepts of diffraction, interference, and structural resonance. To this end, we design a porous cap, called *stencil*, for the microphone with a particular hole pattern at different sides. As shown in Figure 2, sounds coming at a specific angle pass through the unique patterns of holes and combines at the microphone. The holes on the stencils are connected to microstructures of different parameters leading to unique frequency responses.

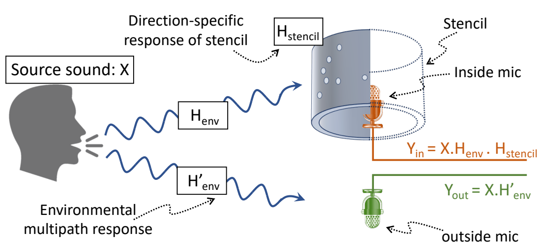

Figure 3: The two microphones model with a microstructure for DoA estimation. The result is comparable with that of 9-microphone array with the MUSIC algorithm.

The stencil forms a metamaterial with internal microstructures that naturally modulates incoming sound to introduce a signature response for an incoming angle. As the impact of the microstructures depends on frequencies, the signature response is basically a vector of complex gains, G_θ, of the frequency response. The concept is explained in Figure 2. Once we collect the frequency responses from different angles at the calibration stage, we train a deep learning model with variations of the signature table of frequency responses and use the pre-processed signal at run-time to get predicted DoA from the model.

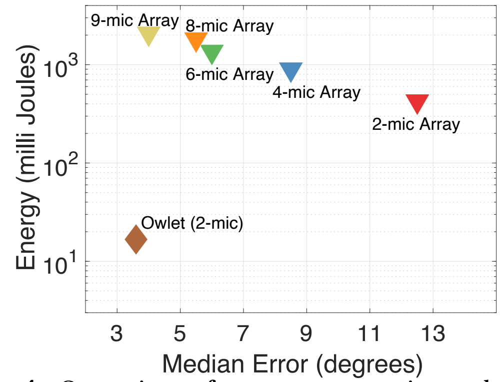

Figure 4: Comparison of energy consumption and median direction-of-arrival error for Owlet and traditional array based techniques.

The source signal can be known for some applications, like in navigations where the robot localizes itself by finding DoA of a known control signal. However, DoA estimation is useful in many other applications including localizing an ambient noise source or finding the user's direction from spoken words. In such scenarios, the source signal is unknown to the system, and it is difficult to separate $H_{stencil}$ from the arbitrary source signal. We eliminate this problem by introducing a secondary microphone to the system which is placed outside the stencil. The incoming sound is recorded simultaneously by these two microphones, but the outside microphone's recording is unaffected by $H_{stencil}$. Figure 3 shows the system in a realistic scenario considering the environmental responses along with the stencil response. An elaborate description of the technique is included in [1].

3 DEMONSTRATION

We demonstrate the feasibility of spatial sensing with the help of carefully designed microstructures. This shows the real-time DoA estimation of sound using the two microphones and a 3D-printed stencil setup of *Owlet*. As shown in Figure 4, *Owlet* can achieve even higher accuracy than a 9-element microphone array (3.6° vs 4° median error) with over 100 times less power consumption. *Owlet* requires 16.69mJ of power while the traditional array consumes over 2078.05mJ. Our demonstration is composed of a $1.5cm \times 1.3cm$ stencil, two microphones, a speaker, a smartphone, and a laptop. The speaker or the smartphone will play different kinds of sounds to the microphones placed inside and outside the stencil. We sample the microphones using a data acquisition system (DAQ) and perform all computations on Matlab. During the demonstration, the system uses the trained neural model for DoA prediction. The user can arbitrarily move the signal source (i.e., the smartphone) within a range of 2 meters from the *Owlet* sensor. The prototype estimates the direction of the sound and shows it on the laptop screen in real-time for visualization.

REFERENCES

[1] Garg, N., Bai, Y., and Roy, N. Owlet: Enabling spatial information in ubiquitous acoustic devices. In *The 19th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '21), June 24–July 2, 2021, Virtual, WI, USA*, ACM.