

Poster: Ultra-low-power Acoustic Imaging

Yang Bai[‡]
yangbai8@umd.edu
UMD, College Park

Nakul Garg[‡]
nakul22@umd.edu
UMD, College Park

Nirupam Roy
niruroy@umd.edu
UMD, College Park

([‡] Co-primary Student Authors)

ABSTRACT

This poster presents the design and implementation of *SPiDR*, an ultra-low-power acoustic imaging system. This imaging system produces a cross-sectional map of the field-of-view using only one speaker/microphone pair. It leverages the fact that sound’s interaction with small structures can project spatially coded signals on a region at a fine granularity. We create a 3D-printed passive filter, called a stencil, that can image the scene with a single omnidirectional source and sensor. With spatially coded signal, the system receives a linear combination of the reflections from nearby objects and applies a novel power-aware depth-map reconstruction algorithm. *SPiDR* consumes only $10mW$ of power to generate a depth-map in real-world scenario with over 80% structural similarity score with the scene.

1 OVERVIEW

Scene imaging for understanding the environment for navigation is crucial for robotics. Widely adopted methods have been developed for scene perception purpose. Depth camera is one of the most effective approach. Other strategies include scanning the surroundings and generating a depth-map of the scene using lidar, radar, or sonar based techniques. However, these techniques require using mechanical maneuver or electronic beam-steering using an array of sensors – leading to higher energy requirements. Therefore, despite tremendous advancement in engineering, these techniques are not directly applicable in micro-robotic systems for their unique constraints of limited energy source, small size, limited computational power, and the requirement of low-cost manufacturing (SWaP-C constraints). A survey indicates that only $1030mW$ of power remains available for sensing in typical micro-robot systems after allocating power to the actuators for locomotion. Needless to say, these systems contain significantly limited computational resources.

In this poster, we present *SPiDR*¹, an ultra-low-power acoustic spatial sensing system capable of generating an accurate depth-map of nearby objects, with only one pair of speaker/microphone. Also, we do not consider mechanical steering for scanning the scene with one sensor. We achieve

¹SPiDR stands for Structure-assisted Perception, Detection, and Ranging

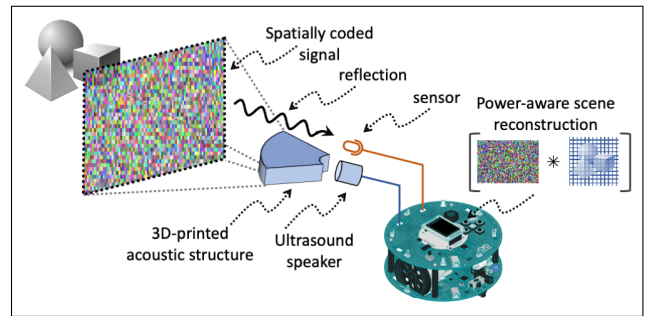


Figure 1: *SPiDR*, an ultra-low-power acoustic spatial sensing system for mobile robots. The system requires a single transducer for sensing and uses a 3D-printed microstructure for projecting spatially coded signals.

this by using a spatially coded signal that sends unique patterns of signal in each direction. Figure 1 gives an overview of the system. We design a 3D-printed cover for the speaker, called a stencil, that divides the speaker’s output into multiple replicas by passing it through small internal tubes. These tubes have different lengths, thus induce diverse phase delays and therefore their superimposition leads to a specific amplitude and phase of the received signal. A unique combination of the path lengths can produce a unique coded pattern in 3D space. The lengths of these internal tubes are carefully calculated to channelize the signals through different time-delayed paths before releasing them through separate output sound holes pointed at different spatial directions. When the signals encounter the edge of the holes, diffraction happens to transmit each replicas to a wide angle. These delayed and diffracted replicas interfere with each other and create complex but predictable patterns at different points of the scene. With this specific code to the 3D points in space, occupied voxels can be separated through processing to convert them to a point cloud of objects in the scene. Our previous work Owllet [2] shows the possibility of using acoustic microstructure to embed directional clues to the signal recorded by a microphone. Owllet detects the angle of arrival of incoming sounds, while *SPiDR* captures a cross-sectional depth image of the scene. Refer to our MobiSys 2022 paper [1] for a detailed system description and evaluation.

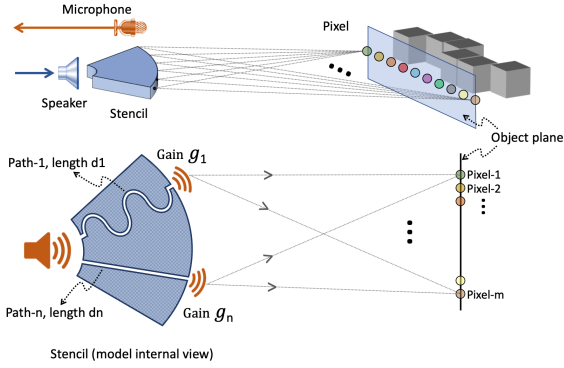


Figure 2: Diversity projection with the stencil with the internal channel to encode unique gains to signals.

2 INTUITIONS AND SYSTEM DESIGN

We aim to design an acoustic structure to produce distinct channels for sound propagation with desired delays and attenuations. With only one pair of speaker/microphone, the signal propagated in space is encoded by location, so that when the signal is reflected by an object in a certain location, as shown in Figure 2, the reflected signals bear the locations of the objects, and can be reconstructed as an image. The stencil is a 3D-printed porous cap that covers the speaker and channelizes the output signal through a number of internal tubes connected to the openings pointed in different angular directions, as shown in Figure 3. The size and length of the tubular paths vary to control the amplitude and relative phase of the signals at the opening. Moreover, diffraction happens when acoustic signal arrives at the edge of the tubes, further increasing the angle of propagation, thus inducing more diversity to the received signals in space. Figure 4 shows that the stencil spreads the signal energy in a wide region.

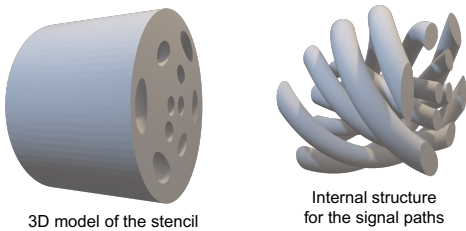


Figure 3: (left) The 3D design of the stencil and (right) the internal structure showing the tubular helical paths.

We discretize the scene to a collection of pixels on a 2D scene we are interested in reconstructing. The probing signal only reflects off the pixels that represent an object and combines linearly at the only microphone used for sensing. The received signal y_{rcv} can be formulated as $y_{rcv} = Hx$. Here H is the collection of ideal reflected signals from

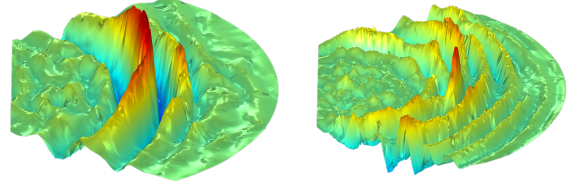


Figure 4: Ultrasound emitted from the speaker without (left) and with the stencil (right). The stencil spreads the signal energy over the region of interest.

each individual pixel organized in columns. The vector x represents the reflectivity status of the pixels indicating the fraction of the ideal signal reflected from each pixel. The values of x are higher when the object at the corresponding location is a good reflector of the signal and a zero value indicates the absence of object at that pixel. In other words, the vector x selects columns of H to map a scene to the weighted sum of reflections y_{rcv} . Our scene reconstruction algorithm aims to recover x from the received signal y_{rcv} .

3 PRELIMINARY EVALUATION

In the *SPiDR* prototype, we embed an ultrasound speaker in a 3D-printed stencil and a microphone placed on the top of the stencil. We also compare the performance and the power consumption with lidar and ultrasound distance sensor. The sound sources are ultrasound 10-cycle tone burst signals with frequencies $38 - 42kHz$ with $1kHz$ apart, with signal strength 40 dB SPL. The size of the stencil is $5 \times 3 \times 2cm$ with internal tubes. We show the images of 4 representative scenes in Figure 5. The real-world scenes are shown in the first row, and the estimated scenes are shown in the second row. We observe that the scenes can be detected with *SPiDR* when up to 70% of the scene is occupied, where the width of the scene is $20cm$. Moreover, *SPiDR* can produce similar accuracy as lidar but with $400\times$ less power consumption.

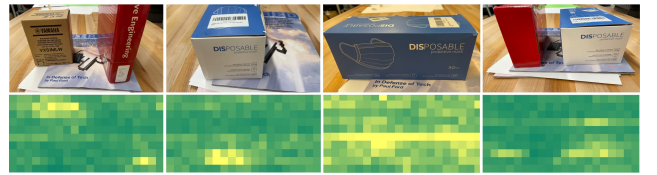


Figure 5: Depth-map reconstruction using *SPiDR* for various real-world scenes.

REFERENCES

- [1] Yang Bai, Nakul Garg, and Nirupam Roy. Spidr: Ultra-low-power acoustic spatial sensing for micro-robot navigation. In *The 20th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '22)*. ACM, 2022.
- [2] Nakul Garg, Yang Bai, and Nirupam Roy. Microstructure-guided spatial sensing for low-power iot. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pages 503–504, 2021.