MEMS Accelerometer based Acoustic Real-time Event Sensors

David Dall'Osto, Ph.D.
Applied Physics Laboratory at the University of Washington – Seattle
1013 NE 40th St, Seattle WA 98105
dallosto@uw.edu
(206) 221-5085

Abstract

A new system, the Acoustic Real-time Event Sensors (ARES), has been developed for autonomously monitoring the directional sound field in urban, transportation and industrial settings, managed within an IoT network. The directional information provided by each sensor enables a fully unattended sparse acoustic monitoring network that can discriminate specific sources of sound from overall noise levels at a measurement location, even if that noise is very close in proximity. While possible to make a network with existing 3D microphone-based acoustic intensity or camera solutions, these options are unwieldy and expensive for real-time deployments. The ARES sensor takes a different approach, measuring the velocity of a small parcel of air surrounding a triaxial accelerometer, from which a vector-based representation of sound intensity is calculated. Each ARES sensor node is compact directional sensor, sensitive to low frequencies with acoustic wavelengths much greater the ARES sensor). Sensitivity to the acoustic field is enabled by using a lightweight MEMS accelerometer paired with a MEMS microphone, both housed within an ultra-lightweight polystyrene sphere. Calibrations are applied to create gap free short time-spectral data stream, which are sampled by a small low power Raspberry Pi (RPi) computer that can manage the data from multiple nodes through a wired CAN bus supporting distances up to 100 m. Each node is synchronized to GNSS time facilitating collaboration among sensors surrounding and/or within a local measurement site. Low power ARES nodes are location aware and estimate azimuth and elevation angles to detected noise sources as a function of frequency. Multiple ARES are combined either in a beamforming capacity with closely spaced nodes, or as a sparse network to triangulate sound source. In the sparse network operation, more widely separated nodes are joined through WWLAN technologies via their local RPi hub. ARES represents an enabling technology that extends the capabilities of sensor network to detect airborne sources that often elude detection on existing sensors networks. Outdoor measurements of impulsive, tonal and broadband noise signatures on a small ARES network in urban settings were collected to demonstrate this enabling technology. This includes examples covering: (1) beamforming on multiple ARES nodes to focus the network for more precise monitoring and identification of noise sources; (2) recreation of calibrated directional audio waveforms from the complex time-frequency spectral data; and (3) and impulsive and continuous noise source trilateration and tracking on a sparse network of sensors.

DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.

1. INTRODUCTION

Passive acoustic noise monitoring is ideally associated with the direction and location of detected acoustic sources. While possible with an omnidirectional acoustic sensor, adding direction information at each sensor improves confidence in locating sources. Microphone arrays, acoustic cameras, and phased array technologies use beamforming or acoustic intensity methods to determine direction, but these solutions occupy space, such that the physical size of the array depends on what frequencies are to be detected. Urban transportation and construction noise occupies fairly low frequencies, and for reasonable spatial resolution the array element separations scale with wavelength, making low-frequency (100 Hz) arrays physically large.

A more environmentally robust alternative to microphone arrays has been developed, using lightweight MEMS microphone and inertial MEMS accelerometers embedded in a lightweight expanded-polystyrene foam sphere to directly sense the 3D acoustic intensity vector. By housing the two sensing elements within the foam, the microphone and inertial have the same reference field point simplifying any interpretation of the measurements and combination. This spherical sensor package has a density roughly 6X that of air. In addition to the sensitivity gained and weather protection offered by the foam enclosure, the microphone placement within provides additional resilience to the sensing orifice of the microphone protected from the external air-turbulence. This makes this technology ideal for outdoor applications and promising for directional sensing on moving platforms.

The accelerometer based acoustic vector sensor (AVS) operates under the principle that it measures the motion of the small parcel of air the device displaces. With acceleration is measured, it is integrated and scaled by the sensitivity factor to provide sensor output as acoustic particle velocity. When combined with a MEMS microphone pressure measurement at the same point in space (a few millimeters apart), a full 3-axis representation of the sound intensity is produced. The main achievements during the sensor design phase were to improve manufacturability, implement the framework and support system, and refine calibration methods to achieve the desired accuracy.

The suspended sensors are exposed to sound fields in three dimensions, which are sampled, and from which calculations are performed locally to scale the data and convert measured acceleration and pressure to acoustic particle velocity and intensity. A micromesh windscreen enclosing the sensor is water repellent, but not resistant. Rainwater and moisture can permeate through the micromesh windscreen and collect on the foam solid body that encloses the sensor board. It is thus convenient that in the accelerometer-based AVS design, both the accelerometer and microphone are encased within the closed-cell foam, which provides good protection from the weather.

In this paper, we first explain the construction of the Acoustic Vector Sensor (AVS). Distinct from delay and sum acoustic arrays, the accelerometer-based sensor formulation of the active intensity vector provides a direct directional mapping of the sound field. Next, we examine the processing of the 4-channel output (3 acceleration and 1 pressure) which is integrated into Acoustic Real-time Event Sensor (ARES) node that synchronizes data collection to a common time-base. Finally, we examine the data collected on ARES100 product (target sensitivity: 100 Hz – 2 kHz) for applications including: directional hearing within a car, source localization of LF impulses, and remote sensing of ground and air-traffic. We conclude with a short discussion and potential market solutions.

2. AVS CONSTRUCTION

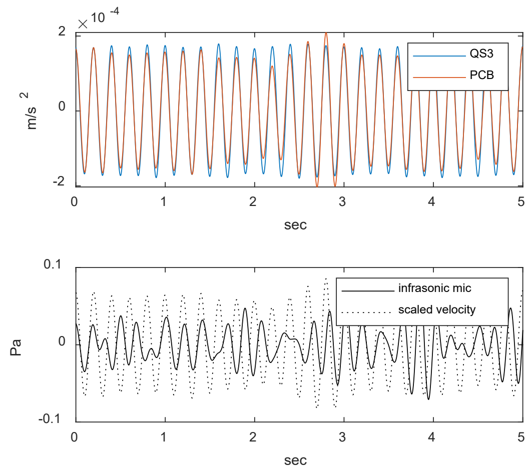
One overarching principle in interpreting the acoustic vector field is the relation between acoustic pressure and particle velocity by the characteristic impedance - in *air* $_{that}$ is 1.2 kg/m^3 x 340 m/s ≈ 400 Rayls (units named after the great 19th century acoustician Lord Rayleigh aka John William Strutt). In the far-field well away from reflecting surfaces, the acoustic field can be described locally as a single progressive plane-wave and the magnitude of pressure scales with the magnitude of velocity as $|v|/|p|$ = 400 Rayls. Arrays are built to interpret this feature of the wave-field, isolating sound in a particular direction by combining the receptions (in particular the phase differences) over an aperture.

One limitation in array processing is the inherent noise affecting low-frequency measurements. As we will see in the next section, this is a problem when arrays have an element spacing much less than a wavelength. To solve this issue requires larger arrays or a direct measure of acoustic velocity with an acoustic velocity sensor (AVS). An AVS + P sensor (acoustic velocity + pressure measurement) has a unique output, where multiplication of pressure (dynamic field) and acoustic velocity (kinematic field) form a second order vector called acoustic intensity. Note, a tetrahedral microphone array generates a similar measure of intensity via finite differencing, a technique forming the basis of many sound power measurement systems. These finite difference methods, like the array-processing limit, require significant aperture at low frequencies.

Accelerometers performance as AVS sensors also suffer at low frequencies. For example, a typical linear accelerometer will measure acceleration levels equally at all frequencies, and thus when integrating acceleration to form the particle velocity, the electronic noise floor scales as 1/f. The sensitivity of the accelerometer to the acoustic particle motion depends on the relative sensor density. The response (relative to dB incident levels) follows a straightforward relationship, relating the density of the sensor package to the air. While maximum sensitivity gained from true neutral buoyancy, this is difficult to achieve in air requiring lighter than air rigid materials. While at some cost in sensitivity, a slightly denser-than-air sensor has some advantages, including sensor robustness, facilitation suspension design and avoiding the complicated sensor response of a truly buoyant in dynamic atmospheric conditions.

Through selection of a lightweight high-sensitivity accelerometers and extremely lightweight foam, we created the ARES100 sensor with a density 6X the density of air, which achieved a reasonable response to low frequency acoustic signals. The fundamental limitation is the noise floor of the sensors. Two combinations have been developed under this project: ARES100 and ARES1. While both are sensitive to 100 Hz, only the most sophisticated (expensive) accelerometers have a noise floor low enough to reliably measure acoustic frequencies down to 1 Hz.

The ARES1 uses an ultra-low noise Sercel QS3 seismic sensor, an imbedded circuit board that implements the optical sensing feedback-loop to improve to “nano-g” sensitivity (1 billionth the acceleration due to gravity). Figure 1 shows a comparison of the QS3 to a more traditional PCB piezoelectric accelerometer, weight 21 g vs. 6 g, respectively, embedded in the ARES1 foam hemispheres. Both accelerometers show the same levels their linear response to low frequencies.



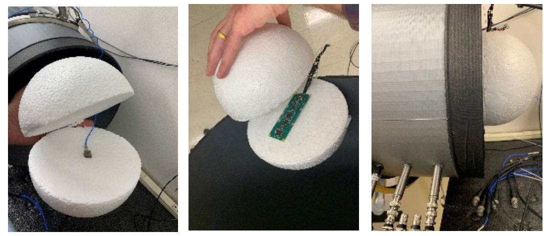


Figure 1. [Top] a comparison of the two accelerometers in embedded in the 8" foam sphere to a 5 Hz sine wave. [Bottom] The corresponding pressure within the tube, showing complexity due low signal levels owing to the reactive nature of the infrasonic field.

Examining the output of the infrasonic ARES1 sensor, numerical integration of the acceleration provides a reliable estimate of the sensor velocity (see Figure 2). The velocity scales with the circular frequency ω, which for the 3 Hz signal is a factor a 6π and for a 1 Hz signal is 2π. Note that numerical integration is subject to noise, and can be better limited by providing the integration in the spectral domain by multiply the acceleration spectra by 1/ω.

The sensing components are interfaced with the ARES processor (establishing an ARES node) which applies calibrations to convert the signal to engineering units and compute the vector intensity. The vector products are passed as complex spectra in real-time, or derivative distillations to generate an event sequence with features (as opposed to continuous acoustic data). Modified complex pressure spectra can feed an audio reproduction presented to a listener, or the acoustic features tabulated over a sensing network can identify and geo-locate sources or sound. Next, we will examine the processing outputs from the ARES100 through a series of examples.

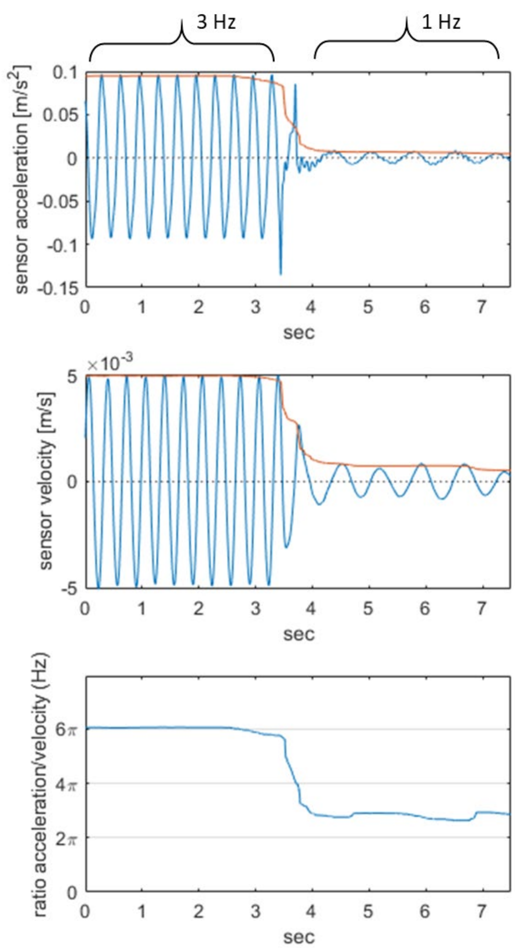


Figure 2. Infrasonic accelerations (top), integrated to velocity (middle) and the ratio (bottom), measured in the acoustic wave tube at frequencies f = 3 Hz, followed by f = 1 Hz. The velocity signal scales with the circular frequency, ω= 2π f.

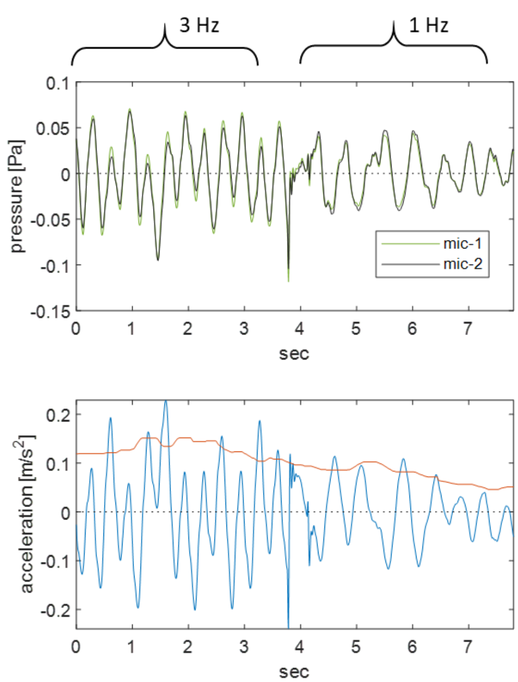
3. ARES PROCESSING

The first layer of processing, post calibration, is to combine the 4-channel data to form the complex intensity vector. The real part of this quantity is the active intensity, which points in the direction of sound flow. While this processing is possible with two microphones, via a finite difference approximation to the acceleration, this conversion is frequency dependent and very sensitive to the phase accuracy in the pressure sensors. In terms of array processing, the direction provided by the real part of the complex intensity vector is the maximum response axis of beam formed data conveniently provided with a simple arctangent computation.

Now we demonstrate the inability for short microphone arrays to accurately measure low-frequencies, the example here is using spatial separation of microphones to approximate the velocity via a finite-difference. The difference between these two follow the relationship,

$$a = -\frac{1}{\rho_o}\nabla p,$$

where the gradient of pressure (∇p) along the axis of two microphones can be approximated by subtracting the two signals and dividing by the separation. The same 3-Hz infrasonic signal shown in Figure 2 is recorded by a pair of microphones separated by 4 inches (equivalent to the radius of the ARES1 sensor). While each microphone produces reliable 3-Hz waveforms individually, the measurements must be sensitive to less than 1/3 degree of phase-shift. As apparent in the comparison in Figure 3, at low frequencies the finite difference method is plagued by noise in this approximation.



Alternatively, with the 4-channel ARES sensor particle velocity and pressure data are combined directly via multiplication of time-series or cross-spectra. As part of the processing stream, the raw acceleration and microphone data are first corrected spectrally within the ARES node from calibration values. The complex spectra are passed downstream for processing off the ARES node. As the phase information is retained in the complex spectral data stream, we can recreate the now calibrated timeseries waveforms via an inverse short-time Fourier transform. We will examine this recreation further in the next section addressing applications.

Figure 3. Infrasonic pressure (top) and particle accelerations estimated by finite difference between the two microphones, at f = 3 Hz followed by f = 1 Hz. Compared to the direct measurements in Fig. 3, acceleration approximation via finite difference of microphones suffers from noise contamination.

The base processing of the ARES node is combines the calibrated spectra of the 3 orthogonal particle velocity components with the complex pressure spectrum, forming the complex intensity vector,

$$\vec{I}_c = p \times \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}^* = \begin{bmatrix} I_x + iQ_x \\ I_y + iQ_y \\ I_z + iQ_z \end{bmatrix},$$

for which the real part is the active (propagating) intensity and the imaginary part is reactive (local) intensity. With a sensor close to a reflecting surface, we expect a degree of reactive intensity due to multipath interference with reflections, and to a lesser degree when the sensor is in close proximity to a sound source (i.e., within its near field). For most sensing applications, the sensor will within a few wavelengths of a horizontal reflecting plane (e.g., the ground) and thus a straightforward interpretation of elevation angle becomes complicated by the multipath interference the sound and its ground reflection arrive simultaneously. The angle in the horizontal plane, the azimuth, is computed from the tangent of the two horizontal components (I_x and I_y, assuming a vertical orientation of I_z) and for outdoor applications give a good sense of direction of the sound propagation.

From the spectra served by the ARES node, we can construct a time-frequency display of sound power (spectrogram) and one for the azimuthal direction (azigram). The azigram provides another feature to pick out different sound sources. The precision of the direction estimates depends on the source type, and importantly the processing applied to extract this data. A Short-Time Fourier Transform (STFT) continuously calculated on the ARES node in 40 ms windows updates every 20 ms. While sufficient to pick out azimuth of impulsive noises, owing to the 25 Hz resolution prescribed by the STFT, the corresponding azigram does not identify low-level tonal signals out of the ambient noise. However, we can reprocess the data by constructing an inverse STFT, stacking and concatenating the reconstructed waveforms to recreate the time-series. From this time-series, we can resample the spectrum with higher resolution (longer window) Fourier transform, or narrowband filter and track tonal data.

To illustrate the process, an ARES100 sensor node located on the rooftop of the Applied Physics Laboratory in Seattle, WA (see figure 4). We examine two recordings: (1) impulsive noise from 4th of July fireworks display on the nearby Lake Union for impulse detection in Figure 5, and (2) tonal noise of a propeller driven floatplane flying overhead departing from Lake Union in Figure 6.

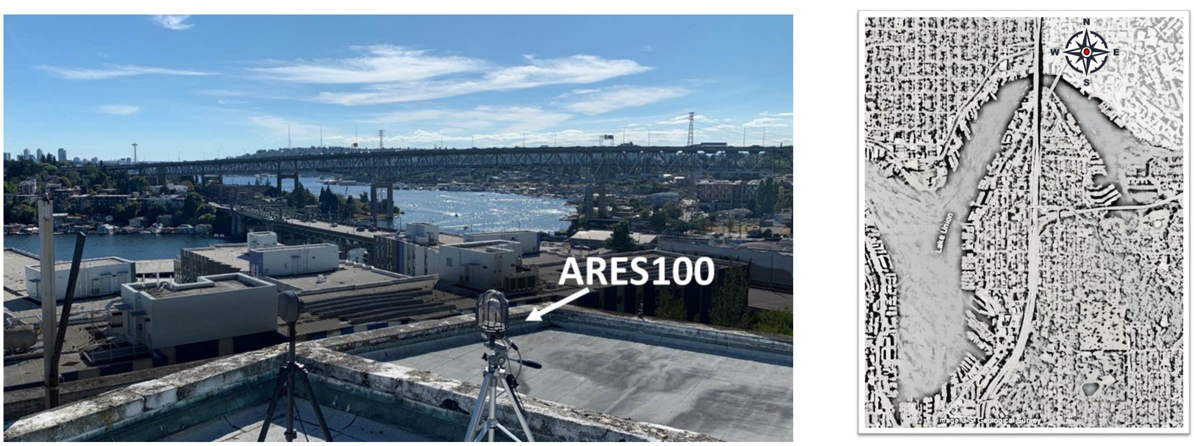


Figure 4. [left] The ARES100 node on top of the APL-UW rooftop, overlooking Lake Union to the Southwest. [right] Map of the sensor location (red dot on the compass rose). The I-5 corridor and Lake Union waterways are to the West of the sensor.

The base processing calculates the azimuthal direction (bearing) of intensity, forming a time-frequency-bearing direction history (azigram) equivalent to the time-frequency-intensity (spectrogram). Over an event-window (user set) the intensity azimuth is binned by angle (i.e., histogram over-bearing angle) weighted by their spectral level to emphasize the direction of loud sounds; a threshold is set to premask the azigram prior to converting to the histogram to remove spectral noise. To improve signal to noise, the output in Figures 5 and 6 show both the Native 40 ms STFT result and reprocessed data, first converted from the STFT to a time series and then recalculating the spectrum content over a longer sliding time-window to improve the frequency resolution.

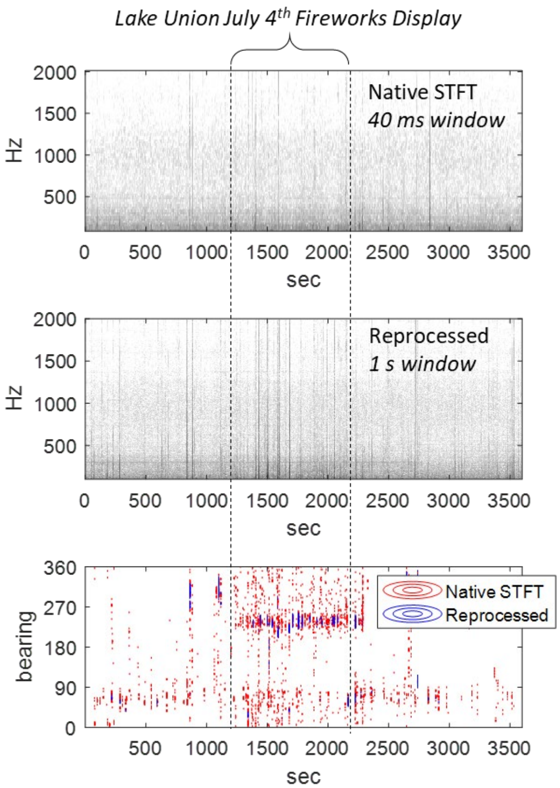


Figure 5: July 4th 10:00-11:00 pm local time record from the ARES100 on the APL roof. [top] The Native STFT intensity spectrum [middle] reprocessed data from reconstructed timeseries using a longer window for more frequency resolution. [bottom] The directional output, clearly showing the impulse signals ~10-minute fireworks display emanating from the South West.

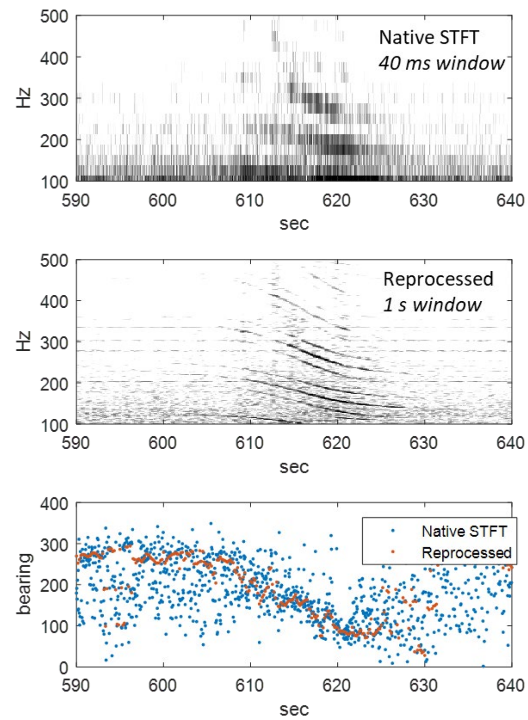


Figure 6: June 28th recording from the ARES100 on the APL of a float plane flying by. [top] The Native STFT intensity spectrum [middle] reprocessed data from reconstructed timeseries using a 1 second windows to pick out the Doppler shifting tonals of the propeller noise. [bottom] The directional output tracking the path of the airplane, with clear enhancement using the high-resolution spectrum.

4. ARES100 APPLICATIONS

Since development of the ARES100, we have identified numerous applications in addition to the directional impulse and tonal tracking demonstrated in the previous section. As the ARES nodes are GNSS synchronized to the same time-base, the absolute time-of-flight is possible on a network of sensors. Source location via trilateration improves with the directional information provided by each node, acting as a no-computational cost filter that can assist in algorithms in associating detections on multiple nodes to the correct events.

This same directional information has great potential in providing input for augmented reality, feeding directional acoustic input from the environment. One example is improved communication and awareness in noisy environments, as well as enhancement of sound emanating from a particular direction. To examine this, we performed an experiment to demonstrate how the ARES100 can identify the location of "speakers" within the cabin of a moving car. Voices generate a tonal aspect, so we reprocess the data with a 100 ms window to discriminate the frequencies. We then can compute the azigram, and only accept sound from a particular direction.

Additional noise reduction is possible, through masking the spectral content. To pick out coherent signals (sources), we mask the signals with a polarization filter – a time-averaged measure of the signal coherence in the complex domain. From this, we establish a threshold to accept or reject the data point using a time-frequency mask. This process has potential for a near real-time feed for augmented reality to focus attention on speakers. Figure 7 demonstrates the effective seperation of two sources and the isolated audio signals reproduction, based on the masking the spectrum by the directional information.

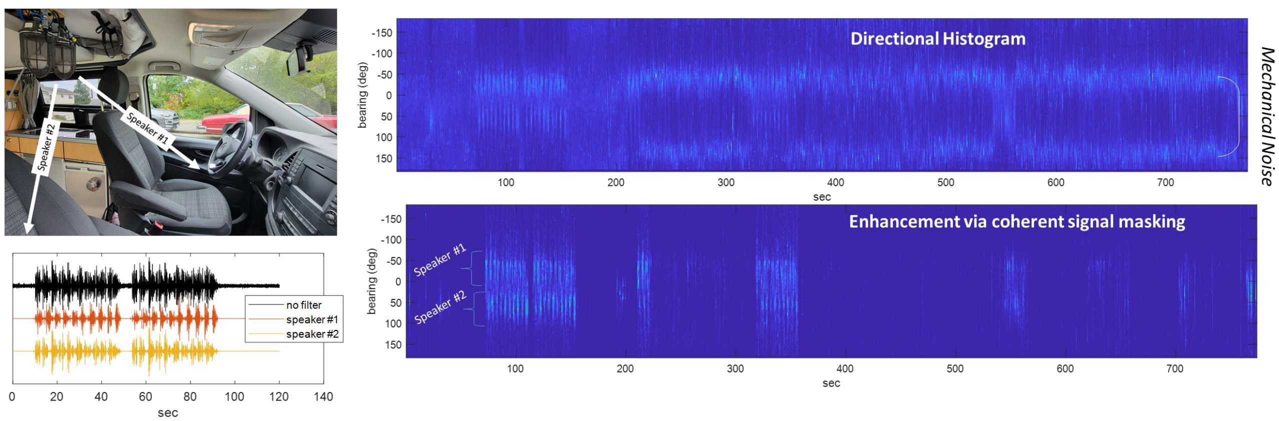


Figure 7. Example of the directional output in a moving vehicle (top right). The speaker #1 and #2 are located on the driver and passenger side, separated by roughly 90 degrees relative to the ARES100 node (top left). A signal polarization "filter-mask" is applied to remove the incoherent mechanical noise (bottom right) and an additional directional filter is applied to isolate the individual speaker audio signal (bottom left).

5. CONCLUSION

The ARES technology as developed has potential to solve many problems, including remote tracking of acoustic sources (including manned and unmanned aircraft), environmental characterization and awareness to identify and localize noise sources, and improved performance in directional sensing in adverse weather conditions. The path to development continues to extend the ARES sensing capability to lower frequencies with ultra-low noise accelerometers, and to higher frequencies with an eye on miniaturizing the device and cutting excess weight. The real-time data feed can be reprocessed to improve the directional focus of the sound field and has potential in augmented reality. Future development will investigate a wearable sensor, with processing that can provide enhanced audio output to a human user (or raw directional information to feed speech and processing algorithms) to improve audio and speech perception in challenging acoustic environments.

6. ACKNOWLEGMENTS

Product development was funded by the ARMY SBIR Phase 2 agreement W911QX21C0020.

Prior Work: U.S. Provisional Patent Application No. 63/335,879, entitled "ACCELEROMETER-BASED INTENSITY VECTOR SENSOR WITH COLLOCATED MEMS MICROPHONE" filed on April 28, 2022
U.S. Patent 11796562, entitled "ACOUSTIC INTENSITY SENSOR USING A MEMS TRIAXIAL ACCELEROMETER AND MEMS MICROPHONES" issued on Oct 24, 2023