INTERNATIONAL STUDENT CHALLENGE PROBLEM IN ACOUSTIC SIGNAL PROCESSING 2023

Jacob Brotman-Krass, Sochima Omenkeukwu, Jenna Rutowski, Augustus Standeven, and Isaac Wang

Dept. of Electrical & Computer Engineering, University of Rochester

Abstract

The International Student Challenge Problem in Acoustic Signal Processing 2023 involves processing acoustic sensor data from a set of hydrophones. The goal is to extract information about the source from its corresponding acoustical data. In this case, a diver swims at a constant speed and depth while passing three evenly spaced hydrophones that are an equal height above the seafloor. The main signal processing tasks were to display the spectrogram of the hydrophones, estimate the breathing rate of the diver, estimate when the diver is closest to the middle hydrophone O, estimate the diver's altitude in relation to the hydrophone array, and estimate the diver's swimming speed [1].

I. Introduction

The three hydrophones, N, O, and P, are located 1m above the seafloor at a depth of 20 m. The hydrophones are each separated by 14m along a straight line. Hydrophone O is located at the center of the array. Each hydrophone is sampled at 250 kHz. The acoustic signature of the diver consists of periodic pulses of white noise which correspond to the diver inhaling.

II. Task 1a

This specific task involved displaying the output spectrogram for the middle hydrophone O and commenting on the spectral properties. A spectrogram is a visual representation of the frequencies which make up a signal as it varies with time. By analyzing the spectrogram of the hydrophone data, it is possible to analyze and identify frequency components present in the signal and observe how they vary with time.

In terms of the scuba diver's acoustic signature, the spectrogram helps to display characteristic frequency components associated with the sound of the diver's breathing and movement. The noise generated by the scuba gear produces a characteristic spectral pattern that is distinguishable from other noise sources.

By analyzing the spectral properties of the acoustic signature, it is possible to identify the sound source, estimate its location, and extract other information about it, such as its movement. In this case, this periodic noise pattern correlates to the breathing rate of the diver.

Our MATLAB [2] code loads the audio file named "SCP23_Hyd O.wav", which contains the recording captured by hydrophone O. The audio file is read using the built in audioread() function, and the hydrophone O channel data and sample rate are extracted. Next, a spectrogram is generated using the built in spectrogram() function. The spectrogram is a 2D representation of the audio signal's frequency content over time. It is created by dividing the audio signal into small, overlapping time segments and computing the frequency spectrum for each section. The resulting spectrogram plots time on the x-axis and frequency on the y-axis. Color intensity of each point represents the magnitude of the frequency content at that specific instance in time. The parameters used for generating the spectrogram are a window size of 1024 and a hop size of 512. The hamming window function is used to reduce spectral leakage and smearing in the frequency domain caused by windowing. Finally, the spectrogram is converted to a logarithmic scale in decibels (dB) using the 20*log10 function. The spectrogram is plotted using the imagesc function. The plot is displayed with the title "Spectrogram of Hydrophone O", and the x- and y-axes labeled with "Time [s]" and "Frequency [Hz]," respectively.

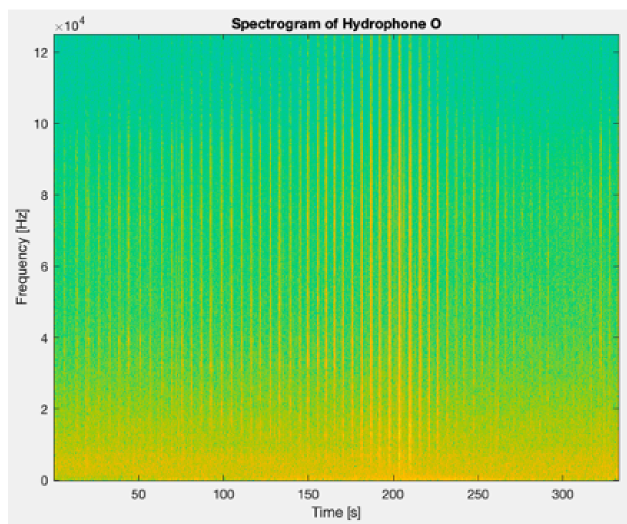

Figure 1: *Spectrogram of Hydrophone O*

Based on the spectrogram of the hydrophone O as seen in Figure 1, we can observe that the scuba diver's acoustic signature has several spectral properties. First, we can see that there are several peaks in the frequency domain that correspond to the various sounds made by the diver. In particular, there is periodic broadband noise corresponding to the diver's breathing rate. This indicates that, barring computational difficulties, a small window size and sizable overlap will give an optimal depiction of the diver's breathing, which requires little spectral precision but significant time resolution. The intensity of this noise appears to change over time, indicating that the diver seems to swim toward and then away from this particular hydrophone.

III. Task 1b

To estimate the breathing rate of the diver from the spectrogram, we need to look for spectral peaks that correspond to the breathing sounds. First, we can visually review the spectrogram to identify the breathing sounds. From the spectrogram, several vertical bands of energy that correspond to the diver's breathing are observed. To find the breathing rate, it is possible to find the time between these peaks and calculate the corresponding rate in Hz. Based on this visual inspection, it can be estimated the breathing time is about 5 seconds, and therefore the breathing rate is approximately 0.2 Hz.

To confirm this observation through data analysis, one possibility is to use a peak-finding algorithm, such as the findpeaks() function built into MATLAB, to locate the peaks in the spectrogram. It is also possible to implement a Fourier-based method to estimate the breathing rate. This would involve performing a Fourier transform on the audio signal to obtain the power spectrum, and then identifying the frequency band that corresponds to the breathing sounds. In this case, the first method described was employed.

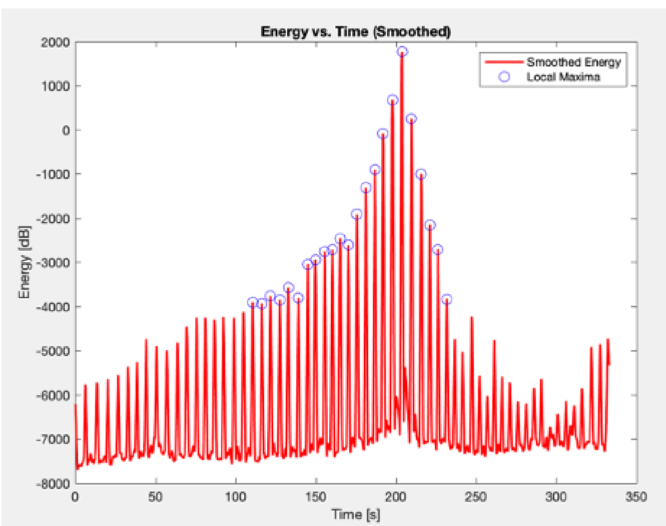

Figure 2: *Finding Maxima*

To find the breathing rate, an audio file was loaded and a spectrogram analysis was performed on the signal for hydrophone O using the built-in spectrogram() function in MATLAB. The energy of the spectrogram was calculated, and a moving average filter was applied to create a resulting energy curve. The smoothed energy curve was used to find local maxima in the energy curve, which are interpreted as peaks in the audio signal. The time intervals between adjacent peaks were calculated using the built-in findpeaks() function and used to estimate the average time interval between peaks. This corresponds to the breathing rate of the diver that produced the audio signal. The code then plotted the smoothed energy curve with the identified local maxima and calculated and displayed the breathing rate in Hz.

Doing this, the average time between breaths was found to be 5.416 seconds. The breathing rate of the diver in Hz was then calculated to be **0.18464 Hz**.

A similar approach also used the spectrogram, which has a very clear visualization of each breath. The breaths correspond with a wideband burst of noise that sustains for about a second. After converting the STFT data to decibels, the bins in each individual frame were summed to get an idea of total spectral intensity. A zoomed-in plot of this is shown below:

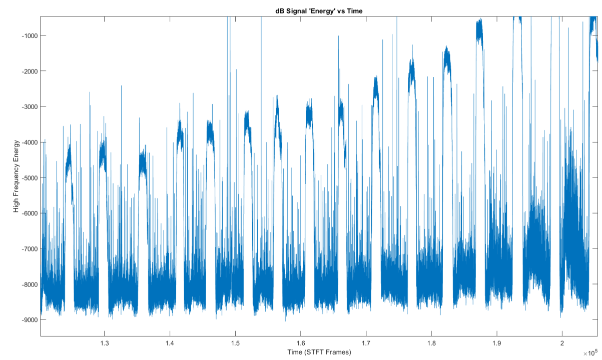

Figure 3: *Spectral 'energy' in dB for each frame*

Clearly, there is a periodicity in this signal, and it appears to correspond to the breathing rate, as confirmed by a listening test. An FFT of the above plot displays a peak of **0.18189 Hz** (see Fig. 4), which seems to line up visually with the STFT and audibly with the sound file. Despite 100x zero padding and the use of a Hanning window, there are significant sidelobes. This can likely be attributed to the nonuniformity of the diver's breathing rate; these other periodicities are likely also present in the breathing rate at one time or another.

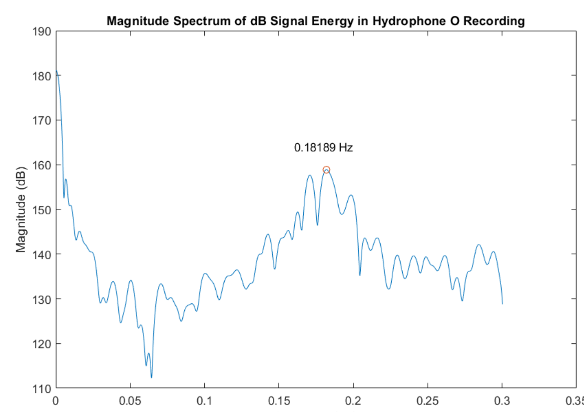

Figure 4: *Magnitude spectrum of the dB signal 'energy' in Hydrophone O*

IV. Task 2a

This task required us to estimate the time at which the diver is closest to hydrophone O. The code used for this section computes the spectral energy of the audio signal received by each of the hydrophones N, O, and P in the frequency range of interest which is expected to contain the dominant components of the diver's acoustic signature.

Starting with the same STFTs that we used in previous tasks, we then focused on the frequency bins that correspond to the frequency range of interest (above 20 kHz and below 125 kHz) and extracts the spectral energy in that range for each hydrophone using sum(). Below 20kHz, there was a lot of non-breathing noise, and above 125 kHz exceeds the Nyquist limit.

Next, our code identifies the time with maximum energy using the max() function. This corresponds to the loudest breath, which we can assume is also the closest breath taken to hydrophone O.

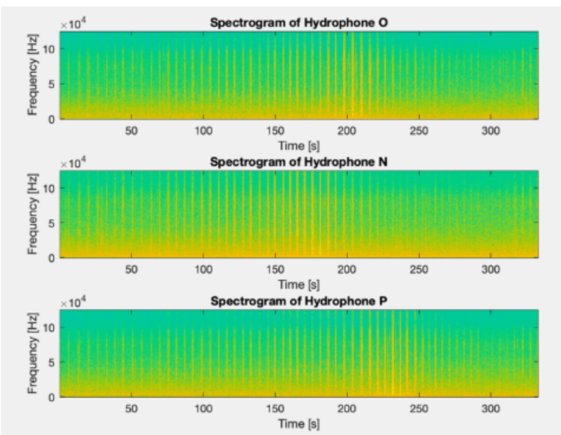

Figure 5: *Spectrograms of Hydrophones O, N, P*

The time of closest diver breath to hydrophone O was found to be 204.04 seconds. This calculation also makes sense visually based on Figure 5 above, as the band of greatest strength for hydrophone O seems to occur around 200s.

However, this calculation does not take into account the time between breaths. Such a discrete approach has a resolution of about 2.5 s. Looking at the signal energy in Figure 6, we can infer that the peak likely falls somewhere between the two largest peaks.

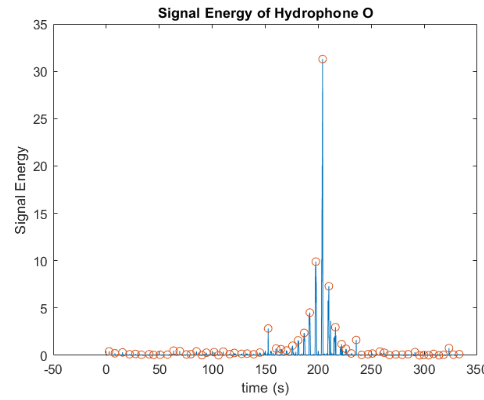

Figure 6: *Signal energy of Hydrophone O with peaks emphasized*

As the diver's breaths are separated by large blocks of silence, it is fairly likely that the actual time of closest approach is in between breaths. For this purpose, spline interpolation was used to predict the actual peak. Figure 7 shows the interpolated signal energy over time. Subsequent usage of Matlab's findpeaks() function shows the time of closest approach to Hydrophone O to be **203.83 seconds**.

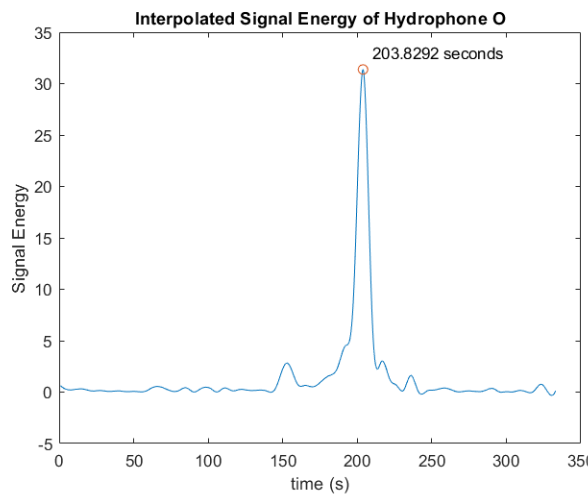

Figure 7: *Interpolated signal energy of Hydrophone O, with peak energy identified at 203.83 s.*

V. Task 2b

Task 2b asks to estimate the diver's altitude above hydrophone O when they are closest to said hydrophone at the time found in task 2a. To estimate the diver's altitude, we need to analyze the time delay between the signals received by each hydrophone. Since the hydrophones are arranged in a straight line, we can find the time delay between signal arrival at adjacent hydrophones and then use Pythagorean Theorem to find the vertical distance between the diver and the hydrophone array.

First, the Pythagorean Theorem was used to put the height component of the diver's position in terms of known variables as seen in Figure 8:

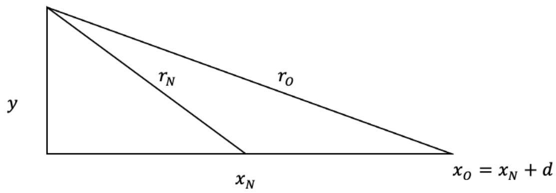

Figure 8a: *Height Diagram*

$$r_N{}^2 = y^2 + x_N{}^2$$

$$r_O{}^2 = y^2 + x_O{}^2 = y^2 + (x_N + d)^2$$

$$r_N = \sqrt{y^2 + x_N{}^2}$$

$$r_O = \sqrt{y^2 + x_N{}^2 + 28x_N + d^2}$$

$$r_O - r_N = c\Delta t = \sqrt{y^2 + x_N{}^2 + 28x_N + d^2} - \sqrt{y^2 + x_N{}^2}$$

$$if\ x_N = 0$$

$$c\Delta t = \sqrt{y^2 + d^2} - y$$

$$(c\Delta t + y)^2 = y^2 + d^2$$

$$c^2\Delta t^2 = y^2 + d^2$$

$$c^2\Delta t^2 + 2c\Delta t y + y^2 = y^2 + d^2$$

$$2c\Delta t y = d^2 - c^2\Delta t^2$$

$$y = \frac{d^2 - c^2\Delta t^2}{2c\Delta t}$$

$$y = \frac{1}{2}\left(\frac{d^2}{c\Delta t} - c\Delta t\right)$$

$$where\ d = 14m, c = 1520\frac{m}{s}, and\ \Delta t\ was\ found\ using\ xcorr()$$

Figure 8b: *Height Derivation*

The time delay between hydrophones N and O and hydrophones O and P can be estimated since the distance between hydrophones and the speed of sound in water are known. Doing this, the altitude above hydrophone O was found to be **4.31 meters.**

For this section of the project we had a lot of difficulty obtaining the time delay between various signal arrivals. The approach our team initially took was to look at the time when the signal was strongest above Hydrophone O (i.e., when the diver is directly above the middle hydrophone). At this time the diver breathes and the signal can be seen on all three hydrophones. If we zoom in on this plot with all three hydrophone plots overlaid, we can theoretically choose a noticeable peak or valley in all three signals and discern the time difference between these signals.

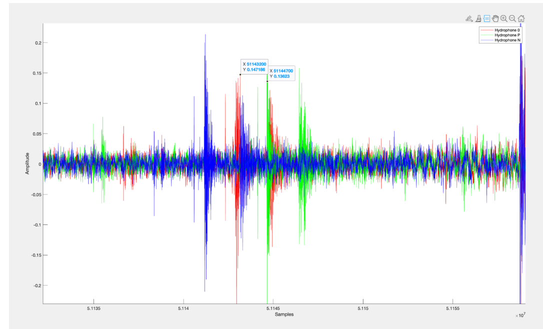

Figure 9: *The Three Hydrophone Plots Overlaid*

Figure 9 displays a noticeable noise spike (not likely a breath) apparent through all three hydrophones. The distance between these spikes appears to be 1500 samples which is about 0.006 seconds.

After further analysis, our team was unsure if these results were a reliable answer for the time delay. This was due to the fact that the noticeable spike appears earlier in time for hydrophone N than hydrophone O. If the noticeable spike was due to the diver (and since we are looking at the time when the diver is directly above hydrophone O and directly inbetween hydrophone N and P) then signals N and P would both be directly after the hydrophone O signal, directly on top of each other. As a result, our team decided to take another approach to find the time delay. We concluded that this spike must be due to some noise farther away, not due to the diver, since it arrives at hydrophone N first.

In an attempt to compare the time of arrivals of every breath at the three hydrophones, we crafted an algorithm to analyze the signal spectrum for sustained energy peaks. Due to the high volume of non-breath-related noise, this required the use of a few methods to filter the signal. First, a high-order Butterworth high-pass filter was applied to remove the noisiest frequency bands (0-20 kHz). The resulting spectrogram is shown in Figure 10.

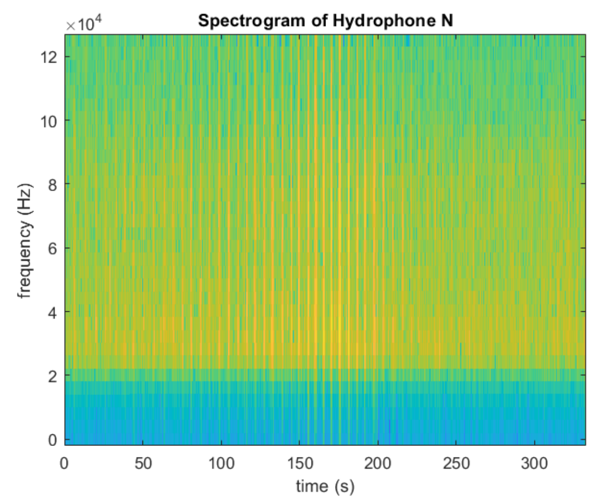

Figure 10: *Spectrogram of Hydrophone N after a 20th-order Butterworth high-pass filter was applied at a cutoff of 20 kHz.*

To extract the clear pattern seen in the spectrogram, a magnitude threshold was set at -39 dB to exclude all low intensity bins. A count of all bins that clear the threshold for each frame is shown below:

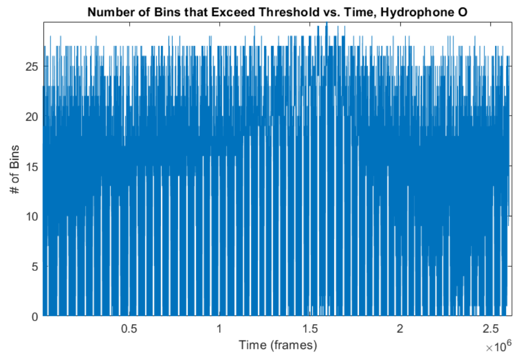

Figure 11: *Plot showing how many bins are greater than the magnitude threshold in each frame. Sustained periods where many bins exceed the threshold correlates well with the breathing rate.*

The noticeable white bands in Figure 11 correspond well to the diver's breaths and appear to show their durations, as well. We devised an algorithm to look ahead/behind each sample and mark spots that begin and end these periods of sustained threshold-clearing. The results, shown in Figure 12, appeared to be promising. However, the level of accuracy needed to properly determine time of arrivals was not achieved with this method. Calculated time delay values gave corresponding heights around 30 m, which was not feasible given the fact that the water was only 20 m deep.

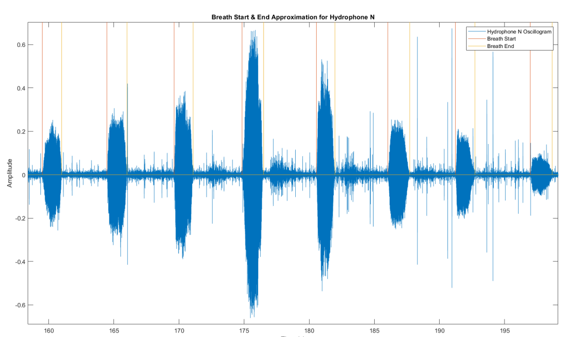

Figure 12: *Amplitude plot of Hydrophone N overlaid with algorithmically derived start and end points of each breath. While initially appearing promising, subsequent height calculation proved this method too inaccurate.*

VI. Task 2c

Task 2c asked to estimate the diver's swimming speed. In this case, the time delays between the hydrophones can still be used. We can also use the time it takes for the diver to move between the hydrophones to estimate their speed. Since we know the spacing of the hydrophones and the time it takes for the diver to move between them, we can calculate the speed of the diver.

To do this, the hydrophone signals were cross correlated to estimate time delays. The time delays between hydrophones were then used in conjunction with the known distances between hydrophones to calculate the speed of the diver and the output was printed. The diver was found to be moving at **0.45 m/s**, which is a leisurely pace, but realistic for a slow moving diver.

Another method used to calculate the velocity was a similar approach to that of finding the time at which the diver was closest to each hydrophone. If we know the times at which the diver is "closest to each hydrophone, and we know the distance between hydrophones, we can use each of these variables to calculate the velocity of the diver, since $v = d / t$.

The same code from task 2a was used but utilized the datasets from each hydrophone. It was found that the diver was closest to hydrophone N at 175.58 seconds, hydrophone O at 204.04 seconds, and hydrophone P at 238.42 seconds. The distance between each hydrophone is 14 meters, making the distance between the outer hydrophones 28 meters. So, the average velocity of the diver was calculated to be **0.41 m/s between hydrophones O and P**, while **0.49 m/s between hydrophones N and O**. For the overall average (between N and P), we calculated a velocity of **0.44 m/s**. This is a plausible velocity for a diver, and about the same as the method which utilized xcorr() previously.

VII. Significance

This challenge problem involved the practical implementation of several key signal processing techniques. Take 1a involved displaying an output spectrogram of the hydrophone signal. Spectrograms are very important in analyzing audio signals since they can provide a visual representation of signal changes over both time and frequency by plotting the intensity of the frequency components. Spectrograms are very useful in identifying and analyzing signals and for feature extraction.

Task 1b used a variety of signal processing techniques in tandem. First, energy calculations were utilized. The energy of a signal can provide information about the signal's power, intensity, and loudness. Normalization was then used, which helps when comparing different signals and making them more consistent for processing. A moving average filter was used to smooth out the energy curve and remove noise that may be present. This helps in identifying maxima more accurately, which is helpful for analyzing any audio signal with any sort of noise. Peak detection was utilized, which can provide valuable information about the characteristics of the signal. It can be used to identify transients, periodic sounds, etc. This process helps to simplify complex signals by identifying the most significant parts. It can also extract important features such as amplitude, frequency, or duration. Time interval calculations are critical for any type of rate or frequency analysis.

Task 2a used similar techniques, including spectrogram generation and peak detection. Other important processes for audio included the dB scale conversion, which is common for audio applications since humans hear on a dB scale, as well as spectral energy extraction, which extracts the spectral energy in a frequency range of interest by summing the power spectrogram values across the relevant frequency bins. This is very useful to focus on specific frequency bands in a signal when noise may want to be disregarded.

In task 2b, time delays were calculated and the between hydrophones were used with propagation speed to calculate the distance traveled by a wave. On a conceptual scale, the propagation of sound in a medium was analyzed to determine the location and qualities of a source. The use of cross-correlation to estimate the time delay between signals is a widely used technique in signal processing, particularly in audio processing and speech recognition. The calculation of distance and altitude based on propagation time and applying relevant mathematical concepts is also important for sonar applications.

Lastly, task 2c applied concepts used in part 2a, plus cross-correlation and the relationship between distance and time to calculate velocity. The concepts used in these tasks can be applied to a wide range of signal processing fields, including medical imaging, speech recognition, and machine health. These are just a few examples of how the signal-processing techniques used in this challenge problem can apply to a wide range of signal-processing applications. Several concepts used throughout this entire semester were conceptually applied in this problem.

VIII. Conclusion

The International Student Challenge Problem in Acoustic Signal Processing 2023 has significant implications in the field of signal processing. This particular problem involves analyzing the acoustic signature of a scuba diver, which consists of periodic broadband emissions. By analyzing the signals received by the three hydrophones arranged in a line array, it is possible to extract properties about the diver and environment.

This problem is important because it highlights the potential applications of signal processing in acoustics. Acoustic signal processing can be used to analyze and interpret acoustic signals from a variety of underwater sources, and these same concepts generally apply to acoustics at large as well. The techniques used in this problem can be applied to a range of other underwater or above-water acoustic sensing applications. Further, this challenge problem provides an opportunity to gain experience in signal processing and to develop approaches for solving practical problems utilizing concepts learned in class. Some important concepts utilized in this problem that apply to a myriad of signal processing applications include the use and interpretation of spectrograms, signal envelopes and filters, power, dB scaling, time of arrival, cross-correlation, Fourier transforms, Doppler effects, and more. Please note full source code is also available as an attached document.

References

[1] Ferguson, B.G., Culver, R.L., & Gemba, K.L. (2023). International Student Challenge Problem in Acoustic Signal Processing 2023. Acoustics Today, 19(1), 52-59.

[2] The MathWorks Inc. (2021). MATLAB (R2021a) [Computer software]. Natick, MA: The MathWorks Inc.

[3] MathWorks. xcorr. Retrieved April 23, 2023, from https://www.mathworks.com/help/matlab/ref/xcorr.html