Time-frequency Resolution, Representation, and Interpretability

MR23-3982

Wendy Newcomb

Georgia Tech Research Institute (GTRI)

Final Project Presentation

August 14, 2025

Project Team

				
Wendy Newcomb **Senior Research Scientist** *GTRI*	**Dr. Alessio Medda** **Principal Research Scientist** *GTRI*	**Dr. Aprameya Satish** **Senior Research Engineer** *GTRI*	**Jayden Smith** **Graduate Student** *GTRI*	**Saaketh Reddy** **Undergraduate Student** *GTRI*
Principal Investigator	Co-Principal Investigator	Technical Contributor	Graduate Assistant	Student Assistant

Bottom Line Up Front

- This research addresses the Statement of Need (SON) MRSEED-23-S1 to develop technologies to classify targets under complex conditions and identify specific types of targets; especially in a cluttered environment
- Specifically, this work investigates time-frequency processing for the purpose of assessing the feasibility of classifying UXO with similar structure and distinguishing material composition or UXO with similar material composition and small variations in structure
- The first half of this work demonstrated a structured quantitative approach to spectral representations
- The second half of this work investigates signal separation and subsequent impacts on spectral descriptors, as well as experiments in classification using different time-frequency data representations

Technical Objective

1. Identify optimal/useful transform for time-frequency representation suitable for feature selection from a time-frequency vs. aspect volume of data
2. Implement signal processing methods that are effective in feature-space reduction in the time-frequency domain for sparse dataset generation to improve classification
3. Apply statistical analysis methods useful in signal separation in the time-frequency domain for generating denoised data representations among different UXOs in a clutter scene
4. Use interpretable DNNs to inform what frequency bands and time steps are more influential in classification to support increased sparsity in datasets used for classification and improve accuracy in classification

Technical Approach

Task 1

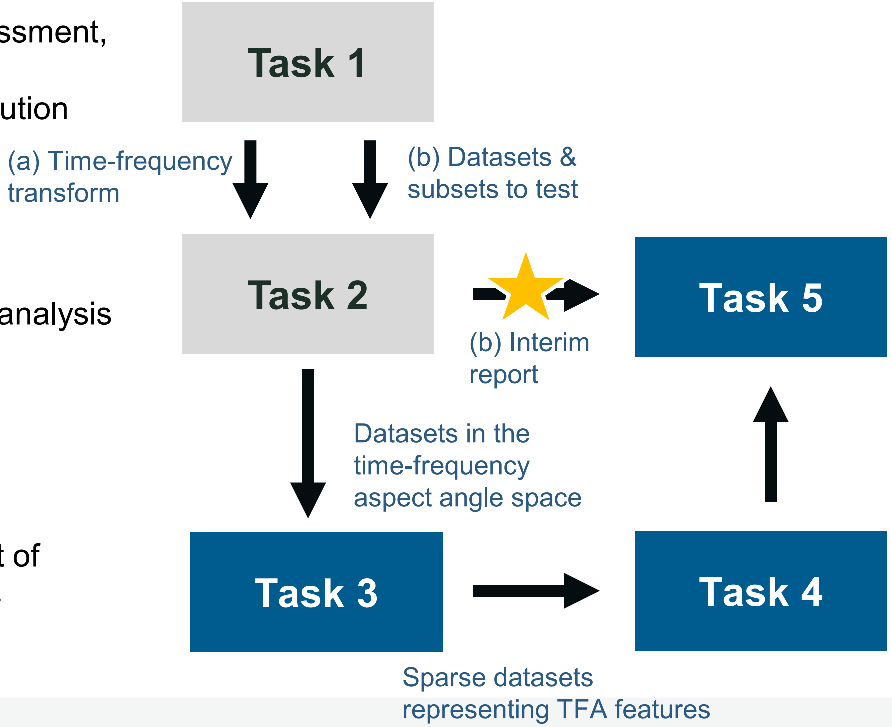
(a) Dataset review, assessment, labeling, & selection
(b) Time-frequency resolution optimization

Task 2

(a) Spectral Descriptors analysis

Task 3

(a) Assessing the impact of denoising techniques



Task 5

(a) Interim report
(b) Final report

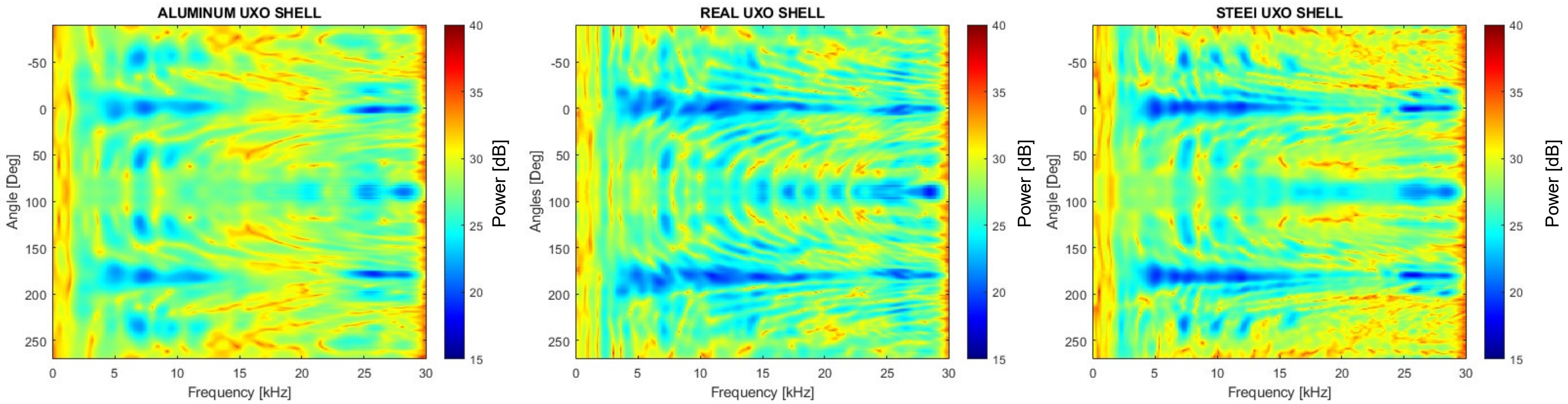
Task 4

(a) Implement deep neural networks (DNNs) for time series analysis
(b) Evaluate interpretable measures from application of DNNs

Overview

- Background: frequency analysis in ordnance characterization
- Time-frequency transform optimization
- Spectral descriptors
- Signal separation in the time-frequency domain- noise removal approaches
- Classification using diverse time-frequency data representations
- Initial interpretable outcomes

PONDEX 2010



- FFT across entire time series
- Analysis in multiple domains:
  - Aspect angle vs. frequency
  - Wavenumber $k_x - k_y$
- High frequency resolution
- Missing time resolution

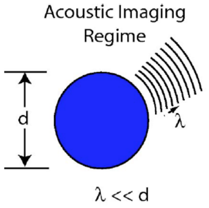
In situ UXO shells from the PONDEX 2009 Experiment [1]

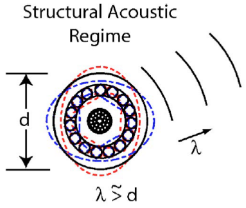
SERDP

Cylinder Wave Physics

The interference between specular and elastic scattering creates an acoustic spectrum that can provide valuable information about that object's shape, composition, and orientation.

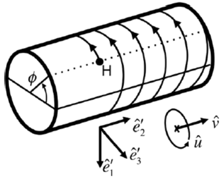
Acoustic scattering "**specular**" in nature- direct reflections as a function of impedance/grazing angle illustrate geometric shape of an object

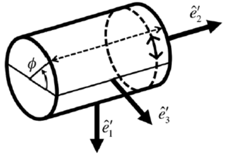




Acoustic scattering related to **vibrational** dynamics of the object, both whole-body and internal structure

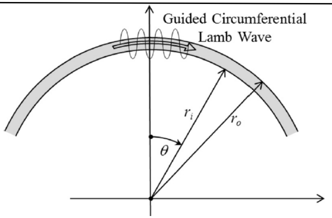
Helical Rayleigh waves: surface elastic waves launched so that they propagate in a spiral down the cylinder

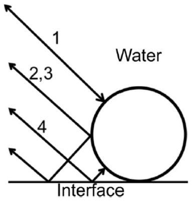




Waves that propagate around the circumference of the cylinder are **circumferential Rayleigh waves**; **meridional Rayleigh waves** propagate parallel to the cylinder axis are.

Guided circumferential lamb waves – elastic wave motion is 2D within the plane defined by wave propagation direction and the plate thickness direction





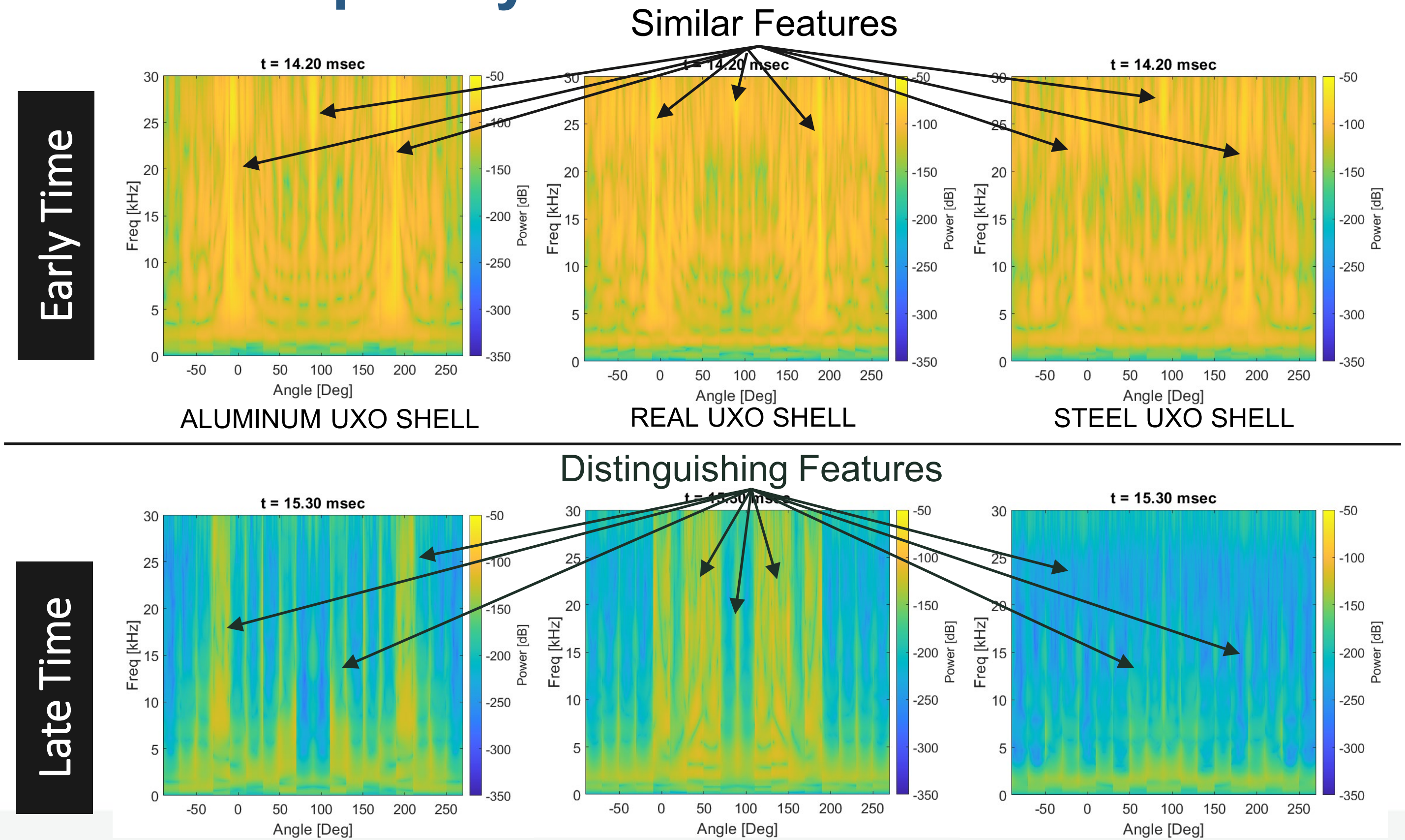
Scattering wave **reflection paths** - the **direct path** 1, and other **multipath directions** 2-4

References: [2]-[5]

Extending Dimensional Analysis in the Time-Frequency Domain

PONDEX 2010

UXO SHELLS



Early time arrivals consist of the direct specular reflection or the initial geometric response from the shape of the object, they have the shortest travel time path and the strongest intensity characteristic of the surface and its impedance or reflectivity.

Late time acoustic wave behaviors are from wave field interactions related to vibrational dynamics of the object, both whole body and internal structure (see next slide)

SERDP

Background

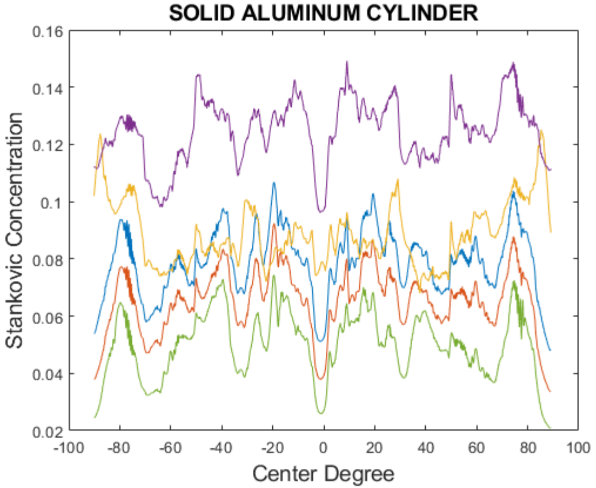
While targets were either proud, half-buried, or flush buried, this work only considers **proud targets on sand**

- Past work by *Morse & Marston [6]* and *Anderson & Sabra [7]* in time-frequency analysis for spherical and cylindrical shells demonstrates the capability to characterize the mid-frequency enhancement waves circumnavigating objects in underwater environments
- This work focuses on data collected from PONDEX 2009/2010, TREX 2013, and CLUTTEREX 2017 experiments.
  - Collections on a 20 meter rail with an LFM chirp 1-30 kHz transmitting at 2 Hz
  - Each object is rotated clockwise by a 20° increment.
  - Datasets are named for the center angle designed to include +/- 10 degree of aperture angle and +/- 5 degree overlap with the next dataset.
  - A complete 180° of angle returns is collected for each target with the assumption that targets have axis symmetry along the longer axis.
  - Targets vary in range from the source by 5 meter increments from 5-40 meters

Time-Frequency Transform Optimization

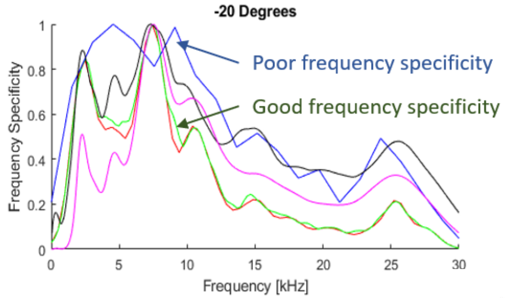
Transform Parameter Optimization

Optimizing Stankovic concentration to inform parameter selection for each transform



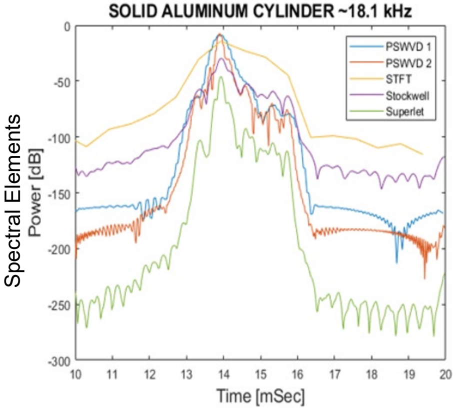
Stockwell | STFT | SPWVD1 | SPWVD2 | Superlet

Performance Analysis



SPWVD1
SPWVD2
STFT
Stockwell
Superlet

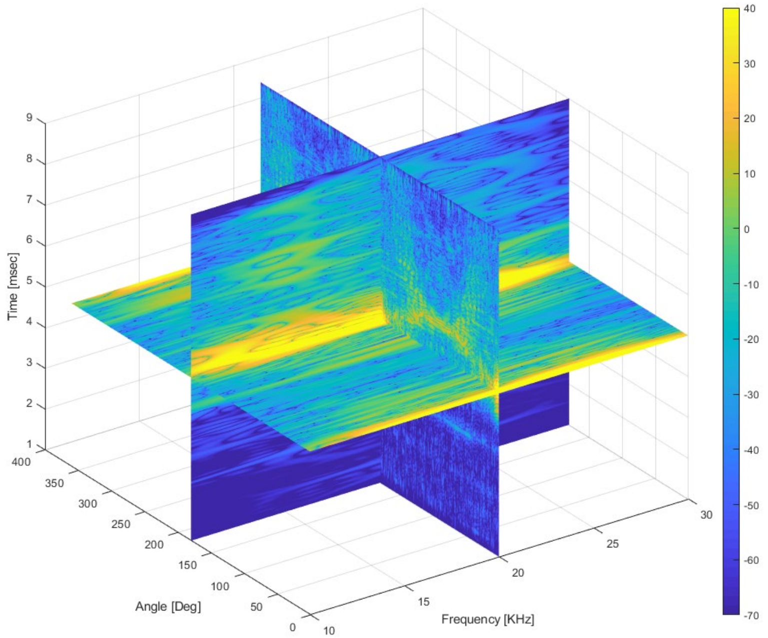
The Superlet transform and smoothed pseudo Wigner-Ville transform have high concentration, good frequency specificity, and better SNR in the spectral elements



SERDP

Time-Frequency-Aspect Angle

Dimensionality Reduction from 3D to 2D



Spectral Descriptors Over Frequency

- Spectral Crest
- Spectral Slope
- Spectral Centroid
- Spectral Spread
- Spectral Skewness

Spectral Descriptors Over Time

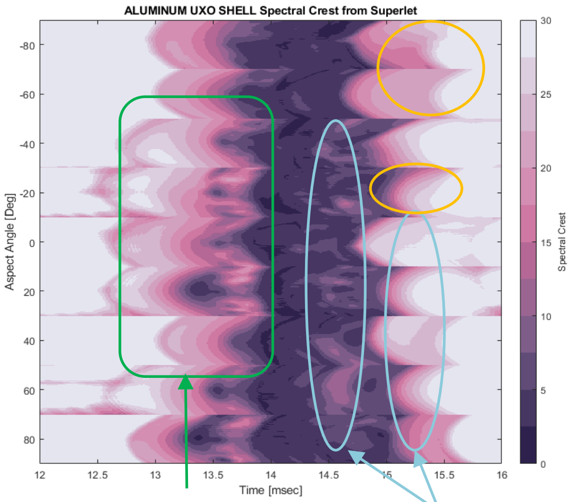
- Normalized Renyi Entropy
- Frequency Variation

SERDP

Spectral Descriptors

PONDEX
ALUMINUM UXO SHELL

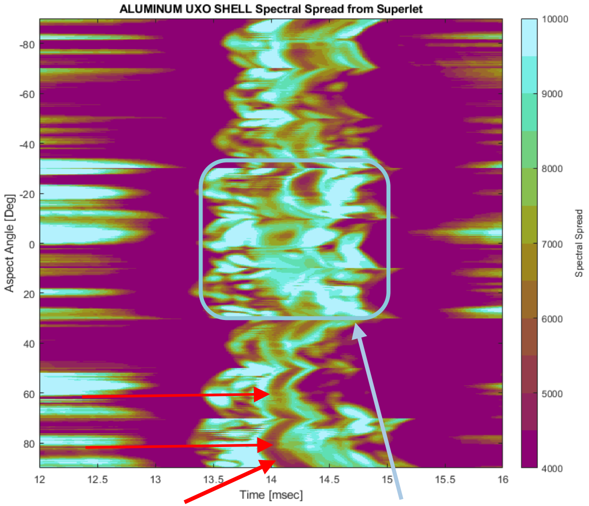
Spectral Crest



Moderate deviations in spectral content

Larger deviations in spectral content

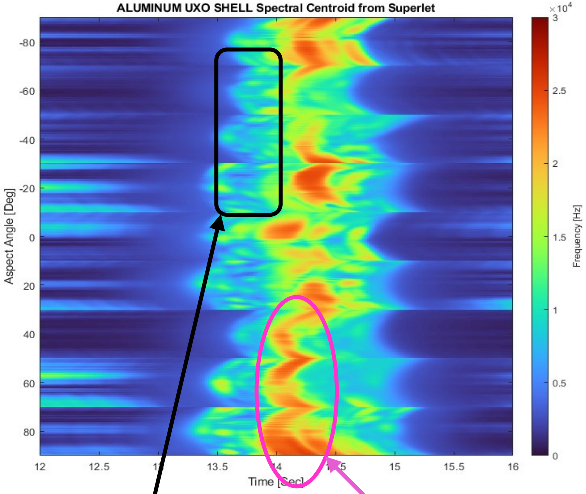
Spectral Spread



Individual tonal sounds have low spectral spread

Noise like signals usually have large spectral spread

Spectral Centroid



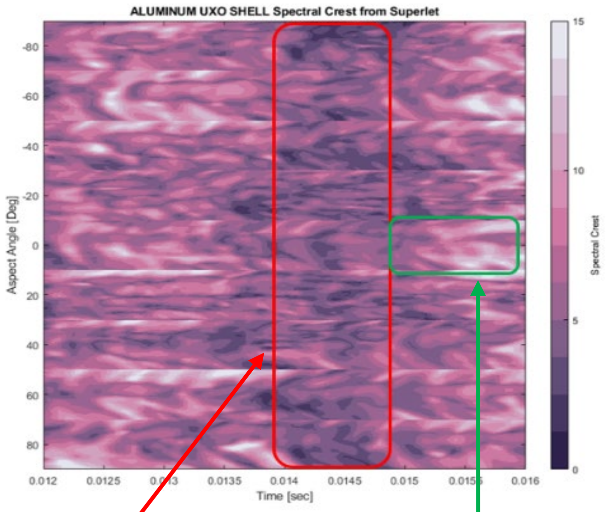
Energy is centered close 10 - 15 kHz

Energy is centered close to 25 - 30 kHz

Spectral Descriptors

TREX
ALUMINUM UXO SHELL

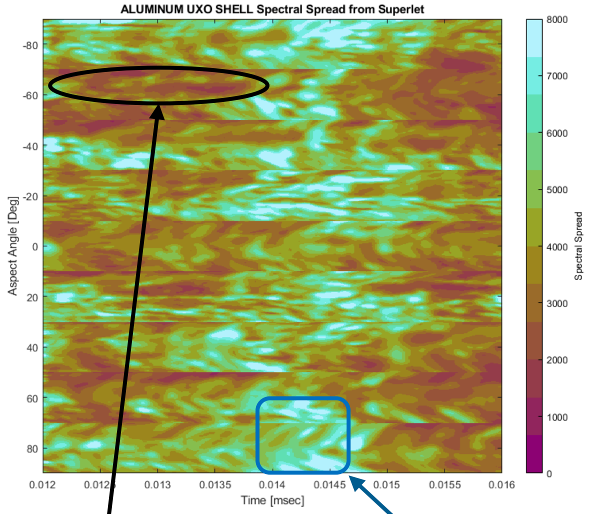
Spectral Crest



Small deviations in spectral content

Moderate deviations in spectral content

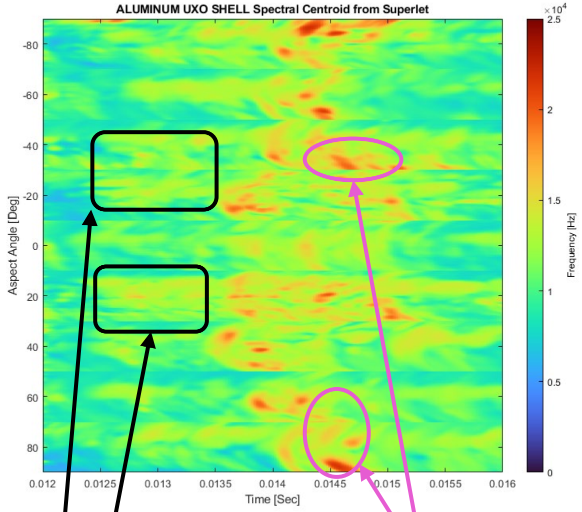
Spectral Spread



Individual tonal sounds have low spectral spread

Noise like signals usually have large spectral spread

Spectral Centroid



Energy is centered close 10 - 15 kHz

Energy is centered close to 25 - 30 kHz

TREX data is severely impacted by noise

SERDP

Signal Separation in Time-Frequency

- Subspace array processing for the smoothed pseudo Wigner-Ville distribution
  - Signal separation using a large spatial time frequency distribution
  - The cross Wigner-Ville distribution provides a means to compute the coherence of two non-stationary signals along both the time and frequency axis
  - This is used for consecutive channels where the primary acoustic return is likely to be coherent across a small subspace of the array, and noise is more likely to be random and vary per channel
  - For a single snapshot of received time series across an array of N receivers, the XPSWVD can be calculated for all the pairwise combinations and combined into a spatial time frequency distribution super matrix (STFD) given as:

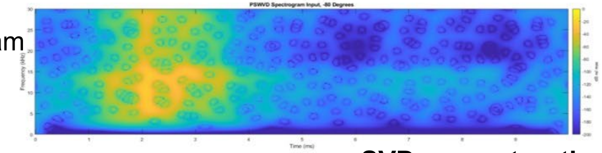
$$\widetilde{WV}_{x_i x_j}(t,f) = \int_{-\infty}^{\infty} h(\tau) \int_{-\infty}^{\infty} g(u-t) x_i\left(u+\frac{\tau}{2}\right) x_j{}^*\left(u-\frac{\tau}{2}\right) e^{-j2\pi f\tau}\, du\, d\tau \longrightarrow STFD = \begin{bmatrix} WV_{x_1x_1} & WV_{x_1x_2} & \cdots & WV_{x_1x_N} \\ \vdots & & \ddots & \vdots \\ WV_{x_Nx_1} & WV_{x_Nx_2} & \cdots & WV_{x_Nx_N} \end{bmatrix}$$

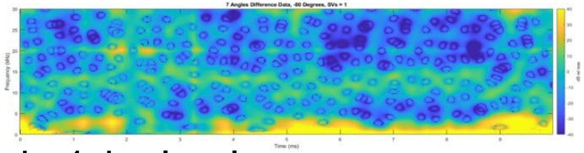
  - Here the diagonal becomes the SPWVD for a single channel and off-diagonals are the XSPWVD for channel pairs
  - Sabra and Anderson [8], [9] propose to apply singular value decomposition to the full STFD matrix to support extraction of the signal components from the noise. This approach leverages the notion that most of the signal power is confined within the first n principal components with the noise dominating the $n\ to\ N$ components.

SERDP

TREX
ALUMINUM UXO SHELL

SPWVD Spectrogram

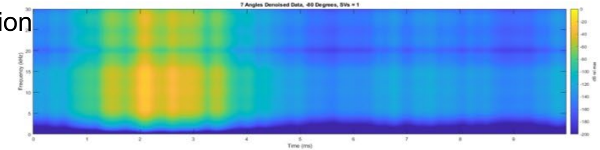


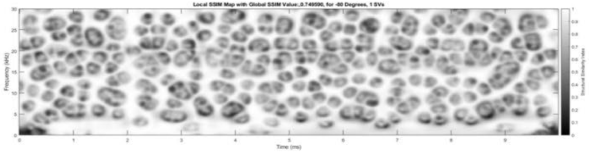


Difference

SVD reconstruction using 1 singular value

After reconstruction

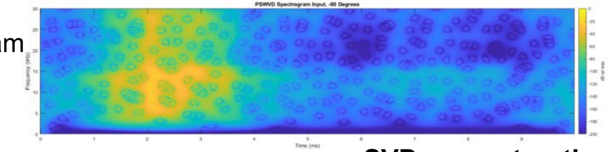


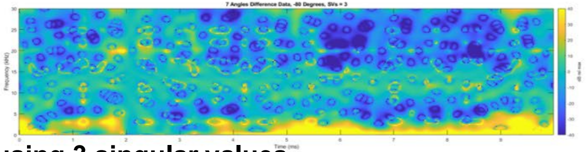


Structural similarity index map

Reconstruction using 1 singular value removes an excess amount of information from the spectrum

SPWVD Spectrogram

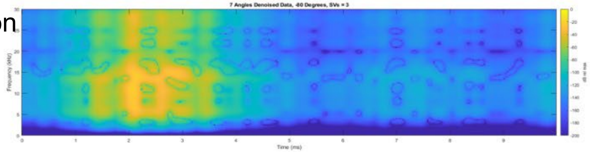


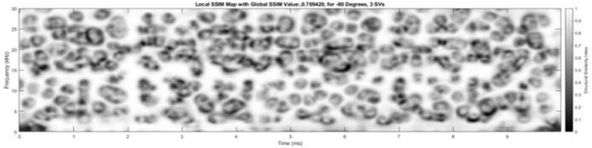


Difference

SVD reconstruction using 3 singular values

After reconstruction

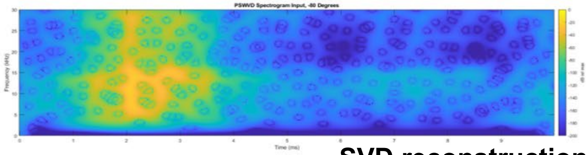


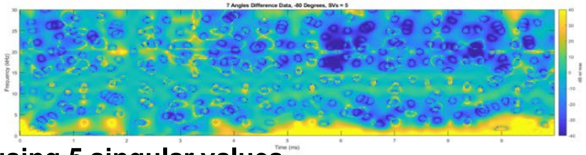


Structural similarity index map

Reconstruction using 3 singular value is more moderate, removes some cross terms, but introduces additional artifacts

SPWVD Spectrogram

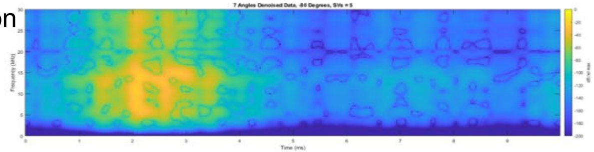


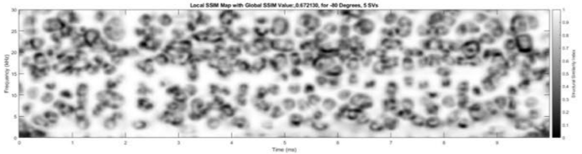


Difference

SVD reconstruction using 5 singular values

After reconstruction

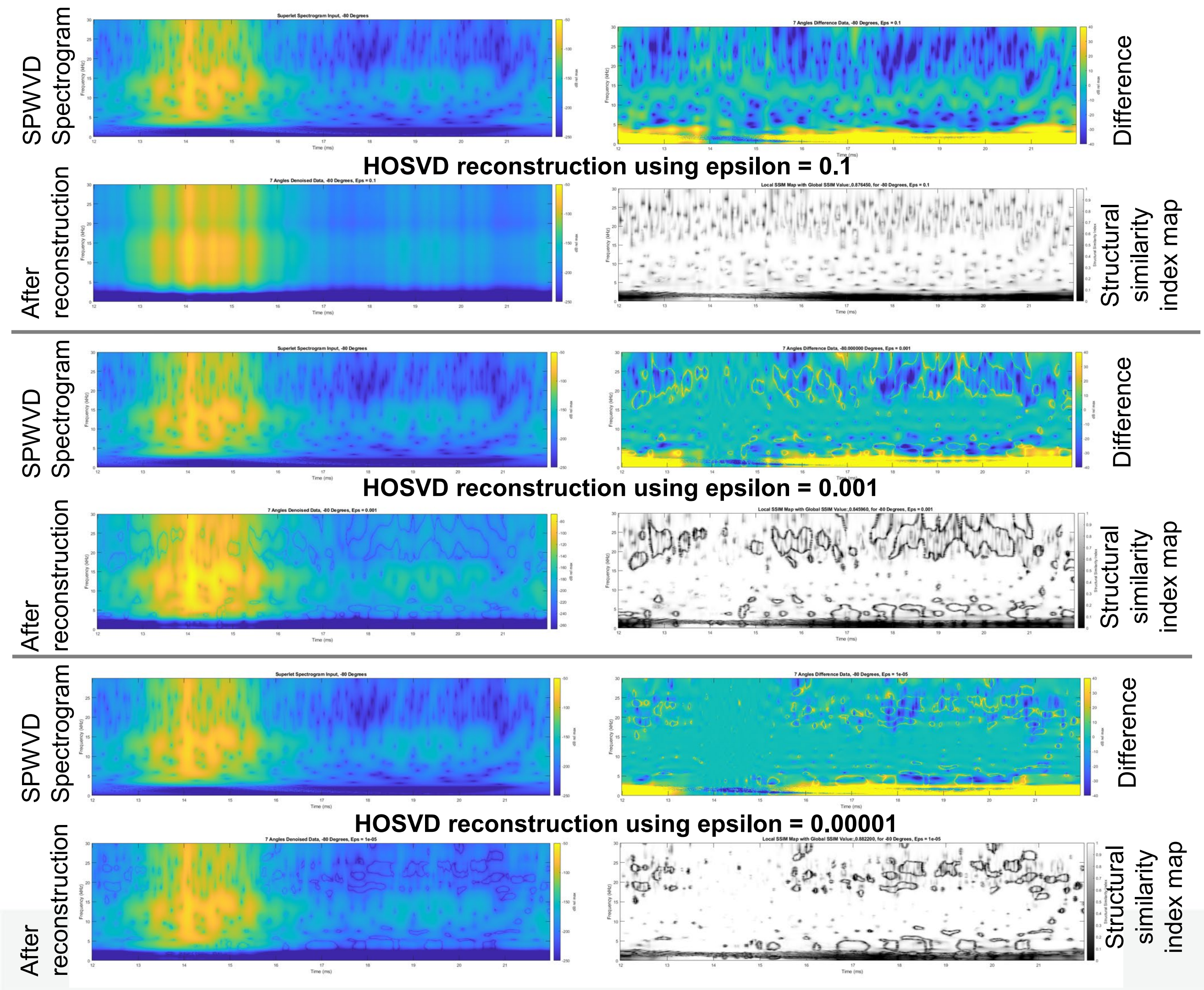




Structural similarity index map

Reconstruction using 5 singular value is extremely similar to 3 SVs. Increasing the # of SVs past 3 yields very little change in outcomes

SERDP



Higher Order Singular Value Decomposition for the Superlet transform (HOSVD)

- *Shift from super matrix to 3D tensor approach*
- *Eigenvalue decomposition with epsilon thresholds set for eigenvalue reconstruction [10]*

$$T_{ijk} = \sum_{I=1}^{M} \sum_{J=1}^{N} \sum_{K=1}^{P} G_{IJK} U_{iL} V_{jJ} W_{kK}$$

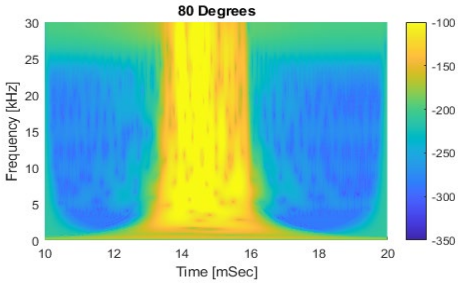
- Reconstruction using eps = 0.1 is too sparse
- Reconstruction using eps = 1e-3 is moderate but introduces some artifacts in the reconstruction
- Reconstruction using eps = 1e-5 removes the least signal content but artifacts are still present in outcomes

SERDP

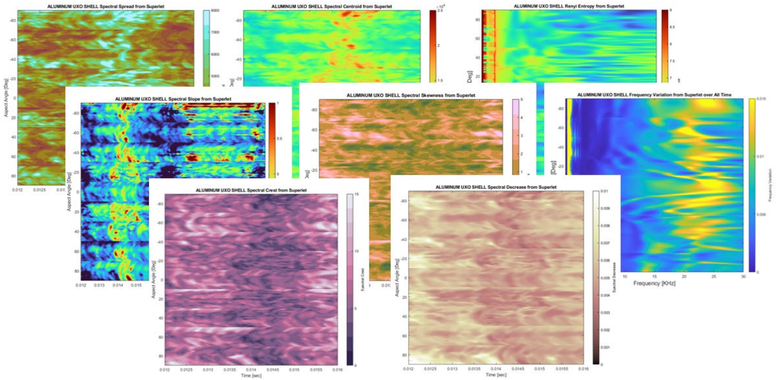
Time-Frequency Classification

Three groups of time-frequency data representation

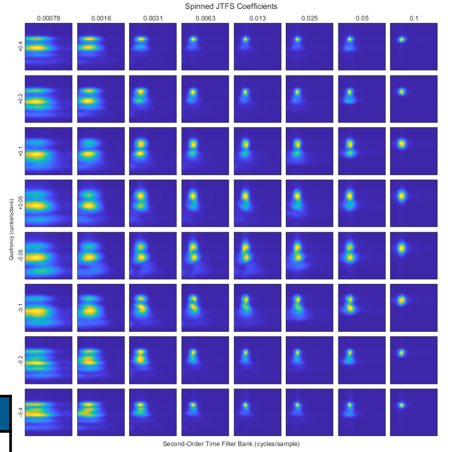
Superlet Spectrograms



Spectral Descriptors from the Superlet Spectrograms



Joint Time Frequency Scattering Coefficients



Classification tests are run on each of three categories of object type

26 Objects	35 Objects	32 Objects
Aluminum	**Howitzer**	**UXO**
Solid Aluminum Cylinder	155 mm Howitzer	Original Material 100 mm UXO
Hollow Aluminum (Aluminum Pipe)	155 mm Howitzer no Collar	Solid Aluminum 100 mm UXO Replica
	155 mm Howitzer no Nose no Collar	Solid Steel 100 mm UXO Replica

SERDP

Neural Network Classification

Classification of Super Spectrogram and Spectral Descriptors	Classification of Joint Time Frequency Spin Up and Spin Down Coefficients [11]
Input is 2D Spectrogram angles selected = -60°, -70°, -80°, -90°	Input is 3D tensor of all 2D coefficients Time series angles selected = -60°, -70°, -80°, -90° for input into JTFS transform
Data split = 70% training, 30% testing	Data split = 70% training, 15% testing, 15% validation
14 layers	13 layers
initial learning rate of .012	initial learning rate of .005
L-2 regularization of 6.4530e-9	L2 regularization of .0001
150 epochs for training	150 epochs for training
Cross entropy loss function	Cross entropy loss function
25 iterations are run for each aspect angle within a category and for each spectral descriptor within a category	25 iterations are run for each aspect angle within a category

Classification Outcomes

Spectral Descriptor Classification Results

Spectral Descriptor	Howitzer	Aluminum	UXO
Spectral Centroid	59.09%	100%	30%
Spectral Crest	45.45%	57.14%	35%
Spectral Decrease	18.18%	42.86%	45%
Spectral Skewness	40.91%	78.57%	35%
Spectral Slope	27.27%	57.14%	35%
Spectral Spread	31.82%	71.43%	35%
Frequency Variation	45.45%	85.71%	35%
Renyi Entropy	50%	85.71%	45%

JTFS Transform Coefficients Classification Results

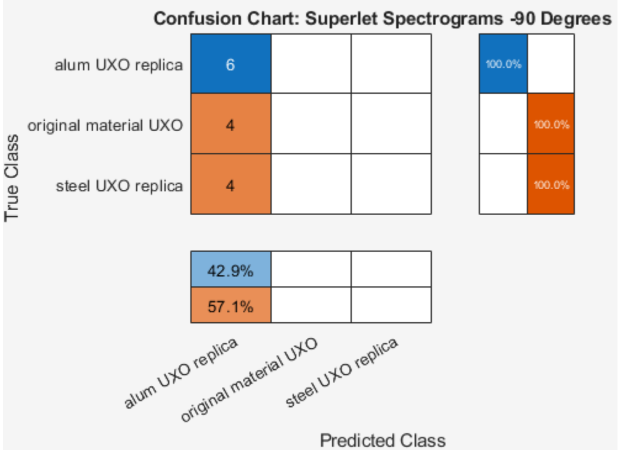
Category	-60°	-70°	-80°	-90°
Howitzer	63.33%	36.67%	43.33%	80%
Aluminum	75%	100%	97.5%	100%
UXO	33.33%	33.33%	35%	33.33%

Superlet Spectrogram Classification Results

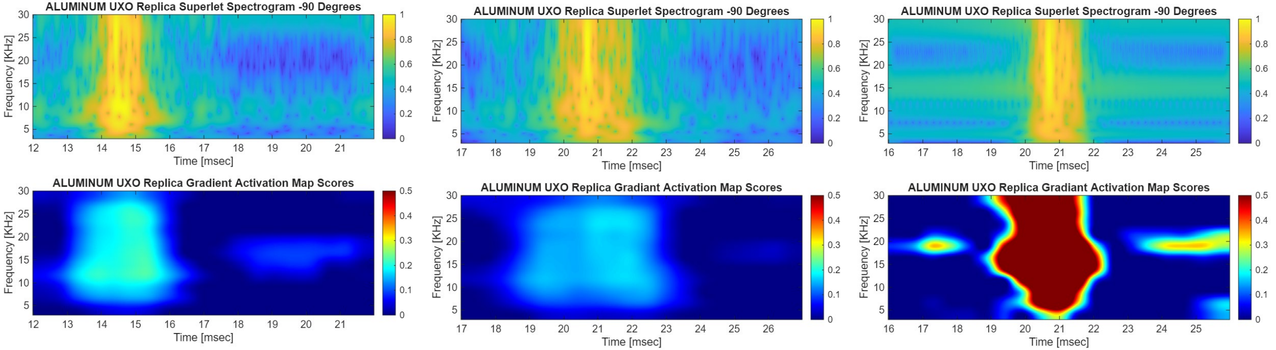
Category	-60°	-70°	-80°	-90°
Howitzer	63.64%	81.82%	72.73%	72.72%
Aluminum	71.43%	85.71%	71.42%	85.71%
UXO	40%	30%	50%	20%

- Spectral descriptor classification results are overall very poor and the least successfully among the three time-frequency dataset representations.
- JTFS transform coefficients results are mixed, lower scores give low confidence in this approach
- Single angle Superlet spectrograms generate the highest classification scores among the three time-frequency dataset representations
- The aluminum category is consistently highest scoring like because there are only two classes
- The UXO category is consistently the lowest scoring – this may indicate that characterizing material composition among the same geometric structure is a more challenging problem to the networks

Interpreting Classification Outcomes

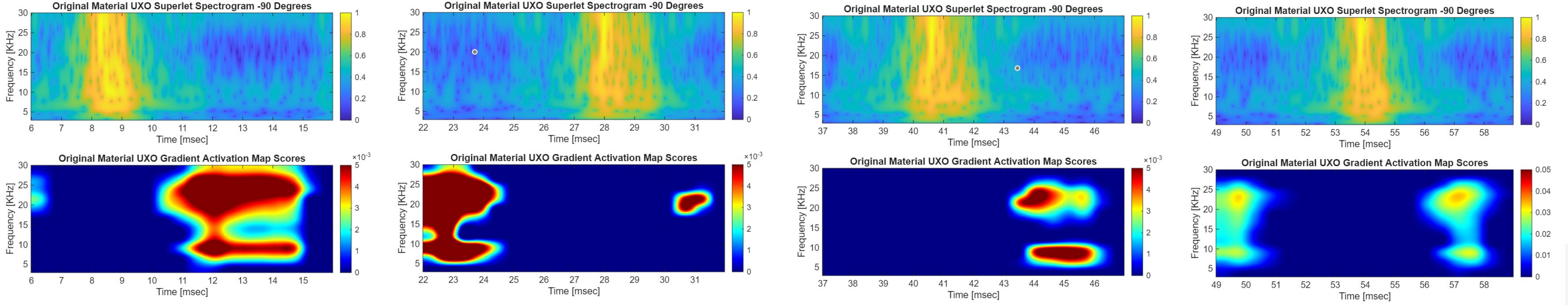


Aluminum UXO Replica – Example with 100% classification rate



Gradient activation weights are higher over times containing primary signal – or higher power time-frequency content from the object's acoustic return

Original UXO Material – Example with 0% classification rate



Conclusions

- The adaptive Superlet transform provides good time-frequency concentration and frequency specificity with higher time-frequency resolution while the smoothed pseudo Wigner-Ville distribution provides moderate time-frequency resolution with the best frequency specificity and higher resolution at the expense of artifact cross-terms.
- Spectral descriptors over time demonstrate small & large deviations in spectral content, increase and decrease of energy with increasing frequency, central frequency, as well as broad and narrow bandwidth features. Spectral descriptors over frequency demonstrate angles & times of high and low spectral concentration and variation/coherence.
- Initial tests of signal separation using SVD methods in the time-frequency domain are inconclusive. Results demonstrate a reduction in signal content with introduction of additional artifacts and uncertainty on the optimal thresholds for reconstruction.
- Classification results demonstrate that simple use of the Superlet spectrogram and single angles performs better than other time-frequency representations. Gradient activation maps illustrate when the object's acoustic return is of importance and leads to success and when the background takes priority and leads to classification failure.

Future Research

- Unexpected classification results shift this area of research in a different direction –focus on pure time-frequency spectral content for TREX & CLUTTEREX data
  - Classification approach should be refined to focus the networks on the acoustic return of the object of interest and steer away from background content
  - Could also consider using early time and late time angle slices from a time-frequency aspect angle (Superlet) cube for classification input
- For volumetric data (MuST or SVSS), a small study is needed to identify the orientation and feasibility of spectrogram representation for buried objects

SERDP

Questions

Special thanks to:

- Applied Physics Lab University of Washington
  Dr. Aubrey Espana, Dr. Steven Kargl, & Dr. Tim Marston.
- Penn State Applied Research Lab
  Dr. Dan Brown, Dr. Daniel Park, & company

BACKUP MATERIAL

MR23-3982 : Time-frequency Resolution, Representation, and Interpretability

Performers: Wendy Newcomb (PI), Alessio Medda (Co-PI), Aprameya Satish (Technical Contributor)

Technology Focus

- *Develop technologies to classify targets under complex conditions and identify specific types of targets; especially in a cluttered environment.*

Research Objectives

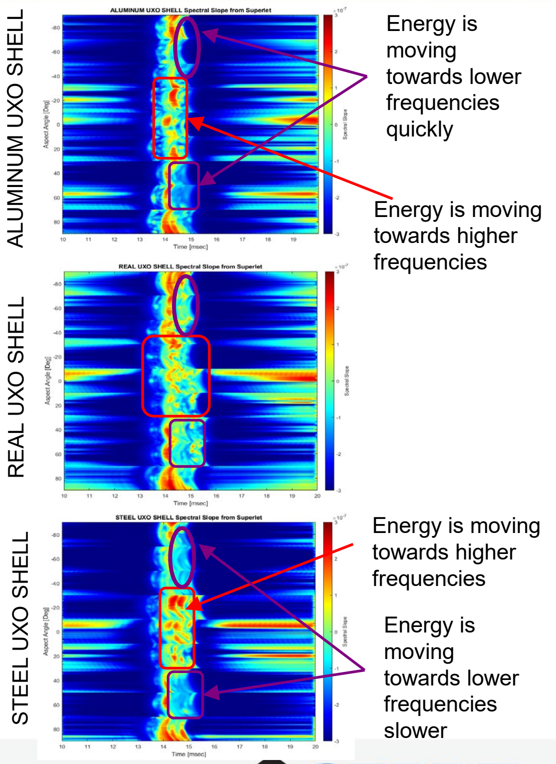
- *Investigate unexploded ordnance detection, characterization and classification through time-frequency-aspect angle analysis*

Project Progress and Results

- *Identified and applied suite of time-frequency transforms optimized for high resolution and concentration*
- *Applied spectral descriptors to time-frequency-aspect angle underwater sonar field datasets*
- *Experimented with subspace array signal separation in the time-frequency domain to try to remove undesirable signal and noise*
- *Completed initial small scale classification experiments across multiple time-frequency data representations*

Technology Transition

- *Fundamental research requiring additional exploration, experimentation, validation, and application to buried ordnance datasets before technology transition*



Plain Language Summary

- Develop technologies to classify targets under complex conditions and ***identify specific types of targets***; especially in a cluttered environment.
- Develop framework for research in time-frequency aspect angle analysis on unexploded ordnance, and demonstrate statistical tools and approaches that can provide distinguishing spectral information among targets to characterize internal structure and potentially material composition
- Current state of the art analysis of frequency spectral returns of targets in acoustic color apply the Fourier transform providing perfect frequency resolution and zero time resolution. The acoustic wave field return of an object on the seafloor varies over time. Those spectral time variations are a function of multipath, environmental conditions, geometrics shape, and wave field interactions with the object that are a reflection of internal structure and material composition etc.
- This work extends beyond traditional acoustic color domain analysis of frequency vs. aspect angle for an entire time series acoustic return by looking at variations, patterns, and correlations that are only visible in the spatial time-frequency domain that may support identification of specific types of target that differ in terms of internal structure or material composition

Impact to DoD Mission

- This is the first SEED project presentation for this work. Providing strong sound logic for time-frequency transform selection should dispel concerns regarding tradeoffs and challenges with time-frequency distributions. In addition, demonstrating statistical methods in the form of well established spectral descriptor methods opens up a new area of research in unexploded ordnance target characterization and classification, advancing the current state of the art in new directions
- Current and historical work in this area has demonstrated the ability to classify objects into to classes: natural and man-made objects
- This work is establishing a new analysis space for making headway in classification between different man-made objects

Publications

- Newcomb, W., et al. (2024). “**Time-frequency spectral optimization of underwater sonar acoustic returns for the purpose of characterization using spectral descriptors**”. In *Proceedings of the Acoustical Society of America 187th Conference*. Acoustical Society of America. https://acousticalsociety.org/asa-virtual-fall-2024/

Literature Cited (continued)

1. Kargle, Steven, and Kevin Williams. “Full Scale Measurement and Modeling of the Acoustic Response of Proud and Buried Munitions at Frequencies from 1-30 KHz .” *Serdp-Estcp*, SERDP, May 2012, serdp-estcp.mil/projects/details/ef9410dc-c3af-4cb4-a990-cd9aafe9acb4/mr-1665-project-overview.

2. Intrator, Nathan, et al. “Feature Extraction from Acoustic Backscattered Signals Using Wavelet Dictionaries.” *Proceedings of SPIE*, vol. 3079, no. 1, SPIE, 1997, pp. 183–90, doi:10.1117/12.280919.

3. Blanford, Thomas E, et al. “Development of an in-Air Circular Synthetic Aperture Sonar System as an Educational Tool.” *The Journal of the Acoustical Society of America*, vol. 145, no. 3, 2019, pp. 1935–1935, doi:10.1121/1.5102049.

4. Park, J.D., Blanford, T.E., Brown, D.C. “Late return focusing algorithm for circular synthetic aperture sonar data”. JASA Express Letters 1, 014801 (2021) https://doi.org/10.1121/10.0003063.

5. Bucaro, J. A., et al. “Broadband Acoustic Scattering Measurements of Underwater Unexploded Ordnance (UXO).” *The Journal of the Acoustical Society of America*, vol. 123, no. 2, Acoustical Society of America, 2008, pp. 738–46, https://doi.org/10.1121/1.2821794.

6. Morse, Scot F., and Philip L. Marston. “Backscattering of Transients by Tilted Truncated Cylindrical Shells: Time-Frequency Identification of Ray Contributions from Measurements.” *The Journal of the Acoustical Society of America*, vol. 111, no. 3, 2002, pp. 1289–94, https://doi.org/10.1121/1.1448518.

7. Anderson, Shaun D., Sabra, K., et al. “Time-Frequency Analysis of the Bistatic Acoustic Scattering from a Spherical Elastic Shell.” *The Journal of the Acoustical Society of America*, vol. 131, no. 1, 2012, pp. 164–73, https://doi.org/10.1121/1.3669995.

8. Sabra, Karim G. “Single Snapshot Spatial Array Processing Using Time-Frequency Distributions.” *The Journal of the Acoustical Society of America*, vol. 133, no. 5_Supplement, 2013, pp. 3436–3436, https://doi.org/10.1121/1.4806064.

9. Sabra, Karim G., and Shaun D. Anderson. “Subspace Array Processing Using Spatial Time-Frequency Distributions: Applications for Denoising Structural Echoes of Elastic Targets.” *The Journal of the Acoustical Society of America*, vol. 135, no. 5, 2014, pp. 2821–35, https://doi.org/10.1121/1.4871183.

Literature Cited

10. Bergqvist, Goran, and Erik G. Larsson. “The Higher-Order Singular Value Decomposition: Theory and an Application [Lecture Notes].” *IEEE Signal Processing Magazine*, vol. 27, no. 3, 2010, pp. 151–54, https://doi.org/10.1109/MSP.2010.936030.

11. Anden, Joakim, et al. “Joint Time-Frequency Scattering.” *IEEE Transactions on Signal Processing*, vol. 67, no. 14, 2019, pp. 3704–18, https://doi.org/10.1109/TSP.2019.2918992.

SERDP