SERDP

FINAL REPORT

Joint Beamforming and Automated Target Recognition for 3D SAS

Suren Jayasuriya
Arizona State University

August 2023

SERDP Project MR21-1334

DISTRIBUTION STATEMENT A
This document has been cleared for public release.

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
15-08-2023

2. REPORT TYPE
SERDP Final Report

3. DATES COVERED *(From - To)*
8/20/2021 - 8/19/2023

4. TITLE AND SUBTITLE
Joint Beamforming and Automated Target Recognition for 3D SAS

5a. CONTRACT NUMBER
W912HQ21P0055

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Suren Jayasuriya

5d. PROJECT NUMBER
MR21-1334

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Arizona State University
ORSPA
660 S MILL AVE STE 312 TEMPE
AZ 85281-3670

8. PERFORMING ORGANIZATION REPORT NUMBER
MR21-1334

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
SERDP Program Office
4800 MARK CENTER DRIVE
SUITE 16F16
Alexandria, VA 22350-3600

10. SPONSOR/MONITOR'S ACRONYM(S)
SERDP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
MR21-1334

12. DISTRIBUTION / AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Unexploded ordnance in cluttered, occluded, and buried environments present a major environmental and security risk for munitions response. Synthetic aperture sonar technology presents an opportunity to mitigate these risks but suffer from reduced fidelity and lowered recognition for potential targets in these challenging environmental conditions. This project develops a suite of algorithms to enhance synthetic aperture sonar beamforming for automated target recognition for buried unexploded ordnance detection. Two innovations are introduced: (1) a new differentiable beamforming method for volumetric synthetic aperture sonar reconstruction, and (2) a machine learning based tone mapper for improving automated target recognition. Both methods were developed using state-of-the-art advances in machine learning and computational imaging and were evaluated on exemplar data from a sub-bottom synthetic aperture sonar of various man-made objects distributed on a lakebed. Benefits to the DoD and the scientific community include new knowledge of how to merge the physics of underwater acoustics and beamforming with machine learning.

15. SUBJECT TERMS
Synthetic Aperture Sonar, Machine Learning, Automated Target Recognition

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Suren Jayasuriya
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	21	19b. TELEPHONE NUMBER *(include area code)*
UU	UU	UU			609-649-4366

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

List of Tables

List of Figures

List of Acronyms

Applied Research Laboratory at Penn State University (ARL-PSU)
Automated Target Recognition (ATR)
Backprojection (BP)
Buried Object Scanning Sonar (BOSS)
Convolutional Neural Networks (CNN)
Department of Defense (DoD)
Dynamic Range Compression (DRC)
Gradient descent (GD)
Inverse Tone Mapping Operator (iTMO)
Linear Frequency Modulation (LFM)
Machine Learning (ML)
Neural Radiance Fields (NeRF)
Peak Signal-to-Noise Ratio (PSNR)
Polar formatting algorithm (PFA)
Sediment Volume Search Sonar (SVSS)
Signal-to-Noise Ratio (SNR)
Strategic Environmental Research and Development Program (SERDP)
Synthetic Aperture Sonar (SAS)
Tone Mapping Operator (TMO)
Unexploded Ordnance (UXO)

Keywords

Synthetic Aperture Sonar, Machine Learning, Automated Target Recognition

Acknowledgements

We would like to thank SERDP and Dr. David Bradley for their support of this project. We would also like to thank graduate students Albert Reed and Gregory Vetaw who conducted most of the experiments and algorithm development that contributed to this report. We would also like to thank collaborators Dan Brown, Thomas Blanford, Benjamin Cowen, and David P. Williams from ARL-PSU as well as Juhyeon Kim and Adithya Pediredla from Dartmouth College for their assistance and co-authorship on several publications related to research from this project.

Most of these technical results are featured in two publications:

1) Reed, A., Juhyeon Kim, Thomas Blanford, Adithya Pediredla, Daniel C. Brown, Suren Jayasuriya, Neural Volumetric Reconstruction for Coherent Synthetic Aperture Sonar, ACM Transactions on Graphics (SIGGRAPH) 2023
2) Vetaw, G., Benjamin Cowen, Daniel C. Brown, David P. Williams, Suren Jayasuriya, Learning-based Tone Mapping to Improve 3D SAS ATR, International Geoscience and Remote Sensing Symposium (IGARSS) 2023
3) Albert Reed, Thomas Blanford, Daniel C. Brown, Suren Jayasuriya, SINR: Deconvolving Circular SAS Images Using Implicit Neural Representations, IEEE Journal of Selected Topics in Signal Processing (special issue) 2023
4) Gregory Vetaw, Albert Reed, Daniel C. Brown, Suren Jayasuriya, A 3D GAN Architecture for Volumetric Synthetic Aperture Sonar, MTS/IEEE Oceans 2021

Abstract

Introduction and Objectives: Unexploded ordnance in cluttered, occluded, and buried environments present a major environmental and security risk for munitions response. Synthetic aperture sonar technology presents an opportunity to mitigate these risks but suffer from reduced fidelity and lowered recognition for potential targets in these challenging environmental conditions. This project develops a suite of algorithms to enhance synthetic aperture sonar beamforming for automated target recognition for buried unexploded ordnance detection. The primary objectives for the project were to merge new advances in physics-based machine learning with three-dimension synthetic aperture sonar to achieve higher resolution, improved imagery, and better task performance in evaluation.

Technical Approach: This project introduces two primary innovations to advance the project's objectives: (1) a new differentiable beamforming method based on neural rendering for volumetric synthetic aperture sonar reconstruction, and (2) a machine learning based tone mapper for improving automated target recognition for buried unexploded ordnance detection. Both methods were developed using state-of-the-art advances in machine learning and computational imaging and were evaluated on exemplar data from a sub-bottom synthetic aperture sonar of various man-made objects distributed on a lakebed.

Results: Differentiable beamforming resulted in improved three-dimensional synthetic aperture sonar volumetric reconstructions including better visualizations of targets relative to clutter for a circular in-air system as well as a bistatic water-based system. Further, machine learning based tone mapping was shown to boost performance in both precision and recall for automated target recognition for partially and fully buried objects.

Benefits: Benefits to the DoD and the scientific community include new knowledge of how to merge the physics of underwater acoustics and beamforming with machine learning. New algorithms and methodologies have been developed and shown via experimental demonstrations as proof-of-concept. As shown in Figure 1 below, these algorithms can help synthetic aperture sonar systems better visualize and recognize three-dimensional targets on the seafloor.

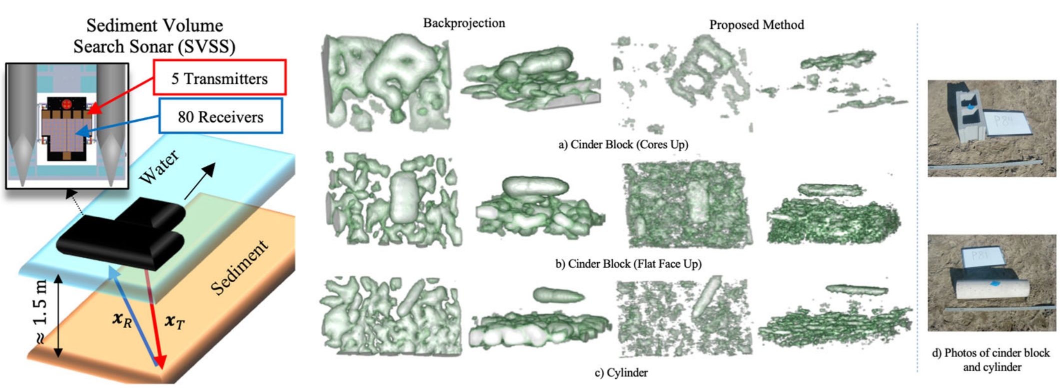

Figure 1: Differentiable beamforming for 3D synthetic aperture sonar - Our proposed algorithms enable better target visualization as compared to background/clutter in shallow water

Executive Summary

Objective

The technical objective for the project was to develop a suite of algorithms for enhanced beamforming and automated target recognition (ATR) for unexploded ordnance (UXO) in cluttered, occluded, and buried environments underwater. Physics-based knowledge including scattering models, target modeling, and environmental characterization would be merged with state-of-the-art machine learning-based ATR algorithms, to enhance their performance while providing interpretability to these detectors/classifiers. An explicit goal was to leverage the paradigm of differentiable programming where machine learning methods can be written alongside physics-based models, and thus complement the relative advantages of each approach. This differentiable programming would allow beamforming and ATR as an end-to-end pipeline to optimize target detection and classification, specifically for buried munitions. This objective would be accomplished over a one-year period to assess the feasibility of this approach, including how it can be scaled up and deployed in future efforts for in-the-field operation.

We aimed to develop solutions for goals issued in the 2017 SERDP Workshop on Acoustic Detection and Classification of Munitions in the Underwater Environment [1]. The main avenues of future work identified in the workshop include: (1) Improve sensor models to handle more complicated environments and munitions geometries including munitions in the vicinity of clutter so that munitions' acoustic response can be simulated at a high fidelity to augment data from experiment and demonstration efforts, and (2) Develop processing for 3D low frequency sonar to produce additional data products for use in classification. Both modeling and previous experimental results can be used as a starting point. As part of this effort carryout configuration studies, e.g., trade-offs in bandwidth, beam width, and number and arrangement of sensors. The proposed research aims to address both these thrusts through the integration of the physics of sonar imaging with modern machine learning ATR algorithms.

There were several technical research questions to be investigated over the course of the SEED project, including: (1) How can automated target recognition algorithms utilize both knowledge of the raw acoustic waveforms as well as the beamformed SAS imagery to better perform detection/classification? (2) How can the physics of acoustic scattering for partially-to-fully buried targets in sediment be leveraged in an ATR pipeline? (3) How do we design optimal beamforming for improved ATR performance for underwater munitions detection and classification? The SEED project aimed to tentatively answer these questions with pilot studies and prototype preliminary algorithms.

In particular, there were three main tasks to be pursued to answer the above research questions. The first task, Task 1: Differentiable Beamforming for 3D SAS, was to show proof-of-concept that differentiable beamforming could be accomplished for 3D SAS. The second task, Task 2: Coupling Beamforming with ATR, was intended to understand which beamforming operations and signal processing were optimal for ATR applications and establish the feasibility of using a differentiable beamformer in an end-to-end fashion. Finally, the last task, Task 3: Enhanced Performance for Buried UXO using Acoustical Scattering Models, aimed to enhance the performance of ATR models using physics-based knowledge to tailor algorithms for specific buried UXO detection and classification. Results for all these tasks were to be demonstrated via simulation, experimental prototyping in-air using a circular SAS sensor for lab bench testing, as well as deploying algorithms on real sensor data captured in the field from existing SERDP funded efforts, namely from the Sediment Volume Search Sonar (SVSS) [2].

Background

In recent years, synthetic aperture sonar (SAS) has proven to be the best method for capturing high resolution images of the seafloor [3, 4]. As such, SAS imaging has found important military applications for detecting and classifying under-water mines. Research has demonstrated the ability of low frequency sonar to generate images through SAS [5, 6], while high frequency sonars have indicated the presence of elastic responses for targets on the seafloor. Since SAS images are captured by systematically scanning large areas of the seafloor, this technology is particularly of interest for wide area coverage of detection and classification of military munitions for a wide variety of aquatic environments including ponds, lakes, rivers, and shallow coastal areas of interest to SERDP.

For SAS imaging, automated target recognition (ATR) algorithms are the most practical and efficient methods of detecting targets in these conditions [7]. There have been numerous advances in ATR technology, particularly using machine learning [8, 9]. Machine learning techniques have achieved state-of-the-art performance, but typically require large amounts of training data and have a "black-boxed" nature to their operation, where decisions made by the algorithms are not human interpretable and uncertainty/confidence estimation is difficult to quantify.

While there has been much success for SAS ATR for targets on the seafloor, the particular problem of partially to fully buried unexploded ordnance (UXO) is challenging. In general, high frequency imaging sonars can visualize 2D images of the water/sediment interface and proud targets while low frequency imaging sonars can create either 2D images of the interface and proud targets or 3D images of the sub-bottom and buried targets [1]. However, the detection of buried UXO is complicated by effects from the sediment interface, volume reverberation, and multipath effects. To mitigate these problems, typically 3D SAS imaging is required to penetrate the sediment layer. An example system is the BOSS system [10], although it has difficulty due to variations in object and radiometric geometries as well as loading effects of sediment due to burial depth. The Applied Research Laboratory at Penn State University (ARL-PSU) has developed a Sediment Volume Search Sonar (SVSS) capable of detecting surficial and buried UXO in water less than 5m in depth [2].

One main technical tool utilized in this SEED project is differentiable programming. Differentiable programming refers to the paradigm of writing algorithms which can be fully differentiated end-to-end using automatic differentiation for any parameter [11, 12]. This has been applied for audio [13], 3D geometry processing [14] and light rendering [15].

Materials and Methods

Our SEED project resulted in two main technical innovations that were investigated to help achieve the technical objectives listed in the Introduction. The first innovation is the use of neural rendering to achieve differentiable 3D SAS beamforming, which primarily achieved the objectives of Task 1. The second innovation is the design of enhanced signal processing and tone mapping for 3D SAS ATR, which achieved superior performance to other competing methods for buried UXO detection and satisfied the objectives of Tasks 2 and 3. In this section, we proceed to describe the methods corresponding to innovations proposed for Tasks 1-3, and the subsequent Results and Discussion section showcases the evaluation of these innovations in both simulation and real experimental data.

Task 1: Differentiable Beamforming for 3D SAS. Conventional image reconstruction for 3D SAS leverage time-domain backprojection where acoustic measurements are replica correlated or matched filtered, and then beamformed (e.g. focused) back into voxels using knowledge of the sonar's geometry relative to the scene. However, these algorithms do not typically handle occlusion, complex scattering effects, and sparse undersampling by the sonar. Further, such algorithms have not been made compatible with deep machine learning methods including neural networks that rely on gradient descent and other iterative solvers to update important parameters in the algorithm.

In this task, the goal was to design a differentiable beamformer that could be compatible with machine learning methods while still reconstructing 3D SAS data. We have achieved this goal through the design of a neural volumetric renderer. Our proposed reconstruction method (shown in Figure 2) consists of two main steps that roughly parallel traditional matched filtering and coherent backprojection steps that are undertaken in conventional 3D beamforming. First, we propose deconvolving given waveforms via an iterative deconvolution optimization rather than performing matched filtering. While matched filtering can be optimized through waveform design to realize a better ambiguity function in cross-correlation (i.e. better range compression), these techniques require a priori knowledge and do not work across a variety of sonar environments. In contrast, we present an adaptable approach to waveform compression where performance can be tuned via sparsity and smoothness priors, which we label pulse deconvolution.

Our second step is an analysis-by-synthesis reconstruction using an implicit neural representation (similar to NeRF in traditional view synthesis [16]). We use a network to predict complex-valued scatterers and use a differentiable forward model to synthesize complex sensor measurements in time. Traditional NeRF scene sampling methods are not directly applicable to our problem since we require sampling the scene points with constant time-of-flight, which correspond to ellipsoids with the transmitter and receiver as foci. Thus, we project rays from the transducer and sample rays at the intersection of ellipsoidal surfaces corresponding to measurement time bins and develop importance sampling methods to determine transmission probabilities for these rays and ellipsoidal surfaces. Finally, we explain how we implement our physics-based priors, such as a Lambertian scattering assumption, and regularization to our analysis-by-synthesis optimization. For technical details about the mathematical formulations and insights into our differentiable forward model, use of implicit neural representations for scene representation, ellipsoid sampling, and physics-based priors for optimization, we refer the reader to our journal paper on the subject [17].

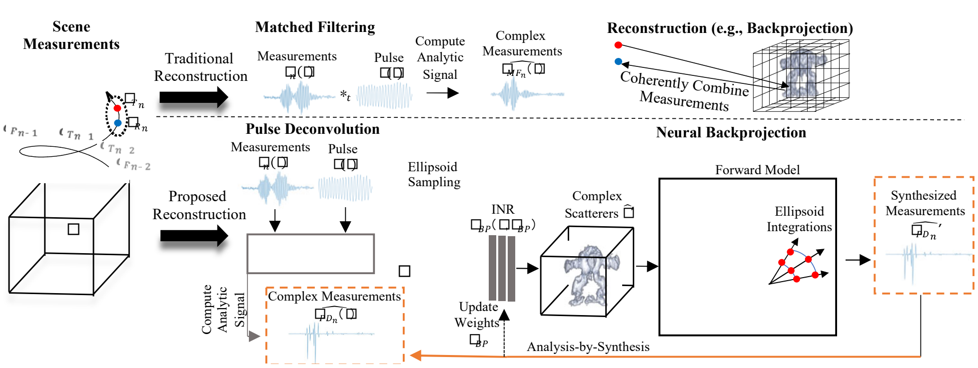

Figure 2: Block diagram of our differentiable beamforming method - Given measurements obtained from n sensor positions, the top row shows the traditional reconstruction pipeline and

the bottom row shows our proposed reconstruction pipeline. Our method applies pulse deconvolution, which is an optimization that deconvolves the transmitted pulse from measurements. We then propose neural backprojection, which uses a neural network to estimate the scene and synthesizes measurements using our differentiable forward model. The network is trained by minimizing a loss between synthesized and pulse deconvolved measurements.

Task 2: Coupling Beamforming with ATR. Our second task focused on how to couple beamforming techniques with ATR, particularly to address the challenges of highly cluttered ocean environments, especially for objects partially or completely buried in the sediment. In this task, we observed that a key stage for processing SAS data is the choice of dynamic range compression used in the pipeline. Conventional dynamic range compression (DRC) techniques such as log-compression, which is a type of tone mapping intended to appeal to the human visual system, can further obscure the sonar signatures of these already physically occluded objects and lead to suboptimal downstream ATR performance, particularly for convolutional neural networks (CNNs).

To address this, we have developed a novel machine learning-based approach for tone mapping sub-bottom SAS imagery as a pre-processing stage in the 3D SAS ATR pipeline. This learned tone mapping function can be jointly optimized with a CNN based ATR algorithm which we show in the Results and Discussion section of this report. Figure 3 shows our proposed end-to-end ATR pipeline. The differentiable pre-processing block takes in the raw amplitude imagery and outputs a tone mapped data cube which is then fed to the 3D CNN for classifying the data as target or clutter. Conventionally, the pre-processing block would perform DRC via the logarithm function with a data normalization method applied afterwards. In the 3D case, implicit assumptions in the normalization procedure with additional data transformations would have to be made.

We propose to make this pre-processing stage in the SAS ATR pipeline learnable with two ML variations of it. The first variation feeds dynamic range compressed data to an inverse tone mapping layer which can either be a global inverse tone mapping function (e.g. sinh, cosh, gamma function) applied uniformly across the data cube or a learned inverse tone mapping function (iTMO) that uses a spatially varying 3D cube of differentiable weights applied across the data cube. The combination of DRC (tone mapping) and inverse tone mapping yields a composite function that tone maps the original raw amplitude data. In the second variation, we utilize machine learning to learn the entire tone mapping directly from the raw amplitude data, which we label as TMO. We utilize a two-layer differentiable tone mapping for this purpose. This variation achieves the highest performance for classification in our experimental results. For more technical description for this algorithm and its variations, we refer the reader to our conference paper that describes the method in complete detail [18].

Task 3: Enhanced Performance for Buried UXO using Acoustical Scattering Models. Our third task focused on enhancing the benefits of Tasks 1 and 2 by focusing primarily on buried UXO and their scattering physics which can be leveraged to improve both visualization and ATR performance. While there was no explicit innovation created to address this task during the SEED project, we do note that our learning-based tone mapping achieved superior performance on buried targets (as we will show in the next section of the report).

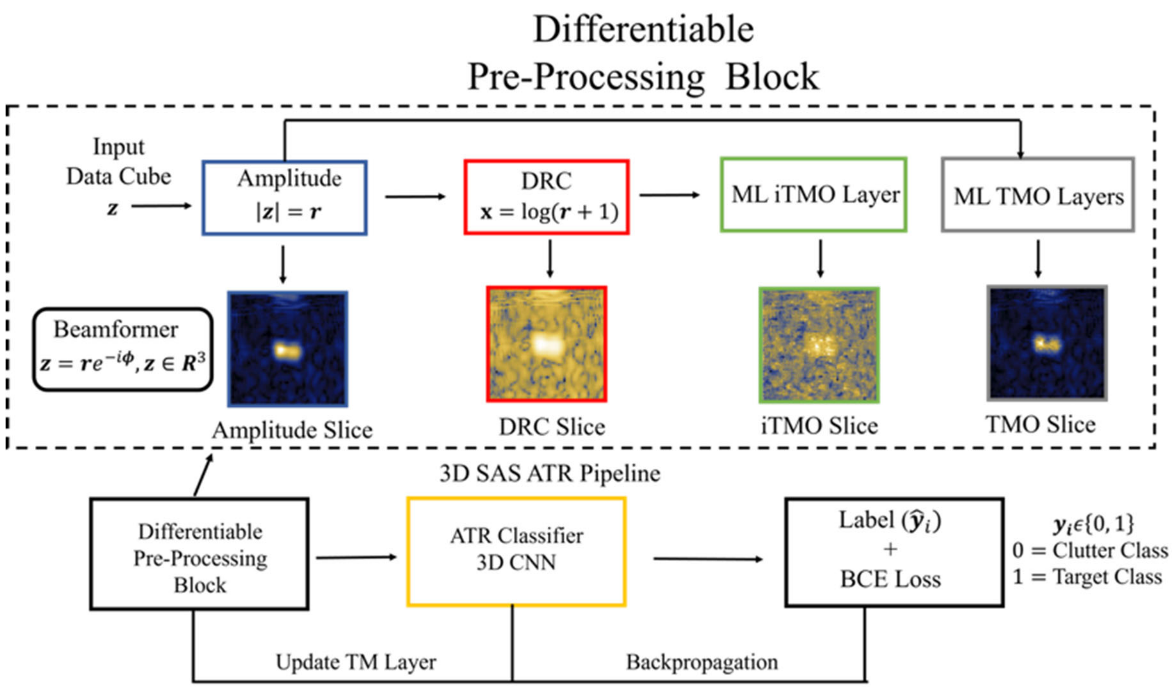

Figure 3: End-to-end 3D SAS ATR pipeline - A 2D slice drawn from each representation of a data cube along the depth channel to highlight how the operations of DRC, inverse tone mapping (iTMO), and tone mapping (TMO) can transform the SAS imagery in the processing chain.

Results and Discussion:

Task 1: Differentiable Beamforming for 3D SAS - Results. In this SEED project, we have successfully developed a differentiable beamforming method for 3D SAS reconstruction leveraging our analysis-by-synthesis pipeline. To validate our method, we tested on simulated data as well as two real-data sources for proof-of-concept. Our simulation results leverage a time-of-flight renderer to simulate direct bounce of the acoustic wave in a scene, and allows us to quantitatively evaluate performance with respect to ground truth scenes. Our first real data source, AirSAS, is a circular in-air SAS in the laboratory for rapidly prototyping reconstructions. Our other real-data source uses measurements captured from a lakebed, the Foster Joseph Sayers Reservoir in Pennsylvania, using the SVSS. Our SVSS results validate our method's applicability to underwater environments and bistatic transducer arrays.

In simulation, we compare our proposed method against backprojection, the polar formatting algorithm, and gradient descent on simulated scenes measured with an LFM chirp of 20KHz center frequency and 20KHz bandwidth, as well as show reconstructions under different noise levels as shown in Figure 4. Table 1 presents quantitative metrics averaged over reconstructions from 8 different objects. Considering the PSNR metric, our approach offers 2 dB improvement over other methods.

In Figure 5, you can see examples of experimental results for SAS in-air and water using our methods. In all cases, our methods provide better reconstructions than traditional SAS reconstruction algorithms. We refer readers to the journal paper for more extensive results and comparisons [17].

Task 1 Assessment: Our self-assessment on the progress made in Task 1 is largely positive. We have been able to publish an iterative, analysis-by-synthesis based method [17], which is compatible with neural networks, and successfully showed 3D reconstruction for SVSS data. This framework can be optimized further with additional priors in optimization and adding more pre-trained neural networks for ATR downstream applications. One bottleneck for the method is the latency (1-2 hours per reconstruction), but there are some methods in the literature to accelerate scene reconstruction that can be pursued as follow-up research.

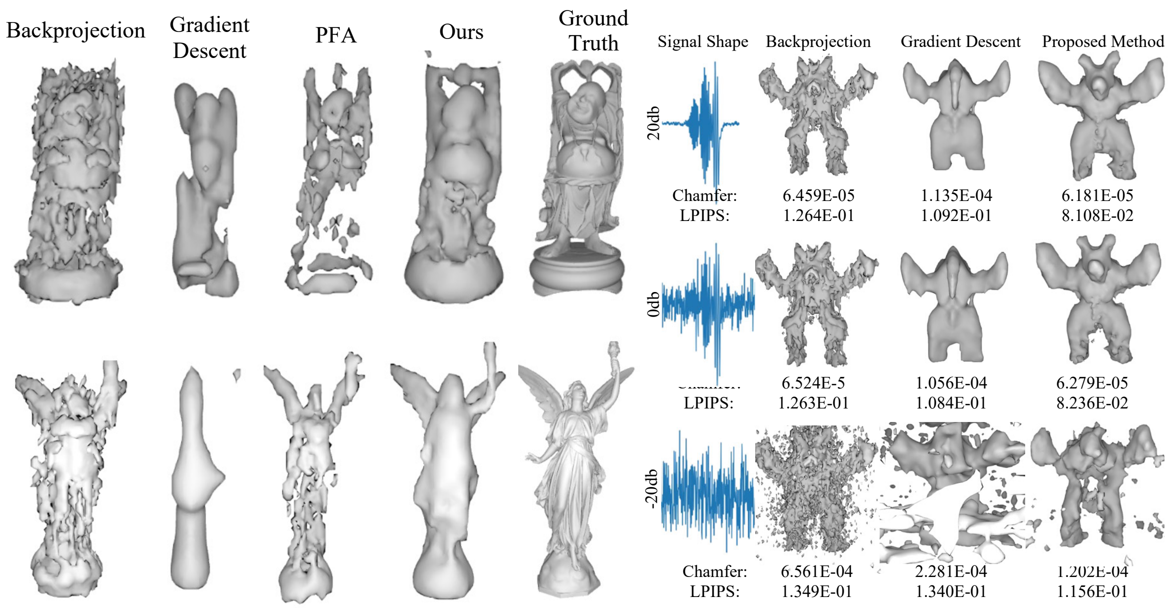

Figure 4: Qualitative results for neural reconstructions on simulated data - (Left) Reconstructions from simulated measurements using backprojection, gradient descent, the polar formatting algorithm (PFA), and our method. Compared to other methods, our reconstructions more accurately match the ground truth geometry. (Right) Simulation results using a 20 kHz LFM showing the reconstructed meshes of an armadillo object at three noise levels. Our method performs decently well even at -20 dB signal-to-noise-ratio on the raw acoustic measurements.

Metric	Chamfer ↓	IOU ↑	LPIPS ↓	PSNR ↑	MSE ↓
BP	1.36E-04	0.2928	0.1215	15.783	5.55E-03
GD	2.21E-04	0.4309	0.1236	15.117	6.13E-03
PFA	2.13E-04	0.3586	0.1238	15.048	6.64E-03
Ours	**1.12E-04**	**0.5194**	**0.0988**	**17.918**	**3.99E-03**

Table 1: Quantitative results for neural reconstructions for simulated data - Simulation results showing the average quantitative metrics for Backprojection (BP), Gradient descent (GD), the Polar formatting algorithm (PFA), and our reconstructions across 8 different meshes.

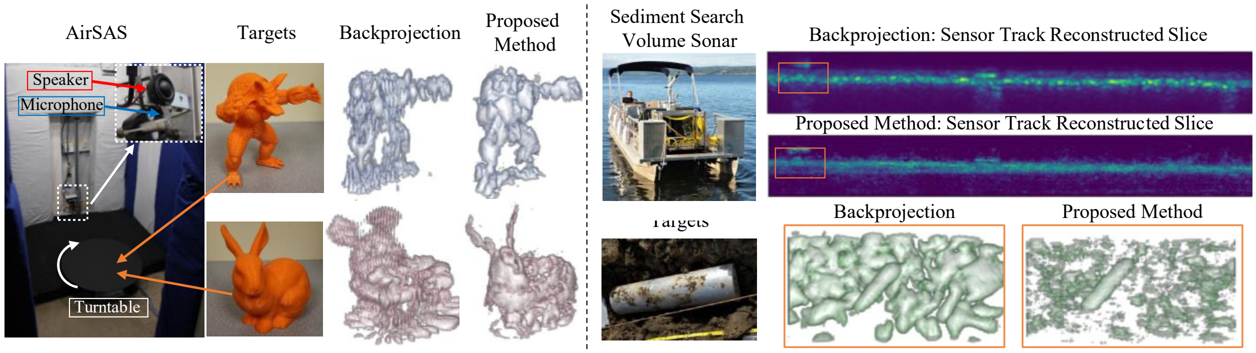

Figure 5: Experimental results for neural reconstructions on real data - We propose an analysis-by-synthesis optimization that leverages techniques from neural rendering to optimize coherent reconstructions of SAS volumetric scenes. We demonstrate our approach on an in-lab circular SAS in air (AirSAS) and in-water bistatic SAS, the sediment volume search sonar (SVSS). On the left side of the figure, we show the AirSAS, 3D printed targets, and reconstructions obtained using backprojection and our proposed method. On the right side, we show 2D maximum intensity projections (MIPs) of the SVSS track and 3D reconstructions of targets highlighted in orange. In many cases, our method produces better reconstructions than traditional SAS reconstruction algorithms, such as backprojection. SVSS hardware photos courtesy of [2].

Task 2: Coupling Beamforming with ATR - Results. We now discuss the results of our learning-based tone mapping coupled with a downstream CNN for ATR. Allowing the CNN network to make predictions based on tone-mapped data as opposed to the traditional log-compressed representation has a direct effect on the classification performance. Table 2 shows area under the curve for precision-recall (AUC-PR) values from testing on the SVSS June dataset with an imbalanced ratio of 3-to-1 after applying 8 different TMOs on the DRC data cubes. The table shows that tone mapping raw or DRC data increases ATR classification performance. Further, the ML approach yields the highest performance for tone mapping buried objects.

For qualitative results, Figure 6 shows 2D slices of a partially buried object. The dynamic range compressed imagery represents the object masked in visual clutter. The global iTMOs attempt to preserve the acoustic energy of the object from the DRC imagery and present a more appealing imagery to human observers than the other forms of the imagery. The ML inverse tone mapping reduces the clutter on a level of a lower order gamma function but does not preserve the original pixel intensity of the object. The ML tone mapping compresses the clutter around the object and roughly preserves high level object details.

Data Representation	Precision-Recall AUC
Raw Data	0.693
DRC	0.785
DRC + $\gamma = 1$	0.806
DRC + $\gamma = 2$	0.835
DRC + Sinh	0.824
DRC + Cosh	0.858
Williams and Brown [9]	0.838
DRC + 1 iTMO Layer	0.851
Raw + 2 TMO Layers	**0.870**

Table 2: AUC-PR scores for 3D SAS ATR on SVSS dataset – We show quantitative AUC-PR scores for our method with respect to other methods, highlighting the performance benefits of learning tone mapping layers for 3D SAS ATR.

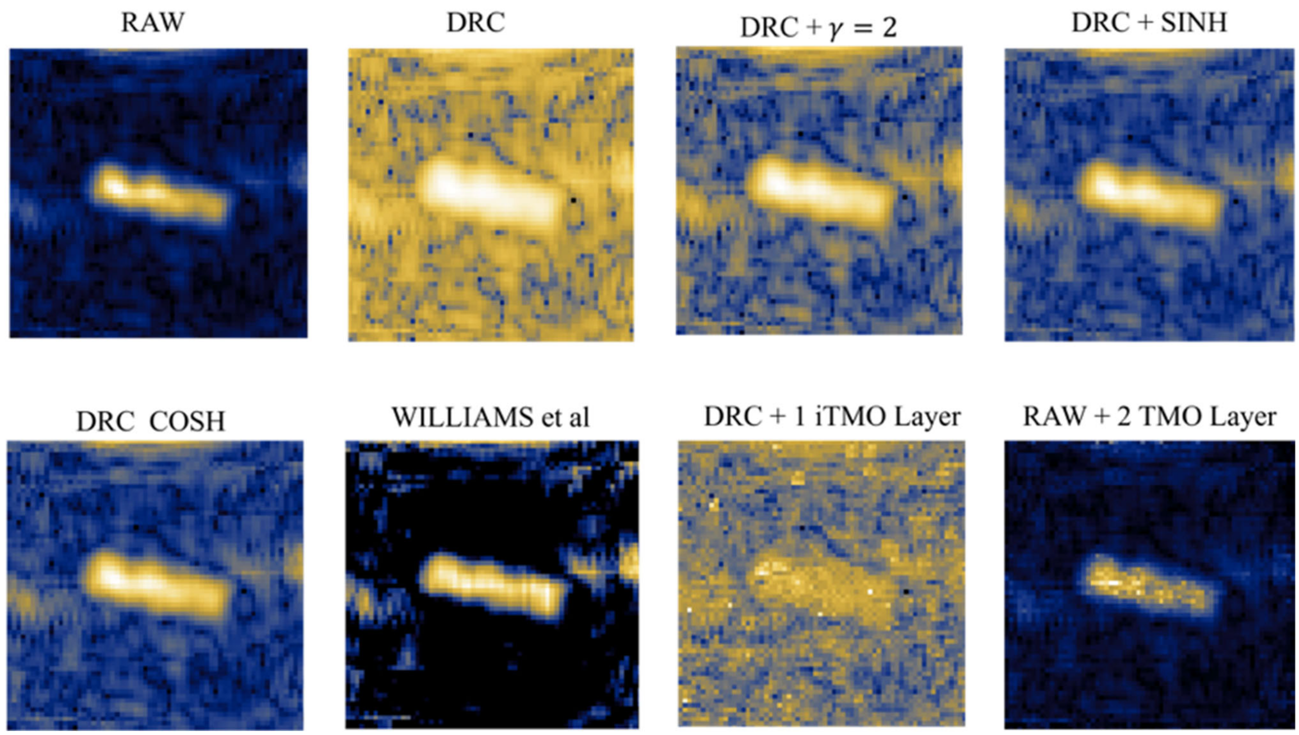

Figure 6: Visualization of tone mapping for 3D SAS data - 8 different representations of a buried object captured in the June 2019 SVSS dataset are shown. Note how the DRC operation submerges the object in sediment clutter as opposed to the tone mapped imagery.

Task 2 Assessment: Our self-assessment for the success of Task 2 objectives is more mixed. We have successfully shown that tone mapping is an important pre-processing step that is key to ATR performance for 3D SAS and developed a technique to learn this tone mapping within the CNN. This enhanced performance as well as visualization of the data. However, we were not able to accomplish joint training of the differentiable beamforming in Task 1 with the ATR from Task 2, as this requires building a large software pipeline and extensive experimentation to get these methods to work together. We identify this as a potential avenue of further research to yield the benefits of differentiable beamforming.

Task 3: Enhanced Performance for Buried UXO using Acoustical Scattering Models – Results. For Task 3, we have conducted preliminary studies on the effectiveness of our ATR methods for buried targets. Our learning-based tone mapping strategy seemed to effectively work for partially buried and buried targets, achieving higher AUC-PR scores in Table 2 compared to the conventional DRC tone mapping. For our differentiable beamforming, our current formulation does not consider advanced scattering models for buried targets or multi-path reflections. For instance, Figure 7 shows that effects such as elastic scattering are not captured in our method's reconstructions as compared to the conventional backprojection.

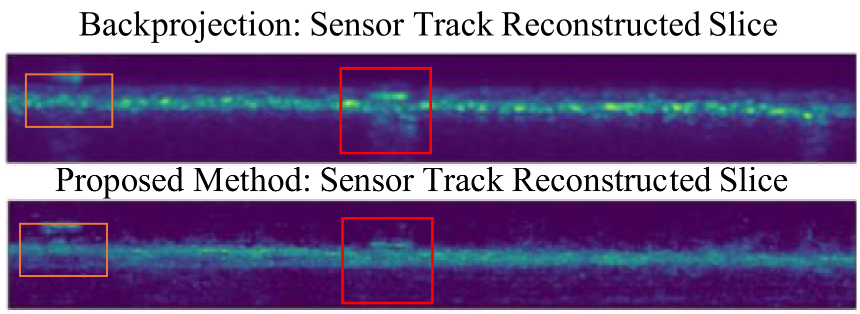

Figure 7: Elastic Scattering – Note that our differentiable beamforming method fails to capture elastic scattering from a target in the red box, while it does resolve higher spatial resolution and detail in the orange box.

Task 3 Assessment: Our success on Task 3 has been limited and requires further study and research. We have clearly identified a link between tone mapping and pre-processing the 3D SAS data and the resulting ATR performance for buried objects. However, the acoustic scattering physics of these buried objects is more advanced than our ray-based models leveraged for differentiable beamforming, and thus our method is unable to capture these effects in reconstruction.

Conclusions and Implications for Future Research: We have introduced two new algorithms to perform differentiable beamforming via neural rendering and ML-based tone mapping respectively for 3D SAS. The feasibility of both these approaches was not certain at the start of the project, and the demonstration of these methods working in real SVSS field data shows the promise of the proposed technical approach. Further, there have been some technical papers published because of this research which disseminates the work widely to both the SERDP community as well as the broader scientific audience.

There are several avenues for potential next steps in realizing these solutions for buried UXO detection. First, these algorithms have not been integrated successfully with state-of-the-art ATR algorithms such as the Mondrian detector or Tiny CNNs [9], which may help boost performance even further. Second, our algorithms can be adapted to a wide variety of 3D SAS platforms (such as the BOSS system) in addition to the SVSS. Finally, there is still room to model more complex physical effects such as resonance and elastic scattering, as well as multipath effects that make the detection of buried UXO more difficult. Follow-up research can help extend these initial efforts in this general direction to validate the robustness of the methods in these challenging use cases.

Literature Cited:

1. SERDP Workshop on Acoustic Detection and Classification of Munitions in the Underwater Environment, Final Report, April 2018.

2. D. Brown, SERDP project MR-2545, “Sediment Volume Search Sonar Development” active project.
3. M. P. Hayes and P. T. Gough, “Synthetic aperture sonar: A review of current status,” IEEE Journal of Oceanic Engineering, vol. 34, pp. 207–224, July 2009
4. R. E. Hansen, “Introduction to synthetic aperture sonar,” in Sonar Systems (N. Z. Kolev, ed.), ch. 1, Rijeka: IntechOpen, 2011.
5. D. D. Sternlicht, J. E. Fernandez, R. Holtzapple, D. P. Kucik, T. C. Montgomery, C. M. Loeffler. “Advanced Sonar Technologies for Autonomous Mine Countermeasures,” Proceedings MTS/IEEE OCEANS, pp. 1-5, October 2011.
6. W. Jans, D.D. Sternlicht, K. L. Williams, M. D. Richardson, “Emerging Technologies in Underwater Munitions Mapping,” Proceedings of the Underwater Acoustics Conference and Exhibition (UACE) 2017, pp. 669-676, Sept. 2017
7. Jason Stack "Automation for underwater mine recognition: current trends and future strategy", Proc. SPIE 8017, Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVI, 80170K (23 May 2011)
8. Bellettine & Pinto (2009) ‘Design and Experimental Results of a 300-kHz Synthetic Aperture Sonar Optimized for Shallow-Water Operations’, IEEE J. Oceanic Eng. 34(3), 285—293.
9. D. Williams, "On the Use of Tiny Convolutional Neural Networks for Human-Expert-Level Classification Performance in Sonar Imagery," IEEE Journal of Oceanic Engineering, in press, February 2020.
10. J. Sara, D. Sternlicht, SERDP project MR-2752, “Next Generation Buried Object Scanning Sonar (BOSS) for detecting buried UXO in Shallow Water,” active project.
11. Baydin, A.G., Pearlmutter, B.A., Radul, A.A., Siskind, J.M.: Automatic differentiation in machine learning: a survey. The Journal of Machine Learning Research 18(1), 5595–5637 (2017)
12. Wang, F., Decker, J., Wu, X., Essertel, G., Rompf, T.: Backpropagation with callbacks: Foundations for efficient and expressive differentiable programming. In: Advances in Neural Information Processing Systems. pp. 10180–10191 (2018)
13. Engel, J., Hantrakul, L.H., Gu, C., Roberts, A.: Ddsp: Differentiable digital signal processing. In: International Conference on Learning Representations (2020)
14. Ravi, N., Reizenstein, J., Novotny, D., Gordon, T., Lo, W.Y., Johnson, J., Gkioxari, G.: Pytorch3d. https://github.com/facebookresearch/pytorch3d (2020)
15. Nimier-David, M., Vicini, D., Zeltner, T., Jakob, W. Mitsuba 2: A retargetable forward and inverse renderer. Trans. on Graphics (Proceedings of SIGGRAPH Asia) 38(6) (2019).
16. Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 65(1), 99-106.
17. Reed, A., Juhyeon Kim, Thomas Blanford, Adithya Pediredla, Daniel C. Brown, Suren Jayasuriya, Neural Volumetric Reconstruction for Coherent Synthetic Aperture Sonar, ACM Transactions on Graphics (SIGGRAPH) 2023
18. Vetaw, G., Benjamin Cowen, Daniel C. Brown, David P. Williams, Suren Jayasuriya, Learning-based Tone Mapping to Improve 3D SAS ATR, International Geoscience and Remote Sensing Symposium (IGARSS) 2023

Supplemental Documentation: References [17] and [18] can be found on the SERDP MR21-1334 project in the Publications folder under the file names: Siggraph_2023_Final.pdf and IGARSS_2023.pdf respectively.