Spring 2016

Characterization of active sonar targets

Daniel Schupp-Omid
University of Iowa

Characterization of Active Sonar Targets

By:
Daniel Schupp-Omid

A thesis submitted in partial fulfilment of the
requirements for the Master of Science
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

May 2016

Thesis Supervisor: Assistant Professor Ananya Sen Gupta

Graduate College

The University of Iowa

Iowa City, Iowa

CERTIFICATE OF APPROVAL

MASTER'S THESIS

This is to certify that the Master's thesis of

Daniel Schupp-Omid

has been approved by the Examining Committee for
the thesis requirement for the Master of Science degree
in Electrical and Computer Engineering at the May 2016 graduation.

Thesis Committee: _____
Ananya Sen Gupta

_____
Mark Andersland

_____
Anton Kruger

Dedicated to my wife, Chloe

Acknowledgements

Thanks go to the University of Washington's Applied Physics Laboratory for providing the sonar target data, to Ivars Kirsteins of NUWC for his collaboration, and, of course, to Ananya Sen Gupta for her support, guidance, and collaboration.

Abstract

Characterization of Active Sonar Targets

by Daniel Schupp-Omid

We present a method of characterization of active sonar target response that makes use of a combination of physics-driven Gabor dictionaries and ellipsoidal feature manifolds. The key innovation is the combination of the useful characteristics of the Gabor wavelet with an empirical approach to dictionary selection for underwater acoustics. The feature manifolds are represented as a set of ellipsoids, each of which geometrically encompasses a unique physical characteristic of the target's response. We demonstrate over experimental field data that this combination of approaches yields significant clustering behavior and high accuracy on real-world datasets.

Public Abstract

Characterization of Active Sonar Targets

by Daniel Schupp-Omid

The problem of characterization of active sonar target response has important applications in many fields, including the currently cost-prohibitive recovery of unexploded ordinance on the ocean floor. We present a method for recognizing these objects using a multidisciplinary approach that fuses machine learning, signal processing, and feature engineering. In short, by taking inspiration from other fields, we solve the problem of object recognition in shallow water in an inexpensive way. These techniques add to the body of explored knowledge in the field of active sonar processing and address real-world problems in the process.

Table of Contents

List of Figures viii

Abbreviations x

1 Introduction 1
1.1 Problem Statement . . . 1
1.1.1 Challenges . . . 1
1.1.1.1 Weak Ground Truths . . . 2
1.1.1.2 Intra-Target Interference . . . 2
1.1.1.3 Boundary Conditions . . . 2
1.1.2 Existing Approaches . . . 3
1.1.2.1 Acoustic Propagation Studies . . . 3
1.1.2.2 Information Bounds . . . 4
1.1.2.3 Data Driven Techniques . . . 5
1.1.2.4 Bio-Inspired Techniques . . . 7
1.2 Applicability . . . 7
1.2.1 Autonomous recovery of UXO . . . 8
1.3 Relevant Technologies and Inquiries . . . 9
1.3.1 Contributions . . . 10

2 Background Information and Theory 11
2.1 K-Nearest Neighbor Clustering . . . 11
2.1.1 The Mahalanobis Distance . . . 12
2.2 Support Vector Machines . . . 12
2.3 Neural Networks . . . 13
2.3.1 Limitations and Potential Improvements . . . 14
2.3.2 Our Implementation . . . 14
2.4 Gabor Filters . . . 15
2.5 Acoustic Physics . . . 16

3 Synopsis of Technical Approach and Data-Driven Investigations 17
3.1 Characterization of UXO Using Active Sonar . . . 17
3.2 Using Ellipsoids as Sonar Feature Manifolds . . . 18
3.3 Using Gabor Wavelets to Derive Sonar Features . . . 19
3.4 Distinguishing Characteristics . . . 20
3.5 Scope of Proposed Work . . . 21

4 Ellipsoid Features 22
4.1 Introduction . . . 22
4.2 Technical Approach . . . 23
4.2.1 Feature Separation and Classification . . . 25
4.2.2 Feature Selection from Overlapping Features . . . 27

4.3 Results . . . 27
4.3.1 Evaluation against Traditional Correlation-Based Matching . . . 30
4.4 Concluding Remarks . . . 32

5 Gabor Feature Investigation 34
5.1 Introduction . . . 34
5.2 Technical Approach . . . 35
5.2.1 Gabor Features . . . 35
5.3 Results . . . 36
5.3.1 Transformation Results and Examples . . . 36
5.3.2 SVM Results and Examples . . . 39
5.4 Conclusions . . . 41
5.4.1 Physics-Based Dictionaries . . . 42
5.4.2 Boosted Classifiers . . . 42

6 Combined Investigation 43
6.1 Introduction . . . 43
6.2 Technical Approach . . . 44
6.3 Results . . . 48

7 Remarks and Conclusions 50
7.1 Collaboration in Works Herein . . . 50
7.2 Scope . . . 50
7.3 Closing statement . . . 50
7.4 Acknowledgment . . . 51

Appendix Gabor Transform Examples 52
A.1 $\lambda = 2$ . . . 52
A.2 $\lambda = 10$ . . . 64
A.3 $\lambda = 4$ . . . 75

Bibliography . . . 86

List of Figures

Figure	Description	Page
3.1	Summary of region of interest and location of investigations at the intersections of acoustic physics, machine learning, and feature engineering.	18
4.1	Peak Finding: Extrema are overlaid (in red) on the cross-range frequency response.	24
4.2	Clustering with Overlay: Clusters overlaid back on the cross-range frequency response and shown from top-view. The centroids of each cluster are marked by green squares.	25
4.3	Ellipsoids: An example of what the ellipsoidal features look like in the 3D sensor position/frequency/intensity space. The orientation noted in the title is of the target, not the sensor. Note that these ellipsoids are well localized in frequency, and so appear large in this image. Compare to ref Fig. 4.6 to see how the ellipsoids project on a single sensor orientation.	25
4.4	Steel UXO (Target A) vs AL UXO (Target B): Feature regions in red separate Steel UXO from Aluminum UXO.	28
4.5	Steel UXO (Target) vs Rock (Non-target): Feature regions (in red) separate Steel UXO from a rock.	28
4.6	Frequency response of a direct hit with Steel UXO ellipsoids overlaid.	29
4.7	Frequency response of a glancing hit with Steel UXO ellipsoids overlaid.	30
4.8	Frequency response of a glancing hit on a rock with Steel UXO ellipsoids overlaid.	31
4.9	Normalized cross correlation of direct-path steel UXO frequency response with rock response.	31
4.10	Comparison of template matching using ellipsoidal features and normalized cross correlation.	32
5.1	Example Gabor kernel.	37
5.2	First rock example.	37
5.3	Second rock example.	38
5.4	First steel UXO example.	38
5.5	Second steel UXO example.	39

5.6 Confusion matrix of $\lambda = .1$, $\theta = 90°$, and $\sigma = 1/6$, with 5-fold cross validation accuracy of 64.61%. 39

5.7 Confusion matrix of $\lambda = .25$, $\theta = 90°$, and $\sigma = 1/6$, with 5-fold cross validation accuracy of 64.44%. 40

5.8 Confusion matrix of $\lambda = .1$, $\theta = 45°$, and $\sigma = 1/10$, with 5-fold cross validation accuracy of 64.32%. 40

5.9 Confusion matrix of $\lambda = .01$, $\theta = 45°$, and $\sigma = 1$, with 5-fold cross validation accuracy of 64.53%. 41

6.1 Ellipsoids generated from the filtered acoustic color features that distinguish between a steel and aluminum UXO. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids. 44

6.2 Ellipsoids generated from the unfiltered acoustic color features that distinguish between a steel and aluminum UXO. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids. 45

6.3 Ellipsoids generated from the filtered features to distinguish an aluminum UXO from a rock. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids. 46

6.4 Ellipsoids generated from the filtered features to distinguish an aluminum UXO from a steel UXO. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids. 46

6.5 Ellipsoids generated from the filtered features to distinguish an steel UXO from a rock. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids. 47

6.6 Confusion matrix for the direct classification approach used in Chapter 5. Accuracy of 64.5%. 48

6.7 Confusion matrix for the results of a neural network trained on a single feature set. Accuracy of 84.3%. 49

6.8 Confusion matrix for the results of a neural network trained using four different Gabor feature filters. Accuracy of 95%. 49

Abbreviations

UXO **U**n E**x**ploded **O**rdinance

AUV **A**utonomous **U**nderwater **V**ehicle

Chapter 1

Introduction

1.1 Problem Statement

The detection and classification of active sonar targets remains an ongoing area of research. This problem has applications in many facets of society, including military, environmental, economic, and emergency response. Specific statements of the problem include autonomous identification of UXO (unexploded ordinance) for retrieval, underwater SLAM (simultaneous localization and mapping), and wildlife survey.

This topic is particularly interesting to me, because of the real-world applications, and the variety of learning techniques that need to be leveraged to build a robust solution to the problem. Target detection and classification has been an active field of study for decades and continues to generate wide-range interest for a variety of open challenges. These are discussed below.

1.1.1 Challenges

Characterization of targets is confounded by several factors:

- Weak ground truths for targets and nonexistent ground truths for unknown targets makes conventional learning techniques untenable [1–5]
- Intra-target interference (in the case of multi-target scenarios) can cause overlapping resonance components that challenges conventional learning techniques [4–7]

- Artifacts due to boundary and channel effects increase target response complexity, making the problem more difficult in general

1.1.1.1 Weak Ground Truths

Conventional learning techniques such as GMM rely on a wealth of data for training. Many learning techniques also suffer from a tendency to overtrain, or are otherwise ill suited to characterization of noisy data. Conventional wisdom would indicate that the data can be massaged and cleaned up prior to applying traditional learning techniques, but the nature of cluttered environments underwater make that desire untenable.

1.1.1.2 Intra-Target Interference

High-activity ocean environments such as a fishing port or busy harbor are full of clutter, and techniques intended for deployment in real-world settings will need to be robust enough to handle the distortions caused by the complex multipath arrivals confounded by non-target objects and clutter. Successful modeling techniques will need to be able to identify meaningful feature groups in spite of the possibility of additional feature groups from additional targets.

1.1.1.3 Boundary Conditions

The boundary conditions imposed on the sonar problem are particularly insidious. Aside of the already sizeable concerns due to the complex surface presented by the sea floor and surface, the boundaries are constantly shifting. This makes characterization not only difficult, but impossible over any time period long enough to be useful for prediction. This additional clutter makes target characterization significantly more difficult, as a

successful method would have to be able to extract meaningful features from stochastic clutter.

1.1.2 Existing Approaches

Current state-of-the-art in buried target detection (using active or passive sonar), signal recovery and source classification techniques combine rich advances in four complementary directions:

- (i) Underwater acoustic propagation studies [8–10], including noise models [11] and acoustic backscatter studies that may be harnessed to achieve buried source localization and signal recovery on physics-based principles;
- (ii) Information bounds and tradeoffs [3, 12–15] that provide mathematical limits on when a buried signal may be recoverable against background noise and clutter;
- (iii) Adaptive signal processing [1–5, 16], learning algorithms [17–20] and object classification techniques (ref. e.g. [4–7] and references therein) that perform data-driven signal recovery. This includes array-based methods) and other techniques that acquire side information [4, 5, 7] to recover buried signals.
- (iv) Bio-inspired sonar signal processing techniques (e.g. [21] among others).

1.1.2.1 Acoustic Propagation Studies

A variety of methods have been put to use in the ongoing attempt to characterize the physical phenomena at play in the acoustic backscattering system. Many attempts at characterization have attempted to characterize the impulse response of a spherical shell. Sessarego et al attempted to characterize the surface waves by analyzing a spherical shell

using the Wigner-Ville distribution [22]. Our work presented in [23] analyzes a similar experimental setup, but develops distributions based on resonance topology groupings, with encouraging preliminary results.

Badiey et al attempted characterization based on a model of multipath arrivals [8]. Baggenstoss developed a model of specular decomposition to yield range and Doppler information by fusing data from multiple transmitted waveforms, specifically, LFM (linear frequency-modulated) and CW(continuous wave) [10]. Early investigations on our part indicated good results for this approach [23], which will proceed in a complementary direction to the provided investigations. In the cited investigation, the model was built on a combination of Rayleigh and WG (whispering gallery) waves generated by the resonance of the target itself rather than using different kinds of transmitted waveforms.

Carter and Dobeck [24] developed a parametric model by modeling the target as an superposition of point scatters. The model simply sums the simplified delta functions, then applies a wavelet transform for feature extraction. This method was targeted for distinguishing a mine from a rock [24]. This paper discussed empirical derivation of wavelet parameters from training data, which influenced the approach taken in the wavelet investigation of chapter 5, though the approaches differed significantly. Carter and Dobeck acknowledge that the model is not perfect, but present compelling evidence that it sufficiently captures enough meaningful features for useful classification.

1.1.2.2 Information Bounds

Significant work has gone into the study of the information limits of the target detection system. Culver and Bradley published a work developing the relationship between signal bandwidth and frequency correlation [12]. This work had important consequences for

frequency-based feature extraction by putting a lower limit on the necessary bandwidth for reliable characterization.

Meng and Buck characterized rate distortion on passive sonar [13, 14]. This work highlights the extreme challenges that passive sonar present for the general problem of object recognition. These challenges are not addressed by the proposed methods herein, as we've opted to make use of active sonar.

Sharma, Buck, and Simmons described the issue of the tradeof between detection versus classification as signal power increases [3]. In short, increasing power has the advantage of making object detection easier while also reducing resolution needed for object classification. The paper builds a model of a VRDR (variable resolution and detection receiver) that is useful for addressing these issues in practice. The specifics of this issue are not addressed by the proposed work, as we've adopted the dataset built by the Applied Physics Laboratory at the University of Washington.

1.1.2.3 Data Driven Techniques

The proposed works fit most within the realm of data-driven signal processing techniques, and are therefor most appropriately compared to other techniques within this category. In Hollinger et al [6], a method of SLAM using surface mapping via Gaussian approximation of surface points is proposed. This approach has some similarities to the investigations that follow, in that they model geometric clusters and use a modeling function to approximate them and develop models of uncertainty and confidence. The primary differentiation of this work against existing art is combining geometric features representative of underlying acoustic physics with sophisticated machine learning and signal processing techniques. The works differ very greatly in execution, however. The

proposed works build a geometric model of the feature vectors generated by seeking extrema in the frequency domain, rather than the space domain. The proposed works further use a Gabor wavelet as a basis function rather than a Gaussian, leveraging the periodicity for improved modeling. The two works also differ in the problem solved, as [6] is targeted towards building models of macro surfaces from the perspective of a small observer, the proposed works here focus on the classification of discrete objects.

Maussang et al [5] tackles the same problem as the proposed work (detection/classification of underwater mines), and uses another data-driven approach. This work highlights the relative difficulty of the problem due to low signal-to-noise. The approach proposed therein builds a model based entirely on the statistical properties of the sonar images, attempting to tackle the problem at the pixel level. The approach proposed here uses the frequency content instead, and aims to deal with the problem of glancing incident angles on the target more directly by building models that take the complete range of sensor angles into account. This problem is not addressed by [5], though their approach to feature grouping provided some inspiration for some of the proposed methods. The density groupings proposed therein are somewhat similar to the ellipsoidal features proposed here, though their work does not evaluate groupings in the geometric approach presented in Chapter 4.

Battiti [20] discusses the process of feature selection for supervised neural net learning. The article discusses the process of feature selection via mutual information. The work defines mutual information based on the uncertainty between feature groupings. The derivation stems from Shannon's information theory to quantify uncertainty in terms of entropy. The proposed work uses the Mahalanobis distance to convert the statistical problem into a geometric one, and defines uncertainty in terms of this geometric rendition of the statistical space. In both cases, features that overlap too much (either

in terms of uncertainty or geometric overlap) are not selected for training of the neural network.

Weiss and Kapouleas [18] provide an empirical evaluation of a number of different learning techniques. The conclusion indicates that a properly trained neural network can outperform other learning techniques, but takes significantly longer to train. In many cases, simpler techniques can be just as successful, but the neural network approach had more success with features that were especially difficult to optimize for.

1.1.2.4 Bio-Inspired Techniques

Hague et al [21] describe a compressive sensing approach to modeling sonar based on bat sonar. This general approach constitutes the fourth major pillar in the field of sonar target identification. The general approach is to take advantage of the sparse representation of bat biosonar to apply compressive sensing techniques to reconstruct the signal from sub-Nyquist sampling rates. While we do not take a compressive sensing approach directly, the applications for signal acquisition are clear. For the work presented here, the data was gathered by the University of Washington's Applied Physics Laboratory.

1.2 Applicability

This problem has far-reaching consequences. Potential applications include

- Autonomous recovery of UXO
- Navigation in cluttered underwater environs
- Surveying, including for salvage or resources

1.2.1 Autonomous recovery of UXO

A report of the Defense Science Board Task Force examined the issue of UXO clearance, including the technologies and costs involved [25]. It is indicated from this report that land-based UXO recover is still implemented by manually searching with metal detectors and engaging in careful excavation. These techniques have been deployed for underwater environs as well, with limited success. Recovery is typically carried out by divers with hand-held magnetometers, assisted by electromagnetic systems mounted on sleds and side scanning sonar and chemical sensors on AUV.

In a test of autonomous technologies, the Navy invited five firms to test their abilities at Mare Island Naval Shipyard in a test referred to as the Validation of Detection Systems Test Program. Just one company was able to meet the goal detection rate of 85% [25]. They managed a detection rate of 99%, but also had a false alarm rate of 18%. Classification proved more difficult, with a peak classification rate of 68%. These approaches were not autonomous, and involved substantial human intervention. The Navy issued a Statement of Need following these tests indicating interest in expansions that address

- access limitations of the sensing equipment
- improved signal processing to aid detection and discrimination
- characterization of UXO migration and burial
- removal and disposal techniques

Our proposed work is applicable primarily to the first two objectives. By enabling AUVs to make more accurate predictions and classifications, discrimination will be enabled in more complex environs.

These applications will increase in importance as research in autonomous underwater vehicles accelerates. AUVs have already been used in the search for downed planes [26], environmental research in harsh/unfriendly environs [27, 28], and augmentation of surveying capabilities [29].

1.3 Relevant Technologies and Inquiries

The research presented here focused on the intersection between signal processing and pattern recognition. To that end, some of the key technologies under investigation were:

- KNN clustering
- support vector machines
- neural networks
- Gabor filters

A variety of techniques were employed as part of research in a supporting role, including:

- L1L2 optimization
- Periodogram analysis

Discussion of these techniques will follow in subsequent chapters.

1.3.1 Contributions

The proposed techniques laid out in both chapter 4 and chapter 5 adapt some common pattern recognition techniques to the underwater object recognition problem. The ellipsoidal manifolds proposed in chapter 4 are similar conceptually to the bag of words technique common in image processing [30]. The author contributed the visualization of the clusters as geometric objects for simplification of conversion from feature to keypoint in the bag of words technique. This technique is used extensively in image processing [31] as well as for the general problem of recognition and classification [32].

Chapter 5 adapts the Gabor wavelet to the characterization of underwater objects. Gabor wavelets are used with success for the characterization of images for analysis and compression [33]. Between the biologically inspired arguments made in [33] and the similarity to the models proposed in [34] for characterization of active sonar acoustic color features, the author pursued the Gabor wavelet as a feature filter.

The proposed techniques stem from an interdisciplinary approach to problem solving, and add to the body of knowledge explored in the field of underwater object recognition using active sonar.

Chapter 2

Background Information and Theory

2.1 K-Nearest Neighbor Clustering

KNN is a relatively simple machine learning algorithm that provides tools for both recognition and regression [35]. At its heart, it is essentially a voting algorithm whereby a given sample is given a label based on the labels of the k nearest samples. Nearest samples are typically weighted by a distance metric to prevent domination by the classification with the plurality in the data set [36]. Common distance metrics include the Euclidean and Mahalanobis distance. A common unsupervised approach to data segmentation involves picking seed centroids and performing the KNN labeling, then iterating to reduce in-class variance [37].

In our investigations, we will take advantage of the KNN algorithm cluster extrema, and use the Mahalanobis distance to characterize the result. We believe this is an appropriate approach because the KNN clustering algorithm uses similar metrics internally to form the clusters in the first place.

2.1.1 The Mahalanobis Distance

We will be referring to the Mahalanobis distance in several places, so a brief aside for definition and discussion is warranted. The Mahalanobis distance is defined as:

$$D_M = \sqrt{(\vec{x} - \vec{\mu})^T \vec{S}^{-1} (\vec{x} - \vec{\mu})} \tag{2.1}$$

where $\vec{x}$ is a point, $\vec{\mu}$ is the mean of the distribution to compute the distance to, and $\vec{S}$ is the covariance matrix of the distribution. In the case of unit variance, the Mahalanobis distance collapses to the Euclidean distance [38].

2.2 Support Vector Machines

SVMs are powerful tools for classification and discrimination. Simply put, they find separating hyperplanes for a given feature set [37]. The hyperplanes are defined to satisfy the equation [39]:

$$\vec{w} \cdot \vec{x} - b = 0 \tag{2.2}$$

This approach, which generated linear separating planes, was expanded by adding kernels to enable nonlinear discrimination [39, 40]. This approach transforms the data using nonlinear kernels rather than the simple dot product originally proposed. This has the effect of transforming the data to a higher dimensional space before performing the optimization. When transformed back into the original space, the separating hyperplane is no longer linear, but maintains the positive properties of the hyperplanes optimized with the original method.

2.3 Neural Networks

ANNs are somewhat controversial tools for classification and discrimination. The algorithm acts as a learning algorithm that can be tailored to a variety of tasks. The networks typically consist of a set of inputs (called input neurons or the input layer), one or more intermediary layers (called hidden layers), and a set of outputs (the output layer). The defining characteristics of an ANN are typically [41, 42]:

- The nodes have a tunable gain, which is trained by a learning algorithm
- The sum of the node operations can approximate complex (even nonlinear) functions

The nodes can be connected in a variety of ways, but a common method (and the one we employ) is complete connectivity. Under this schema, all nodes in any layer are connected to all nodes in any adjacent layer. The connections are only "severed" by dropping the gain on the corresponding input in the next layer to zero. Each node contains an activation function that transforms the input to an output. A common choice is the hyperbolic tangent, which provides a transition region over which the behavior can be considered roughly linear and approximately steady state regions over which the behavior is approximately binary [37].

The network only describes the relations between the nodes (i.e. the mapping from input to output). The training paradigm for the node weights (and even the decision functions) can vary wildly. A common method is the back propagation method [37, 42], which is a deterministic method that maps output error backwards through the system to determine which direction the nodes should be adjusted to reduce the error. Stochastic

techniques can also be employed, which do not necessarily suffer from the trap of local minima, but the convergence time is poor and optimality is still not guaranteed [43]

2.3.1 Limitations and Potential Improvements

ANNs are frequently criticized for some serious shortcomings. Arguably the most important limitation of an ANN is the tendency towards overtraining (i.e. excellent performance on the training set, with poor performance on real world data due to overfitting) [44]. This limitation can be alleviated by increasing the amount of training data (and guaranteeing that the training set is representative of real world data) and by guaranteeing the training and verification sets are disjoint (to prevent reporting unrealistic accuracies) [44]. We employ both methods in our work.

2.3.2 Our Implementation

The ANN used in this work is composed of three layers, meaning one hidden layer along with the input layer and output layer. The network is initialized to be fully connected by initializing the arrays to unity value. At each layer, the activation function (tanh) is applied to the value at each node after multiplying the input by the weight.

If we ignored the activation function, we could model a pxq ANN with matrix operations quite easily: $input * L_{input} * L_{hidden} * L_{output} = outputlabels$, where the input is 1xp, L_{input} is pxp, L_{hidden} is pxp, and L_{output} is pxq. Adding in the activation function modifies this equation to: $tanh(tanh(tanh(input * L_{input}) * L_{hidden}) * L_{output}) = outputlabels$. This notation is a bit unwieldy, which is why the network notation is typically used to describe the ANN.

The values at each output node are a sort of "confidence" in that label being correct. Training is conducted via backpropagation of errors. The error at the output is propagated backwards. Gradient descent is used to improve the modeling of the target function. The gradient is taken of the tanh function which leads to a step of $\alpha * (1 - tanh^2(error))$. This process is repeated back to the input layer.

2.4 Gabor Filters

The Gabor wavelet is a useful function for modeling a variety of physical processes for several key reasons [45]. Among other things, the wavelet provides high time/frequency localization, which allows for high quality feature compression. This localization means the Gabor wavelet (when acting as a basis function) doesn't fall victim to the issues of the impulse or cosine functions which have difficulty representing periodic and aperiodic signals (respectively) using a finite number of coefficients. Gabor filters are commonly used in image processing for edge finding [46]. In this application, the 2-D version of the Kernel is used, which is defined as:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}) \exp(i * 2\pi \frac{x'}{\lambda} + \psi) \tag{2.3}$$

where:

$$x' = x \cos\theta + y \sin\theta; y' = -x \sin\theta + y \cos\theta \tag{2.4}$$

It is clear from the above that the 2-D Gabor kernel is simply the combination of a 2-D Gaussian and a plane wave. The Gabor transform was originally considered a sort of windowed Fourier transform, which makes sense intuitively given the definition of the wavelet.

Gabor wavelets have been adapted for use in sonar target response characterization [47–49]. The authors therein claim the feature space is desirable due to the potential for dimensionality reduction and quality of features for object representation.

2.5 Acoustic Physics

The sonar target response of an object is largely due to two major components: backscatter from the surface and elastic waves within the object [50]. The elastic waves can be modeled as a combination of Rayleigh and Whispering Gallery waves [23]. Rayleigh waves are surface distortions in the solid, while whispering gallery waves are due to internal resonance where the concave surface acts as a wave guide.

Chapter 3

Synopsis of Technical Approach and Data-Driven Investigations

3.1 Characterization of UXO Using Active Sonar

The research was focused primarily on the problem of characterization and classification of the active sonar signature of several underwater objects, including steel and aluminum UXO. The data was generously provided by the Applied Physics Laboratory at the University of Washington, and is in the public domain for the validation of techniques such as the ones we've developed.

We investigated several approaches to characterization and classification of active sonar targets in cluttered environments. These approaches are potentially complementary, but were evaluated separately. The first approach used KNN clustering on frequency extrema, then characterized the resulting clusters using the Mahalanobis distance. The result was classified using an ANN due to the flexibility of the network in terms of inputs and tuning. The second approach used Gabor filters for feature extraction and used SVMs for classification of the resulting features.

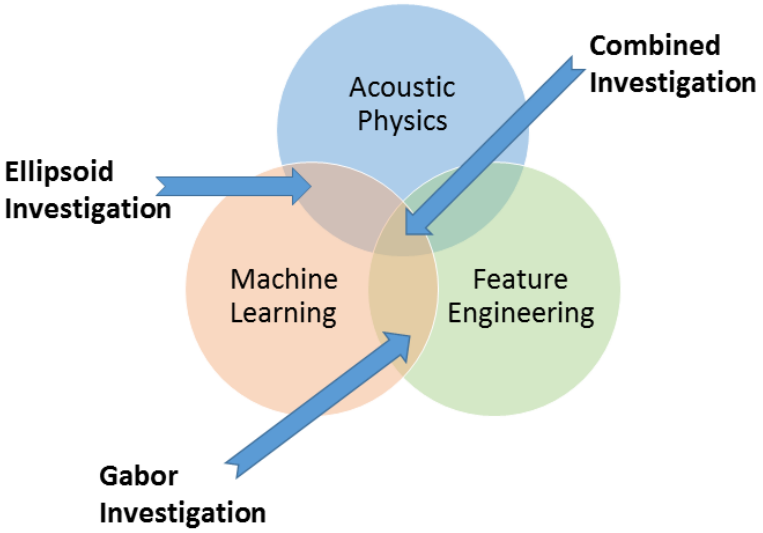


FIGURE 3.1: Summary of region of interest and location of investigations at the intersections of acoustic physics, machine learning, and feature engineering.

3.2 Using Ellipsoids as Sonar Feature Manifolds

As explained earlier, unsupervised methods of KNN clustering with distance metrics tend to minimize the variance of the distribution, subject to normalization using the Euclidean distance. As the Mahalanobis distance is effectively a statistical extension of the Euclidean distance, it was chosen as a natural metric for the formation of geometric feature manifolds. These feature manifolds can be described in terms of overlap and separation relatively easily. This allowed us to build feature models by constructing lists of feature manifolds from the clustered result of the frequency extrema. The manifold lists were then culled by searching for significant inter-class overlap. This overlap would indicate poor separability for frequency features in these regions.

The remaining "high quality" feature manifolds could be fed into a neural network. For our purposes, we treated the manifolds as input nodes to our network. If an input had extrema matching a given manifold, that input node activated. The ANN was trained with cross validation, and it was found that the resulting classifiers had superior accuracy on the testing sets, especially in terms of rejection of superfluous features that would cause a false positive. This approach is discussed in more detail in Chapter 4.

3.3 Using Gabor Wavelets to Derive Sonar Features

We opted to take advantage of the nature of Gabor Wavelets, in part thanks to a collaboration with NUWC (the National Underwater Warfare Center). We investigated a broad approach in which a variety of kernels (varying the wavelet parameters) were generated and used to transform the data. Each transform was evaluated by training an SVM for classification, and the kernels that yielded the highest accuracy in the classification phase were selected as the most promising candidates. The investigation yielded the tools and a framework for evaluating this approach that will work well for future collaborations with physicists, as they will enable swift iteration and testing for the kernel parameters suggested by physical models.

The approach (detailed further in Chapter 5) was evaluated using (among other things) a set of confusion matrices. These matrices allow quick visualization of a multi-label system's accuracy in and between each class. This enabled us to give a graphical indication of quality that makes it easy to see that Rocks are the hardest to accurately distinguish (which is unsurprising due to their non-standard shapes and compositions), and aluminum cylinders are difficult to distinguish from aluminum UXO. In each of the examples presented, the true class had the plurality of predictions.

3.4 Distinguishing Characteristics

The application of neural networks and wavelets to the underwater target classification is not new [47–49]. The works presented here bear some resemblance to the approaches attempted in the past but differ in several key areas. As the investigations presented here have been split, with separate investigations for the role of neural networks and the role of Gabor wavelets, the distinctions will also be separated thusly.

Our investigation involving neural networks differs from existing approaches such as [47] primarily in the choice of inputs. In Chapter 4, we discuss the use of ellipsoidal geometric manifolds for feature characterization. These manifolds are conducive to feature culling, and the inputs to the neural network are selected from the generated manifolds. While existing approaches make use of the neural network by feeding the entirety of the feature vector (in [47] this is the result of Gabor transformation cross correlated with the Gabor transform of the input signal to the system), our investigation predicts high-quality features and allows for variance based on the feature clusters across all sensor orientation angles. The crux of the difference here is that the use of geometric feature manifolds in our paper improve the range of sensor angles over which a classification can be achieved that is well defined and confident.

Our investigation involving Gabor features also differs from existing approaches such as [47, 48] in terms of the choice of input. In Chapter 5, we discuss the application of the 2D Gabor kernel to the periodogram of the received signal across all sensor orientation angles. The existing methods use a 1D Gabor transform on the received signal for the objects. The method presented in Chapter 5 uses multiple sensor orientations in the evaluation of each set of Gabor parameters. The accuracies are quoted for the entire

dataset, including all sensor orientation angles. The verification step uses the data one-dimensionally (i.e. treating each sensor position as a separate received signal, using the same Gabor parameters to characterize).

While real-world systems can incorporate other techniques for localization in conjunction with identification, it is also valuable to have identification techniques that are built with multiple sensor orientations in mind, as glancing results will constitute a significant portion of received signals.

3.5 Scope of Proposed Work

The investigations to follow are focused on the problem of target recognition from a machine learning perspective, informed by the physics of the system. The data collection mechanism is not investigated, as the data set has been inherited from the Applied Physics Laboratory at the University of Washington. The investigations will include discussions of the theory and application of several machine learning techniques, along with some novel approaches to cluster identification and analysis. The investigations will also make use of standard k-fold validation techniques.

Chapter 4

Ellipsoid Features

4.1 Introduction

Detection and classification of buried and partially buried targets in active sonar against background interference from the environment have been a long-standing challenge for decades. Current state-of-the-art in buried target detection (using active or passive sonar), signal recovery and source classification techniques combine rich advances in four complementary directions: (i) Underwater acoustic propagation studies [8–10], including noise models [11] and acoustic backscatter studies that may be harnessed to achieve buried source localization and signal recovery on physics-based principles; (ii) Information bounds and trade-offs [3, 12–15] that provide mathematical limits on when a buried signal may be recoverable against background noise and clutter; (iii) Adaptive signal processing [1–5, 16], learning algorithms [17–20] and object classification techniques (ref. e.g. [4–7] and references therein) that perform data-driven signal recovery. This includes array-based methods) and other techniques that acquire side information [4, 5, 7] to recover buried signals. (iv) Bio-inspired sonar signal processing techniques (e.g. [21] among others).

Our contribution in this work falls in the third category above, with strong potential of combining our geometry driven strategy with physics-inspired approaches in (i). Our

methods provide for improved *object recognition*, and have a high potential for applicability for *image processing* applications. Specifically, our intellectual contributions may be summarized as:

- Formulate the sonar target response as a feature network of overlapping ellipsoids in a geometric space defined by the average intensity, frequency, and time delay in the frequency space representation
- Cluster and classify features based on intra-class overlap (measured using Mahalanobis distance) between the featured ellipsoids
- Identify methods to promote features that stay relatively invariant over orientation angle for the same target while culling features that are heavily shared with other objects, or otherwise fail to enable inter-class discrimination

In essence, we augment two-dimensional (time delay/frequency) into a third dimension with sensor position to augment, then use those features to characterize future samples that consist only of the time delay/frequency pair. We validate our method over experimental field data ranging between different orientation angles for different man-made targets and naturally occurring non-targets (e.g. rock).

4.2 Technical Approach

This approach effectively models the sonar channel as a series of frequency groupings in space, along a cross-range scan. We adopt a geometric interpretation of intensity extrema clusters by adapting the Mahalanobis distance as a statistical metric for achieving this grouping. We consider this geometric manifold technique useful because it allows for

the augmentation of frequency extrema using variable sensor position, while preserving computational simplicity. We now provide details below.

We first determine the local maxima of the sonar response in the frequency domain by employing a standard nearest neighbor comparison. Thresholding is applied to avoid low-intensity variations (ref Fig. 4.1). We represent the maxima as a list of tuples representing the (i) rail position, (ii) frequency, and (iii) amplitude at each maximum. We now cluster the maxima via k-means [51], and convert the resulting clusters into ellipsoidal features(ref Fig. 4.2 and 4.3).

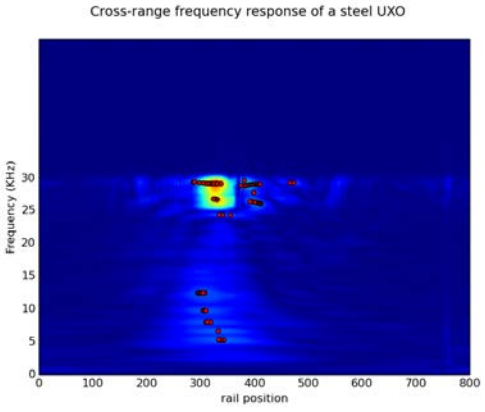


FIGURE 4.1: *Peak Finding:* Extrema are overlaid (in red) on the cross-range frequency response.

The three-dimensional ellipsoidal features are composed of: (i) a center, with three sub-features (average cross range sensor location, average intensity of the group, and average frequency of the group); and (ii) three radii (range of sensor positions, intensity variation of the group, and frequency range of the group). These features characterize the acoustic impulses received in terms of frequency and sensor position, and thus carry the sonar signature of the target. Our preliminary results indicate (ref Fig. 4.3) key ellipsoidal features remain invariant through a variety of orientation angles and thus, allow distinguishing a target (e.g. unexploded ordinance) against non-targets (e.g. a

rock, ref Fig. 4.5). We generate as many separable features as possible, and cull the result in the feature separation step.

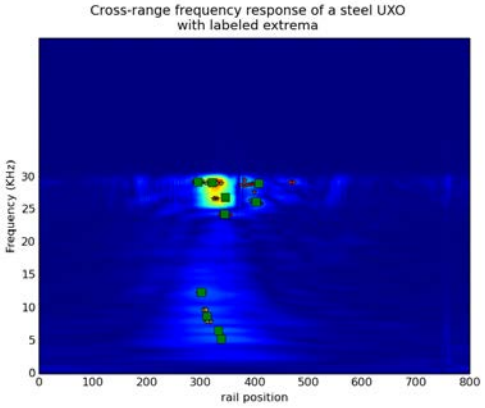


FIGURE 4.2: *Clustering with Overlay:* Clusters overlaid back on the cross-range frequency response and shown from top-view. The centroids of each cluster are marked by green squares.

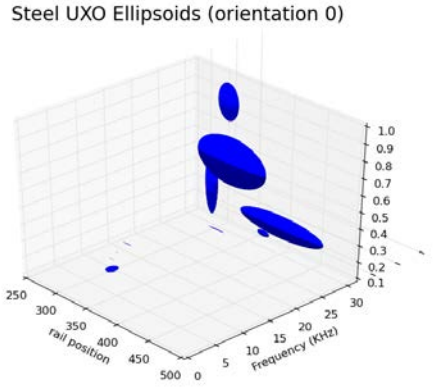


FIGURE 4.3: *Ellipsoids:* An example of what the ellipsoidal features look like in the 3D sensor position/frequency/intensity space. The orientation noted in the title is of the target, not the sensor. Note that these ellipsoids are well localized in frequency, and so appear large in this image. Compare to ref Fig. 4.6 to see how the ellipsoids project on a single sensor orientation.

4.2.1 Feature Separation and Classification

Resulting ellipsoidal features are compared between potential targets using geometric means. The centroid-to-centroid distance (Equation 4.2), normalized by the radii, is analogous to the Mahalanobis distance. We cull features that have a normalized distance to another feature of less than unity. This isolates feature subsets where centroids of

Algorithm 1 Feature Generation

```
0: procedure ExtremaSearch(cross range response)
0:   for railposition/frequency voxel in periodogram do
0:     if I(voxel) > I(neighbors) & threshold then
0:       append voxel sensor pos/frequency/intensity
   return extrema list
0: procedure KMeans(extrema)
0: procedure EllipsoidGeneration(clusters)
0:   for each cluster c do
0:     E ← new ellipsoid container
0:     E.cx ← average rail position of c
0:     E.cy ← frequency centroid of c
0:     E.cz ← intensity centroid of c
0:     E.rx ← variance of rail positions for c
0:     E.ry ← frequency variance of c
0:     E.rz ← intensity variance of c
  =0
```

at least one of the features is in the variation envelope of the another. This ensures overlapping features are classified into subsets that are most likely separable as a group within the target's acoustic signature. Mathematically, this may be stated as

$$D_M = \sqrt{(\vec{x} - \vec{\mu})^T \vec{S}^{-1} (\vec{x} - \vec{\mu})} \tag{4.1}$$

where $\vec{x}$ is a point, $\vec{\mu}$ is the mean of the distribution to compute the distance to, and $\vec{S}$ is the covariance matrix of the distribution. In the case of unit variance, the Mahalanobis distance collapses to the Euclidean distance [38]. We adopt the l_1 norm for computational efficiency. The inter-feature distance from any given ellipsoidal feature to any other feature is given by:

$$D_E = (\vec{\mu_1} - \vec{\mu_2}) \vec{R_2}^T \tag{4.2}$$

where $\vec{R}$ is the array of radii of the ellipsoid. Due to the nature of ellipsoid generation, this is analogous to the covariance matrix.

$$\vec{R} = \begin{pmatrix} E.rx^{-1} & E.ry^{-1} & E.rz^{-1} \end{pmatrix}$$

4.2.2 Feature Selection from Overlapping Features

We adopt a subtractive approach to feature selection. If a feature has $|D_E| < 1$ to any other feature, that feature is considered overlapping. We consider these overlapping, as the distance is normalized by ellipsoid radius, thus $|D_E| < 1$ indicates ellipsoids whose centroids fall within the bounds of each other. After generating a large number of ellipsoidal features, the list is searched for inter-class overlap. Those features are removed from the list of distinguishing features. The remaining list consists only of features that do not significantly overlap with known features in other classes. *Intra-class overlap indicates features that remain constant for different target orientations.*

4.3 Results

We provide representative results from our investigations which indicate promising feature separation, despite obfuscation due to sensor orientation. For the purpose of visualization, all feature centroids are plotted as blue points, and the regions of separability are plotted as red ellipsoids. Features that fall in these regions distinguish between the given classes. We have provided results that offer separability between two distinct targets, both unexploded ordinances (UXOs), as well as between a target and a non-target (rock).

Even for similarly shaped objects such as a steel and aluminum UXO, there is distinguishability as is evident in Fig. 4.4. The steel UXO exhibits high frequency content,

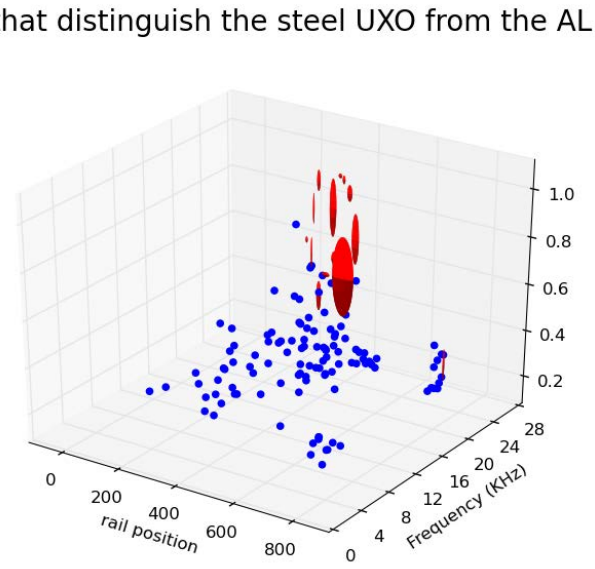


FIGURE 4.4: *Steel UXO (Target A) vs AL UXO (Target B):* Feature regions in red separate Steel UXO from Aluminum UXO.

especially on the direct path arrival (16% of generated features exhibit no overlap). Red regions from ref Fig. 4.4 indicate distinguishability and separability (10% of features can be found on the direct path that are unique to the steel UXO) between the two distinct targets.

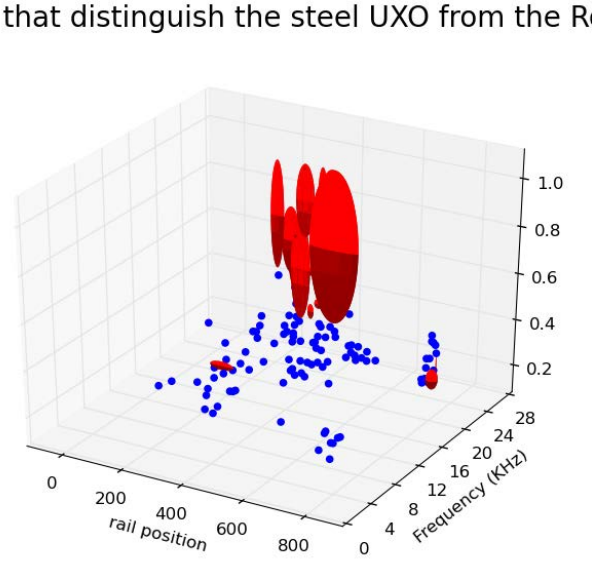


FIGURE 4.5: *Steel UXO (Target) vs Rock (Non-target):* Feature regions (in red) separate Steel UXO from a rock.

Separability between steel UXO and the rock is significantly better, as evidenced by the increased localization of features in the direct arrival (approx. rail position 300). The rock does not exhibit any significant high intensity feature localization, while the UXO shows significantly more intense and tightly grouped 30KHz components over approximately 15% of rail positions. This is further evidenced by Figs. 4.6 and 4.7 in the sequel.

Projecting our in-class ellipsoidal features onto the intensity/frequency plane, we observe visual peak overlap in the one-dimensional data with the generated features (ref Fig. 4.6). We further observe that the features capture the target's response as it changes with orientation angles, e.g. compare between ref Figs. 4.6 and 4.7.

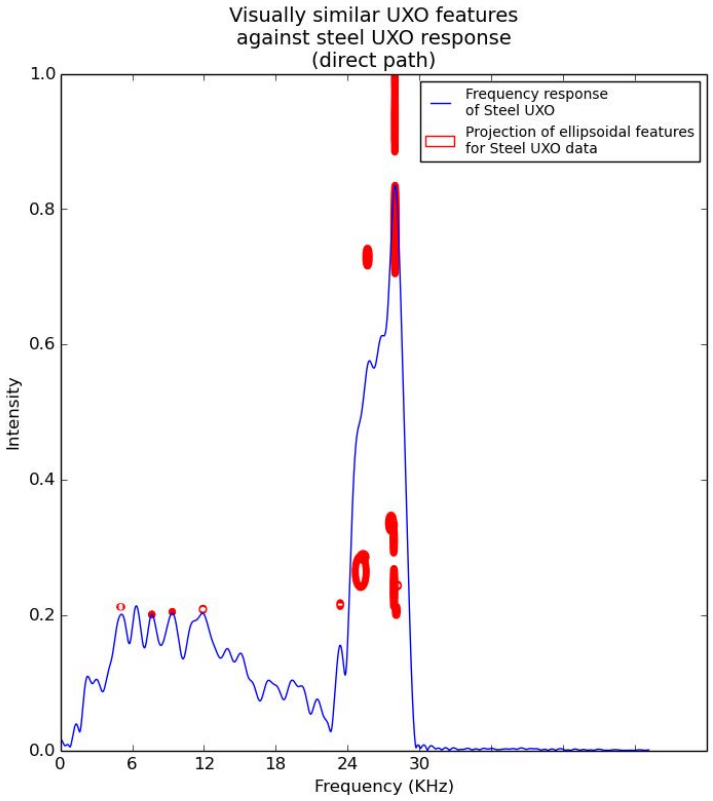


FIGURE 4.6: *Frequency response of a* direct *hit with Steel UXO ellipsoids overlaid.*

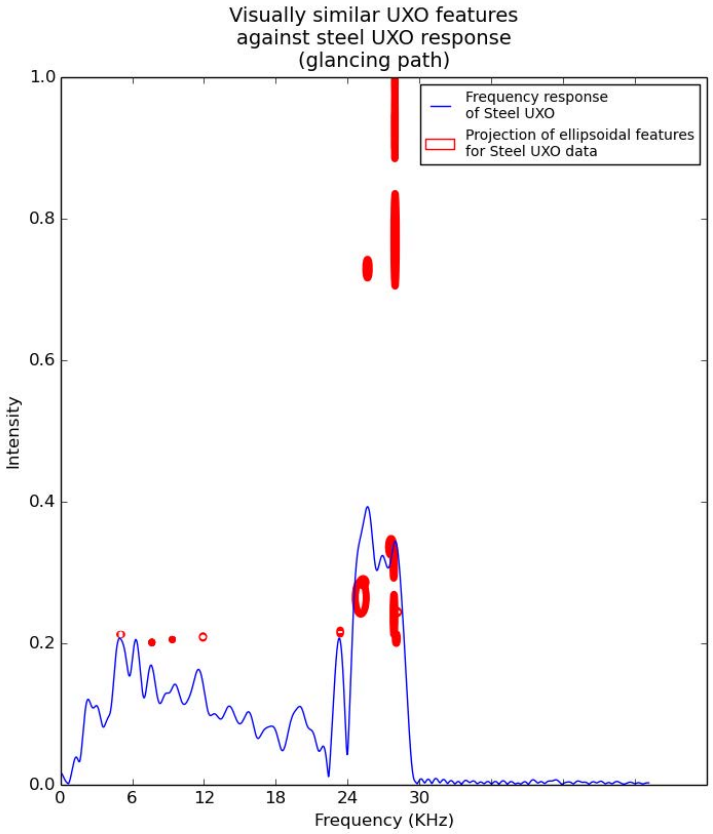


FIGURE 4.7: *Frequency response of a* glancing *hit with Steel UXO ellipsoids overlaid.*

4.3.1 Evaluation against Traditional Correlation-Based Matching

We compared the technique against normalized cross correlation, the basis of matched field processing [1], with favorable results. The DC component of the cross correlation of the received signal of the steel UXO with the rock data shows poor visual separability (see ref Fig. 4.9) across all sensor angles. *Quantitatively we observed that with the steel UXO as reference, matched processing using direct-hit response attains* 90.7% *and* 94.7% *match respectively with Aluminium UXO and the rock. ref Fig. 4.10 shows zero percentage overlap between ellipsoidal features from the steel UXO or the rock that match the extrema received steel UXO data across all sensor angles.* Only the features that have one (and only one) match are counted as matching distinguishing features. The percent of matching ellipsoids is much lower, but the separation is visually distinguishable,

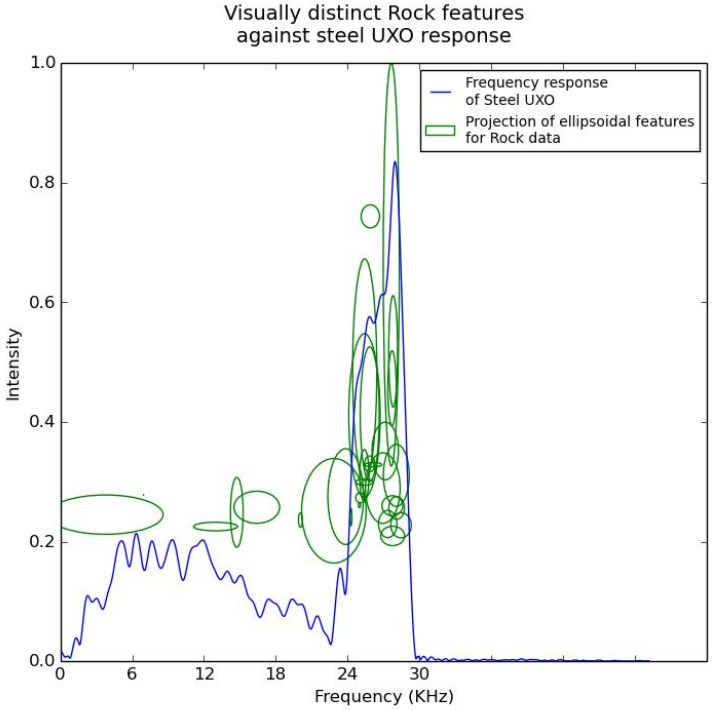


FIGURE 4.8: *Frequency response of a glancing hit on a rock with Steel UXO ellipsoids overlaid.*

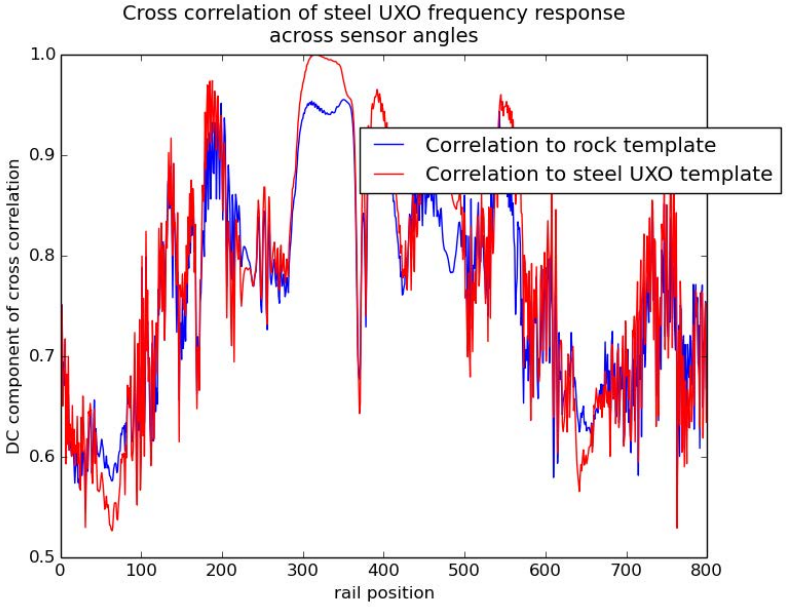


FIGURE 4.9: *Normalized cross correlation of direct-path steel UXO frequency response with rock response.*

especially around the direct-path sensor orientation.

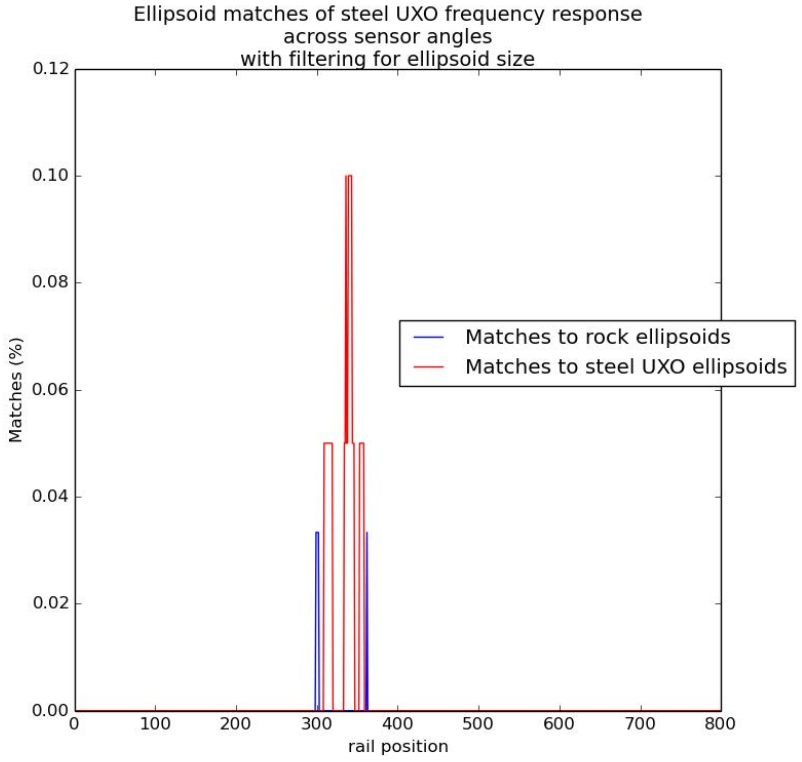


FIGURE 4.10: *Comparison of template matching using ellipsoidal features and normalized cross correlation.*

4.4 Concluding Remarks

We have presented a novel visualization of target response to active sonar that exploits intra-class feature geometry to achieve enhanced inter-class feature distinction. Specifically, we have modeled key features as ellipsoids in the time-frequency domain and derived classification techniques that utilize their overlap or lack thereof, to cluster the response from an unknown target into known target profiles. We have provided data-driven validation of our techniques over two well-known targets and one non-target using public domain sonar data collected by the Applied Physics Laboratory, University of Washington. Our methods indicate high feature separation between a given target and other targets as well as non-targets over established state-of-the-art (matched field processing [2, 16]). Continuing and future directions include validating this promising

approach over a broader set of field data including more targets and non-targets, as well as connecting the ellipsoidal features of a given target to its inherent scattering physics.

Chapter 5

Gabor Feature Investigation

5.1 Introduction

The problem of target identification using active sonar continues to be a challenge. This problem has far-reaching implications including military, environmental, economic, and emergency response. Characterization of targets is confounded by several factors:

- Weak ground truths for targets and nonexistent ground truths for unknown targets makes conventional learning techniques untenable [1–5]
- Intra-target interference (in the case of multi-target scenarios) can cause overlapping resonance components that challenges conventional learning techniques [4–7]
- Artifacts due to boundary and channel effects increase target response complexity, making the problem more difficult in general

To meet these challenges, we adopt a data-transformative approach using Gabor wavelets for feature generation. These wavelets are ideal for modeling the system, and the timefrequency localization characteristic enables high data compression (well-defined features). Our intellectual contributions are summarized as:

- Generate a series of features that
  - Take advantage of the physical properties of the system

  - Exhibit strong data compression via time-frequency locality, enabling well-defined, distinguishable features

- Make use of a cascaded SVM classifier using our features that yields high accuracy predictions for target identification [17–19]

The characterization of sonar target response using Gabor features enables the leveraging of the very convenient characteristics of the Gabor wavelet which aren't just limited to timefrequency localization. The Gabor wavelet is also particularly useful because the tunable parameters enable the wavelet to emulate many meaningful physical characteristics of the system. These include

- Resonance due to Rayleigh waves and whispering gallery waves [52]
- Boundary effects due to object size/shape

Two-dimensional Gabor wavelets are suitable for modeling the responses of acoustic waves due to their flexibility in representing sharp features (major resonance peaks) as well as smoother acoustic topography in the time-frequency domain.

5.2 Technical Approach

5.2.1 Gabor Features

The active sonar response of a variety of targets is transformed into the frequency space. The resulting signal is convolved with a Gabor wavelet, which is defined as

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}) \exp(i * 2\pi \frac{x'}{\lambda} + \psi) \tag{5.1}$$

where:

$$x' = x\cos\theta + y\sin\theta; y' = -x\sin\theta + y\cos\theta \tag{5.2}$$

It is clear from Eq. 5.1 that the wavelet is a combination of a Gaussian and a plane wave. By tuning λ, we can alter the captured frequencies, and by tuning σ, we can alter the captures scales (the other coefficients play their own role in feature capture). As a proof of concept, a variety of Gabor kernels were generated, and the data set was transformed by each. The transformation was performed as a 2-D kernel convolution across the presented data.

The resulting vectors were used to train SVMs to test discriminating ability. For the purpose of this kernel selection stage, a linear C-SVC was adopted. The training and tests sets were generated via random selection and guaranteed to be disjoint. The resulting SVMs were evaluated via a 5-fold cross validation for robustness. The accuracies were between 50% and 65%. Preliminary attempts at boosting the results together indicated that the classifiers were suitable for combination for higher accuracy.

5.3 Results

We tested our method over the experimental field data collected by the Applied Physics Laboratory, at the University of Washington.

5.3.1 Transformation Results and Examples

A variety of kernels were generated, and their performance was evaluated by attempting to train and test a simple linear SVM. Fig. 5.1 is an example of a generated kernel with $\lambda = .1$, $\theta = 90°$, and $\sigma = 1/6$.

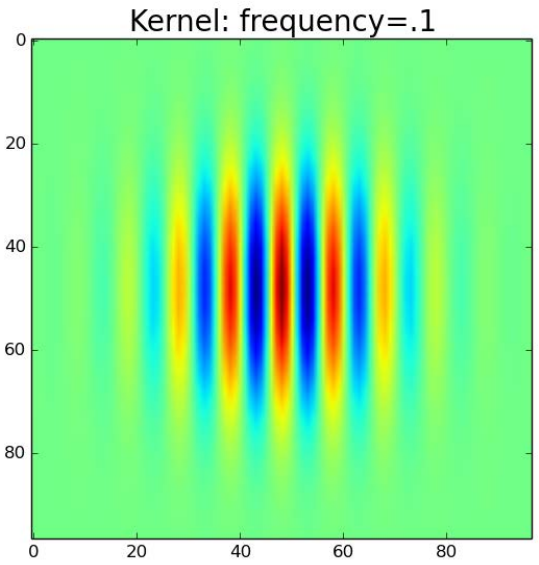


FIGURE 5.1: *Example Gabor kernel.*

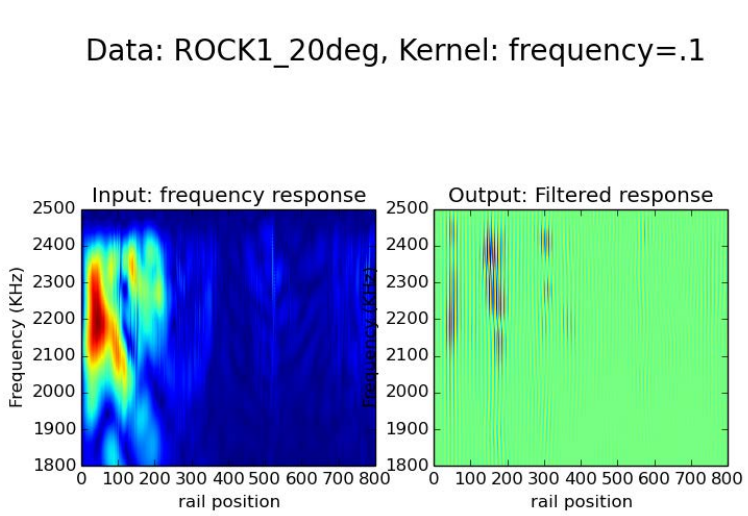


FIGURE 5.2: *First rock example.*

It is clear from Figs. 5.2-5.5 that the transform makes the features more visually distinct at least. Figs. 5.2 and 5.3 are both examples of a transformed rock, while Figs. 5.4 and 5.5 are both examples of a steel UXO (Unexploded Ordinance). These are examples taken from different orientations of the target object, yet the transformed features (from the convolution) bear striking visual similarity. Despite significant visual

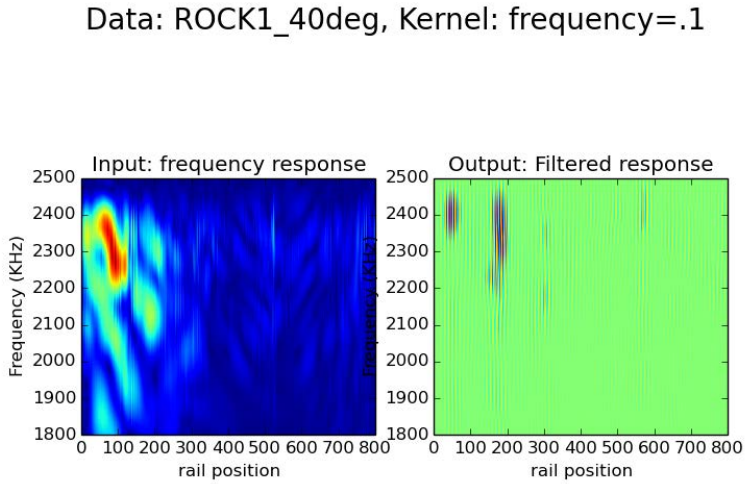


FIGURE 5.3: *Second rock example.*

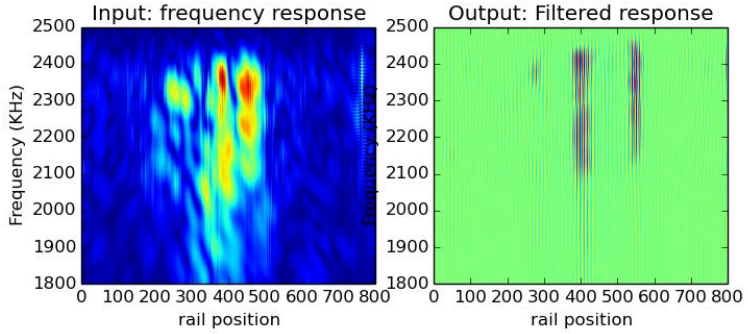


FIGURE 5.4: *First steel UXO example.*

differences in the frequency response of the steel UXO at the different orientations, the transformed result (in Figs. 5.4 and 5.5) remains strikingly similar. To quantify our results, we used a simple SVM with a linear kernel to attempt discrimination.

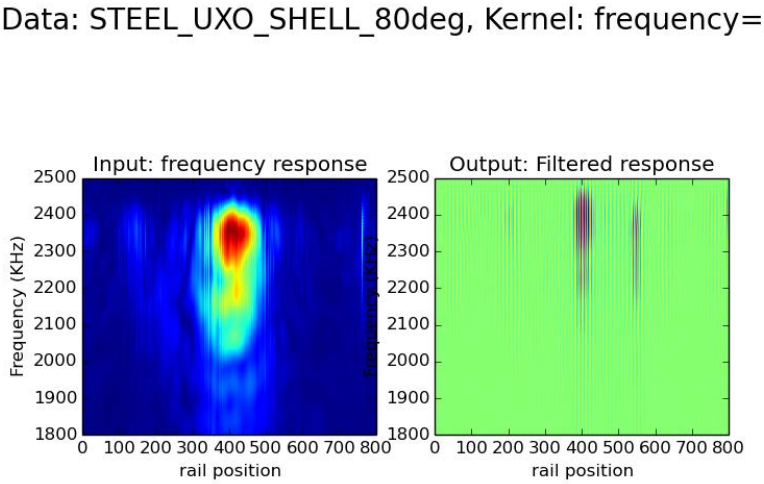


FIGURE 5.5: *Second steel UXO example.*

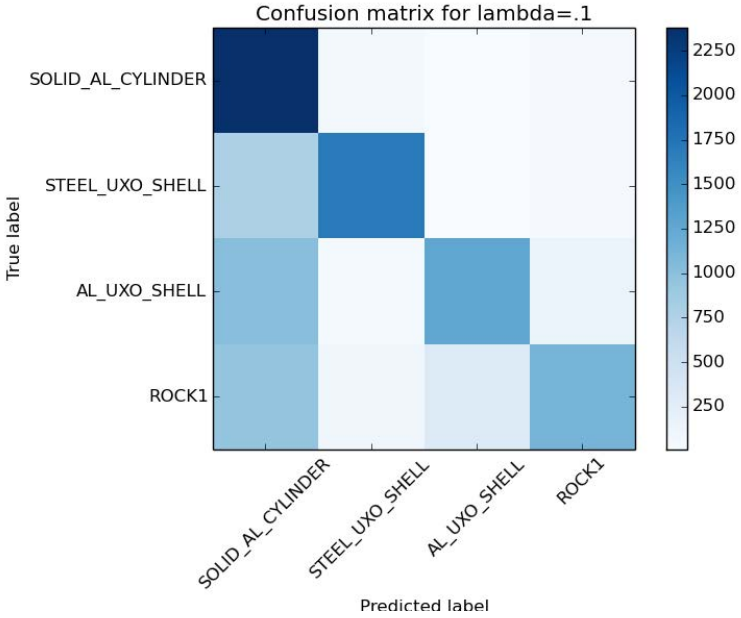


FIGURE 5.6: *Confusion matrix of $\lambda = .1$, $\theta = 90°$, and $\sigma = 1/6$, with 5-fold cross validation accuracy of 64.61%.*

5.3.2 SVM Results and Examples

The confusion matrices (which measure predicted label vs true label) indicate clear preference towards in-class identification with several artifacts, owing to the large number of harmonics exhibited by the aluminum cylinder. Our 5-fold cross validation provides

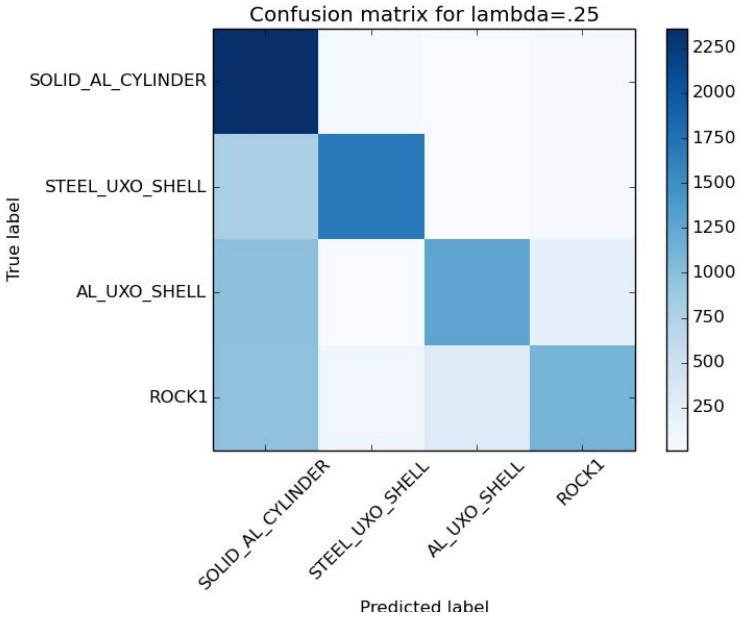


FIGURE 5.7: *Confusion matrix of* $\lambda = .25$, $\theta = 90°$, *and* $\sigma = 1/6$, *with 5-fold cross validation accuracy of 64.44%.*

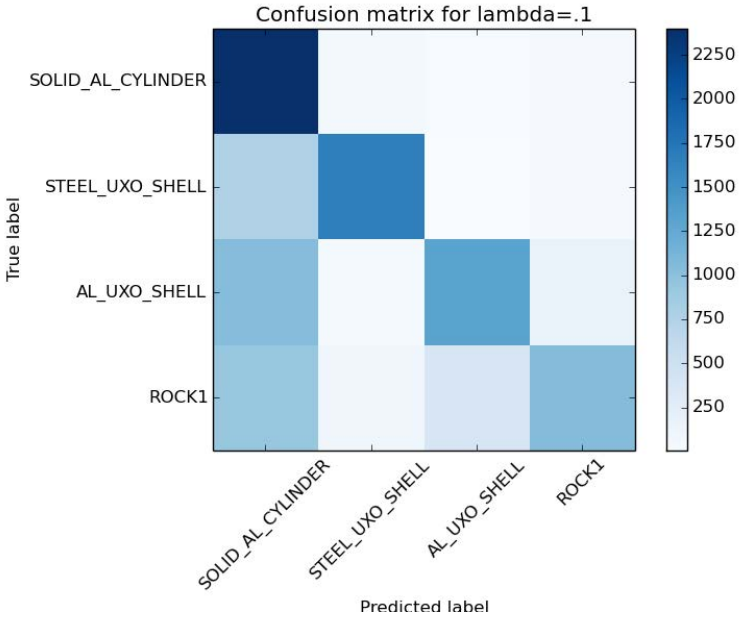


FIGURE 5.8: *Confusion matrix of* $\lambda = .1$, $\theta = 45°$, *and* $\sigma = 1/10$, *with 5-fold cross validation accuracy of 64.32%.*

high confidence for the reported accuracies of each classifier. The accuracies ranged from 50% to 65%, but we have many classifiers at the high end of the range to choose from.

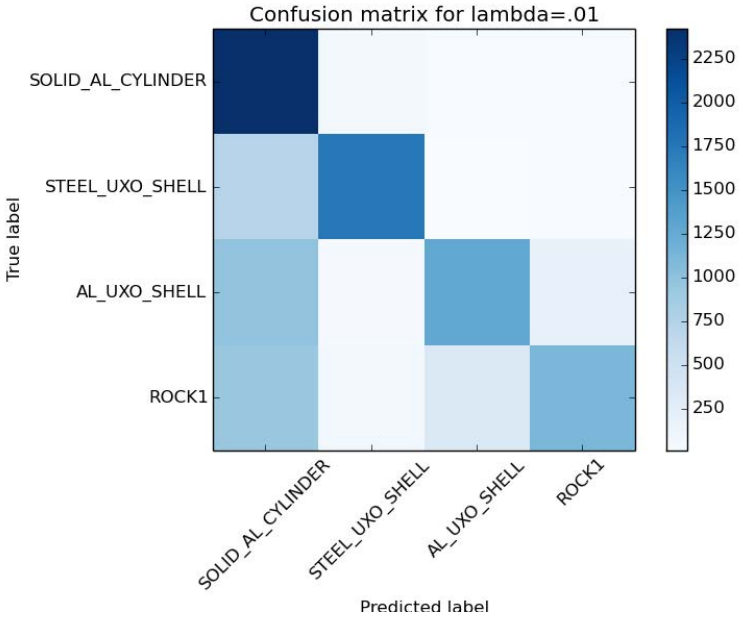


FIGURE 5.9: *Confusion matrix of $\lambda = .01$, $\theta = 45°$, and $\sigma = 1$, with 5-fold cross validation accuracy of 64.53%.*

5.4 Conclusions

The Gabor features exhibit strong localization and data compression, presenting exciting preliminary results for characterization, discrimination, and classification. It is clear from the results of the training stage that there is potential for high-quality discrimination that can be built upon with future work.

These features address the challenges presented earlier by

- Compressing complex responses into well defined, distinct features that enable high quality discrimination with simple learning techniques
- Enabling characterization using a superposition of these features, mitigating intra-target interference
- Increasing classifier robustness, mitigating the effects of noise by capturing harmonic and resonance effects

5.4.1 Physics-Based Dictionaries

By combining this approach with a characterization of the physical processes in the system, we can make more intelligent predictions as a starting point for our kernel selection. The wavelet can emulate many relevant features including frequency harmonics and direct-path arrival impulses. We will evaluate several physical models for accuracy using our transformation and testing framework.

5.4.2 Boosted Classifiers

Preliminary work indicated that a combination of the classifiers improved discrimination performance. By building distributions based on distance to the separating plane, it would be possible to convert that distance into a confidence (instead of a binary decision), which would allow for the use of a Bayesian cascade for classification refinement. By combining physically predicted models and empirical results, we will present a comprehensive and robust characterization scheme for active sonar targets.

Chapter 6

Combined Investigation

6.1 Introduction

Target identification using active sonar is a compound problem. The problem is difficult to solve with physics-agnostic machine learning approaches due to the dearth of data. Combined with the complex and cluttered nature of the available samples, traditional machine learning approaches can suffer from a variety of issues including overtraining. Physics cognizant matched filter approaches have other issues. The non-stationarity of the environment as well as environmental clutter are difficult to model. The target response will change drastically with orientation angle, making reliable dictionaries difficult to build comprehensively.

The proposed approach is to combine the approaches presented in Chapters 4 and 5, building a physics-cognizant machine learning approach. By using machine learning methods to select optimal parameters for feature filtering on the acoustic color features, we can create a feature set with less dependence on target and sensor orientation. By leveraging the powerful clustering presented in Chapter 4, optimal feature manifolds can then be identified from the superset of features generated using the method of Chapter 5.

6.2 Technical Approach

A large number of candidate Gabor kernels were generated for evaluation as potential feature filters per the approach laid out in Chapter 5. The resulting filtered features were then clustered per the approach laid out in Chapter 4 to generate ellipsoidal feature manifolds in the transformed space. The resulting ellipsoids are pictured in figure 6.1. The same data set without the feature filtering is presented in figure 6.2. The red regions denote distinguishing characteristics in this new transformed space. These new ellipsoids look visually striking and seem to be significantly more comprehensive of the range of possible sensor positions.

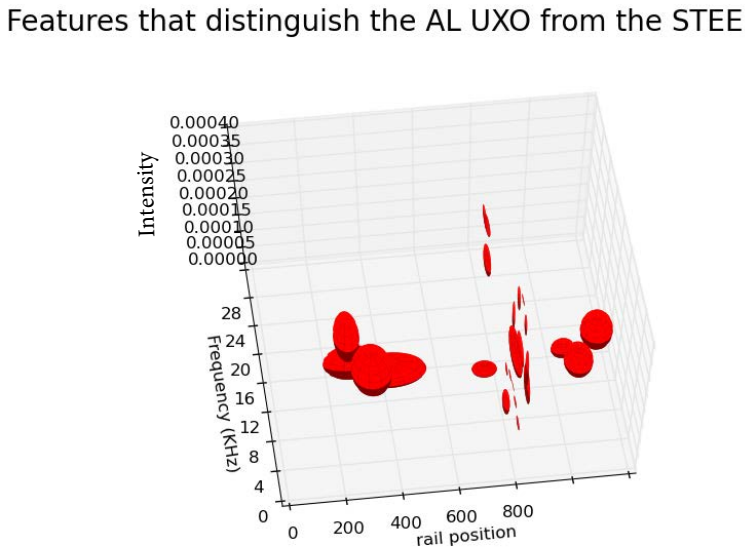


FIGURE 6.1: *Ellipsoids generated from the filtered acoustic color features that distinguish between a steel and aluminum UXO. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids.*

The ellipsoids for the selected Gabor kernel were all selected and culled using the method described in Chapter 4. Following the machine learning approach described there, several neural networks were trained. The first neural networks were trained exactly as in Chapter 4. The ellipsoids from one set of filtered features were used as the

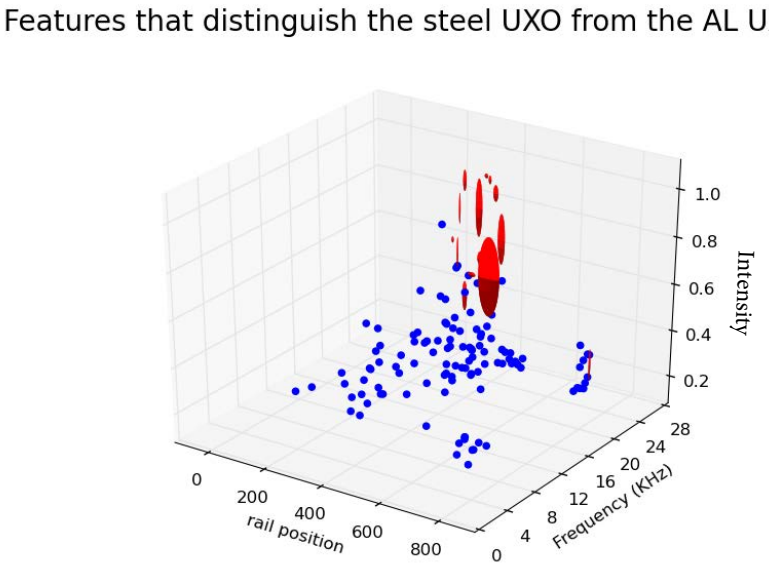


FIGURE 6.2: *Ellipsoids generated from the unfiltered acoustic color features that distinguish between a steel and aluminum UXO. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids.*

input neurons. The different filtered features were then combined into one meta neural network, with input neurons representing the ellipsoids from several filtered sets.

Other Ellipsoids are provided for example purposes here. Figures 6.3, 6.4, and 6.5 show the distinct set of ellipsoids that differentiate some of the classes.

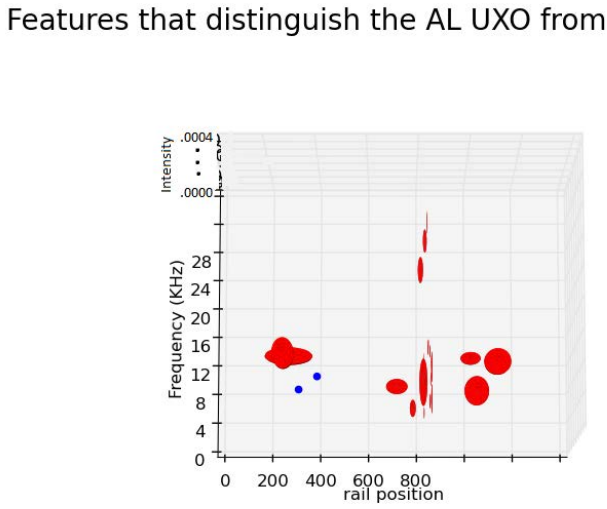


FIGURE 6.3: *Ellipsoids generated from the filtered features to distinguish an aluminum UXO from a rock. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids.*

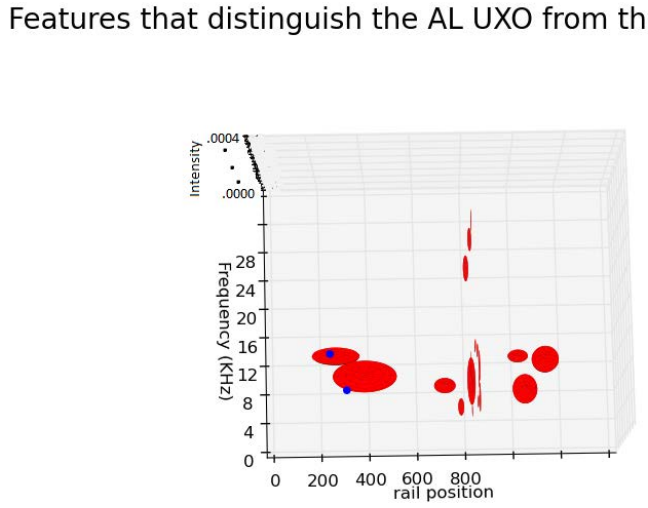


FIGURE 6.4: *Ellipsoids generated from the filtered features to distinguish an aluminum UXO from a steel UXO. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids.*

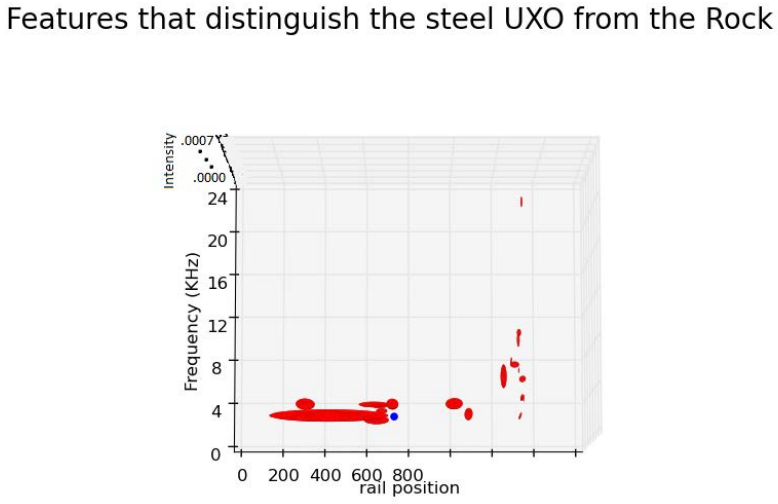


FIGURE 6.5: *Ellipsoids generated from the filtered features to distinguish an steel UXO from a rock. Red regions represent distinguishing ellipsoids. Blue regions represent discarded ellipsoids.*

6.3 Results

All neural networks were evaluated via a 5-fold cross validation. Under 5-fold cross validation, the data is randomly divided into five equally sized subsets. Each of the sets is held out for evaluation while the other four are used for training. This helps to reduce possible sources of data bias. The results are presented in confusion matrices to provide a clear picture of accuracy including false positives and negatives. All confusion matrices are presented for the full range of sensor positions. Figure 6.7 shows the improvement provided by the ellipsoid feature clustering over the original direct SVM approach presented in Chapter 5, included here in figure 6.6. Four of the transformed sets were combined with one neural network, and the results are presented in figure 6.8. These results indicate excellent classification performance across a wide variety of sensor and object orientations.

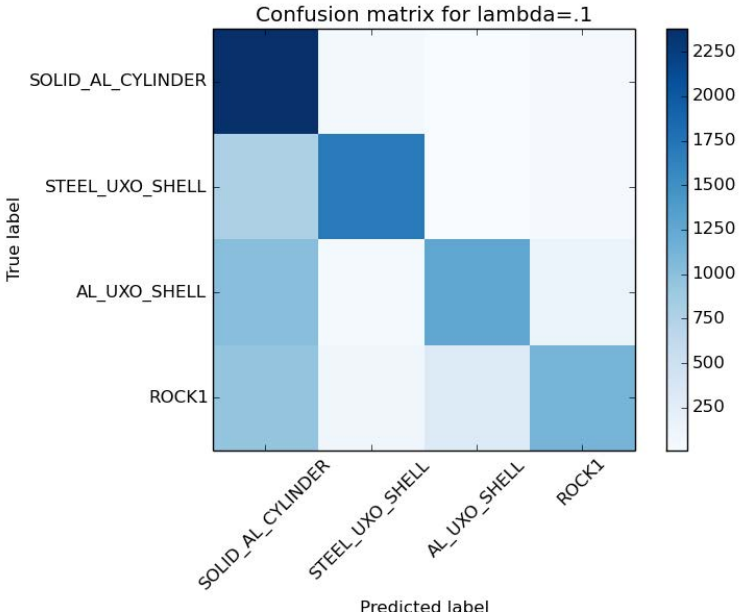


FIGURE 6.6: *Confusion matrix for the direct classification approach used in Chapter 5. Accuracy of 64.5%.*

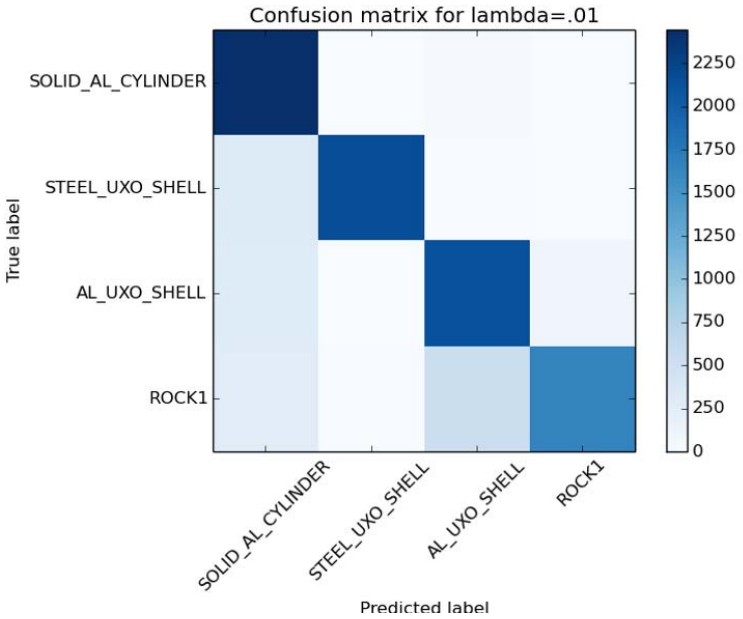


FIGURE 6.7: *Confusion matrix for the results of a neural network trained on a single feature set. Accuracy of 84.3%.*

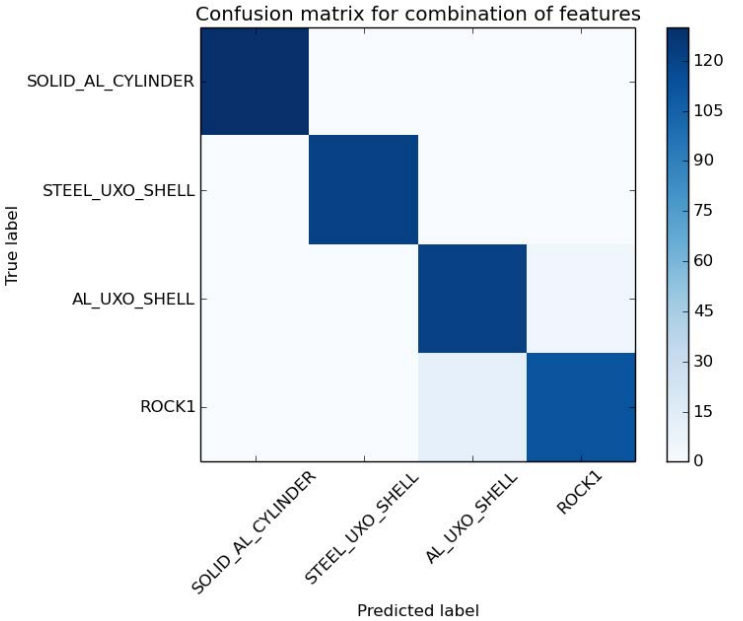


FIGURE 6.8: *Confusion matrix for the results of a neural network trained using four different Gabor feature filters. Accuracy of 95%.*

Chapter 7

Remarks and Conclusions

7.1 Collaboration in Works Herein

Chapters 4 and 5 are both primarily derived from papers submitted for publishing. These papers were written in collaboration with Prof. Ananya Sen Gupta (University of Iowa) and Dr. Ivars Kiersteins (NUWC). The works presented in this thesis represent the contribution of the primary author.

7.2 Scope

As previously established, the data collection itself was out of the scope of this work. As a result, problems like channel disambiguation and beam forming, while important to the problem at large, have not been addressed.

7.3 Closing statement

Our investigations proceeded in several orthogonal directions and explored complementary approaches to characterization and classification. The work took advantage of an understanding of a variety of state-of-the-art technologies and techniques while also enabling collaboration with other institutions yielding publishable material. The work is

documented and archived in the group git server for the support of future work in the environmental signal processing group.

7.4 Acknowledgment

The authors would like to thank the Applied Physics Laboratory, University of Washington for providing public-domain active sonar field data used for validation of our techniques in this work.

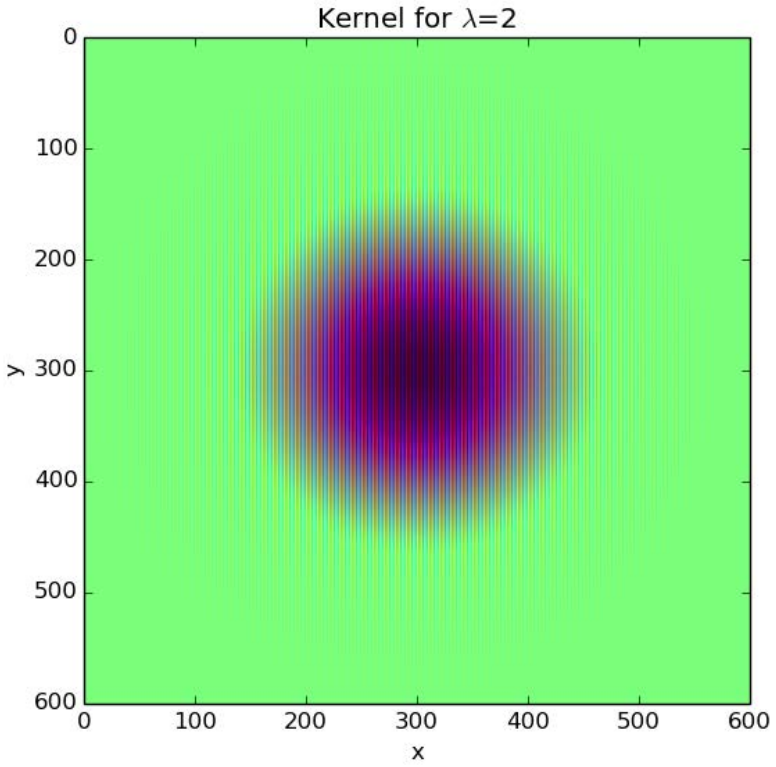
Appendix

Gabor Transform Examples

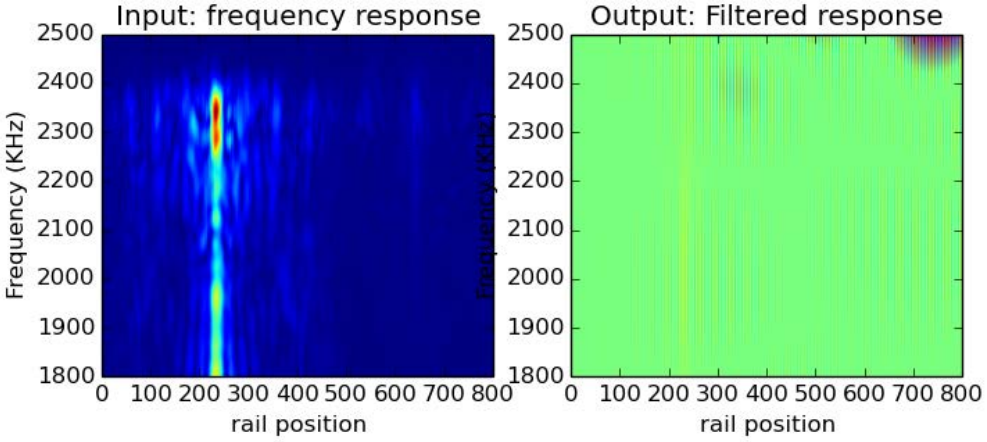
The investigation in Chapter 5 yielded a significant number of candidate kernels. For the purpose of brevity, only a few were included in the chapter for illustrative purposes. Here is presented a more representative sample of the kernels that were evaluated positively by our evaluation technique as described in Chapter 5. All figures include the input, output, and kernel. The input is the periodogram of the signal across a range of sensor positions (i.e. a sensor sweep). The targets include UXOs (unexploded ordinance), cylinders, and rocks. The angle specified in the title refers to target orientation with respect to the sensor. Some objects have large changes in acoustic color response depending on the face presented to the sensor. The kernel is a 2D Gabor kernel generated as described in Eq. 5.1. Most of the kernel parameters were kept constant across these examples, with the only exception being $f = \frac{1}{\lambda+\psi}$, which was varied and found different significant features. Other combinations of parameters (e.g. θ, σ, γ) were not as successful.

Per Eq. 5.1, $\sigma = 1$ and $\theta = 90°$, and $\gamma = 1$.

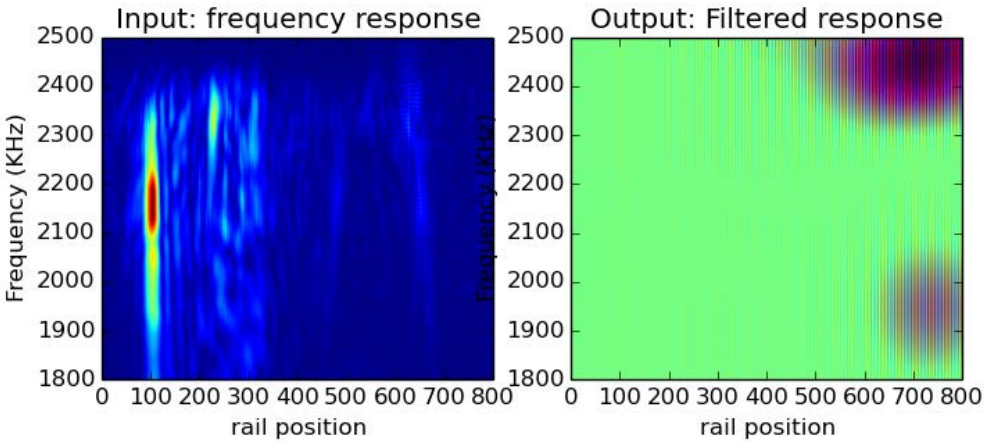
A.1 $\lambda = 2$



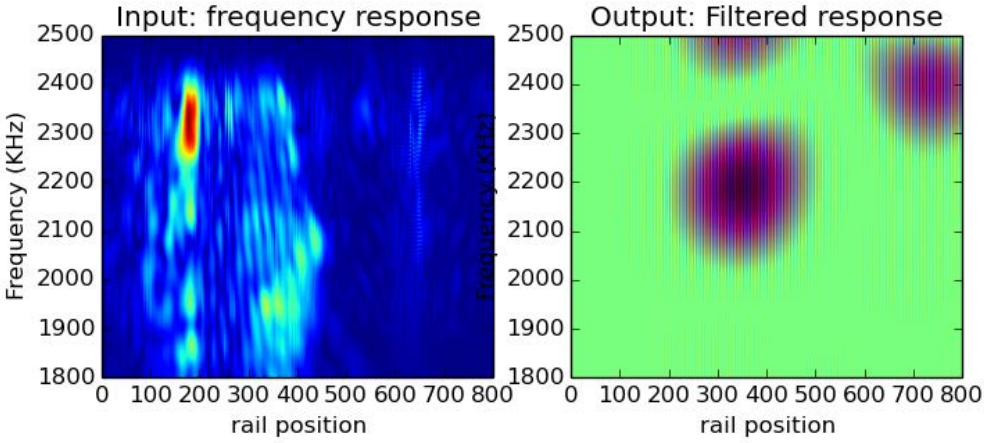
AL Cylinder at 0°, λ=2



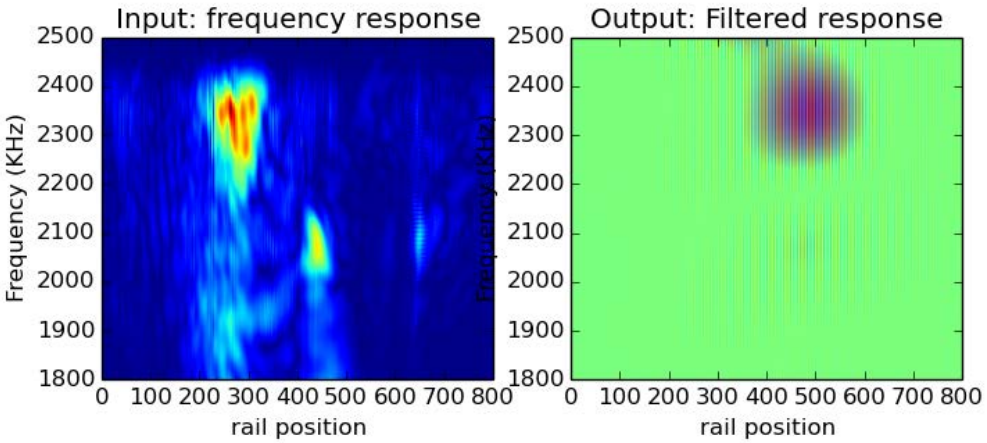
AL Cylinder at 20°, λ=2



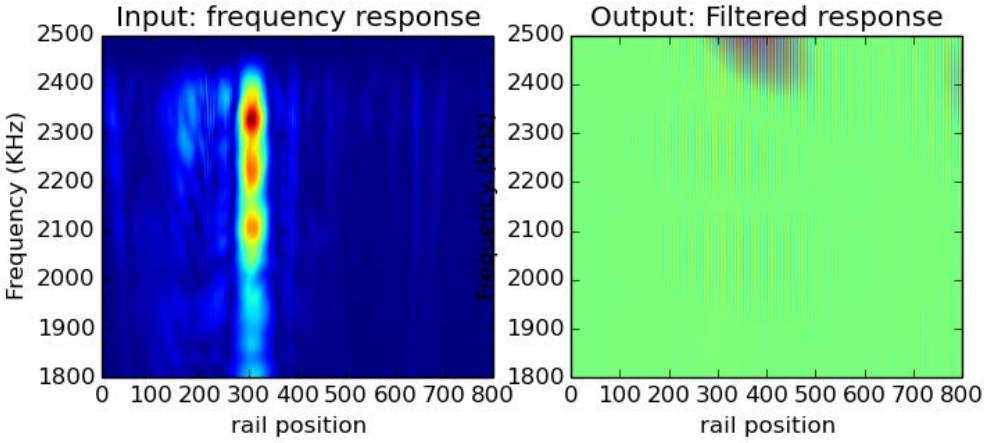
AL Cylinder at 40°, λ=2



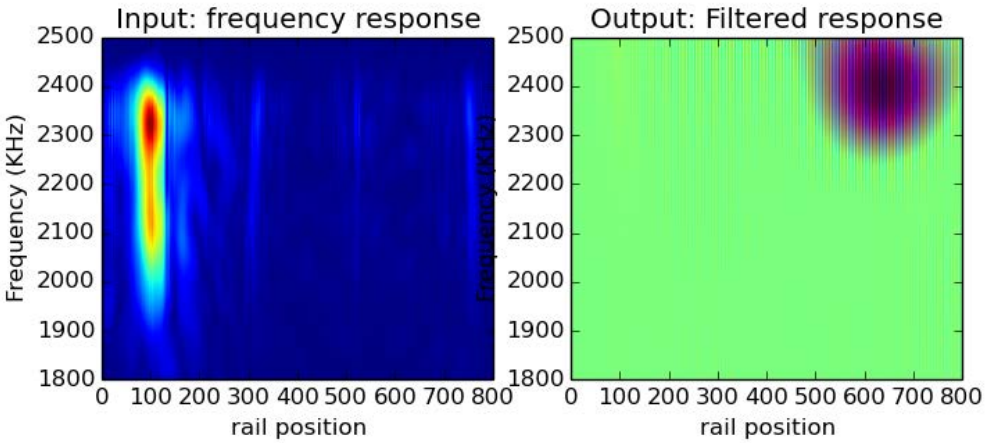
AL Cylinder at 60°, λ=2



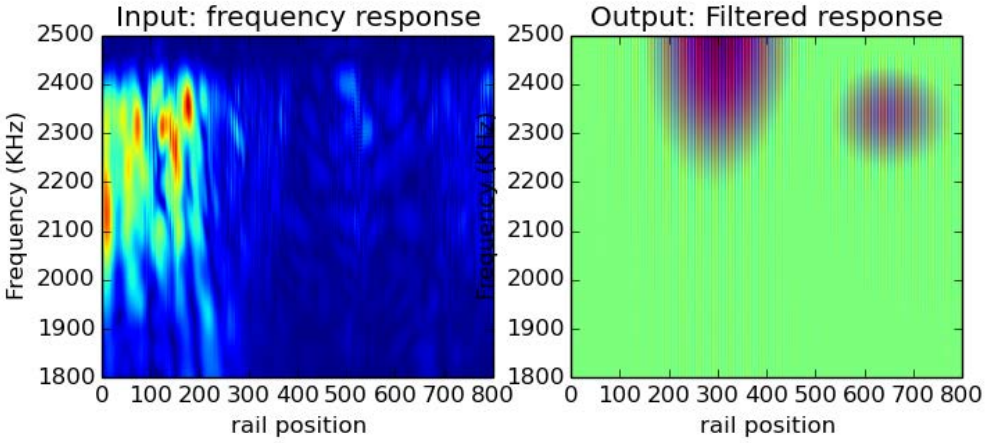
AL Cylinder at 80°, λ=2



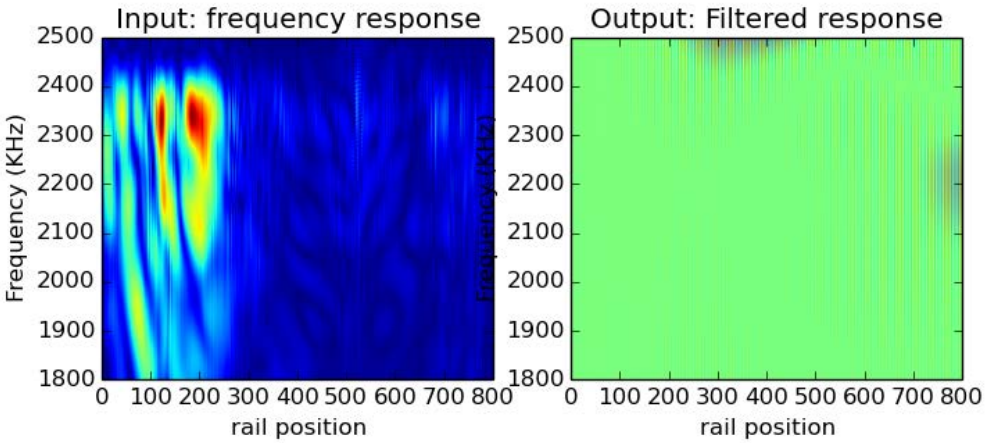
AL UXO at 0°, λ=2



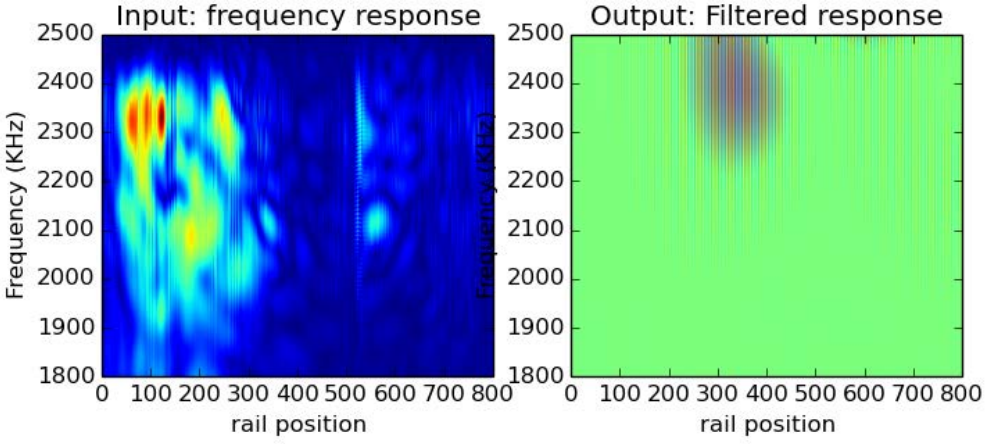
AL UXO at 20°, λ=2



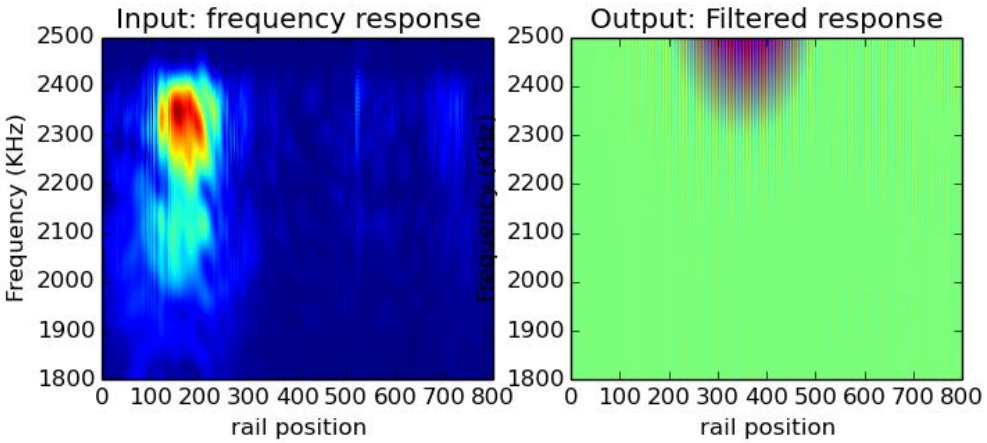
AL UXO at 40°, λ=2



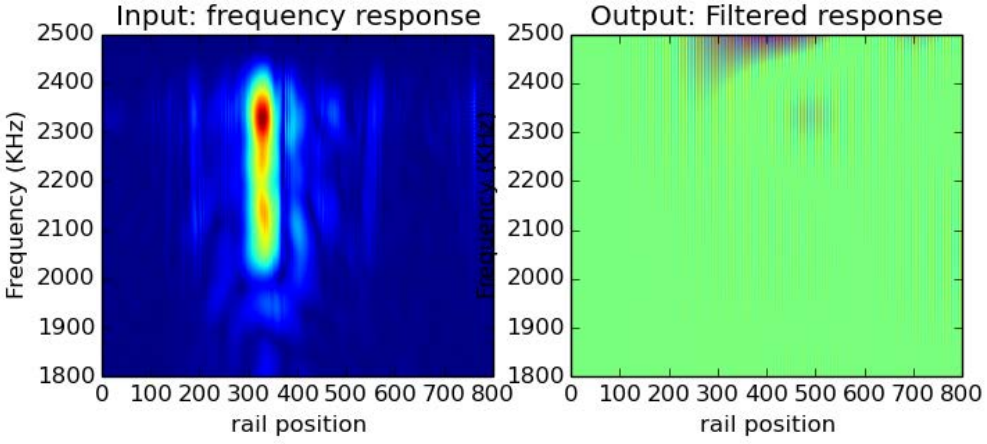
AL UXO at 60°, λ=2



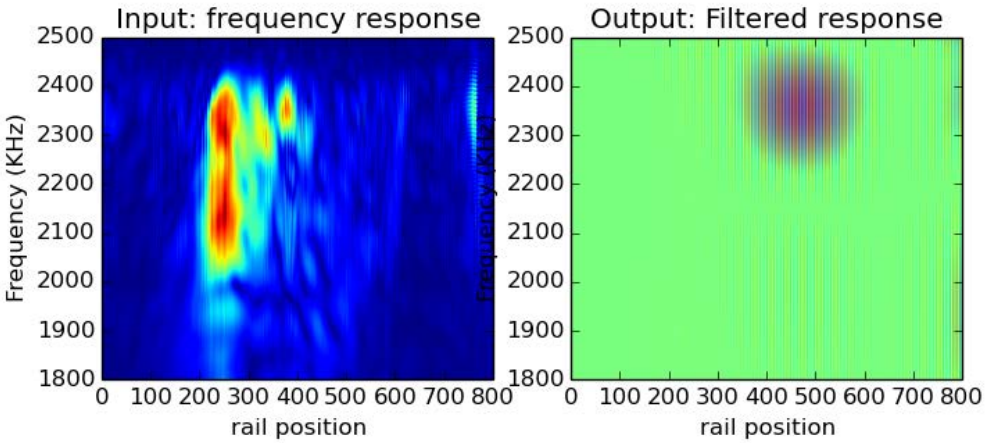
AL UXO at 80°, λ=2



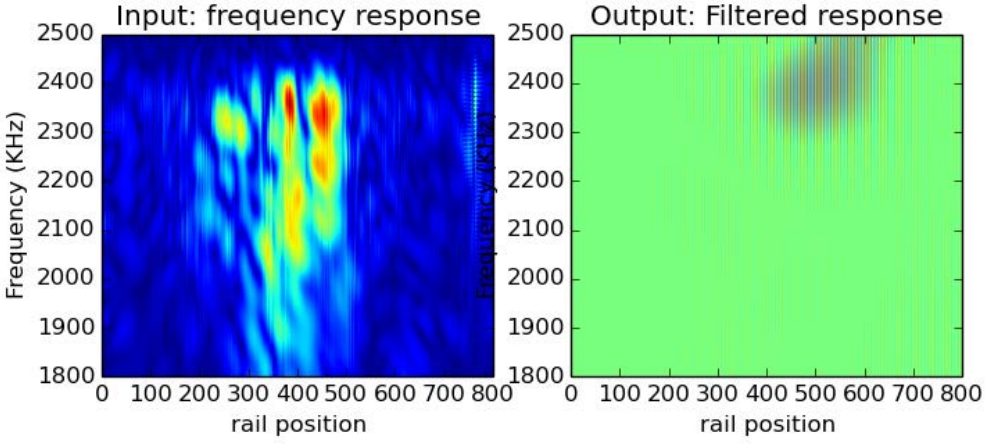
Steel UXO at 0°, λ=2



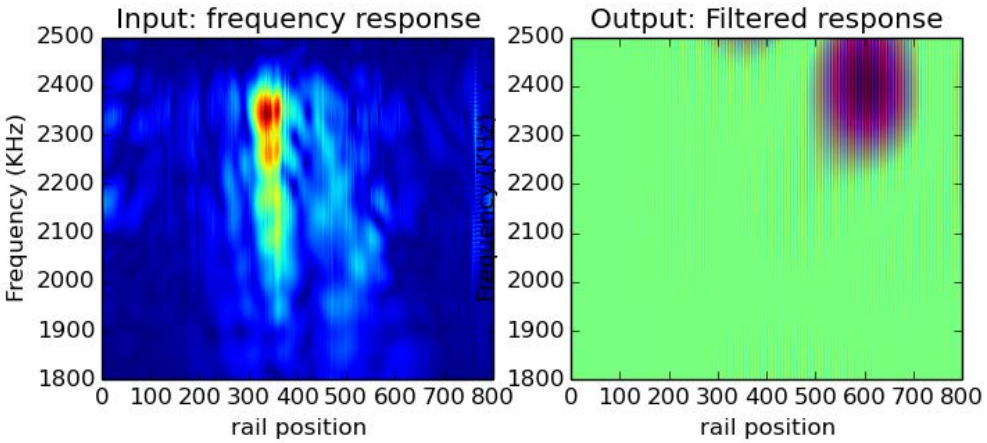
Steel UXO at 20°, λ=2



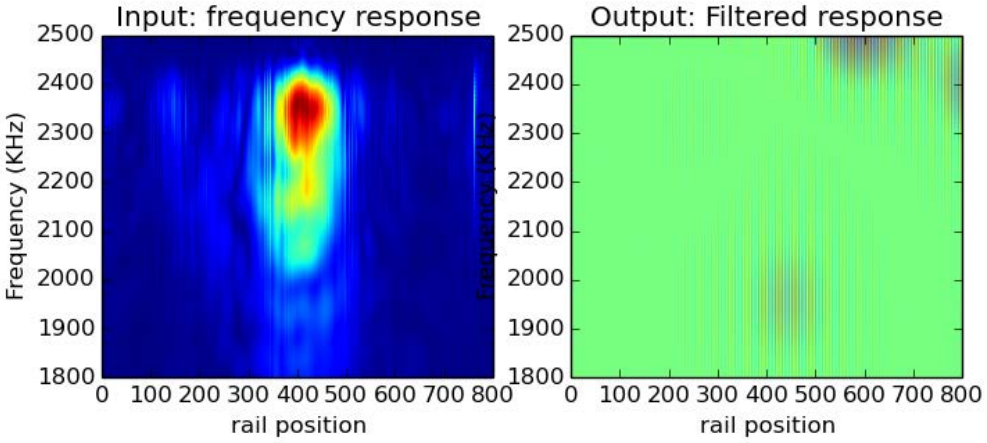
Steel UXO at 40°, λ=2



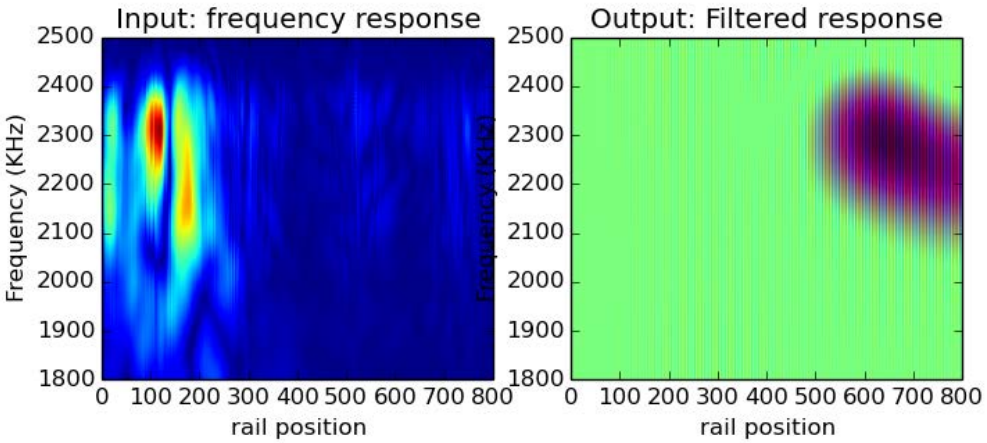
Steel UXO at 60°, λ=2



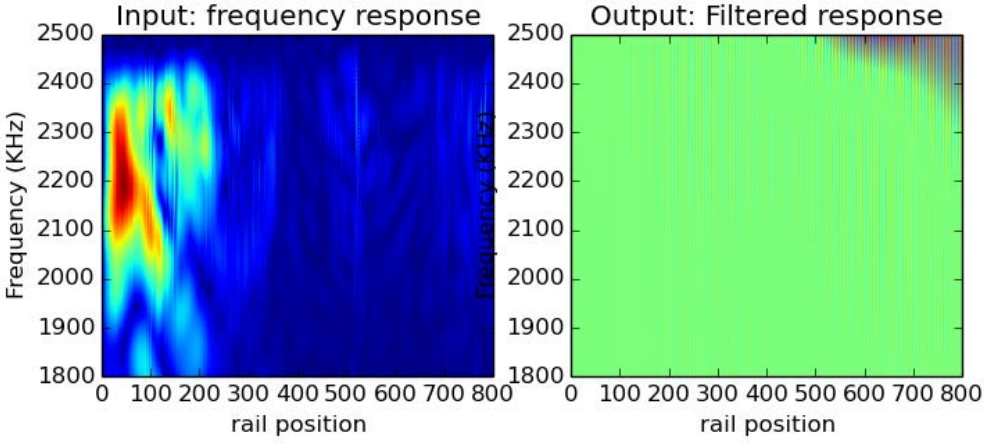
Steel UXO at 80°, λ=2



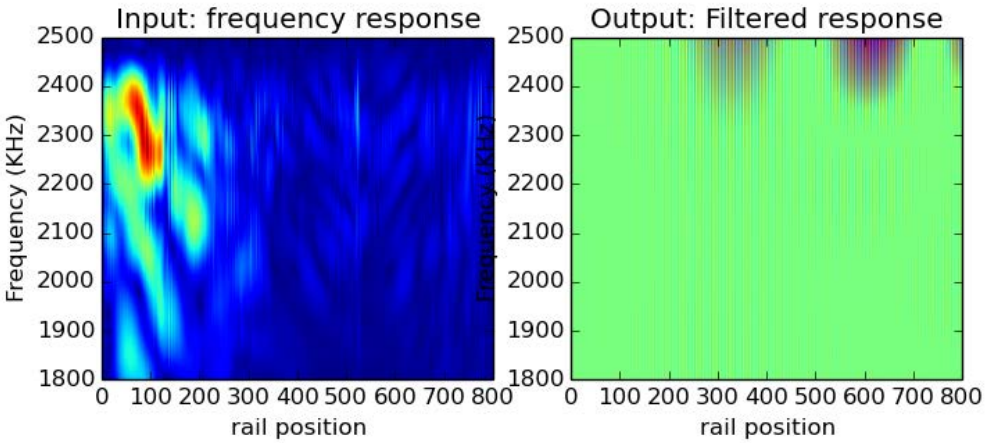
Rock at 0°, λ=2



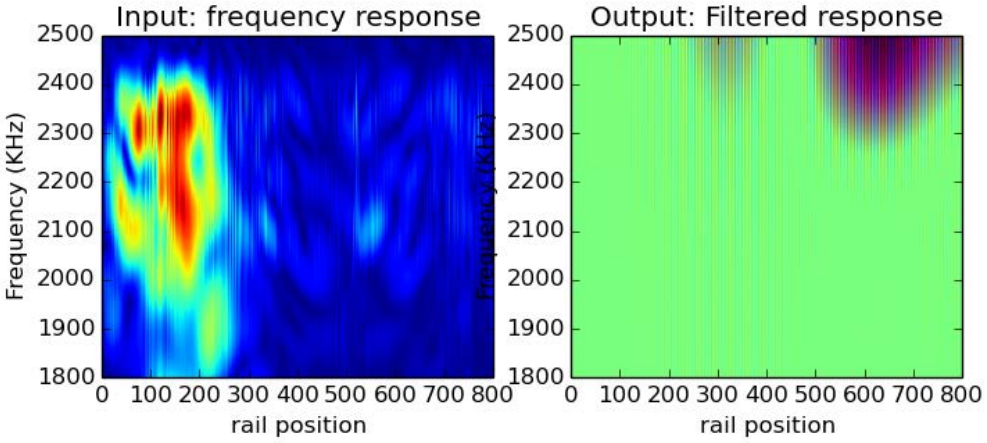
Rock at 20°, λ=2



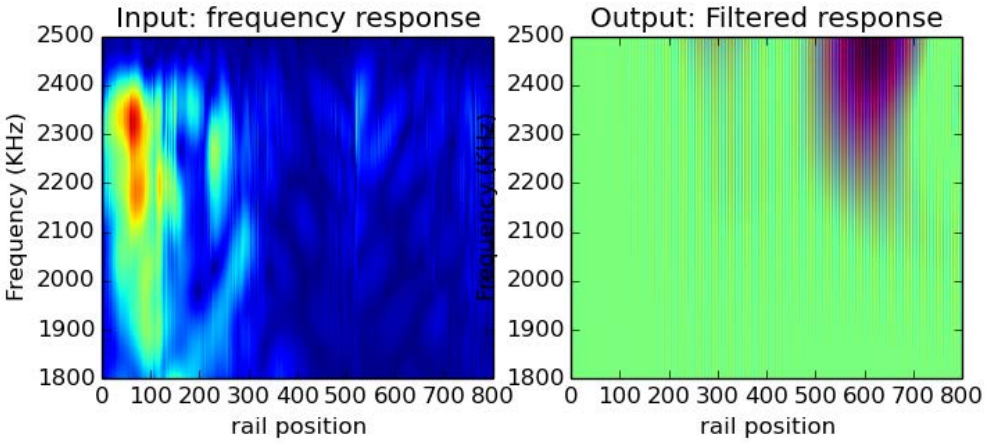
Rock at 40°, λ=2



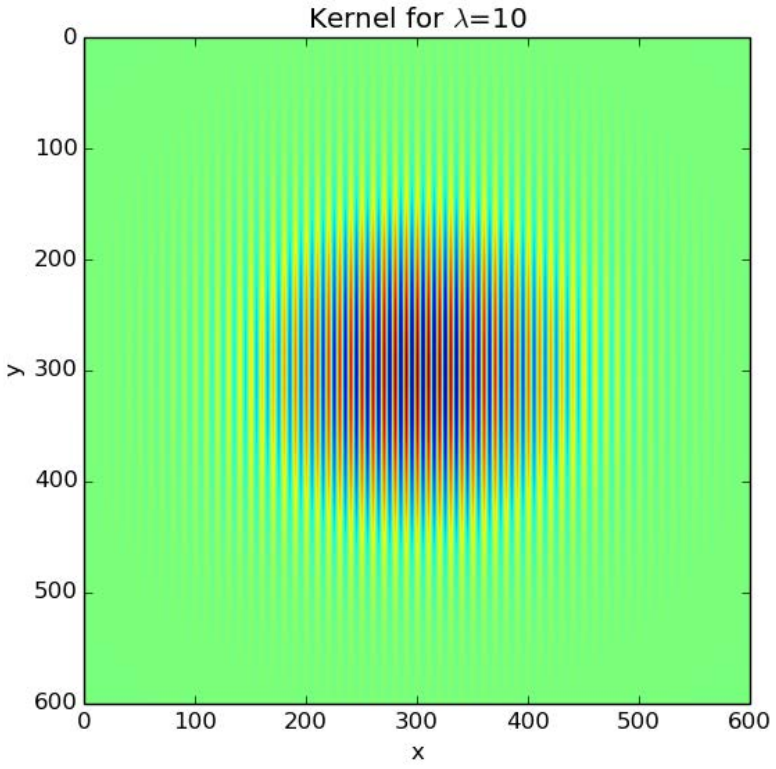
Rock at 60°, λ=2



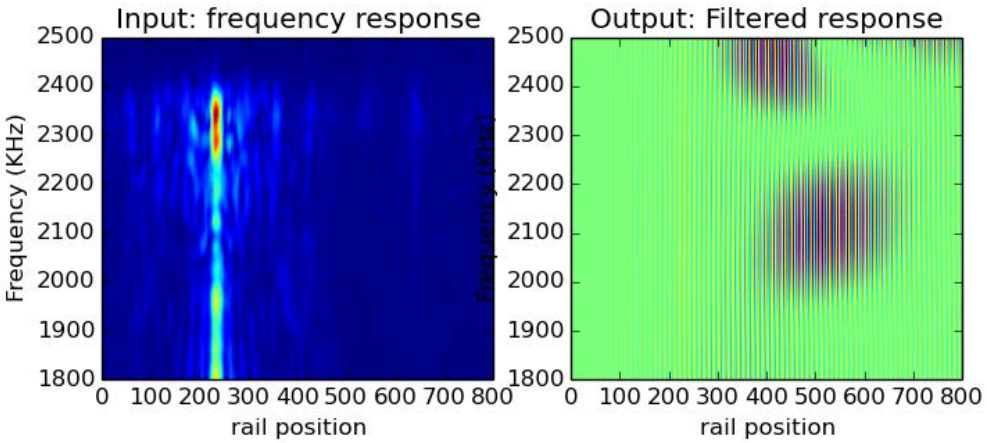
Rock at 80°, λ=2



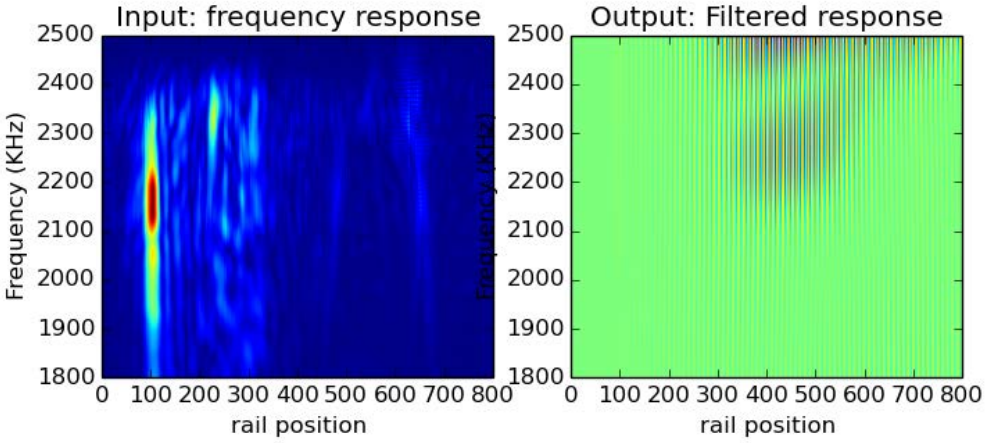
A.2 $\lambda = 10$


Kernel for λ=10
0
100
200
300
400
500
600
y
0
100
200
300
400
500
600
x

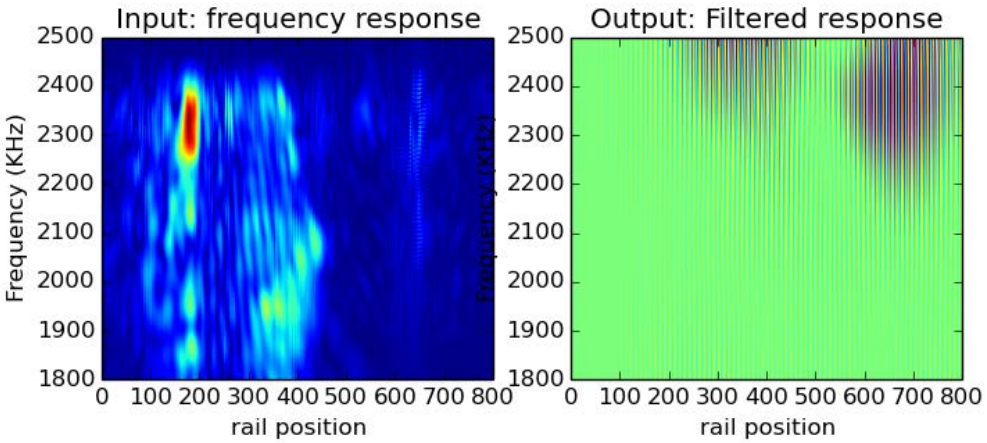
AL Cylinder at 0°, λ=10



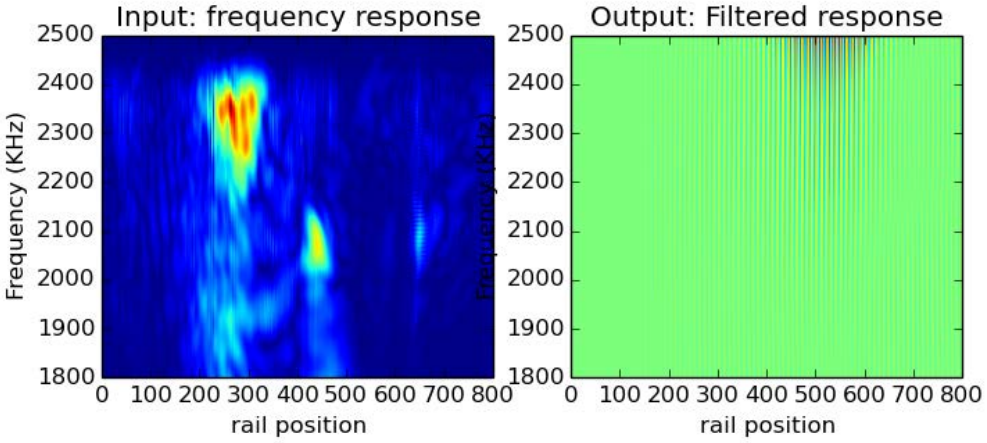
AL Cylinder at 20°, λ=10



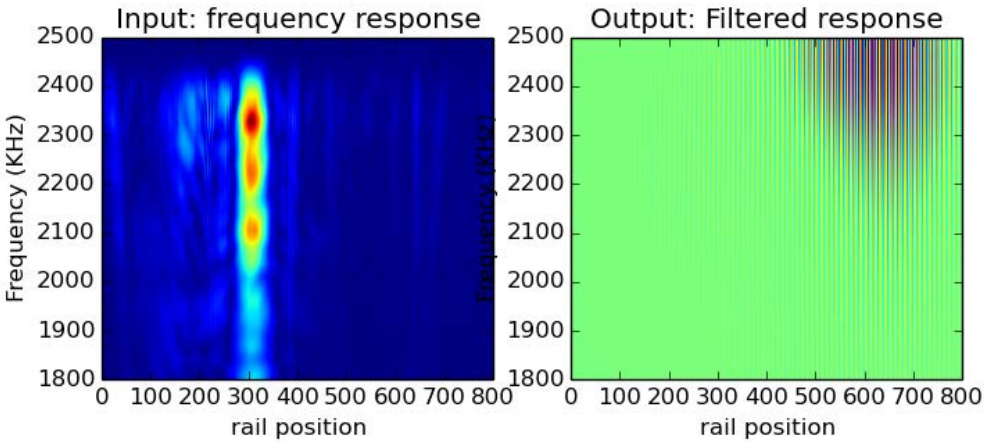
AL Cylinder at 40°, λ=10



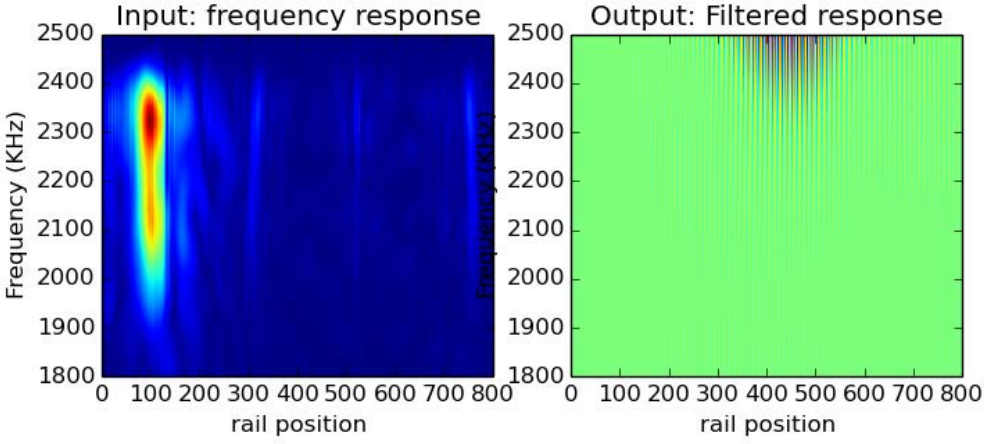
AL Cylinder at 60°, λ=10



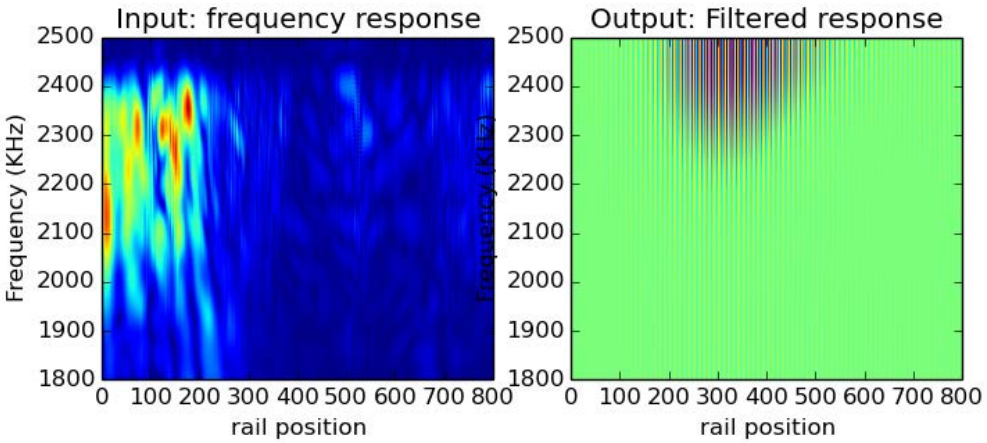
AL Cylinder at 80°, λ=10



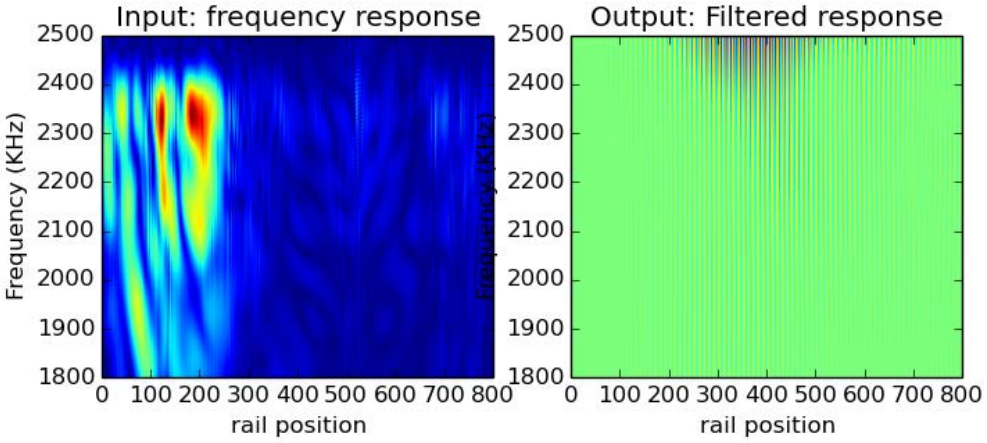
AL UXO at 0°, λ=10



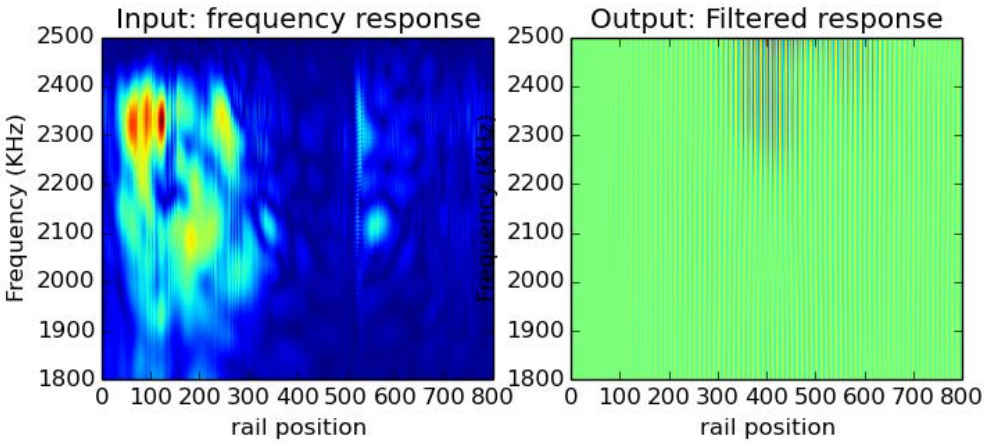
AL UXO at 20°, λ=10



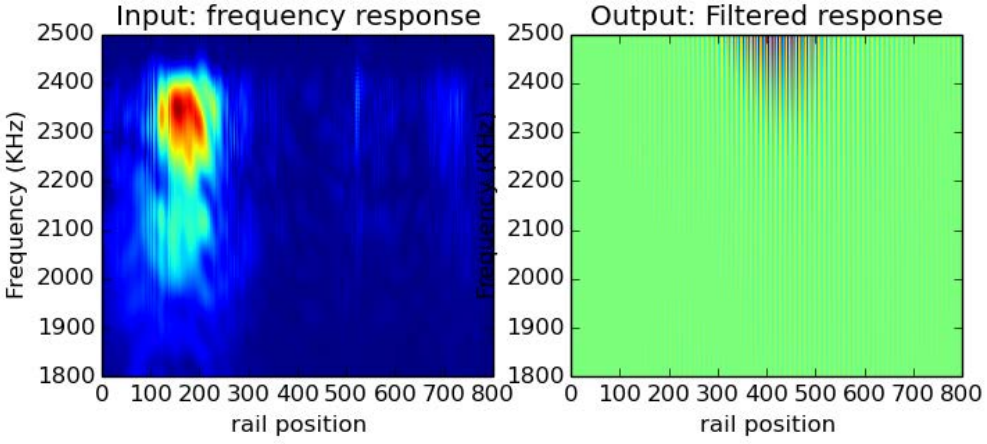
AL UXO at 40°, λ=10



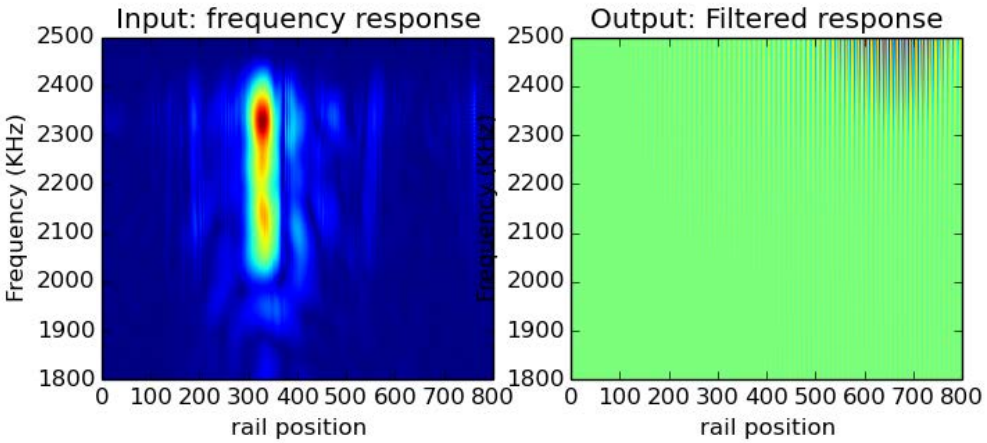
AL UXO at 60°, λ=10



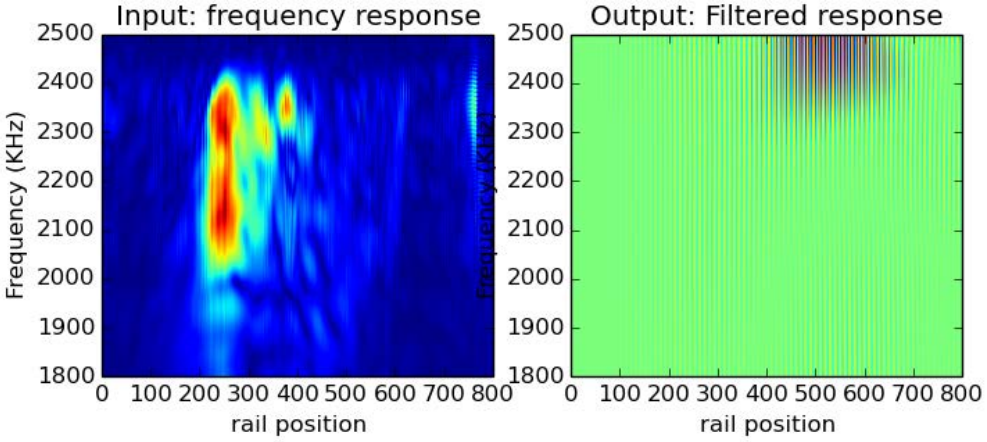
AL UXO at 80°, λ=10



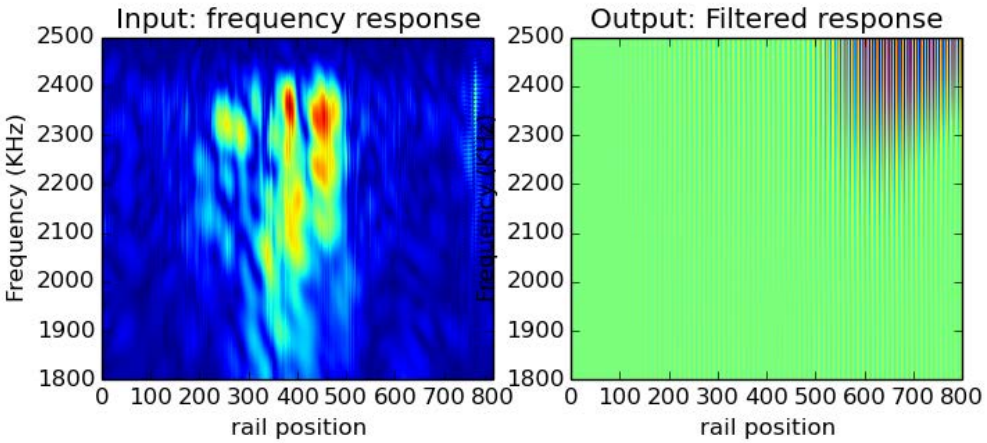
Steel UXO at 0°, λ=10



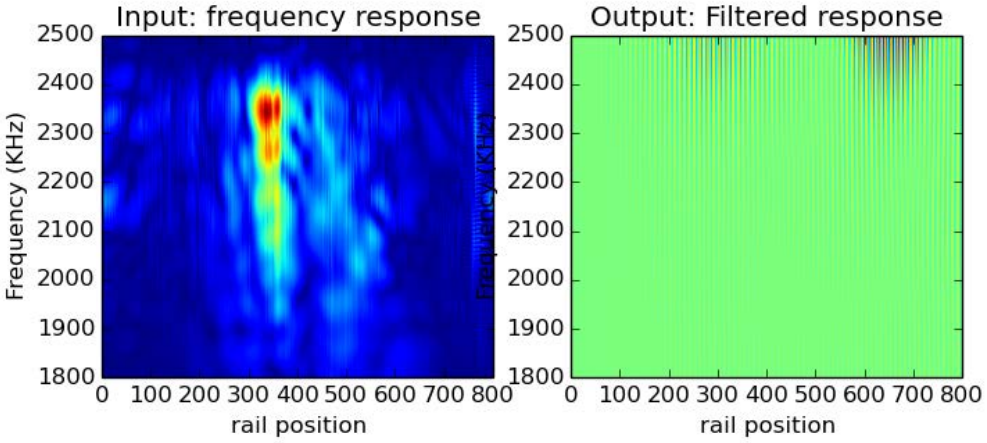
Steel UXO at 20°, λ=10



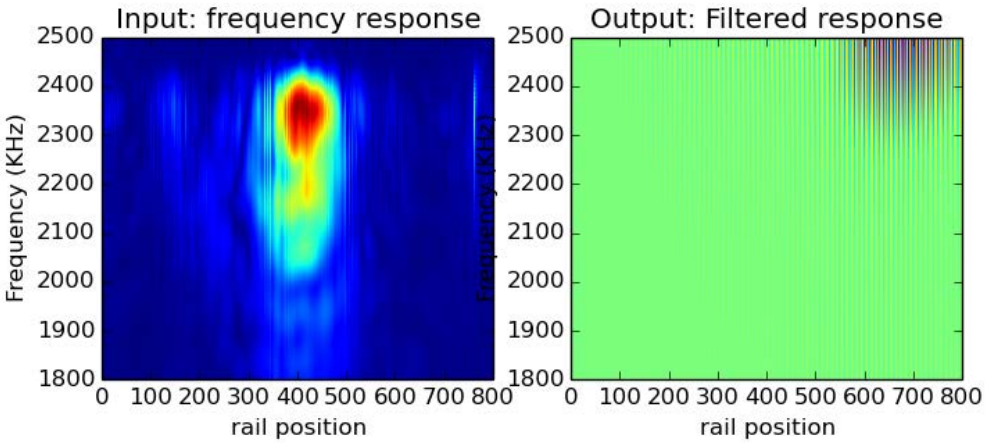
Steel UXO at 40°, λ=10



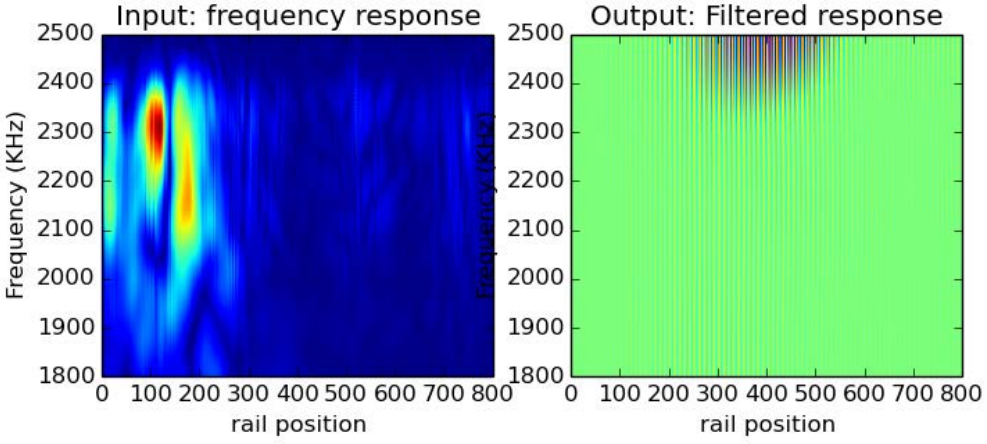
Steel UXO at 60°, λ=10



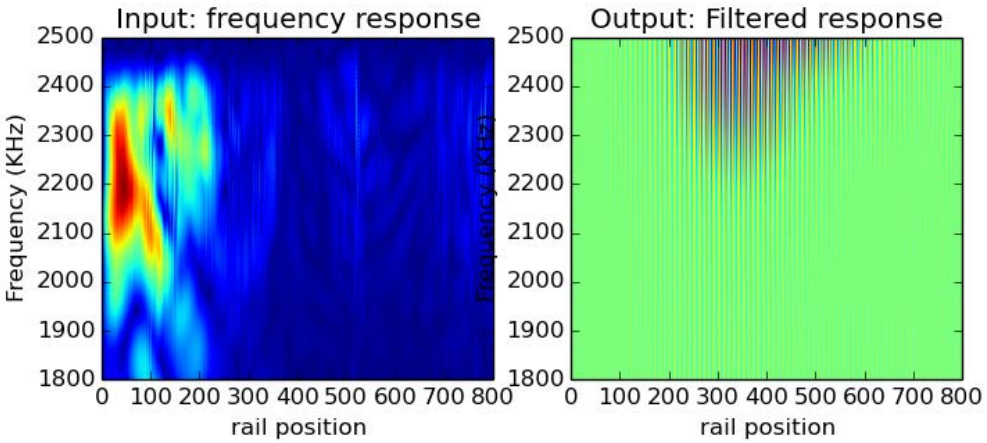
Steel UXO at 80°, λ=10



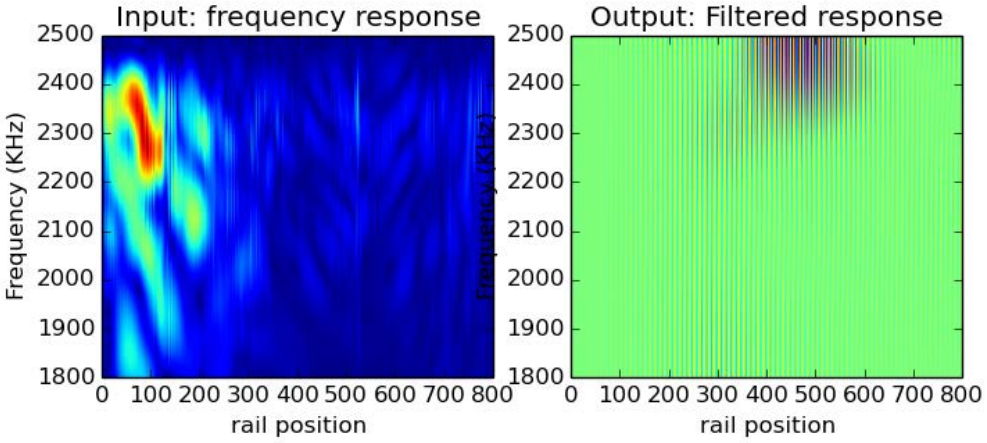
Rock at 0°, λ=10



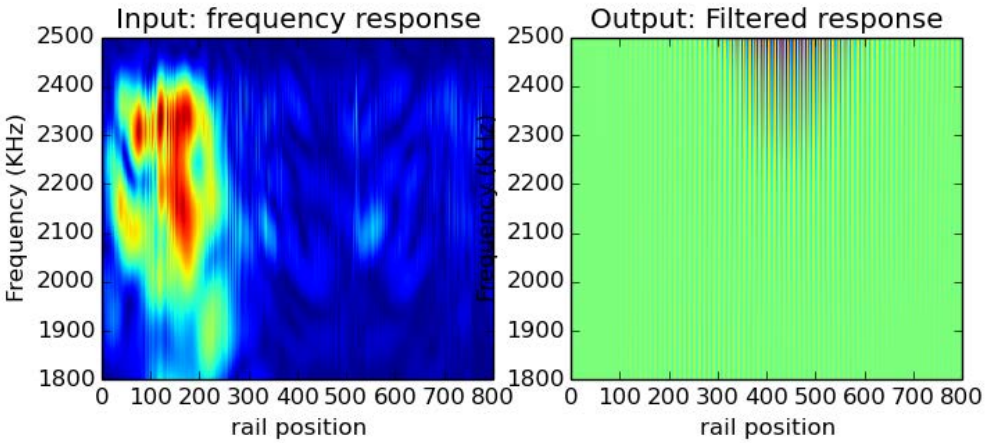
Rock at 20°, λ=10



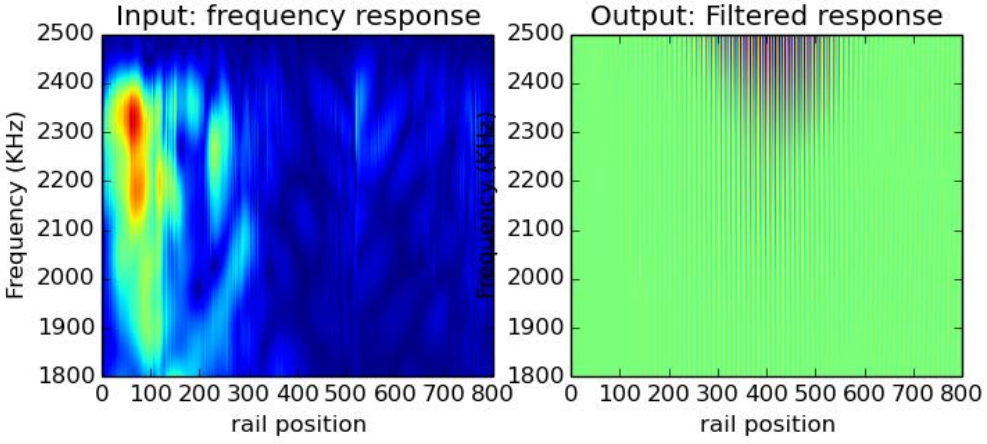
Rock at 40°, λ=10



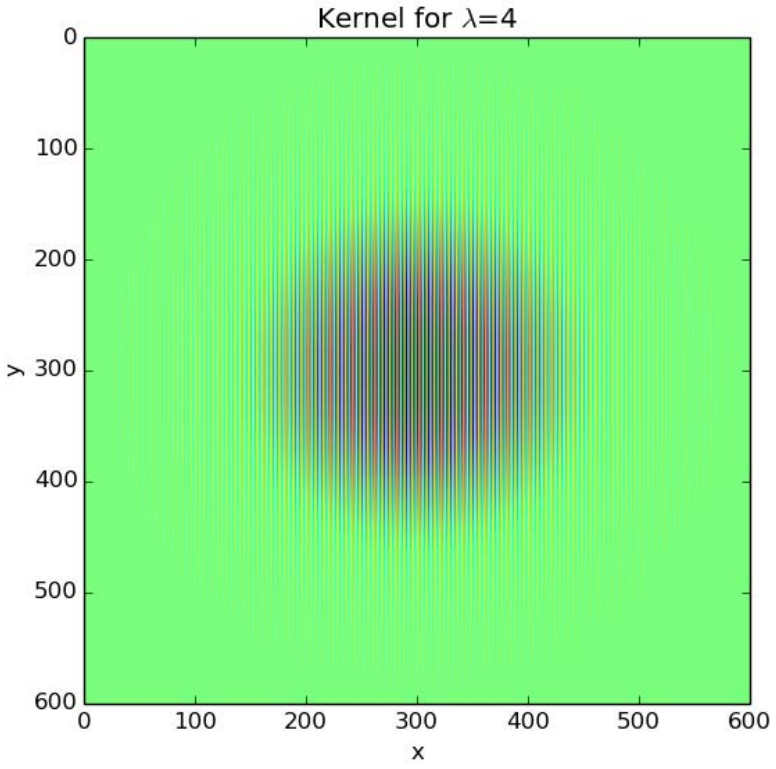
Rock at 60°, λ=10



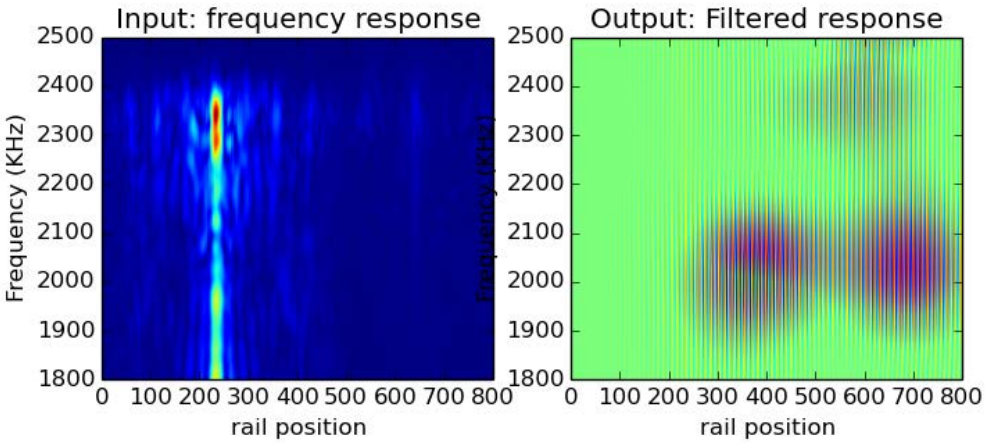
Rock at 80°, λ=10



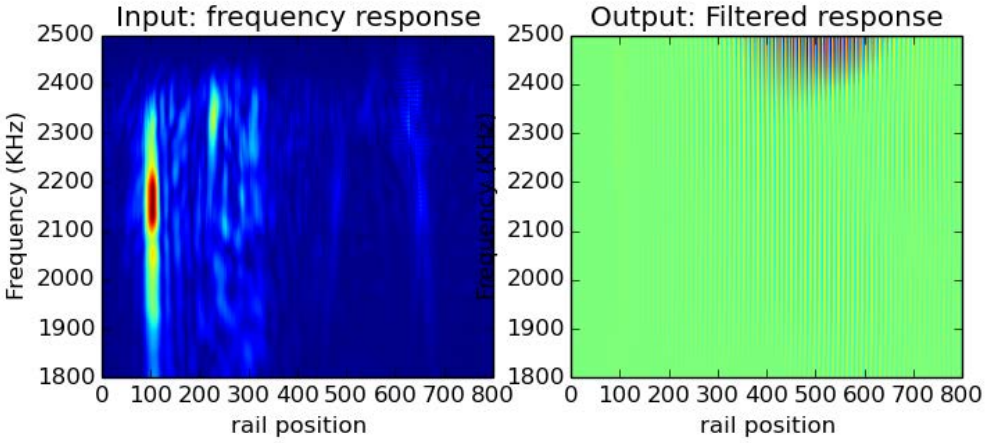
A.3 $\lambda = 4$


Kernel for λ=4
0
100
200
300
400
500
600
y
0
100
200
300
400
500
600
x

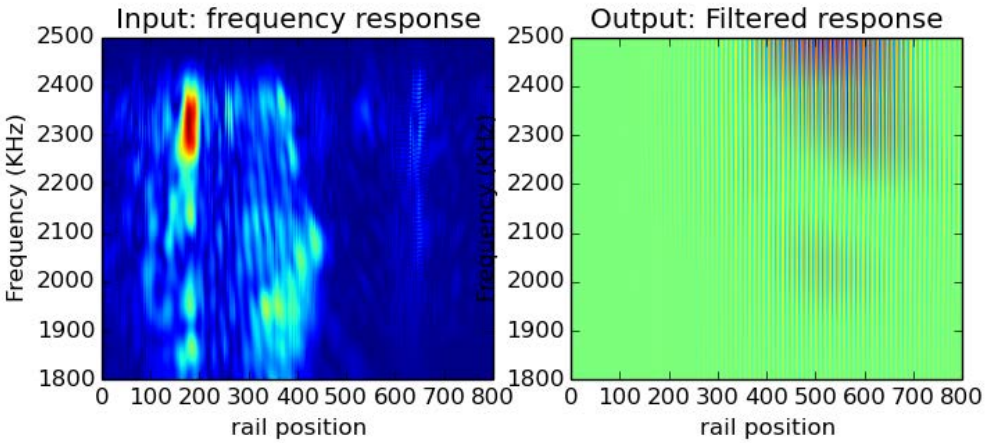
AL Cylinder at 0°, λ=4



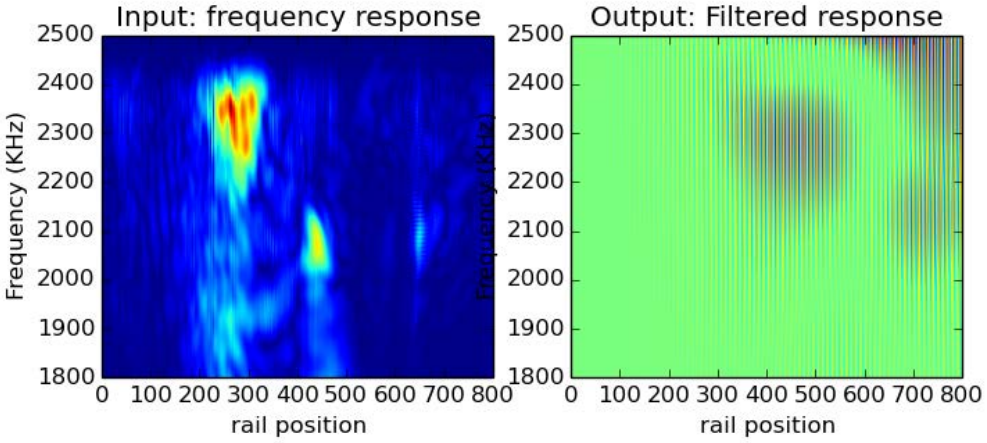
AL Cylinder at 20°, λ=4



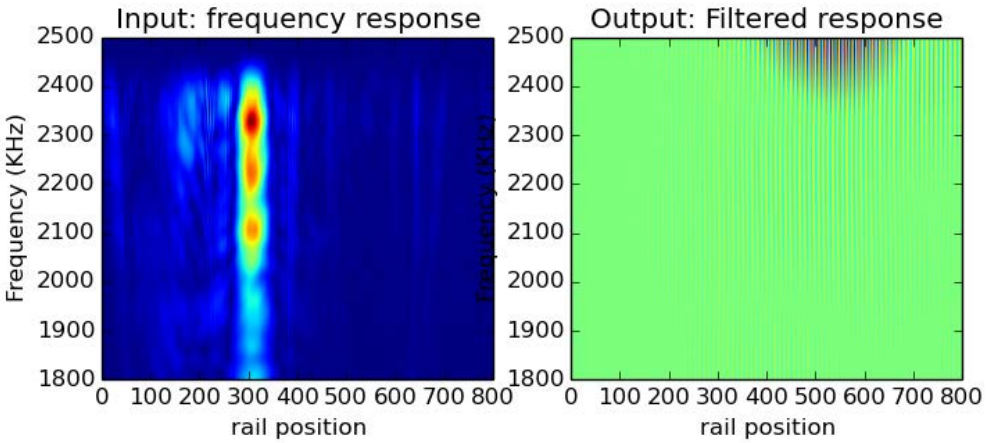
AL Cylinder at 40°, λ=4



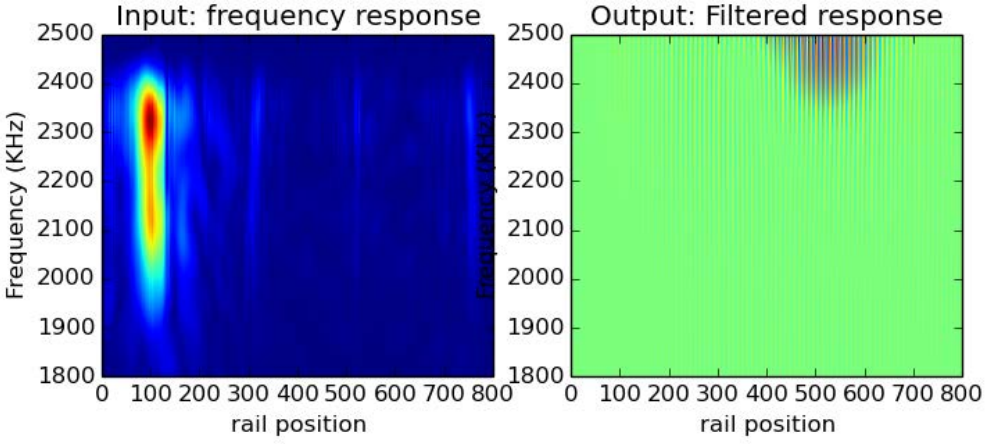
AL Cylinder at 60°, λ=4



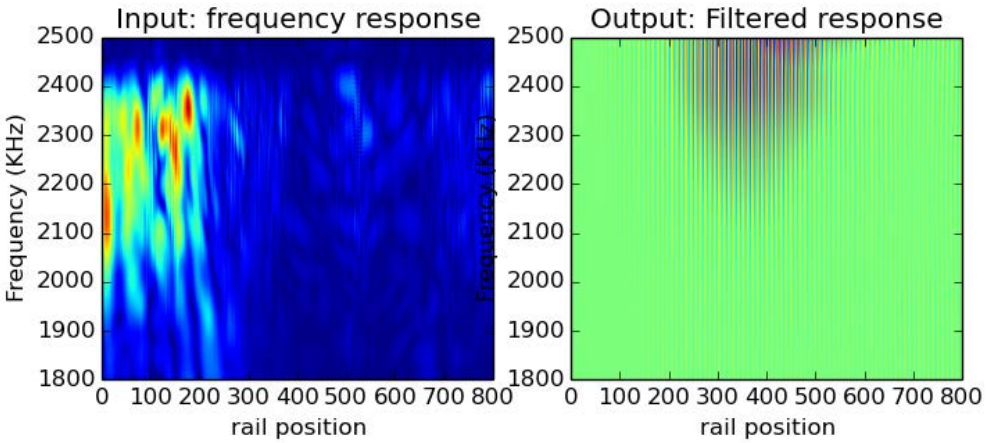
AL Cylinder at 80°, λ=4



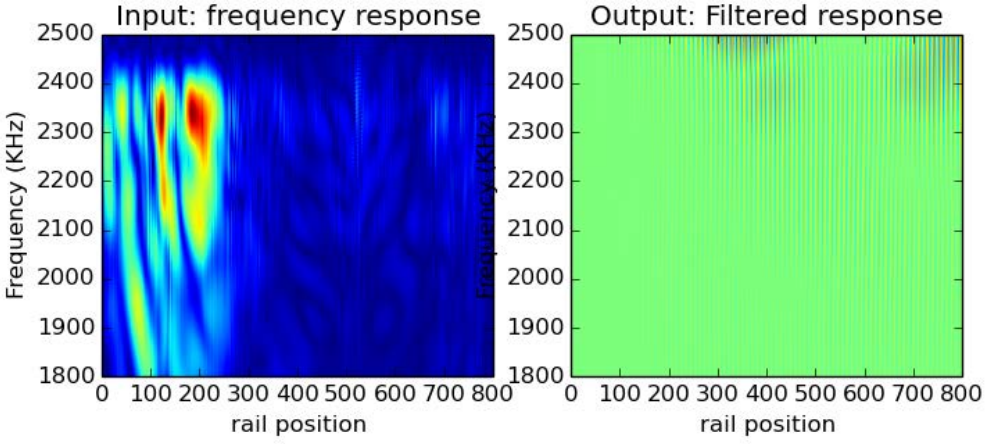
AL UXO at 0°, λ=4



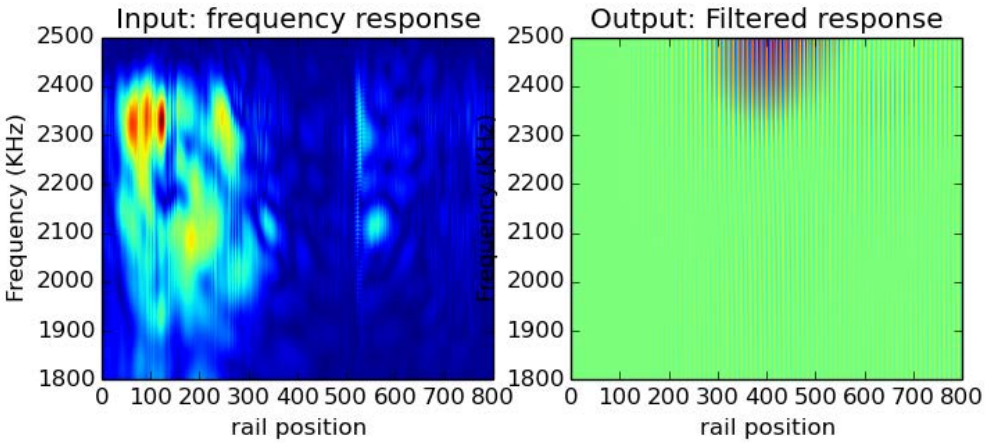
AL UXO at 20°, λ=4



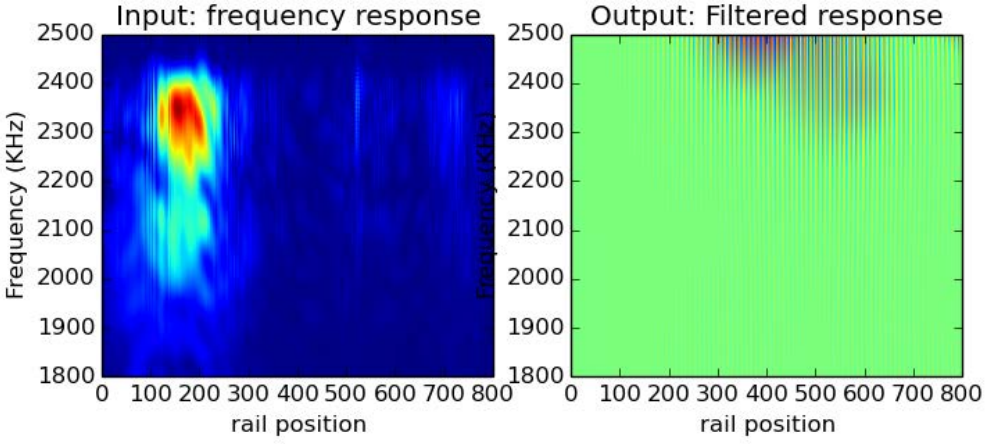
AL UXO at 40°, λ=4



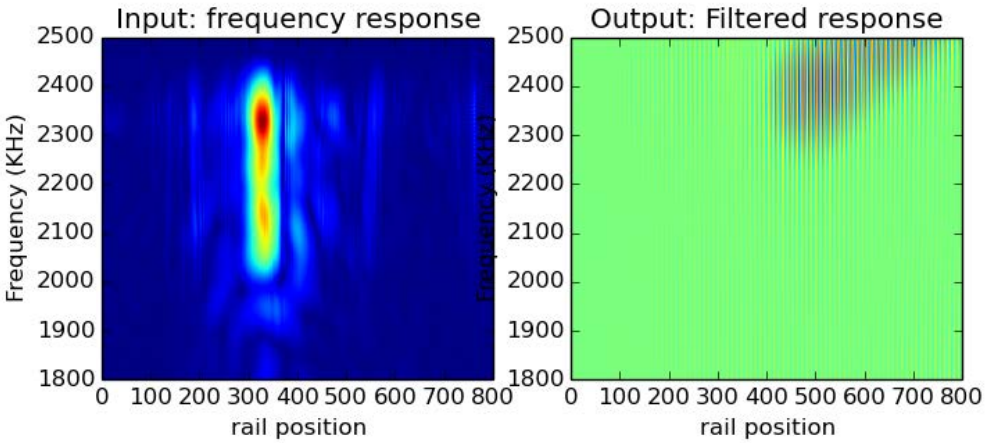
AL UXO at 60°, λ=4



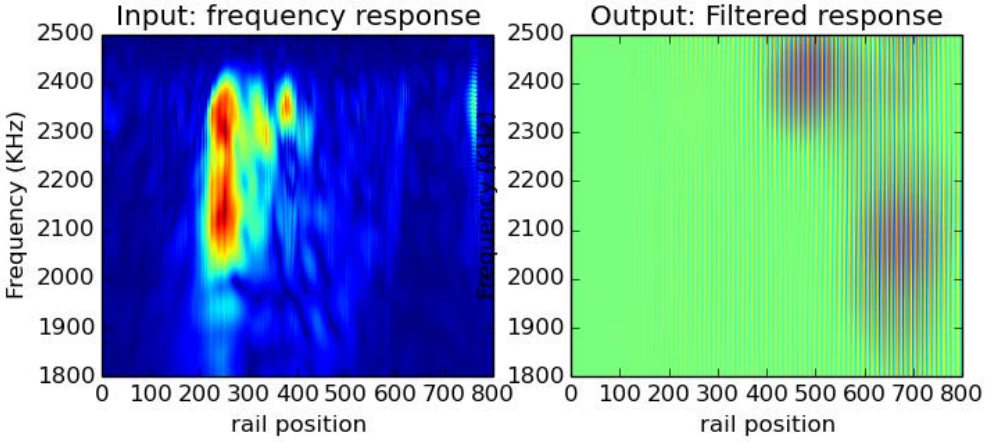
AL UXO at 80°, λ=4



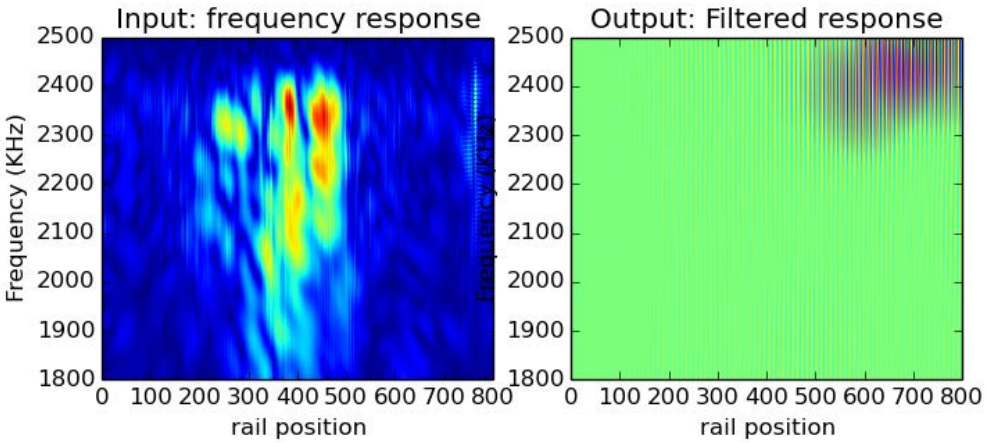
Steel UXO at 0°, λ=4



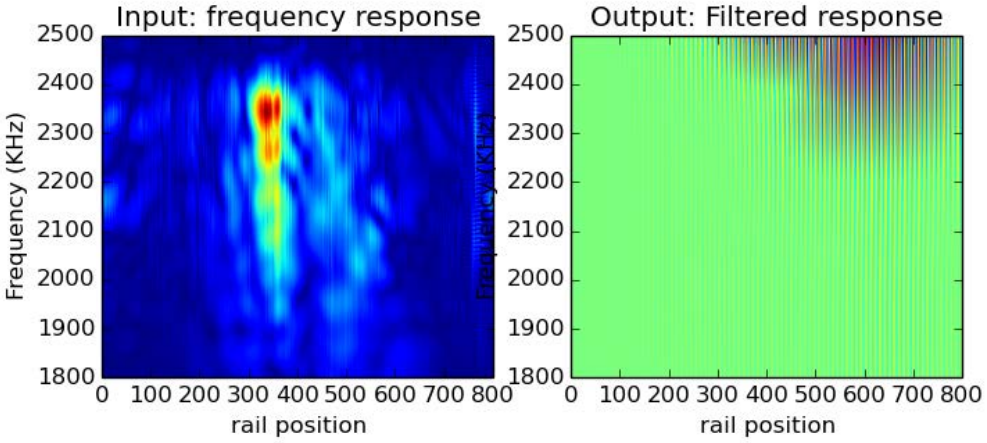
Steel UXO at 20°, λ=4



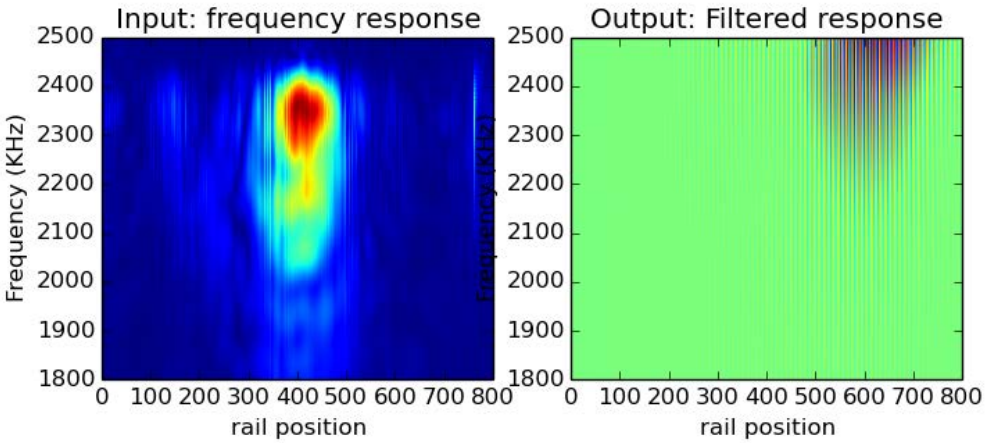
Steel UXO at 40°, λ=4



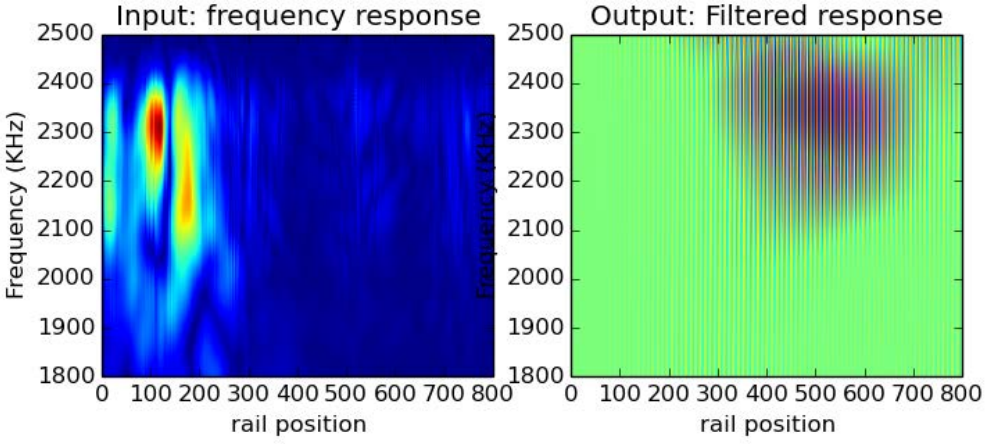
Steel UXO at 60°, λ=4



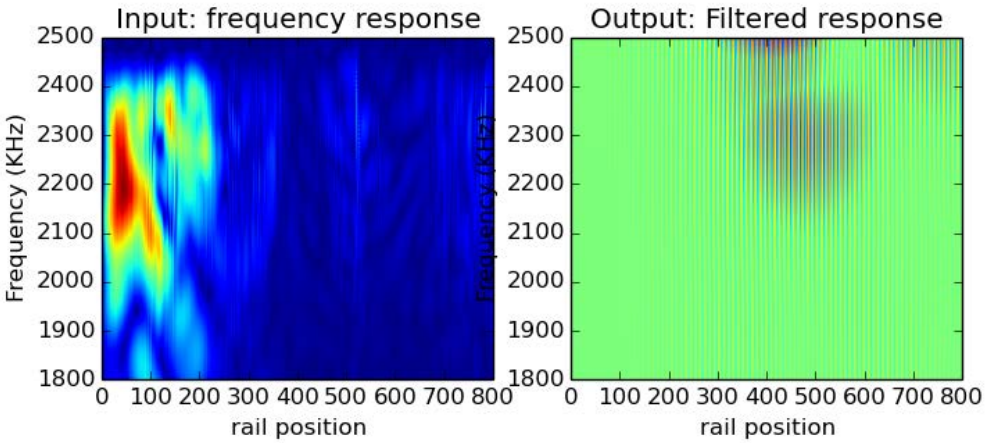
Steel UXO at 80°, λ=4



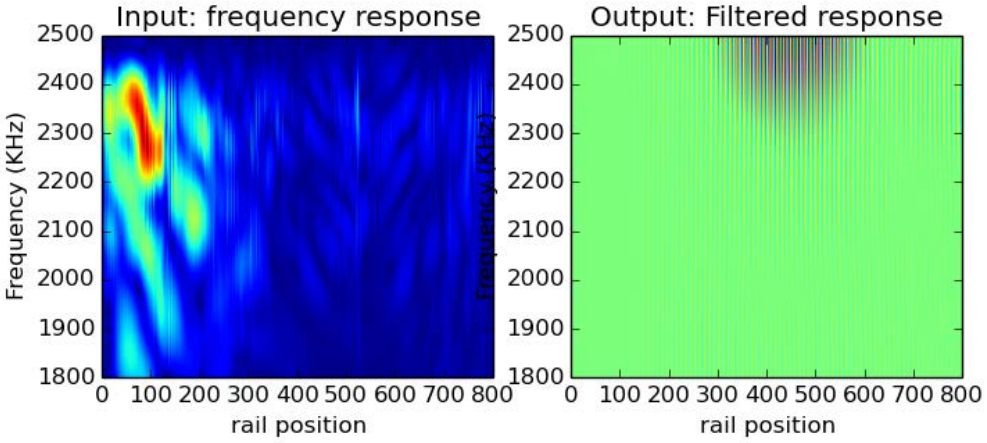
Rock at 0°, λ=4



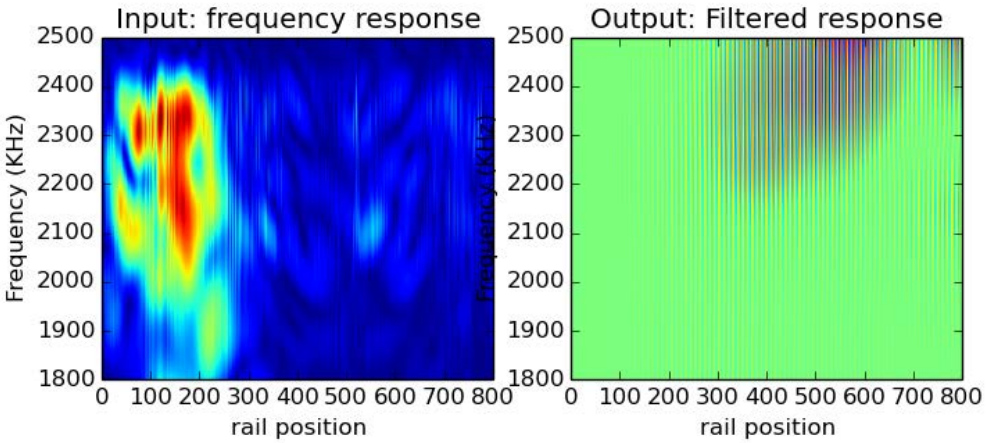
Rock at 20°, λ=4



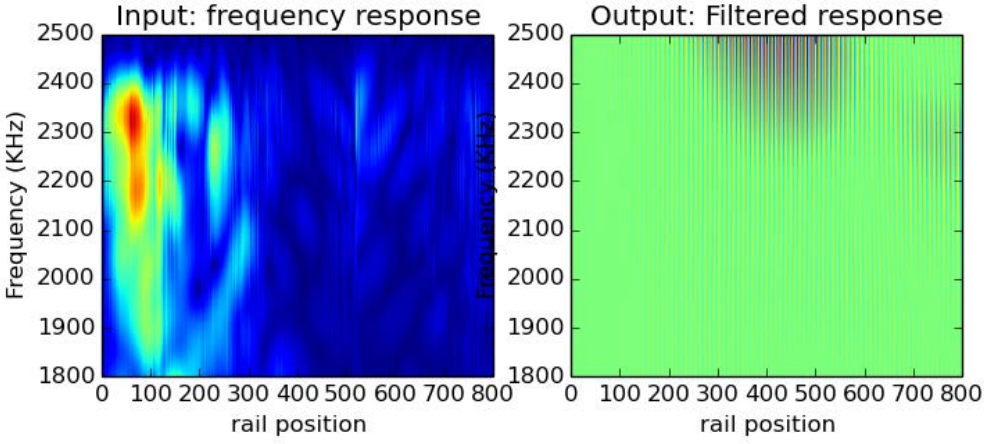
Rock at 40°, λ=4



Rock at 60°, λ=4



Rock at 80°, λ=4



Bibliography

[1] W. Knight, R. Pridham, and S. Kay. Digital signal processing for sonar. *Proc. of IEEE*, 69(11):1451–1506, November 1981.

[2] A. B. Baggeroer, W. A. Kuperman, and Henrik Schmidt. Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem. *The Journal of the Acoustical Society of America*, 83.2:571–587, 1988.

[3] Nabin S. Sharma, John R. Buck, and James A. Simmons. Trading detection for resolution in active sonar receivers. *The Journal of the Acoustical Society of America*, 130.3:1272–1281, 2011.

[4] Ph. Blondel. Automatic mine detection by textural analysis of cots sidescan sonar imagery. *International Journal of Remote Sensing*, 21.16:3115–3128, 2000.

[5] F. Maussang, M. Rombaut, J. Chanussot, A. Hetet, and M. Amate. Fusion of local statistical parameters for buried underwater mine detection in sonar imaging. *EURASIP Journal on Advances in Signal Processing*, 2008.

[6] Geoffrey A. Hollinger, Brendan Englot, Franz Hover, Urbashi Mitra, and G. Sukhatme. Uncertainty-driven view planning for underwater inspection. *In Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 4884–4891, May 2012. doi: 10.1109/ICRA.2012.6224726.

[7] N.; Mitra Kumar and U.; Narayanan S. Robust object classification in underwater sidescan sonar images by using reliability-aware fusion of shadow features. *Oceanic Engineering, IEEE Journal of*, PP(99):1,15.

[8] et al. Badiey, Mohsen. Acoustic multipath arrivals in the horizontal plane due to approaching nonlinear internal waves. *The Journal of the Acoustical Society of America*, 129(4):EL141–EL147, 2011.

[9] A. D. Pierce and Acoustics: An. Introduction to its physical principles and applications acoustical society of america. *Acoustical Society of America*, pages 428–429, 1991.

[10] P. M. Baggenstoss. Specular decomposition of active sonar returns using combined waveforms. *Aerospace and Electronic Systems, IEEE Transactions on , vol*, 49(4): 2509–2521, October 2013.

[11] Ian M. Rooney, John R. Buck, and Kathleen E. Wage. Implementing physical constraints for noise only normal mode shape estimation. *Proceedings of Meetings on Acoustics*, 19(1):–, 2013. doi: http://dx.doi.org/10.1121/1.4801396. URL `http://scitation.aip.org/content/asa/journal/poma/19/1/10.1121/1.4801396`.

[12] R. Lee Culver and David L. Bradley. On the relationship between signal bandwidth and frequency correlation for ocean surface forward scattered signals. *The Journal of the Acoustical Society of America*, 118.1:129–138, 2005.

[13] Tianzhu Meng and John R. Buck. Rate distortion bounds on passive sonar performance. *Signal Processing, IEEE Transactions on*, 58.1:326–336, 2010.

[14] John R. Buck. Fading channel capacity and passive sonar performance bounds. *Sensor Array and Multichannel Processing, 2006. Fourth IEEE Workshop on*, pages 294–298, July 2006. doi: 10.1109/SAM.2006.1706140.

[15] C. W.; Culver Jemmott, R. L.; Bissinger, and B. E.; Gaumond C. An information theoretic performance bound for passive sonar localization of a moving source.

Information Sciences and Systems (CISS), 2010 44th Annual Conference on, 2010 (44):1,5,17–19, March 2010.

[16] William S. Hodgkiss. Matched field processing. *SCRIPPS INSTITUTION OF OCEANOGRAPHY LA JOLLA CA MARINE PHYSICAL LAB*, (MPL-U-53/91), 1991.

[17] T. G. Dietterich. Ensemble methods in machine learning. In *Multiple classifier systems*, pages 1–15. Springer Berlin Heidelberg, 2000.

[18] S. M. Weiss and I. Kapouleas. An empirical comparison of pattern recognition, neural nets and machine learning classification methods. *Readings in machine learning*, pages 177–183, 1990.

[19] Cullen Schaffer. Selecting a classification method by cross-validation. *Machine Learning*, 13(1):135–143, 1993.

[20] R. Battiti. Using mutual information for selecting features in supervised neural net learning. *Neural Networks, IEEE Transactions on*, 5(4):537–550, 1994.

[21] David A. Hague, John R. Buck, and Igal Bilik. A deterministic filterbank compressive sensing model for bat biosonar. *The Journal of the Acoustical Society of America*, 129.4:2628–2628, 2011.

[22] J. P. Sessarego et al. Time-frequency analysis of signals related to scattering problems in acoustics—part i: Wigner–ville analysis of echoes scattered by spherical shell. *Wavelets, New York: Springer-Verlag*, 1987.

[23] Ivars Kirsteins Ananya Sen Gupta, Daniel Schupp. Dynamic target identification and classification based on resonance topography grouping. *Proceedings of the Asilomar Conference on Signals, Systems, and Computers (to appear)*, 2014.

[24] P. H. Carter and G. Dobeck. Classification of acoustic backscatter using the generalized target description. *Proc. SPIE Int. Symp. Aerospace/Defense Sensing Contr*, 2765:190–200, 1996.

[25] Sherry McCahill et al. Report of the defense science board task force on unexploded ordnance. *Office of the Under Secretary of Defense For Acquisition, Techno logy, and Logistics*, nov 2003.

[26] Paul Marks. Robot subs seek a downed plane's secrets.(autonomous underwater vehicle). *New Scientist*, 209(2801):22, 2011. Web.

[27] Mark Otero Lisa Hazard William Middleton Peter Rogowski, Eric Terrill. Ocean outfall plume characterization using an autonomous underwater vehicle. *Water science and technology : a journal of the International Association on Water Pollution Research*, 67(4):925–933, 2013.

[28] Gwyn ; Challenor Peter Brito, Mario Paulo ; Griffiths. Risk analysis for autonomous underwater vehicle operations in extreme environments. *Risk Analysis*, 30(12):771–1788, 2010.

[29] P. Ryan Jackson. Integrated synoptic surveys using an autonomous underwater vehicle and manned boats. *Geological Survey*, 2013. Web.

[30] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. *In Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22, 2004.

[31] Andrew Zisserman Josef Sivic. Efficient visual search of videos cast as text retrieval. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(4):591–606, 2009.

[32] J. Weinberger, K. Q.; Dasgupta A.; Langford J.; Smola A.; Attenberg. Feature hashing for large scale multitask learning. *Proceedings of the 26th Annual International Conference on Machine Learning*, page 1113–1120, 2009.

[33] John D Daugman. Complete discrete 2-d gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7), 1988.

[34] William Visscher. Scattering of rayleigh surface waves from partly-closed surface-breaking cracks. Technical report, Los Alamos National Laboratory, 1983.

[35] N. S Altman. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185, 1992.

[36] D. Coomans; D.L. Massart. Alternative k-nearest neighbour rules in supervised pattern recognition : Part 1. k-nearest neighbour classification by using alternative voting rules. *Analytica Chimica Acta*, 136:15–27, 1982.

[37] Christopher Bishop. Pattern recognition and machine learning. *Springer*, 2007.

[38] Geoffrey McLachlan. *Discriminant Analysis and Statistical Pattern Recognition.* Wiley-Interscience, aug 2004. pp 15-25.

[39] E. M. BRAVERMAN AIZERMAN, M. A. and L. I. ROZONOER. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, 25:821–837, 1964.

[40] I. M.; Vapnik V. N. Boser, B. E.; Guyon. A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory*, page 144, 1992.

[41] F Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 1958.

[42] P. Werbos. Beyond regression: new tools for prediction and analysis in the behavioral sciences. 1974.

[43] Jorge Nocedal. *Numerical Optimization*. Springer, 2006.

[44] Yoshua Bengio and Yann LeCun. Scaling learning algorithms towards ai. *Large-Scale Kernel Machines*, 2007.

[45] Wade Trappe and Joseph D. Lakey. Composite wavelet transform and frames. *The Journal of the Acoustical Society of America*, 1995.

[46] D Fogel, I.; Sagi. Gabor filters as texture discriminator. *Biological Cybernetics*, 61 (2), 1989.

[47] IEEE De Yao Qiang Huang Mahmood R. Azimi-Sadjadi, Senior Member and Gerald J. Dobeck. Underwater target classification using wavelet packets and neural networks. *IEEE TRANSACTIONS ON NEURAL NETWORKS*, 11(3), 2000.

[48] J. Wilbur and S. G. Kargl. Application of wavelets to acoustic resonance- elastic targets surrounded by biologics. *Proc. ICASSP*, 1993.

[49] H. Szu B. Telfer and G. Dobeck. Adaptive wavelet classification of acoustic backscatter. *Proc. SPIE*, 2242:661–668, 1994.

[50] R. P. Hodges. *Active Target Strength, in Underwater Acoustics: Analysis, Design and Performance of Sonar*, volume ch9. John Wiley & Sons, Ltd, Chichester, UK, 2010.

[51] Tapas Kanungo, D.M. Mount, N.S. Netanyahu, C.D. Piatko, R. Silverman, and AY. Wu. An efficient k-means clustering algorithm: analysis and implementation.

Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(7):881–892, Jul 2002. ISSN 0162-8828. doi: 10.1109/TPAMI.2002.1017616.

[52] I Kirsteins A Sen Gupta, D Schupp. Dynamic target identification and classification based on resonance topography grouping. *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, 2014.