Handling Acoustic Scattering via Scattering Transforms: Robust classification of objects under geometric deformations from acoustic wavefields

Naoki Saito and David Weber

Department of Mathematics
University of California, Davis

Applied Mathematics Seminar
Department of Mathematics
Stanford University
November 29, 2017

Outline

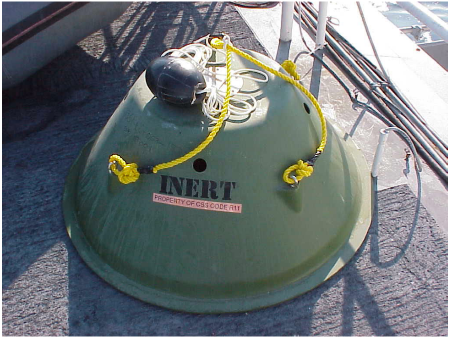
Acknowledgment

- ONR Grants:N00014-12-1-0177; N00014-16-1-2255
- NSF Grant: DMS-1418779
- Vincent Bodin (former intern from Ecole Polytechnique; now at Sinequa)
- Jim Bremer (UC Davis)
- Dan Cook, Frank Crosby, Gerry Dobeck, Carrie Dowdy, Julia Gazagnaire, Ross Goroshin, Richard Holtzapple, Quyen Huynh, Joe Lopes (NSWC-PCD, FL)
- Bradley Marchand (formerly Ph.D. student, UC Davis; now at NSWC-PCD, FL)
- Ian Sammis (formerly postdoc @ UC Davis; currently Google)
- `ScatNet` (S. Mallat and his group)
- `glmnet` (T. Hastie and his group)

Outline

Objectives

- Analyze acoustic wavefields scattered from underwater objects via the wideband FM (synthetic aperture) sonar system, in particular, how waveforms change relative to geometric transformation of those objects, e.g.: *translations, rotations, change of material inside objects*
- Classify underwater objects using those waveforms
- Apply the Scattering Transform (ST) and examine its effectiveness



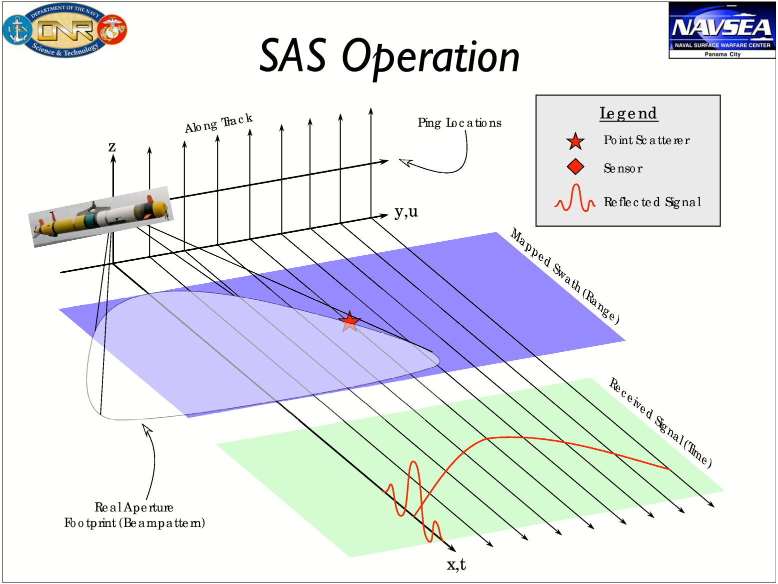
(a) Manta

(b) Pdm1

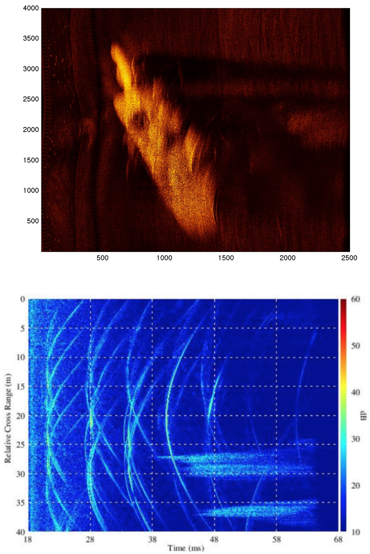
Outline

Synthetic Aperture Sonar System (Courtesy: D. Cook)

SAS Operation (Courtesy: D. Cook) . . .

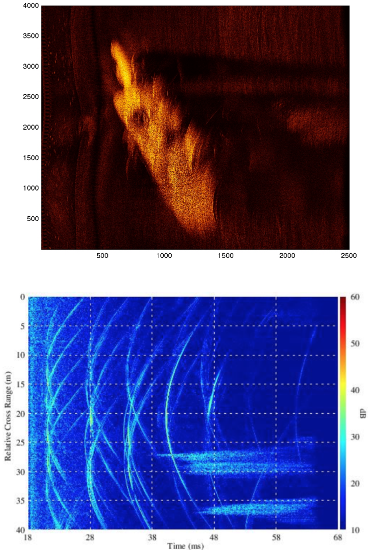


Waveforms $\Longrightarrow$ Images (Courtesy: R. Goroshin; S. Kargl)

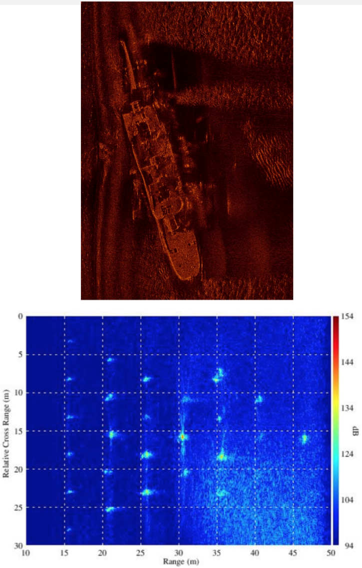


(a) Waveforms

Waveforms $\Longrightarrow$ Images (Courtesy: R. Goroshin; S. Kargl)

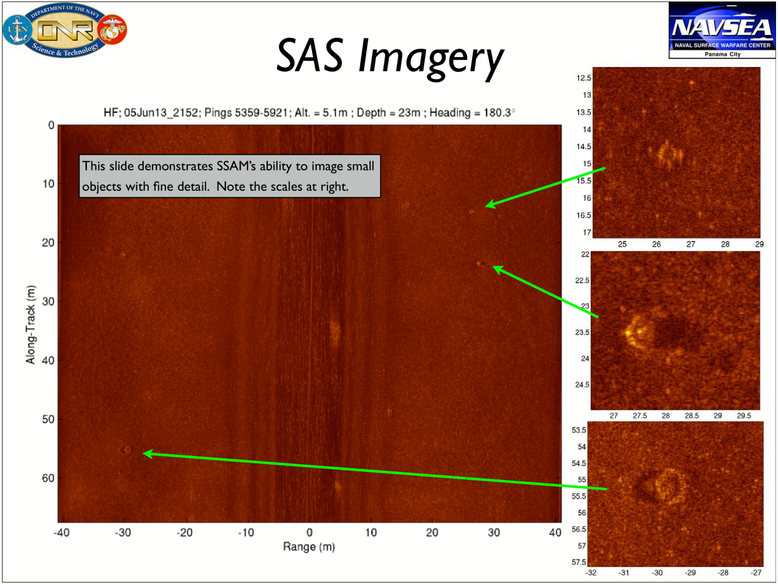


(a) Waveforms

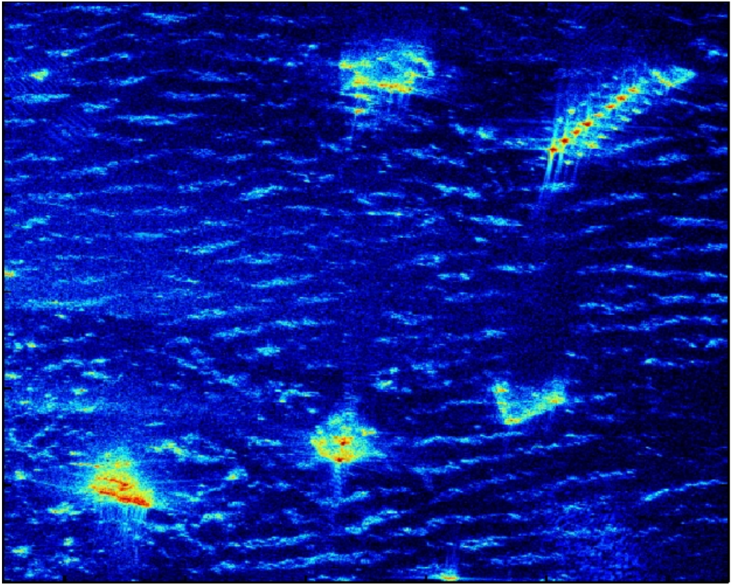


(b) Reconstructed Images

Several more real images (Courtesy: D. Cook)



Ambiguous Objects! (Courtesy: R. Goroshin)



Observations

- Shape information alone extracted from images (generated by the SAS imaging algorithm applied to sonar waveforms) is *ambiguous* for object classification
- Better to examine the *raw waveforms and the entire wavefield* scattered from an object for classification
- Do dolphins always reconstruct images from the returns of their clicks from objects in their brains??
- Dolphins do use some features of the waveforms returned from fishes to estimate their locations, fish species/sizes (via their *swim bladder shapes*; see, e.g., Yovel and Au, *PLoS ONE*, 2010)

Figure: Malene Thyssen, http://commons.wikimedia.org/wiki/User:Malene

Outline

Problem Setup

- Consider an object (a domain) $\Omega \in \mathbb{R}^2$ of arbitrary shape whose acoustic velocity is c_m, which is immersed in the surrounding material (i.e., water) of acoustic velocity $c_w = 1503(\mathrm{m/s})$.
- For a single frequency source $\mathrm{e}^{\mathrm{i}\omega t}$, this situation can be described by the *Helmholtz equations with transmission boundary conditions*:

$$\begin{aligned} \Delta u + k_1^2 u = 0 &\quad \text{in } \Omega \\ \Delta v + k_2^2 v = 0 &\quad \text{in } \Omega^c \\ u - v = g &\quad \text{on } \partial\Omega \\ \partial_\nu u - \partial_\nu v = \partial_\nu g &\quad \text{on } \partial\Omega \\ \sqrt{|x|}\left(\partial_{|x|} - \mathrm{i}k_2\right) v(x) \to 0 &\text{ as } |x| \to \infty \end{aligned}$$

  where $k_1 = \omega / c_m$ and $k_2 = \omega / c_w$.
- The above BVP provides a slightly more realistic model of acoustic scattering than simple Dirichlet or Neumann conditions.
- Our efficient numerical solver based on *boundary integral equations* allows singularities (e.g., corners, cusps) in the boundary curve $\partial\Omega$.

Responses to Multifrequency Transmitters

- A more realistic source waveform consists of multiple frequency sinusoids (e.g., chirps or truncated sinusoids).
- To do so, we employ the so-called *frequency-domain modeling*:
  1. Decompose a source signature $s(t_j),\ t_j = j\Delta t,\ j = 0, 1, \ldots, N-1$, $\Delta t := T/N$ into the (discrete) Fourier series $\sum_{n=0}^{N-1} \hat{s}_n \mathrm{e}^{\mathrm{i}\omega_n t_j}$, $\omega_n := 2\pi n/T$
  2. For each frequency ω_n , solve the system of the Helmholtz equations to obtain the wave $v_n^{\mathrm{obj}}(t_j) = a_n \mathrm{e}^{\mathrm{i}(\omega_n t_j + \theta_n)}$ scattered from the object Ω
  3. Synthesize the total response by summing all the individual responses with the appropriate coefficients: $v^{\mathrm{obj}}(t_j) := \sum_{n=0}^{N-1} \hat{s}_n v_n^{\mathrm{obj}}(t_j)$
- Note that in the transmission Helmholtz equations in the previous page, the total potential v in Ω^c is split into two pieces $v = v^{\mathrm{src}} + v^{\mathrm{obj}}$ where v^{src} is part of v purely due to the source in the absence of Ω and v^{obj} is part of v solely due to the presence of Ω.

Responses to Multifrequency Transmitters . . .

- In reality, one needs to be very careful about the *silent* trailing period of the input source signal in order to avoid the interference between the wave sent at $t = 0$ and that at $t = T$.
- This leads to some intricate choice of the silent period and zero padding in the DFT, and extraction of the output signal of period T, etc. But we will not discuss these details here.

A Fast Helmholtz Solver via BIEs

- The transmission Helmholtz BVP can be reformulated as a system of *boundary integral equations* (BIEs).
- Let u and v be represented as

$$\begin{aligned} u &= D_{k_1}\sigma + S_{k_1}\tau \\ v &= D_{k_2}\sigma + S_{k_2}\tau \end{aligned}$$

where

$$\begin{aligned} S_k f(\boldsymbol{x}) &:= \frac{\mathrm{i}}{4}\int_{\partial\Omega} H_0(k|\boldsymbol{x}-\boldsymbol{y}|)\, f(\boldsymbol{y})\,\mathrm{d}s(\boldsymbol{y}) \\ D_k f(\boldsymbol{x}) &:= \frac{\mathrm{i}}{4}\int_{\partial\Omega} k|\boldsymbol{x}-\boldsymbol{y}| H_1(k|\boldsymbol{x}-\boldsymbol{y}|)\, f(\boldsymbol{y})\frac{(\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{\nu}_y}{|\boldsymbol{x}-\boldsymbol{y}|^2}\,\mathrm{d}s(\boldsymbol{y}) \end{aligned}$$

where $H_\alpha(\cdot)$ are the *Hankel* functions of the first kind of order $\alpha = 0, 1$.

A Fast Helmholtz Solver via BIEs ...

- Then, the transmission Helmholtz BVP can be written as a system of the BIEs as:

$$\begin{bmatrix} D_{k_1} - D_{k_2} - I & S_{k_1} - S_{k_2} \\ D'_{k_1} - D'_{k_2} & S'_{k_1} - S'_{k_2} + I \end{bmatrix} \begin{bmatrix} \sigma \\ \tau \end{bmatrix} = \begin{bmatrix} g \\ \partial_\nu g \end{bmatrix}$$

where

$$\begin{aligned} S'_k f(\boldsymbol{x}) &:= \frac{\mathrm{i}}{4} \int_{\partial\Omega} k|\boldsymbol{x}-\boldsymbol{y}| H_1(k|\boldsymbol{x}-\boldsymbol{y}|) f(\boldsymbol{y}) \frac{(\boldsymbol{y}-\boldsymbol{x}) \cdot \boldsymbol{\nu}_x}{|\boldsymbol{x}-\boldsymbol{y}|^2} \,\mathrm{d}s(\boldsymbol{y}) \\ D'_k f(\boldsymbol{x}) &:= \frac{\mathrm{i}}{4} \int_{\partial\Omega} \left(\partial_{\nu_x} \partial_{\nu_y} H_0(k|\boldsymbol{x}-\boldsymbol{y}|)\right) f(\boldsymbol{y}) \,\mathrm{d}s(\boldsymbol{y}) \end{aligned}$$

- The advantages of solving Helmholtz BVPs using BIEs include:
  - dimension reduction $\Longleftarrow$ the integral is taken over $\partial\Omega$, not in Ω
  - well-conditioned systems (even for singular domains)
  - tamer singularities than in FDM/FEM
  - solvable via a *direct* (i.e., non-iterative) method

A Few *More* Words about Our Fast Solver

- Our solver is *direct* as opposed to iterative. This means that we form a *compressed* representation of the inverse of the matrix discretizing the relevant integral operator in order to solve the associated system of equations. In other words, we form a *scattering matrix* for the problem.
- Solvers of this type have a number of advantages; among them, resistance to ill-conditioning and the ability to solve for multiple right-hand sides efficiently.
- The ability to rapidly solve for multiple right-hand sides allows us to conduct our simulations efficiently.
- Computational cost of our 2D solver for a single frequency source is:
  - $O(N)$ to form a scattering matrix of size $N_{\text{out}} \times N_{\text{in}}$ where N is the number of discretization of the boundary curve $\partial\Omega$
  - $O(N_{\text{out}} \times N_{\text{in}})$ to build a solution for a given right-hand side.
  - $O(N_{\text{in}})$ to evaluate $v^{\text{out}}(x)$ for x far away from Ω.

Simulation Setup

- Three simple geometric shapes as Ω: triangle; shark fin; rectangle
- The range of c_m: 1, 500, 1000, . . . , 5000 (m/s)
- The geometry of measurements was similar to the real experiments conducted by NSWC-PCD: 481 transmitter/receiver location along a straight line (rail system)
- Each object was rotated $360°$ with $10°$ increment and the resulting wavefield was computed
- To speed up the wavefield synthesis, a database of a set of single frequency responses were created at the frequency range from 156.25 Hz to 50,000 Hz with 156.25 Hz increment (320 frequencies in total).

Simulation Results: Triangle

Triangle

Play/Pause | 1m/s | 2000m/s

Simulation Results: Shark Fin

Shark Fin

Play/Pause | 1m/s | 2000m/s

Outline

Effects of Object Deformation to Waveforms

- *Translation* of an object $\Longrightarrow$ *translation* and *amplitude decay* of the waveforms
- *Rotation* of an object $\Longrightarrow$ *translation*, *amplitude decay*, and *shifts in receiver indices* if measured under the *straight line* receiver arrays
- Changes of *object material* or *source pulse shape* $\equiv$ changes in $k = \omega / c_m \Longrightarrow$ *nonlinear changes in amplitude and phase* of the waveforms

Effects of Object Translation to Waveforms

- Translation of an object $\Longrightarrow$ *translation* and *amplitude decay* of the waveforms
- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be the coordinates of the receiver location and a point scatter, respectively. Then, the scattered wave arrives at this receiver at time $\frac{2|\boldsymbol{x}-\boldsymbol{y}|}{c_w}$.
- If the object location is translated to $\boldsymbol{y}+\Delta\boldsymbol{y}$, then the arrival time changes to $\frac{2|\boldsymbol{x}-\boldsymbol{y}-\Delta\boldsymbol{y}|}{c_w}$. Let θ be an angle between $\boldsymbol{y}-\boldsymbol{x}$ and $\Delta\boldsymbol{y}$.
- If $|\Delta\boldsymbol{y}| \ll 1$, the arrival time difference is $\approx \frac{2|\Delta\boldsymbol{y}|}{c_w}\cos\theta$.
- This arrival difference depends both on $\boldsymbol{x}$, $\Delta\boldsymbol{y}$, for a fixed $\boldsymbol{y}$.
- On the other hand, the amplitude decays approximately with the factor $\exp\left(-\alpha_w|\Delta\boldsymbol{y}|\cos\theta\right)$ where $\alpha_w = \frac{2\eta\omega^2}{3\rho c_w^3}$ is the attenuation coefficients of the water according to *Stoke's law*; it's frequency dependent! $\eta = 8.90 \times 10^{-4}$ Pa and $\rho = 1000\mathrm{kg/m}^3$ are the dynamic viscosity coefficient and the density of the water, respectively; but they also depends on temperature and salinity of the water.

Effects of Object Rotation to Waveforms

- Rotation of an object $\Longrightarrow$ may result in drastic changes in waveforms if the object has *sigularities* in $\partial\Omega$
- The *geometry of a measurement system* becomes quite important!
- If an array of receivers surround the object completely, a rotation of the object simply amounts to a circular shift of the receiver indices.
- An expected geometry of a receiver array is, however, *a straight line*.
- Hence, a rotation of the object amounts to the changes of translation, amplitude decay, and shifts in receiver indices.

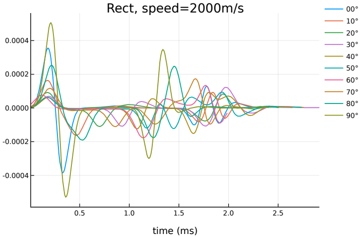


Figure: Signals scattered from a rectangle with varying view angles.

Effects of Change of Material/Source Pulse to Waveforms

- Change of material inside an object and/or change of the source pulse shape $\equiv$ change in $k = \omega / c_m$
  $\Longleftrightarrow$ *nonlinear* changes in amplitude and phase of the waveforms

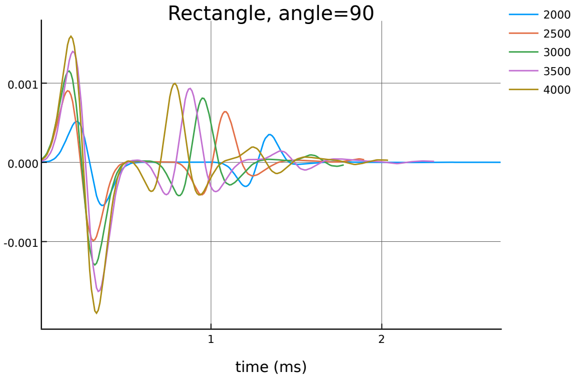


Figure: Signals scattered from rectangles of varying speed of sound.

Comments on Invariant Features: Amari & Otsu

- Amari (late 60's) and Otsu (mid 70's) already worked on *invariant feature extractors* based on the "Invariant Theory."
- They first considered the *linear* feature extractors (LFEs) as linear functionals in a Hilbert space: $\rho[f] := \langle f, \rho \rangle = \int f(x)\overline{\rho(x)}\,\mathrm{d}x$.
- For a given pattern deformation T_λ (i.e., a composite of continuous transformation groups in $\mathbb{R}^1$ or $\mathbb{R}^2$), an LFE is defined as $\rho[T_\lambda f] = \eta(\lambda)\rho[f]$. If $\eta(\lambda) \equiv 1$, then the LFE ρ is called *absolute*; otherwise called *relative*.
- If T_λ = an additive group (e.g., translations), then the invariant LFEs must be of the *Fourier-Laplace* transform type: $\rho[f] = \int f(x) c_1 \mathrm{e}^{\gamma_1 x}\,\mathrm{d}x$.
- If T_λ = a multiplicative group (e.g., dilations), then the invariant LFEs must be of the *Mellin* transform type (including *moment* feature extractors): $\rho[f] = \int f(x) c_2 x^{\gamma_2 - 1}\,\mathrm{d}x$.
- They showed that the absolute LFEs are so limited ($\gamma_1 = 0$, $\gamma_2 = 0$) due to *Haar's theorem* that they are basically meaningless.
- Hence, they suggested that in order to find more useful invariant feature extractors, one must seek *nonlinear* operators.

Scattering Transform

- A *scattering transform* (ST; proposed by S. Mallat and further developed by him and his group) can maintain *Lipschitz continuity* relative to a small deformation applied to an input signal! In other words, the ST representation is *stable* relative to such small deformations.
- An example of a small deformation close to a translation is: $T_\tau f(x) := f(x - \tau(x))$ where $\tau(\cdot) \in C^2\left(\mathbb{R}^d\right)$ is a displacement field.
- An operator $\Psi : L^2\left(\mathbb{R}^d\right) \to \mathscr{H}$ is said to be *translation invariant* if $\Psi[T_c f] = \Psi[f]$, for every constant vector $c \in \mathbb{R}^d$.
- A translation invariant operator Ψ is *Lipschitz continuous* relative to T_τ if $\forall \Omega \in \mathbb{R}^d$: compact, $\exists C > 0$ such that $\forall f \in L^2(\mathbb{R}^d)$, $\operatorname{supp} f \subset \Omega$, $\forall \tau \in C^2\left(\mathbb{R}^d\right)$,

$$\left\| \Psi[f] - \Psi[T_\tau f] \right\|_{\mathscr{H}} \le C \left\| f \right\| (\|\nabla \tau\|_\infty + \|H\tau\|_\infty),$$

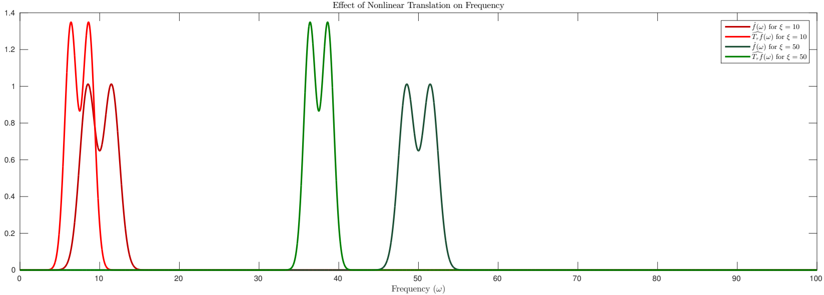
where $H\tau$ is the Hessian tensor of τ.

Why not just use the Fourier Transform?

The magnitude of the Fourier transform is translation invariant, but the Lipschitz-continuity is not preserved for deformations:
Let $\tau(t) = st$, with $|s| < 1$, and $f(t) = \mathrm{e}^{\mathrm{i}\xi t}\theta(t)$, where θ is even and $O\left(\mathrm{e}^{-x^2}\right)$ then $T_\tau f(t) = f((1-s)t)$ translates the central frequency ξ to $(1-s)\xi$

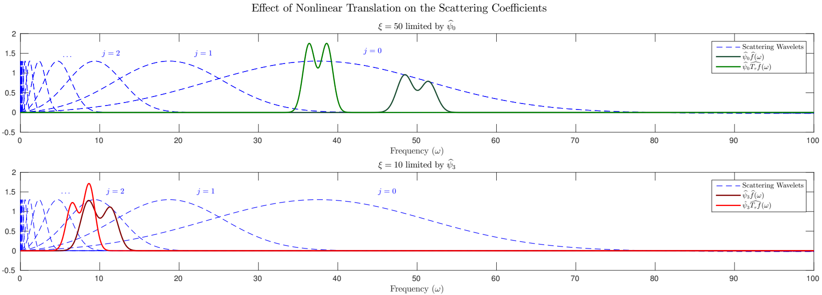
$$\left\| \widehat{T_\tau f} - \widehat{f} \right\| \sim |s||\xi| \, \|\theta\| = |\xi| \, \left\| f \right\| \, \|\nabla\tau\|_\infty$$

No universal bound for arbitrary ξ!



How about Wavelet Transforms?

In the Fourier domain, a wavelet transform $\psi_j \star f$ bandpasses the signal over windows whose bandwidth decreases exponentially with j, so that both f and $T_\tau f$ are captured within the same wavelet, regardless of ξ.



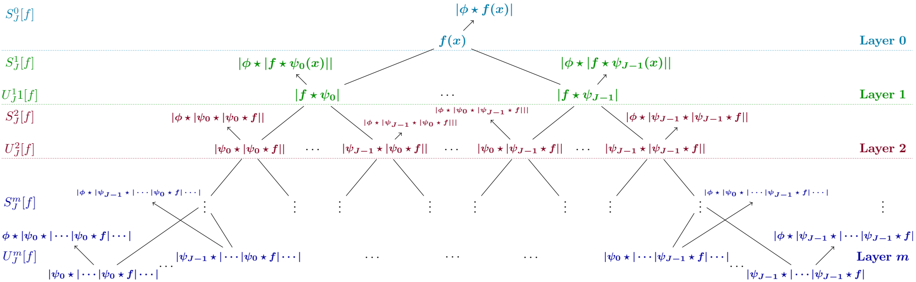
- Discrete Orthonormal Wavelet Transform ≡ not invariant at all
- Stationary Wavelet Transform ≡ relatively invariant but not absolutely invariant
- *Averaging* after Stationary Wavelet Transform improves invariance

Scattering Transform

A single propagating layer $U_J^m[f]$ of the scattering transform is a vector consisting of alternating *convolution with wavelets* $\widehat{\psi}_j(\omega) = \hat{\psi}\left(2^{j/Q_m}\omega\right)$ and a *modulus* $|\cdot|$ with the scale in each layer varying from the finest scale of 0 up to $J-1$:

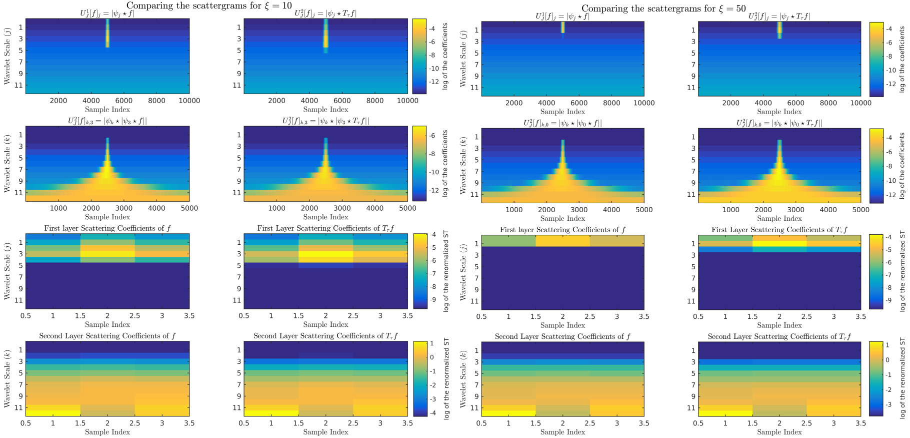
$$U_J^1[f] := \left(|\psi_0 \star f|, \ldots, |\psi_{J-1} \star f|\right)$$

$$U_J^2[f] := \left(|\psi_0 \star |\psi_0 \star f||, \ldots, |\psi_{J-1} \star |\psi_0 \star f||, \ldots,\right.$$
$$\left.\ldots, |\psi_0 \star |\psi_{J-1} \star f||, \ldots, |\psi_{J-1} \star |\psi_{J-1} \star f||\right)$$



The output $S_J^m[f]$ is taken by *averaging* every term of $U_J^m[f]$ with *the father wavelet* ϕ corresponding to ψ, followed by *subsampling*.

Scattering Transform comparison of f and $T_\tau f$



Useful Properties

Theorem (Energy conservation, Mallat (2012))

For all $f \in L^2(\mathbb{R}^d)$, if (ψ, ϕ) are admissible, then

$$\|f\|_2 = \|S_J[f]\|_2 \quad \text{where} \quad S_J[f] := \left(S_J^0[f], S_J^1[f], \ldots, S_J^m[f], \ldots\right),$$
$$\|S_J[f]\|_2^2 := \sum_{m=0}^{\infty} \|S_J^m[f]\|_2^2$$

In addition to preserving the energy, as the scale goes to infinite resolution, the scattering transform is translation invariant

Theorem (Limit Translation Invariance, Mallat (2012))

For all $f \in L^2(\mathbb{R}^d)$ and $c \in \mathbb{R}^d$, if (ψ, ϕ) are admissible, then

$$\lim_{J \to -\infty} \|S_J[f] - S_J[T_c f]\|_2 = 0$$

Useful Properties

Theorem (Energy conservation, Mallat (2012))

For all $f \in L^2(\mathbb{R}^d)$, if (ψ, ϕ) are admissible, then

$$\|f\|_2 = \|S_J[f]\|_2 \quad \textit{where} \quad S_J[f] := \left(S_J^0[f], S_J^1[f], \ldots, S_J^m[f], \ldots\right),$$
$$\|S_J[f]\|_2^2 := \sum_{m=0}^{\infty} \|S_J^m[f]\|_2^2$$

In addition to preserving the energy, as the scale goes to infinite resolution, the scattering transform is translation invariant

Theorem (Limit Translation Invariance, Mallat (2012))

For all $f \in L^2\left(\mathbb{R}^d\right)$ and $c \in \mathbb{R}^d$, if (ψ, ϕ) are admissible, then

$$\lim_{J \to -\infty} \|S_J[f] - S_J[T_c f]\|_2 = 0$$

Theorem (Lipschitz Continuity, Mallat (2012))

For all compactly supported $f \in L^2(\mathbb{R}^d)$ satisfying $\left\|\sum_m U_J^m f\right\|_1 < \infty$ and $\tau \in C^2(\mathbb{R}^d)$ where $\|\nabla\tau\|_\infty \le \frac{1}{2}$ and $\|\tau\|_\infty / \|\nabla\tau\|_\infty \le 2^J$, there is a C such that:

$$\left\|S_J[f] - S_J[T_\tau f]\right\|_2 \le C \left\|\sum_m U_J^m f\right\|_1 (\|\nabla\tau\|_\infty + \|H\tau\|_\infty)$$

A more recent result: for general frames, and not just admissible wavelets, increasing the depth m improves translation invariance property:

Theorem (Depth Translation Invariance, Wiatowski–Bölcskei (2015))

If R_n is the subsampling rate at layer n, as long as the wavelets have frame bounds B_n satisfying $\max\{B_n, B_n R_n^d\} \le 1$, the features at depth m satisfy:

$$S_J^m[T_c f] = T_{\frac{c}{R_1 \cdots R_m}} S_J^m[f]$$

Theorem (Lipschitz Continuity, Mallat (2012))

For all compactly supported $f \in L^2(\mathbb{R}^d)$ *satisfying* $\left\|\sum_m U_J^m f\right\|_1 < \infty$ *and* $\tau \in C^2(\mathbb{R}^d)$ *where* $\|\nabla\tau\|_\infty \le \frac{1}{2}$ *and* $\|\tau\|_\infty / \|\nabla\tau\|_\infty \le 2^J$*, there is a* C *such that:*

$$\left\|S_J[f] - S_J[T_\tau f]\right\|_2 \le C \left\|\sum_m U_J^m f\right\|_1 (\|\nabla\tau\|_\infty + \|H\tau\|_\infty)$$

A more recent result: for general frames, and not just admissible wavelets, increasing the depth m improves translation invariance property:

Theorem (Depth Translation Invariance, Wiatowski–Bölcskei (2015))

If R_n *is the subsampling rate at layer* n*, as long as the wavelets have frame bounds* B_n *satisfying* $\max\{B_n, B_n R_n^d\} \le 1$*, the features at depth* m *satisfy:*

$$S_J^m[T_c f] = T_{\frac{c}{R_1 \cdots R_m}} S_J^m[f]$$

Implementation Details

- As m increases, the remaining energy is concentrated at coarser scales, so only those S_J^m with increasing scales at deeper layers are kept for computational reasons (e.g., $|\phi \star |\psi_4 \star |\psi_1 \star f|||$ is kept, while $|\phi \star |\psi_1 \star |\psi_4 \star f|||$ is not).
- This project so far has focused on using *Morlet Wavelets* (*almost analytic*):

$$\psi(t) = c_\xi \mathrm{e}^{-t^2/2}\left(\mathrm{e}^{\mathrm{i}\xi t} - \kappa_\xi\right) \quad \Longleftrightarrow \quad \widehat{\psi}(\omega) = c_\xi\left(\mathrm{e}^{-(\omega-\xi)^2/2} - \kappa_\xi \mathrm{e}^{-\omega^2/2}\right),$$

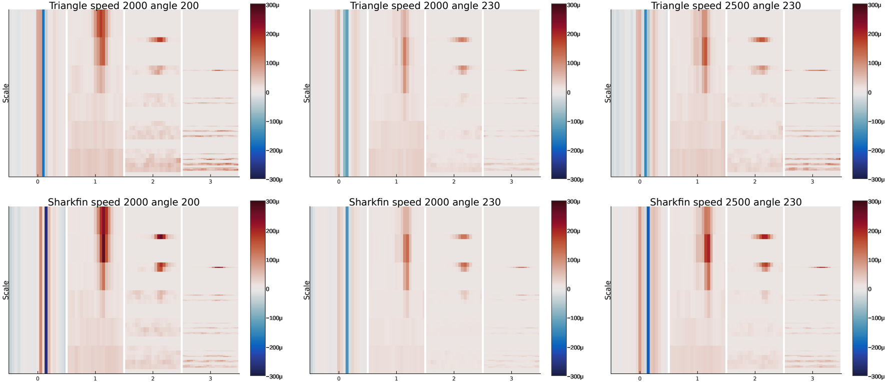
where κ_ξ is a constant for ψ to be admissible and c_ξ is a normalization constant.

Outline

Synthetic Data Setup I: Target Shape Discrimination

- Classification of Triangle object vs Shark Fin object via waveforms
- Each object is rotated by 10° increment to cover 0° to 350°
- Speed of sound in both objects is the same, $2000\mathrm{m/s}$.
- Hence, this classification is about the object *shape* through the scattered waveforms *regardless of rotations*
- Each signal is normalized so the maximum amplitude is 1
- The white Gaussian noise $\mathcal{N}(0, 10^{-5})$ is added to the waveforms, i.e., the average SNR is about 12dB.
- For each angle for each object, 481 waveforms with 641 time samples are generated.
- Multiclass logistic regression with Lasso (via `glmnet`) is used as a feature extractor and a classifier.
- Perform twofold cross validation 10 times, i.e., repeats the classification 10 times by randomly splitting the whole dataset into training and test sets with 50/50.

Examples of Scattering Transform Coefficients



The results from a depth 3 scattering transform with ($Q_1 = Q_2 = Q_3 = 1$, $0 \le m \le 3$) on various materials, shapes, and angles.

Target Shape Discrimination: Results

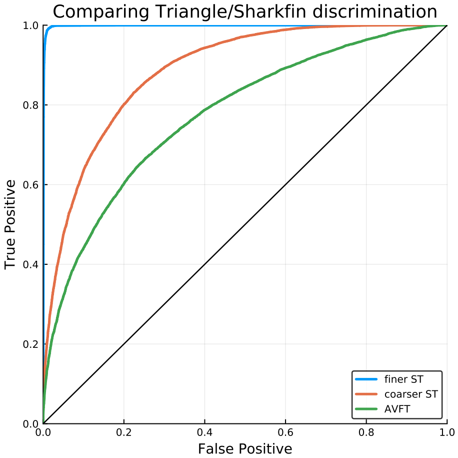


Figure: The ROC curve for discriminating a sharkfin from a triangle. Finer ST: $(Q_1, Q_2, Q_3) = (8, 4, 4)$; Coarser ST: $(Q_1, Q_2, Q_3) = (1, 1, 1)$. The AUC values of Finer ST, Coarser ST, AVFT are: 0.998, 0.886, and 0.775, respectively.

Target Shape Discrimination: ST Coefficients

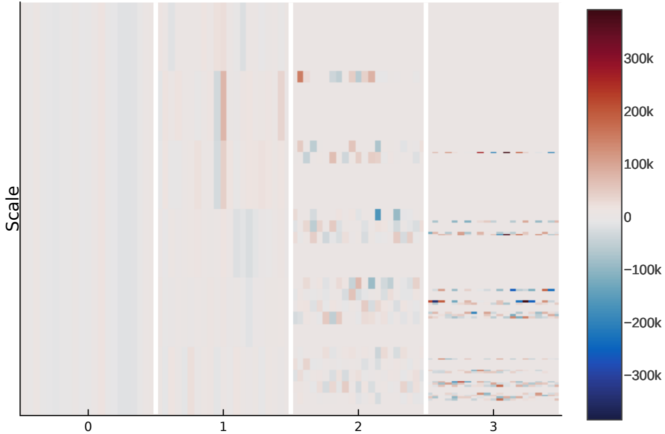


Figure: The glmnet coefficients selected in one run. Red coefficients correspond to the triangle class, while blue to the sharkfin

Synthetic Data Setup II: Target Material Discrimination

- Classification of two Triangle objects each of which has a different speed of sound.
- Each object is rotated by 10° increment to cover 0° to 350°
- Speed of sound in these two objects are set to $2000\mathrm{m/s}$ and $2500\mathrm{m/s}$.
- Hence, this classification is about the object *material* through the scattered waveforms *regardless of rotations*.
- The other classification setup is the same as the shape discrimination case.

Target Material Discrimination: Results

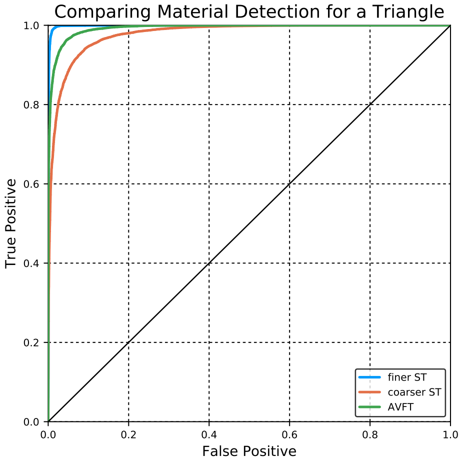


Figure: The ROC curve for detecting the material difference in a triangle, for speeds of sound $c_1 = 2000\mathrm{m/s}$ and $c_1 = 2500\mathrm{m/s}$. The AUC values of Finer ST, Coarser ST, AVFT are: 0.99994, 0.97778, and 0.992837, respectively.

Target Material Discrimination: ST Coefficients

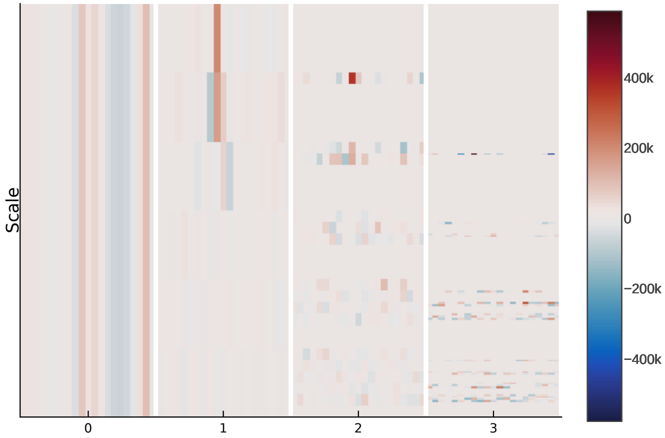


Figure: The `glmnet` coefficients selected in one run. Red coefficients correspond to the speed 2000 class, while blue to the speed 2500

Outline

About the BAYEX14 Dataset

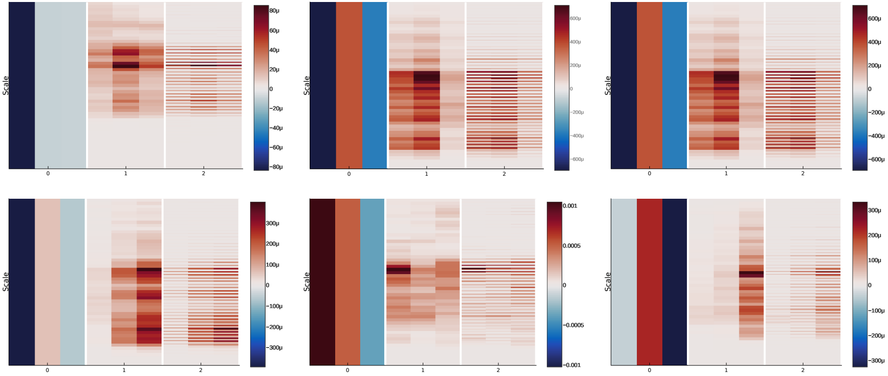
- The Bay Experiment 2014 (BAYEX14) was conducted from 29 April 2014 through 1 June 2014 at St. Andrews Bay (Panama City, FL).
- 22 objects were placed on the ocean floor (about 8m deep).
- Each object was placed at different grid cell of the ocean floor.

Figure: Some of the targets used in BAYEX14

About the BAYEX14 Dataset . . .

- Each SAS measurement ran along the sonar rail (42m long) with spatial sampling rate 2.5cm.
- Each object was rotated through a set of 9 angles with respect to the rail (-80° to 80° with 20° increment) by divers!
- We received the waveforms scattered from 14 objects.
- For each angle for each object, about 1600 waveforms with 1400 time samples were recorded.
- Due to some amplitude bursts of some waveforms, we normalized each waveform so that it has the unit ℓ^2 norm.
- We have split the data into two classes: UXOs (or their replicas) and non-UXOs (including natural rock, water-filled drum and tank).
- Between classes, there are no similar shapes, but there are two with the same material (aluminum UXO replica vs aluminum non-UXO pipe).
- Performed 10-fold cross validation, i.e., repeats the classification 10 times by randomly splitting the whole dataset into training and test sets with 50/50.

Examples of Scattering Transform Coefficients



The results from a depth 2 scattering transform with $(Q_1 = 8; Q_2 = 1, 0 \leq m \leq 2)$. Top row: UXOs; Bottom row: non-UXOs.

The BAYEX14 Dataset: Results

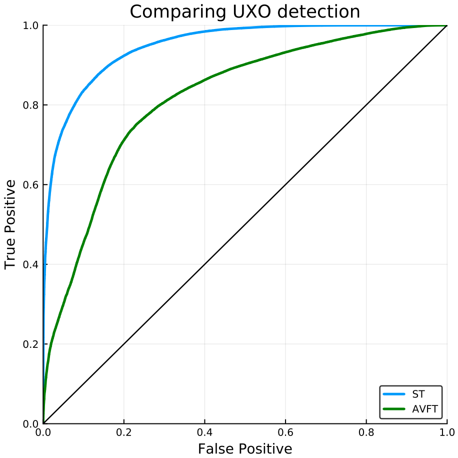


Figure: The ROC curve for detecting UXOs. ST: $(Q_1, Q_2) = (8, 1)$; AUC=0.9487; AVFT: AUC=0.8186.

UXO vs non-UXO Discrimination: ST Coefficients

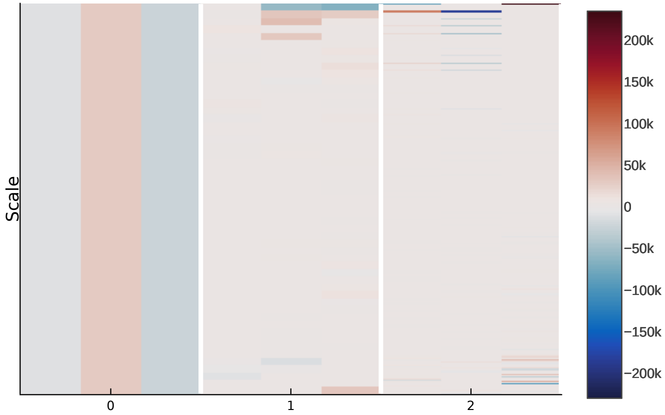


Figure: The `glmnet` coefficients selected in one run. Red coefficients correspond to the UXO class, while blue to the non-UXO

Outline

Summary

- Our preliminary results indicated *the robustness of the ST representations* under the SAS setup, both synthetic and real.
- As predicted, the ST coefficients at *deeper layers* turned out to be more useful for classification of signals with various deformations thanks to their quasi-invariance to those deformations.
- The ST-based representations performed better than the the modulus of the Fourier transforms (AVFTs), confirming that the deformations in our signals are not simple constant shifts of template/prototype signals.

Future Plan

- Examine how to present the selected features (ST coefficients) in an *intuitive* manner (perhaps, via reconstructing the waveforms from the selected ST coefficients, which requires estimating the *phase* information lost due to the modulus operations).
- Investigate if waveform deformations due to target rotations, material changes, and geometry of measurements can be formulated via nonlinear displacement field $\tau(\boldsymbol{x})$; though it may not be in $C^2(\mathbb{R}^d)$, and the Lipschitz continuity may not be preserved.
- Use the whole 2D wavefields as training signals, i.e., view each *2D wavefield* as a special *image* and use 2D Scattering Transform based on *Shearlets*.
- Examine the simulation results for manifold learning, e.g., how to learn rotation angles of an object or medium velocity changes purely from the scattered wavefields.

Shearlets (Guo, Kutyniok, Labate, ...)

A 2D frame based around *dilation* and *shearing* of a mother "wavelet":

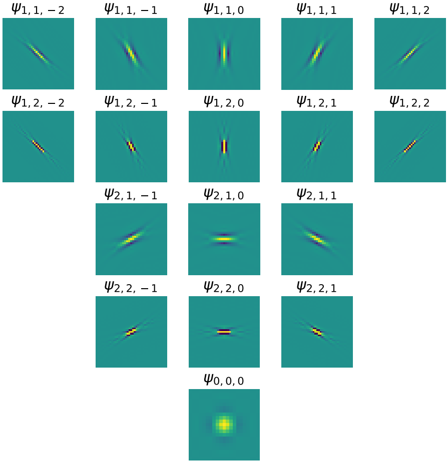


Figure: The real part of a shearlet system with $J = 2$, where $\psi_{i,j,k}$ is in cone i, with scale j and shearing k, and ϕ is the averaging function.

"Shattering" Transform

- A *shattering* (or shearlet scattering) transform is a generalized scattering transform using shearlets as the frames.
- For the nonlinearity operator, we use $\sigma = |\cdot|$ here, but the other possibilities, e.g., the complex extension of the Rectifier Linear Unit, i.e., $\mathrm{ReLU}(z) := \mathrm{ReLU}(\mathrm{Re}(z)) + \mathrm{ReLU}(\mathrm{Im}(z))\mathrm{i}$, where $\mathrm{ReLU}(x) := \max\{0, |x|\}$ for $x \in \mathbb{R}$.
- For the theoretical results to apply, we will eventually need to show that

$$\max\left\{B, \frac{B\gamma^2}{R^2}\right\} \leq 1 \tag{1}$$

  where γ is the Lipschitz constant for σ, which is at worst 2 for the above examples, R is the subsampling rate, and B is the frame bound for the system of shearlets. The last is the only tricky one.
- To improve computation time and to increase the parallelism with CNNs (convolutional neural networks), we could only use the output from the final layer, but that might be suboptimal ...

Outline

References

- Check http://www.math.ucdavis.edu/~saito/publications/ and http://www.math.ucdavis.edu/~bremer/, e.g.,
- J. Bremer, V. Rokhlin, & I. Sammis: "Universal quadratures for boundary integral equations on two-dimensional domains with corners," *J. Comput. Phys.*, vol. 229, pp. 8259–8280, 2010.
- J. Bremer: "A fast direct solver for the integral equations of scattering theory on planar curves with corners," *J. Comput. Phys.*, vol. 231, pp. 1879–1899, 2012.
- B. Marchand & N. Saito: "Earth Mover's Distance based local discriminant basis," in *Multiscale Signal Analysis and Modeling* (X. Shen & A. I. Zayed, eds.), Springer, Chap. 12, pp. 275–294, 2013.
- N. Saito & D. Weber: " Underwater object classification using scattering transform of sonar signals," in *Wavelets and Sparsity XVII* (Y. M. Lu, D. Van De Ville, and M. Papadakis, eds.), Proceedings of SPIE, vol. 10394, Paper #10394117.

Thank you very much for your attention!