Underwater Object Classification Using Scattering Transform of Sonar Signals

David Weber and Naoki Saito

Department of Mathematics
University of California, Davis

August 6, 2017

Signal Invariants

How to construct a feature extractor that is invariant to task-irrelevant deformations in the data? Some examples:

Translation	$T_c[f] = f(x-c)$
Modulation	$M_\omega[f] = e^{i\omega t} f(x)$
Scaling	$\mathscr{S}_a[f] = f\left(x/a\right)$
Amplitude	$A_a[f] = a f(x)$

Signal Invariants

How to construct a feature extractor that is invariant to task-irrelevant deformations in the data? Some examples:

Translation	$T_c[f] = f(x-c)$
Modulation	$M_\omega[f] = e^{i\omega t} f(x)$
Scaling	$\mathscr{S}_a[f] = f\left(x/a\right)$
Amplitude	$A_a[f] = af(x)$

[Amari, 1968] and [Otsu, 1973] demonstrated that only trivial linear features are absolutely invariant to even just translation, so they use *relative invariance* of feature extractor ρ:

$$\rho[T_c f] = \eta(c)\rho[f]$$

Signal Invariants

How to construct a feature extractor that is invariant to task-irrelevant deformations in the data? Some examples:

Translation	$T_c[f] = f(x-c)$
Modulation	$M_\omega[f] = \mathrm{e}^{\mathrm{i}\omega t} f(x)$
Scaling	$\mathscr{S}_a[f] = f\left(x/a\right)$
Amplitude	$A_a[f] = a f(x)$

[Amari, 1968] and [Otsu, 1973] demonstrated that only trivial linear features are absolutely invariant to even just translation, so they use *relative invariance* of feature extractor ρ:

$$\rho[T_c f] = \eta(c)\rho[f]$$

They establish that the only linear feature extractors $\rho[f] = \langle f, \rho\rangle$ that are relatively invariant w.r.t. both amplitude and translation deformations are Fourier-Laplace type, i.e. for some $z \in \mathbb{C}^d$

$$\int_{\mathbb{R}^d} f(x) c_1 \mathrm{e}^{z\cdot x} \, \mathrm{d}x$$

Generalized Scattering Transform

A single propagating layer $u[q_i]f$ of the generalized scattering transform is a vector consisting of semi-discrete shift invariant frame transforms $\psi_{\lambda_i^{(m)}} \star f$ indexed by $\lambda_i^{(m)} \in \Lambda_m$, a pointwise nonlinearity σ_m with Lipschitz constant γ_m, and a subsampling factor $r_m \geq 1$.

Generalized Scattering Transform

A single propagating layer $u[q_i]f$ of the generalized scattering transform is a vector consisting of semi-discrete shift invariant frame transforms $\psi_{\lambda_i^{(m)}} \star f$ indexed by $\lambda_i^{(m)} \in \Lambda_m$, a pointwise nonlinearity σ_m with Lipschitz constant γ_m, and a subsampling factor $r_m \geq 1$.Putting together k indices in a path $q = (\lambda_{i_m}^{(m)}, \ldots, \lambda_{i_1}^{(1)})$ gives the propagating value at layer m:

$$u[\lambda_i]f := \sigma\big(\psi_{\lambda_i} \star f\big)(r_1 \cdot) \qquad m = 1$$

$$u[\lambda_i^{(2)}, \lambda_j^{(1)}]f := \sigma\Big(\psi_{\lambda_i^{(2)}} \star \sigma\big(\psi_{\lambda_j^{(1)}} \star f\big)(r_1 \cdot)\Big)(r_2 \cdot) \quad m = 2$$

And so on, until the desired depth.

Generalized Scattering Transform

A single propagating layer $u[q_i]f$ of the generalized scattering transform is a vector consisting of semi-discrete shift invariant frame transforms $\psi_{\lambda_i^{(m)}} \star f$ indexed by $\lambda_i^{(m)} \in \Lambda_m$, a pointwise nonlinearity σ_m with Lipschitz constant γ_m, and a subsampling factor $r_m \geq 1$.Putting together k indices in a path $q = (\lambda_{i_m}^{(m)}, \ldots, \lambda_{i_1}^{(1)})$ gives the propagating value at layer m:

$$u[\lambda_i]f := \sigma\big(\psi_{\lambda_i} \star f\big)(r_1 \cdot) \qquad m = 1$$

$$u[\lambda_i^{(2)}, \lambda_j^{(1)}]f := \sigma\Big(\psi_{\lambda_i^{(2)}} \star \sigma\big(\psi_{\lambda_j^{(1)}} \star f\big)(r_1 \cdot)\Big)(r_2 \cdot) \quad m = 2$$

And so on, until the desired depth. The output $s^m[f]$ is taken by averaging $u[f]$ for every path q of depth m with one atom ϕ_m, and then subsampling:

$$s_m[\lambda_{i_m}^{(m)}, \ldots, \lambda_{i_1}^{(1)}]f := \phi_m \star u[\lambda_{i_m}^{(m)}, \ldots \lambda_{i_1}^{(1)}]f$$

Generalized Scattering Transform

A single propagating layer $u[q_i]f$ of the generalized scattering transform is a vector consisting of semi-discrete shift invariant frame transforms $\psi_{\lambda_i^{(m)}} \star f$ indexed by $\lambda_i^{(m)} \in \Lambda_m$, a pointwise nonlinearity σ_m with Lipschitz constant γ_m, and a subsampling factor $r_m \geq 1$.Putting together k indices in a path $q = (\lambda_{i_m}^{(m)}, \ldots, \lambda_{i_1}^{(1)})$ gives the propagating value at layer m:

$$u[\lambda_i]f := \sigma\big(\psi_{\lambda_i} \star f\big)(r_1 \cdot) \qquad m = 1$$

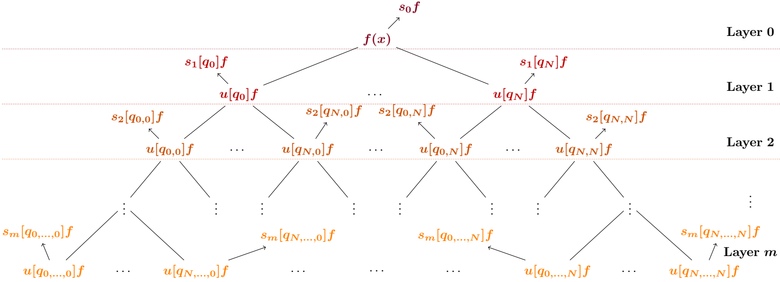
$$u[\lambda_i^{(2)}, \lambda_j^{(1)}]f := \sigma\Big(\psi_{\lambda_i^{(2)}} \star \sigma\big(\psi_{\lambda_j^{(1)}} \star f\big)(r_1 \cdot)\Big)(r_2 \cdot) \quad m = 2$$

And so on, until the desired depth. The output $s^m[f]$ is taken by averaging $u[f]$ for every path q of depth m with one atom ϕ_m, and then subsampling:

$$s_m[\lambda_{i_m}^{(m)}, \ldots, \lambda_{i_1}^{(1)}]f := \phi_m \star u[\lambda_{i_m}^{(m)}, \ldots \lambda_{i_1}^{(1)}]f$$

If we write s without an index, this is the collection of outputs at all layers up to some desired M: $0, 1, \ldots, M$.

Generalized Scattering Transform



The original scattering transform specifies that $\sigma_m = |\cdot|$, indexes by $\Lambda_m = \{a^{j/Q_m} h\}_{j>-J_m, h\in H_m}$ for some rotation h in the discrete rotation group H_m, subsamples only the output, and has strong conditions on the parent wavelets ψ and ϕ. Q_m is the quality factor, which can vary by layer.

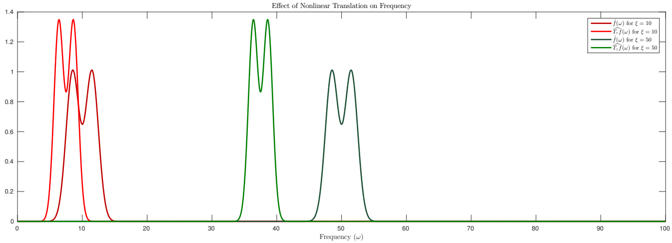
The Fourier transform is relatively invariant to translation, but it is *not* relatively invariant to non-constant translation $F_\tau[f] = f(t - \tau(t))$:

Heuristics
Problems of the Fourier Transform

The Fourier transform is relatively invariant to translation, but it is *not* relatively invariant to non-constant translation $F_\tau[f] = f(t - \tau(t))$:
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 $T_\tau[f](t) = f\big((1-s)t\big)$ translates the central frequency ξ to $(1-s)\xi$

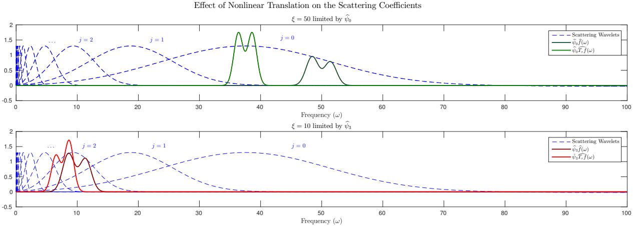
$$\|\widehat{T_\tau f} - \widehat{f}\| \sim |s||\xi|\|\theta\| = |\xi|\|f\|\|\nabla\tau\|_\infty$$

No universal bound for arbitrary ξ!



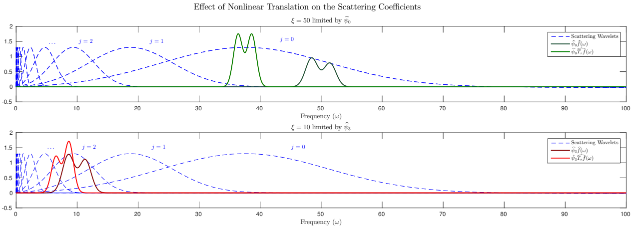
Heuristics
Wavelet Transform & T_τ

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



Heuristics
Wavelet Transform & T_τ

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



A Wavelet transform isn't translation invariant, but it does commute with the translation operator, i.e., if $W[j]f(n) = f \star \psi_{j,n}$, then

$$W[j]T_c f(n) = T_c W[j] f(n)$$

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

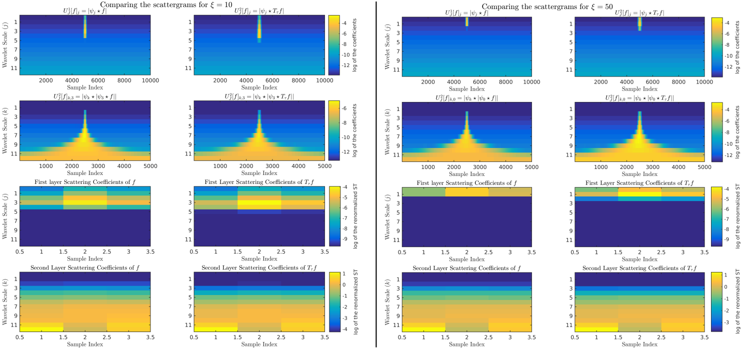


Figure: Output of the Scattering Transform in the first layers for $\xi = 10$ on the left and $\xi = 50$ on the right. Upper 2 rows are u for the first and second layer, while the bottom 2 rows are the actual outputs for the first and second layers

The first results on Scattering transforms were from [Mallat, 2012] and his group. A more recent generalization for "weakly admissible" frames, and not just wavelets, that increasing the depth m increases translation invariance:

Theorem (Depth translation invariance, [Wiatowski and Bölcskei, 2015])

As long as the frames have upper frame bounds b_m satisfying $\max\{b_m, \gamma_m b_m / r_m^d\} \leq 1$, the features at depth m satisfy:

$$S^m[T_c f] = T_{\frac{c}{r_1 \cdots r_{m-1}}} S^m[f]$$

Previous Theory
Non-uniform Translation and Modulation

In addition to this quasi-translation invariance, this generalized scattering transform is stable under space and frequency modulations:

$$F_{\tau,\omega}[f](x) = \mathrm{e}^{\mathrm{i}\omega(x)} f\big(x - \tau(x)\big)$$

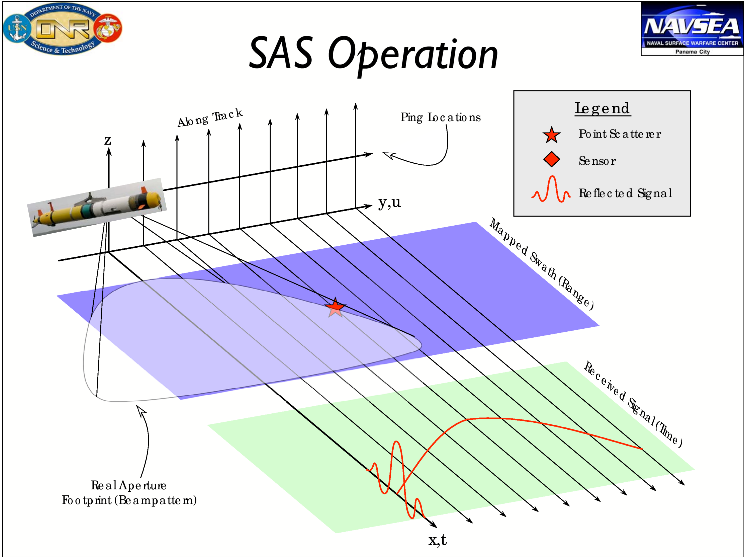
Theorem (Stability, [Wiatowski and Bölcskei, 2015])

If f is a band limited function, ω and τ are continuous, τ is once differentiable and $\|\nabla\tau\|_\infty \leq \frac{1}{2d}$, there is a C independent of S so that

$$\Big\| S[f] - S\big[F_{\tau,\omega}[f]\big] \Big\|_2 \leq C\|f\|_2\big(R\|\tau\|_\infty + \|\omega\|_\infty\big)$$

where the norm on S is just $\|\cdot\|_2$ on each output element

Sonar Scattering



Sonar Scattering

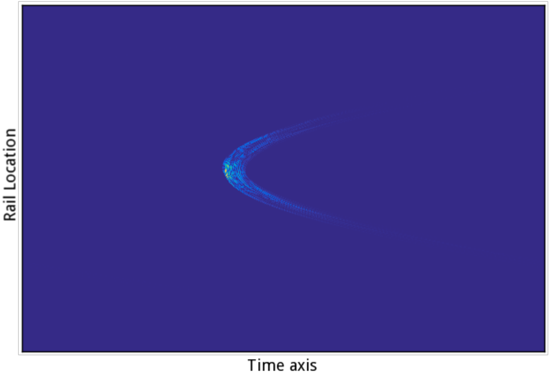
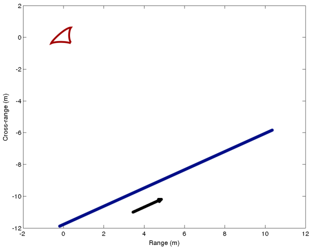


Figure: The scattering off of a 155mm Howitzer shell

1D Classifier

Consider each 2D wavefield as a set of 1D signals, perform a scattering transform using Morlet Wavelets, and use the resulting vectors as the input to a linear classifier, in our case Sparse logistic regression.



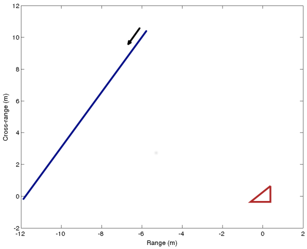


Figure: The triangle, the sharkfin and observation paths

Synthetic Experiments
Material Discrimination

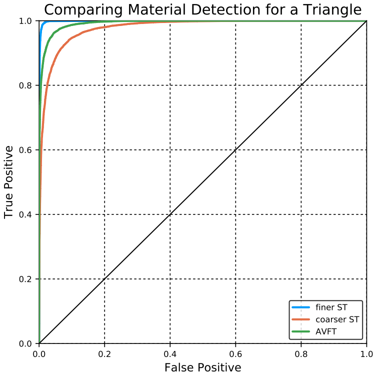


Figure: The ROC curve for detecting the material difference in a triangle, for speeds of sound $c_1 = 2000\text{m/s}$ and $c_1 = 2500\text{m/s}$.

Synthetic Experiments
Shape Discrimination

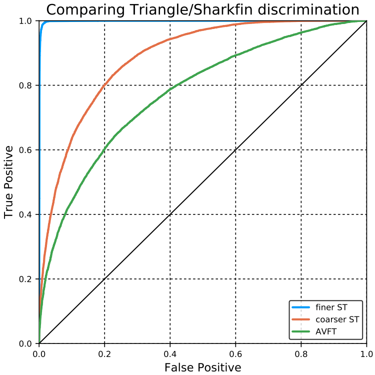


Figure: The ROC curve for discriminating a shark-fin from a triangle where both have a speed of sound fixed at $2000\mathrm{m/s}$.

Real Experiments
UXO Detection

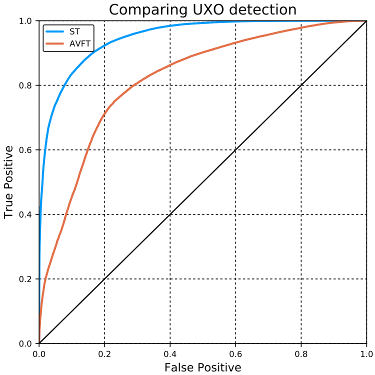
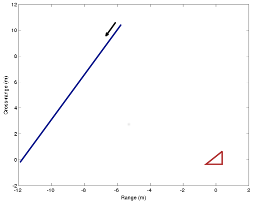
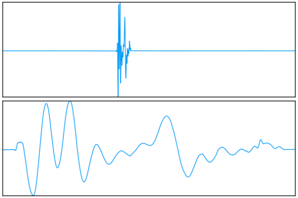


Figure: The ROC curve for detecting UXOs. The scattering transform has two layers, with quality factors $Q_1 = Q_2 = 8$

Object Domain vs Signal Domain



(a) object domain



(b) signal domain

The invariants discussed in the first part of the talk are in the signal domain $f(t-c)$. What happens when we move or deform the triangle?

- Translation perpendicular to the rail
- Translation along rail
- Rotation
- Shape deformation
- Material deformation

Acknowledgements

- ONR Grants: N00014-12-1-0177; N00014-16-1-2255
- Vincent Bodin (former intern with Saito, responsible for initial database code)
- Jim Bremer and Ian Sammis (UC Davis, wrote the fast Helmholtz solver)
- Bradley Marchand
- Frank Crosby, Julia Gazagnaire (NSWC-PCD, FL, real dataset)
- GLMNET (T. Hastie and his group)
  - Simon Kornblith for his Julia wrapper for the fortran GLMNET
- SCATNET (S. Mallat and his group)

References I

Amari, S. (1968).
Invariant structures of signal and feature space in pattern recognition problems.
RAAG Memoirs, 4(1-2):553–566.

Mallat, S. (2012).
Group invariant scattering.
Comm. Pure Appl. Math., 65(10):1331–1398.

Otsu, O. (1973).
An invariant theory of linear functionals as linear feature extractors.
Bulletin of the Electrotechnical Laboratory, 37(10):893–913.

Wiatowski, T. and Bölcskei, H. (2015).
Deep convolutional neural networks based on semi-discrete frames.
IEEE Int. Symp. on Info. Theory pages 1212–1216.

Real Experiments
Examples

Figure: Various unexploded ordinance (UXO), replicas, and other sea debris.

Synthetic Experiments
Setup Description

We use Mallat's framework with Morlet Wavelets, and compare with the absolute value of the Fourier transform (AVFT). We use two scattering transforms:

Type	Q_1	Q_2	Q_3
Finer	8	8	1
Coarser	8	4	4

- Each signal is normalized so the maximum amplitude is 1
- White Gaussian noise is added to get average SNR is about 5dB.
- Multiclass logistic regression with Lasso (via `GLMNET`) is used as a feature extractor and a classifier.
- Perform 10-fold cross validation, i.e., repeat the classification 10 times by randomly splitting the whole dataset into training and test sets with 50/50.