Leveraging Machine Learning for High-Resolution Restoration of Satellite Imagery

Daniel L. Pimentel-Alarcón, Ashish Tiwari
Georgia State University, Atlanta, GA

Douglas A. Hope
Hope Scientific Renaissance LLC, Colorado Springs, CO

Stuart M. Jefferies
Georgia State University, Atlanta, GA,
Institute for Astronomy, University of Hawaii, HI

ABSTRACT

High-resolution imaging of satellites through the Earth's turbulent atmosphere plays a pivotal role in space situational awareness. Imaging of this type requires screening thousands of images to select a few promising ones that can be combined into a restored image. The quality of the restoration depends on the selected images.

We propose a novel image selection method, based on machine learning techniques. In particular, we select a small set of images that well-represent the estimated Fourier coefficients of the satellite images. This is done using the least absolute shrinkage and selection operator (Lasso). We compare the reconstruction based on our selection method with the current state-of-the-art technique, which is based on a greedy Fourier analysis.

1 INTRODUCTION

A major goal for ground-based space situational awareness (SSA) is to achieve uninterrupted coverage of the entire sky above each of the Air Force's surveillance telescopes, and to be able to monitor all objects from low-Earth orbit to geosynchronous orbit (GEO), ideally, through high-resolution, multi-wavelength imagery. Realizing this goal requires developing the capability to observe faint objects through strong atmospheric turbulence.

While high-resolution imaging of objects in GEO remains elusive, there has been some recent progress in our ability to achieve imagery at diffraction-limited resolution through strong turbulence [1]. It appears that relief with this problem is afforded by turning to multi-aperture imaging systems. However, such systems can generate extremely large volumes of data, and removing the deleterious effects of atmospheric blur from the raw imagery requires digital restoration algorithms that can accommodate 10^3 - 10^5 images (or more, e.g., when looking at faint objects that remain stationary for extended periods of time). Processing this volume of data will require finding solutions to not only effectively and efficiently extract information from the data, but also to overcome practical issues such as finite computer memory and the optimization of the millions of parameters required to describe the imagery and wave front sensing data used in the restoration process.

There are a number of image restoration algorithms used in SSA. Arguably the most popular is multi-frame blind deconvolution (MFBD) [2]. This algorithm has been in use for SSA since the early 2000s. In 2011, a variation of the classic MFBD algorithm was introduced that required solving for fewer variables than in the classic algorithm, but which still used all the available data to constrain the solution [3]. The initial application of the new approach, referred to as compact multi-frame blind deconvolution (CMFBD), was found to be significantly faster than MFBD, and showed an indication that it may be able to provide restorations of higher quality (fewer artifacts). Since its introduction, the CMFBD approach has become the foundation of several MFBD-based algorithms that have been developed for applications such as high-accuracy wave front sensing from imaging Shack-Hartmann sensor data, and imaging through strong turbulence: both of which contribute to space situational awareness by providing high-quality imagery through a wide range of observing conditions, thus increasing the area of sky available for surveillance.

The first step in the CMFBD algorithm requires identifying and selecting the "best frames" from a data set. However, what constitutes the *best* frames remains an open question. At this time the selection is made by finding a small number of data frames that provide sufficiently high SNR over a given coverage of spatial frequencies in the target object: typically 70%. However, there may be a better approach. In this paper we use a widely used statistical machine learning approach to selecting significant components, namely the *least absolute shrinkage and selection operator* (Lasso) [4].

2 PROBLEM FORMULATION

Suppose we observe m × n matrices $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$, given by

$$\mathbf{Y}_i \;=\; \mathbf{A}_i * \mathbf{X}, \tag{1}$$

where $\mathbf{A}_i$ is an unknown m × n transformation matrix modeling atmospheric turbulence, $*$ denotes the convolution operator, and $\mathbf{X}$ is an m × n image matrix. Our goal is to recover the image $\mathbf{X}$ based on the observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$. See Figure 1 below to build some intuition.

As mentioned before, CMFBD can perform this type of recovery. However, CMFBD relies on selecting a *good* subset of the matrices $\mathbf{Y}_i$ that allows a good reconstruction of $\mathbf{X}$. The challenge is to select the right matrices $\mathbf{Y}_i$, i.e., to select a good subset $\mathbf{\Omega} \subset \{1, \ldots, N\} =: [N]$.

2.1 Frame Selection Based on Greedy Fourier Plane Coverage

Current state-of-the-art selects frames based on a 3-step process. The first step is to determine the coverage of the spatial frequencies of the target under observation. This is done by adding the power spectra of a large number of data frames. More precisely, let $\mathcal{Y}_i$, $\mathcal{A}_i$, and $\mathcal{X}$ denote the Fourier transforms of $\mathbf{Y}_i$, $\mathbf{A}_i$, and $\mathbf{X}$, respectively. Since convolution in time-domain equates to multiplication in frequency domain, it is easy to see that

$$\mathcal{Y}_i \;=\; \mathcal{A}_i \odot \mathcal{X},$$

where $\odot$ denotes the Hadamard (point-wise) product. Since atmospheric transformations are isotropic,

$$\bar{\mathcal{Y}} \;:=\; \frac{1}{N}\sum_{i=1}^{N} \mathcal{Y}_i \;=\; \mathcal{X} \odot \frac{1}{N}\sum_{i=1}^{N} \mathcal{A}_i \;\xrightarrow{N\to\infty}\; \sigma\mathcal{X},$$

where σ is a constant modeling noise power. In other words, the effect of the signal modulation by atmospheric turbulence will average out, and so $\bar{\mathcal{Y}}$ provides an estimator of $\mathcal{X}$.

Because the atmospheric modulation transfer function always reduces the signal, it is reasonable to assume that the highest signal-to-noise sample at a given discrete spatial frequency in the entire dataset is the one with the highest signal level. Hence, the second step is to generate a binary mask $\mathbf{M}_i \in \{0,1\}^{m\times n}$ for each frame $\mathbf{Y}_i$. An entry/pixel in $\mathbf{M}_i$ takes the value 1 if and only if the i^{th} frame $\mathcal{Y}_i$ has the highest signal in the corresponding Fourier frequency/pixel in the entire dataset. The third step is to then rank the frames according to the number of 1's in each individual binary mask $\mathbf{M}_i$. The set $\mathbf{\Omega}$ is formed with the s frames whose binary masks have the largest number of 1's (which are arguably the most promising frames), where s is a tuning parameter.

2.2 Frame Selection Based on Lasso

Recall that our goal is to select a subset $\mathbf{\Omega}$ of the N frames $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ to use in the CMFBD reconstruction of $\mathbf{X}$. The downside of the greedy approach discussed in Section 2.1 is that the selected frames may contain redundant frequencies, and so they may either leave some frequencies in $\mathbf{X}$ unrepresented, or require a large number of frames to represent $\mathbf{X}$ accurately.

Here we will use a more structured strategy to search for the smallest subset of $\mathbf{Y}_i$'s that represent $\mathbf{X}$ accurately enough. Ideally we would like to find the minimal set $\mathbf{\Omega}$ such that $\mathcal{X}$ can be written as a linear

combination of the $\mathbf{\mathcal{Y}}_i$'s indicated in $\mathbf{\Omega}$. We can write this as the following optimization:

$$\underset{\mathbf{\Omega}\subset[N],\boldsymbol{\omega}\in\mathbb{R}^N}{\arg\min} \ |\mathbf{\Omega}| \quad \text{subject to} \quad \mathbf{\mathcal{X}} = \sum_{i\in\mathbf{\Omega}} \omega_i \mathbf{\mathcal{Y}}_i. \tag{2}$$

where $\boldsymbol{\omega} := [\omega_1 \ \cdots \ \omega_N]^T$ denotes the vector of coefficients of the linear combination of the $\mathbf{\mathcal{Y}}_i$'s that will produce $\mathbf{\mathcal{X}}$. Notice that if $\omega_i = 0$ for every $i \notin \mathbf{\Omega}$, then $\sum_{i\in\mathbf{\Omega}} \omega_i \mathbf{\mathcal{Y}}_i = \sum_{i=1}^N \omega_i \mathbf{\mathcal{Y}}_i$. Hence, we can simplify (2) in terms of only $\boldsymbol{\omega}$, as follows:

$$\underset{\boldsymbol{\omega}\in\mathbb{R}^N}{\arg\min} \ \|\boldsymbol{\omega}\|_0 \quad \text{subject to} \quad \mathbf{\mathcal{X}} = \sum_{i=1}^N \omega_i \mathbf{\mathcal{Y}}_i, \tag{3}$$

where $\|\cdot\|_0$ denotes the so-called ℓ_0-norm, given by the number of non-zero entries. Of course, it may be overly ambitious to pretend to represent $\mathbf{\mathcal{X}}$ exactly with only a small subset of $\mathbf{\mathcal{Y}}_i$'s. Hence, we can relax (3) to only require that our linear combination of $\mathbf{\mathcal{Y}}_i$'s is approximately equal to $\mathbf{\mathcal{X}}$:

$$\underset{\boldsymbol{\omega}\in\mathbb{R}^N}{\arg\min} \ \|\boldsymbol{\omega}\|_0 \quad \text{subject to} \quad \left\| \mathbf{\mathcal{X}} - \sum_{i\in\mathbf{\Omega}} \omega_i \mathbf{\mathcal{Y}}_i \right\|_F < \epsilon, \tag{4}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, given by the sum of the squared entries in a matrix, and $\epsilon \geq 0$ denotes the tolerance (error) that we will allow. Unfortunately we do not know $\mathbf{\mathcal{X}}$. However we know from our discussion in Section 2.1 that $\bar{\mathbf{\mathcal{Y}}} = \frac{1}{N}\sum_{i=1}^N \mathbf{\mathcal{Y}}_i$ provides an estimator of $\mathbf{\mathcal{X}}$. Hence we can reformulate (4) as:

$$\underset{\boldsymbol{\omega}\in\mathbb{R}^N}{\arg\min} \ \|\boldsymbol{\omega}\|_0 \quad \text{subject to} \quad \left\| \bar{\mathbf{\mathcal{Y}}} - \sum_{i\in\mathbf{\Omega}} \omega_i \mathbf{\mathcal{Y}}_i \right\|_F < \epsilon. \tag{5}$$

Unfortunately, minimizing (5) turns out to be a non-convex combinatorial NP-hard problem (the so-called ℓ_0-norm in fact not a norm, as it does not satisfy norm's second property, namely it is not absolutely homogeneous). To address this, following what is now a standard approach in machine learning and statistics, we will instead aim to optimize its closest convex relaxation

$$\underset{\boldsymbol{\omega}\in\mathbb{R}^N}{\arg\min} \ \|\boldsymbol{\omega}\|_1 \quad \text{subject to} \quad \left\| \bar{\mathbf{\mathcal{Y}}} - \sum_{i=1}^N \omega_i \mathbf{\mathcal{Y}}_i \right\|_F < \epsilon, \tag{6}$$

where $\|\cdot\|_1$ denotes the ℓ_1-norm, given by the sum of absolute values. The difference between (5) and (6) is that the former will produce a *hard* sparse solution vector $\boldsymbol{\omega}$ with exact-zero entries, while the later will produce a *soft* sparse solution vector $\boldsymbol{\omega}$ with many near-zero entries, and a few significant ones. The sparsity pattern reveals the most important coefficients, which in turn provides the desired set $\mathbf{\Omega}$ indicating the $\mathbf{Y}_i$'s to use in the CMFBD reconstruction. We can further rewrite (6) in Lagrangian form, to obtain the following Lasso-type formulation:

$$\underset{\boldsymbol{\omega}\in\mathbb{R}^N}{\arg\min} \ \|\boldsymbol{\omega}\|_1 \ + \ \frac{\lambda}{2} \left\| \bar{\mathbf{\mathcal{Y}}} - \sum_{i=1}^N \omega_i \mathbf{\mathcal{Y}}_i \right\|_F, \tag{7}$$

where $\lambda > 0$ is a proxy for ϵ.

3 Experiments

In our experiments we compare our Lasso approach with the current state-of-the-art (greedy) reconstruction algorithm. To this end we generated a 630×840 satellite image $\mathbf{X}$, emulating the satellite image that we

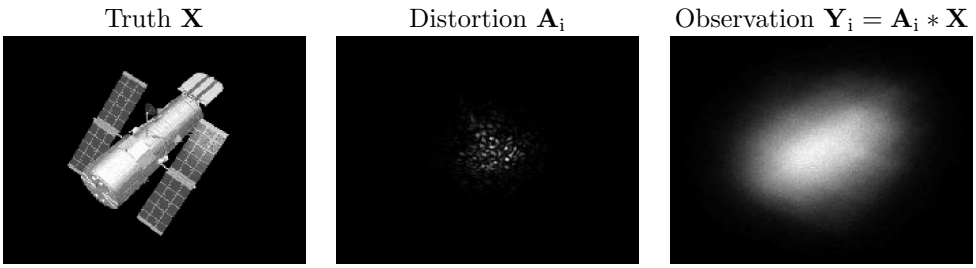

Figure 1: **Left:** Satellite image $\mathbf{X}$ that we aim to recover. **Center:** One point-spread function $\mathbf{A}_\mathrm{i}$ showing how the atmosphere would transform a single pixel image at time i. **Right:** Distorted observation $\mathbf{Y}_\mathrm{i} = \mathbf{A}_\mathrm{i} * \mathbf{X}$. We observe multiple images $\mathbf{Y}_1, \ldots, \mathbf{Y}_\mathrm{N}$, each distorted by a different transformation $\mathbf{A}_1, \ldots, \mathbf{A}_\mathrm{N}$, and our goal is to recover $\mathbf{X}$.

want to recover. Next we generated $\mathrm{N} = 2,500$ atmospheric transformations $\mathbf{A}_1, \ldots, \mathbf{A}_\mathrm{N}$, as well as the corresponding images $\mathbf{Y}_1, \ldots, \mathbf{Y}_\mathrm{N}$ according to (1), emulating the observations we would obtain from a 3.5m telescope observing at 0.5 μm. Figure 1 shows an example.

Next we computed the Fourier transforms of $\mathbf{Y}_1, \ldots, \mathbf{Y}_\mathrm{N}$ to obtain $\mathcal{Y}_1, \ldots, \mathcal{Y}_\mathrm{N}$, and vectorized (stacked the columns of) each of these images to obtain vectors $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_\mathrm{N}$, each with $\mathrm{D} := (630)(840) = 529200$ entries. Letting $\boldsymbol{\mathcal{V}} := [\boldsymbol{y}_1 \ \ldots \ \boldsymbol{y}_\mathrm{N}]$ be the $\mathrm{D} \times \mathrm{N}$ matrix containing all the vectorized observations, and letting $\bar{\boldsymbol{y}} := \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \boldsymbol{y}_\mathrm{i}$ be the vectorization of $\bar{\mathcal{Y}}$, we can rewrite (7) in the standard Lasso form:

$$\underset{\boldsymbol{\omega} \in \mathbb{R}^\mathrm{N}}{\arg\min} \ \|\boldsymbol{\omega}\|_1 \ + \ \frac{\lambda}{2} \|\bar{\boldsymbol{y}} - \boldsymbol{\mathcal{V}}\boldsymbol{\omega}\|_2 \,, \tag{8}$$

where $\|\cdot\|_2$ denotes the euclidean norm, given by the sum of the squared entries in a vector. At this point, we are able to use any standard Lasso package to solve (8). In our experiments we used the LassoRegression class from the Linear Model package of the Scikit-learn Python library [5].

Recall that the goal is to determine a subset $\boldsymbol{\Omega}$ of the N frames $\mathbf{Y}_1, \ldots, \mathbf{Y}_\mathrm{N}$ to use for the CMFBD reconstruction of $\mathbf{X}$. Depending on $\boldsymbol{\Omega}$, the reconstruction of $\mathbf{X}$ may be better or worse. The current state-of-the-art (greedy) approach selects s frames that best cover the Fourier spectra of $\bar{\mathcal{Y}}$, where s is a tuning parameter. To compare things vis-à-vis, we fixed $\mathrm{s} = 20$, and tuned λ in (8) such that it produced a solution $\boldsymbol{\omega}$ with 20 non-zero values indicating the 20 $\mathbf{Y}_\mathrm{i}$'s to use in the reconstruction. Figure 2 shows a comparison of the reconstructions obtained with both methods. In contrast, Figure 3 shows the reconstructions obtained with 20 randomly selected $\mathbf{Y}_\mathrm{i}$'s.

For a more objective comparison, we also computed the root mean squared error (RMSE) of each reconstruction, given by $\|\mathbf{X} - \hat{\mathbf{X}}\|_\mathrm{F}$, where $\hat{\mathbf{X}}$ denotes the reconstruction of $\mathbf{X}$. The greedy approach achieved an RMSE of 0.352, the Lasso approach achieved 0.337, and random selections achieved an average of 0.431 with a standard deviation of of 0.04. This is indicative of the promising results that can be attained with

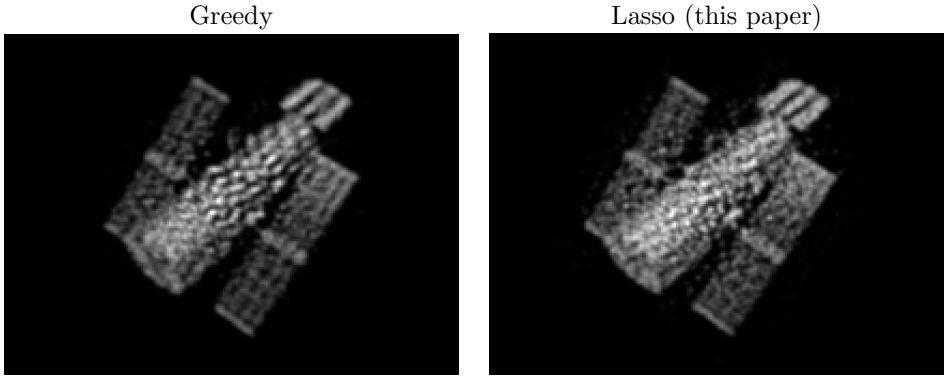

Figure 2: Reconstructed images using CMFBD in combination with the state-of-the-art greedy Fourier coverage approach, and our Lasso approach.

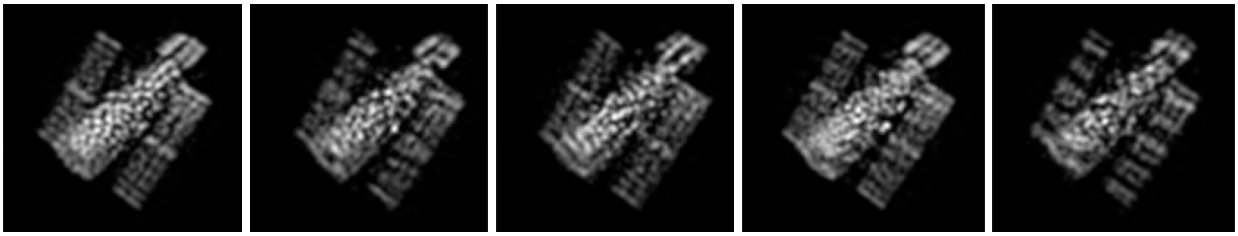

Figure 3: Reconstructed images using CMFBD with a random selection of $\mathbf{Y}_i$'s.

the Lasso approach, which achieves much higher performance than random selection, and outperforms the current state-of-the-art, even with a suboptimal λ. Our future work will focus on adequately tuning λ to produce the number of frames that optimize the reconstruction of $\mathbf{X}$.

4 Acknowledgements

This work was supported by award FA9550-14-1-0178 from the Air Force Office of Scientific Research.

References

[1] D. Hope, S. M. Jefferies, M. Hart, and J. Nagy, *High-resolution speckle imaging through strong atmospheric turbulence*, Optics Express, 24, 12116-12129, 2016.

[2] S. M. Jefferies and J. C. Christou, 1993, *Restoration of Astronomical Images by Iterative Blind Deconvolution*, Astrophysical Journal, 415, 862-864, 1993.

[3] D. A. Hope and S. M. Jefferies, *Compact multi-frame blind deconvolution*, Optics Letters, 36, 867-869, 2011.

[4] R. Tibshirani, *Regression Shrinkage and Selection via the Lasso*, Journal of the Royal Statistical Society, 1996.

[5] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research, `http://scikit-learn.org/stable`, 2011.