

Fast PSF reconstruction using the frozen flow hypothesis

James Nagy

*Mathematics and Computer Science
Emory University
Atlanta, GA 30322, USA
nagy@mathcs.emory.edu*

Stuart Jefferies

*Institute for Astronomy
University of Hawaii
Pukalani, HI 96768, USA
stuartj@ifa.hawaii.edu*

Qing Chu

*Mathematics and Computer Science
Emory University
Atlanta, GA 30322, USA
qchu@mathcs.emory.edu*

ABSTRACT

When imaging space objects from ground-based telescopes, observed images are degraded by atmospheric blurring. If an accurate estimate of the point spread function (PSF) is known, then deconvolution algorithms can be used to restore the image. Wavefront sensors (WFS) collect gradients of the wavefront, which can then be used to estimate the PSF. However, the relatively coarse grid used by a typical WFS limits the accuracy of the PSF estimate, especially when there is severe atmospheric turbulence. Using the frozen flow hypothesis, it is possible to capture the inherent temporal correlations present in wavefronts in consecutive frames of data. Exploiting these correlations can lead to more accurate estimation of the PSF. Here we address the computational aspects of the problem. Specifically we show that the process of extracting additional information from the correlated WFS data can be done by solving a sparse linear least squares problem.

1. INTRODUCTION

When viewing objects in space using a ground-based telescope, optical aberrations introduced by turbulence in the atmosphere cause the observed image to be blurred. We assume that the image formation model of this situation is

$$g(x, y) = k(x, y) * f(x, y) + \varepsilon_g(x, y) \quad (1)$$

where f is the true object, k is the point spread function (PSF), $*$ denotes convolution, g is the observed blurred image, and ε_g is additive noise. Using a Fourier optics model for atmospheric turbulence, the PSF k can be expressed in terms for the incoming wavefront of light, ϕ that reaches the telescope mirror,

$$k(x, y) = \left| \mathcal{F}^{-1} \{ \mathcal{P}(x, y) e^{\iota \phi(x, y)} \} \right|^2. \quad (2)$$

Here \mathcal{F}^{-1} is the inverse Fourier transform, $\mathcal{P}(x, y)$ is the mirror aperture function, and $\iota = \sqrt{-1}$.

Image restoration algorithms compute an approximation of f from the given blurred image g and the PSF k . The quality of the restoration depends on how accurately the PSF, or equivalently the wavefront

ϕ , can be estimated. Approximations of ϕ can be computed from gradient measurements obtained from a wavefront sensor (WFS) on the telescope [8, 9]. Mathematically the wavefront reconstruction problem can be modeled as

$$\begin{bmatrix} s_x \\ s_y \end{bmatrix} = \begin{bmatrix} D_x \\ D_y \end{bmatrix} \phi + \epsilon_s, \quad (3)$$

where s_x and s_y are discrete, noisy measurements of the horizontal and vertical derivatives of ϕ ; D_x and D_y are discrete, horizontal and vertical derivative operators. The precise structure of D_x and D_y depends on the sensor geometry. Common wavefront sensor geometries discussed in the literature are those of Hudgin [6] and Fried [4]. In our work we assume a Fried geometry because it is more commonly used in operational adaptive optics systems. However, our methods can be implemented with any geometry.

A difficulty with using equation (3) is that the gradient measurements are given on a relatively coarse grid compared to the observed image data. Interpolation of the gradient data to a fine grid can be used to reconstruct the wavefront and corresponding PSFs. Unfortunately such an approach may not provide enough accuracy in the PSF to allow for quality restorations, especially when seeing conditions are poor (e.g., during daylight hours, or when looking close to the horizon).

A recent approach proposed by Jefferies and Hart [7] uses multiple frames of data and a frozen flow hypothesis (FFH) of the atmosphere to construct a composite, higher resolution grid of gradient measurements. The purpose of this paper is to describe an efficient implementation of this approach, and to show that it can be easily adapted to situations when there are nonlinear changes in the wind velocity from frame to frame.

This paper is organized as follows. In Section 2 we describe the FFH, and show that it can be modeled as a linear system involving sparse matrices. A regularized linear least squares approach is then proposed to reconstruct high resolution gradients from coarse grid WFS measurements. In Section 3 we provide some experiments on simulated data to illustrate the effectiveness of our approach. Some concluding remarks are given in Section 4.

2. FROZEN FLOW HYPOTHESIS

The FFH assumes that atmospheric turbulence can be modeled by a series of independent static layers, each moving across the telescope aperture with the prevailing wind at the altitude of the layer. Because of its simplicity, the FFH is frequently used as the basis for numerical studies of telescope imaging performance, particularly in the modeling of adaptive optics (AO) systems. While the FFH is observed not to hold in the real world over long time scales, a number of studies have shown that it is a reasonable approximation for short but still interesting periods [5, 10, 11].

To use the FFH to reconstruct wavefront gradients, several frames of data are collected over a short time period, each giving gradient measurements at a different set of grid points. This is illustrated in Figure 1 for two different velocity profiles. For ease of presentation, we consider only one layer; clearly multiple, overlapping layers will provide even more grid points in regions where the various layers overlap each other. Note that the composite grid resolution depends on the velocity profile; in the example shown in the top part of Figure 1, the velocity (direction and magnitude) remains constant from frame to frame, and the magnitude of the velocity is relatively small. A more extreme situation is illustrated in the bottom part of Figure 1, where there is a nonlinear change in the velocity from frame to frame.

If the composite grid is uniform, then the high resolution wavefront gradient measurements can be obtained from the composite grid through a simple extraction process. However, a situation such as this would only occur when there is a very specific, constant velocity from frame to frame. For example, consider the simple situation illustrated in Figure 2, where an underlying uniform grid is represented by x's, and two coarse grids are denoted by o's. Although one frame falls directly on the underlying uniform grid, the other frame does not. Including rotation in the velocity profile will introduce additional irregularities in the composite grid of data. The computationally challenging step is to construct the composite gradient measurements on a uniform high resolution grid from the given, nonuniform, coarse grid information.

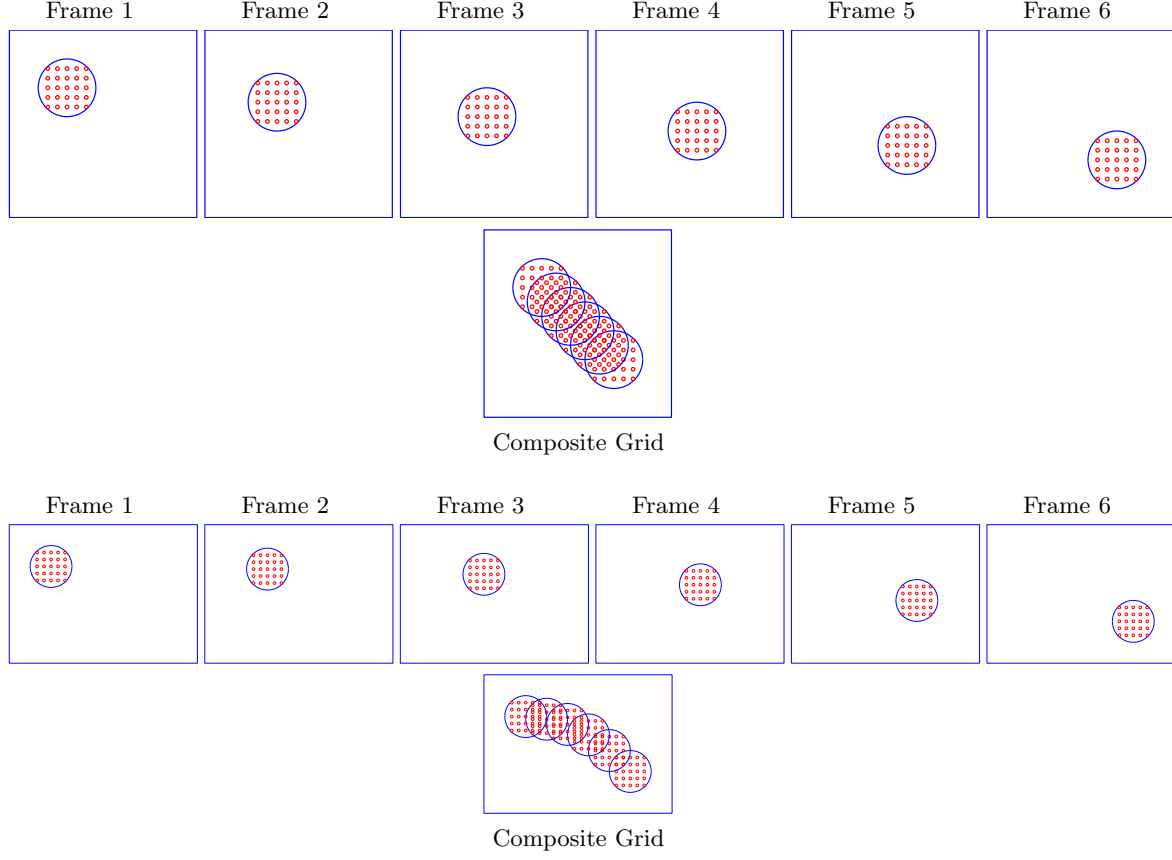


Figure 1. Illustrations of building a composite, high resolution grid using gradient grid points from several frames. The first two rows illustrate a situation when the velocity remains constant from frame to frame. The bottom two rows illustrate a situation when the velocity changes nonlinearly from frame to frame.

2.1. Linear model of wavefront motion

In this subsection we describe the approach we use to model motion of the wavefront or, equivalently, motion of the gradient fields. To simplify notation, we consider a general situation where $s(x, y)$ is a function and let S be an array of discrete samples of $s(x, y)$; that is, the (i, j) entries of S are given by

$$S(i, j) = s(x_i, y_j), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, n.$$

Using the FFH, we can assume changes in the wavefront and corresponding gradients from frame to frame are modeled as a rigid movement of $s(x, y)$. Suppose \hat{S} is a discretization of $s(x, y)$ after a rigid movement; this can be described through a 3×3 affine transformation:

$$\hat{S}(\hat{i}, \hat{j}) = s(\hat{x}_{\hat{i}}, \hat{y}_{\hat{j}}), \quad \text{where} \quad \begin{bmatrix} \hat{x}_{\hat{i}} \\ \hat{y}_{\hat{j}} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_j \\ 1 \end{bmatrix}. \quad (4)$$

Note that because the function s is not known at every point (x, y) (all that is known is the discrete values S), it may not be possible to evaluate $s(\hat{x}_{\hat{i}}, \hat{y}_{\hat{j}})$, unless $\hat{x}_{\hat{i}} = x_{\hat{i}}$ and $\hat{y}_{\hat{j}} = y_{\hat{j}}$ for integers \hat{i} and \hat{j} , $1 \leq \hat{i} \leq n$ and $1 \leq \hat{j} \leq n$. However, an approximation of $s(\hat{x}_{\hat{i}}, \hat{y}_{\hat{j}})$ can be computed by interpolating known values of s near $s(\hat{x}_{\hat{i}}, \hat{y}_{\hat{j}})$. Suppose, as illustrated in Figure 3, that $s(x_{\hat{i}}, y_{\hat{j}})$, $s(x_{\hat{i}+1}, y_{\hat{j}})$, $s(x_{\hat{i}}, y_{\hat{j}+1})$

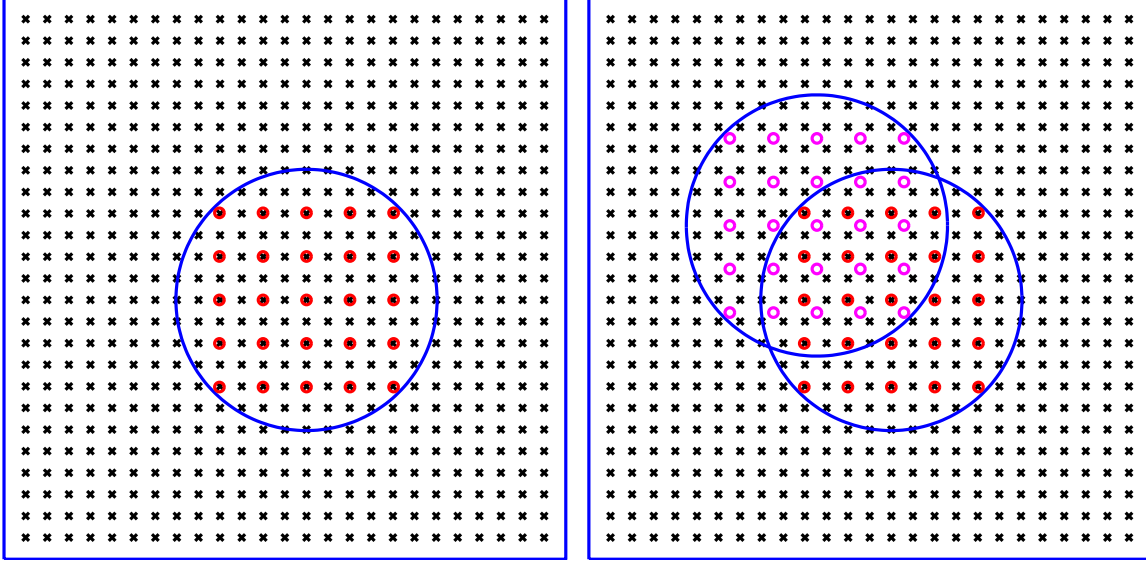


Figure 2. A single frame of data can be aligned to an underlying, uniform fine grid, as illustrated with the image on the left. However, additional frames are likely to move to locations that do not fall directly on the uniform grid; this is illustrated with the image on the right. In this figure, the x's denote points on the underlying uniform grid, and the o's denote points on the coarse grid of each frame of WFS data.

and $s(x_{i+1}, y_{j+1})$ are four known pixel values surrounding the unknown value $s(\hat{x}_i, \hat{y}_j)$. Using bilinear interpolation, a weighted average of the four pixels surrounding $s(\hat{x}_i, \hat{y}_j)$ is used for the approximation:

$$\begin{aligned} \hat{S}(i, j) &= s(\hat{x}_i, \hat{y}_j) \\ &\approx (1 - \Delta x_i)(1 - \Delta y_j)s(x_{\hat{i}}, y_{\hat{j}}) + (1 - \Delta x_i)\Delta y_j s(x_{\hat{i}}, y_{\hat{j}+1}) \\ &\quad + \Delta x_i(1 - \Delta y_j)s(x_{\hat{i}+1}, y_{\hat{j}}) + \Delta x_i\Delta y_j s(x_{\hat{i}+1}, y_{\hat{j}+1}). \end{aligned}$$

where $\Delta x_i = \hat{x}_i - x_{\hat{i}}$ and $\Delta y_j = \hat{y}_j - y_{\hat{j}}$.

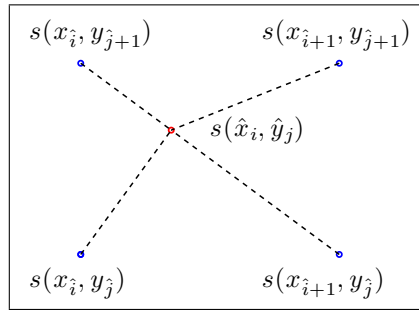


Figure 3. Illustration of bilinear interpolation, where a weighted average of the four known discrete values is used to approximate $s(\hat{x}_i, \hat{y}_j)$.

If we define vectors $\mathbf{s} = \text{vec}(S)$ and $\hat{\mathbf{s}} = \text{vec}(\hat{S})$ from the discrete data arrays (e.g., through lexicographical ordering), we can write

$$\hat{\mathbf{s}} = \mathbf{A}\mathbf{s}$$

where \mathbf{A} is a sparse matrix that contains the interpolation weights. Specifically, the k th row of \mathbf{A} contains the weights for the pixel in the k th entry of $\hat{\mathbf{s}}$. That is, in the case of bilinear interpolation, there are at most four nonzero entries per row, given by

$$(1 - \Delta x_i)(1 - \Delta y_j), \quad (1 - \Delta x_i)\Delta y_j, \quad \Delta x_i(1 - \Delta y_j), \quad \Delta x_i\Delta y_j.$$

We emphasize that by using a sparse data format (e.g., compressed row [2]) to represent \mathbf{A} , we need only keep track of the nonzero entries and their locations in the matrix \mathbf{A} . Moreover, this discussion assumes that the affine transformation used in equation (4) is known from wind velocity information.

Our discussion to this point assumes the discretization grid for $\hat{\mathbf{s}}$ is the same as \mathbf{s} . If we want $\hat{\mathbf{s}}$ to represent a subsampling of a small section of \mathbf{s} after movement, then this can be modeled as

$$\hat{\mathbf{s}} = \mathbf{R}\mathbf{W}\mathbf{A}\mathbf{s}$$

where \mathbf{W} is an indicator matrix that grabs a specified section of \mathbf{s} , and \mathbf{R} is a sparse downsampling (or restriction) matrix that transforms high resolution data to a lower resolution.

2.2. Least squares reconstruction of high resolution wavefront gradients

In this subsection we set up the linear model for reconstructing the high resolution wavefront gradients from coarse grid measurements. Let $\mathbf{s}_x^{(k)}$ and $\mathbf{s}_y^{(k)}$ be the given coarse grid wavefront gradient measurements for frames $k = 1, 2, \dots, m$. Assume that each of the gradient measurements is obtained from a moving, frozen wavefront, and that there is some overlap of information in adjacent frames, as depicted in Figure 1. Further denote the composite, high resolution gradient fields by \mathbf{s}_x and \mathbf{s}_y . The aim is, given $\mathbf{s}_x^{(k)}$ and $\mathbf{s}_y^{(k)}$, $k = 1, 2, \dots, m$, compute \mathbf{s}_x and \mathbf{s}_y . With the FFH, and the notation introduced in the previous subsection, we can describe the relationship between the high resolution information to the measured, low resolution gradient fields of each frame as

$$\mathbf{s}_x^{(k)} = \mathbf{R}\mathbf{W}\mathbf{A}_k\mathbf{s}_x \quad \text{and} \quad \mathbf{s}_y^{(k)} = \mathbf{R}\mathbf{W}\mathbf{A}_k\mathbf{s}_y \quad (5)$$

where \mathbf{A}_k is a sparse matrix that describes the motion of wavefront (and corresponding gradient fields) to those corresponding to the k th frame, and \mathbf{W} and \mathbf{R} are the indicator and downsampling matrices described in the previous subsection.

With this model we can set up linear systems of the form:

$$\begin{bmatrix} \mathbf{R}\mathbf{W}\mathbf{A}_1 \\ \mathbf{R}\mathbf{W}\mathbf{A}_2 \\ \vdots \\ \mathbf{R}\mathbf{W}\mathbf{A}_m \end{bmatrix} \mathbf{s}_x = \begin{bmatrix} \mathbf{s}_x^{(1)} \\ \mathbf{s}_x^{(2)} \\ \vdots \\ \mathbf{s}_x^{(m)} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{R}\mathbf{W}\mathbf{A}_1 \\ \mathbf{R}\mathbf{W}\mathbf{A}_2 \\ \vdots \\ \mathbf{R}\mathbf{W}\mathbf{A}_m \end{bmatrix} \mathbf{s}_y = \begin{bmatrix} \mathbf{s}_y^{(1)} \\ \mathbf{s}_y^{(2)} \\ \vdots \\ \mathbf{s}_y^{(m)} \end{bmatrix} \quad (6)$$

We remark that this linear model is very similar to that used in digital super-resolution applications; see for example [1, 3]. As in digital super-resolution, regularization is needed to stabilize the inversion process. In this work we use standard Tikhonov regularization, and thus we need to solve the following least squares problems:

$$\min_{\mathbf{s}_x} \left\| \begin{bmatrix} \mathbf{s}_x^{(1)} \\ \mathbf{s}_x^{(2)} \\ \vdots \\ \mathbf{s}_x^{(m)} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{R}\mathbf{W}\mathbf{A}_1 \\ \mathbf{R}\mathbf{W}\mathbf{A}_2 \\ \vdots \\ \mathbf{R}\mathbf{W}\mathbf{A}_m \\ \alpha \mathbf{I} \end{bmatrix} \mathbf{s}_x \right\|_2 \quad \text{and} \quad \min_{\mathbf{s}_y} \left\| \begin{bmatrix} \mathbf{s}_y^{(1)} \\ \mathbf{s}_y^{(2)} \\ \vdots \\ \mathbf{s}_y^{(m)} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{R}\mathbf{W}\mathbf{A}_1 \\ \mathbf{R}\mathbf{W}\mathbf{A}_2 \\ \vdots \\ \mathbf{R}\mathbf{W}\mathbf{A}_m \\ \alpha \mathbf{I} \end{bmatrix} \mathbf{s}_y \right\|_2. \quad (7)$$

Computational efficiency is obtained by exploiting the sparse structure of the matrices [2].

3. NUMERICAL EXPERIMENTS

In this section we report on some numerical experiments that illustrate the effectiveness of using the FFH to extract additional gradient information, as well as our linear approach for the reconstruction process. To generate a realistic simulation, we begin with a large, high resolution global wavefront, and corresponding gradient fields. We then propagate the wavefront using two different velocity profiles: first using a linear, constant velocity for each frame, and second a nonlinear change in velocity from frame to frame. In each case we generate 16 frames of data, and the measured gradient fields are obtained by downsampling the high resolution gradient fields to a 32×32 grid.

3.1. Linear constant velocity

In this first example we consider a linear constant velocity from frame to frame, similar to the illustration shown in the top part of Figure 1. Figure 4 shows the composite gradient fields obtained by solving the least squares problems given in equation (7). From these we extract high resolution estimates of the gradients for each frame. Figure 4 shows these estimates for frame 1, along with the corresponding true gradients and naive estimates obtained by up-sampling the measured low resolution gradients. In addition, we also display the corresponding reconstructed PSFs for frame 1. Rather than displaying the results for all 16 frames, we instead show in Figure 5 a plot of the relative errors for all reconstructed PSFs. Because the velocity profile in this example is similar to the illustration shown in the top part of Figure 1, where we have more high resolution grid information for the middle frames than at the beginning and end, we expect to get better estimates of the PSFs for the middle frames. The PSF errors shown in Figure 5 clearly verifies this expectation. Finally, from Figure 6 we see that the better PSF estimates, when used in a deconvolution algorithm, produces a clearly superior restoration.

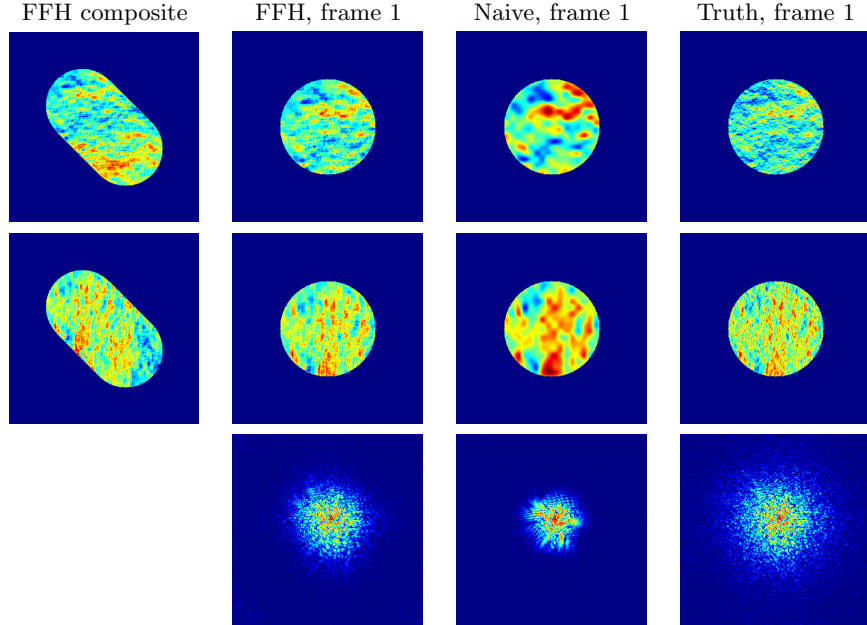


Figure 4. Comparison of reconstructed gradients for one frame. Top row shows the x gradient results, the middle row shows the y gradient results, and the last row shows the corresponding PSFs.

3.2. Nonlinear velocity

In this example velocity changes nonlinearly from frame to frame, similar to the illustration shown in the bottom part of Figure 1. Figure 7 shows the composite gradient fields obtained by solving the least squares problems given in equation (7). From these we extract high resolution estimates of the gradients

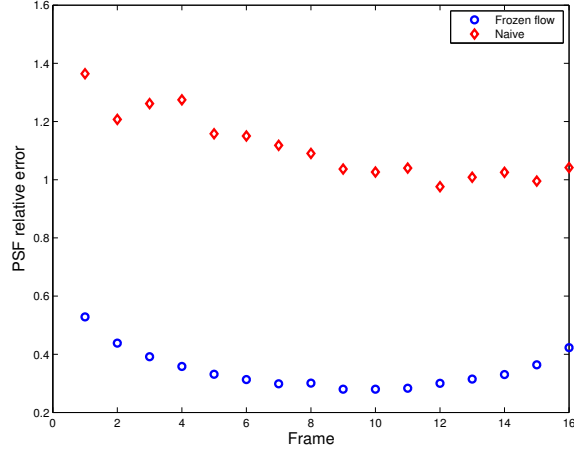


Figure 5. Plot of the errors for all reconstructed PSFs.

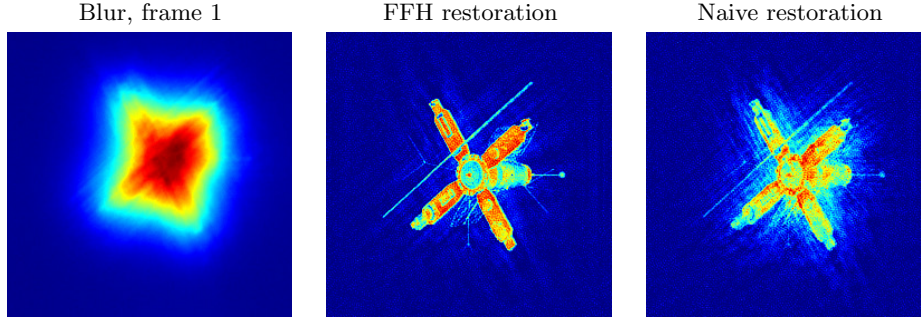


Figure 6. Comparison of the reconstructed images.

for each frame. Figure 7 shows these estimates for frame 1, along with the corresponding true gradients and naive estimates obtained by up-sampling the measured low resolution gradients. In addition, we also display the corresponding reconstructed PSFs for frame 1. Figure 8 shows a plot of the relative errors for all reconstructed PSFs. Because the velocity profile in this example is similar to the illustration shown in the bottom part of Figure 1, where we have more high resolution grid information for the first few frames, and less towards the end, we expect to get better estimates of the PSFs for the first few frames. The PSF errors shown in Figure 8 clearly verifies this expectation. Finally, from Figure 9 we see that the better PSF estimates, when used in a deconvolution algorithm, produces a clearly superior restoration.

4. CONCLUDING REMARKS

The frozen flow hypothesis captures the inherent temporal correlations present in wavefronts in consecutive frames of data. Exploiting these correlations can lead to more accurate PSF estimations, as illustrated by the numerical experiments in this paper. We have shown that reconstructing high resolution gradient fields from coarse grid WFS measurements can be modeled as a linear system involving sparse matrices. Because our approach requires only to solve linear least squares problems involving sparse matrices, we can compute PSF estimates very efficiently.

Acknowledgments

The work of S. Jefferies is supported by the Air Force Office of Scientific Research grant FA9550-09-1-0216. The work of J. Nagy is supported by the Air Force Office of Scientific Research grant FA9550-09-1-0487.

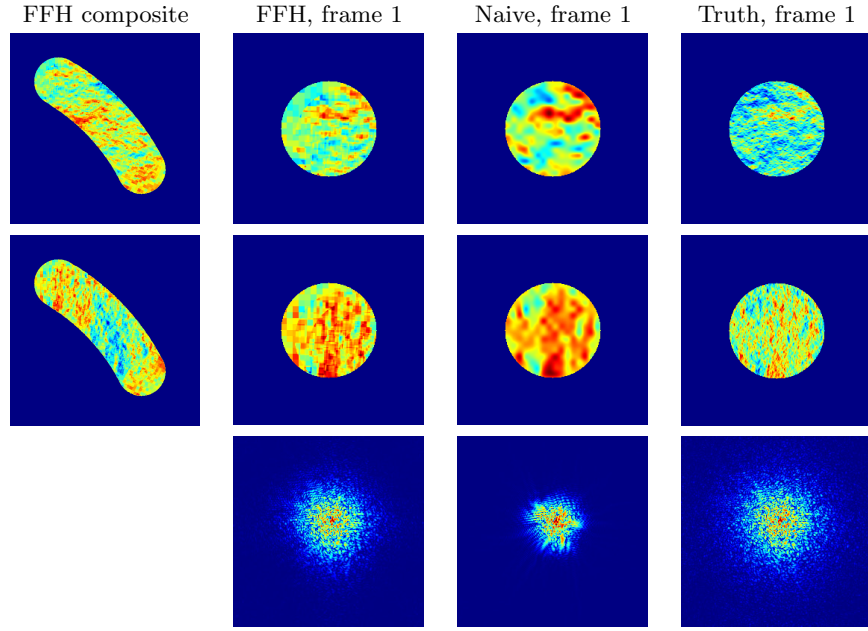


Figure 7. Comparison of reconstructed gradients for one frame. Top row shows the x gradient results, the middle row shows the y gradient results, and the last row shows the corresponding PSFs.

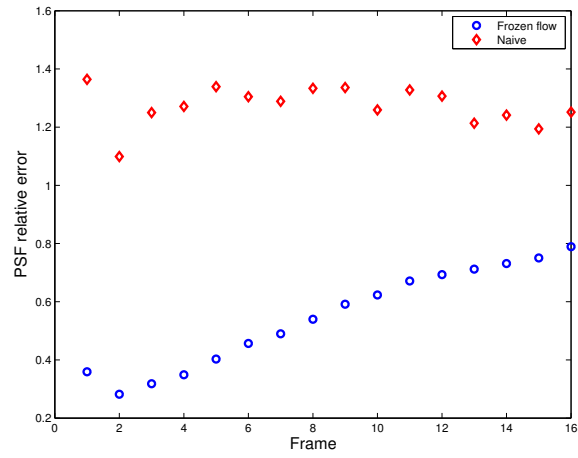


Figure 8. Plot of the errors for all reconstructed PSFs.

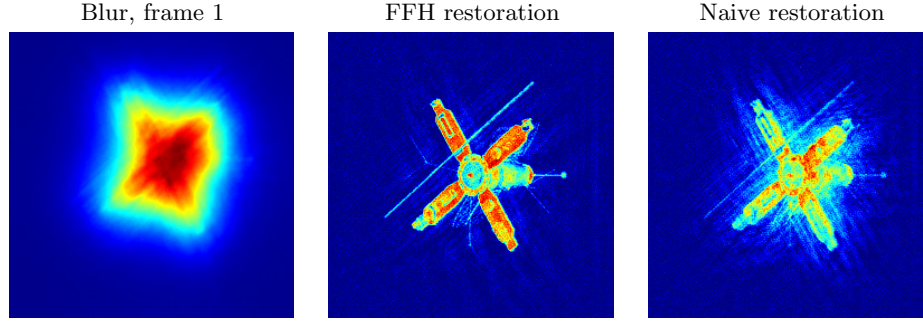


Figure 9. Comparison of the reconstructed images.

REFERENCES

1. J. M. Chung, E. Haber, and J. G. Nagy. Numerical methods for coupled super-resolution. *Inverse Problems*, 22:1261–1272, 2006.
2. T. A. Davis. *Direct Methods for Sparse Linear Systems*. SIAM, Philadelphia, PA, 2006.
3. S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Advances and challenges in super-resolution. *Int. J. Imaging Syst. Technol.*, 14(2):47–57, 2004.
4. D. L. Fried. Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements. *J. Opt. Soc. Am.*, 67(3):370–375, 1977.
5. E. Gendron and P. Léna. Single layer atmospheric turbulence demonstrated by adaptive optics observations. *Astrophysics and Space Science*, 239:221–228, 1996.
6. R. H. Hudgin. Wave-front reconstruction for compensated imaging. *J. Opt. Soc. Am.*, 67(3):375–378, 1977.
7. S. M. Jefferies and M. Hart. Deconvolution from wave front sensing using the frozen flow hypothesis. submitted, 2010.
8. S. M. Jefferies, M. Lloyd-Hart, E. K. Hege, and J. Georges. Sensing wave-front amplitude and phase with phase diversity. *Appl. Opt. IP*, 41:2095–2102, 2002.
9. L. M. Mugnier, C. Robert, J. M. Conan, V. Michau, and S. Salem. Myopic deconvolution from wave-front sensing. *J. Opt. Soc. Am. A*, 12:862–872, 2001.
10. L. A. Poyneer, M. van Dam, and J. P. Véran. Experimental verification of the frozen flow atmospheric turbulence assumption with use of astronomical adaptive optics telemetry. *J. Opt. Soc. Am. A*, 26:833–846, 2009.
11. M. Schöck and E. J. Spillar. Method for a quantitative investigation of the frozen flow hypothesis. *J. Opt. Soc. Am. A*, 17:1650–1658, 2000.