MFBD and the Local Minimum Trap

James G. Nagy

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

Veronica Mejia-Bustamante

Mathematics and Computer Science
Emory University
Atlanta, GA 30322, USA
vmejia@emory.edu

ABSTRACT

Obtaining high resolution images of space objects from ground based telescopes involves using a combination of sophisticated hardware and computational post processing techniques. An important, and often highly effective, computational post processing tool is multi-frame blind deconvolution (MFBD) for image restoration. One difficulty with using MFBD algorithms is that the nonlinear inverse problem they are designed to solve may have many local minima. Standard optimization methods that use the gradient to search for a minimum may get trapped in a local minimum, resulting in a less than optimal restored image. Moreover, in some cases the global minimum may not correspond to the best reconstructed image compared to a different local minimum. In this paper we consider a flexible optimization approach that can easily incorporate various filtering schemes and prior information, resulting in a scheme that is computationally more tractable. Practical implementation details are discussed.

1. INTRODUCTION

Image restoration is the process of reconstructing an approximation of an image from blurred and noisy measurements. The image formation process is typically modeled as a convolution equation,

$$g(x,y) = p(x,y) \otimes f(x,y) + \eta(x,y)\,,$$

where $f(x,y)$ is a function representing the true image, which is convolved with a point spread function (PSF), $p(x,y)$. Some noise, $\eta(x,y)$, is then added to obtain the observed image, $g(x,y)$. Since the convolution equation assumes the blur is spatially invariant, it cannot be used to model more difficult problems, such as when the blur is spatially variant. In addition, it does not indicate that it is often the case that the PSF can be represented in terms of only a few parameters.* We therefore consider the general discrete image formation model:

$$\mathbf{g} = \mathbf{H}(\mathbf{p}_{\text{true}})\,\mathbf{f}_{\text{true}} + \boldsymbol{\eta} \tag{1}$$

where $\mathbf{g}$ is a vector representing the observed, blurred and noisy image, and $\mathbf{f}_{\text{true}}$ is a vector representing the unknown true image we wish to reconstruct. $\mathbf{H}(\mathbf{p}_{\text{true}})$ is an ill-conditioned matrix defined by a vector of parameters, $\mathbf{p}_{\text{true}}$. $\mathbf{H}$ may be sparse and/or structured. For example, if the blur is spatially invariant and periodic boundary conditions are imposed, then $\mathbf{H}$ has a circulant matrix structure. $\mathbf{p}_{\text{true}}$ is a vector of parameters defining the (true) blurring operation. For example, in the case of spatially invariant

*It should be noted that it may also be possible to put $f(x,y)$ in a parametric form. However, this is more difficult and restrictive; it requires having a significant amount of information about the object being imaged. This is an interesting problem, which may be considered in the class of *sparse sensing*.

blurs, $\mathbf{p}_{\text{true}}$ could simply be the pixel (image space) values of the PSF. Or $\mathbf{p}_{\text{true}}$ could be a small set of parameters that define the PSF, such as with a Zernike-based representation [29]. In general we assume that the number of parameters defining $\mathbf{p}_{\text{true}}$ is significantly smaller than the number of pixels in the observed image. $\boldsymbol{\eta}$ is a vector that represents unknown additive noise in the measured data. Generally $\boldsymbol{\eta}$ is a combination of background and readout noise, where the background noise is modeled as a Poisson random process with fixed Poisson parameter β, and the readout noise is modeled as a Gaussian random process with mean 0 and fixed variance σ^2.

The aim of computational postprocessing is to reconstruct an approximation, $\mathbf{f}$, of the "true" image $\mathbf{f}_{\text{true}}$. The term *deconvolution* is typically used when $\mathbf{p}_{\text{true}}$ (i.e., the true blurring operator) is known, whereas *blind deconvolution* implies that $\mathbf{p}_{\text{true}}$ is not known. In either case, the computational problem involves solving an optimization problem,

$$\underline{\text{Deconvolution}} \qquad \min_{\mathbf{f}} \phi(\mathbf{f}) \qquad\qquad \underline{\text{Blind Deconvolution}} \qquad \min_{\mathbf{f},\mathbf{p}} \phi(\mathbf{f},\mathbf{p})$$

where ϕ must be specified, usually incorporating regularization and possibly including additional constraints, such as nonnegativity.

For example, in the case of deconvolution with known blurring operator, we might use a least squares fit to data term with Tikhonov regularization,

$$\phi(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda^2 \|\mathbf{L}\mathbf{f}\|_2^2 \,.$$

With an *a priori* chosen regularization operator, $\mathbf{L}$, and a prespecified regularization parameter, λ, this is a simple linear least squares problem. More difficult is determining what regularization operator and parameter is "optimal" for a given set of data. Poorly chosen $\mathbf{L}$ and λ will result in an optimization problem whose global minimum is far from the actual true solution. Moreover, even with good choices for $\mathbf{L}$ and λ, additional constraints may be needed to ensure a good reconstruction is computed. Finally, we remark that the specific form of ϕ may depend on statistical properties of the noise and data, such as when a maximum likelihood approach is used, or it may depend on the regularization approach, such as in the case of total variation [1, 2, 4, 6, 8, 13, 14, 16, 24, 30, 32, 33, 34, 39, 43]. Thus, even the "simple" problem of deconvolution with known blurring operator is nontrivial.

Blind deconvolution is much more challenging. If we have a good estimate of the blurring operator, then it may be appropriate to use a regularized total least squares model [38]. However, these techniques are not applicable for general blind deconvolution problems, and it is necessary to consider other optimization approaches. For example, we might consider (a now nonlinear) least squares data fit with Tikhonov regularization,

$$\phi(\mathbf{f},\mathbf{p}) = \|\mathbf{g} - \mathbf{H}(\mathbf{p})\mathbf{f}\|_2^2 + \lambda_f^2 \|\mathbf{L}_f \mathbf{f}\|_2^2 + \lambda_p^2 \|\mathbf{L}_p \mathbf{p}\|_2^2 \,.$$

Even if very good regularization operators and parameters are known, the problem of minimizing ϕ over $\mathbf{f}$ and $\mathbf{p}$ is extremely underdetermined; there are infinitely many solutions. The nonuniqueness difficulty of blind deconvolution problems have been well documented; see, for example, [9, 12, 15, 17, 18, 19, 20, 21, 23, 27, 29, 28, 37]. To overcome this difficulty, it is necessary to include additional constraints. Although nonnegativity is a strong constraint, it is typically not enough to ensure a good reconstruction, and so additional constraints, such as support of the object and PSF, are often needed. Another approach that can be very effective is to construct a specific ϕ for particular classes of blurs and objects [5].

Multi-frame blind deconvolution (MFBD) [25, 26, 27, 29, 31, 41, 42] reduces some of the nonuniqueness problems by using multiple images of the same object, but with different blurring operators. Specifically, one obtains a set of (for example, m) observed images,

$$\mathbf{g}^{(i)} = \mathbf{H}(\mathbf{p}_{\text{true}}^{(i)})\mathbf{f}_{\text{true}} + \boldsymbol{\eta}^{(i)} \,, \quad i = 1, 2, \ldots, m, \tag{2}$$

which can be put in our general discrete model (1) by setting

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}^{(1)} \\ \vdots \\ \mathbf{g}^{(m)} \end{bmatrix}, \quad \mathbf{p}_{\text{true}} = \begin{bmatrix} \mathbf{p}_{\text{true}}^{(1)} \\ \vdots \\ \mathbf{p}_{\text{true}}^{(m)} \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{\eta}^{(1)} \\ \vdots \\ \boldsymbol{\eta}^{(m)} \end{bmatrix}.$$

Although the multiple frames reduces, to some extent, the nonuniqueness problem, it does not completely eliminate it. In addition, the problem becomes more difficult in two respects. First, the nonlinearity involving $\mathbf{p}_{\text{true}}$ increases, thus increasing the risk of an optimization algorithm becoming trapped in local minima. In addition, the computational complexity of processing the large, multiple data sets significantly increases.

In this paper we describe optimization methods that can be used to solve the MFBD problem, and consider a reduced parameter space formulation to help alleviate some of the difficulties associated with the local minimum trap. The approach we use is general (e.g., can be used for spatially invariant as well as spatially variant blurs), and flexible (e.g., various filtering approaches and constraints can be incorporated into the algorithm). The paper is organized as follows. In Section 2 we describe a general objective function that can be used for MFBD, and a variety of approaches to compute its minima. Section 3 illustrates the challenges to finding the local minimum that corresponds to the best reconstructed image. Additional numerical results are given in Section 4, and concluding remarks are given in Section 5.

2. OPTIMIZATION METHODS

To develop an MFBD algorithm, we need first to define an objective function $\phi(\mathbf{f}, \mathbf{p})$, and then we need to use an appropriate optimization scheme to find its minima. In this paper we consider the general Tikhonov regularized least squares problem:

$$\min_{\mathbf{f},\mathbf{p}} \frac{1}{2} \left\{ \|\mathbf{H}(\mathbf{p})\,\mathbf{f} - \mathbf{g}\|_2^2 + \lambda^2 \|\mathbf{f}\|_2^2 \right\} = \min_{\mathbf{f},\mathbf{p}} \frac{1}{2} \left\| \begin{bmatrix} \mathbf{H}(\mathbf{p}) \\ \lambda \mathbf{I} \end{bmatrix} \mathbf{f} - \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix} \right\|_2^2 . \tag{3}$$

A variety of algorithms can be used to solve this problem. In this section we outline three approaches, and and conclude that a variable projection Gauss-Newton method provides an extremely computationally convenient framework for MFBD problems.

2.1. Fully Coupled Problem

To simplify notation, we rewrite the nonlinear least squares problem given in equation (3) as

$$\min_{\mathbf{z}} \phi(\mathbf{z}) = \min_{\mathbf{z}} \frac{1}{2} \|\boldsymbol{\rho}(\mathbf{z})\|_2^2 , \quad \text{where} \quad \boldsymbol{\rho}(\mathbf{z}) = \boldsymbol{\rho}(\mathbf{f}, \mathbf{p}) = \begin{bmatrix} \mathbf{H}(\mathbf{p}) \\ \lambda \mathbf{I} \end{bmatrix} \mathbf{f} - \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix}, \tag{4}$$

and $\mathbf{z}^T = [\, \mathbf{f}^T \;\; \mathbf{p}^T \,]$. Nonlinear least squares problems are solved iteratively, with algorithms having the general form:

General Iterative Algorithm

choose initial $\mathbf{z}_0 = \begin{bmatrix} \mathbf{f}_0 \\ \mathbf{p}_0 \end{bmatrix}$

for $k = 0, 1, 2, \ldots$

- choose a direction, $\mathbf{d}_k$, in which the objective can be improved
- determine how far to step, α_k, in the direction $\mathbf{d}_k$
- update the solution: $\mathbf{z}_{k+1} = \mathbf{z}_k + \alpha_k \mathbf{d}_k$
- stop when a minimum of the objective is obtained

end

Typically the direction $\mathbf{d}_k$ is chosen to approximate the Newton direction,

$$\mathbf{d}_k = -(\widehat{\phi}''(\mathbf{z}_k))^{-1}\phi'(\mathbf{z}_k)\,,$$

where $\widehat{\phi}''$ is an approximation of ϕ'', $\phi' = \mathbf{J}_\phi^T\boldsymbol{\rho}$, and $\mathbf{J}_\phi$ is the Jacobian matrix

$$\mathbf{J}_\phi = \begin{bmatrix} \boldsymbol{\rho}_f & \boldsymbol{\rho}_p \end{bmatrix} = \begin{bmatrix} \dfrac{\partial\boldsymbol{\rho}(\mathbf{f},\mathbf{p})}{\partial\mathbf{f}} & \dfrac{\partial\boldsymbol{\rho}(\mathbf{f},\mathbf{p})}{\partial\mathbf{p}} \end{bmatrix}.$$

In the case of the Gauss-Newton method, which is often recommended for nonlinear least squares problems, $\widehat{\phi}'' = \mathbf{J}_\phi^T\mathbf{J}_\phi$.

This general Gauss-Newton approach can work well, but constructing and solving the linear systems required to update $\mathbf{d}_k$ can be very expensive. Note that the dimension of the matrix $\mathbf{J}_\phi$ corresponds to the number of pixels in the image plus the number of parameters in $\mathbf{p}$, and thus $\mathbf{J}_\phi$ may be on the order of $10^6 \times 10^6$. Thus, instead of using Gauss-Newton, it might be preferable to use a low storage scheme such as the (nonlinear) conjugate gradient method. But there is a tradeoff – although the cost per iteration is reduced, the number of iterations needed to attain a minimum can increase significantly.

There are other difficulties with using this fully coupled optimization approach. For example, it requires specifying *a priori* the regularization parameter λ. Although there are well known approaches for estimating a good regularization parameter for linear problems, the situation is much more difficult for nonlinear problems. Moreover, the rate of convergence of the linear and nonlinear terms may be quite different, and this general Gauss-Newton approach does not exploit this fact. Convergence difficulties such as these have been documented in the case of blind deconvolution; see for example Biggs [3].

2.2. Decoupled Problem

Probably the simplest idea to solve the nonlinear least squares problem is to decouple it into two problems, one involving $\mathbf{f}_k$ and the other involving $\mathbf{p}_k$. Specifically, the approach would have the form:

Coordinate Descent Iterative Algorithm

choose initial $\mathbf{p}_0$

for $k = 0, 1, 2, \ldots$

- choose λ_k and solve the linear problem:

$$\mathbf{f}_k = \arg\min_{\mathbf{f}} \|\mathbf{H}(\mathbf{p}_k)\mathbf{f} - \mathbf{g}\|_2^2 + \lambda_k^2\|\mathbf{f}\|_2^2$$

- solve the nonlinear problem:

$$\mathbf{p}_{k+1} = \arg\min_{\mathbf{p}} \|\mathbf{H}(\mathbf{p})\mathbf{f}_k - \mathbf{g}\|_2^2 + \lambda_k^2\|\mathbf{f}_k\|_2^2$$

- stop when objectives are minimized

end

The advantage of this approach is that there are many well known approaches to solve the linear problem, including methods to determine λ. The nonlinear problem involving $\mathbf{p}$ requires using another iterative method, such as the Gauss-Newton method. However, there are significantly fewer parameters than in the fully coupled approach discussed in the previous subsection. Thus, a Gauss-Newton method to update $\mathbf{p}_{k+1}$ at each iteration is significantly more computationally tractable.

The disadvantage to this approach, which is known in the optimization literature as coordinate descent, is that it is not clear what are the practical convergence properties of the method. As mentioned in the previous subsection, the rate of convergence of the linear and nonlinear terms may be quite different [3]. Moreover, if the method does converge, it will typically be very slow, especially for tightly coupled variables [35].

2.3. Variable Projection Method

The variable projection method [10, 11, 22, 36, 40] exploits structure in the nonlinear least squares problem (3). The approach exploits the fact that $\phi(\mathbf{f}, \mathbf{p})$ is linear in $\mathbf{f}$, and that $\mathbf{p}$ contains relatively few parameters compared to $\mathbf{f}$. However, rather than explicitly separating variables $\mathbf{f}$ and $\mathbf{p}$ as in coordinate descent, variable projection implicitly eliminates the linear parameters $\mathbf{f}$, obtaining a reduced cost functional that depends only on $\mathbf{p}$. We then apply a Gauss-Newton method to the reduced cost functional. Specifically, consider

$$\psi(\mathbf{p}) \equiv \phi(\mathbf{f}(\mathbf{p}), \mathbf{p})$$

where $\mathbf{f}(\mathbf{p})$ is a solution of

$$\min_{\mathbf{f}} \phi(\mathbf{f}, \mathbf{p}) = \min_{\mathbf{f}} \left\| \begin{bmatrix} \mathbf{H}(\mathbf{p}) \\ \lambda \mathbf{I} \end{bmatrix} \mathbf{f} - \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix} \right\|_2^2 . \tag{5}$$

To use the Gauss-Newton algorithm to minimize the reduced cost functional $\psi(\mathbf{p})$, we need to compute $\psi'(\mathbf{p})$. Note that because $\mathbf{f}$ solves (5), it follows that $\phi_{\mathbf{f}} = 0$, and thus

$$\psi'(\mathbf{p}) = \frac{d\mathbf{f}}{d\mathbf{p}} \phi_{\mathbf{f}} + \phi_{\mathbf{p}} = \phi_{\mathbf{p}} = \boldsymbol{\rho}_{\mathbf{p}}^T \boldsymbol{\rho} ,$$

where the Jacobian of the reduced cost functional is given by $\mathbf{J}_\psi = \boldsymbol{\rho}_{\mathbf{p}} = \dfrac{\partial(\mathbf{H}(\mathbf{p})\mathbf{f})}{\partial \mathbf{p}}$. Thus, a Gauss-Newton method applied to the reduced cost functional has the basic form:

Variable Projection Gauss-Newton Algorithm

choose initial $\mathbf{p}_0$

for $k = 0, 1, 2, \ldots$

- choose λ_k
- $\mathbf{f}_k = \arg\min_{\mathbf{f}} \left\| \begin{bmatrix} \mathbf{H}(\mathbf{p}_k) \\ \lambda_k \mathbf{I} \end{bmatrix} \mathbf{f} - \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix} \right\|_2$
- $\mathbf{r}_k = \mathbf{g} - \mathbf{H}(\mathbf{p}_k)\, \mathbf{f}_k$
- $\mathbf{d}_k = \arg\min_{\mathbf{d}} \|\mathbf{J}_\psi \mathbf{d} - \mathbf{r}_k\|_2$
- determine step length α_k
- $\mathbf{p}_{k+1} = \mathbf{p}_k + \alpha_k \mathbf{d}_k$

end

Although computing $\mathbf{J}_\psi$ is nontrivial, it is often much more tractable than constructing $\mathbf{J}_\phi$. In addition, the problem of variable convergence rates for the two sets of parameters, $\mathbf{f}$ and $\mathbf{p}$, has been eliminated. Another big advantage of using the variable projection method for large scale inverse problems is that we can use standard approaches to solve the regularized least squares problem at each iteration. Good methods for estimating regularization parameters, such as generalized cross validation, have been well studied, and good preconditioners have been developed for these types of linear inverse problems.

We remark that there are many options to solve the linear subsystem for $\mathbf{f}_k$, ranging from a simple Wiener filter, to more sophisticated iterative solvers. Iterative solvers are essential for spatially variant problems. But this general variable projection Gauss-Newton setup provides many options, depending on the structure of the blurring operator and the PSF. Consider, for example, that in the computation of $\mathbf{f}_k$ we can replace $\lambda_k \mathbf{I}$ with $\lambda_k \mathbf{L}$, where $\mathbf{L}$ provides additional (e.g., statistical or edge) information about the data, and we can easily include prior knowledge about the object by replacing the $\mathbf{0}$ vector with a (weighted) target object, $\lambda_k \mathbf{f}_*$.

Finally, we mention that this approach significantly reduces the parameter search space to one that involves only the nonlinear parameters $\mathbf{p}$. Thus the challenge of finding the "correct" local minimum should be more tractable.

3. LOCAL MINIMUM TRAP

A difficult challenge for MFBD algorithms is that the objective function typically has many local minima, and, moreover, the global minimum may not correspond to an optimally reconstructed image. We illustrate this behavior with a simple one-dimensional deconvolution problem involving a gamma ray spectra and a Gaussian convolution kernel. Figure 1 shows objective functions for this problem, for one and two frame data. Also shown in the plots are the true parameters $\mathbf{p}$ (denoted by the black dot) and the solution computed by the variable projection Gauss-Newton method (magenta dot).

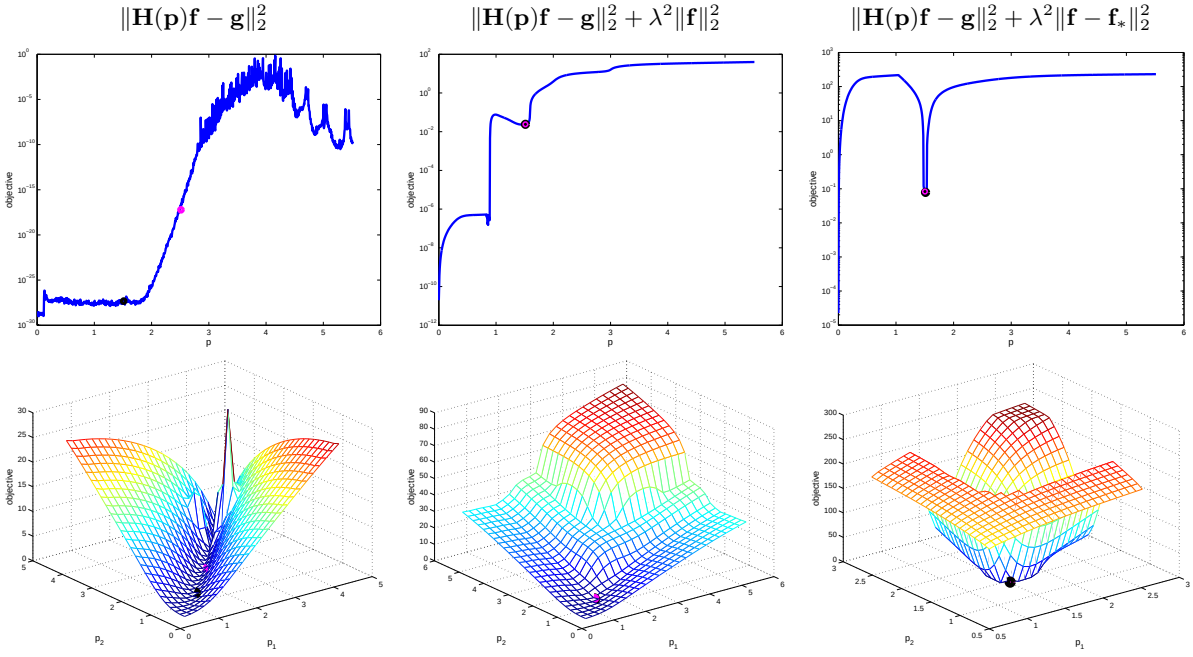

Figure 1. Three different objective functions for single frame blind deconvolution (top row) and two-frame blind deconvolution (bottom row).

We see from this figure that inclusion of more information, including multiple frames, makes the objective function more amenable to determining local optimal solutions. We also see that with just the inclusion of some regularization, the variable projection Gauss-Newton method is very effective at finding the correct local minimum. In the next section we provide further illustration of the effectiveness of the optimization on more realistic image data. However, it is important to note that it is not always the case that the global minimum provides the best reconstructed image. Indeed, it is often the case

that the global minimum occurs when $\mathbf{H}(\mathbf{p}) = \mathbf{I}$ (that is, the PSF is a delta function), and the resulting "restored" image is simply the given blurred data.

4. NUMERICAL RESULTS

In order to illustrate the performance of the variable projection Gauss-Newton method for MFBD, we observe some reconstructions of $\mathbf{f}_{\text{true}}$ in the general discrete model described by (1). In this example we assume the presence of a general Gaussian blur, whose PSF is 256×256 with entries of the form

$$p_{ij} = \exp\left(\frac{-(i-k)^2 s_2^2 - (j-l)^2 s_1^2 + 2(i-k)(j-l)s_3^2}{2s_1^2 s_2^2 - 2s_3^4}\right)$$

and centered at (k, l) (location of point source). To generate simulated MFBD data, we construct three PSFs of this form using three different values for s_1, s_2, and s_3. Specifically,

True PSF parameters

	Frame 1	Frame 2	Frame 3
s_1	6.0516	5.4016	5.7347
s_2	5.8419	4.3802	6.8369
s_3	2.2319	2.1562	2.7385

Simulated blurred image data was generated convolving the PSFs constructed from these sets of parameters with a known true image, and then adding 1% white noise. The resulting simulated observed image frames are shown in Figure 2.

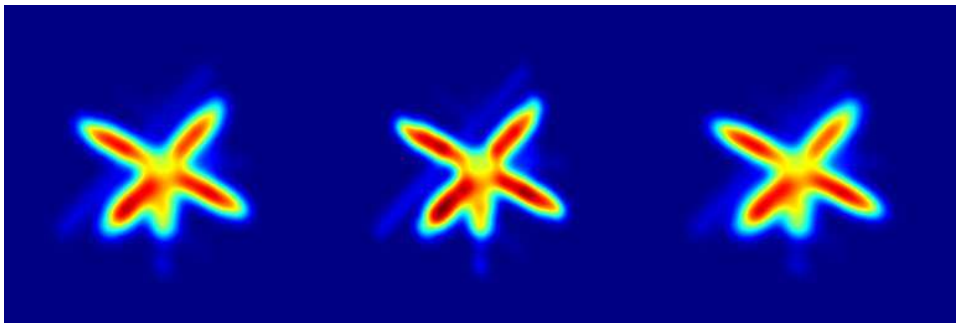

Figure 2. Simulated MFBD data.

To apply the variable projection Gauss-Newton algorithm to this example we adapt the general discrete model of (1) to include the information of the three frames. As described in Section 1, the model for the algorithm then becomes that of equation (2) where $m = 3$. Note, in particular, that since we now have three frames, the PSF parameter vector $\mathbf{p}$ will contain a total of nine entries.

Next, we need to find the Jacobian $\mathbf{J}_\psi$ required for the variable projection Gauss-Newton method. $\mathbf{J}_\psi$ can be constructed analytically using the chain rule as :

$$\mathbf{J}_\psi = \frac{\partial}{\partial \mathbf{p}} \{ \mathbf{H}(\mathbf{P}(\mathbf{p}))\mathbf{f} \} = \frac{\partial}{\partial \mathbf{P}} \{ \mathbf{H}(\mathbf{P}(\mathbf{p}))\mathbf{f} \} \cdot \frac{\partial}{\partial \mathbf{p}} \{ \mathbf{P}(\mathbf{p}) \} = \mathbf{H}(\mathbf{F}) \cdot \frac{\partial}{\partial \mathbf{p}} \{ \mathbf{P}(\mathbf{p}) \}$$

where $\mathbf{f} = \text{vec}(\mathbf{F})$. It should be noted that, for this example, the use of a finite difference approach to approximate $\mathbf{J}_\rho$ can also work very well.

To choose λ_k and to solve the linear subproblem, we use the hybrid Tikhonov-conjugate gradient method, HyBR; see [7] for further details. The step length α_k is chosen using an Armijo rule [35]. One

should try to choose a good initial guess for $\mathbf{p}_0$, but in any case, we recommend over estimating the values $\mathbf{p}_{\text{true}}$.

The results in Figure 3 show the convergence behavior in terms of relative (mean square) error at each iteration of the variable projection Gauss-Newton algorithm for this example. The left plot shows the convergence history of $\mathbf{p}_k$, and the right plot shows the convergence history of $\mathbf{f}_k$. Note that the convergence behavior of both terms is very similar, and that the algorithm is converging to the correct local minimum. Figure 4 shows the reconstructed image after the first variable projection Gauss-Newton iteration (i.e., the initial reconstruction) and the reconstructed image after the last iteration of the algorithm. Note that we have not tried to include any support constraints, prior knowledge, or nonnegativity into the algorithm, but still the results are quite good. Better reconstructions could be obtained if additional prior knowledge is included in the algorithm.

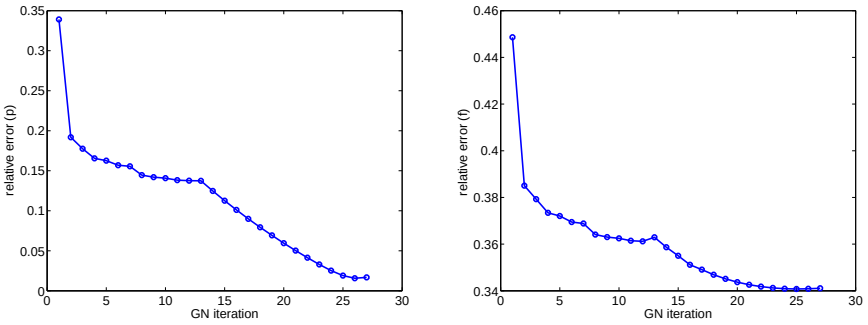

Figure 3. Convergence results for MFBD. Left: Relative error of the estimated PSF parameters. Right: Relative error of the reconstructed image

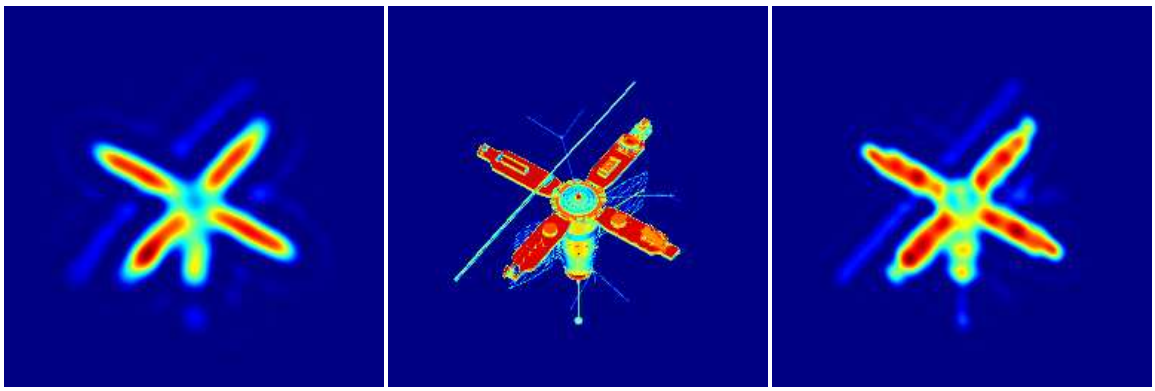

Figure 4. Left: Initial reconstructed image. Middle: True image. Right: Final reconstructed image

5. CONCLUDING REMARKS

In this paper we have provided a general optimization framework based on a variable projection Gauss-Newton method for multi-frame blind deconvolution. The approach is flexible and computationally tractable. In addition, we have illustrated that a hybrid Tikhonov-conjugate gradient method is effective at simultaneously determining the amount of regularization and efficiently solving the linear subproblem in the optimization scheme. Additional work is needed to further improve convergence behavior, such as through preconditioning, as well as to efficiently incorporate bound constraints.

6. ACKNOWLEDGMENTS

This work was supported by the Air Force Office of Scientific Research grant FA9550-09-1-0487.

REFERENCES

1. H.-M. Adorf. Towards HST restoration with space-variant PSF, cosmic rays and other missing data. In R. J. Hanisch and R. L. White, editors, *The Restoration of HST Images and Spectra II*, pages 72–78, 1994.
2. H. Andrews and B. Hunt. *Digital Image Restoration*. Prentice-Hall, Englewood Cliffs, NJ, 1977.
3. D. S. C. Biggs. *Accelerated Iterative Blind Deconvolution*. PhD thesis, University of Auckland, Aukland, New Zealand, 1998.
4. A. F. Boden, D. C. Redding, R. J. Hanisch, and J. Mo. Massively parallel spatially-variant maximum likelihood restoration of Hubble Space Telescope imagery. *J. Opt. Soc. Am. A*, 13:1537–1545, 1996.
5. A. Carasso. Direct blind deconvolution. *SIAM J. Appl. Math.*, 61:1980–2007, 2001.
6. K. Castleman. *Digital Image Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1996.
7. J. Chung, J. G. Nagy, and D. P. O'Leary. A weighted GCV method for Lanczos hybrid regularization. *Elec. Trans. Num. Anal.*, 28:149–167, 2008.
8. M. Faisal, A. D. Lanterman, D. L. Snyder, and R. L. White. Implementation of a modified Richardson-Lucy method for image restoration on a massively parallel computer to compensate for space-variant point spread function of a charge-coupled device camera. *J. Opt. Soc. Am. A*, 12:2593–2603, 1995.
9. D. R. Gerwe, M. M. Johnson, and B. Calef. Local minima analysis of phase diverse phase retrieval using maximum likelihood. In *AMOS Technical Conference, Kihei HI*, September 2008.
10. G. Golub and V. Pereyra. The differentiation of pseudo-inverses and nonlinear least squares whose variables separate. *SIAM J. Numer. Anal.*, 10:413–432, 1973.
11. G. Golub and V. Pereyra. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Problems*, 19:R1–R26, 2003.
12. R. A. Gonsalves. Phase diversity in adaptive optics. *Opt. Eng.*, 21:829–832, 1982.
13. R. Gonzalez and R. Woods. *Digital Image Processing*. Addison-Wesley, Reading, MA, 1992.
14. P. C. Hansen, J. G. Nagy, and D. P. O'Leary. *Deblurring Images: Matrices, Spectra and Filtering*. SIAM, Philadelphia, PA, 2006.
15. E. K. Hege, M. Cheselka, M. Lloyd-Hart, P. Hinz, P. Hoffmann, and S. M. Jefferies. Astronomical results using physically-constrained iterative deconvolution. In *Signal Recovery and Synthesis, OSA Technical Digest Series*, volume 11, page 23, Washington DC, 1998. Optical Society of America.
16. A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1989.
17. S. M. Jefferies, K. J. Schulze, and C. L. Matson. Blind deconvolution with the use of a phase constraint. In *AMOS Technical Conference, Kihei HI*, September 2002.
18. S. M. Jefferies, K. J. Schulze, C. L. Matson, M. Giffin, and J. Okada. Improved blind deconvolution methods for objects imaged through turbid media. In *AMOS Technical Conference, Kihei HI*, September 2002.
19. S. M. Jefferies, K. J. Schulze, C. L. Matson, E. K. Hege, and K. Stoltenberg. Imaging through turbid media: post processing using blind deconvolution. In A. R. Pirich, P. L. Repak, P. S. Idell, and S. R. Czyzak, editors, *Multifrequency Electronic/Photonic Devices and Systems for Dual-Use Applications*, pages 282–289. SPIE, 2001.
20. S. M. Jefferies, K. J. Schulze, C. L. Matson, K. Stoltenberg, and E. K. Hege. Blind deconvolution in optical diffusion tomography. In P. Kervin, L. Bragg, and S. Ryan, editors, *Maui Economic Development Board, Kihei HI*, pages 513–519. AMOS, 2001.
21. S. M. Jefferies, K. J. Schulze, C. L. Matson, K. Stoltenberg, and E. K. Hege. Blind deconvolution in optical diffusion tomography. *Optics Express*, 10:46–53, 2002.
22. L. Kaufman. A variable projection method for solving separable nonlinear least squares problems. *BIT*, 15:49–57, 1975.

23. D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, pages 43–64, May 1996.
24. R. L. Lagendijk and J. Biemond. *Iterative Identification and Restoration of Images*. Kluwer Academic Publishers, Boston/Dordrecht/London, 1991.
25. R. G. Lane. Blind deconvolution of speckle images. *J. Opt. Soc. Am. A*, 9:1508–1514, 1992.
26. N. F. Law and D. T. Nguyen. Multiple frame projection based blind deconvolution. *Electronic Letters*, 31:1734–1735, 1995.
27. M. G. Löfdahl. Multi-frame blind deconvolution with linear equality constraints. In Fiddy & Millane Bones, editor, *Image Reconstruction from Incomplete Data II*, volume 4792-21. SPIE, 2002.
28. C. L. Matson and K. Borelli. Parallelization and automation of a blind deconvolution algorithm. In *HPCMP Users Group Conference*, pages 327–332. IEEE, 2006.
29. C. L. Matson, K. Borelli, S. Jefferies, C. C. Becnker, E. K. Hege, and M. Lloyd-Hart. Fast and optimal multiframe blind deconvolution algorithm for high-resolution ground-based imaging of space objects. *Applied Optics*, 48:A75–A92, 2009.
30. S. R. McNown and B. R. Hunt. Approximate shift-invariance by warping shift-variant systems. In R. J. Hanisch and R. L. White, editors, *The Restoration of HST Images and Spectra II*, pages 181–187, 1994.
31. N. Miura and N. Baba. Segmentation-based multiframe blind deconvolution of solar images. *J. Opt. Soc. Am. A*, 12:1858–1866, 1995.
32. J. G. Nagy and D. P. O'Leary. Fast iterative image restoration with a spatially varying PSF. In F. T. Luk, editor, *Advanced Signal Processing Algorithms, Architectures, and Implementations VII*, volume 3162, pages 388–399. SPIE, 1997.
33. J. G. Nagy and D. P. O'Leary. Restoring images degraded by spatially-variant blur. *SIAM J. Sci. Comput.*, 19:1063–1082, 1998.
34. J. G. Nagy, K. M. Palmer, and L. Perrone. Iterative methods for image deblurring: A Matlab object oriented approach. *Numerical Algorithms*, 36:73–93, 2004.
35. J. Nocedal and S. Wright. *Numerical Optimization*. Springer, New York, 1999.
36. M. R. Osborne. Separable least squares, variable projection, and the Gauss-Newton algorithm. *Elec. Trans. Numer. Anal.*, 28:1–15, 2007.
37. R. G. Paxman, B. J. Thelen, and J. H. Seldin. Phase-diversity correction of turbulence-induced space-variant blur. *Optics Letters*, 19(16):1231–1233, 1994.
38. A. Pruessner and D. P. O'Leary. Blind deconvolution using a regularized structured total least norm algorithm. *SIAM J. Matrix Anal. Appl.*, 24:1018–1037, 2003.
39. M. C. Roggemann and B. Welsh. *Imaging Through Turbulence*. CRC Press, Boca Raton, FL, 1996.
40. A. Ruhe and P. Wedin. Algorithms for separable nonlinear least squares problems. *SIAM Review*, 22:318–337, 1980.
41. T. J. Shulz. Multiframe blind deconvolution of astronomical images. *J. Opt. Soc. Am. A*, 10:1064–1073, 1993.
42. T. J. Shulz, B. E. Stribling, and J. J. Miller. Multiframe blind deconvolution with real data: imagery of the Hubble Space Telescope. *Optics Express*, 1:355–362, 1997.
43. H. J. Trussell and S. Fogel. Identification and restoration of spatially variant motion blurs in sequential images. *IEEE Trans. Image Proc.*, 1:123–126, 1992.