Technical Report

TR-2007-023

Nonlinear Least Squares and Super Resolution

by

Julianne Chung, James G. Nagy

Mathematics and Computer Science

EMORY UNIVERSITY

Nonlinear Least Squares and Super Resolution

Julianne Chung and James G. Nagy

Department of Mathematics and Computer Science
Emory University
Atlanta, GA, 30322, USA

E-mail: jmchung@mathcs.emory.edu, nagy@mathcs.emory.edu

Abstract. Digital super resolution is a term used to describe the inverse problem of reconstructing a high resolution image from a set of known low resolution images, each of which is shifted by subpixel displacements. Simple models assume the subpixel displacements are known, but if the displacements are not known, then nonlinear approaches must be used to jointly find the displacements and the reconstructed high resolution image. Furthermore, regularization is needed to stabilize the inversion process. This paper describes a separable nonlinear least squares formulation and a solution scheme based on the Gauss-Newton method. In addition, an approach is proposed to choose appropriate regularization parameters at each Gauss-Newton iteration.

1. Introduction

Digital super resolution is a term used to describe the inverse problem of reconstructing a high resolution image from a set of known low resolution images, each of which is shifted by subpixel displacements [1, 2]. Suppose we measure m low resolution images, $\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \ldots, \mathbf{b}^{(m)}$, which could be obtained either simultaneously from multiple sensors targeted at the same object, or from a single sensor that captures images of the same object at multiple time instances. In either case, it is assumed that each low resolution image is shifted by subpixel displacements from a particular reference image. These subpixel displacements suggest that each low resolution image contains different information about the same object. The aim of super resolution image reconstruction is to fuse the different image information in order to create one high resolution image. The problem can be modeled as follows. Each low resolution image can be represented as

$$\mathbf{b}^{(i)} = \mathbf{D}\,\mathbf{S}(\mathbf{y}^{(i)})\,\mathbf{x} + \boldsymbol{\eta}^{(i)}$$

where $\boldsymbol{\eta}^{(i)}$ is additive noise, $\mathbf{D}$ is a decimation matrix that transforms a high resolution image into a low resolution image, and $\mathbf{S}$ is a sparse matrix that performs a geometric distortion (e.g., shift, rotation, etc.) of the high resolution image, $\mathbf{x}$. The geometric distortion, and hence $\mathbf{S}$, is defined by the parameter vector, $\mathbf{y}^{(i)}$. (To simplify this discussion we assume no blurring in the recorded images, but such effects can easily be incorporated into the model.) The reconstruction problem is then given by

$$\mathbf{b} = \mathbf{A}(\mathbf{y})\,\mathbf{x} + \boldsymbol{\eta}\,,$$

where

$$\mathbf{b} = \begin{bmatrix} \mathbf{b}^{(1)} \\ \mathbf{b}^{(2)} \\ \vdots \\ \mathbf{b}^{(m)} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \mathbf{y}^{(1)} \\ \mathbf{y}^{(2)} \\ \vdots \\ \mathbf{y}^{(m)} \end{bmatrix}, \quad \mathbf{A}(\mathbf{y}) = \begin{bmatrix} \mathbf{D}\,\mathbf{S}(\mathbf{y}^{(1)}) \\ \mathbf{D}\,\mathbf{S}(\mathbf{y}^{(2)}) \\ \vdots \\ \mathbf{D}\,\mathbf{S}(\mathbf{y}^{(m)}) \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{\eta}^{(1)} \\ \boldsymbol{\eta}^{(2)} \\ \vdots \\ \boldsymbol{\eta}^{(m)} \end{bmatrix},$$

Note that if we assume each low resolution image is only shifted horizontally and vertically, then each $\mathbf{y}^{(i)}$ contains only two values (the horizontal and vertical displacements). If we want to consider more complicated movement (such as rotation), then each $\mathbf{y}^{(i)}$ might contain up to six values that define, for example, general linear affine transformations. In either case, clearly there are significantly fewer parameters defining $\mathbf{y}$ than the number of pixel values defining $\mathbf{x}$.

The computational problem requires estimating the parameters $\mathbf{y}$ that define $\mathbf{A}(\mathbf{y})$ as well as the unknown image $\mathbf{x}$. In this paper we formulate the reconstruction process as a nonlinear least squares problem,

$$\min_{\mathbf{x},\mathbf{y}} \phi(\mathbf{x},\mathbf{y}) = \min_{\mathbf{x},\mathbf{y}} \|\mathbf{A}(\mathbf{y})\,\mathbf{x} - \mathbf{b}\|_2^2 \,, \tag{1}$$

which can be solved using a Gauss-Newton approach [3]. We can reduce the computational effort by observing: (i) the variables $\mathbf{x}$ and $\mathbf{y}$ separate; (ii) the problem is linear in $\mathbf{x}$; (iii) there are significantly fewer parameters defining $\mathbf{y}$ than there are defining $\mathbf{x}$. The *variable projection method* [4] exploits these observations by implicitly separating the variables $\mathbf{x}$ and $\mathbf{y}$, and then uses a Gauss-Newton method to minimize a reduced cost functional that depends only on the nonlinear parameters $\mathbf{y}$.

Each iteration of the Gauss-Newton method requires computing the Jacobian of the reduced cost functional. Computation of the Jacobian is nontrivial and involves solving a least squares problem, in which regularization must be incorporated to stabilize the inversion process. In [5] we showed how to efficiently compute the Jacobian in the case where subpixel displacements are defined by general affine transformations. However, in that paper we needed to select, *a priori*, regularization parameters. This approach has two potential disadvantages. First, heuristics must be used when choosing the regularization parameters. Second, the regularization parameters remain fixed for each Gauss-Newton iteration.

In this paper we use Tikhonov regularization to solve the linear subproblems at each Gauss-Newton iteration. However, instead of specifying *a priori* regularization parameters, we use a Lanczos hybrid bidiagonalization regularization (HyBR) method. HyBR is an efficient iterative method that computes a Tikhonov regularized solution for linear inverse problems, and, more importantly, it can automatically choose regularization parameters based on the data [6]. This paper is outlined as follows. In Section 2 we discuss the separable nonlinear least squares formulation we use to solve the optimization problem (1), and in Section 3 we describe the HyBR algorithm. Numerical examples are given in Section 4, and conclusions are summarized in Section 5.

2. Separable Nonlinear Least Squares Formulation

The nonlinear least squares problem (1) can be solved using a Gauss-Newton algorithm [7, 3, 8]. To do this, we need to compute $\phi' = \mathbf{J}_\phi^T \mathbf{f}$, where $\mathbf{f}(\mathbf{x},\mathbf{y}) = \mathbf{A}(\mathbf{y})\,\mathbf{x} - \mathbf{b}$, and $\mathbf{J}_\phi = \begin{bmatrix} \mathbf{f}_x & \mathbf{f}_y \end{bmatrix}$. Setting $\mathbf{r} = -\mathbf{f} = \mathbf{b} - \mathbf{A}(\mathbf{y})\,\mathbf{x}$, each Gauss-Newton iteration requires solving a least squares problem of the form

$$\min_{\mathbf{d}} \|\mathbf{J}_\phi \mathbf{d} - \mathbf{r}\|_2 \,.$$

Because the super resolution problem is an ill-posed inverse problem, we need to incorporate regularization. In this paper we consider Tikhonov regularization. To summarize, then, a Gauss-Newton method applied to (1) has the basic form:

General Gauss-Newton Algorithm

choose initial $\mathbf{z}_0 = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{bmatrix}$

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

$\quad \mathbf{r}_k = \mathbf{b} - \mathbf{A}(\mathbf{y}_k)\,\mathbf{x}_k$

$\quad \mathbf{d}_k = \arg\min_{\mathbf{d}} \left\| \begin{bmatrix} \mathbf{J}_\phi \\ \alpha \mathbf{I} \end{bmatrix} \mathbf{d} - \begin{bmatrix} \mathbf{r}_k \\ \mathbf{0} \end{bmatrix} \right\|_2$

$\quad \mathbf{z}_{k+1} = \mathbf{z}_k + \mathbf{d}_k$

end

This general Gauss-Newton approach has some disadvantages. First, it does not take algorithmic advantage of the fact that the problem is strongly convex in $\mathbf{x}$. Thus, the Gauss-Newton algorithm may take very small steps due to the nonlinearity induced by $\mathbf{y}$. Moreover, construction of $\mathbf{J}_\phi$ can be very difficult, and solving the associated least squares problem can be very expensive.

One approach to exploiting structure in the problem is to completely decouple the unknowns $\mathbf{x}$ and $\mathbf{y}$ and use the simple idea of alternating between minimization of the two sets of variables. This approach is referred to as coordinate descent in the optimization literature [3], and previously has been used for super resolution [9, 10, 11]. The advantage of this process is that we can use standard algorithms for constrained least squares problems. The difficulty, as is well known with coordinate descent type methods, is that it is not clear what are the practical convergence properties of the method. Moreover, if the method does converge, it will typically be very slow, especially for tightly coupled variables. For an illustration of this behavior in the context of super resolution, see [5]. Thus alternative approaches should be considered.

Because ϕ is linear in $\mathbf{x}$, and because $\mathbf{y}$ contains relatively few parameters compared to $\mathbf{x}$, it can be advantageous to consider separation of variables. However, rather than using an explicit approach, such as coordinate descent, we consider an implicit approach where we mathematically eliminate one set of parameters (specifically, $\mathbf{x}$) to obtain a reduced cost functional that depends only on $\mathbf{y}$ [12, 4, 13, 14]. A Gauss-Newton type method can then be used to solve the resulting optimization problem. Because there are several difficult issues that must be considered when solving large scale inverse problems, we briefly outline the basic idea.

Consider

$$\psi(\mathbf{y}) = \phi(\mathbf{x}(\mathbf{y}), \mathbf{y})$$

where $\mathbf{x}(\mathbf{y})$ is a solution of

$$\min_{\mathbf{x}} \phi(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x}} \|\mathbf{A}(\mathbf{y})\,\mathbf{x} - \mathbf{b}\|_2^2 \,. \tag{2}$$

To use the Gauss-Newton algorithm to minimize the reduced cost functional $\psi(\mathbf{y})$, we need to compute $\psi'(\mathbf{y})$. Note that because $\mathbf{x}$ solves (2), it follows that $\phi_{\mathbf{x}} = 0$, and thus

$$\psi'(\mathbf{y}) = \frac{d\mathbf{x}}{d\mathbf{y}}\,\phi_{\mathbf{x}} + \phi_{\mathbf{y}} = \phi_{\mathbf{y}} = \mathbf{f}_y^T \mathbf{f}\,,$$

where the Jacobian of the reduced cost functional is given by $\mathbf{J}_\psi = \mathbf{f}_y = \mathbf{A}'(\mathbf{y})\mathbf{x}$. Thus, a Gauss-Newton method applied to the reduced cost functional has the basic form:

Reduced Gauss-Newton Algorithm

choose initial $\mathbf{y}_0$
for $k = 0, 1, 2, \ldots$

$$\mathbf{x}_k = \arg\min_{\mathbf{x}} \left\| \begin{bmatrix} \mathbf{A}(\mathbf{y}_k) \\ \alpha\mathbf{I} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix} \right\|_2$$

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}(\mathbf{y}_k)\,\mathbf{x}_k$$

$$\mathbf{d}_k = \arg\min_{\mathbf{d}} \|\mathbf{J}_\psi \mathbf{d} - \mathbf{r}_k\|_2$$

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \mathbf{d}_k$$

end

Although computing $\mathbf{J}_\psi$ is nontrivial, it is often much more tractable than constructing $\mathbf{J}_\phi$. An efficient approach to construct $\mathbf{J}_\psi$ for super resolution, in the case of subpixel displacements described by affine transformations, was presented in [5]. However, in that paper, we specified the regularization parameter α and held it constant for each Gauss-Newton iteration. But this is clearly not an optimal situation because regularization parameters are problem and data dependent. In the next section we describe an approach that simultaneously chooses regularization parameters and solves regularized least squares problems, and which can be used in the reduced Gauss-Newton algorithm.

3. Lanczos Hybrid Bidiagonalization Regularization

In this section we describe an iterative scheme called *Hybrid Bidiagonalization Regularization* (HyBR). This is essentially a conjugate gradient type algorithm that solves a Tikhonov regularized least squares problem

$$\min_{\mathbf{x}} \left\| \begin{bmatrix} \mathbf{A} \\ \alpha\mathbf{I} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix} \right\|_2, \tag{3}$$

but has the advantage that it can refine the choice of the regularization parameter during the iteration process.

The scheme is based on Lanczos bidiagonalization, which can be described as follows. Given an $n \times n$ matrix $\mathbf{A}$ and vector $\mathbf{b}$, the jth-iteration of Lanczos bidiagonalization ($j = 1, \ldots, n$) computes an $n \times (j+1)$ matrix $\mathbf{U}_j$, an $n \times j$ matrix $\mathbf{V}_j$, an $n \times 1$ vector $\mathbf{v}_{j+1}$, and a $(j+1) \times j$ bidiagonal matrix $\mathbf{B}_j$ such that

$$\begin{aligned} \mathbf{A}^T\mathbf{U}_j &= \mathbf{V}_j\mathbf{B}_j^T + \gamma_{j+1}\mathbf{v}_{j+1}\mathbf{e}_{j+1}^T \\ \mathbf{A}\mathbf{V}_j &= \mathbf{U}_j\mathbf{B}_j, \end{aligned}$$

where $\mathbf{e}_{j+1}$ denotes the $(j+1)$st unit vector and $\mathbf{B}_j$ has the form

$$\mathbf{B}_j = \begin{bmatrix} \gamma_1 & & & \\ \beta_2 & \gamma_2 & & \\ & \ddots & \ddots & \\ & & \beta_j & \gamma_j \\ & & & \beta_{j+1} \end{bmatrix}.$$

Matrices $\mathbf{U}_j$ and $\mathbf{V}_j$ have orthonormal columns, and the first column of $\mathbf{U}_j$ is $\mathbf{b}/\|\mathbf{b}\|_2$.

Given these relations, we can approximate the Tikhonov regularization problem by the *projected* LS problem

$$\begin{aligned}\min_{\mathbf{x}\in R(\mathbf{V}_j)} \left\{ \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2 + \alpha^2\|\mathbf{x}\|_2^2 \right\} &= \min_{\hat{\mathbf{x}}} \left\{ \|\mathbf{B}_j\hat{\mathbf{x}} - \mathbf{U}_j^T\mathbf{b}\|_2^2 + \alpha^2\|\hat{\mathbf{x}}\|_2^2 \right\} \\ &= \min_{\hat{\mathbf{x}}} \left\{ \|\mathbf{B}_j\hat{\mathbf{x}} - \beta\mathbf{e}_1\|_2^2 + \alpha^2\|\hat{\mathbf{x}}\|_2^2 \right\} \qquad (4)\end{aligned}$$

where $\beta = \|\mathbf{b}\|_2$, and choose our approximate solution as $\mathbf{x}_j = \mathbf{V}_j\hat{\mathbf{x}}$. Thus we can build an iterative method where at each iteration we solve a regularized least squares problem involving a bidiagonal matrix $\mathbf{B}_j$. Notice that since the dimension of $\mathbf{B}_j$ is very small compared to $\mathbf{A}$, it is much easier to solve for $\hat{\mathbf{x}}$ in equation (4) than it is to solve for $\mathbf{x}$ in the full Tikhonov regularized problem (3). More importantly, when solving equation (4) we can use sophisticated parameter choice methods to find α at each iteration.

This projection based approach was used by Paige and Saunders [15] to develop the LSQR algorithm, which is an implementation of the conjugate gradient method for least squares problems without regularization. The idea of using regularization for the projected problem was proposed independently by O'Leary and Simmons [16] and Björck [17]. A variety of authors have considered computational issues [17, 18, 19]. Recent work has also been done on robust methods for choosing regularization parameters and reliable stopping iterations [6].

Returning to the Gauss-Newton method applied to the reduced cost functional, we simply replace steps that require solving regularized least squares problems with calls to a HyBR algorithm. Thus, the Gauss-Newton method will have the basic form:

Reduced Gauss-Newton with HyBR Algorithm

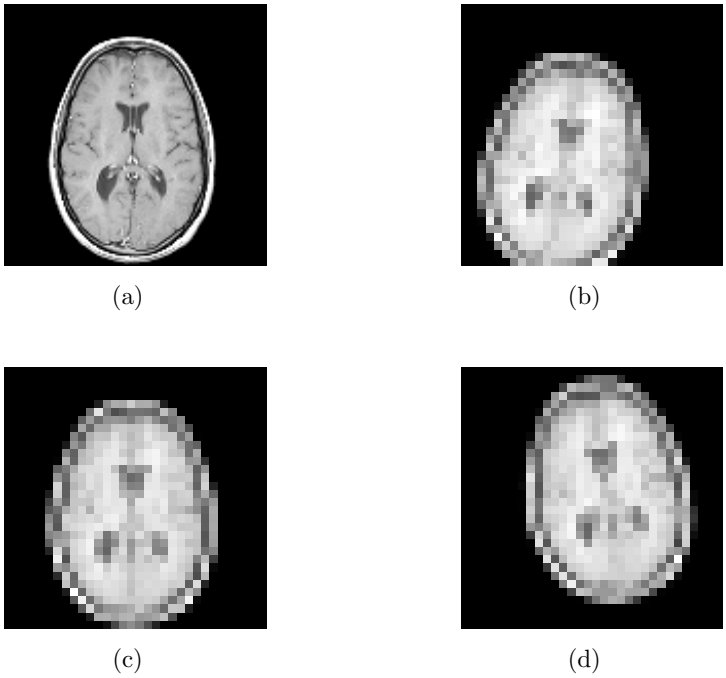
choose initial $\mathbf{y}_0$
for $k = 0, 1, 2, \ldots$
  $\mathbf{x}_k = \text{HyBR}(\mathbf{A}(\mathbf{y}_k), \mathbf{b})$
  $\mathbf{r}_k = \mathbf{b} - \mathbf{A}(\mathbf{y}_k)\,\mathbf{x}_k$
  $\mathbf{d}_k = \arg\min_{\mathbf{d}} \|\mathbf{J}_\psi\mathbf{d} - \mathbf{r}_k\|_2$
  $\mathbf{y}_{k+1} = \mathbf{y}_k + \mathbf{d}_k$
end

It should be noted that it may be necessary to include a line search with the Gauss-Newton method [3]. A careful implementation of HyBR, that incorporates a scheme to choose regularization parameters and reliable stopping criteria is nontrivial. A MATLAB implementation of HyBR is described in [6], and is available at `www.mathcs.emory.edu/~nagy/WGCV`. In the next section we illustrate the effectiveness of using HyBR in the Gauss-Newton algorithm.

4. Numerical Results

In this section, we demonstrate that using HyBR to solve the regularized least squares problem at each Gauss-Newton iteration can be beneficial to super resolution applications. More specifically, we show that one can achieve sufficient objective function and gradient norm decrease, as well as more accurate parameter estimation by using the HyBR method in a reduced Gauss-Newton framework. Furthermore, we illustrate that sufficient reconstructions can be computed without having to make an *a priori* selection of a regularization parameter.

The results presented in this section correspond to a magnetic resonance (MR) image, in which the goal is to recover an approximation of the original high-resolution image with 128^2



(a) (b)

(c) (d)

Figure 1. The high resolution image is shown in (a), and three selected low resolution images are shown in (b-d).

pixels from a set of 32 low-resolution images of 32^2 pixels. Figure 1 contains the true high resolution image along with three sample low-resolution images.

The low resolution images were generated using a sequence of rotations and translations of the original image, followed by a decimation or averaging operator. 1% random noise was then added to each low resolution image. All computations were done in MATLAB, using IEEE double precision arithmetic.

We consider solving problem (1) using the reduced Gauss-Newton algorithm with regularization parameters of $\alpha = 0.0005$, $\alpha = 0.01$, and $\alpha = 0.25$. The motivation for choosing $\alpha = 0.01$ was that it seemed to produce adequate results in numerical experiments. However, we also present results for regularization parameters of 0.0005 and 0.25 to illustrate the potential disadvantages of an inaccurate regularization parameter. An initial guess, $\mathbf{y}_0$, was found by minimizing the displacements of each of the coarse images from a reference image.

To evaluate the success of an algorithm, we examine a variety of measures at each iteration, including the relative objective function, the relative gradient value and the relative error of the rotation and translation parameters. More precisely, this last measure is computed as

$$\Delta \mathbf{y} = \frac{\|\mathbf{y}_k - \mathbf{y}_{\text{true}}\|}{\|\mathbf{y}_{\text{true}}\|}.$$

The results for the reduced Gauss-Newton approach can be found in Table 1.

We then consider the reduced Gauss-Newton with HyBR approach described in Section 3 and present these results in Table 2. The additional column in this table reports the regularization parameter selected by HyBR at each Gauss-Newton iteration. For visual inspection, a specific region of the true and reconstructed images after 6 Gauss-Newton iterations in shown in Figure 2.

Table 1. Convergence of iterations for reduced Gauss-Newton approach.

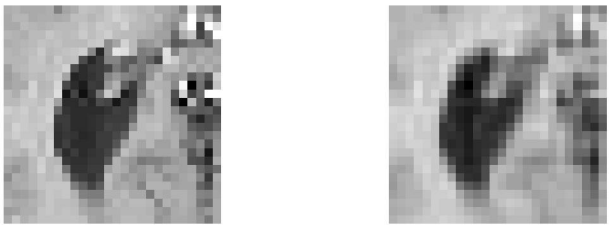
$\alpha = 0.0005$			
iteration	relative objective	relative gradient	$\Delta\mathbf{y}$
0	1.0000	1.0000	0.5784
1	0.6981	0.6298	0.4703
2	0.4876	0.4319	0.3661
3	0.3712	0.2932	0.2830
4	0.3006	0.2325	0.2142
5	0.2508	0.2250	0.1573
6	0.2164	0.1732	0.1153

$\alpha = 0.01$			
iteration	relative objective	relative gradient	$\Delta\mathbf{y}$
0	1.0000	1.0000	0.5784
1	0.8743	0.5670	0.3871
2	0.8195	0.3492	0.2517
3	0.7937	0.2532	0.1598
4	0.7804	0.1899	0.1014
5	0.7745	0.1418	0.0693
6	0.7717	0.1066	0.0531

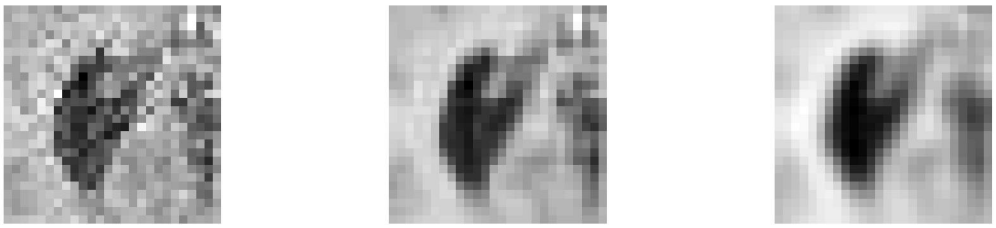
$\alpha = 0.25$			
iteration	relative objective	relative gradient	$\Delta\mathbf{y}$
0	1.0000	1.0000	0.5784
1	0.9894	0.5958	0.3298
2	0.9838	0.5094	0.1791
3	0.9797	0.4780	0.1137
4	0.9761	0.4575	0.1168
5	0.9727	0.4429	0.1401
6	0.9695	0.4313	0.1660

Table 2. Convergence of iterations for reduced Gauss-Newton approach with HyBR.

iteration	relative objective	relative gradient	$\Delta\mathbf{y}$	HyBR computed α
0	1.0000	1.0000	0.5784	0.0125
1	0.6584	0.5342	0.3404	0.0094
2	0.4925	0.3152	0.2012	0.0075
3	0.3774	0.2188	0.1199	0.0060
4	0.3063	0.1528	0.0770	0.0049
5	0.2526	0.1112	0.0563	0.0041
6	0.2271	0.0812	0.0472	0.0037



(a) True (b) GN - HyBR



(c) GN - $\alpha = .0005$ (d) GN - $\alpha = .01$ (e) GN - $\alpha = .25$

Figure 2. Comparison of Reconstructed Images. The images shown are only a specific region of the entire image.

Notice that using HyBR in the Gauss-Newton iteration gave similar or slightly better results than the reduced Gauss-Newton approach, with the great advantage of not having to to make an *a priori* selection of a regularization parameter. An imprecise value of the regularization parameter may result in poor convergence behavior, as demonstrated by $\alpha = 0.0005$ and $\alpha = 0.25$. A regularization parameter that is chosen too small will result in a noisy image such as that in Figure 2(c), while one chosen too large results in a solution that is too smooth, like that in Figure 2(e).

Another observation is the positive effect of allowing α to change at each Gauss-Newton iteration. At iteration 0 of the reduced Gauss-Newton with HyBR algorithm, HyBR selected a regularization parameter of 0.0125, but each additional iteration resulted in smaller and smaller computed parameters. Compared to the case of $\alpha = 0.01$ in Table 1, where the same regularization parameter was used at each iteration, we see that not only was the relative objective and relative gradient descent faster for the HyBR approach, but also the parameter estimation was more accurate. Thus, the reduced Gauss-Newton with HyBR approach can be an effective scheme for jointly estimating displacement parameters and the reconstructed high resolution image in super resolution imaging applications.

5. Concluding Remarks

We described a separable nonlinear least squares formulation of an inverse problem that arises in super resolution image reconstruction. A reduced cost functional was formulated by separation of variables, and a Gauss-Newton approach was used to solve the resulting optimization problem. We described the HyBR algorithm, which is an efficient iterative method to compute Tikhonov regularized solutions of linear inverse problems, and, more importantly, which can automatically choose regularization parameters. We showed that HyBR is an effective approach to solve the linear subproblems that arise at each Gauss-Newton iteration.

Acknowledgments
Research of JC is supported by a DOE Computational Sciences Graduate Fellowship, and the research of JGN is supported by the NSF under grant DMS-05-11454.

References

[1] Park S C, Park M K and Kang M G 2003 *IEEE Signal Processing Magazine* **20** 21–36
[2] Farsiu S, Robinson D, Elad M and Milanfar P 2004 *Int. J. Imaging Syst. Technol.* **14** 47–57
[3] Nocedal J and Wright S 1999 *Numerical Optimization* (New York: Springer)
[4] Golub G and Pereyra V 2003 *Inverse Problems* **19** R1–R26
[5] Chung J, Haber E and Nagy J 2006 *Inverse Problems* **22** 1261–1272
[6] Chung J, Nagy J G and O'Leary D P 2007 A weighted GCV method for Lanczos hybrid regularization Tech. Rep. TR-2007-004-A Mathematics and Computer Science Department, Emory University. To appear in Elec. Trans. Num. Analysis
[7] Dennis J E and Schnabel R B 1996 *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Philadelphia: SIAM)
[8] Ortega J M and Rheinboldt W C 2003 *Iterative Solution of Nonlinear Equations in Several Variables (2nd edition)* (Philadelphia: SIAM)
[9] Cheeseman P, Kanefsky B, Kraft R, Stutz J and Hanson R Dec, 1994 Super-resolved surface reconstruction from multiple images Tech. Rep. Tech. Rep. FIA-94-12 NASA Ames Research Center Moffett Field, CA
[10] Hardie R, Barnard K J and Armstrong E E 1997 *IEEE Trans. Image Proc.* **6** 1621–1633
[11] Tom B C and Katsaggelos A K 1995 *Proc. IEEE Int. Conf. Image Processing* (Washington, DC)
[12] Golub G and Pereyra V 1973 *SIAM J. Numer. Anal.* **10** 413–432
[13] Kaufman L 1975 *BIT* **15** 49–57
[14] Ruhe A and Wedin P 1980 *SIAM Review* **22** 318–337
[15] Paige C C and Saunders M A 1982 *ACM Trans. Math. Soft.* **8** 43–71
[16] O'Leary D P and Simmons J A 1981 *SIAM J. Sci. Stat. Comp.* **2** 474–489
[17] Björck Å 1988 *BIT* **28** 659–670
[18] Kilmer M E and O'Leary D P 2001 *SIAM J. Matrix Anal. Appl.* **22** 1204–1221
[19] Hanke M 2001 *BIT* **41** 1008–1018