2014 Canadian Conference on Computer and Robot Vision

Projected Barzilai-Borwein Method with Infeasible Iterates for Nonnegative Least-Squares Image Deblurring

Kathleen Fraser
Department of Computer Science
University of Toronto
Toronto, Ontario
kfraser@cs.toronto.edu

Dirk V. Arnold
Faculty of Computer Science
Dalhousie University
Halifax, Nova Scotia
dirk@cs.dal.ca

Graham Dellaire
Departments of Pathology
and Biochemistry & Molecular Biology
Dalhousie University
Halifax, Nova Scotia
dellaire@dal.ca

***Abstract*—We present a non-monotonic gradient descent algorithm with infeasible iterates for the nonnegatively constrained least-squares deblurring of images. The skewness of the intensity values of the deblurred image is used to establish a criterion for when to enforce the nonnegativity constraints. The approach is observed on several test images to either perform comparably to or to outperform a non-monotonic gradient descent approach that does not use infeasible iterates, as well as the gradient projected conjugate gradients algorithm. Our approach is distinguished from the latter by lower memory requirements, making it suitable for use with large, three-dimensional images common in medical imaging.**

***Keywords*-Image processing, image restoration; Inverse problems, deconvolution.**

I. Introduction

Image restoration is a common and essential task in astronomy, medical imaging, and digital photography. Image degrading blur may result from sources including optical aberrations in lens systems, atmospheric turbulence in the case of astronomical or satellite images, out of focus contributions of light in widefield microscopy, and motion blur. Removing or reducing the amount of blur is often required before being able to successfully apply image analysis techniques.

Blur is often modelled as being linear [1], and it is assumed that the point spread function (PSF) that mathematically describes the blur is available, either through modelling or measurement. The blurred image is the convolution of the "true" image with the PSF. Deblurring is accomplished through deconvolution.

The deconvolution problem is often formulated as a least-squares problem. Solving that problem directly does not yield useful results as noise that is inevitably present in digital images is amplified in the inverse filtering process. Instead, various regularization approaches that weigh low least-squares values against desirable image properties (e.g., the absence of excessive high frequency content) have been proposed [1]. A commonly used regularization technique is to employ an iterative solver and to terminate after some number of iterations. It is typically observed that image quality initially improves before the deblurred image becomes overly contaminated by noise and its quality deteriorates. The phenomenon is referred to as semi-convergence [2].

Iterative solvers that have been applied to image deblurring include monotonic and non-monotonic gradient descent algorithms as well as conjugate gradient approaches. Ill-conditioning of the deblurring problem usually causes monotonic gradient descent approaches, such as the steepest descent method, to converge prohibitively slowly unless a good preconditioner can be found. Non-monotonic gradient descent approaches, such as the Barzilai-Borwein method [3], offer an interesting alternative that is less sensitive to ill-conditioning and the performance of which can without preconditioning be comparable with that of conjugate gradient approaches. Compared with the latter, non-monotonic gradient descent approaches are characterized by virtue of lower memory requirements, making them suitable for the deblurring of large, three-dimensional images common in medical imaging.

It has been found that in images with large dark areas, which are often encountered in astronomy and medical imaging, imposing nonnegativity constraints on pixel intensities can provide a stabilizing effect and result in higher quality solutions [4], [5], [6], [7], [8]. Constrained variants of iterative solvers that have been applied to nonnegatively constrained least-squares deblurring include projected steepest descent, gradient projected conjugate gradients, and constrained variants of non-monotonic gradient descent strategies. To the best of our knowledge, all of the approaches that have been proposed are feasible-iterates algorithms and either project onto the feasible region or limit step sizes in every iteration.

In this paper, we propose a Barzilai-Borwein algorithm with infeasible iterates and apply it to the nonnegatively constrained least-squares deblurring of images. The approach benefits from the frequently observed efficiency of the non-monotonic search while providing the beneficial stabilizing effects of nonnegativity constraints. We experimentally compare the performance of our algorithm with the projected Barzilai-Borwein method without infeasible iterates as well as with the gradient projected conjugate gradients strategy.

978-1-4799-4337-1/14 $31.00 © 2014 IEEE
DOI 10.1109/CRV.2014.33

We find the infeasible-iterates approach to either perform comparably to or to outperform the other algorithms on several test instances.

The remainder of this paper is organized as follows. Section II introduces formalism and notational conventions and briefly discusses related approaches. Section III motivates and describes BBII, a Barzilai-Borwein algorithm with infeasible iterates for the constrained least-squares deblurring of images. Section IV presents an experimental evaluation of BBII using several test images, and it compares its performance with that of other approaches. Section V concludes with a brief discussion.

II. Background

Image degradation is commonly modelled as

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \boldsymbol{\eta}, \tag{1}$$

where for an image consisting of $m \times n$ pixels, column vector $\mathbf{x} \in \mathbb{R}^{mn}$ contains the true image's pixel intensities, assuming periodic boundary conditions matrix $\mathbf{A} \in \mathbb{R}^{(mn)\times(mn)}$ is a block circulant matrix with circulant blocks that implements the blur, $\boldsymbol{\eta} \in \mathbb{R}^{mn}$ represents noise, and $\mathbf{b} \in \mathbb{R}^{mn}$ contains the pixel intensities of the observed, blurred and noisy image [1]. Due to its size, storage of matrix $\mathbf{A}$ as an $(mn) \times (mn)$ array is impractical for all but very small images. However, multiplication with $\mathbf{A}$ as well as with its inverse can be accomplished in the frequency domain by multiplication with the optical transfer function (OTF), which is the discrete Fourier transform of the PSF, with computational cost linear in mn. The dominant computational cost of deblurring algorithms is that of transforming between the spatial and frequency domains by means of the fast Fourier transform (FFT).

The quality of a candidate solution $\mathbf{y}$ to the image restoration problem is commonly quantified by the relative error

$$e(\mathbf{y}) = \frac{\|\mathbf{y} - \mathbf{x}\|}{\|\mathbf{x}\|}, \tag{2}$$

where $\mathbf{x}$ is the "true" image before blurring and noise and $\|\cdot\|$ denotes the Euclidean length. Obviously, the relative error of a solution cannot be computed in practice as the true solution is unknown; however, it is a useful quantity for comparing deblurring algorithms on test instances where optimal solutions are known [4], [9], [7], [10].

Many approaches to image restoration rely on minimization of objective function

$$\phi(\mathbf{y}) = \tfrac{1}{2}\|\mathbf{A}\mathbf{y} - \mathbf{b}\|^2. \tag{3}$$

That is, they strive to compute the image $\mathbf{y}$ that, if subjected to the (known) blur, yields optimal agreement with the observed blurred and noisy image. However, due to the presence of noise, the exact solution $\mathbf{y} = \mathbf{A}^{-1}\mathbf{b}$ is worthless as division by the OTF acts to amplify the noise present in $\mathbf{b}$.

A commonly used regularization approach is to minimize Eq. (3) iteratively and to terminate the minimization once it is observed that further iterations result in an undue amplification of noise [9], [7]. Gradient descent approaches generate sequences of solutions by iterating

$$\mathbf{y}_{t+1} = \mathbf{y}_t - \alpha_t \mathbf{g}_t, \tag{4}$$

where subscripts denote iteration number,

$$\mathbf{g}_t = \nabla\phi(\mathbf{y}_t) = \mathbf{A}^{\mathrm{T}}(\mathbf{A}\mathbf{y}_t - \mathbf{b}) \tag{5}$$

is the gradient, and α_t is the step size. If proceeding using Eq. (4), the gradient can be updated incrementally according to

$$\mathbf{g}_{t+1} = \mathbf{g}_t - \alpha_t \mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{g}_t, \tag{6}$$

effectively reducing the number of FFTs required per iteration. The steepest descent algorithm (SD) uses step size

$$\alpha_t^{\mathrm{SD}} = \frac{\mathbf{g}_t^{\mathrm{T}}\mathbf{g}_t}{\mathbf{g}_t^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{g}_t}, \tag{7}$$

requiring two FFTs per iteration if the incremental gradient update is used. The Barzilai-Borwein method (BB) defines $\Delta\mathbf{y} = \mathbf{y}_t - \mathbf{y}_{t-1}$ and $\Delta\mathbf{g} = \mathbf{g}_t - \mathbf{g}_{t-1}$ and uses either

$$\alpha_t^{\mathrm{BB1}} = \frac{\Delta\mathbf{y}^{\mathrm{T}}\Delta\mathbf{y}}{\Delta\mathbf{y}^{\mathrm{T}}\Delta\mathbf{g}} \tag{8}$$

or

$$\alpha_t^{\mathrm{BB2}} = \frac{\Delta\mathbf{y}^{\mathrm{T}}\Delta\mathbf{g}}{\Delta\mathbf{g}^{\mathrm{T}}\Delta\mathbf{g}} \tag{9}$$

as the step size. Fletcher [11] states that there is some evidence that the two rules are "not all that dissimilar", and only the former is considered in what follows. Eq. (8) can equivalently be written as

$$\alpha_t^{\mathrm{BB1}} = \frac{\mathbf{g}_{t-1}^{\mathrm{T}}\mathbf{g}_{t-1}}{\mathbf{g}_{t-1}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{g}_{t-1}}, \tag{10}$$

and comparison of Eq. (10) with Eq. (7) has resulted in the Barzilai-Borwein method being referred to as a lagged steepest descent approach [12]. A BB iteration using either Eqs. (8) or (9) can be accomplished using a single FFT. Implementing the variant given by Eq. (10) instead requires two FFTs per iteration.

Nagy and Palmer [9] consider the problem of deblurring images and present results that suggest that due to favourable stability properties, preconditioned SD methods may be an attractive alternative to conjugate gradient approaches when solving linear systems arising from the discretization of ill-posed problems. Huang and Ascher [13] apply the BB method to the problem of least-squares image deblurring and find that the number of iterations required to obtain good solutions is often much smaller than in an unpreconditioned SD approach. This is reflected in Fig. 1, which shows that for the case of the commonly employed satellite test image (compare Section IV), zero-mean Gaussian noise, and a blurred signal to noise ratio (BSNR) of 20dB, the non-monotonic BB approach generates solutions of a quality comparable to those generated using SD, but after a much smaller number of iterations. (The best solution generated using SD has a relative error value of 0.383 after 881 FFTs, while the BB approach generates a solution with relative error 0.382 after only 74 FFTs.) The sequence of error values generated using the BB approach exhibits semi-convergence behaviour (i.e., it decreases before increasing) interrupted by frequent spikes where large step sizes result in poor solutions that the approach typically is able to quickly recover from. The reasons for that type of behaviour are poorly understood [11].

Projected gradient descent approaches that enforce non-negativity constraints $\mathbf{y} \geq \mathbf{0}$ employ update rule

$$\mathbf{y}_{t+1} = \mathcal{P}\left[\mathbf{y}_t - \alpha_t \mathbf{g}_t\right], \tag{11}$$

where $\mathcal{P}[\cdot]$ denotes projection onto the feasible region. Using Eq. (11) with step size rule Eq. (7) yields the projected steepest descent method (pSD). The pSD method requires three FFTs in those iterations where an initially infeasible solution needs to be projected, as projection precludes the incremental gradient update from Eq. (6). Using Eq. (11) with Eqs. (8), (9), or (10) yields variants of the projected Barzilai-Borwein method (pBB), which require two FFTs per iteration in the former cases and three FFTs in the latter one in those steps where the solution needs to be projected.

Dai and Fletcher [14] study pBB methods with update rules Eqs. (8) and (9) as well as several further variants for large-scale box-constrained quadratic programming. They show that global convergence can be established if a non-monotone line search is incorporated in the algorithm, and they propose an adaptive line search approach that appears to not lead to a significant deterioration of performance. They also note that failure of the method to converge without a line search is unlikely to be observed in practice. However, the task in [14] is minimization of $\phi(\cdot)$ and thus differs from the one considered here in that the objective in image restoration is to minimize relative error values $e(\cdot)$. Wang and Ma [10] apply a pBB approach with step size rule Eq. (10) and the adaptive Armijo-Goldstein inexact line search by Dai and Fletcher [14] to the image restoration problem and report to observe good performance.

Included in Fig. 1 are curves representing the result of applying pSD as well as pBB with step sizes according to Eq. (8) to the problem of deblurring the satellite image. It can be seen that enforcing nonnegativity constraints allows generating solutions with much smaller relative error values than the best values observed without projection. The pSD method generates solutions with relative error values of 0.332 after 4000 FFTs, and iterating further would result in further improvements. The pBB method, while converging more slowly than BB, generates solutions with superior relative error values much faster then pSD and achieves

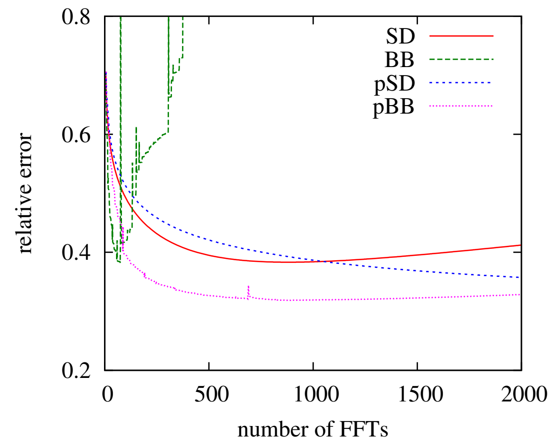

Figure 1. Relative restoration error plotted against the number of FFTs for the steepest descent, Barzilai-Borwein, projected steepest descent, and projected Barzilai-Borwein methods applied to the deblurring of the satellite image with $\text{BSNR} = 20\text{dB}$.

relative error values of 0.319 after 827 FFTs. Not shown in the figure, the pBB approach with step sizes according to Eq. (10) exhibits behaviour that is almost indistinguishable from that of pSD. It is thus not considered in what follows.

III. Algorithm

We present a Barzilai-Borwein based algorithm for the iterative, nonnegatively constrained least-squares deblurring of images that allows infeasible iterates with the goal of combining the fast initial progress of the BB method with the ability of the pBB approach to generate superior solutions. The algorithm determines when the constraints should be enforced by considering the skewness

$$\gamma(\mathbf{y}) = \frac{\frac{1}{mn}\sum_{i=1}^{mn}(y_i - \bar{y})^3}{\left(\frac{1}{mn}\sum_{i=1}^{mn}(y_i - \bar{y})^2\right)^{3/2}} \tag{12}$$

of the pixel intensities in the candidate solution image $\mathbf{y} = (y_i)$. Here, mn is the total number of pixels and $\bar{y} = \sum_{i=1}^{mn} y_i/(mn)$ is the average pixel intensity across the image. Figure 2 shows the relative error as well as the skewness of the deblurred image plotted against the iteration number for the BB method applied to the least-squares deblurring of the satellite test image with $\text{BSNR} = 20\text{dB}$. As seen in Fig. 1, the error curve exhibits the familiar semi-convergence behaviour, interrupted by spikes characteristic of the BB method. The best reconstruction is obtained after 69 iterations. Regularization, e.g., by enforcing nonnegativity constraints, would be required in order to obtain further improvements in solution quality. The skewness curve in Fig. 2 exhibits the opposite behaviour. It increases initially, but later decreases as the solution becomes increasingly contaminated by noise. The same qualitative behaviour can be observed for other test images and other PSFs as well,

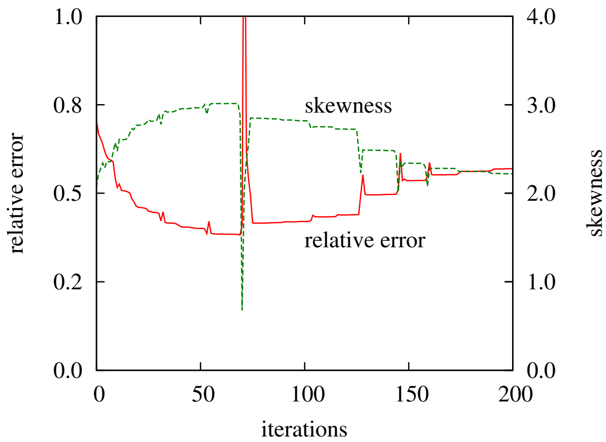

Figure 2. Relative error and skewness plotted against the iteration number for the BB method applied to the least-squares deblurring of the satellite image with $\mathrm{BSNR} = 20\mathrm{dB}$.

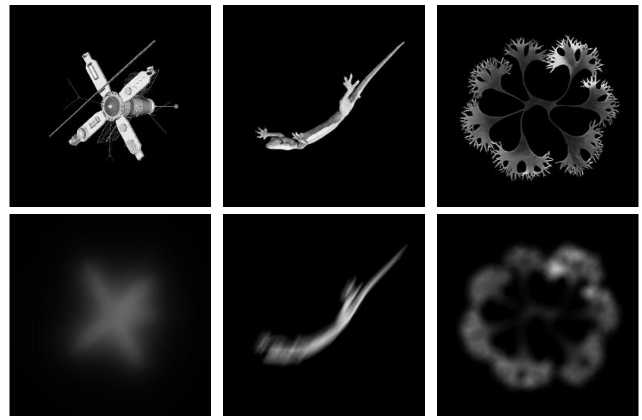

Figure 3. Satellite, gecko, and Irish moss test images before and after blurring.

suggesting that nonnegativity constraints should be enforced when a systematic decrease in skewness of the pixel intensities is observed. In contrast to the relative error value, which is not observable in practice, the skewness can easily be computed.

Our Barzilai-Borwein approach with infeasible iterates (BBII) performs either BB or pBB steps with step sizes according to Eq. (8). It chooses to project and thereby enforce nonnegativity constraints if and only if in an iteration all three of the following conditions hold:

1) The solution that has been generated in the gradient descent step described by Eq. (4) is infeasible.
2) Its skewness is lower than the skewness observed in the previous iteration.
3) No more than two increases in skewness have been observed in the past eight iterations.

The third condition protects against noise in the skewness and prevents the enforcement of the constraints solely due to a spike in the step size. The parameters in it have been obtained with a small amount of experimentation and are not critical to the functioning of the algorithm.

IV. Evaluation

We evaluate BBII using three test image/PSF pairs:

- the commonly used satellite image with a PSF modelling atmospheric turbulence; the pair is available as part of the *RestoreTools* package[1]
- an image of a gecko gliding in a wind tunnel, subject to motion blur generated by linear motion by 24 pixels with an angle of 60°; the original image is courtesy of T. Libby/UC Berkeley
- an image of Irish moss (*chondrus crispus*), with a disk shaped PSF of radius 8; the original image is courtesy of A. Otteson/Adjunct Assistant Professor, Dept. of Plant Sciences and Landscape Architecture, UMD College Park.

Each test image is monochrome, consists of 256×256 pixels, and has a black background. Test images before and after blurring are shown in Fig. 3. For each case we create three test instances by adding Gaussian white noise with blurred signal to noise ratios of $20\mathrm{dB}$, $30\mathrm{dB}$, and $40\mathrm{dB}$. To each of the nine test instances, we apply BBII as described in Section III, BB, pBB with step size rule Eq. (8), and the gradient projected conjugate gradients (GPCG) approach by Moré and Toraldo [15] using an implementation based on code by Bardsley that is available through [16]. All runs are limited to 400 FFTs, and all algorithms are implemented in *Matlab*.

Traces of all runs are shown in Fig. 4; the best solution generated by BBII for each test instance is shown in Fig. 5. It can be seen that BBII either performs comparably to or outperforms both pBB and GPCG for all of the test instances. The advantage of BBII over the other approaches is particularly pronounced for the satellite test image as well as for the lower noise cases of the Irish moss test image. Comparing with the BB curves, it appears that BBII begins to enforce the nonnegativity constraints in the vicinity of the point where the quality of the BB solutions begins to deteriorate, and that the "late" enforcement of the constraints compared to pBB does not negatively affect the quality of the solutions generated.

V. Concluding Remarks

We have presented a Barzilai-Borwein based non-monotone gradient descent method with infeasible iterates for the nonnegative least-squares deblurring of images. The algorithm enforces the nonnegativity constraints only when

[1] http://www.mathcs.emory.edu/~nagy/RestoreTools

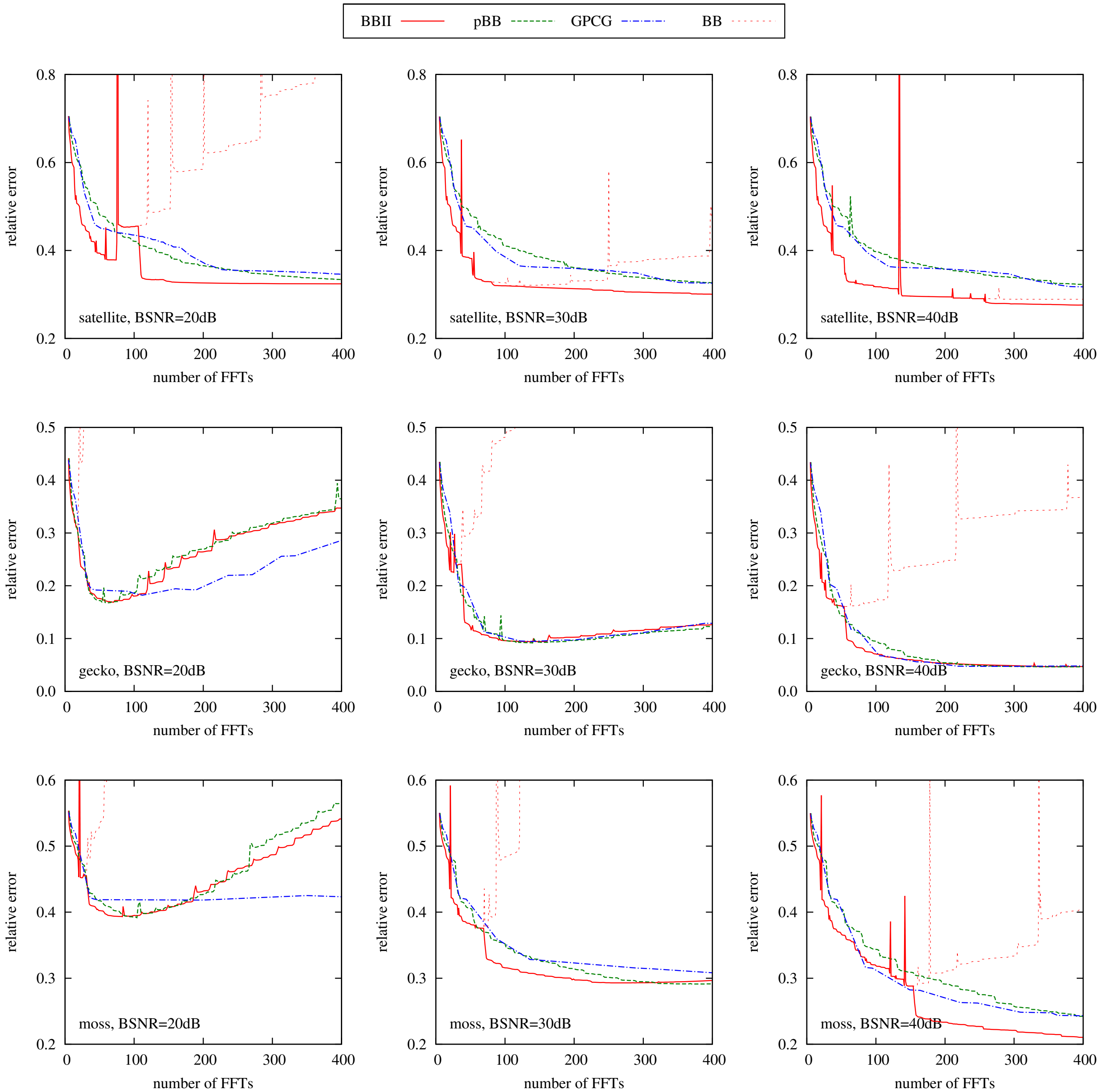

Figure 4. Relative error plotted against the number of FFTs for BBII, BB, pBB, and GPCG. Test images are, from top to bottom, satellite, gecko, and Irish moss. Blurred signal to noise ratios are, from left to right, 20dB, 30dB, and 40dB.

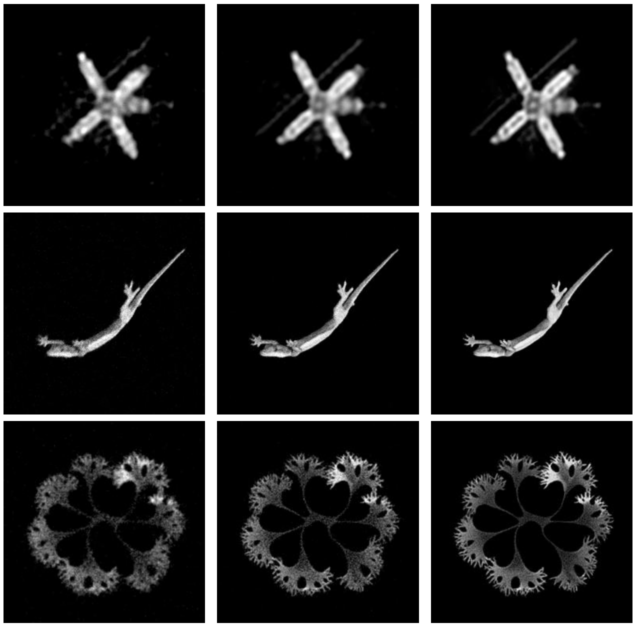

Figure 5. Deblurred images. Shown are the best results obtained using BBII for, from left to right, blurred signal to noise ratios of 20dB, 30dB, and 40dB.

an observed decrease in the skewness of pixel intensities across the image suggests that regularization is required in order to generate further improvement in the solution. Fast minimization of ϕ does not necessarily imply that error values e behave favourably, but the experimental data we have presented suggest that BBII may be a useful alternative to other image restoration approaches, especially in connection with images that require memory efficient algorithms.

VI. Acknowledgement

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Sprectra, and Filtering*. SIAM, 2006.

[2] P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms*. SIAM, 2010.

[3] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.

[4] J. M. Bardsley, J. K. Merikoski, and R. Vio, "The stabilizing properties of nonnegativity constraints in least-squares image reconstruction," *International Journal of Pure and Applied Mathematics*, vol. 43, no. 1, pp. 95–110, 2008.

[5] P. Favati, G. Lotti, O. Menchi, and F. Romani, "Performance analysis of maximum likelihood mathods for regularization problems with nonnegativity constraints," *Inverse Problems*, vol. 26, no. 8, p. 085013, 2010.

[6] M. Hanke, J. G. Nagy, and C. Vogel, "Quasi-Newton approach to nonnegative image restorations," *Linear Algebra and its Applications*, vol. 316, no. 1-3, pp. 223–236, 2000.

[7] J. G. Nagy and Z. Strakos, "Enforcing nonnegativity in image reconstruction algorithms," in *Mathematical Modeling, Estimation, and Imaging*, D. C. Wilson *et al.*, Eds. Society of Photo Optical, 2000, pp. 182–190.

[8] M. Rojas and T. Steihaug, "Large-scale optimization techniques for nonnegative image restorations," in *Proceedings of SPIE, 4791*, 2002, pp. 233–242.

[9] J. G. Nagy and K. M. Palmer, "Steepest descent, CG, and iterative regularization of ill-posed problems," *BIT Numerical Mathematics*, vol. 43, pp. 1003–1017, 2003.

[10] Y. Wang and S. Ma, "Projected Barzilai-Borwein method for large-scale nonnegative image restoration," *Inverse Problems in Science and Engineering*, vol. 15, no. 6, pp. 559–583, 2007.

[11] R. Fletcher, "On the Barzilai-Borwein method," in *Optimization and Control with Applications*, L. Qi *et al.*, Eds. Springer Verlag, 2005, pp. 235–256.

[12] K. van den Doel and U. Ascher, "The chaotic nature of faster gradient descent methods," *Journal of Scientific Computing*, vol. 51, no. 3, pp. 560–581, 2012.

[13] H. Huang and U. Ascher, "Faster gradient descent and the efficient recovery of images," *Vietnam Journal of Mathematics*, 2014, to appear.

[14] Y.-H. Dai and R. Fletcher, "Projected Barzilai-Borwein methods for large-scale box-constrained quadratic programming," *Numerische Mathematik*, vol. 100, no. 1, pp. 21–47, 2005.

[15] J. J. Moré and G. Toraldo, "On the solution of large quadratic programming problems with bound constraints," *SIAM Journal on Optimization*, vol. 1, pp. 93–113, 1991.

[16] C. R. Vogel, *Computational Methods for Inverse Problems*. SIAM, 2002.