Numer. Math. Theor. Meth. Appl. Vol. **2**, No. 4, pp. 427-438
doi: 10.4208/nmtma.2009.m9005s November 2009

Adaptive Parameter Selection for Total Variation Image Deconvolution

You-Wei Wen[1,2,*] and Andy M. Yip[3]

[1] *Faculty of Science, South China Agricultural University, Guangzhou, China.*
[2] *Temasek Laboratories, National University of Singapore, 3, Science Drive 2, 117543, Singapore.*
[3] *Department of Mathematics, National University of Singapore, 3, Science Drive 2, 117543, Singapore.*

Received 13 March 2009; Accepted (in revised version) 11 June 2009

Abstract. In this paper, we propose a discrepancy rule-based method to automatically choose the regularization parameters for total variation image restoration problems. The regularization parameters are adjusted dynamically in each iteration. Numerical results are shown to illustrate the performance of the proposed method.

AMS subject classifications: 65K10, 68U10
Key words: Image restoration, regularization parameter, total variation.

1. Introduction

Digital image deconvolution plays an important part in various areas of applied sciences such as medical and astronomical imaging, and film restoration. The observed image is often degraded by blurring operations and additive noise. The blurring of images is often caused by a relative motion between the camera and the original scene, the defocusing of the lens system, or atmospheric turbulence.

In digital image processing, images are represented by vectors and matrices. In this format, one-dimensional vectors express two-dimensional images. These vectors are formed by stacking the image column by column. Without loss of generality, we assume that the size of the image is $n \times n$, but all discussions can be applied to images of size $n \times m$. Hence the original and observed images f_{true} and g are expressed by the $n^2 \times 1$ vectors $\boldsymbol{f}_{true}$ and $\boldsymbol{g}$ respectively, and their relationship can be expressed as follows

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f}_{true} + \boldsymbol{n}.$$

Here $\boldsymbol{H}$ is a blurring matrix and $\boldsymbol{n}$ is a vector of zero-mean Gaussian white noise with variance σ^2. The main aim of image deconvolution is to recover the image $\boldsymbol{f}$ from the observed image $\boldsymbol{g}$ such that $\boldsymbol{f} \approx \boldsymbol{f}_{true}$.

*Corresponding author. *Email addresses:* `tslwyw@nus.edu.sg` (Y.-W. Wen), `andyyip@nus.edu.sg` (A. M. Yip)

The challenge in image restoration is that it is an ill-posed problem. The simple approach of performing the inverse transformation to the observed image is not feasible since either there doesn't exist an inverse transformation or the inverse transformation is very ill-conditioned; a small perturbation in the observed image can produce a large perturbation in the restored image.

Regularization theory is often used to handle such ill-conditioned problems. One usual approach is to determine the restored image by minimizing the following energy functional

$$\min_{\boldsymbol{f}} \left\|\boldsymbol{H}\boldsymbol{f}-\boldsymbol{g}\right\|_2^2+\beta\mathscr{R}(\boldsymbol{f}), \tag{1.1}$$

where β is called the regularization parameter and $\mathscr{R}$ is the regularization term. Numerous expressions for $\mathscr{R}$ have been used in literature, such as Tikhonov regularization [12, 24], Total variation (TV) regularization [22], Wavelet regularization [4, 9, 11], *etc*. The energy functional is a weighted sum of the two terms. The first term is the data fitting term and the second term is the regularization term which contains some prior information about the original image to alleviate the problem of ill-conditioning. By adjusting the regularization parameter, a compromise is achieved to suppress the noise and preserve the nature of the original image. Usually, the regularization parameter β is determined by trial-and-error method, the generalized cross validation method [12,13], discrepancy principle [17] or the L-curve method [14]. Also, the regularization parameter can be regarded as the Lagrange multiplier of the constrained minimization problem [3]

$$\min_{\boldsymbol{f}} \mathscr{R}(\boldsymbol{f})$$

subject to

$$\left\|\boldsymbol{H}\boldsymbol{f}-\boldsymbol{g}\right\|_2^2=\mathrm{E}[\|\boldsymbol{n}\|_2^2]=\sigma^2 n^2, \tag{1.2}$$

where $\mathrm{E}[\cdot]$ denotes the expectation operator. The variance σ^2 of the noise can be estimated using the median rule [16].

Applying deblurring and denoising independently is a relatively prevalent concept. Its success is due to the facts that the methods are easy to implement, and solving large linear systems is avoided. In [18, 21, 26], the authors proposed a two-step approach to recover the image when a pilot image $\boldsymbol{u}_{pilot}$ is available. This approach can be formulated as the following consecutive minimization problem:

$$\begin{aligned}
\mathscr{S}_\alpha(\boldsymbol{u}_{pilot}) &= \mathrm{argmin}_f \left\|\boldsymbol{H}\boldsymbol{f}-\boldsymbol{g}\right\|_2^2+\alpha\left\|\boldsymbol{R}\left(\boldsymbol{f}-\boldsymbol{u}_{pilot}\right)\right\|_2^2,\\
\mathscr{T}_\beta(\mathscr{S}_\alpha(\boldsymbol{u}_{pilot})) &= \mathrm{argmin}_u \frac{1}{2}\left\|\mathscr{S}_\alpha(\boldsymbol{u}_{pilot})-\boldsymbol{u}\right\|_2^2+\beta\mathscr{R}(\boldsymbol{u}).
\end{aligned}$$

where $\boldsymbol{R}$ is the regularization matrix, α and β are a regularization parameters. Usually, $\boldsymbol{R}$ is the identity matrix, in which a minimum residual on $\boldsymbol{f}$ subject to a noise constraint is sought, or $\boldsymbol{R}$ is the finite difference matrix, in which the smoothness of the restored image is enhanced. The pilot image can be set to $\boldsymbol{u}_{pilot}=0$ or the restored image obtained by other methods.

In [15,25,27], the authors consider the case $\boldsymbol{R}$ being the identity matrix and developed an iterative method. The iterative method is given as follows: starting an initial image $\boldsymbol{u}_0$, which is served as a pilot image, the process can be done iteratively by using an image $\boldsymbol{u}_{k-1}$ as a pilot image, i.e., $\boldsymbol{f}_k = \mathscr{S}_\alpha(\boldsymbol{u}_{k-1})$ and $\boldsymbol{u}_k = \mathscr{T}_\beta(\boldsymbol{f}_k)$. This approach is equivalent to using alternating minimization algorithm to the minimization problem

$$\min_{\boldsymbol{f},\boldsymbol{u}} \mathscr{J}(\boldsymbol{f},\boldsymbol{u}),$$

where

$$\mathscr{J}(\boldsymbol{f},\boldsymbol{u}) = \|\boldsymbol{H}\boldsymbol{f}-\boldsymbol{g}\|_2^2 + \alpha\left\|\boldsymbol{f}-\boldsymbol{u}\right\|_2^2 + 2\beta\mathscr{R}(\boldsymbol{u}). \tag{1.3}$$

In this approach, a trial-and-error method is commonly used to find the optimal regularization parameters α and β such that the restored image $\boldsymbol{u}$ is visually appealing.

In this paper, we consider the model (1.3) rather than (1.1), this model has two parameters that need to be specified. Thus, most existing works on parameter estimation for Tikhonov-type models do not apply to the current setting. There are three components in the deblurred image $\mathscr{S}_\alpha(\boldsymbol{f})$:

$$\mathscr{S}_\alpha(\boldsymbol{f}) = \boldsymbol{f}_{true} + \boldsymbol{n}_f + \alpha\left(\boldsymbol{H}^T\boldsymbol{H}+\alpha\boldsymbol{I}\right)^{-1}\left(\boldsymbol{u}-\boldsymbol{f}_{true}\right),$$

where $\boldsymbol{n}_f = (\boldsymbol{H}^T\boldsymbol{H}+\alpha\boldsymbol{I})^{-1}\boldsymbol{H}^T\boldsymbol{n}$ is the residual noise. The choice of α is very important since the value of α determines how sensitive the solution of $\mathscr{S}_\alpha(\boldsymbol{f})$ is to the noise and how close $\mathscr{S}_\alpha(\boldsymbol{f})$ is to the true image $\boldsymbol{f}_{true}$. Small values of α are desirable to minimize the loss of image features during the first minimization step, but large values of α result in smaller $\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$ and thereby facilitate better estimation of image components via subsequent denoise processing. Our approach is to use two discrepancy rules to determine the two parameters.

The outline of this paper is as follows. In Section 2, we develop a discrepancy rule-based method to calculate the regularization parameters adaptively in each iteration. In Section 3, numerical examples are presented to demonstrate the performance of the proposed method. Finally, concluding remarks are given in Section 4.

2. Adaptive deconvolution method

The main objective of this section is to propose a discrepancy rule-based method to automatically choose the regularization parameters. Our idea is to choose a suitable value of α such that the discrepancy $\boldsymbol{e} = \boldsymbol{H}\boldsymbol{f}-\boldsymbol{g}$ satisfies

$$\|\boldsymbol{e}\|_2 = cn\sigma$$

for some constant c in each iteration. Since $\|\boldsymbol{e}\|_2 \geq \|(\boldsymbol{H}\boldsymbol{H}^\dagger - \boldsymbol{I})\boldsymbol{g}\|_2 = \|(\boldsymbol{H}\boldsymbol{H}^\dagger - I)\boldsymbol{n}\|_2$, where $\boldsymbol{H}^\dagger$ is the pseudo inverse of $\boldsymbol{H}$, the lower bound of c is given by $c \geq \|(\boldsymbol{H}\boldsymbol{H}^\dagger - I)\boldsymbol{n}\|_2/(n\sigma)$. If we follow the Morozov's discrepancy rule, then we have $c=1$. But we found empirically that slightly adjusting the value of c according to the blurred signal-to-noise ratio (BSNR), which is defined as $10\log_{10}(\|g\|_2^2/(n^2\sigma^2))$, can lead to better results. In this paper, we set $c = -0.006 \times BSNR + 1.09$, which is obtained by fitting experimental data with a straight line.

Next we calculate the residual noise $\boldsymbol{n}_f$ in $\boldsymbol{f}$. For each fixed $\boldsymbol{f}$, we propose to set the regularization parameter β such that $\boldsymbol{u}$ satisfies the constraint

$$\left\|\boldsymbol{f}-\boldsymbol{u}\right\|_2^2 = \mathrm{E}[\|\boldsymbol{n}_f\|_2^2].$$

An analytical expression determined for α and β is intractable. Our approach is, instead of using constant regularization parameters α and β, to employ different regularization parameters α_k and β_k in each iteration, which are based on the discrepancy rules.

The resulting algorithm is summarized as follows.

Algorithm 2.1. Framework

input $\boldsymbol{g}$.
estimate σ from $\boldsymbol{g}$ using the median rule.
$c = -0.006 \times BSNR + 1.09$.
$M = cn\sigma$.
initialize $\boldsymbol{u}$.
while stopping criterion is not satisfied
  compute α and $\boldsymbol{f} = \mathscr{S}_\alpha(\boldsymbol{u})$ such that $\left\|\boldsymbol{Hf}-\boldsymbol{g}\right\|_2^2 = M^2$.
  compute $\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$ using (2.4).
  compute β and $\boldsymbol{u} = \mathscr{T}_\beta(\boldsymbol{f})$ such that $\left\|\boldsymbol{f}-\boldsymbol{u}\right\|_2^2 = \mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$.
end
return $\boldsymbol{u}$.

2.1. Solution characteristics of $\boldsymbol{f}$

In the iterative scheme, the first step is to perform the deblurring. The subproblem in the first step can be written as

$$\min_{\boldsymbol{f}} \mathscr{J}(\boldsymbol{f},\boldsymbol{u}) = \min_{\boldsymbol{f}} \mu\left\|\boldsymbol{Hf}-\boldsymbol{g}\right\|_2^2 + \left\|\boldsymbol{f}-\boldsymbol{u}\right\|_2^2, \tag{2.1}$$

where $\mu = 1/\alpha$. We will consider the parameter μ rather than α. This replacement has some advantages, which we will demonstrate below.

The minimization problem (2.1) has a closed-form solution

$$\boldsymbol{f} = \left(\mu\boldsymbol{H}^T\boldsymbol{H}+\boldsymbol{I}\right)^{-1}\left(\mu\boldsymbol{H}^T\boldsymbol{g}+\boldsymbol{u}\right). \tag{2.2}$$

Notice that $\boldsymbol{f}$ is a function of μ and $\boldsymbol{u}$. For convenience, we denote $\boldsymbol{f}(\mu,\boldsymbol{u})$ as $\boldsymbol{f}$ when there is no ambiguity. We remark that the coefficient matrix $\mu\boldsymbol{H}^T\boldsymbol{H}+\boldsymbol{I}$ is always invertible even if $\boldsymbol{H}^T\boldsymbol{H}$ is singular. Under the assumption of periodic boundary condition or symmetric point spread function and Neumann boundary condition, (2.2) is solved using three Fast Fourier Transforms (FFT) in $\mathcal{O}(n^2 \log n)$ operations for an n-by-n restored image, see for instance [19].

The following lemma can be obtained by directly computing the first and second derivatives of $\mathscr{K}(\mu, \boldsymbol{u})$ w.r.t. μ. Similar results can be found in [6, 23].

Lemma 2.1. *Let $\boldsymbol{r} = \boldsymbol{H}\boldsymbol{u} - \boldsymbol{g}$ and $\mathscr{K}(\mu, \boldsymbol{u}) = \|\boldsymbol{H}\boldsymbol{f}(\mu, \boldsymbol{u}) - \boldsymbol{g}\|_2^2$. Then $\mathscr{K}(\mu, \boldsymbol{u})$ is a strictly positive, and monotonically decreasing convex function of μ, and the equation*

$$\mathscr{K}(\mu, \boldsymbol{u}) = M^2 \tag{2.3}$$

has a unique solution $\mu \in (0, \infty)$, for any M satisfying $\|\boldsymbol{r}_0\|_2^2 \leq M^2 \leq \|\boldsymbol{r}\|_2^2$, where $\boldsymbol{r}_0$ denotes the orthogonal projection of $\boldsymbol{r}$ onto the null space of $\boldsymbol{H}\boldsymbol{H}^T$.

We remark the function $\mathscr{K}(1/\alpha, \boldsymbol{u})$ is in general non-convex. This is a reason for using the parameter μ instead of $\alpha = 1/\mu$. A unique solution of μ is guaranteed by the strict convexity of $\mathscr{K}$. Now we turn to computing the regularization parameter μ which satisfies (2.3). Notice that the problem in (2.3) is nonlinear. We apply Newton method to solve the problem (2.3).

We consider periodic boundary condition. As mentioned before, the matrix $\boldsymbol{H}$ can be diagonalized using Fast Fourier Transform. Let $\hat{H}(s,t)$ be the $((s-1)n+t)$-th eigenvalue of $\boldsymbol{H}$ and $\widehat{\cdot}$ be the 2-D discrete Fourier transform defined as

$$\widehat{f}(s,t) = \sum_{j,k} f(j,k) e^{-\frac{2\pi i}{n}(sj+tk)}.$$

Now, we consider the discrepancy $\boldsymbol{e} = \boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}$. Substituting (2.2) into the discrepancy, we have

$$\boldsymbol{e} = \boldsymbol{H}\left(\mu \boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{I}\right)^{-1}\left(\mu \boldsymbol{H}^T\boldsymbol{g} + \boldsymbol{u}\right) - \boldsymbol{g}.$$

Using the identity

$$\mu \boldsymbol{H}\left(\mu \boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{I}\right)^{-1}\boldsymbol{H}^T - \boldsymbol{I} = -\left(\mu \boldsymbol{H}\boldsymbol{H}^T + \boldsymbol{I}\right)^{-1}$$

and noticing that $\boldsymbol{r} = \boldsymbol{H}\boldsymbol{u} - \boldsymbol{g}$, we obtain

$$\boldsymbol{e} = \left(\mu \boldsymbol{H}\boldsymbol{H}^T + \boldsymbol{I}\right)^{-1}\boldsymbol{r}.$$

Therefore, $\mathscr{K}(\mu, \boldsymbol{u})$ and $\frac{\partial}{\partial \mu}\mathscr{K}(\mu, \boldsymbol{u})$ are given by

$$\mathscr{K}(\mu, \boldsymbol{u}) = \|\boldsymbol{e}\|_2^2 = \sum_{s,t} \frac{|\widehat{r}(s,t)|^2}{\left(\mu \left|\widehat{H}(s,t)\right|^2 + 1\right)^2}$$

and

$$\frac{\partial}{\partial \mu}\mathscr{K}(\mu, \boldsymbol{u}) = -2\sum_{s,t} \frac{\left|\widehat{H}(s,t)\widehat{r}(s,t)\right|^2}{\left(\mu \left|\widehat{H}(s,t)\right|^2 + 1\right)^3}.$$

Under the assumption of symmetric point spread function and Neumann boundary condition, the discussion is very similar to the above. While in other cases, by using the Lanczos bidiagonalization and Gauss quadrature approach in [6], we can determine an approximate solution of the regularization parameter μ.

2.2. Analysis of the error

The regularization parameter μ is determined by the constant M. The analysis of the error in $\boldsymbol{f}(\mu,\boldsymbol{u})$ can tell how to determine the constant M.

The error between the deblurred image and the original image is given by

$$\begin{aligned}\boldsymbol{f}-\boldsymbol{f}_{true} &= \left(\mu\boldsymbol{H}^T\boldsymbol{H}+\boldsymbol{I}\right)^{-1}\left(\mu\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{f}_{true}+\boldsymbol{n}\right)+\boldsymbol{u}\right)-\boldsymbol{f}_{true}\\ &= \left(\mu\boldsymbol{H}^T\boldsymbol{H}+\boldsymbol{I}\right)^{-1}\left(\mu\boldsymbol{H}^T\boldsymbol{n}+\boldsymbol{u}-\boldsymbol{f}_{true}\right)\\ &= \boldsymbol{n}_f+\delta\boldsymbol{f},\end{aligned}$$

where

$$\boldsymbol{n}_f=\mu\left(\mu\boldsymbol{H}\boldsymbol{H}^T+\boldsymbol{I}\right)^{-1}\boldsymbol{H}^T\boldsymbol{n},\quad \delta\boldsymbol{f}=\left(\mu\boldsymbol{H}^T\boldsymbol{H}+\boldsymbol{I}\right)^{-1}\left(\boldsymbol{u}-\boldsymbol{f}_{true}\right)$$

denote the residual noise and distortion error in $\boldsymbol{f}$, respectively. Hence $\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$ is given by

$$\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]=\sigma^2\sum_{s,t}\frac{|\hat{H}(s,t)|^2}{\left(|\hat{H}(s,t)|^2+1/\mu\right)^2}. \tag{2.4}$$

And the distortion error is given by

$$\|\delta\boldsymbol{f}\|_2^2=\sum_{s,t}\frac{|\hat{u}(s,t)-\widehat{f}_{true}(s,t)|^2}{\left(\mu|\hat{H}(s,t)|^2+1\right)^2}.$$

Notice that we treat $\boldsymbol{u}$ and $\boldsymbol{f}_{true}$ as fixed constants, therefore the mean squared error (MSE) is given by

$$\mathrm{E}[\|\boldsymbol{f}-\boldsymbol{f}_{true}\|_2^2]=\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]+\|\delta\boldsymbol{f}\|_2^2.$$

Taking the first derivative with respect to μ in $\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$ and $\|\delta\boldsymbol{f}\|_2^2$ yields

$$\frac{\partial}{\partial\mu}\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]=2\sigma^2\sum_{s,t}\frac{\mu|\hat{H}(s,t)|^2}{\left(\mu|\hat{H}(s,t)|^2+1\right)^3}$$

and

$$\frac{\partial}{\partial\mu}\|\delta\boldsymbol{f}\|_2^2=-2\sum_{s,t}\frac{|\hat{H}(s,t)|^2|\hat{u}(s,t)-\hat{f}_{ture}(s,t)|^2}{\left(\mu|\hat{H}(s,t)|^2+1\right)^3}.$$

We conclude that $\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$ is a strictly positive, and monotonically increasing function of μ for $\mu>0$ and that $\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]=0$ when $\mu=0$. Similarly, we conclude that the distortion error is a positive and monotonically decreasing function of μ for $\mu>0$ and that $\|\delta\boldsymbol{f}\|_2^2=\|\boldsymbol{u}-\boldsymbol{f}_{true}\|_2^2$ when $\mu=0$ and $\|\delta\boldsymbol{f}\|_2^2=0$ when $\mu=\infty$. Thus the total error is a coercive function of μ. This suggests that some intermediate value of μ will minimize the total error.

2.3. Total variation penalty

Since TV minimization can preserve sharp edges while reducing the noise and other oscillations in the reconstruction, Rudin, Osher and Fatemi [22] proposed to use TV regularization to solve image denoising problems. In this subsection, we use total variation as a penalty term. The total variation denoising model is formulated as

$$\min_{\boldsymbol{u}} \frac{1}{2} \left\|\boldsymbol{f} - \boldsymbol{u}\right\|_2^2 + \lambda \|\boldsymbol{u}\|_{TV} . \tag{2.5}$$

Here $\|\cdot\|_{TV}$ is the discrete TV norm. While we focus on total variation regularization algorithm, all discussions can be applied to other regularization functions too. The problem in (2.5) can be solved by many TV denoising methods such as Chambolle's projection algorithm [7], semismooth Newton's method [20], multilevel optimization method [8] and graph-based optimization method [10]. In this paper, we employ the Chambolle's projection algorithm in the denoising step for its simplicity and efficiency. This algorithm can be used to compute the minimizer of (2.5) and the parameter β simultaneously by using the constraint that $\|\boldsymbol{f} - \boldsymbol{u}\|_2^2 = \mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$ and setting $\lambda = \mu\beta$.

Let us define the discrete gradient operator $\nabla : \mathbb{R}^{n^2} \to \mathbb{R}^{2n^2}$ by

$$(\nabla \boldsymbol{u})_{r,s} = \Big((\nabla^x u)_{r,s}, (\nabla^y u)_{r,s}\Big)$$

with $(\nabla^x \boldsymbol{u})_{r,s} = u_{r+1,s} - u_{r,s}$ and $(\nabla^y \boldsymbol{u})_{r,s} = u_{r,s+1} - u_{r,s}$ for $r,s = 1,\cdots,n$. We use the reflective boundary condition such that $u_{0,s} = u_{1,s}$, $u_{n+1,s} = u_{n,s}$, $u_{r,0} = u_{r,1}$ and $u_{r,n+1} = u_{r,n}$. Here $u_{r,s}$ refers to the $((r-1)n+s)$th entry of the vector $\boldsymbol{u}$ (it is the (r,s)th pixel location of the image). The discrete TV of $\boldsymbol{f}$ is defined by

$$\|\boldsymbol{u}\|_{TV} = \sum_{1\le r,s\le n} \left|(\nabla \boldsymbol{u})_{r,s}\right| = \sum_{1\le r,s\le n} \sqrt{\left|(\nabla^x \boldsymbol{u})_{r,s}\right|^2 + \left|(\nabla^y \boldsymbol{u})_{r,s}\right|^2}.$$

Here $|\cdot|$ is the Euclidean norm in $\mathbb{R}^2$.

Lemma 2.2 ([7]). *Let $q(\lambda) = \|\boldsymbol{u} - \boldsymbol{f}\|_2$ where $\boldsymbol{u} = \boldsymbol{u}(\lambda)$ is the solution of* (2.5)*, then $q(\lambda)$ maps $[0,\infty)$ onto $[0, \|\boldsymbol{f} - \langle \boldsymbol{f}\rangle\|_2]$. Here $\langle \boldsymbol{f}\rangle$ denotes the average value of the pixel $f_{i,j}$. It is non-decreasing, while the function $\lambda \mapsto q(\lambda)/\lambda$ is non-increasing.*

According to Lemma 2.2, there exists a unique λ such that

$$q(\lambda)^2 = \|\boldsymbol{u} - \boldsymbol{f}\|_2^2 = \mathrm{E}[\|\boldsymbol{n}_f\|_2^2]$$

provided that $q(\lambda) \in [0, \|\boldsymbol{f} - \langle \boldsymbol{f}\rangle\|_2]$. An algorithm to find $\boldsymbol{u}$ was proposed in [7]. It was proven in [7] that such a scheme converges to the desired solution. Here, we restate the algorithm in our context. The tolerance ϵ is set to 10^{-4}.

Algorithm 2.2. Constrained total variation method to find β

ConstrainedTV_Method($\boldsymbol{f},\mu$)

set $j=0$, $\beta_0=1$.
compute $C=\sqrt{\mathrm{E}[\|\boldsymbol{n}_f\|_2^2]}$ by using (2.4).
compute the minimizer $\boldsymbol{u}=\boldsymbol{u}(\mu\beta_0)$ of (2.5).
set $\phi_0=\|\boldsymbol{f}-\boldsymbol{u}\|_2$.
while $|\phi_j-C|\leq\epsilon$
 $j=j+1$.
 $\beta_j=\frac{C}{\phi_{j-1}}\beta_{j-1}$.
 compute the minimizer $\boldsymbol{u}=\boldsymbol{u}(\mu\beta_j)$ of (2.5).
 $\phi_j=\|\boldsymbol{f}-\boldsymbol{u}\|_2$.
end
return $\boldsymbol{u}$ and β_j.

2.4. Relation to other works

Recent years have witnessed a growing interest in the research of image restoration and a lot of fast algorithms have been independently proposed and analyzed. Figueiredo and Nowak studied the expectation minimization framework for wavelet based deconvolution in [11]. Daubechies, Defrise and De Mol studied an optimization transfer approach in [9]. Bect *et al.* studied a projection based l_1 unified variational framework in [2], etc. These algorithms have used iterative methods, in which an image denoising scheme is applied in combination with Tikhonov method during the iterations. The basic idea is to suppress the residual noise between iterations. Though these algorithms used different approaches, they have a common iterative scheme as

$$\boldsymbol{u}_{k+1}=\mathscr{T}_\beta\left(\boldsymbol{u}_k-\lambda_k\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{u}_k-\boldsymbol{g}\right)\right) \tag{2.6}$$

for $k=0,1,\cdots$, starting from an initial guess $\boldsymbol{u}_0$. Here the parameter λ_k is positive and serves as the step size at iteration k.

Now we examine our iterative scheme,

$$\begin{cases}\boldsymbol{f}_{k+1}=\left(\mu\boldsymbol{H}^T\boldsymbol{H}+\boldsymbol{I}\right)^{-1}\left(\mu\boldsymbol{H}^T\boldsymbol{g}+\boldsymbol{u}_k\right),\\ \boldsymbol{u}_{k+1}=\mathscr{T}_{\mu\beta}(\boldsymbol{f}_{k+1}).\end{cases}$$

Introducing a matrix $\boldsymbol{P}=(\boldsymbol{H}^T\boldsymbol{H}+\mu^{-1}\boldsymbol{I})^{-1}$ and noticing that

$$\left(\mu\boldsymbol{H}^T\boldsymbol{H}+\boldsymbol{I}\right)^{-1}\left(\mu\boldsymbol{H}^T\boldsymbol{g}+\boldsymbol{u}_k\right)=\boldsymbol{u}_k-\boldsymbol{P}\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{u}_k-\boldsymbol{g}\right),$$

we can rewrite our iterative scheme as

$$\boldsymbol{u}_{k+1}=\mathscr{T}_{\mu\beta}\left(\boldsymbol{u}_k-\boldsymbol{P}\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{u}_k-\boldsymbol{g}\right)\right). \tag{2.7}$$

We notice that there is a difference between (2.6) and (2.7). (2.7) can be regarded as the version replacing $\lambda_k \boldsymbol{I}$ in (2.6) with $\boldsymbol{P}$, which gives an entirely different algorithm. (2.6) converges to the solution of the minimization problem in (1.1), whereas (2.7) converges to the solution of the minimization problem in (1.3). The matrix $\boldsymbol{P}$ can be regarded as a preconditioner to accelerate the convergence speed.

3. Numerical Results

In this section, numerical results are presented to demonstrate the performance of the proposed algorithm for parameter estimation. The results are compared with those obtained by Bayesian approaches in [1, 5]. Bioucas-Dias *et al.* [5] adopted majorization-minimization approach to estimate the original image and the regularization parameter which is assumed to follow the Jeffeys' distribution. Babacan *et al.* [1] considered the Gamma distribution for the hyperiors of the regularization parameter. For simplicity, we call the approach of Bioucas-Dias *et al.* [5] as BFO and the approach of Babacan *et al.* [1] as BMA.

The Improved Signal-to-Noise Ratio (ISNR) is used to measure the quality of the

Table 1: Gaussian blur with variance 9.

		Lena	Cameraman	Shepp-Logan
BSNR (dB)	Method	ISNR (dB)	ISNR (dB)	ISNR (dB)
20	BFO [5]	2.99	2.21	4.24
	BMA [1]	2.87	1.72	1.85
	Proposed	3.27	2.59	5.45
30	BFO [5]	3.82	3.59	7.21
	BMA [1]	3.87	2.63	4.31
	Proposed	3.83	3.58	7.40
40	BFO [5]	4.41	5.78	10.27
	BMA [1]	4.78	3.39	6.69
	Proposed	5.74	5.90	11.19

Table 2: 9×9 uniform blur.

		Lena	Cameraman	Shepp-Logan
BSNR (dB)	Method	ISNR (dB)	ISNR (dB)	ISNR (dB)
20	BFO [5]	4.05	3.27	6.25
	BMA [1]	3.72	2.42	3.01
	Proposed	4.36	3.80	6.93
30	BFO [5]	5.43	5.69	10.49
	BMA [1]	5.89	5.41	7.77
	Proposed	5.96	5.75	11.28
40	BFO [5]	6.22	8.46	16.39
	BMA [1]	8.42	8.57	13.69
	Proposed	8.34	8.59	17.10

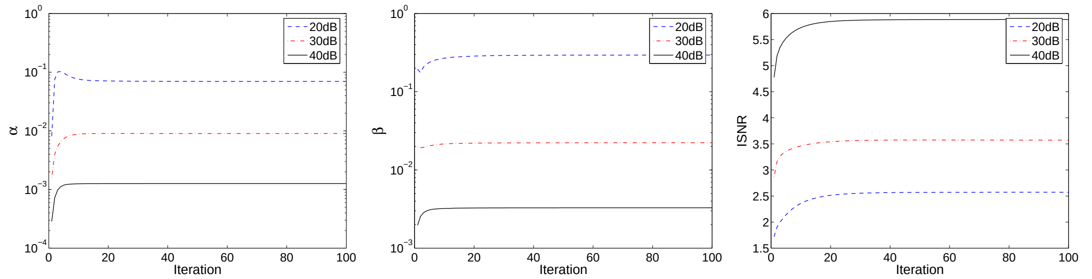

Figure 1: α, β and ISNR vs. the iteration of the *Cameraman* image for different noise levels. The image is degraded by a Gaussian blur with variance 9.

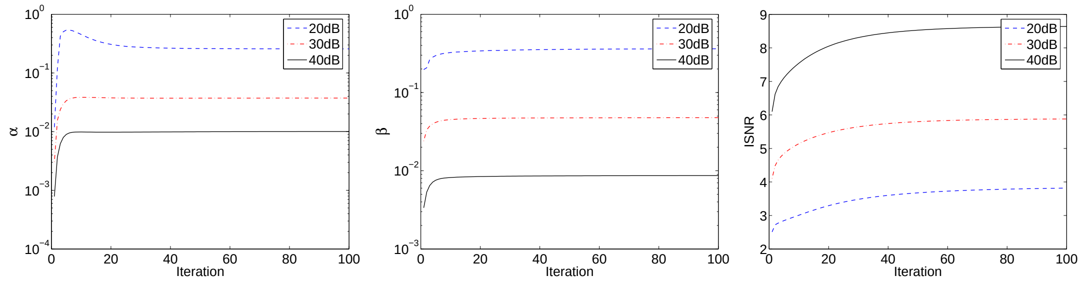

Figure 2: α, β and ISNR vs. the iteration of the *Cameraman* image for different noise levels. The image is blurred by a uniform blur with size 9×9.

restoration results. It is defined as

$$\text{ISNR} = 10 \log_{10} \left(\|\boldsymbol{g} - \boldsymbol{f}_{true}\|_2^2 / \|\boldsymbol{u} - \boldsymbol{f}_{true}\|_2^2 \right),$$

where $\boldsymbol{u}$ is the estimated image, $\boldsymbol{f}_{true}$ is the original clean image. The stopping criterion is set to

$$\|\boldsymbol{u}_{k+1} - \boldsymbol{u}_k\|_2 / \|\boldsymbol{u}_k\|_2 < 1 \times 10^{-4}$$

and the maximum number of iterations is set to 100. We experiment with three noise levels, corresponding to BSNR of 20, 30 and 40dB. The variance σ^2 is estimated using the median rule [16]. The *Lena*, *Cameraman* and *Shepp-Logan phantom* images are used in the tests. The initial image $\boldsymbol{u}_0$ is chosen as $\boldsymbol{u}_0 = 0$.

As mentioned above, the regularization parameter $\alpha = 1/\mu$ is chosen such that $\|\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}\|_2 = cn\sigma$. The smaller the value c, the smaller the distortion error and the larger the expectation of $\|\boldsymbol{n}_f\|_2^2$. Since the smaller the BSNR in the observed image, the larger the variance of the noise, we choose the parameter c according to BSNR. For a smaller BSNR, we choose a larger c. The parameter c is set to $c = -0.006 \times BSNR + 1.09$, which is obtained by fitting experimental data with a linear model.

In the first set of experiments, the images were degraded by a Gaussian blur with variance 9 and white Gaussian noise. The ISNR values are shown in Table 1. When the noise level is set to 30dB, the ISNRs of the restored image obtained by the proposed algorithm and BFO are about the same. For the other levels, the ISNR of the restored images obtained by the proposed algorithm are better than those obtained by the BFO method and the BMA method. The plots of α, β and ISNR vs. the iteration of the *Cameraman* image are shown in Fig. 1. We observe that the regularization parameters stabilize after about 20 iterations.

In the second set of experiments, we restore blurred images degraded by a uniform blur with size 9×9 and white Gaussian noise. The degraded images have a BSNR of 20, 30, 40dB respectively. The ISNR values are shown in Table 2. The plots of α, β and ISNR vs. the iteration of the *Cameraman* image are shown Fig. 2. We observe that the regularization parameters again stabilize after about 20 iterations. Our experimental results show that the ISNRs obtained by the proposed method are competitive to the tested methods.

4. Conclusion

In this paper, we have extended the first author's recent work on decoupling of deblurring and denoising steps in the restoration process. In particular, we have adopted a discrepancy rule-based method to automatically choose the regularization parameters in each iteration. Numerical results were used to demonstrate the performance of proposed method. The ISNRs obtained by the proposed method are competitive with the tested methods.

Acknowledgments This research was supported in part by NSFC Grant No. 60702030 and 10871075 and the wavelets and information processing program under a grant from DSTA, Singapore.

References

[1] S.D. Babacan, R. Molina, and A.K. Katsaggelos. Parameter estimation in TV image restoration using variational distribution approximation. *IEEE Trans. Image Process.*, 17(3):326–339, 2008.

[2] J. Bect, L. Blanc-Feraud, G. Aubert, and A. Chambolle. A l^1-unified variational framework for image restoration. *Lecture Notes in Computer Sciences*, 3024(IV):1–13, 2004.

[3] D. P. Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Computer Science and Applied Mathematics. Academic Press Inc., New York, 1982.

[4] J.M. Bioucas-Dias. Bayesian wavelet-based image deconvolution: A GEM algorithm exploiting a class of heavy-tailed priors. *IEEE Trans. Image Process.*, 15(4):937–951, 2006.

[5] J.M. Bioucas-Dias, M.A.T. Figueiredo, and J.P. Oliveira. Adaptive total variation image deconvolution: A majorization-minimization approach. In *Proceedings of the European Signal Processing Conference (EUSIPCO2006)*, 2006.

[6] D. Calvetti and L. Reichel. Tikhonov regularization of large linear problems. *BIT*, 43(2):263–283, 2003.

[7] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20(1-2):89–97, 2004.

[8] T. Chan and K. Chen. An optimization-based multilevel algorithm for total variation image denoising. *Multiscale Model. Simul.*, 5(2):615–645, 2006.

[9] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57:1413–1457, 2004.

[10] F. de Hoog. A new algorithm for solving Toeplitz systems of equations. *Linear Alg. Appl.*, 88-89:123–138, 1987.

[11] M. Figueiredo and R. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.*, 12(8):906–916, 2003.

[12] N. P. Galatsanos and A. K. Katsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Trans. Image Process.*, 1(3):322–336, 1992.

[13] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.

[14] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.*, 34(4):561–580, 1992.

[15] Y. Huang, M. Ng, and Y. Wen. A fast total variation minimization method for image restoration. *Multiscale Model. Simul.*, 7(2):774–795, 2008.

[16] S. Mallat. *A Wavelet Tour of Signal Processing*. 2nd edition. Academic Press: San Diego, 1999.

[17] V. A. Morozov. *Methods for solving incorrectly posed problems*. Springer-Verlag, New York, 1984. Translated from the Russian by A. B. Aries, Translation edited by Z. Nashed.

[18] R. Neelamani, H. Choi, and R. Baraniuk. ForWaRD: Fourier-Wavelet Regularized Deconvolution for ill-conditioned systems. *IEEE Trans. Signal Process.*, 52:418–433, 2004.

[19] M. Ng, R. Chan, and W. Tang. A fast algorithm for deblurring models with Neumann boundary conditions. *SIAM J. Sci. Comput.*, 21(3):851–866, 2000.

[20] M. Ng, L. Qi, Y. Yang, and Y. Huang. On semismooth Newton's methods for total variation minimization. *J. Math. Imaging Vision*, 27:265–276, 2007.

[21] R. Nowak and M. Thul. Wavelet-vaguelette restoration in photon-limited imaging. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 5:2869–2872, 1998.

[22] S. Osher, L. I. Rudin, and E. Fatem. Nonlinear total variation based noise removal algorithms. *Physics D*, 60:259–268, 1992.

[23] L. Reichel and A. Shyshkov. A new zero-finder for tikhonov regularization. *BIT Numerical Mathematics*, 48:627–643, 2008.

[24] A. Tikhonov. Solution of incorrectly formulated problems and regularization method. *Soviet Math. Dokl*, 4:1035–1038, 1963.

[25] Y. Wen. *Fast Solvers for Toeplitz Systems with Applications to Image Restoration*. PhD thesis, Department of Mathematics, the University of Hong Kong, 2006.

[26] Y. Wen, W. Ching, M. Ng, and H. Liu. A hybrid algorithm for spatial and wavelet domains image restoration. In *Proceedings of SPIE, Visual Communications and Image Processing 2005*, 5960:2004–2011, 2005.

[27] Y. Wen, M. Ng, and W. Ching. Iterative algorithms based on decoupling of deblurring and denoising for image restoration. *SIAM J. Sci. Comput.*, 30(5):2655–2674, 2008.