Selection of regularization parameter in total variation image restoration

Haiyong Liao,[1] Fang Li,[2] and Michael K. Ng[1,*]

[1]*Center for Mathematical Imaging and Vision and Department of Mathematics, Hong Kong Baptist University, Hong Kong, China*
[2]*Department of Mathematics, East China Normal University, Shanghai, China*
[*]*Corresponding author: mng@math.hkbu.edu.hk*

Received April 8, 2009; revised August 13, 2009; accepted September 8, 2009; posted September 15, 2009 (Doc. ID 109847); published October 9, 2009

We consider and study total variation (TV) image restoration. In the literature there are several regularization parameter selection methods for Tikhonov regularization problems (e.g., the discrepancy principle and the generalized cross-validation method). However, to our knowledge, these selection methods have not been applied to TV regularization problems. The main aim of this paper is to develop a fast TV image restoration method with an automatic selection of the regularization parameter scheme to restore blurred and noisy images. The method exploits the generalized cross-validation (GCV) technique to determine inexpensively how much regularization to use in each restoration step. By updating the regularization parameter in each iteration, the restored image can be obtained. Our experimental results for testing different kinds of noise show that the visual quality and SNRs of images restored by the proposed method is promising. We also demonstrate that the method is efficient, as it can restore images of size 256×256 in ≈20 s in the MATLAB computing environment.

OCIS codes: 100.1758, 100.1830, 100.2000, 100.3020.

1. INTRODUCTION

Digital image restoration and reconstruction play an important part in various areas of applied sciences such as medical and astronomical imaging, film restoration, and image and video coding. In this paper, we focus on a common degradation model: an ideal image $f\in\mathbb{R}^{n^2}$ is observed in the presence of a spatially invariant point spread function (PSF) and noise $n\in\mathbb{R}^{n^2}$. Thus the observed image $g\in\mathbb{R}^{n^2}$ is modeled as

$$g=Hf+n. \tag{1}$$

Mathematically, our aim is to determine f given a blurring matrix H generated by a spatially invariant PSF and an observed image g. Traditionally, the standard image restoration methods involve computation in the frequency domain, facilitated by efficient fast Fourier transform (FFT) algorithms. There has been new movement toward more variational approaches [1]. The variational approaches are often designed to process certain desirable geometrical properties.

A. Total Variation Image Restoration

It is well known that restoring an image f is a very ill-conditioned problem. A regularization method should be used in the image restoration process. The main disadvantage of using quadratic regularization functionals such as the H^1 semi-norm of f (in the continuous setting),

$$\int (f_x^2+f_y^2)\mathrm{d}x\mathrm{d}y,$$

is the inability to recover sharp discontinuities (edges) [2]. Mathematically, this is clear because discontinuous functions do not have a bounded H^1 semi-norm. To remedy this, Rudin *et al.* [2] proposed the use of the total variation (TV)

$$\int \sqrt{f_x^2+f_y^2}\mathrm{d}x\mathrm{d}y$$

as a regularization functional. The reason for using this functional is that it measures the jumps of f, even if it is discontinuous. The main advantage of the TV formulation is the ability to preserve edges in the image as a result of the piecewise smooth regularization property of the TV norm.

In this paper, we consider the following TV-based image restoration problem by

$$\min_f \frac{1}{\alpha}\|Hf-g\|+\|f\|_{TV}, \tag{2}$$

where $\|x\|$ is equal to $\|x\|_2^2\equiv\Sigma_i|x_i|^2$ or $\|\cdot\|_1\equiv\Sigma_i|x_i|$. Here α is a positive regularization parameter that measures the trade-off between a good fit and a TV regularization. The discrete version of TV regularization term $\|\cdot\|_{TV}$ is given as follows. The discrete gradient operator $\nabla:\mathbb{R}^{n^2}\to\mathbb{R}^{n^2}$ is defined by

$$(\nabla f)_{j,k}=((\nabla f)_{j,k}^x,(\nabla f)_{j,k}^y)$$

with

$$(\nabla f)_{j,k}^x=\begin{cases} f_{j+1,k}-f_{j,k}, & \text{if } j<n \\ 0, & \text{if } j=n \end{cases},$$

$$(\nabla f)^{v}_{j,k} = \begin{cases} f_{j,k+1} - f_{j,k}, & \text{if } k < n \\ 0, & \text{if } k = n \end{cases},$$

for $j,k=1,\ldots,n$. Here $f_{j,k}$ refers to the $(jn+k)$th entry of the vector f [it is the (j,k)th pixel location of the image]. The discrete total variation of f is defined by

$$\|f\|_{TV} := \sum_{1\leqslant j,k\leqslant n} |(\nabla f)_{j,k}|_2 = \sum_{1\leqslant j,k\leqslant n} \sqrt{|(\nabla f)^{x}_{j,k}|^2 + |(\nabla f)^{y}_{j,k}|^2}.$$

Here $|\cdot|_2$ is the Euclidean norm in $\mathbb{R}^2$.

Numerical results have shown that TV regularization is useful in image restoration; see [2,3]. There are some papers [4–7] that describe why and how TV regularization works. In [4], Bellettini *et al.* analyzed conservation of shape and scale parameters controlling evolution within the TV-flow context. In [6], Dobson and Santosa studied how TV regularization affects frequency distributions to show that the effectiveness of TV image restoration depends on the size and gray-scale intensity of the image geometric feature, and TV regularization is particularly effective for restoring blocky images but not for textured images. In [7], Strong and Chan showed that TV regularization tends to preserve edge locations and that restored intensity change is exactly inversely proportional to local feature scale. In addition, Strong and Chan investigated how the total variation image restoration is affected by the regularization parameter α. Their results include that for smooth radially symmetric function features, function intensity change is inversely proportional to radial distance and directly proportional to α. Based on their results, they can choose the values of regularization parameters to be spatially varying in Eq. (2) for spatially adaptive image restoration.

TV regularization schemes have been extended to the case of vector-valued (e.g., color) images [8], superresolution [9], blind deconvolution [10], and other applications for restoring nonflat image features [11]. We refer readers to [12] for recent developments of TV image restoration applications.

A number of numerical methods have been proposed for solving Eq. (2). These methods include partial-differential-equation-based methods [2,3,13,14], primal-dual methods [15–17], and Newton-type methods [18–20]. Recently, Huang *et al.* [21] proposed and developed a fast TV minimization method for the TV image restoration model. In their paper, an auxiliary variable and an additional quadratic term is added to the objective function (2). In the alternative minimization framework, minimizing the new objective function can be interpreted as a two-stage process, i.e., denoising and blurring. Experimental results in [21] have shown that the quality of images restored by their method is competitive with the quality of those restored by existing TV restoration methods. This approach has also been successfully used for l_1-norm data-fitting term, i.e., $\|Hf-g\|_1$ in [22]. Wang *et al.* [23,24] also proposed and developed the alternating minimization algorithm for deblurring and denoising jointly by solving a TV regularization problem. Their algorithm is derived from the well-known variable-splitting and penalty techniques in optimization. The idea is that at each pixel an auxiliary variable is introduced to transfer its gradient out of nondifferentiable term. Extensive numerical results show that their algorithm performs favorably in comparison with several state-of-the-art algorithms.

B. Selection of Regularization Parameters

It is well known that blurring matrices are in general ill-conditioned and that deblurring algorithms will be extremely sensitive to noise; see, for instance, Gonzalez and Woods [25]. The ill-conditioning of the blurring matrices stems from the wide range of magnitude of their eigenvalues; see Engl *et al.* [26]. Therefore, excess amplification of the noise at small eigenvalues can occur. The method of regularization is used to achieve stability for image restoration problems. In Eq. (2), the restored image is attained by using regularization, which restricts the set of admissible solutions.

There are several possible strategies for choosing regularization parameters, e.g., the discrepancy principle [26,27], the generalized cross-validation method [28,29], and the L-curve method [30,31]. The discrepancy principle is an *a posteriori* strategy for choosing a regularization parameter as a function of the error level (the input error level must be known). The generalized cross-validation method can be applied to determine regularization parameters when the input error level is unknown. The L-curve method is based mainly on the plot of the norm of the regularized solution versus the norm of the corresponding data-fitting residual. The method is used to choose a regularization parameter related to the characteristic L-shaped "corner" of the plot. However, to our knowledge, these regularization-parameter-selection methods have not been developed for TV image restoration problems. Since the TV formulation in Eq. (2) is nonlinear, the generalized cross-validation evaluation formula cannot be derived explicitly. Recently, Lin and Wohlberg [32] extended the method of unbiased predictive risk estimator (UPRE) to the TV regularization problem by considering a quadratic approximation to the TV term. However, they reported the experimental results only on a small problem.

Recently, Babacan *et al.* [33] proposed a variational method for solving deconvolution problems with TV regularization. Their idea is to use a stochastic method approximating *a posteriori* distribution by a product of distributions using Kullback–Leibler divergence. Their resulting method is simple and efficient, and it can update a regularization parameter in TV image restoration in an iterative manner. Molina *et al.* [34] also studied this kind of variational method for blind deconvolution applications.

C. Outline

The main aim of this paper is to solve the problem of regularization parameter selection in Eq. (2) to restore the original image f from an observed blurred and noisy image g in the alternative minimization framework. The method exploits the generalized cross-validation technique to determine how much regularization is used in each restoration step. By updating these regularization parameters in the iterative procedure, the restored image becomes increasingly close to the original image. Our experimental results show that the quality of images re-

stored by the proposed method with no prior knowledge of noise is quite good. We will also demonstrate that the proposed method is also very efficient.

The outline of this paper is as follows. In Section 2, an automatic method for selection of regularization parameters in TV image restoration is presented. In Section 3, numerical examples are given to demonstrate the effectiveness of the proposed method. Finally, concluding remarks are given in Section 4.

2. SELECTION OF REGULARIZATION PARAMETERS

We first present a variable-splitting method and an alternating minimization scheme to solve Eq. (2) [21,22].

When the data-fitting term is measured in l_2-norm instead of Eq. (2), we minimize the following functional:

$$\min_{f,u} \frac{1}{\alpha}\|Hf-g\|_2^2+\frac{\beta}{2}\|u-f\|_2^2+\|u\|_{TV}. \tag{3}$$

Here we add an auxiliary variable u in Eq. (3) and assume that the data-fitting term is measured in the square of the Euclidean norm. We note that when β is sufficiently large, we force f close to u. Thus the minimization problem in Eq. (3) is close to the original problem:

$$\min_{f} \frac{1}{\alpha}\|Hf-g\|_2^2+\|f\|_{TV}. \tag{4}$$

When the data-fitting term is measured in the l_1-norm, we minimize the following functional:

$$\min_{f,u,v} \frac{1}{\alpha}\left(\frac{\beta}{2}\|v-(Hf-g)\|_2^2+\|v\|_1\right)+\left(\frac{\beta}{2}\|u-f\|_2^2+\|u\|_{TV}\right). \tag{5}$$

Here we use two auxiliary variables u and v in the algorithm. In this case, when β is sufficiently large, we force f close to u and $(Hf-g)$ close to v. Thus the minimization problem in Eq. (5) is close to the original problem:

$$\min_{f} \frac{1}{\alpha}\|Hf-g\|_1+\|f\|_{TV}. \tag{6}$$

An alternative minimization method is employed to solve Eq. (3) by iteratively minimizing the two subproblems. One subproblem is solved for determination of f, and the other subproblem is solved for determination of u. In the algorithm, we can simultaneously increase the value of β so that the original problem in Eq. (4) can be approximately solved. The algorithm can be described as follows.

Algorithm 1

Step 0. Initialize $\hat{\beta}$ and u;

Step 1. Fix u and solve f:

$$\min_{f}\|Hf-g\|_2^2+\gamma\|f-u\|_2^2, \tag{7}$$

where γ is estimated by the generalized cross-validation method that will be discussed in the next subsection. [We note that γ can be related to α and β in Eq. (3) by the formula $\gamma=\alpha\beta/2$.] The solution of f can be determined by solving a linear system:

$$(H^TH+\gamma I)f=H^Tg+\gamma u.$$

Such a linear system can be solved efficiently by using fast transform-based methods and preconditioning techniques for Toeplitz-like matrices; see [35];

Step 2. Fix f and solve u:

$$\min_{u} \frac{\hat{\beta}}{2}\|u-f\|_2^2+\|u\|_{TV}.$$

The solution u can be solved by many TV denoising methods such as Chambolle's projection algorithm [36] or semismooth Newton's method [37]. The algorithm goes to Step 1 if the convergence criterion is not satisfied, otherwise, it goes to Step 3;

Step 3. The value of $\hat{\beta}$ is increased by a factor $\theta>1$, i.e., $\hat{\beta}:=\hat{\beta}\theta$. The algorithm goes to Step 1.

Similarly, an alternative minimization method is employed to solve Eq. (5) by iteratively minimizing the following three subproblems and simultaneously increasing the value of β. The algorithm can be described as follows.

Algorithm 2

Step 0. Initialize $\hat{\beta}$ and f;

Step 1. Fix f and u, solve v:

$$\min_{v} \frac{\hat{\beta}}{2}\|v-(Hf-g)\|_2^2+\|v\|_1.$$

There is an explicit solution of v:

$$v=\mathrm{sign}(Hf-g)\circ\max\left\{|Hf-g|-\frac{1}{\hat{\beta}},0\right\},$$

where $\circ$ denotes the elementwise product, and the convention $0\cdot(0/0)=0$ is required.

Step 2. Fix f and v, solve u:

$$\min_{u} \frac{\hat{\beta}}{2}\|u-f\|_2^2+\|u\|_{TV}.$$

The solution u can be solved by many TV denoising methods such as Chambolle's projection algorithm [36] or the semismooth Newton method [37].

Step 3. Fix u and v and solve f:

$$\min_{f}\|v-(Hf-g)\|_2^2+\gamma\|u-f\|_2^2, \tag{8}$$

where γ is estimated by the generalized cross-validation method that will be discussed in the next subsection. The solution of f can be determined by solving a linear system:

$$(H^TH+\gamma I)f=H^T(g+v)+\gamma u.$$

The algorithm goes to Step 1 if the convergence criterion is not satisfied, otherwise, it goes to Step 4.

Step 4. The value of $\hat{\beta}$ is increased by a factor $\theta>1$, i.e., $\hat{\beta}:=\hat{\beta}\theta$. The algorithm goes to Step 1.

In the above two iterative algorithms, we need to input the values of the regularization parameters in Eqs. (7) and (8). We propose to employ the generalized cross-validation (GCV) method to determine a suitable value of the regularization parameter in the process of iteration.

A. Generalized Cross-Validation Method

Generalized cross-validation [28] is a technique that estimates a regularization parameter directly without requiring an estimate of the noise variance. It is based on the concept of prediction errors. The basic idea is to take kth observation out of all observed data, and then to use the remaining observations to predict the kth observation. If the regularization parameter is a good choice, the kth component of the fitted data should be a good predictor for kth observation on average.

For the minimization problems in Eqs. (7) and (8), it is necessary to solve the problem in the following form:

$$\min_f \|Hf - z\|_2^2 + \gamma\|f - u\|_2^2,$$

where $z=g$ for Algorithm 1 and $z=g+v$ for Algorithm 2. Based on this form, we let $f_\gamma^{(k)}$ $(k=1,\cdots,n^2)$ be the vector that minimizes the error measure:

$$\sum_{j=1,j\neq k}^{n^2} ([z-Hu]_j - [Hf]_j)^2, \tag{9}$$

where $[\cdot]_j$ is the jth element of a vector. If γ is such that $f_\gamma^{(k)}$ is a good estimate of f, then $[Hf_\gamma^{(k)}]_k$ should be a good approximation of $[z-Hu]_k$ on average. For a given γ, the average squared error between the predicted value $[Hf_\gamma^{(k)}]_k$ and the observed value $[z-Hu]_k$ is given by

$$\frac{1}{n^2}\sum_{k=1}^{n^2} ([z-Hu]_k - [Hf_\gamma^{(k)}]_k)^2.$$

The generalized cross-validation is a weighted version of the above error:

$$\tau(\gamma) \equiv \frac{1}{n^2}\sum_{k=1}^{n^2} ([z-Hu]_k - [Hf_\gamma^{(k)}]_k)^2 \left[\frac{1-m_{k,k}(\gamma)}{1-\frac{1}{n^2}\sum_{j=1}^{n^2} m_{j,j}(\gamma)}\right],$$

where $m_{j,j}(\gamma)$ is the (j,j)th entry of the so-called influence matrix

$$M(\gamma) = H(H^TH + \gamma I)^{-1}H^T.$$

In [28], Golub *et al.* have shown that $\tau(\gamma)$ can be written as

$$\tau(\gamma) = n^2\frac{\|(I-M(\gamma))(z-Hu)\|_2^2}{\mathrm{trace}(I-M(\gamma))^2}.$$

The optimal regularization parameter γ is chosen to be the one that minimizes $\tau(\gamma)$. Since $\tau(\gamma)$ is a nonlinear function, the minimizer usually cannot be determined analytically. However, if H is a blurring matrix generated by a symmetric PSF, H can be diagonalized by a fast transform matrix [35]; then we can rewrite $\tau(\gamma)$ as

$$\tau(\gamma) = \frac{n^2\sum_{j=1}^{n^2}\left\{\mathrm{diag}\left(\frac{\gamma}{\lambda_j^2+\gamma}\right)\Phi[z-Hu]\right\}_j^2}{\left(\sum_{j=1}^{n^2}\frac{\gamma}{\lambda_j^2+\gamma}\right)^2}, \tag{10}$$

where λ_j is the jth eigenvalue of H and Φ is the corresponding discrete transform matrix (e.g., discrete cosine transform matrix or discrete Fourier matrix). We recall that λ_j can be obtained by taking the fast transform of the first column of H [35].

When H cannot be diagonalized by a fast transform matrix, we can use the Hutchinson stochastic estimator to compute an approximation of $\tau(\gamma)$. It has been shown in [38] that the term $\mathrm{trace}(I-M(\gamma))$ can be rewritten as follows: $\gamma\cdot\mathrm{trace}((H^TH+\gamma I)^{-1})$. Let x denote a random vector for the Hutchinson estimator. We can define

$$\tilde{\tau}(\gamma) \equiv n^2\frac{\|(I-M(\gamma)(z-Hu)\|_2^2}{\gamma^2(x^T(H^TH+\gamma I)^{-1}x)^2}. \tag{11}$$

We employ $\tilde{\tau}(\gamma)$ to approximate $\tau(\gamma)$ and estimate a suitable value of the regularization parameter to be used in the image restoration. We note that the evaluation of $\tau(\gamma)$ involves the solution of linear systems with the matrix $H^TH+\gamma I$, which can be solved by the conjugate gradient method [39]. The conjugate gradient method can be improved by using preconditioning techniques. Transform-based preconditioning techniques have been proved to be very successful [35].

We remark that numerical approximation techniques have been proposed and developed to further reduce the computational complexity of determination of the optimal regularization parameter for the minimization of $\tau(\gamma)$ or $\tilde{\tau}(\gamma)$; see [40]. In [38], the approach employs Gauss quadrature to compute lower and upper bounds on the generalized cross-validation function. This method requires matrix-vector multiplications, and the factorization of large matrices can be avoided. Some recent methods are proposed and studied in [41,42].

3. EXPERIMENTAL RESULTS

In this section, we illustrate the performance of the proposed algorithms. Signal-to-noise ratio (SNR), blurred signal-to-noise ratio (BSNR), and improvement in signal-to-noise ratio (ISNR) are used to measure the quality of the restored images. They are defined as follows:

$$\mathrm{SNR} = 20\log_{10}\left(\frac{\|f-\mathrm{mean}(f)\|_2}{\|u-f\|_2}\right),$$

$$\mathrm{BSNR} = 20\log_{10}\left(\frac{\|g\|_2}{\|n\|_2}\right),$$

$$\mathrm{ISNR} = 20\log_{10}\left(\frac{\|f-g\|_2}{\|f-u\|_2}\right),$$

where f, g, u, and n are the original image, observed image, recovered image, and the noise vector, respectively,

and *mean*(*f*) denotes the mean of *f*. Three images are used to test the proposed algorithms. The first two images are of size 256×256, and the third one is of size 301×301. These images are shown in Fig. 1.

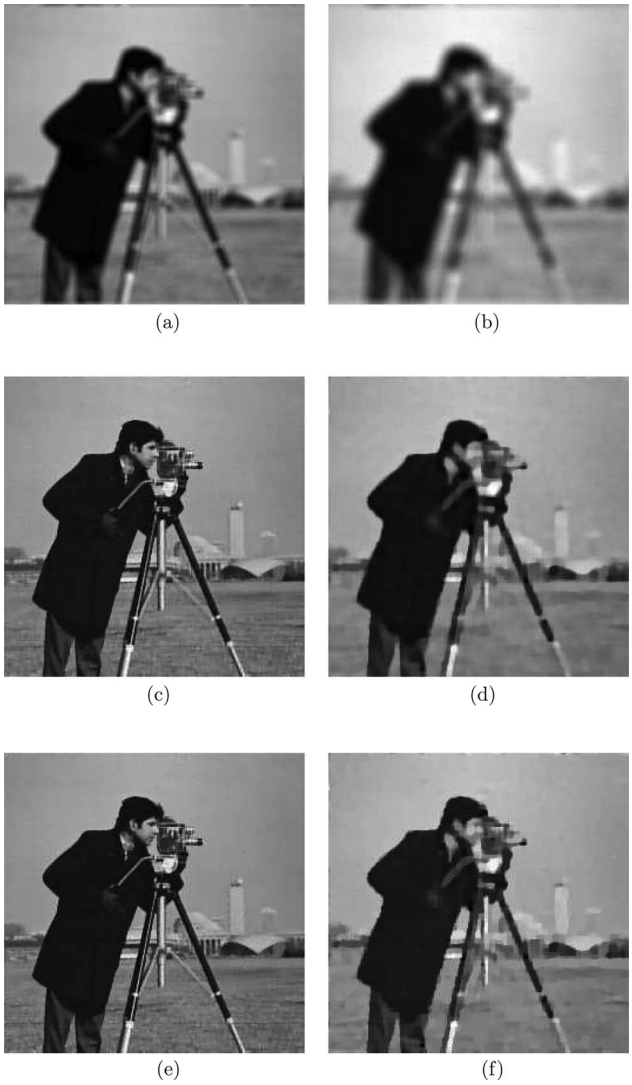
A. Experiment 1

The "Cameraman" image is blurred by a Gaussian PSF of size 7 with the standard derivation of 5, the "Barbara" image is blurred by out-of-focus PSFs with radii of 3 and 7, and the "CarNo" image is blurred by motion PSFs of length 9 and 15 pixels. Then a Gaussian noise is added to each blurred image, and the resulting observed and noisy images are shown in Figs. 2–4 (a) and (b). We also display the images restored by Algorithm 1 in Figs. 2–4 (c) and (d), respectively. It can be seen that the restored images are quite good visually. In the tests, we set the number of iterations for Steps 1 and 2 in the inner loop to 5, θ to 2, and the number of overall iterations in Algorithm 1 is 20.

To evaluate the performance of the proposed algorithms, we compare the restoration results with the restoration method in [21] by choosing the the "optimal" parameters α and β in Eq. (3). We tested 10,000 combinations of α and β; 100 values of α were uniformly and logarithmically sampled from 10^{-9} to 10^{0}, and 100 values of β were sampled starting from 1 with a step size of 10. We selected the result with the highest SNR among the tested values, and report them in Table 1. We note that it is very time-consuming to try a lot of different combinations of α and β to find the optimal regularization parameters. Indeed, we see from Table 1 that the computation time required by such calculation is very large. The main issue of this approach is that the original image is assumed to be known in the selection of regularization parameters in the method by [21]. However, in practice the original image is not known.

(a) (b) (c)

Fig. 1. Original images. (a) Cameraman. (b) Barbara.



(a) (b) (c) (d) (e) (f)

Fig. 2. Restored Cameraman images. (a) Gaussian PSF of (hsize=7, σ=5), BSNR=40 dB. (b) Gaussian PSF of (hsize=14, σ=5), BSNR=30 dB. (c) Image restored by Algorithm 1 for 2(a). (d) Image restored by Algorithm 1 for 2(b). (e) Image restored by the method in [26] for 2(a). (f) Image restored by the method in [26] for 2(b).

In Table 2, we summarize the restoration results using Algorithm 1. We also test the Hutchinson estimator for computing regularization parameters in the proposed algorithm when the blurring matrices are not diagonalized; see Table 3. We see from the tables that the performance of Algorithm 1 is about the same as that of Algorithm 1 with the Hutchinson estimator in terms of SNR and computation time. When we compare the restoration method in [21] in Table 1, it is clear that the computation time of testing 10,000 different combinations of two regularization parameters is much more than that of Algorithm 1. However, the image restoration results (in SNR) of the restoration method in [21] and Algorithm 1 are not significantly different. Visually, the restored images by the restoration method in [21] and Algorithm 1 are about the same; see Figs. 2–4. The main advantage of the proposed image restoration method is that the original image is not required in the selection of regularization parameters.

Fig. 3. Restored Barbara images. (a) Out-of-focus PSF of radius=3, BSNR=40 dB. (b) Out-of-focus PSF of radius=7, BSNR=30 dB. (c) Image restored by Algorithm 1 to 3(a). (d) Image restored by Algorithm 1 to 3(b). (e) Image restored by the method in [26] to 3(a). (f) Image restored by the method in [26] to 3(b).

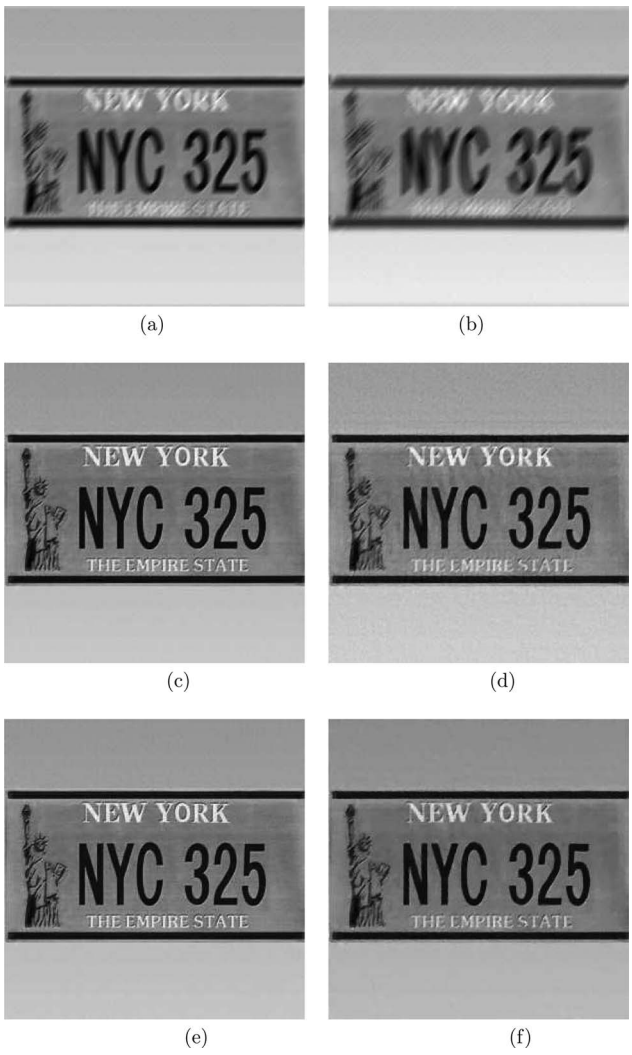


Fig. 4. Restored CarNo images. (a) Motion PSF of length 9, BSNR=40 dB. (b) Motion PSF of length 15, BSNR=30 dB. (c) Image restored by Algorithm 1 to 4(a). (d) Image restored by Algorithm 1 to 4(b). (e) Image restored by the method in [21] to 4(a). (f) Image restored by the method in [21] to 4(b).

On the other hand, we give in Table 4 the two regularization parameters searched from 10,000 different combinations of α and β by the image restoration method in [21]. In the proposed method, we use the GCV method to estimate the regularization parameter γ [see Eq. (7)], and the corresponding parameter $\hat{\alpha}$ can be calculated by $2\gamma/\hat{\beta}$. We note that $\hat{\beta}$ in Algorithm 1 increases with respect to the iterations ($\hat{\beta}=1,2,4,8,16,32,64,128,256,512$), and the method in [21] searches the 100 values of β starting from 1 with a step size of 10. Therefore, in Table 4 we can show only the parameters $\hat{\alpha}$ where the tested value of $\hat{\beta}$ is close to the best regularization parameter β searched by the method in [21]. We see from the table that $\hat{\alpha}$ computed by the proposed method is not too bad, as the restored images are visually quite good (Figs. 2–4). We also find in Table 4 that the parameters $\hat{\alpha}$ determined by Algorithm 1 are close to those determined by Algorithm 1 with the Hutchinson estimator.

Next we compare the proposed method with ALG1 in [33], where Babacan *et al.* use a stochastic method approximating *a posteriori* distribution by a product of distributions using Kullback–Leibler divergence for estimating the regularization parameter and restoring an image. In [33], Babacan *et al.* reported that this is the best among all their tested variational methods. In Table 5, we

Table 1. Restoration Results Using the Restoration Method in [21]

Image	PSF/BSNR	SNR (dB)	Time (s)
Cameraman	Gaussian (7,5)/40	18.13	4.38×10^4
Cameraman	Gaussian (14,5)/30	12.10	4.01×10^4
Barbara	Out-of-focus (3)/40	17.99	3.98×10^4
Barbara	Out-of-focus (7)/30	13.49	4.62×10^4
CarNo	Motion (9)/40	22.03	1.05×10^4
CarNo	Motion (15)/30	16.75	1.21×10^5

Table 2. Restoration Results Using Algorithm 1

Image	PSF/BSNR	SNR (dB)	Time (s)
Cameraman	Gaussian (7,5)/40	17.47	14.13
Cameraman	Gaussian (14,5)/30	11.65	16.22
Barbara	Out-of-focus (3)/40	17.40	14.25
Barbara	Out-of-focus (7)/30	13.30	14.34
CarNo	Motion (9)/40	20.41	25.69
CarNo	Motion (15)/30	15.65	25.70

Table 3. Restoration Results Using Algorithm 1 with the Hutchinson Estimator

Image	PSF/BSNR	SNR (dB)	Time (s)
Cameraman	Gaussian (7,5)/40	17.35	17.82
Cameraman	Gaussian (14,5)/30	11.74	19.39
Barbara	Out-of-focus (3)/40	17.40	17.48
Barbara	Out-of-focus (7)/30	13.34	17.44
CarNo	Motion (9)/40	20.40	30.97
CarNo	Motion (15)/30	15.64	31.00

list the restoration results (ISNRs) of the ALG1 and the proposed method. We see from the table that the proposed method is competitive with the ALG1.

B. Experiment 2

The Barbara image is blurred by a Gaussian PSF of size 7 with a standard derivation of 5. In the tests, salt-and-pepper, uniform, Laplace, partial Gaussian, and random-valued impulse noises are added to the blurred image. For these blurred and noisy images, it is more suitable to use the data-fitting measured in l_1-norm [20]. In Fig. 5, we first show the performance of Algorithm 2 for different levels of salt-and-pepper noise. According to the figures, the restoration results are very promising. Here we set the number of iterations for Steps 1–3 in the inner loop to be 5, θ to be 2, and the number of overall iterations in Algorithm 2 to be 20. In Fig. 6, we show the images restored by using Algorithm 2 for the other kinds of noise. We see from the figures that the visual quality of the restored images is good. We note that the proposed algorithm does not require any prior knowledge of the original image.

Table 5. Comparison between Algorithm 1 and ALG1 in [33]

Image	PSF/BSNR	ALG1 (ISNR)	Algorithm 1 (ISNR)
Lena	Gaussian (3)/40	4.78	5.44
Lena	Gaussian (3)/30	3.87	4.17
Lena	Gaussian (3)/20	2.87	2.57
Cameraman	Gaussian (3)/40	3.39	5.02
Cameraman	Gaussian (3)/30	2.63	3.43
Cameraman	Gaussian (3)/20	1.72	1.82
Lena	Uniform (9)/40	8.42	6.92
Lena	Uniform (9)/30	5.89	4.43
Lena	Uniform (9)/20	3.72	3.15
Cameraman	Uniform (9)/40	8.57	7.86
Cameraman	Uniform (9)/30	5.41	5.57
Cameraman	Uniform (9)/20	2.42	2.88

In Table 6, we show the restoration results of Algorithm 2 for different kinds of noise. The SNRs show that the restored images are acceptable. It is important to note that in addition, the computation time required by Algorithm 2 is very small. These results demonstrate that the proposed algorithm is effective and efficient.

C. Experiment 3

In Fig. 7, we show the convergence of Algorithms 1 and 2 in terms of SNRs. Both plots show that the SNRs of the restored images increase with respect to increasing value of $\hat{\beta}$. Here we set the increasing rate of $\hat{\beta}$ to 2. We recall that the new $\hat{\beta}$ value is equal to the old $\hat{\beta}$ value times θ, i.e., $\theta=2$. When $\hat{\beta}$ reaches a certain value, the SNRs of the restored images are kept at about the same value. These results suggest that the proposed algorithm can be stopped early even when the $\hat{\beta}$ value is not sufficiently large.

In Figs. 8 and 9, we show the SNRs of the images restored by Algorithms 1 and 2, respectively. In Figs. 8(a)

Table 4. Parameters by Different Methods

		Method in [21]		Algorithm 1		Algorithm 1 with the Hutchinson Estimator	
Image	PSF/BSNR	α	β	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
Cameraman	Gaussian (7,5)/40	8.11×10^{-5}	231	7.59×10^{-5}	256	7.31×10^{-5}	256
Cameraman	Gaussian (14,5)/30	2.85×10^{-5}	81	2.84×10^{-4}	64	1.31×10^{-4}	64
Barbara	Out-of-focus (3)/40	1.00×10^{-4}	981	1.95×10^{-3}	1024	1.95×10^{-3}	1024
Barbara	Out-of-focus (7)/30	3.51×10^{-4}	351	1.29×10^{-4}	256	2.45×10^{-4}	256
CarNo	Motion (9)/40	1.23×10^{-4}	341	2.90×10^{-4}	256	2.90×10^{-4}	256
CarNo	Motion (15)/30	6.58×10^{-4}	141	1.22×10^{-3}	128	1.41×10^{-3}	128

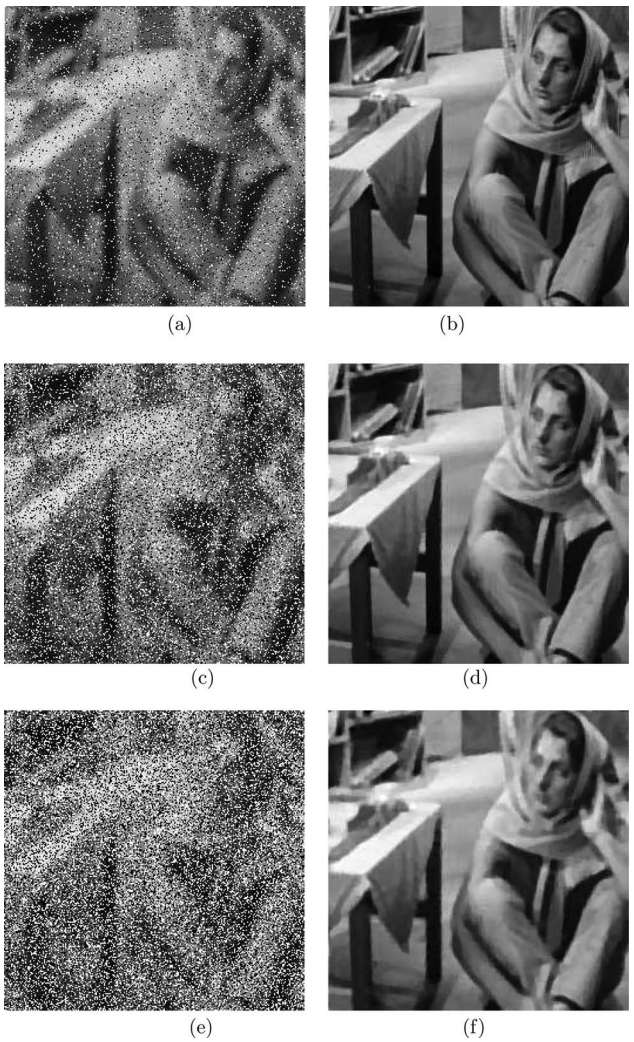


Fig. 5. Restored images of different levels of salt-and-pepper noise. (a) 10%. (b) Image restored to 5(a). (c) 30%. (d) Image restored to 5(c). (e) 50%. (f) Image restored to 5(e).

and 9(a), we fix the initial value of $\hat{\beta}$ to be 1 and check the SNRs of the restored images for different values of $\theta = 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10$. We see from the figures that they are not very sensitive to θ. For Algorithms 1 and 2, we find that the maximum differences among the SNRs of the restored images over their average SNRs are 2.71% and 1.89%. These results show that for different values of θ, the resulting SNRs of the restored images are not significantly affected. On the other hand, we fix the value θ to be 2 and check the SNRs of the restored images for different values of initial $\hat{\beta} = 1, 2, 4, 8, 16, 32, 64, 128, 256, 512$ in Figs. 8(b) and 9(b). We see from the figures that they can be sensitive to large initial value of $\hat{\beta}$; for instance, the SNRs of the images restored by Algorithm 1 drop after use of the initial $\hat{\beta} = 64$; the SNRs of the images restored by Algorithm 2 drop after use of the initial $\hat{\beta} = 32$. In both Algorithms 1 and 2, we need to solve the minimization problem $\min_u (\hat{\beta}/2)\|u - f\|_2^2 + \|u\|_{TV}$. It is clear that when $\hat{\beta}$ is large, the denoising task cannot be performed effectively. Therefore, we do not recommend use of a large initial value of $\hat{\beta}$ in the proposed image restoration process. According to our results in Figs. 8(b) and 9(b), the initial value of $\hat{\beta}$ can be set to be 1.

Fig. 6. Images restored for other kinds of noise. (a) Uniform noise. (b) Laplace noise. (c) Partial Gaussian noise (50%). (d) Random-valued noise (20%).

4. CONCLUDING REMARKS

In this paper, we studied selection of regularization parameters in total variation (TV) image restoration. We develop fast TV image restoration methods with an automatic selection of the regularization parameter scheme to restore blurred and noisy images. The methods exploit the generalized cross-validation (GCV) technique to determine inexpensively how much regularization is used in each restoration step. By updating the regularization parameter in each iteration, the restored image can be obtained. Our experimental results for different kinds of noise have shown that the visual quality of images restored by the proposed methods are quite promising even without prior knowledge of the original image. We have

Table 6. Restored Images for Different Kinds of Noise Using Algorithm 2

Noise (BSNR)	SNR (dB)	Time (s)
Uniform (40 dB)	17.68	10.82
Uniform (30 dB)	15.76	10.10
Laplace (40 dB)	15.83	10.03
Laplace (30 dB)	14.52	8.96
Partial Guassian (50% and 40 dB)	16.29	10.33
Partial Guassian (50% and 30 dB)	15.11	9.83
Random-valued impulse (20%)	15.42	10.13
Random-valued impulse (40%)	13.89	9.91

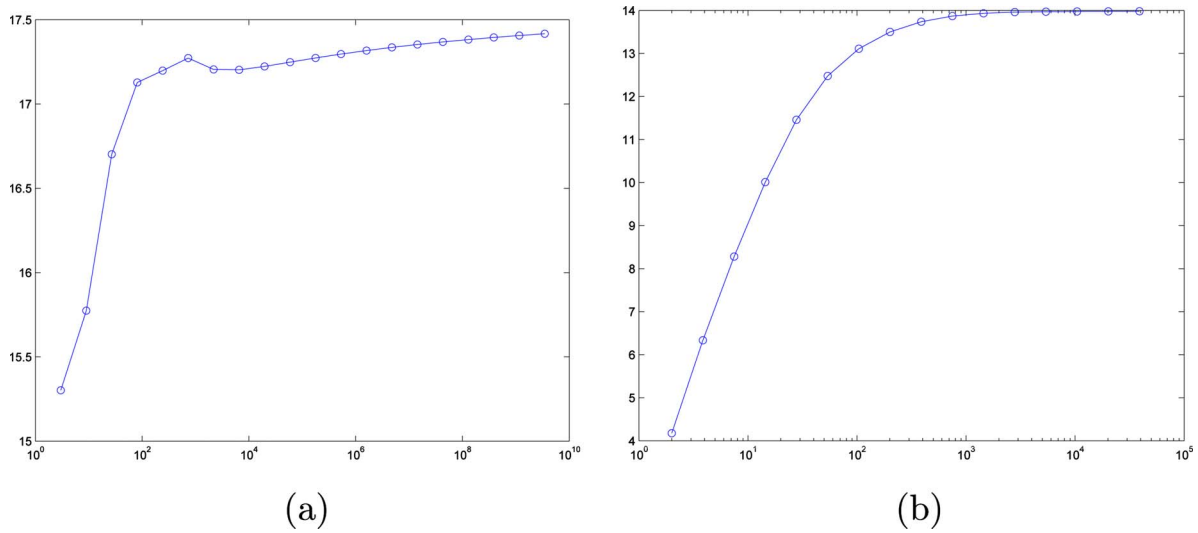


Fig. 7. (Color online) Test of convergence of Algorithms 1 and 2. (a) Test of Algorithm 1 on the Cameraman image using the Gaussian PSF of (hsize=7, σ=5, BSNR=40 dB). (b) Test of Algorithm 2 on the Barbara image using the Gaussian PSF of (hsize=7, σ=5), 40% of random-valued impulse noise.

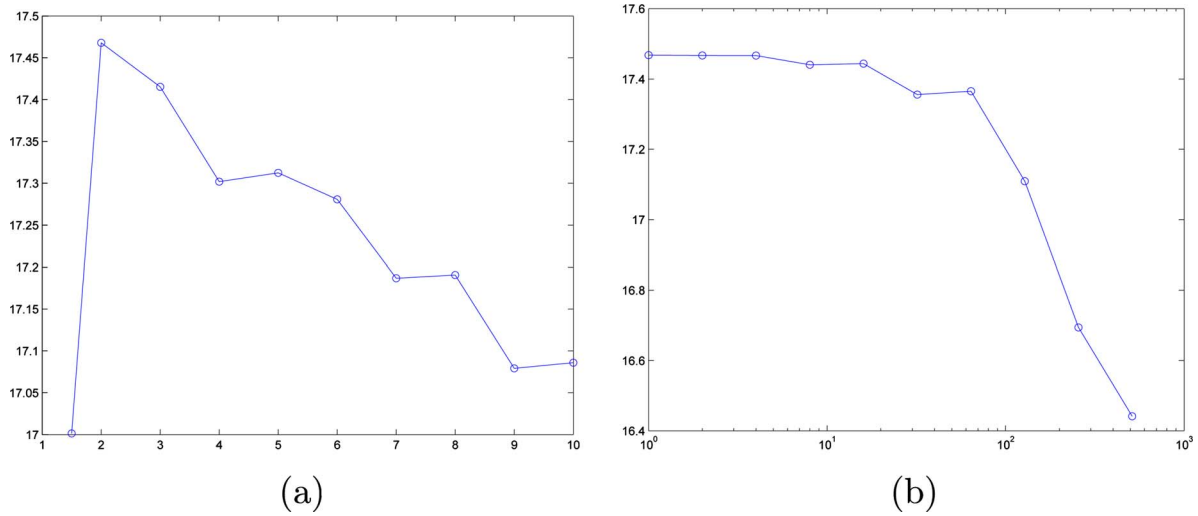


Fig. 8. (Color online) Test of Algorithm 1 on the Gaussian PSF of (hsize=7, σ=5), BSNR=40 dB. (a) SNRs of the restored images with respect to values of θ, initial β=1. (b) SNRs of the restored images with respect to initial values of β, θ=2.

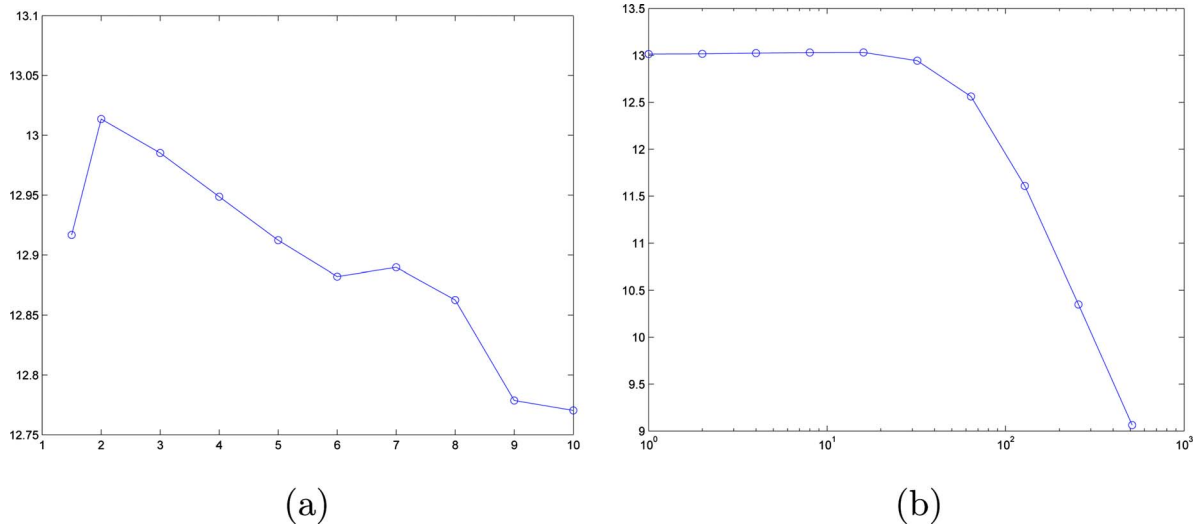


Fig. 9. (Color online) Test of Algorithm 2 on the out-of-focus PSF of radius=7, BSNR=30 dB. (a) SNRs of the restored images with respect to values of θ, initial β=1. SNRs of the restored images with respect to initial values of β, θ=2.

also demonstrated that the proposed algorithms are very efficient for image restoration.

ACKNOWLEDGMENTS

This work is supported by Hong Kong Research Grant Council 201508 and Hong Kong Baptist University, Faculty Research Grants.

REFERENCES

1. J. M. Morel and S. Solimini, *Variational Methods for Image Segmentation* (Birkhauser, 1995).
2. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D **60**, 259–268 (1992).
3. C. Vogel and M. Oman, "Iterative method for total variation denoising," SIAM J. Sci. Comput. (USA) **17**, 227–238 (1996).
4. G. Bellettini, V. Caselles, and M. Novaga, "The total variation flow in R^N," J. Differ. Equations **184**, 475–525 (2002).
5. A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," Numer. Math. **76**, 167–188 (1997).
6. D. Dobson and F. Santosa, "Reovery of blocky images from noisy and blurred data," SIAM J. Appl. Math. **56**, 1181–1198 (1996).
7. D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," Inverse Probl. **19**, 165–187 (2003).
8. P. Blomgren and T. Chan, "Color TV: total variation methods for restoration of vector-valued images," IEEE Trans. Image Process. **7**, 304–309 (1998).
9. T. Chan, M. Ng, A. Yau, and A. Yip, "Superresolution image reconstruction using fast inpainting algorithms," Appl. Comput. Harmonic Anal. **23**, 3–24 (2007).
10. T. Chan and C. Wong, "Total variation blind deconvolution," IEEE Trans. Image Process. **7**, 370–375 (1998).
11. T. Chan and J. Shen, "Variational restoration of nonflat image features: models and algorithms," SIAM J. Appl. Math. **61**, 1338–1361 (2001).
12. N. Paragios, Y. Chen, and O. Faugeras, *Handbook of Mathematical Models in Computer Vision* (Springer, 2006).
13. D. Krishnan, P. Lin, and X. Tai, "An efficient operator splitting method for noise removal in images," Commun. Comput. Phys. **1**, 847–858 (2006).
14. M. Lysaker and X. Tai, "Noise removal using smoothed normals and surface fitting," IEEE Trans. Image Process. **13**, 1345–1357 (2004).
15. M. Hintermüller and K. Kunisch, "Total bounded variation regularization as a bilaterally constrained optimization problem," SIAM J. Appl. Math. **64**, 1311–1333 (2004).
16. M. Hintermüller and G. Stadler, "An infesible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration," SIAM J. Sci. Comput. (USA) **28**, 1–23 (2006).
17. T. Chan, G. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," SIAM J. Sci. Comput. (USA) **20**, 1964–1977 (1999).
18. M. Hintermüller, K. Ito, and K. Kunisch, "The primal-dual active set strategy as a semismooth Newton's method," SIAM J. Optim. **13**, 865–888 (2003).
19. Y. Li and F. Santosa, "A computational algorithm for minimizing total variation in image reconstruction," IEEE Trans. Image Process. **5**, 987–995 (1996).
20. H. Fu, M. Ng, M. Nikolova, and J. Barlow, "Efficient minimization methods of mixed l1-l1 and l2-l1 norms for image restoration," SIAM J. Sci. Comput. (USA) **27**, 1881–1902 (2006).
21. Y. Huang, M. Ng, and Y. Wen, "A fast total variation minimization method for image restoration," SIAM J. Multiscale Model. Simul. **7**, 774–795 (2008).
22. X. Guo, F. Li, and M. Ng, "A fast l_1-TV algorithm for image restoration," SIAM J. Sci. Comput. (USA) (to be published).
23. Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," SIAM J. Imaging Sci. **1**, 248–272 (2008).
24. J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," SIAM J. Imaging Sci. **2**, 569–592 (2008).
25. R. Gonzalez and R. Woods, *Digital Image Processing* (Addison Wesley, 1992).
26. H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems* (Kluwer Academic, 1996).
27. P. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems* (SIAM, 1998).
28. G. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," Technometrics **21**, 215–223 (1979).
29. P. Hansen, J. Nagy, and D. O'leary, *Deblurring Images: Matrices, Spectra, and Filtering* (SIAM, 2006).
30. P. Hansen and D. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," SIAM J. Sci. Comput. (USA) **14**, 1487–1503 (1993).
31. C. Vogel, *Computational Methods for Inverse Problems* (SIAM, 1998).
32. Y. Lin and B. Wohlberg, "Application of the UPRE method to optimal parameter selection for large scale regularization problems," in *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation* (IEEE, 2008), pp. 89–92.
33. S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Parameter estimation in TV image restoration using variation distribution approximation," IEEE Trans. Image Process. **17**, 326–339 (2008).
34. R. Molina, J. Mateos, and A. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image and blur estimation," IEEE Trans. Image Process. **15**, 3715–3727 (2006).
35. M. Ng, *Iterative Methods for Toeplitz Systems* (Oxford Univ. Press, 2004).
36. A. Chambolle, "An algorithm for total variation minimization and applications," J. Math. Imaging Vision **20**, 89–97 (2004).
37. 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).
38. G. Golub and U. Matt, "Generalized cross-validation for large scale problems," J. Comput. Graph. Stat. **6**, 1–34 (1997).
39. M. Ng, R. Chan, and W. Tang, "A fast algorithm for deblurring models with Neumann boundary conditions," SIAM J. Sci. Comput. **21**, 851–866 (2000).
40. G. Golub and G. Meurant, "Matrices, moments and quadrature," in *Numerical Analysis*, D. Griffiths and G. Watson, eds. (Longman, 1994), pp. 105–156.
41. L. Reichel and H. Sadok, "A new L-curve for ill-posed problems," J. Comput. Appl. Math. **219**, 493–508 (2008).
42. L. Reichel, H. Sadok, and A. Shyshkov, "Greedy Tikhonov regularization for large linear ill-posed problems," Int. J. Comput. Math. **84**, 1151–1166 (2007).