ITERATIVE ALGORITHMS BASED ON THE DECOUPLE OF DEBLURRING AND DENOISING FOR IMAGE RESTORATION

YOU-WEI WEN*, MICHAEL K. NG †, AND WAI-KI CHING ‡

Abstract. In this paper, we propose iterative algorithms for solving image restoration problems. The iterative algorithms are based on the decouple of deblurring and denoising steps in the restoration process. In the deblurring step, an efficient deblurring method using fast transforms can be employed. In the denoising step, effective methods like wavelet shrinkage denoising method or total variation denoising method can be used. The main advantage of this proposal is that the resulting algorithms can be very efficient, and can produce better restored images in visual quality and signal-to-noise ratio than those by the restoration methods using the combination of a data-fitting term and a regularization term together. The convergence of the proposed algorithms is shown in the paper. Numerical examples are also given to demonstrate the effectiveness of these algorithms.

Keywords: Iterative algorithms, image restoration, deblurring, denoising, wavelet, total variation

1. Introduction. Digital image restoration and reconstruction plays an important part in various areas of applied sciences such as medical and astronomical imaging, film restoration, image and video coding. In this paper, we focus on a common degradation model [25]: an ideal image $\boldsymbol{f} \in \mathbb{R}^N$ is observed in the presence of Toeplitz matrix $\boldsymbol{H} \in \mathbb{R}^{N\times N}$ arising from a spatial-invariant blur, and an additive zero-mean Gaussian white noise $\boldsymbol{n} \in \mathbb{R}^N$ of standard deviation σ. Thus the observed image $\boldsymbol{g} \in \mathbb{R}^N$ is obtained by:

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{n}. \tag{1.1}$$

It is well-known that restoring an image $\boldsymbol{f}$ is a very ill-conditioned problem. A regularization method should be used in the image restoration process. One usual approach is to determine the restored image by minimizing a cost function consisting of a data-fitting term and a regularization term:

$$\min_{\boldsymbol{f}} \|\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \alpha\mathcal{R}(\boldsymbol{f}). \tag{1.2}$$

Here $\|\cdot\|_2$ is the Euclidean norm in $\mathbb{R}^N$, $\mathcal{R}(\boldsymbol{f})$ is the regularization term and α is a positive regularization parameter. Numerous expressions for $\mathcal{R}$ have been used in literature:

- Tikhonov regularization [15, 31]: $\mathcal{R}(\boldsymbol{f}) = \|\boldsymbol{f}\|_2^2$ or $\|\nabla\boldsymbol{f}\|_2^2$, where ∇ stands for discrete gradient. It has been shown in [24] that an efficient image restoration method based fast transforms can be developed, and the computational cost is $O(N \log N)$ operations. However, the drawback of the Tikhonov regularization is that image edges cannot be preserved in the restoration process.
- Total variation (TV) regularization [27]: $\mathcal{R}(\boldsymbol{f}) = \|\nabla\boldsymbol{f}\|_2$. The distinctive feature of TV regularization is that image edges can be preserved. Thus TV regularization is in general more suitable than the Tikhonov regularization for

* Faculty of Science, South China Agricultural University, Wushan, Guangzhou, P. R. China. E-mail: wenyouwei@graduate.hku.hk. Research supported in part by NNSFC Grant No. 60702030.

†The Corresponding Author. Centre for Mathematical Imaging and Vision, and Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong. E-mail: mng@math.hkbu.edu.hk. Research supported in part by RGC 7045/04P and 7045/05P, and HKBU FRGs.

‡Advanced Modeling and Applied Computing Laboratory, Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong. E-mail: wching@hkusua.hku.hk.

image restoration purpose. We refer readers to [28] for recent developments of TV image restoration.

- Wavelet regularization [2, 10, 13]: $\mathcal{R}(\boldsymbol{f}) = \sum_k \lambda_k \phi(\langle \boldsymbol{f}, \boldsymbol{\psi}_k \rangle)$, where $\phi(\cdot)$ is a penalty function, λ_k denotes a weighted coefficient and $\{\boldsymbol{\psi}_k\}$ denotes a wavelet orthonormal basis. In image restoration, $\phi(x) = |x|^p$, where $1 \le p \le 2$ is commonly used.

In this paper, we propose iterative algorithms for solving image restoration problems. The iterative algorithms are based on the decouple of deblurring and denoising steps in the image restoration process. The motivations of this approach are that

(M1) an efficient deblurring method based on fast transforms can be used in the deblurring step;

(M2) effective denoising methods like wavelet shrinkage method and total variation method can be employed in the denoising step;

(M3) the resulting algorithms can be very efficient, and can produce better restored images than those of the restoration method using data-fitting and regularization terms together in (1.2).

We consider and study the following iterative algorithms based on the decouple of deblurring and denoising steps (an initial guess $\boldsymbol{f}^{(0)}$ is used).

Algorithm1 :

$$\hat{\boldsymbol{f}}^{(i)} = \operatorname{argmin}_{\boldsymbol{f}} \|\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \alpha_1 \|\boldsymbol{f} - \boldsymbol{f}^{(i-1)}\|_2^2 \tag{1.3}$$

$$\boldsymbol{f}^{(i)} = \operatorname{argmin}_{\boldsymbol{f}} \|\boldsymbol{f} - \hat{\boldsymbol{f}}^{(i)}\|_2^2 + \sum_k \lambda_k \phi(\langle \boldsymbol{f}, \boldsymbol{\psi}_k \rangle) \tag{1.4}$$

Algorithm2 :

$$\hat{\boldsymbol{f}}^{(i)} = \operatorname{argmin}_{\boldsymbol{f}} \|\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \alpha_1 \|\boldsymbol{f} - \boldsymbol{f}^{(i-1)}\|_2^2 \tag{1.5}$$

$$\boldsymbol{f}^{(i)} = \operatorname{argmin}_{\boldsymbol{f}} \|\boldsymbol{f} - \hat{\boldsymbol{f}}^{(i)}\|_2^2 + \alpha_2 TV(\boldsymbol{f}) \tag{1.6}$$

We can also apply the wavelet denoising method and then the TV denoising method in the denoising step, The main reason is that the Gibbs-type oscillation artifacts may be introduced in the restored signal by using wavelet denoising methods, see for instance Figure 1.3(b) in the next subsection). When the TV denoising method is used after the wavelet denoising method, oscillation artifacts can be suppressed, see for instance Figure 1.3(d) in the next subsection.

Algorithm3 :

$$\hat{\boldsymbol{f}}^{(i)} = \operatorname{argmin}_{\boldsymbol{f}} \|\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \alpha_1 \|\boldsymbol{f} - \boldsymbol{f}^{(i-1)}\|_2^2 \tag{1.7}$$

$$\tilde{\boldsymbol{f}}^{(i)} = \operatorname{argmin}_{\boldsymbol{f}} \|\boldsymbol{f} - \hat{\boldsymbol{f}}^{(i)}\|_2^2 + \sum_k \lambda_k \phi(\langle \boldsymbol{f}, \boldsymbol{\psi}_k \rangle) \tag{1.8}$$

$$\boldsymbol{f}^{(i)} = \operatorname{argmin}_{\boldsymbol{f}} \|\boldsymbol{f} - \tilde{\boldsymbol{f}}^{(i)}\|_2^2 + \alpha_2 TV(\boldsymbol{f}) \tag{1.9}$$

In the above three algorithms, α_1 and α_2 are the positive parameters for deblurring and denoising respectively.

We note that the computational cost of the deblurring step in (1.3), (1.5) and (1.7) is $O(N \log N)$ operations [24], and both wavelet denoising methods [11, 16, 21] and the TV denoising method [5, 6, 25] in the denoising steps are also very efficient. These refer to our motivations (M1) and (M3). In the next subsection, we consider an

one-dimensional signal restoration example to further illustrate our motivations (M2) and (M3).

We remark that we use wavelet shrinkage or/and total variation in the denoising step. However, the other effective denoising methods can be used in the framework if they are available and suitable for the purpose.

1.1. An Example. The original signal is shown in Figure 1.1. The following experiments are studied.

(i) The original signal is blurred by a motion blur of length 9 and degraded by the Gaussian noise with a zero mean and a variance of 64. The blurred and noisy signal is then shown in Figure 1.2(a). Figures 1.2(b) and 1.2(c) show the restored signals by minimizing a cost function (cf. (1.2)) consisting of a data-fitting term and a regularization term of the Haar wavelet (hard thresholding) regularization method [13] and of the TV regularization method [27] respectively. Here trial by error for different values of α in (1.2) is used. In the figures, the least relative errors of the restored signals among the tested values of α are reported.

(ii) The original signal is degraded only by the Gaussian noise with a zero mean and a variance of 64. The noisy signal is then shown in Figure 1.3(a). Figures 1.3(b) and 1.3(c) show the restored signals by minimizing a cost function

$$\min_{\boldsymbol{f}} \|\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \alpha \mathcal{R}(\boldsymbol{f}) \tag{1.10}$$

($\boldsymbol{H} = \boldsymbol{I}$, i.e., no blur) using the Haar wavelet hard thresholding denoising method and the TV denoising method respectively. We observe that the Haar wavelet hard thresholding method [21] (Figure 1.3(b)) and the TV method (Figure 1.3(c)) are more effective in denoising (both methods have smaller relative errors) than in the combination of deblurring and denoising together (Figures 1.2(b)-(c)). This illustrates our motivation (M2).

(iii) For the motivation (M3), we used the three proposed iterative algorithms to restore the blurred and noisy signal in Figure 1.2(a). Figures 1.4(a), 1.4(b) and 1.4(c) show the restored signals by using Algorithms 1, 2 and 3 respectively. In these tests, Haar wavelet hard thresholding denoising method [21] is used. We see from the figures that the relative errors of the restored signals by Algorithms 1 and 2 are less than those of Figures 1.2(b) and 1.2(c). We also find that the relative errors of the restored signals by Algorithm 3 is the least one among all the restored signals. These examples show that the decouple of deblurring and denoising for image restoration can be an effective approach. In Section 4, more numerical examples will be given to demonstrate the effectiveness of this approach.

1.2. Outline. The outline of this paper is given as follows. In Section 2, we review deblurring and denoising methods. In Section 3, we show the convergence of Algorithms 1, 2 and 3. In Section 4, numerical examples are given to demonstrate the effectiveness of the proposed methods. Finally, concluding remarks are given in Section 5.

2. Deblurring and Denoising Methods. In this section, we review deblurring and denoising methods.

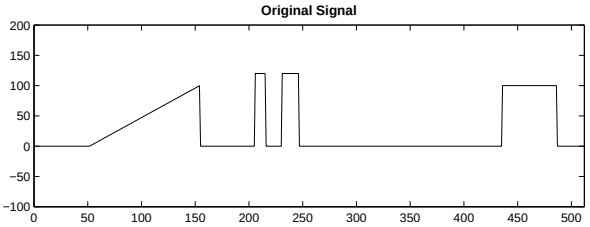

FIGURE 1.1. *The original signal.*

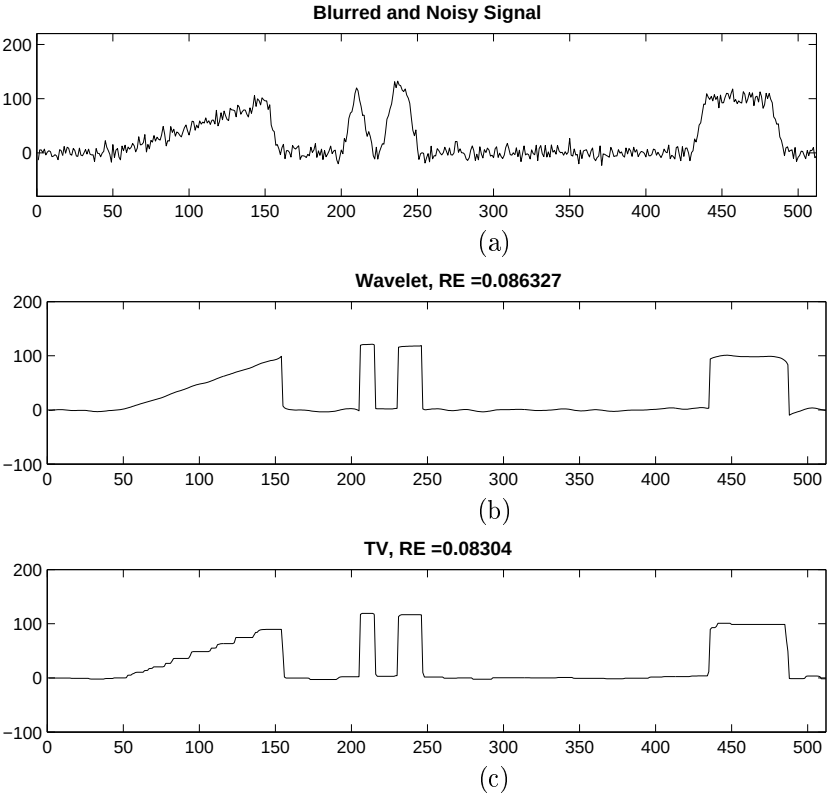

FIGURE 1.2. *The blurred and noisy signal (a); the restored signal by the Haar wavelet regularization method (b); the restored signal by the TV regularization method (c).*

2.1. Deblurring Methods. In order to determine $\hat{\boldsymbol{f}}^{(i)}$ in (1.3), (1.5) and (1.7), it is required to solve the following linear system:

$$\left(\boldsymbol{H}^T\boldsymbol{H} + \alpha_1\boldsymbol{I}\right)\hat{\boldsymbol{f}}^{(i)} = \boldsymbol{H}^T\boldsymbol{g} + \alpha_1\boldsymbol{f}^{(i-1)}. \tag{2.1}$$

Because of the regularized term $\alpha_1\boldsymbol{I}$, the coefficient matrix $(\boldsymbol{H}^T\boldsymbol{H} + \alpha_1\boldsymbol{I})$ is always invertible even if $\boldsymbol{H}^T\boldsymbol{H}$ is singular.

We note that $\boldsymbol{H}$ is a block-circulant with circulant-block (BCCB) matrix when periodic boundary conditions are applied to the image boundary. The matrix $\boldsymbol{H}$ can be diagonalized by the discrete Fast Fourier Transform (FFT) matrix, and therefore the cost of solving (2.1) is $O(N \log N)$ operations where N is the number of pixels of the restored image. When zero boundary conditions are applied to the image

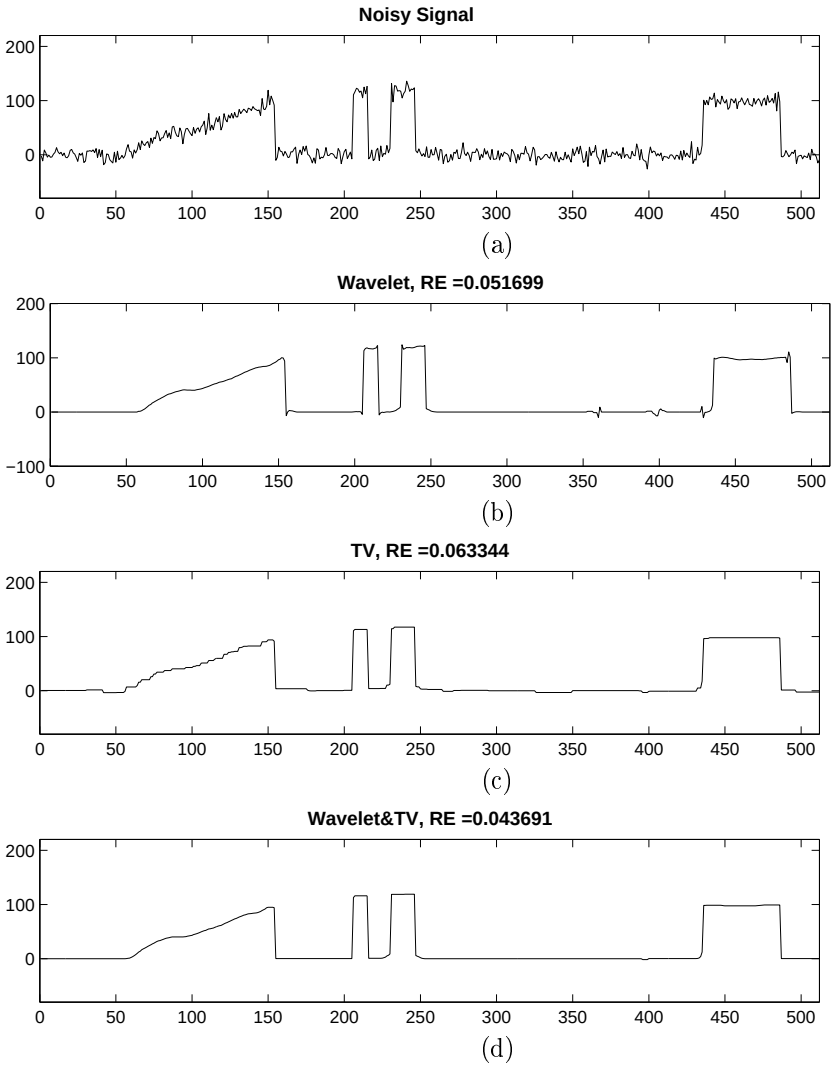

FIGURE 1.3. *The noisy signal (a); the restored signal by the Haar wavelet regularization method (b); the restored signal by the TV regularization method (c); the restored signal by the Haar wavelet regularization method and then the TV regularization method (d).*

boundary, $\boldsymbol{H}$ is a block-Toeplitz with Toeplitz-block (BTTB) matrix. The conjugate gradient method can be used to solve (2.1). The convergence of the conjugate gradient method can be improved using preconditioning techniques. Transformation-based preconditioning techniques have been proved to be very successful [23]. When reflective boundary conditions applied to the image boundary, $\boldsymbol{H}$ is a block-Toeplitz-plus-Hankel with Toeplitz-plus-Hankel-block (BTHTHB) matrix. If the blurring function is symmetric, such block-Toeplitz-plus-Hankel with Toeplitz-plus-Hankel-block matrix can be diagonalized by the discrete cosine transform (DCT) matrix [24]. Therefore the cost of solving the corresponding linear system is $O(N \log N)$ operations.

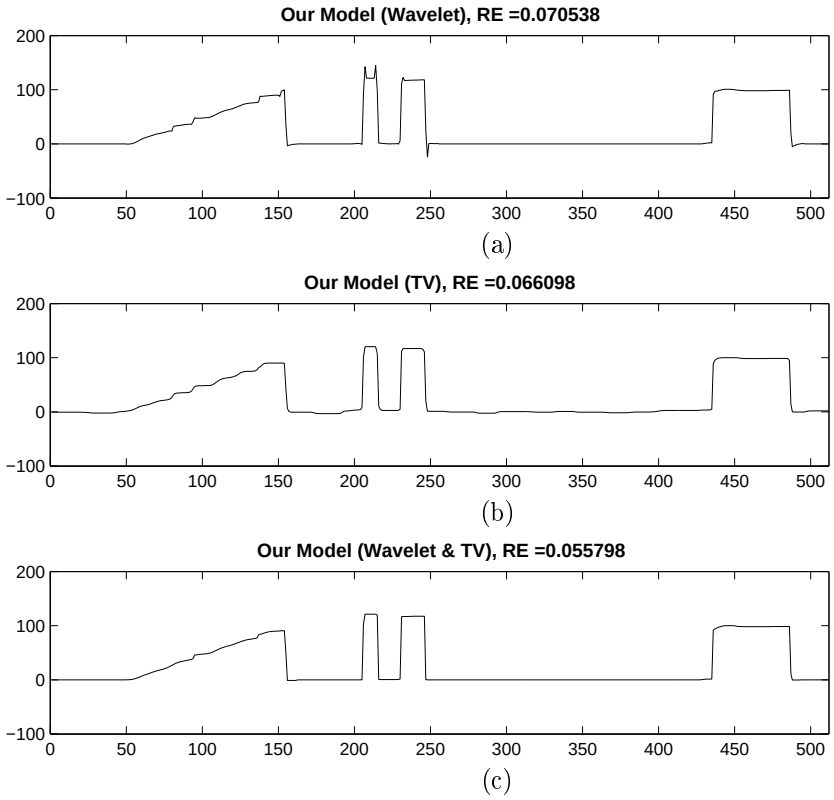

FIGURE 1.4. *The restored signal by Algorithm 1 (a); the restored signal by the TV regularization method (b); the restored signal by Algorithm 2 (b); the restored signal by Algorithm 3(c).*

We note from (1.1) and (2.1) that

$$\hat{\boldsymbol{f}}^{(i)} = \left(\boldsymbol{H}^T\boldsymbol{H} + \alpha_1\boldsymbol{I}\right)^{-1}\left(\boldsymbol{H}^T\boldsymbol{g} + \alpha_1\boldsymbol{f}^{(i-1)}\right) = \boldsymbol{f} + \boldsymbol{\epsilon}_f^{(i)} + \boldsymbol{\epsilon}_n, \tag{2.2}$$

where

$$\boldsymbol{\epsilon}_f^{(i)} = \alpha_1\left(\boldsymbol{H}^T\boldsymbol{H} + \alpha_1\boldsymbol{I}\right)^{-1}\left(\boldsymbol{f}^{(i-1)} - \boldsymbol{f}\right) \quad \text{and} \quad \boldsymbol{\epsilon}_n = \left(\boldsymbol{H}^T\boldsymbol{H} + \alpha_1\boldsymbol{I}\right)^{-1}\boldsymbol{H}^T\boldsymbol{n}.$$

Here the variance of $\boldsymbol{\epsilon}_n$ is given by $\sigma^2\left(\boldsymbol{H}^T\boldsymbol{H} + \alpha_1\boldsymbol{I}\right)^{-1}\boldsymbol{H}^T\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H} + \alpha_1\boldsymbol{I}\right)^{-1}$. According to (2.2), we will choose a small value of α_1 in order to reduce the error $\boldsymbol{\epsilon}_f^{(i)}$. However, the residual noise $\boldsymbol{\epsilon}_n$ will be amplified by using a small value of α_1. In the denoising step of Algorithms 1, 2 and 3, we apply wavelet shrinkage or/and TV regularization denoising methods to suppress the residual noise.

2.2. Wavelet Denoising Methods. Wavelet-based noise reduction algorithms [11, 12, 21, 33] are asymptotically near optimal for a wide class of signals corrupted by additive Gaussian white noises. These algorithms also work well when the noise is neither white noise nor Gaussian, see for instance [17].

By using the unitary invariance property of the 2-norm, one can rewrite

$$\|\boldsymbol{f} - \hat{\boldsymbol{f}}^{(i)}\|_2^2 = \sum_k \left(\langle \boldsymbol{f}, \boldsymbol{\psi}_k\rangle - \langle \hat{\boldsymbol{f}}^{(i)}, \boldsymbol{\psi}_k\rangle\right)^2$$

in (1.4), (1.6) and (1.8). Therefore we have

$$\boldsymbol{f}^{(i)} = \text{argmin}_{\boldsymbol{f}} \sum_k \left(\left(\langle \hat{\boldsymbol{f}}^{(i)}, \boldsymbol{\psi}_k \rangle - \langle \boldsymbol{f}, \boldsymbol{\psi}_k \rangle \right)^2 + \lambda_k |\langle \boldsymbol{f}, \boldsymbol{\psi}_k \rangle|^p \right). \tag{2.3}$$

It is easy to see that the wavelet coefficients of $\boldsymbol{f}^{(i)}$ can be obtained independently for each $\langle \boldsymbol{f}^{(i)}, \boldsymbol{\psi}_k \rangle$ as a scalar optimization problem. In particular, when $p = 1$, the solution in (2.3) is just the popular soft thresholding scheme [12]:

$$\langle \hat{\boldsymbol{f}}^{(i)}, \boldsymbol{\psi}_k \rangle = \begin{cases} \langle \hat{\boldsymbol{f}}^{(i)}, \boldsymbol{\psi}_k \rangle - \lambda_k/2, & \text{if } \langle \hat{\boldsymbol{f}}^{(i)}, \boldsymbol{\psi}_k \rangle \geq \lambda_k/2, \\ \langle \hat{\boldsymbol{f}}^{(i)}, \boldsymbol{\psi}_k \rangle + \lambda_k/2, & \text{if } \langle \hat{\boldsymbol{f}}^{(i)}, \boldsymbol{\psi}_k \rangle \leq -\lambda_k/2, \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

Usually, λ_k is set to be a constant related to the variance σ_k of the residual noise $\boldsymbol{\epsilon}_n$ in its k-th level wavelet decomposition, see [11]. In our setting, the variance σ_k can be estimated as follows:

$$\sigma_k^2 = \mathbf{E}\{|\langle \boldsymbol{\epsilon}_n, \boldsymbol{\psi}_k \rangle|^2\} = \sum_{u,v} \frac{\sigma^2 |H(u,v)|^2 |\Psi(k,u,v)|^2}{(|H(u,v)|^2 + \alpha_1)^2}, \tag{2.5}$$

where $\mathbf{E}(\cdot)$ is the expectation operator, $H(u,v)$ and $\Psi(k,u,v)$ are the discrete Fourier transforms of $\boldsymbol{H}$ and $\boldsymbol{\psi}_k$ respectively. Both $H(u,v)$ and $\Psi(k,u,v)$ can be computed by using the FFT in $O(N \log N)$ operations. This optimization step can be done independently for each wavelet coefficient and therefore the cost of wavelet denoising the image solution is in $O(N \log N)$ operations. Other wavelet shrinkage denoising schemes like hard thresholding method [21] and Garrote thresholding method [16] can be used in the denoising step of the proposed algorithms.

2.3. The TV Denoising Method. The use of TV regularization for image processing has been proposed by Rudin, Osher and Fatemi [27]. The distinctive feature of TV regularization is that edges can be preserved. We remark that Steidl et.al. [29] have shown the equivalence of soft wavelet shrinkage, total variation diffusion and total variation regularization. Recent developments of TV image restoration can be found in [28].

The Euler-Lagrange equation for the total variation denoising problem is given by the following a nonlinear system of equations. The artificial time marching scheme proposed in Rudin et al. [27] obtains the solution of this nonlinear system of equations as the steady state of a parabolic partial differential equation. Indeed, it is equivalent to employing the following gradient descent method to solve the minimization problem. The drawback of this method is that the convergence rate of explicit methods can be very slow due to stability constraints. Implicit methods can be applied but then one has to deal with the nonlinearity problem and the solution of the resulting linear systems. In [32], the lagged diffusivity fixed point iteration is introduced. This method consists of linearizing the nonlinear differential term in the nonlinear system of equations by lagging the diffusion coefficient one iteration behind. The resulting linear system can be solved efficiently by multigrid methods [32]. The algorithm can be interpreted in the framework of generalized Weiszfeld's methods, see for instance [7]. As proved in [7], this method is monotonically convergent, in the sense that the objective function values evaluated at the iterates form a monotonically decreasing sequence, and the convergence rate is linear. In practice, this method is very robust. Recently, Chan and Chen [6] proposed a very fast multilevel method using primal

relaxation for total variation denoising problems. Chambolle [5] considered a dual formulation of total variation denoising problem, and used the projected gradient method to solve the resulting constrained optimization problem. Based on the theory on semismooth operators, Ng et al. [25] studied semismooth Newton's methods for computing the nonlinear projection. The convergence and numerical results have shown that the proposed algorithm is quite effective.

3. Convergence Analysis. Let us first introduce the following notations in the following discussion:

$$\boldsymbol{S}_h(\boldsymbol{x}) \equiv \text{argmin}_{\boldsymbol{f}} \ \|\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}\|_2^2 + \alpha_1\|\boldsymbol{f} - \boldsymbol{x}\|_2^2$$

$$\boldsymbol{S}_w(\boldsymbol{x}) \equiv \text{argmin}_{\boldsymbol{f}} \ \|\boldsymbol{f} - \boldsymbol{x}\|_2^2 + \sum_k \lambda_k \phi(\langle \boldsymbol{f}, \boldsymbol{\psi}_k \rangle)$$

$$\boldsymbol{S}_{tv}(\boldsymbol{x}) \equiv \text{argmin}_{\boldsymbol{f}} \ \|\boldsymbol{f} - \boldsymbol{x}\|_2^2 + \alpha_2 TV(\boldsymbol{f})$$

The main aim of this section is to show the convergence of Algorithm 1:

$$\boldsymbol{f}^{(i)} = \boldsymbol{S}_w(\boldsymbol{S}_h(\boldsymbol{f}^{(i-1)}))$$

Algorithm 2:

$$\boldsymbol{f}^{(i)} = \boldsymbol{S}_{tv}(\boldsymbol{S}_h(\boldsymbol{f}^{(i-1)}))$$

and Algorithm 3:

$$\boldsymbol{f}^{(i)} = \boldsymbol{S}_{tv}(\boldsymbol{S}_w(\boldsymbol{S}_h(\boldsymbol{f}^{(i-1)})))$$

for $i = 1, 2, \cdots$. We will use the techniques of non-expansive operators to show the convergence. Let us first introduce some definitions.

DEFINITION 3.1. *[8] An operator $\boldsymbol{P} : \mathbb{R}^N \to \mathbb{R}^N$ is called non-expansive if for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^N$, we have*

$$||\boldsymbol{P}(\boldsymbol{x}_1) - \boldsymbol{P}(\boldsymbol{x}_2)||_2 \leq ||\boldsymbol{x}_1 - \boldsymbol{x}_2||_2.$$

If there exists a number $\beta \in (0, 1)$ and an non-expansive operator $\boldsymbol{T} : \mathbb{R}^N \to \mathbb{R}^N$ such that $\boldsymbol{P} = (1-\beta)\boldsymbol{I} + \beta\boldsymbol{T}$ is non-expansive, then $\boldsymbol{P}$ is called β-averaged. In particular, when $\beta = 1/2$, $\boldsymbol{P}$ is called a firmly non-expansive operator.

We remark that an equivalent form of firmly non-expansive operator is defined as

$$\|\boldsymbol{P}(\boldsymbol{x}_1) - \boldsymbol{P}(\boldsymbol{x}_2)\|_2^2 \leq [\boldsymbol{P}(\boldsymbol{x}_1) - \boldsymbol{P}(\boldsymbol{x}_2)]^T (\boldsymbol{x}_1 - \boldsymbol{x}_2), \tag{3.1}$$

see for instance [9]. In fact, let $\boldsymbol{T}$ be a non-expansive operator such that $\boldsymbol{P} = \frac{1}{2}\boldsymbol{I} + \frac{1}{2}\boldsymbol{T}$. For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$, we have

$$\|\boldsymbol{T}(\boldsymbol{x}) - \boldsymbol{T}(\boldsymbol{y})\|_2^2 = 4\|\boldsymbol{P}(\boldsymbol{x}) - \boldsymbol{P}(\boldsymbol{y})\|_2^2 + \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 - 4(\boldsymbol{P}(\boldsymbol{x}) - \boldsymbol{P}(\boldsymbol{y}))^T(\boldsymbol{x} - \boldsymbol{y}).$$

Since $\boldsymbol{T}$ is non-expansive, i.e.,

$$\|\boldsymbol{T}(\boldsymbol{x}) - \boldsymbol{T}(\boldsymbol{y})\|_2^2 \leq \|\boldsymbol{x} - \boldsymbol{y}\|_2^2,$$

we obtain the inequality (3.1). On the other hand, given an operator $\boldsymbol{P}$ with (3.1), we have

$$\|(2\boldsymbol{P}-\boldsymbol{I})(x)-(2\boldsymbol{P}-\boldsymbol{I})(y)\|_2^2 \leq \|\boldsymbol{x}-\boldsymbol{y}\|_2^2.$$

We can construct $\boldsymbol{T}=2\boldsymbol{P}-\boldsymbol{I}$ and the above formula shows that $\boldsymbol{T}$ is non-expansive. The following lemma shows that the product of two averaged non-expansive is also averaged non-expansive.

LEMMA 3.2. *Let $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ be β_1- and β_2-averaged non-expansive operators respectively. Then $\boldsymbol{P}_1\boldsymbol{P}_2$ is $(\beta_1+\beta_2-\beta_1\beta_2)$-averaged non-expansive.*

Proof. Since $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are β_1- and β_2-averaged non-expansive operators, there exist non-expansive operators $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ such that

$$\boldsymbol{P}_i=(1-\beta_i)\boldsymbol{I}+\beta_i\boldsymbol{T}_i$$

for $i=1,2$. Thus we obtain

$$\boldsymbol{P}_1\boldsymbol{P}_2=(1-\beta_1-\beta_2+\beta_1\beta_2)\boldsymbol{I}+(1-\beta_1)\beta_2\boldsymbol{T}_2+(1-\beta_2)\beta_1\boldsymbol{T}_1+\beta_1\beta_2\boldsymbol{T}_1\boldsymbol{T}_2.$$

Set $\beta_3=\beta_1+\beta_2-\beta_1\beta_2$ and

$$\boldsymbol{T}=\frac{1}{\beta_1+\beta_2-\beta_1\beta_2}\left((1-\beta_1)\beta_2\boldsymbol{T}_2+(1-\beta_2)\beta_1\boldsymbol{T}_1+\beta_1\beta_2\boldsymbol{T}_1\boldsymbol{T}_2\right).$$

It is easy to check that $\beta_3\in(\beta_2,1)$ with $\beta_1,\beta_2\in(0,1)$ and $\boldsymbol{T}$ is non-expansive. □

Now we show that $\boldsymbol{S}_h$, $\boldsymbol{S}_w$ and $\boldsymbol{S}_{tv}$ are firmly non-expansive. For any $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^N$, we have

$$\boldsymbol{S}_h(\boldsymbol{x})-\boldsymbol{S}_h(\boldsymbol{y})=\alpha_1\left(\boldsymbol{H}^T\boldsymbol{H}+\alpha_1\boldsymbol{I}\right)^{-1}(\boldsymbol{x}-\boldsymbol{y}).$$

It is straightforward to obtain the following inequality:

$$\begin{aligned}\|\boldsymbol{S}_h(\boldsymbol{x})-\boldsymbol{S}_h(\boldsymbol{y})\|_2^2&=\left(\boldsymbol{S}_h(\boldsymbol{x})-\boldsymbol{S}_h(\boldsymbol{y})\right)^T\alpha_1\left(\boldsymbol{H}^T\boldsymbol{H}+\alpha_1\boldsymbol{I}\right)^{-1}(\boldsymbol{x}-\boldsymbol{y})\\&\leq\left(\boldsymbol{S}_h(\boldsymbol{x})-\boldsymbol{S}_h(\boldsymbol{y})\right)^T(\boldsymbol{x}-\boldsymbol{y}).\end{aligned}$$

It follows that $\boldsymbol{S}_h$ is firmly non-expansive. For $\boldsymbol{S}_w$ and $\boldsymbol{S}_{tv}$, we can use the following results in [9].

LEMMA 3.3. *[9, Lemma 2.4] Let γ be a positive number and φ be a convex and semi-continuous function. Suppose*

$$\widehat{\boldsymbol{x}}\equiv\operatorname{argmin}_{\boldsymbol{x}}\ \|\boldsymbol{y}-\boldsymbol{x}\|_2^2+\gamma\varphi(\boldsymbol{x}), \tag{3.2}$$

and define $\boldsymbol{S}:\mathcal{R}^N\to\mathbb{R}^N$ such that $\widehat{\boldsymbol{x}}=\boldsymbol{S}(\boldsymbol{y})$. Then $\boldsymbol{S}$ is firmly non-expansive.

We note that $|x|^p$ $(1\leq p\leq 2)$ is a convex and semi-continuous function. By using the results in Lemma 3.3, both operators $\boldsymbol{S}_w$ and $\boldsymbol{S}_{tv}$ are firmly non-expansive.

In order to establish the convergence of the proposed iterative scheme, we study the following property of β-averaged non-expansive operators, see for instance [4].

LEMMA 3.4. *Let $\boldsymbol{P}:\mathbb{R}^N\to\mathbb{R}^N$ be an β-averaged non-expansive operator. For a given $\boldsymbol{z}^{(0)}\in\mathbb{R}^N$, define the sequence $\boldsymbol{z}^{(k)}=\boldsymbol{P}(\boldsymbol{z}^{(k-1)})$ for $k=1,2,\cdots$. If the set of fixed points of $\boldsymbol{P}$ is nonempty, then $\boldsymbol{P}$ is asymptotically regular, i.e.,*

$$\lim_{k\to\infty}||\boldsymbol{z}^{(k+1)}-\boldsymbol{z}^{(k)}||_2=\lim_{k\to\infty}||\boldsymbol{P}^{k+1}(\boldsymbol{z}^{(0)})-\boldsymbol{P}^k(\boldsymbol{z}^{(0)})||_2=0.$$

Proof. Let $\boldsymbol{z}$ be a fixed point of $\boldsymbol{P}$. We denote $\boldsymbol{g} = \boldsymbol{I} - \boldsymbol{P}$. Since $\boldsymbol{P}$ is an β-averaged non-expansive operator, there exist a number $\beta \in (0,1)$ and a non-expansive operator $\boldsymbol{T}$ such that $\boldsymbol{P} = (1-\beta)\boldsymbol{I} + \beta\boldsymbol{T}$. It also implies that $\boldsymbol{g} = \beta(\boldsymbol{I} - \boldsymbol{T})$. For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$, we have

$$\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 - \|\boldsymbol{T}(\boldsymbol{x}) - \boldsymbol{T}(\boldsymbol{y})\|_2^2 = \frac{2}{\beta}\left(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\right)^T(\boldsymbol{x}-\boldsymbol{y}) - \frac{1}{\beta^2}\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_2^2 \geq 0.$$

or

$$\left[\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\right]^T(\boldsymbol{x}-\boldsymbol{y}) \geq \frac{1}{2\beta}\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_2^2.$$

By using the fact that $\boldsymbol{z} = \boldsymbol{P}(\boldsymbol{z})$, we have

$$\begin{aligned}\|\boldsymbol{z}-\boldsymbol{z}^{(k)}\|_2^2 - \|\boldsymbol{z}-\boldsymbol{z}^{(k+1)}\|_2^2 &= \|\boldsymbol{z}-\boldsymbol{z}^{(k)}\|_2^2 - \|\boldsymbol{P}(\boldsymbol{z}) - \boldsymbol{P}(\boldsymbol{z}^{(k)})\|_2^2 \\ &= 2\left(\boldsymbol{g}(\boldsymbol{z}) - \boldsymbol{g}(\boldsymbol{z}^{(k)})\right)^T\left(\boldsymbol{z}-\boldsymbol{z}^{(k)}\right) - \|\boldsymbol{g}(\boldsymbol{z}) - \boldsymbol{g}(\boldsymbol{z}^{(k)})\|_2^2 \\ &\geq \left(\frac{1}{\beta} - 1\right)\|\boldsymbol{g}(\boldsymbol{z}) - \boldsymbol{g}(\boldsymbol{z}^{(k)})\|_2^2.\end{aligned}$$

We note that $\boldsymbol{g}(\boldsymbol{z}) = \boldsymbol{z} - \boldsymbol{P}(\boldsymbol{z}) = 0$ and $\boldsymbol{g}(\boldsymbol{z}^{(k)}) = \boldsymbol{z}^{(k)} - \boldsymbol{z}^{(k+1)}$. Therefore, we obtain

$$\|\boldsymbol{z}-\boldsymbol{z}^{(k)}\|_2^2 - \|\boldsymbol{z}-\boldsymbol{z}^{(k+1)}\|_2^2 \geq \left(\frac{1}{\beta} - 1\right)\|\boldsymbol{z}^{(k)} - \boldsymbol{z}^{(k+1)}\|_2^2.$$

Since $(1/\beta - 1) > 0$, one can immediately deduce that $\sum_k \|\boldsymbol{z}^{(k)} - \boldsymbol{z}^{(k+1)}\|_2^2$ is bounded. Then the result follows. □

Now we consider the well-known Opial theorem [26]. It states that if the set of the fixed points of $\boldsymbol{P}$ is nonempty, and $\boldsymbol{P}$ is non-expansive and asymptotically regular, then the sequence $\boldsymbol{x}^{(k)}$ where $\boldsymbol{x}^{(k)} = \boldsymbol{P}(\boldsymbol{x}^{(k-1)})$ for $k = 1, 2, \ldots$ converges weakly to a fixed point. Using the results in Lemma 3.4, we have the following theorem.

THEOREM 3.5. *[4, Theorem 2.1] Let $\boldsymbol{P} : \mathbb{R}^N \to \mathbb{R}^N$ be an β-averaged non-expansive operator. If the set of the fixed points of $\boldsymbol{P}$ is nonempty, then for any $\boldsymbol{z}^{(0)}$, the sequence $\{\boldsymbol{z}^{(k)}\}$ where $\boldsymbol{z}^{(k)} = \boldsymbol{P}(\boldsymbol{z}^{(k-1)})$ for $k = 1, 2, \cdots$ converges to a fixed point in $\mathbb{R}^N$.*

According to Lemma 3.2, since the operators $\boldsymbol{S}_h, \boldsymbol{S}_w$ and $\boldsymbol{S}_{tv}$ are firmly non-expansive, the product $\boldsymbol{S}_{tv}\boldsymbol{S}_w\boldsymbol{S}_h$ is β-averaged non-expansive for some $\beta \in (0,1)$. By applying Theorem 3.5, we have our main results of this paper.

THEOREM 3.6. *(i) Assume that the set of the fixed points of $\boldsymbol{S}_w\boldsymbol{S}_h$ is nonempty, Algorithm 1 converges to a fixed point.*

(ii) Assume that the set of the fixed points of $\boldsymbol{S}_{tv}\boldsymbol{S}_h$ is nonempty, Algorithm 2 converges to a fixed point.

(iii) Assume that the set of the fixed points of $\boldsymbol{S}_{tv}\boldsymbol{S}_w\boldsymbol{S}_h$ is nonempty, Algorithm 3 converges to a fixed point.

Before we end this section, we give several remarks about the convergence of the proposed algorithm.

- When $\boldsymbol{H}$ is rank-deficient, $\boldsymbol{S}_w\boldsymbol{S}_h$, $\boldsymbol{S}_{tv}\boldsymbol{S}_h$, and $\boldsymbol{S}_{tv}\boldsymbol{S}_w\boldsymbol{S}_h$ may not be strictly contraction operators. However, we can add a projection operator $\boldsymbol{S}_p$ at each iteration such that the values of the entries of $\boldsymbol{f}^{(i+1)}$ is projected to $[0, 255]$. It follows from the Banach-Algoglu theorem that the operators $\boldsymbol{S}_p\boldsymbol{S}_w\boldsymbol{S}_h$,

$\boldsymbol{S}_p\boldsymbol{S}_{tv}\boldsymbol{S}_h$, and $\boldsymbol{S}_p\boldsymbol{S}_{tv}\boldsymbol{S}_w\boldsymbol{S}_h$ have a weak accumulation point, see [10] for details. Therefore, their corresponding set of fixed points are non-empty. By using Theorem 3.6, we can show that Algorithms 1, 2 and 3 converges to a fixed point.

- When $\boldsymbol{H}$ is full-rank, i.e., the smallest eigenvalue of $\boldsymbol{H}^T\boldsymbol{H}$ is larger than 0, it is easy to check that for all $\boldsymbol{x}, \boldsymbol{y}$, there exists a number $r \in (0,1)$ such that

$$\|\boldsymbol{S}_h(\boldsymbol{x}) - \boldsymbol{S}_h(\boldsymbol{y})\|_2 \leq r\|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

It implies that $\boldsymbol{S}_w\boldsymbol{S}_h$, $\boldsymbol{S}_{tv}\boldsymbol{S}_h$, and $\boldsymbol{S}_{tv}\boldsymbol{S}_w\boldsymbol{S}_h$ are strictly contraction operator. According to Banach-Picard theorem, for any initial vector $\boldsymbol{f}^{(0)}$, Algorithms 1, 2 and 3 converges to a unique fixed point.

4. Numerical Results. In this section, we illustrate the performance of Algorithms 1, 2 and 3 for solving image restoration problems. Three indices are used to measure the quality of the restored images by different algorithms. They are the Signal-to-Noise Ratio (SNR), the Improvement in SNR (ISNR), and the Blurred SNR (BSNR):

$$\text{SNR} = 10\log_{10}\frac{\|\boldsymbol{f}\|_2^2}{\|\boldsymbol{f} - \widetilde{\boldsymbol{f}}\|_2^2}, \quad \text{ISNR} = 10\log_{10}\frac{\|\boldsymbol{f} - \boldsymbol{g}\|_2^2}{\|\boldsymbol{f} - \widetilde{\boldsymbol{f}}\|_2^2}$$

and

$$\text{BSNR} = 10\log_{10}\frac{\|\boldsymbol{g}\|_2^2}{\|\boldsymbol{n}\|_2^2},$$

where $\boldsymbol{f}, \boldsymbol{g}, \widetilde{\boldsymbol{f}}$ and $\boldsymbol{n}$ are the original image, the blurred and noisy image, the restored image and the noise vector, respectively. In our algorithms, the stopping criterion is

$$\text{RE} = \frac{\|\boldsymbol{f}^{(k-1)} - \boldsymbol{f}^{(k)}\|_2}{\|\boldsymbol{f}^{(k-1)}\|_2} \leq 10^{-3}$$

We use the output of the ForWaRD restoration method [22] as an initial image. Trial by error for α_1 and α_2 is used in the tests. In the figures, we report the highest SNR of the restored image over all tested values of α_1 and α_2. Our codes are written in Matlab. The computaational results by ForWaRD were generated by using the software "ForWarRD" which can be downloaded from http://www.dsp.rice.edu. Research codes accompanying this work will be made availabe at http://wenyouwei.googlepages.com/softwares.

Similar to [13, 22], we employ the Haar wavelet shrinkage denoising methods. The shrinkage methods include soft thresholding method [12], hard thresholding method [21] and Garrote thresholding method [16]. The program for TV denoising method is based on the code [1] written by Vogel [32]. A very small parameter (1×10^{-6}) is set in the total variation term to avoid singularities in the calculation, for details see [32].

In the first experiment, we use a synthetic 128×128 image as shown as in Figure 4.1. The image is blurred by a 7×7 point box-car blur [1, 22] defined as follows:

$$h(i,j) = \begin{cases} \frac{1}{q^2}, & 0 \leq i,j \leq q-1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

[1] http://www.math.montana.edu/∼vogel/Book/

TABLE 4.1
The Blurs, the variances of noises and the original images

	Blur	Noise Variance	Image	Size
Example 1	Blur in (4.1) ($q = 9$)	0.31	cameraman	256×256
Example 2	Blur in (4.2)	2	cameraman	256×256
Example 3	Blur in (4.2)	8	cameraman	256×256
Example 4	Blur in (4.3)	49	lena	256×256

with $q = 7$. White Gaussian noises are added to the blurred image. The blurred and noisy images are shown as in Figures 4.2 (left) and 4.2 (right) for BSNR = 20dB and BSNR = 30dB respectively.

Figures 4.2 and 4.12 show the restored images by using ForWaRD [22], the wavelet regularization in (1.2) [13] and the TV regularization in (1.2) for the restoration of blurred noisy images with BSNR = 20dB and BSNR = 30dB respectively. Figures 4.3 and 4.13 show the restored images by using Algorithm 1 with hard thresholding, soft thresholding and Garrote thresholding for the restoration of blurred noisy images with with BSNR = 20dB and BSNR = 30dB respectively. Figures 4.4 and 4.14 show the restored image by using Algorithm 2 for the blurred noisy images with with BSNR = 20dB and BSNR = 30dB respectively. Figures 4.5 and 4.15 shows the restored images by using Algorithm 3 with hard thresholding, soft thresholding and Garrote thresholding for the restoration of blurred noisy images with with BSNR = 20dB and BSNR = 30dB respectively. In Figures 4.6– 4.8 and 4.16– 4.18, we compare the restoration results in iteration number vs. SNRs. In Figures 4.9– 4.11 and 4.19– 4.21, we compare the restoration results in iteration number vs.relative errors (RE). We find in Figures 4.13– 4.15 and 4.3– 4.5 that the proposed algorithms can produce better restored images in visual quality and signal-to-noise ratio than those by the restoration methods using the combination of a data-fitting term and a regularization term together.

In the second experiment, we test the proposed algorithms and compare with other restoration methods using several different blurs studied in [13]:

$$h(i,j) = \begin{cases} \frac{1}{1+i^2+j^2}, & -7 \le i,j \le 7, \\ 0, & \text{otherwise;} \end{cases} \tag{4.2}$$

$$h(i,j) = [1,4,6,4,1]^T[1,4,6,4,1]/256. \tag{4.3}$$

In the tests, we set the variances σ of noises that have been used in the published papers [1, 2, 13, 14, 18, 19, 20, 22]. The detailed information about the tests is given in Table 4.1. There are several image restoration methods [1, 2, 13, 14, 18, 19, 20, 22] used for comparison. Their restoration results in their published work under the same setting are listed in Table 4.2. In particular, we show the ISNRs of the proposed algorithms and other restoration methods. The symbol '-' in the table denotes that the settings are not tested in the corresponding published work. According to Table 4.2, we observe that the performance of our proposed algorithms are quite good and are comparable to the other restoration methods in literature.

5. Concluding Remarks. We have proposed iterative algorithms for solving image restoration problems. The iterative algorithms are based on the decouple of deblurring and denoising steps in the restoration process. The main advantage of

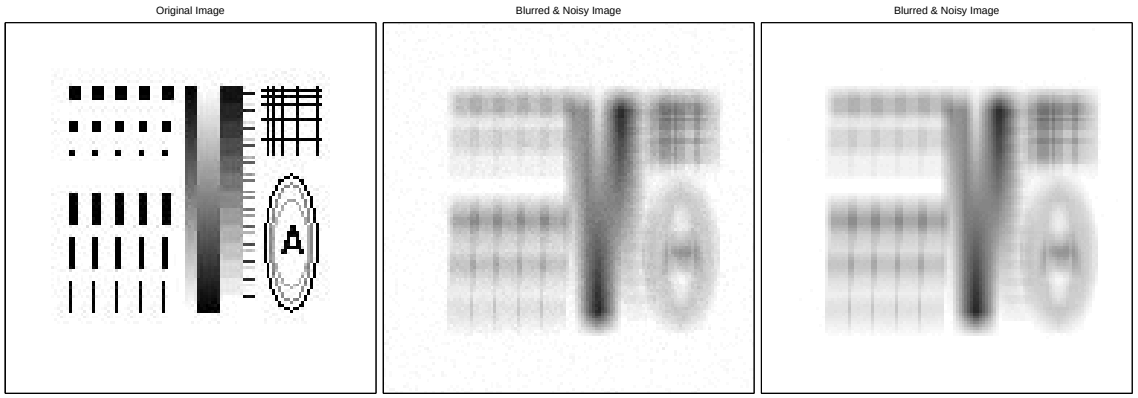

FIGURE 4.1. *The original image, and the blurred and noisy images for BSNR=20dB (left) and BSNR = 30dB (right).*

TABLE 4.2
SNR Improvement in different examples.

Method	Ex.1	Ex.2	Ex.3	Ex.4
Algorithm 1 with hard thresholding	8.46	7.53	5.64	2.90
Algorithm 1 with soft thresholding	8.57	7.53	5.56	2.81
Algorithm 1 with Garrote thresholding	8.77	7.58	5.69	2.90
Algorithm 2	8.81	7.52	5.56	2.58
Algorithm 3 with hard thresholding	8.91	7.74	5.69	3.04
Algorithm 3 with soft thresholding	8.51	7.39	5.46	2.74
Algorithm 3 with Garrote thresholding	8.87	7.60	5.69	2.91
TV [3]	8.41	-	-	2.80
ForWaRD [22]	7.34	6.72	4.95	2.26
EM [13]	7.59	6.93	4.88	2.94
BOA [14]	8.16	7.46	5.24	2.84
GEM [2]	8.10	7.40	5.15	2.85
Banham & Katsaggelos [1]	6.70	-	-	-
Jalobeanu et al. [18]	-	6.75	4.85	-
Liu & Moulin [20]	-	-	-	1.08
RI-RWI [19]	7.84	7.31	5.54	-

this proposal is that the resulting algorithms can be very efficient, and can produce better restored images in visual quality and signal-to-noise ratio than those by the restoration methods using the combination of a data-fitting term and a regularization term together. Numerical results have shown that the proposed algorithms are quite effective.

REFERENCES

[1] M. Banham and A. Katsaggelos. Spatially adaptive wavelet-based multiscale image restoration. *IEEE Transactions on Image Processing*, 5(4):619–634, 1996.

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

[3] J. Bioucas-Dias, M. Figueiredo, and R. Nowak. Total variation-based image deconvolution:

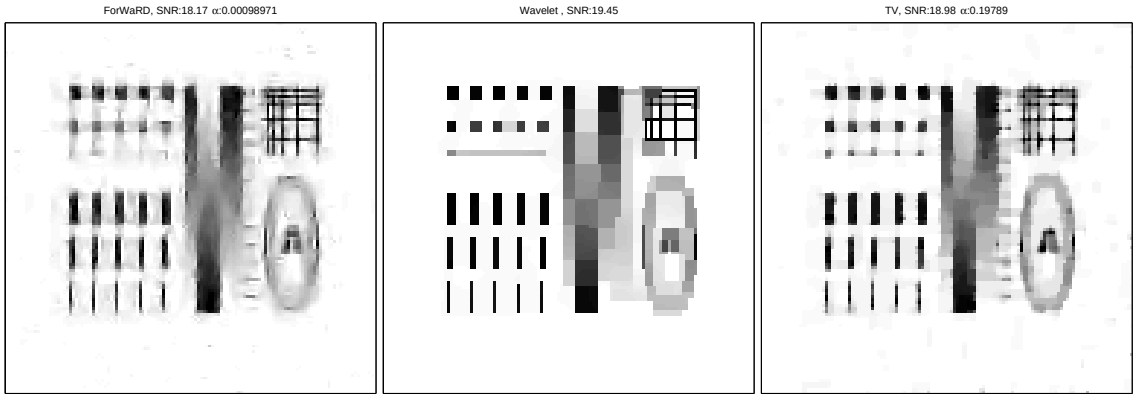

FIGURE 4.2. *The restored images by using ForWaRD [22] [SNR = 18.17dB] (left), wavelet regularization in 1.2 [13] [SNR = 19.45dB] (middle), TV regularization in 1.2 [SNR = 18.98dB] (right).*

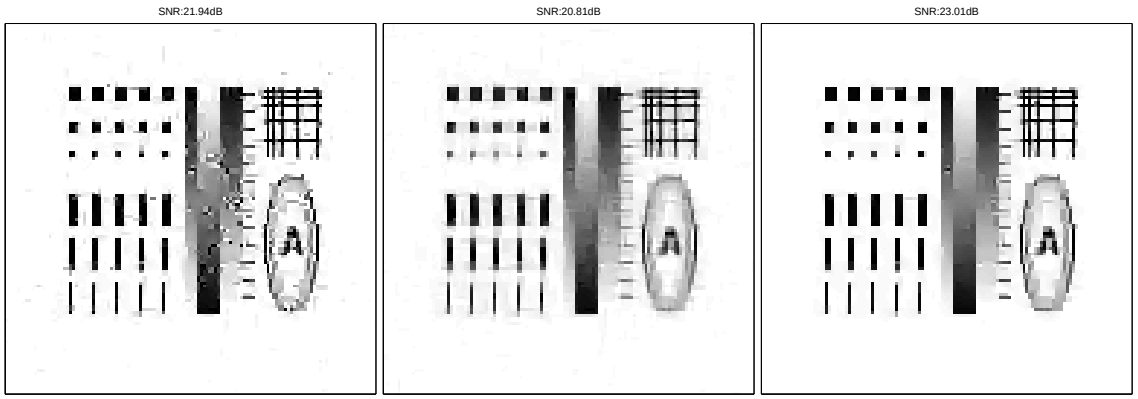

FIGURE 4.3. *The restored images by using Algorithm 1 with hard thresholding [$\alpha_1 = 0.018$, SNR = 21.94dB] (left), soft thresholding [$\alpha_1 = 0.06$, SNR = 20.81dB] (middle), Garrote shrinkage [$\alpha_1 = 0.06$, SNR = 23.01dB] (right).*

A majorization-minimization approach. In *IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP'2006*, Toulouse, France, 2006.

[4] C. Byrne. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Problems*, 20(1):103–120, 2004.

[5] A. Chambolle. An algorithm for total variation minimization and applications, *Journal of Mathematical Imaging and Vision.* , 20:89–97, 2004.

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

[7] T. Chan and P. Mulet. On the convergence of the lagged diffusivity fixed point method in total variation image restoration. *SIAM Journal on Numerical Analysis*, 36(2):354–367, 1999.

[8] P. L. Combettes. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, 53(5–6):475–504, 2004.

[9] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Journal on Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005.

[10] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57:1413–1457, 2004.

[11] D. Donoho. Nonlinear solution of linear inverse problems by wavelet-vaguelette decompositions. *Applied and Computational Harmonic Analysis*, 1:100–115, 1995.

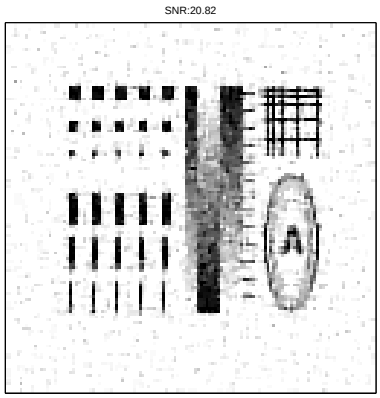

FIGURE 4.4. *The restored image by using Algorithm 2 [SNR = 20.82dB].*

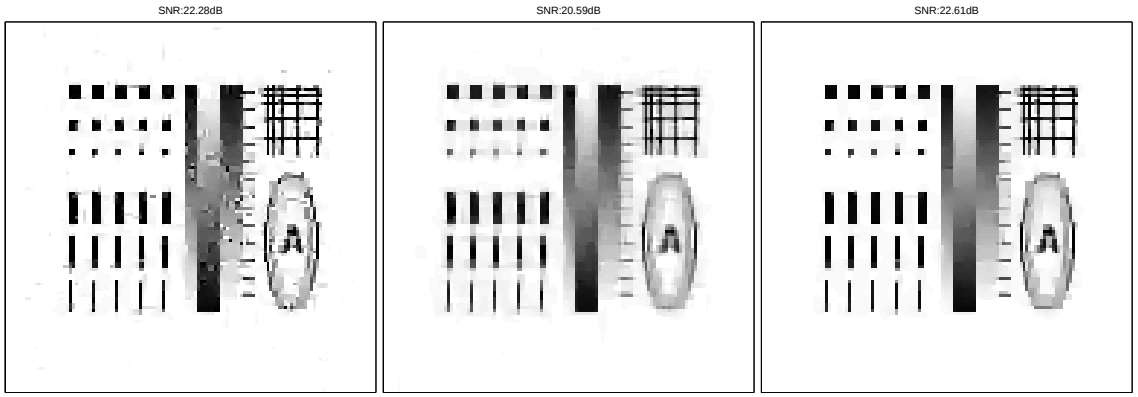

FIGURE 4.5. *The restored images by using Algorithm 3 with hard thresholding [$\alpha_1 = 0.018$, SNR = 22.28dB] (left), soft thresholding [$\alpha_1 = 0.06$, SNR = 20.59dB] (middle), Garrote shrinkage [$\alpha_1 = 0.06$, SNR = 22.61dB] (right).*

[12] D. Donoho. Denoising by soft-thresholding. *IEEE Transactions on Information Theory*, 41, 1995.

[13] M. Figueiredo and R. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Transactions on Image Processing*, 12(8), 2003.

[14] M. Figueiredo and R. Nowak. A bound optimization approach to wavelet-based image deconvolution. In *IEEE International Conference on Image Processing - ICIP'2005*, Genoa, Italy, 2005.

[15] N. Galatsanos and A. Katsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, 1, 1992.

[16] H. Gao. Wavelet shrinkage denoising using the non-negative garrote. *Jounaal of Computational and Graphical Statistics*, 7(4):469–488, 1998.

[17] R. Gopinath, M. Lang, H. Guo, and J. Odegard. Enhancement of decompressed images at low bit rates. *in SPIE Mathematical Imaging: Wavelet Applications in Signal And Image Processing*, 2303:366–377, 1994.

[18] A. Jalobeanu, N. Kingsbury, and J. Zerubia. Image deconvolution using hidden Markov tree modeling of complex wavelet packets. In *IEEE International Conference Image Processing, Thessaloniki, Greece*, 2001.

[19] V. Katkovnik, K. Egiazarian, and J. Astola. A spatially adaptive nonparametric regression image deblurring. *IEEE Transactions on Image Processing*, 14(10):1469–1478, 2005.

[20] J. Liu and P. Moulin. Complexity regularized image restoration. In *Proceedings on IEEE*

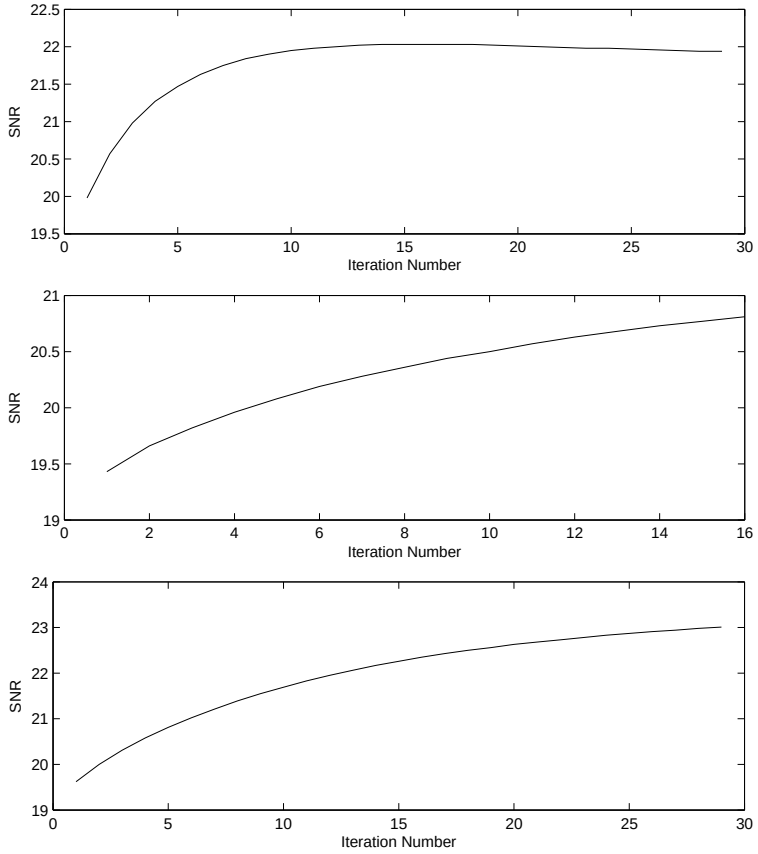

FIGURE 4.6. *xIteration number vs. SNR by using Algorithm 1 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

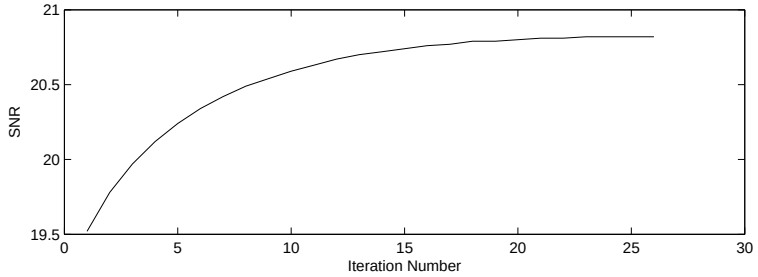

FIGURE 4.7. *Iteration number vs. SNR by using Algorithm 2 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

International Conference on Image Processing-ICIP'98, volume 1, pages 555–559, 1998.

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

[22] R. Neelamani, H. Choi, and R. Baraniuk. ForWaRD: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Transactions on Signal Processing*, 52:418–433, 2004.

[23] M. Ng. *Iterative Methods for Toeplitz Systems.* Oxford University Press, 2004.

[24] M. Ng, R. Chan, and W. Tang. A fast algorithm for deblurring models with Neumann boundary conditions. *SIAM Journal on Scientific Computing*, 21(3):851–866, 2000.

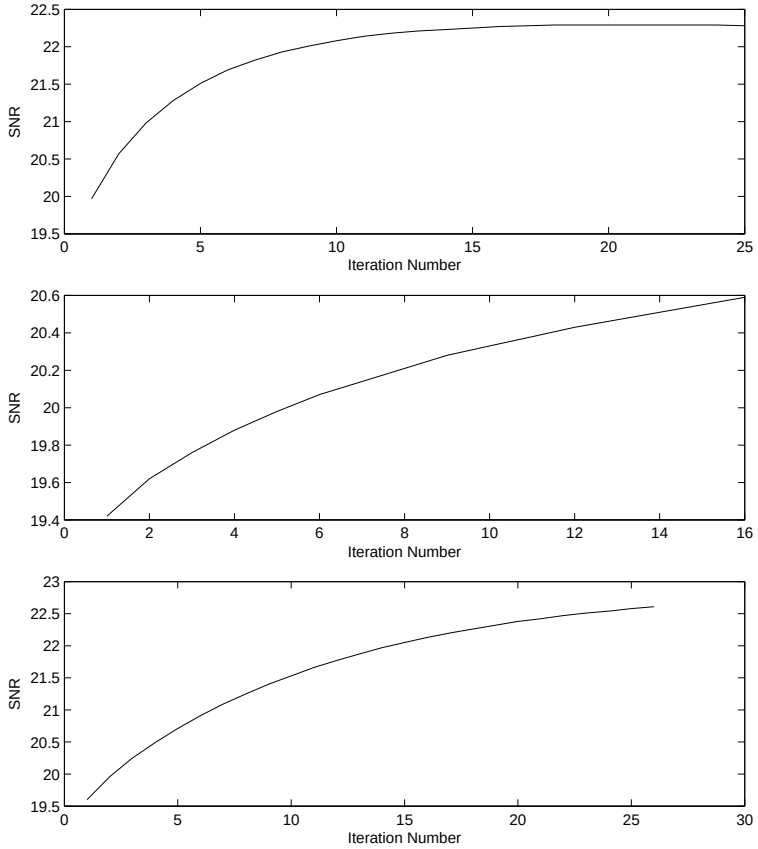

FIGURE 4.8. *Iteration number vs. SNR by using Algorithm 3 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

[25] M. Ng, L. Qi, Y. Yang and Y. Huang, On semismooth Newton's methods for total variation minimization, *Journal of Mathematical Imaging and Vision*, 27:265-276, 2007.

[26] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bullet in American Mathematical Society*, 73:591–597, 1967.

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

[28] N. Paragios, C. Chen, and O. Faugeras. *Handbook of mathematical models in computer vision*, Springer, 2006.

[29] G. Steidl, J. Weickert, T.Brox, P. Mrazek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs. *SIAM Journal on Numerical Analysis*, 42(2):686–713, 2004.

[30] D. Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8:413–424, 1986.

[31] A. Tikhonov. Solution of incorrectly formulated problems and regularization method. *Outlines Joint Sympos. Partial Differential Equations*, 4:1035–1038, 1963.

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

[33] J. B. Weaver and Y. Xu and D. M. Healy and L. D. Cromwell. *Filtering noise from images with wavelet transforms*, Magnetic Resonance in Medicine, 1991(21), pp.288-295.

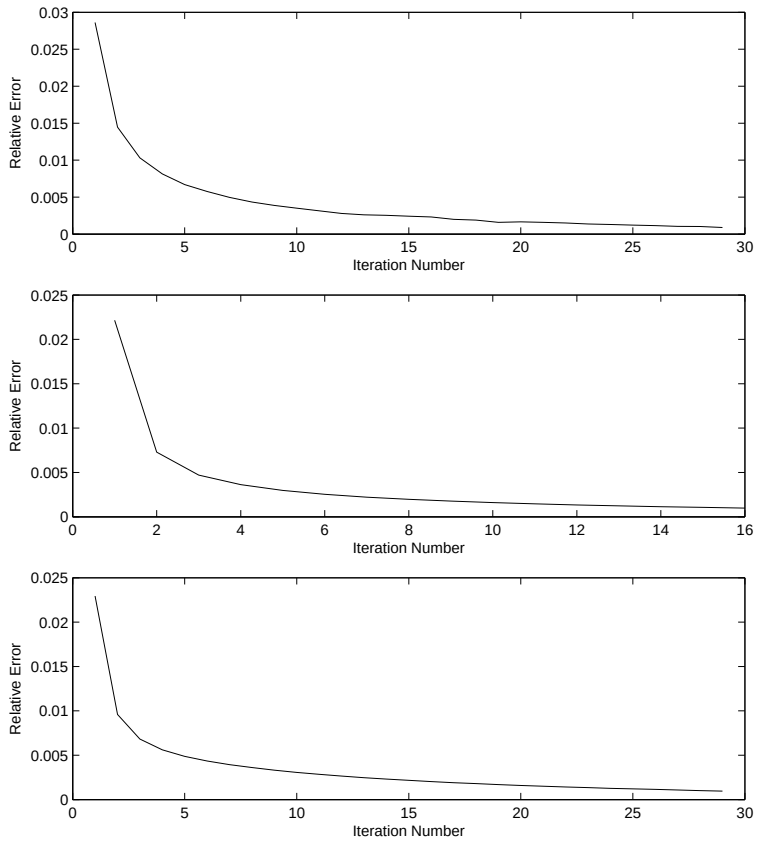

FIGURE 4.9. *Iteration number vs. relative error by using Algorithm 1 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

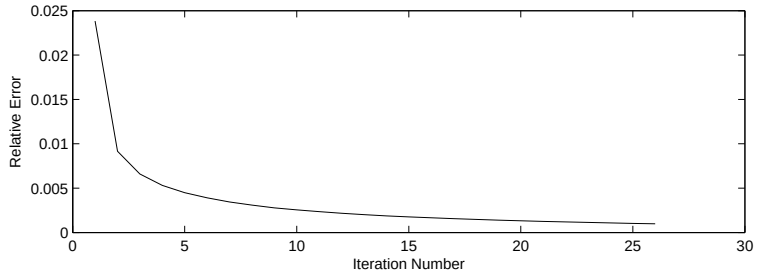

FIGURE 4.10. *Iteration number vs. relative error by using Algorithm 2 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

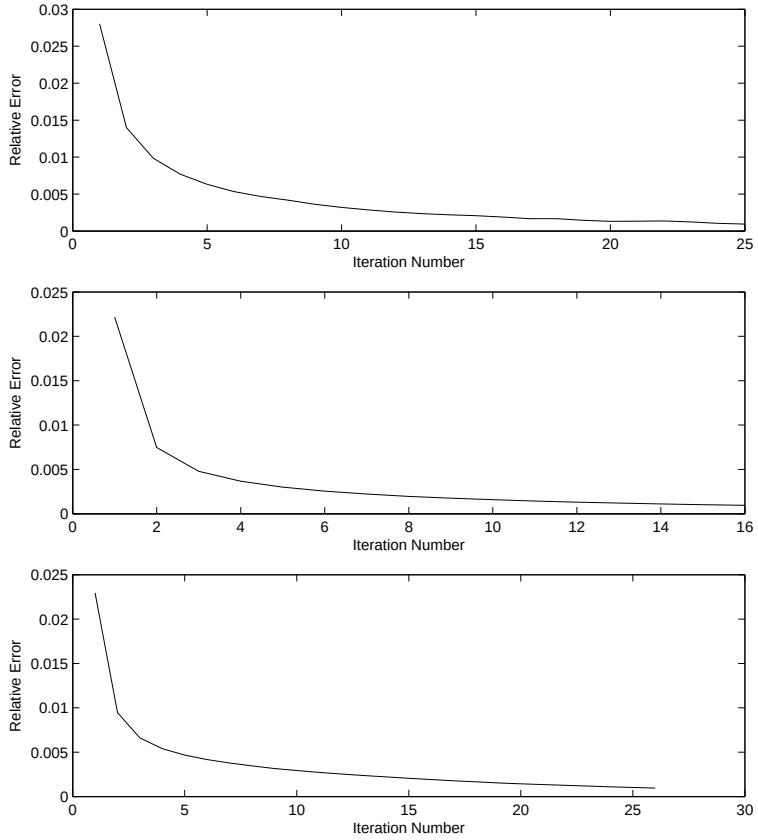

FIGURE 4.11. *Iteration number vs. relative error by using Algorithm 3 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

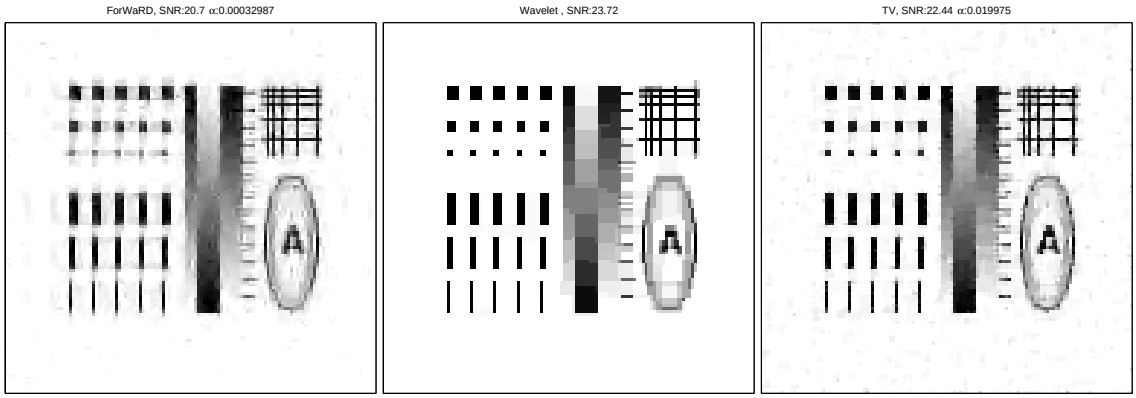

FIGURE 4.12. *The restored images by using ForWaRD [22] [SNR = 20.70dB] (left), wavelet regularization in 1.2 [13] [SNR = 23.72dB] (middle), TV regularization in 1.2 [SNR = 22.44dB] (right).*

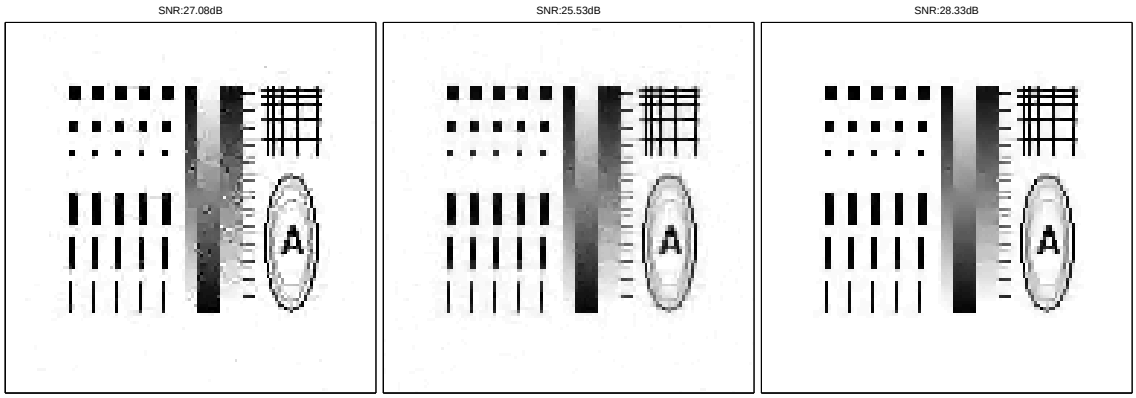

FIGURE 4.13. *The restored images by using Algorithm 1 with hard thresholding [$\alpha_1 = 0.005$, SNR = 27.08dB] (left), soft thresholding [$\alpha_1 = 0.005$, SNR = 25.53dB] (middle), Garrote shrinkage [$\alpha_1 = 0.005$, SNR = 28.33dB] (right).*

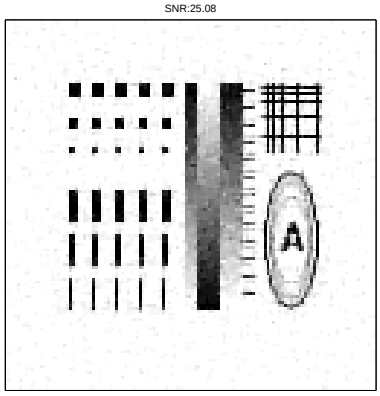

FIGURE 4.14. *The restored image by using Algorithm 2 [SNR = 25.08dB].*

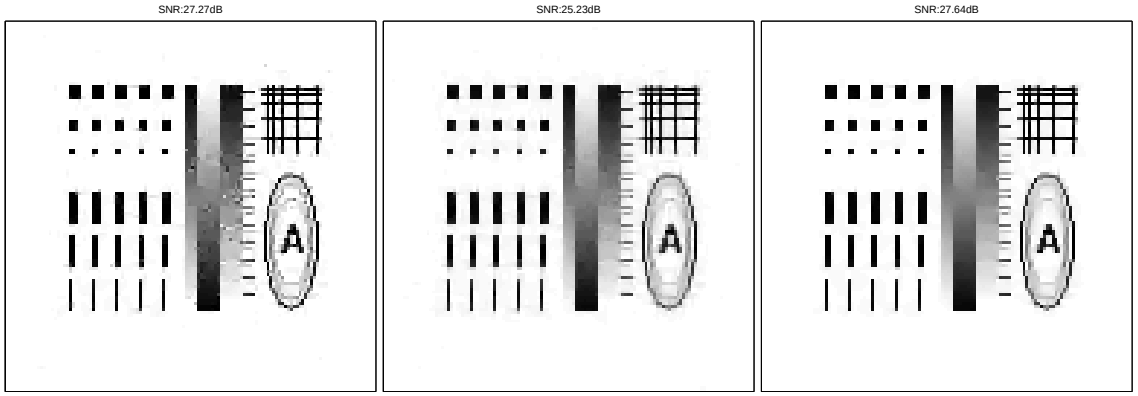

FIGURE 4.15. *The restored images by using Algorithm 3 with hard thresholding [$\alpha_1 = 0.005$, SNR = 27.27dB] (left), soft thresholding [$\alpha_1 = 0.005$, SNR = 25.23dB] (middle), Garrote shrinkage [$\alpha_1 = 0.005$, SNR = 27.64dB] (right).*

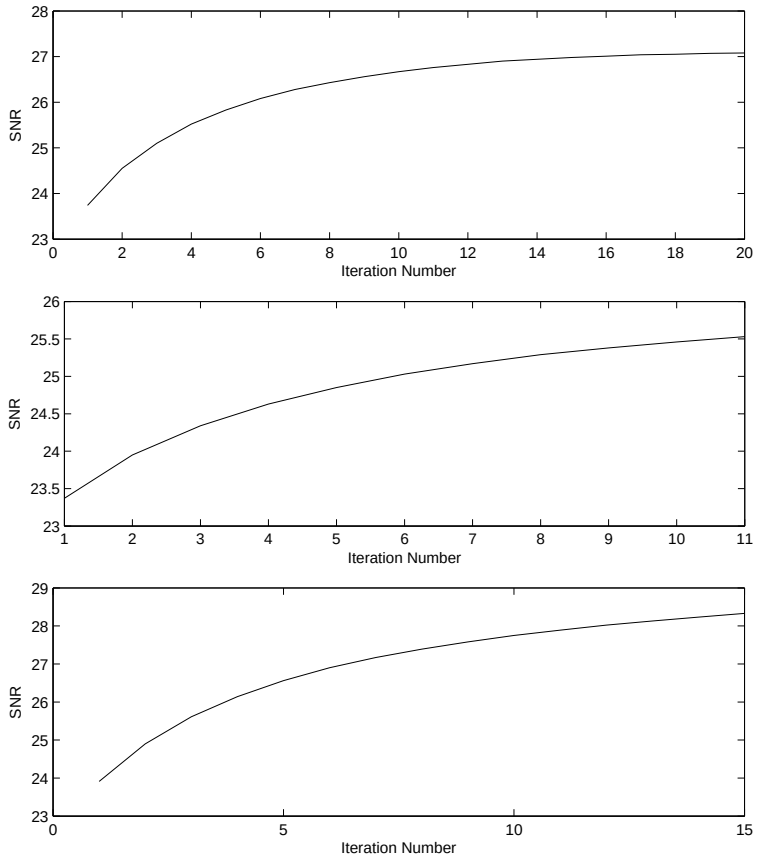

FIGURE 4.16. *Iteration number vs. SNR by using Algorithm 1 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

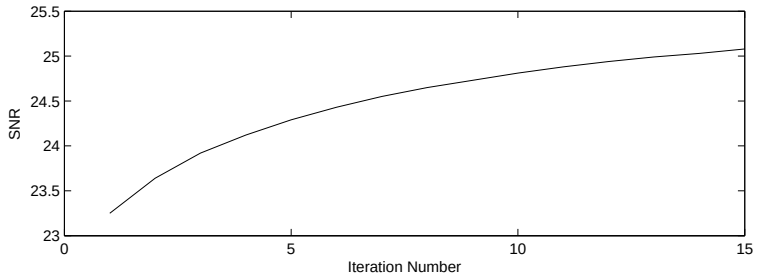

FIGURE 4.17. *Iteration number vs. SNR by using Algorithm 2 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

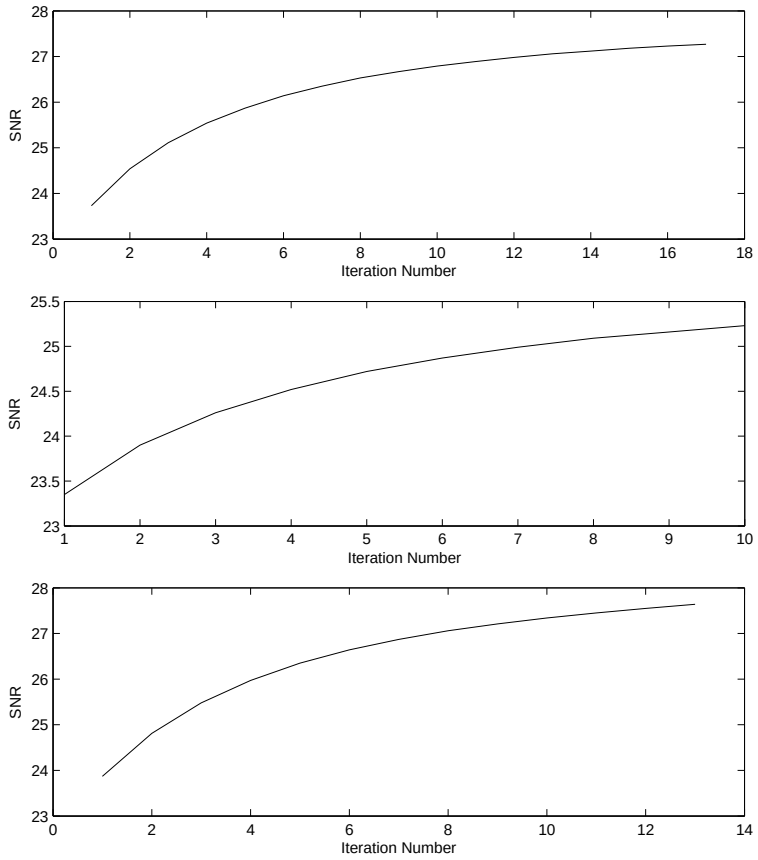

FIGURE 4.18. *Iteration number vs. SNR by using Algorithm 3 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

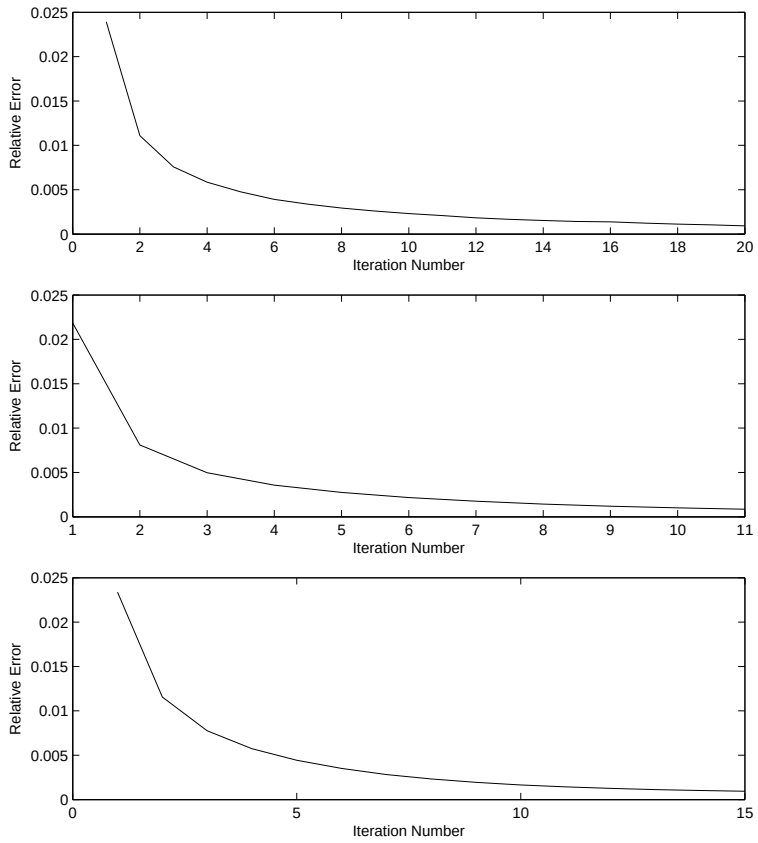

FIGURE 4.19. *Iteration number vs. relative error by using Algorithm 1 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

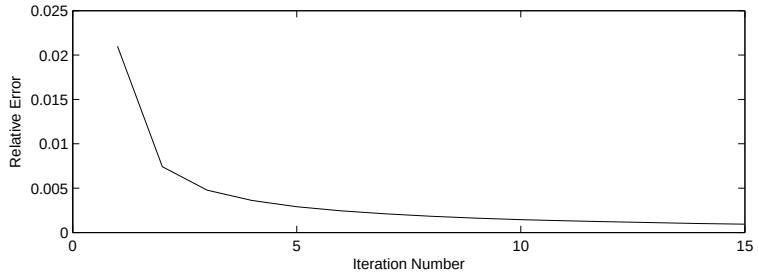

FIGURE 4.20. *Iteration number vs. relative error by using Algorithm 2 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*

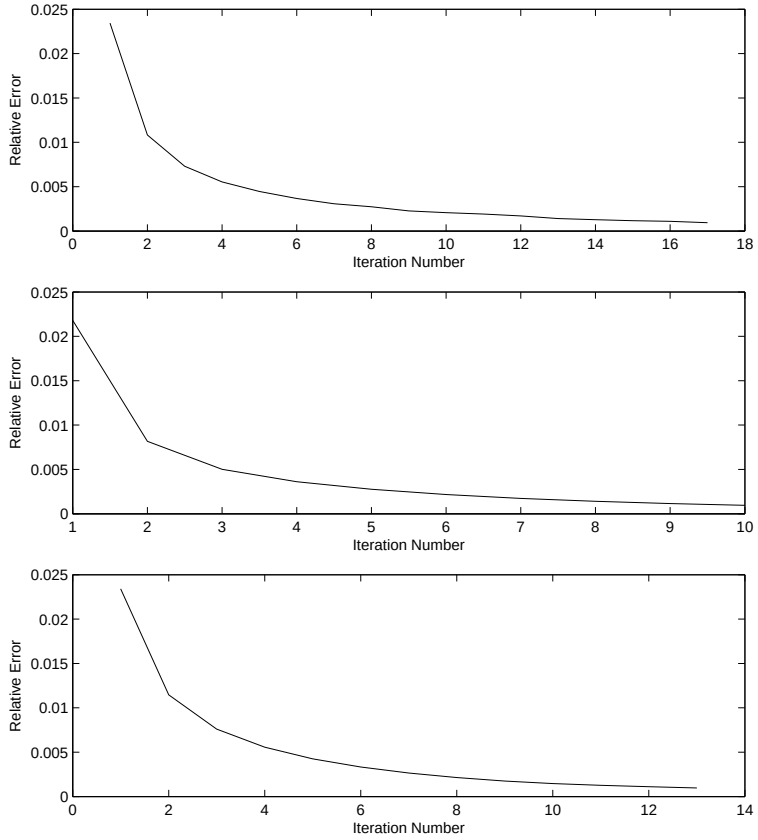

FIGURE 4.21. *Iteration number vs. relative error by using Algorithm 3 with hard thresholding (upper), soft thresholding (middle), Garrote shrinkage (bottom).*