SIAM J. SCI. COMPUT.
Vol. 35, No. 6, pp. B1304–B1320

TOTAL VARIATION STRUCTURED TOTAL LEAST SQUARES METHOD FOR IMAGE RESTORATION*

XI-LE ZHAO†, WEI WANG‡, TIE-YONG ZENG§, TING-ZHU HUANG¶, AND MICHAEL K. NG‖

Abstract. In this paper, we study the total variation structured total least squares method for image restoration. In the image restoration problem, the point spread function is corrupted by errors. In the model, we study the objective function by minimizing two variables: the restored image and the estimated error of the point spread function. The proposed objective function consists of the data-fitting term containing these two variables, the magnitude of error and the total variation regularization of the restored image. By making use of the structure of the objective function, an efficient alternating minimization scheme is developed to solve the proposed model. Numerical examples are also presented to demonstrate the effectiveness of the proposed model and the efficiency of the numerical scheme.

Key words. structured total least squares, total variation, regularization, alternating minimization, image restoration

AMS subject classifications. 65F22, 65J22, 65K10, 65K15, 68U10

DOI. 10.1137/130915406

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 an image degradation model: an ideal image f is observed in the presence of a spatial-invariant point spread function (PSF) h and an additive zero-mean Gaussian white noise η_1. Thus the observed image $\tilde{g}$ is obtained by:

$$\tilde{g} = g + \eta_1 = h \star f + \eta_1, \tag{1.1}$$

where $\star$ is the convolution operator and g is the convolution of h and f.

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 total variation (TV) regularization (ROF) model, proposed by Rudin, Osher, and Fatemi [16],

*Submitted to the journal's Computational Methods in Science and Engineering section April 2, 2013; accepted for publication (in revised form) October 1, 2013; published electronically December 5, 2013.

http://www.siam.org/journals/sisc/35-6/91540.html

†School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, 611731, People's Republic of China, and Centre for Mathematical Imaging and Vision, Hong Kong Baptist University, Hong Kong (xlzhao122003@163.com).

‡Department of Mathematics, Tongji University, Shanghai, China (weiwamng@163.com). This author's research was supported by the National Natural Science Foundation of China (11201341) and China Postdoctoral Science Foundation funded project (2012M511126 and 2013T60459).

§Department of Mathematics, Hong Kong Baptist University, Kowloon, Hong Kong (zeng@hkbu.edu.hk).

¶School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, 611731, People's Republic of China (tingzhuhuang@126.com). This author's research was supported by 973 Program (2013CB329404), NSFC (61170311, 61370147), Chinese Universities Specialized Research Fund for the Doctoral Program (20110185110020), Sichuan Province Scientific and Technical Research Project (2012GZX0080).

‖Corresponding author. Centre for Mathematical Imaging and Vision, and Department of Mathematics, Hong Kong Baptist University, Kowloon, Hong Kong (mng@math.hkbu.edu.hk). This author's research was supported by RGC GRF Grant 202013 and HKBU FRG Grant FRG/12-13/065.

has become very popular for this purpose. The main advantage of the TV formulation is the ability to preserve edges in the image due to the piecewise smooth regularization property of the TV norm. The unconstrained TV image restoration problem [5, 16] is given by

$$\min_f \ \alpha \int_\Omega (h \star f - \tilde{g})^2 dx + \int_\Omega |\nabla f| dx, \tag{1.2}$$

where Ω is the image domain, and α is a positive parameter which measures the trade off between a good fit and a regularized solution.

In classical image restoration, the PSF h is assumed to be known or adequately sampled [1]. However, in practice, one is often faced with imprecise knowledge of the PSF. In the literature, blind deconvolution methods [3] have been developed to estimate both the original image and the PSF from the observed image g. In order to obtain a reasonable restored image, these methods require one to impose suitable constraints on both the PSF and the image. TV-based blind deconvolution has been developed in [4]:

$$\min_{f,h} \frac{\tau_1}{2} \int_\Omega [h \star f - \tilde{g}]^2 dx + \frac{\tau_2}{2} \int_\Omega |\nabla h| dx + \int_\Omega |\nabla f| dx,$$

where τ_1 and τ_2 are two positive regularization parameters. In this model, the TV regularization terms of f and h are employed.

In our setting, we are interested in studying the case when the PSF is not known exactly (e.g., h is corrupted by errors):

$$\tilde{h} = h + \eta_2,$$

where η_2 is assumed to be zero-mean Gaussian white noise; see, for instance, [9, 10, 11, 14]. For instance, in two-dimensional deconvolution problems arising in ground-based atmospheric imaging, the problem consists of an image received by a ground-based imaging system, together with an image of a guide star observed under the effects of atmospheric turbulence. Empirical estimates of the PSF can sometimes be obtained by imaging a relatively bright, isolated point source. The point source might be a natural guide star or a guide star artificially generated using range-gated laser backscatter [2]. We remark that this setting is different from blind image deconvolution. In our setting, we would like to recover the original image according to the given observed image and noisy PSF.

Since the noise η_2 is added to the PSF directly, the error appearing in the corresponding blurring matrix follows the convolution structure instead of being arbitrary. In the literature, this kind of problem refers to a constrained/structured total least squares problem; see [17] for details. In [9], the maximum a posteriori estimation (MAP) technique is applied to the problem of restoring images distorted by noisy PSFs and additive noise. The steepest descent method is employed to solve the corresponding optimization problem. For image restoration problems, it has been demonstrated [6, 10, 11, 14, 15] that structured total least squares methods yield better results than the least squares method when error exists in PSFs.

Assume that $\Omega \subset R^2$ is a connected bounded set with compact Lipschitz boundary. In this paper, we study the following objective function for image restoration problems with noisy PSFs:

$$J(f, e) \equiv \frac{\alpha_1}{2} \int_\Omega [(\tilde{h} + e) \star f - \tilde{g}]^2 dx + \frac{\alpha_2}{2} \int_\Omega e^2 dx + \int_\Omega |\nabla f| dx, \tag{1.3}$$

where α_1 and α_2 are the two positive parameters to control the data-fitting term and the magnitude of error variable e in the PSF, respectively. We note that the proposed objective function can be rewritten as

$$J(f,e) = \int_\Omega \Big(\frac{\alpha_1}{2} r^2 + \frac{\alpha_2}{2} e^2\Big)\, dx + \int_\Omega |\nabla f| dx$$

subject to

$$r = (\tilde{h} + e) \star f - \tilde{g}.$$

Thus the proposed objective function is equivalent to the constrained/structured total least squares formulation with the use of the TV regularization. In the previous papers [6, 10, 11, 15], the TV regularization is not used for regularized structured total least squares (RSTLS) in image restoration.

In [14], Pruessner and O'Leary formulated this image deconvolution problem as a structured least squares problem by adding a regularization. In their algorithm, they minimize

$$\int_\Omega \Big(\frac{\alpha_1}{2} |r|^p + \frac{\alpha_2}{2} |we|^p\Big)\, dx, \tag{1.4}$$

where w is the pointwise weight to e. Correspondingly, a diagonal weighting matrix is attached to the discretized error variable of the PSF. They employ iteration methods to solve the minimization problem by linearizing (1.4). A regularization $\alpha_3 \|f\|^p$ is also added to the resulting least squares problem in each linearized iteration. In their algorithm, they considered $p = 1, 2, \infty$. When $p = 2$, an efficient solver for regularized least squares problem can be derived via fast Fourier transform. When $p = 1$ or $p = \infty$, small testing examples are studied due to the expense of solving the linear programming problems in their numerical tests. We remark that our formulation in (1.3) is different from that in [14]. We make use of the TV regularization for image restoration which is not considered in [14]. The TV regularization is important in image restoration for edge recovery. The proposed setting is that the L_1-norm is used for the regularization term $\int_\Omega |\nabla f| dx$ and the L_2-norm is used for the constrained total least squares term $\int_\Omega (\frac{\alpha_1}{2} r^2 + \frac{\alpha_2}{2} e^2) dx$. By making use of this setting, we can develop an efficient TV regularized structured least squares algorithm for image restoration problems with noisy PSFs. We also note that an interior point algorithm for L_1-norm regularization and L_2-norm data fitting was developed in [7]. However, the main computational task of such interior point algorithm is to solve a very ill-conditioned linear system. The cost of this method is significantly larger than the proposed method. Moreover, the method is applied to the convex optimization only. We note that the structured least squares problem in (1.3) or (1.4) is nonconvex.

The rest of this paper is organized as follows. In section 2, we give some theoretical results for the proposed model and present an alternating minimization scheme for solving the proposed model and discuss the convergence of the numerical scheme. In section 3, we provide numerical examples to demonstrate the effectiveness of the proposed model and the efficiency of the proposed numerical scheme. Finally, some concluding remarks are given in section 4.

2. Existence of minimizers. We first study solutions of (1.3) in the following function space. We assume the restored image pixel values are bounded in the normalized range $[0, 1]$. Let

$$\Lambda = \{(f,e) | (f,e) \in (BV(\Omega) \cap L^\infty(\Omega)) \times L^2(\Omega), 0 \le f \le 1\},$$

where $L^\infty(\Omega)$ and $L^2(\Omega)$ are the space of functions with $\|f\|_\infty < \infty$ and $\|f\|_2 < \infty$, respectively, $BV(\Omega)$ is a space of bounded variation functions. The associated minimization problem is given as follows:

$$\min_{(f,e)\in\Lambda} J(f,e). \tag{2.1}$$

Note $J(f,e)$ is convex with respect to f and e, respectively, and is not convex jointly for (f,e).

Now we show the existence result for the solution of the minimization problem in (2.1).

THEOREM 2.1. *Let $\tilde{g} \in L^\infty(\Omega)$, then the problem in* (2.1) *has at least one solution.*

Proof. First, if we let f and e be constants, the energy will be finite, which implies that the problem in (2.1) is the correct setting.

Suppose $\{(f_i, e_i)\} \subset \Lambda$ is a minimizing sequence of the problem in (2.1), then there exists a constant $M > 0$ such that

$$J(f_i, e_i) \le M.$$

The above inequality reads as

$$\frac{\alpha_1}{2}\int_\Omega (\tilde{h} * f_i + e_i \star f_i - \tilde{g})^2 dx + \frac{\alpha_2}{2}\int_\Omega {e_i}^2 dx + \int_\Omega |\nabla f_i| dx \le M.$$

Noting that $f_i \in L^\infty(\Omega)$ and $0 \le f_i \le 1$, then $\{f_i\}$ is uniformly bounded in $L^2(\Omega)$. Combining it with the boundedness of the TV term, we get that $\{f_i\}$ is uniformly bounded in $BV(\Omega)$. Therefore, there exists $\hat{f} \in BV(\Omega)$ such that, up to a subsequence,

$$f_i \to \hat{f} \in L^1(\Omega) \quad \text{and} \quad f_i \rightharpoonup \hat{f} \in L^2(\Omega). \tag{2.2}$$

As a consequence of the lower semicontinuity of $BV(\Omega)$,

$$\liminf_{i\to\infty} \int_\Omega |\nabla f_i|\, dx \ge \int_\Omega \left|\nabla \hat{f}\right| dx. \tag{2.3}$$

The boundedness of $\int_\Omega {e_i}^2$ guarantees that $\{e_i\}$ is uniformly bounded in $L^2(\Omega)$. We deduce that, up to a subsequence,

$$e_i \rightharpoonup \hat{e} \in L^2(\Omega). \tag{2.4}$$

Then we can obtain

$$\liminf_{i\to\infty} \int_\Omega {e_i}^2 dx \ge \int_\Omega \hat{e}^2 dx \tag{2.5}$$

by using the lower semicontinuity for the $L^2(\Omega)$ norm. Noting the boundedness of $\{f_i\}$ and the convergence results (2.2) and (2.4), $(\tilde{h} + e_i) * f_i \rightharpoonup (\tilde{h} + \hat{e}) \star \hat{f} \in L^2(\Omega)$, then we have

$$\liminf_{i\to\infty} \int_\Omega (\tilde{h} \star f_i + e_i \star f_i - \tilde{g})^2 dx \ge \int_\Omega (\tilde{h} \star \hat{f} + \hat{e} \star \hat{f} - \tilde{g})^2 dx \tag{2.6}$$

by recalling the lower semicontinuity for the $L^2(\Omega)$ norm. Combining (2.3), (2.5), and (2.6), we have

$$\min_{(f,e)\in\Lambda} J(f,e) = \liminf_{i\to\infty} J(f_i,e_i) \geq J(\hat{f},\hat{e}).$$

We can easily deduce $(\hat{f},\hat{e}) \in \Lambda$ by using the convergence results (2.2) and (2.4). This completes the proof. ∎

2.1. The alternating minimization scheme. In this subsection, we propose the numerical procedure for solving (2.1). Our idea is to use the alternating minimization scheme which is described in Algorithm 1.

ALGORITHM 1. THE ALTERNATING MINIMIZATION METHOD.

Input: $\tilde{h}$, $\tilde{g}$, α_1, α_2, and f_0.
Output: f_i.
while not converged **do**
Step 1. Given f_{i-1}, computing e_i by solving

$$\min_e \frac{\alpha_1}{2}\int_\Omega (\tilde{h}\star f_{i-1} + e\star f_{i-1} - \tilde{g})^2 dx + \frac{\alpha_2}{2}\int_\Omega e^2 dx. \tag{2.7}$$

Step 2. Given e_i, computing f_i by solving

$$\min_{0\leq f\leq 1} \frac{\alpha_1}{2}\int_\Omega (\tilde{h}\star f + e_i\star f - \tilde{g})^2 dx + \int_\Omega |\nabla f| dx. \tag{2.8}$$

end while

Next we would like to establish the convergence of Algorithm 1.

THEOREM 2.2. *Let $\{(f_i,e_i)\}$ be the sequence derived from Algorithm* 1. *Then $\{(f_i,e_i)\}$ converges to $(\hat{f},\hat{e})$ (up to a subsequence), and for any $(f,e)\in\Lambda$, we have*

$$J(\hat{f},\hat{e}) \leq J(f,\hat{e}), J(\hat{f},\hat{e}) \leq J(\hat{f},e).$$

Proof. First, we can easily deduce the following inequality from Algorithm 1,

$$J(f_{i+1},e_{i+1}) \leq J(f_i,e_{i+1}) \leq J(f_i,e_i).$$

Then $J(f_i,e_i)$ is bounded; we can find a subsequence (which is still noted as $\{(f_i,e_i)\}$), $(\hat{f},\hat{e})\in\Lambda$ in the same way as in Theorem 2.1 such that it satisfies

$$e_i \to \hat{e} \quad a.e. \text{ in } \Omega,$$
$$f_i \to \hat{f} \in L^1(\Omega) \text{ and } f_i \to \hat{f} \quad a.e. \text{ in } \Omega.$$

Recall that $J(f_i,e_i) > 0$ is bounded; then there exists $m > 0$ such that

$$m = \lim_{i\to\infty} J(f_i,e_i).$$

We can prove that $m = J(\hat{f},\hat{e})$. First,

$$m = \liminf_{i\to\infty} J(f_i,e_i) \geq J(\hat{f},\hat{e}).$$

By considering the inequalities

$$J(f_{i+1},e_{i+1}) \leq J(f_i,e_{i+1}) \leq J(f_i,e_i) \leq J(\hat{f},e_i),$$

and

$$J(f_{i+1}, e_{i+1}) \leq J(f_i, e_{i+1}) \leq J(f_i, \hat{e}),$$

we have

$$2J(f_{i+1}, e_{i+1}) \leq J(\hat{f}, e_i) + J(f_i, \hat{e}). \tag{2.9}$$

By rewriting the right-hand side of (2.9), we obtain

$$\begin{aligned} J(\hat{f}, e_i) + J(f_i, \hat{e}) = {} & \frac{\alpha_1}{2} \int_\Omega (\tilde{h} \star \hat{f} + e_i \star \hat{f} - \tilde{g})^2 dx + \frac{\alpha_2}{2} \int_\Omega {e_i}^2 dx + \int_\Omega |\nabla \hat{f}| dx \\ & + \frac{\alpha_1}{2} \int_\Omega (\tilde{h} \star f_i + \hat{e} \star f_i - \tilde{g})^2 dx + \frac{\alpha_2}{2} \int_\Omega \hat{e}^2 dx + \int_\Omega |\nabla f_i| dx. \end{aligned}$$

We note that

$$\begin{aligned} & \int_\Omega (\tilde{h} \star \hat{f} + e_i \star \hat{f} - \tilde{g})^2 dx + \int_\Omega (\tilde{h} \star f_i + \hat{e} \star f_i - \tilde{g})^2 dx \\ & = \int_\Omega (\tilde{h} \star \hat{f} + \hat{e} \star \hat{f} - \tilde{g})^2 dx + \int_\Omega (\tilde{h} \star f_i + e_i \star f_i - \tilde{g})^2 dx \\ & \quad + \int_\Omega 2(e_i - \hat{e}) \star (f_i - \hat{f}) \tilde{g} dx + \int_\Omega [(\tilde{h} + e_i) \star \hat{f}]^2 dx + \int_\Omega [(\tilde{h} + \hat{e}) \star f_i]^2 dx \\ & \quad - \int_\Omega [(\tilde{h} + e_i) \star f_i]^2 dx - \int_\Omega [(\tilde{h} + \hat{e}) \star \hat{f}]^2 dx. \end{aligned}$$

By combining with (2.9), we have

$$\begin{aligned} 2J(f_{i+1}, e_{i+1}) \leq {} & J(\hat{f}, \hat{e}) + J(f_i, e_i) + \int_\Omega [(\tilde{h} + e_i) \star \hat{f}]^2 dx + \int_\Omega [(\tilde{h} + \hat{e}) \star f_i]^2 dx \\ & - \int_\Omega [(\tilde{h} + e_i) \star f_i]^2 dx - \int_\Omega [(\tilde{h} + \hat{e}) \star \hat{f}]^2 dx \\ & + \int_\Omega 2(e_i - \hat{e}) \star (f_i - \hat{f}) \tilde{g} dx. \end{aligned}$$

Because $\tilde{h}$ is bounded and f_i and e_i are uniformly bounded in $L^2(\Omega)$, we deduce the following inequality by letting i tend to infinity and using the dominated convergence theorem:

$$2m \leq J(\hat{f}, \hat{e}) + m,$$

i.e., $m \leq J(\hat{f}, \hat{e})$, and therefore, $m = J(\hat{f}, \hat{e})$.

On the other hand, for any $f \in BV(\Omega) \bigcap L^\infty(\Omega)$ and $0 \leq f \leq 1$, we have

$$J(f_{i+1}, e_{i+1}) \leq J(f_i, e_i) \leq J(f, e_i)$$

and

$$J(f_{i+1}, e_{i+1}) \leq J(f_i, e_{i+1}) \leq J(f_i, \hat{e}).$$

By combining the above two inequalities, we obtain

$$\begin{aligned}2J(f_{i+1},e_{i+1}) \le J(f,\hat{e}) + J(f_i,e_i) + \int_\Omega [(\tilde{h}+e_i)\star f]^2 dx + \int_\Omega [(\tilde{h}+\hat{e})\star f_i]^2 dx \\ - \int_\Omega [(\tilde{h}+\hat{e})\star f]^2 dx - \int_\Omega [(\tilde{h}+e_i)\star f_i]^2 dx \\ + \int_\Omega 2(\hat{e}-e_i)\star(f-f_i)\tilde{g}dx.\end{aligned}$$

Again letting i tend to infinity and and using the dominated convergence theorem, we obtain

$$J(\hat{f},\hat{e}) \le J(f,\hat{e}).$$

Similarly, for any $e \in L^2(\Omega)$, we have

$$J(\hat{f},\hat{e}) \le J(\hat{f},e).$$

The result follows. ∎

To solve the discrete version of the e-subproblem in (2.7), we calculate e_i in the frequency domain by using fast Fourier transform:

$$e_i = \mathcal{F}^{-1}\left(\frac{\alpha_1 \mathcal{F}^{-1}(f_{i-1})\mathcal{F}(\tilde{g}-\tilde{h}\star f_{i-1})}{\alpha_1 \mathcal{F}^{-1}(f_{i-1})\mathcal{F}(f_{i-1})+\alpha_2}\right),$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transform and the inverse Fourier transform, respectively. In the discretized setting, at every iteration of Algorithm 1, the computational task of solving the e-subproblem in (2.7) involves four fast two-dimensional transforms and its complexity is $O(N^2 \log N)$ for an N-by-N restored image.

To solve the discrete version of the f-subproblem in (2.8), we use the alternating direction method of multipliers (ADMM) [13] by introducing variables w and v as follows:

$$\begin{aligned}&\min_{f,w,v} \frac{\alpha_1}{2}||\tilde{h}\star f + e_i \star f - \tilde{g}||_2^2 + ||w||_2 + \chi_K(v) \\ &\text{subject to} \qquad w = \nabla f \quad \text{and} \quad v = f,\end{aligned} \tag{2.10}$$

where

$$\chi_\kappa(f_{i,j}) = \begin{cases} 0 & \text{if } f_{i.j} \in [0,1], \\ \infty & \text{otherwise.}\end{cases}$$

The problem fits the framework of the alternating direction method by choosing

$$J_1(f) = \frac{\alpha_1}{2}||\tilde{h}\star f + e_i \star f - \tilde{g}||_2^2 \quad \text{and} \quad J_2(w,v) = ||w||_2 + \chi_K(v).$$

For the above constraints, we express them as follows:

$$Bf + Cz := \begin{bmatrix} \nabla \\ I \end{bmatrix} f - \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

By attaching the Lagrangian multiplier

$$\psi = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

to the linear constraints, the augmented Lagrangian function of (2.10) is given by

$$\mathcal{L}(f,z,\psi) = J_1(f) + J_2(z) + <\psi, Bf + Cz> + \frac{\beta}{2}\|Bf + Cz\|_2^2, \tag{2.11}$$

where $\beta > 0$ is the penalty parameter for the linear constraints to be satisfied. More precisely, the computational procedure for solving (2.10) is presented in Algorithm 2.

ALGORITHM 2. THE ALTERNATING DIRECTION METHOD OF MULTIPLIERS FOR SOLVING (2.10).

Input: $\tilde{h}$, $\tilde{g}$, α_1, β, e_i, $w^{(0)}$, $v^{(0)}$, and $\psi^{(0)}$.
Output: $f_{i+1} = f_i^{(k)}$.
while not converged **do**
Step 1. Given $w^{(k-1)}$, $v^{(k-1)}$, and $\psi^{(k-1)}$, computing $f_i^{(k)}$ by

$$f_i^{(k)} = \mathcal{F}^{-1}\left(\frac{\alpha_1 \mathcal{F}^{-1}(\tilde{h} + e_{i-1})\mathcal{F}(\tilde{g}) + \mathcal{F}^{-1}(\nabla)\mathcal{F}(\beta w^{(k-1)} - \mu_1^{(k-1)}) + \mathcal{F}(\beta v^{(k-1)} - \mu_2^{(k-1)})}{\alpha_1 \mathcal{F}^{-1}(\tilde{h} + e_{i-1})\mathcal{F}(\tilde{h} + e_{i-1}) + \beta \mathcal{F}^{-1}(\nabla)\mathcal{F}(\nabla) + \beta}\right). \tag{2.12}$$

Step 2. Given $f_i^{(k)}$ and $\mu_1^{(k-1)}$, update $w^{(k)}$ by using two-dimensional shrinkage formula as

$$w^{(k)} = \max\left\{\left\|\nabla f_i^{(k)} + \frac{\mu_1^{(k-1)}}{\beta}\right\|_2 - \frac{1}{\beta}, 0\right\} \frac{\nabla f_i^{(k)} + \frac{\mu_1^{(k-1)}}{\beta}}{\left\|\nabla f_i^{(k)} + \frac{\mu_1^{(k-1)}}{\beta}\right\|_2}. \tag{2.13}$$

Step 3. Given $f_i^{(k)}$ and $\mu_2^{(k-1)}$, update $v^{(k)}$ by

$$v^{(k)} = P_{[0,1]}\left(f_i^{(k)} + \frac{\mu_2^{(k-1)}}{\beta}\right), \tag{2.14}$$

where $P_{[0,1]}(x)$ is the projection of x on the box $[0,1]$.
Step 4. Update $\psi^{(k)}$ by

$$\psi^{(k)} = \psi^{(k-1)} + \beta(Bf_i^{(k)} - z^{(k)}). \tag{2.15}$$

end while

We observe that only two terms in the right-hand side of (2.12), $\beta w^{(k-1)} - \mu_1^{(k-1)}$ and $\beta v^{(k-1)} - \mu_2^{(k-1)}$, are changed at each iteration of Algorithm 2. In the discretized setting, the main computational task involves three fast two-dimensional transforms, and its complexity is $O(N^2 \log N)$ operations for an N-by-N restored image. We also see that every step of ADMM for solving the discrete version of (2.10) has a closed form solution. Thus the method can be efficiently implementable. The convergence of the alternating direction method for a convex objective function of separable variables with linear constraints is guaranteed; see, for instance, [8, 12]. To get an exact solution of the discrete version of (2.8), we need to choose $f_{i+1} = f_i^{(+\infty)}$. Since the estimated error e_i may not be accurate enough, in practice, it is not necessary to use an exact solution of the discrete version of (2.8). As for computational efficiency, we implement an inexact version of Algorithm 2 which performs several iterations only. Numerical examples will be given in the next section to demonstrate the efficiency of the inexact version of Algorithm 2.

It is clear that discrete versions of subproblems in (2.8) and (2.7) can be solved efficiently. Moreover, the total cost of computing f and e at Step 1 and Step 2 in Algorithm 1 is $O(N^2 \log N)$ for an N-by-N restored image by the proposed total variation regularized structured total least squares (TV-RSTLS) method.

In [10, 11], it has been shown that the use of Tikhonov regularization in the RSTLS problem for image restoration can be decoupled into N^2 one-variable nonlinear equations, where each equation can be solved efficiently. However, when the TV regularization is employed instead of Tikhonov regularization, the problem cannot be reduced to N^2 one-variable nonlinear equations. More precisely, their approach will lead to solving a sequence of N^2-by-N^2 non-Toeplitz linear systems. Therefore, the corresponding computational cost would be very expensive.

3. Numerical experiments. In this section, we present numerical results to demonstrate the performance of the proposed model. The quality of recovered images is measured by the relative error (ReErr), the signal-to-noise ratio (SNR), and the highest peak signal-to-noise ratio (PSNR). They are defined as

$$\text{ReErr} = \frac{||f - f_c||_2}{||f||_2}, \quad \text{SNR} = 10 \log_{10} \frac{\left\|f - \bar{f}\right\|_2^2}{||f - f_c||_2^2},$$

and

$$\text{PSNR} = -20 \log_{10} \frac{||f - f_c||_2}{N},$$

where f, $\bar{f}$ and f_c are the original image, the mean intensity value of f, and the restored image, respectively.

In the tests, for a given PSF h, we generate $\tilde{h}$ by adding a Gaussian white noise with zero mean and standard derivation (std) to h: $\tilde{h} = h + \eta_2$. The blurred and noisy image $\tilde{g}$ is generated by $h \star f + \eta_1$, where f is the original image and η_1 is also a Gaussian white noise with zero mean. We compare the proposed model (TV-RSTLS) in (1.3) with the following four models.

- ROF model with the exact PSF (ROF-exact):

$$\min_f \frac{\alpha_1}{2} \int_\Omega (h \star f - \tilde{g})^2 dx + \int_\Omega |\nabla f| dx \quad \text{subject to} \quad 0 \le f \le 1.$$

  This is the reference model where there is no noise in the PSF.
- ROF model with the noisy PSF (ROF-noisy):

$$\min_f \frac{\alpha_1}{2} \int_\Omega (\tilde{h} \star f - \tilde{g})^2 dx + \int_\Omega |\nabla f| dx \quad \text{subject to} \quad 0 \le f \le 1.$$

  This is the model where the structured total least squares method is not used.
- Chan–Wong model [4]:

$$\min_{f,h} \frac{\tau_1}{2} \int_\Omega [h \star f - \tilde{g}]^2 dx + \frac{\tau_2}{2} \int_\Omega |\nabla h| dx + \int_\Omega |\nabla f| dx \quad \text{subject to} \quad 0 \le f \le 1.$$

  This is the blind deconvolution model where the noisy PSF is used for an initial guess.
- Tikhonov-RSTLS:

$$\min_{f,e} \frac{\alpha_1}{2} \int_\Omega [(\tilde{h}+e) \star f - \tilde{g}]^2 dx + \frac{\alpha_2}{2} \int_\Omega e^2 dx + \int_\Omega |\nabla f|^2 dx \quad \text{subject to} \quad 0 \le f \le 1.$$

  This is the structured total least squares model with Tikhonov regularization.

For the ROF-exact and ROF-noisy models, we use the alternating direction method [13] to obtain the restored image. For the blind deconvolution models, we employ an alternating minimization scheme [4] to solve for f. As for the comparison, we make use of the noisy PSF with exact support as an initial guess in the blind deconvolution algorithm. We note that the Tikhonov-RSTLS model is nonconvex. For the Tikhonov-RSTLS model, we use Algorithms 1 and 2 except we deal with the term $|\nabla f|^2$ instead of $|\nabla f|$ in the restoration process. For the Chan–Wong model, we use the alternating minimization method to obtain the restored image and the estimated PSF and each subproblem of the Chan–Wong model is solved by the alternating direction method. In all the algorithms, the stopping criterion is measured on the relative error of the successive restored image given as follows:

$$\frac{\|f_{i-1} - f_{i-2}\|_2}{\|f_{i-1}\|_2} \leq \epsilon,$$

where ϵ is a given tolerance. In all the tests, the stopping tolerance is set to be 10^{-3} and β is set to be 10. Several values of regularization parameters, $\alpha_1 = 10^i$ $(i = 1, \ldots, 4)$ and $\alpha_2 = 10^j$ $(j = 0, \ldots, 10)$ are tested, and the restoration results with the smallest ReErr among the tested values are reported. All the tests are performed under Windows 7 and MATLAB Version 7.10 (R2010a) running on a desktop with an Intel(R) Core(TM) 2 Quad CPU at 2.66 GHz and 4 GB of memory.

3.1. Test 1. We test the 256×256 "Cameraman" image in Figure 3.1(a). We employ the Gaussian kernel (MATLAB command: `fspecial('Gaussian',[21 21],1.5)`) and an std of 5×10^{-3} to generate a noisy PSF. The exact PSF and the observed noisy PSF are shown in Figure 3.2. A Gaussian white noise of 40 dB is further added to the blurred image. The corresponding blurred and noisy "Cameraman" image is shown in Figure 3.1(b) with SNR = 11.45 dB. Similarly, we test the 256×256 "Bar" image in Figure 3.3(a) and consider the Moffat kernel (MATLAB command: `psfMoffat([23 23],3,5)` which is a function in the HNO package[1]) and an std of 5×10^{-3} to generate a noisy PSF. The exact PSF and the observed noisy PSF are shown in Figure 3.4. A Gaussian white noise of 40 dB is further added to the blurred image. The corresponding blurred and noisy "Bar" image is shown in Figure 3.3(b) with SNR = 9.22 dB. Here we set the number of iterations to be three in Algorithm 2 and the initial guess to be the observed image. The ROF-exact, ROF-noisy, and Tikhonov-RSTLS methods use the same initial guess for a fair comparison.

In Figures 3.1(c)–(g) and 3.3(c)–(g), we display the restored images by different methods. It is clear from the figures the TV-RSTLS method restores better images than the Tikhonov-RSTLS, ROF-noisy, and Chan–Wong methods. Also the SNRs of the restored images by the TV-RSTLS are higher than those by the Tikhonov-RSTLS, ROF-noisy, and Chan–Wong methods. Moreover, the zoom-in regions of recovered images are further shown in Figures 3.5 and 3.6 to illustrate clearly the superiority of the proposed TV-RSTLS method. It is also clear that the edges restored by the TV-RSTLS method are much better than those by the Tikhonov-RSTLS method.

3.2. Test 2. In this test, we evaluate the performance of the proposed method by using different PSFs such as Gaussian kernel, Moffat kernel, and truncated nonsymmetric kernel Nonsym.[2] Their MATLAB commands are `fspecial('gaussian',`

[1] http://www2.imm.dtu.dk/~pch/HNO/.

[2] RestoreTools Package, http://www.mathcs.emory.edu/~nagy/RestoreTools/index.html.

FIG. 3.1. (a) *The original image,* (b) *the blurred and noisy image* ($SNR = 11.45$ *dB*), (c) *ROF-exact* ($SNR = 15.48$ *dB,* $\alpha_1 = 10^4$). *The restored images by* (d) *Tikhonov-RSTLS* ($SNR = 12.43$ *dB,* $\alpha_1 = 10^3$, *and* $\alpha_2 = 10^5$); (e) *ROF-noisy* ($SNR = 4.53$ *dB,* $\alpha_1 = 10^3$); (f) *Chan–Wong* ($SNR = 12.18$ *dB,* $\alpha_1 = 10^3$, *and* $\alpha_2 = 10^2$); *and* (g) *TV-RSTLS* ($SNR = 12.84$ *dB,* $\alpha_1 = 10^3$, *and* $\alpha_2 = 10^5$).

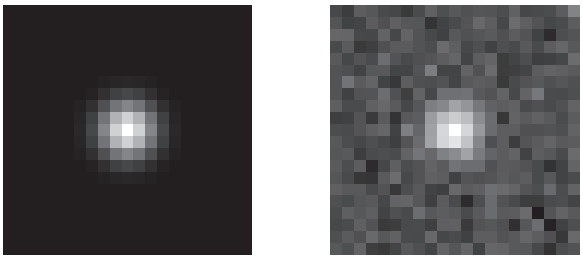


FIG. 3.2. *The exact PSF (left) and observed noisy PSF (right).*

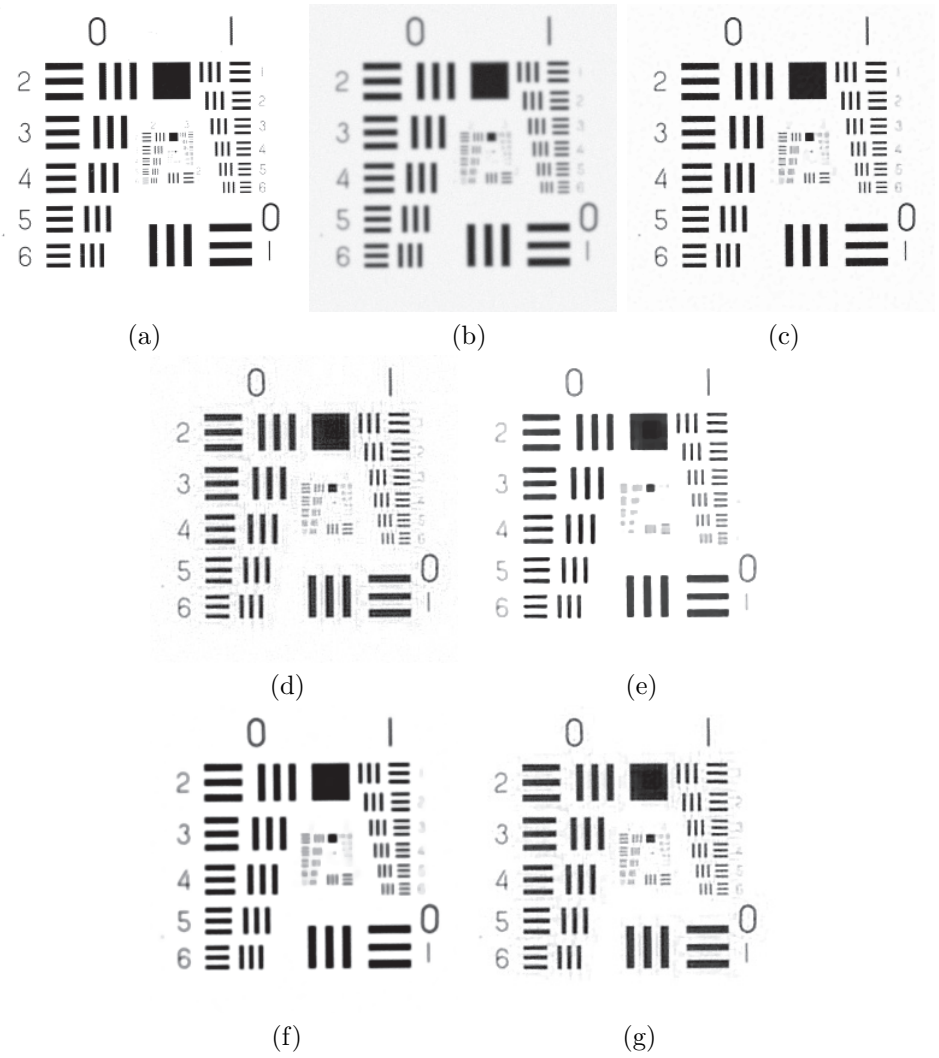


FIG. 3.3. (a) *The original image,* (b) *the blurred and noisy image* ($SNR = 9.22$ *dB*), (c) *ROF-exact* ($SNR = 20.21$ *dB,* $\alpha_1 = 10^3$). *The restored images by* (d) *Tikhonov-RSTLS* ($SNR = 12.44$ *dB,* $\alpha_1 = 10^3$, *and* $\alpha_2 = 10^6$), (e) *ROF-noisy* ($SNR = 10.83$ *dB,* $\alpha_1 = 10^2$), (f) *Chan–Wong* ($SNR = 11.65$ *dB,* $\alpha_1 = 10^2$, *and* $\alpha_2 = 10^1$) *and* (g) *TV-RSTLS* ($SNR = 13.51$ *dB,* $\alpha_1 = 10^3$, *and* $\alpha_2 = 10^6$).

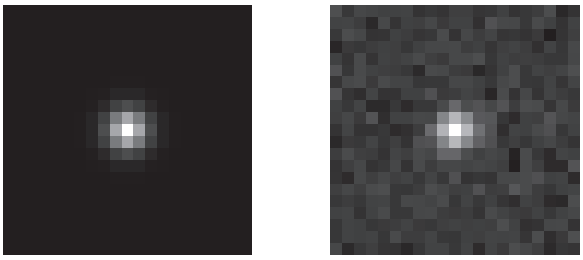


FIG. 3.4. *The exact PSF (left) and observed noisy PSF (right).*

FIG. 3.5. *The zoomed-in regions of the restored images in Figure* 3.1. (a) *The original image,* (b) *the blurred and noisy image,* (c) *ROF-exact. The restored images by* (d) *Tikhonov-RSTLS,* (e) *ROF-noisy,* (f) *Chan–Wong, and* (g) *TV-RSTLS.*

`[hsize,hsize],1.5)` and `psfMoffat([hsize,hsize],3,5)`.[3] Several values of hsize are studied, which is used to control the support PSFs. Different values of variance of the noise in PSFs are considered. The "Cameraman" image is used to test the performance of the proposed method. A Gaussian white noise of 40 dB is further added to the blurred image. Here we set the number of iterations to be three in Algorithm 2, and the initial guess to be the observed image. The restoration results in ReErr, SNR, PSNR, and computational times are shown in Table 3.1. In each entry of the table, the number is the average value of the 20 random noise cases in PSFs and blurred images of the corresponding setting. We see from the table that the performance of the proposed model (the numbers in bold face) is better than those of the Tikhonov-RSTLS model, the ROF-noisy model, and the Chan–Wong model.

[3]HNO package, http://www2.imm.dtu.dk/∼pch/HNO/.

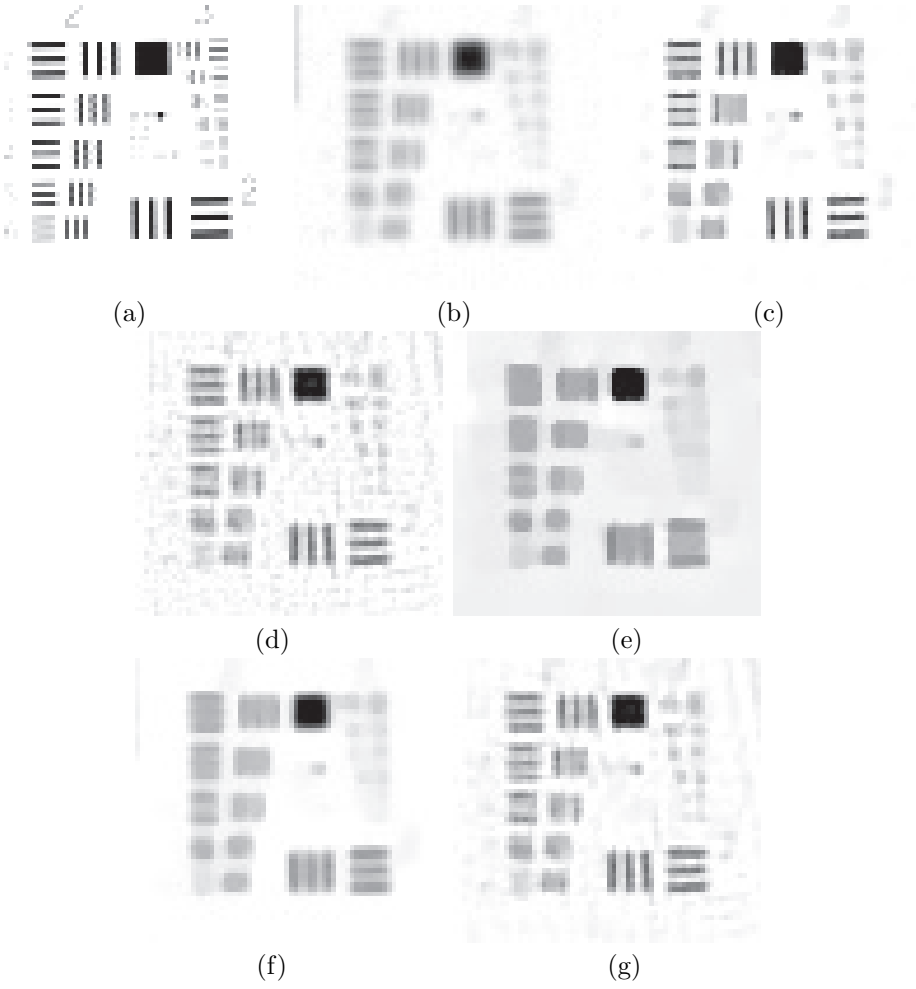


FIG. 3.6. *The zoomed-in regions of the restored images in Figure* 3.3. (a) *The original image,* (b) *the blurred and noisy image,* (c) *ROF-exact. The restored images by* (d) *Tikhonov-RSTLS,* (e) *ROF-noisy,* (f) *Chan–Wong, and* (g) *TV-RSTLS.*

3.3. Test 3. In this example, we study the effect of the parameters in the proposed model. We test their values as follows: $\alpha_1 = 10^i$ for $(i = 1, \ldots, 4)$ and $\alpha_2 = 10^j$ for $(j = 0, 1, \ldots, 10)$. We remark that when α_1 is larger than 10^4 or smaller than 10, the restoration results are not good. In this experiment, we consider the "Cameraman" images blurred by a Gaussian PSF (hsize = 21 and $\sigma = 1.5$) with a Gaussian white noise of mean zero and std 5×10^{-3}. The Gaussian white noise of 40 dB is further added to the blurred image. Here we set the number of iterations to be three in Algorithm 2, and the initial guess to be the observed image. The summary of the restoration results based on SNRs is shown in Table 3.2. We observe from Table 3.2 that the best restoration results are obtained when α_2 is not too small or too large. For illustration, we show the plots of the SNRs with respect to different values of α_2 in Figure 3.7. When α_2 is large, the variable e is not effective and thus it behaves like the ROF-noisy model. On the other hand, when α_2 is small and the initial guess is the observed image, the PSF error variable e is approximately $1 - \tilde{h}$ and thus it

TABLE 3.1

The average results of the TV-RSTLS model compared with the Tikhonov-RSTLS model, the ROF-noisy model, and the Chan–Wang model where 40 dB Gaussian noise is added to the blurred image.

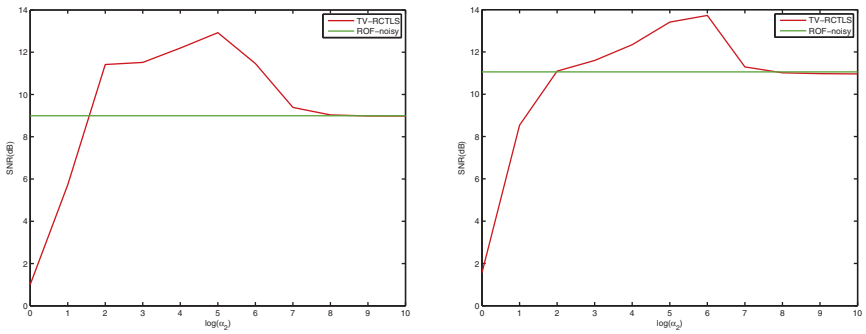
			Tikhonov–RSTLS				ROF-noisy			
PSF	std	hsize	ReErr	SNR	PSNR	Time	ReErr	SNR	PSNR	Time
Gaussian	0.005	19	0.1107	12.48	24.72	4.55	0.1430	10.88	23.11	0.51
		23	0.1138	12.24	24.47	6.14	0.1778	9.06	21.30	0.52
		27	0.1147	12.16	24.40	7.76	0.1811	8.85	21.08	0.46
	0.0025	19	0.0999	13.37	25.61	3.58	0.1039	13.16	25.39	0.63
		23	0.0998	13.37	25.60	4.28	0.1150	12.31	24.55	0.63
		27	0.1016	13.21	25.45	4.64	0.1166	12.18	24.42	0.59
Moffat	0.005	19	0.0868	14.61	26.84	3.07	0.1229	12.57	24.80	0.51
		23	0.0873	14.53	26.77	5.54	0.1616	10.19	22.43	0.42
		27	0.0900	14.26	26.50	6.33	0.1653	9.90	22.13	0.49
	0.0025	19	0.0720	16.20	28.44	3.32	0.0799	15.61	27.84	0.54
		23	0.0738	15.98	28.22	3.78	0.0935	14.29	26.53	0.50
		27	0.0759	15.75	27.98	4.00	0.0953	14.11	26.34	0.49
Nonsym	0.005	19	0.1392	10.52	22.75	17.53	0.1644	9.47	21.70	0.52
		23	0.1663	8.96	21.20	29.79	0.1726	8.69	20.93	0.47
		27	0.1668	8.95	21.18	13.39	0.2153	7.08	19.31	0.52
	0.0025	19	0.1226	11.62	23.86	16.61	0.1212	11.76	24.00	0.57
		23	0.1272	11.26	23.50	15.11	0.1357	10.80	23.03	0.58
		27	0.1356	10.72	22.96	24.65	0.1433	10.31	22.54	0.52
			Chan–Wong				TV-RSTLS			
PSF	std	hsize	ReErr	SNR	PSNR	Time	ReErr	SNR	PSNR	Time
Gaussian	0.005	19	0.1106	12.48	24.72	7.94	**0.1063**	**12.83**	**25.07**	3.01
		23	0.1144	12.18	24.42	5.97	**0.1097**	**12.55**	**24.79**	3.73
		27	0.1146	12.16	24.40	5.08	**0.1125**	**12.37**	**24.60**	3.87
	0.0025	19	0.1039	13.16	25.39	0.63	**0.0973**	**13.59**	**25.82**	2.54
		23	0.1150	12.31	24.55	0.63	**0.0983**	**13.51**	**25.74**	3.27
		27	0.1166	12.18	24.42	0.59	**0.0987**	**13.47**	**25.70**	3.42
Moffat	0.005	19	0.0923	14.05	26.28	5.21	**0.0821**	**15.08**	**27.31**	3.06
		23	0.0927	14.00	26.24	4.85	**0.0830**	**15.01**	**27.24**	3.33
		27	0.0933	13.95	26.18	4.35	**0.0887**	**14.43**	**26.66**	3.13
	0.0025	19	0.0799	15.61	27.84	0.54	**0.0713**	**16.29**	**28.52**	2.48
		23	0.0935	14.29	26.53	0.50	**0.0717**	**16.26**	**28.50**	3.11
		27	0.0953	14.11	26.34	0.49	**0.0738**	**16.00**	**28.24**	3.02
Nonsym	0.005	19	0.1644	9.47	21.70	0.52	**0.1355**	**10.75**	**22.99**	8.55
		23	0.1726	8.69	20.93	0.47	**0.1575**	**9.46**	**21.69**	9.21
		27	0.1950	7.55	19.79	10.07	**0.1598**	**9.32**	**21.55**	18.65
	0.0025	19	0.1212	11.76	24.00	0.57	**0.1165**	**12.03**	**24.26**	8.18
		23	0.1357	10.80	23.03	0.58	**0.1236**	**11.51**	**23.74**	9.94
		27	0.1433	10.31	22.54	0.52	**0.1320**	**10.94**	**23.18**	13.55

behaves like the TV denoising model. We also remark that the above observations are also valid for the other PSFs.

3.4. Test 4. As we have mentioned, $J(f, e)$ is convex with respect to f and e separately, but is not jointly convex for (f, e). To study the sensitivity of the proposed model to initial guesses, we solve the proposed model by the alternating minimization scheme with different initial guesses f_0 including a black image, a white image, and a random image and the observed image. In Table 3.3, we show the average restoration results for 20 cases of the "Cameraman" image blurred by a noisy Gaussian PSF with std $= 0.005$, and further corrupted by a Gaussian noise of 40 dB. We see from Table 3.3 that the observed image as the initial guess gives more information and thus yields better results compared with other initial guesses. As a reference, we also show the corresponding restoration results by the ROF-noisy model using the observed image

TABLE 3.2
The behavior of the proposed model with different parameters for the noise level 40 dB.

		α_2										
Std	α_1	1	10^1	10^2	10^3	10^4	10^5	10^6	10^7	10^8	10^9	10^{10}
0.005	10	−5.83	−4.50	−3.73	−2.32	0.53	3.56	6.64	8.24	8.62	8.72	8.74
	10^2	5.59	2.15	−3.44	9.96	6.46	11.98	10.32	9.04	8.78	8.74	8.73
	10^3	1.00	5.73	11.42	11.52	12.20	**12.63**	11.46	9.39	9.03	8.98	8.98
	10^4	11.46	11.45	11.45	11.46	11.82	12.17	11.85	9.69	8.62	8.41	8.38
0.0025	10	−5.71	−4.55	−3.72	−2.34	0.38	3.36	6.40	8.00	8.42	8.51	8.54
	10^2	2.60	−1.59	8.80	10.45	11.52	12.18	11.34	10.47	10.28	10.25	10.24
	10^3	1.57	8.54	11.09	11.60	12.34	13.41	**13.73**	11.29	11.02	10.97	10.96
	10^4	11.45	11.44	10.94	11.40	11.91	11.66	12.38	11.41	10.59	10.41	10.38



(a) $std = 5 \times 10^{-3}$ (b) $std = 2.5 \times 10^{-3}$

FIG. 3.7. *The SNRs with respect to the values of* α_2.

TABLE 3.3
The average restoration results for the proposed model with different initial guesses.

Models	Initial guesses	ReErr	SNR	PSNR	Time
ROF-noisy	Observed image	0.1735	8.80	21.07	0.92
TV-RSTLS	Observed image	0.1087	12.63	24.90	7.08
	Black image	0.1189	11.91	24.18	15.51
	White image	0.1191	11.89	24.16	16.56
	Random image	0.1425	10.41	22.68	47.93

as an initial guess. We note that the objective function is convex in the ROF-noisy model. We also remark that the above observations are also valid for the other PSFs.

On the other hand, we study the sensitivity of the number of iterations to be set in Algorithm 2. By using the above "Cameraman" blurred and noisy image setting, we report in Table 3.4 the average relative errors, the SNR values, the PSNR values, and the CPU-times for different numbers of iterations in Algorithm 2 of the TV-RSTLS method over 20 noisy cases. Here the initial guess is set to be the observed image. We see from the table that there are very small differences in the relative errors, the SNR values, the PSNR values, and the CPU-times for different numbers of iterations. It is obvious that when the number of iterations is higher, these measures are slightly better, but the overall computational time increases. According to the results in the previous subsections, it is sufficient to set the number of iterations to be three in Algorithm 2.

TABLE 3.4
The average restoration results for the proposed model with different iteration numbers in Algorithm 2.

Iteration number	ReErr	SNR	PSNR	Time
3	0.1087	12.63	24.90	7.08
6	0.1075	12.73	25.00	10.97
9	0.1072	12.75	25.02	15.87
12	0.1072	12.76	25.03	16.72
15	0.1072	12.76	25.03	16.75
18	0.1072	12.76	25.03	17.11
21	0.1072	12.76	25.03	17.31

4. Concluding remarks. In this paper, we have proposed and developed the TV-RSTLS method for image restoration. In the model, we studied the objective function by minimizing two variables: the restored image and the estimated error of the PSF. The proposed objective function contains the data-fitting term of the two variables, the magnitude of noise in the PSF, and the TV regularization term of the restored image. We studied an efficient alternating minimization scheme for solving the TV-RSTLS model. Extensive numerical experiments are provided to illustrate the effectiveness and robustness of the proposed model.

REFERENCES

[1] H. ANDREWS AND B. HUNT, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
[2] T. BELL, *Electronics and the stars*, IEEE Spectrum, 32 (1995), pp. 16–24.
[3] P. CAMPISI AND K. EGIAZARIAN, *Blind Image Deconvolution: Theory and Applications*, CRC Press, Boca Raton, FL, 2007.
[4] T. CHAN AND C. WONG, *Total variation blind deconvolution*, IEEE Trans. Image Process., 7 (1998), pp. 370–375.
[5] T. CHAN, S. ESEDOGLU, F. PARK, AND A. YIP, *Total variation image restoration: Overview and recent developments*, in Handbook of Mathematical Models in Computer Vision, Springer, New York, 2006, pp. 17–31.
[6] H. FU AND J. BARLOW, *A regularized structured total least squares algorithm for high-resolution image reconstruction*, Linear Algebra Appl., 391 (2004), pp. 75–98.
[7] H. FU, M. K. NG, M. NIKOLOVA, AND J. L. BARLOW, *Efficient minimization methods of mixed L2-L1 and L1-L1 norms for image restoration*, SIAM J. Sci. Comput., 27 (2006), pp. 1881–1902
[8] R. GLOWINSKI, *Numerical Methods for Nonlinear Variational Problems*, Springer Verlag, Berlin, 2008.
[9] L. GUAN, AND R. WARD, *Restoration of randomly blurred images via the maximum a posteriori criterion*, IEEE Trans. Image Process., 1 (1992), pp. 256–262.
[10] V. MESAROVIC, N. GALATSANOS, AND A. KATSAGGELOS, *Regularized constrained total least squares image restoration*, IEEE Trans. Image Process., 4 (1995), pp. 1096–108.
[11] M. K. NG, R. PLEMMONS, AND F. PIMENTEL, *A new approach to constrained total least squares image restoration*, Linear Algebra Appl., 316 (2000), pp. 237–258.
[12] M. K. NG, F. WANG, AND X. YUAN, *Inexact alternating direction methods for image recovery*, SIAM J. Sci. Comput., 33 (2011), pp. 1643–1668.
[13] M. NG, P. WEISS, AND X. YUAN, *Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods*, SIAM J. Sci. Comput., 32 (2010), pp. 2710–2736.
[14] A. PRUESSNER AND D. P. O'LEARY, *Blind deconvolution using a regularized structured total least squares norm algorithm*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 1018–1037.
[15] J. B. ROSEN, H. PARK, AND J. GLICK, *Total least norm formulation and solution for structured problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 110–126.
[16] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[17] S. VAN HUFFEL AND J. VANDEWALLE, *The Total Least Squares Problem: Computational Aspects and Analysis*, Frontiers in Applied Mathematics, SIAM, Philadelphia, 1991.