

Alternating Direction Method of Multipliers for Nonlinear Image Restoration Problems

Chuan Chen, Michael K. Ng, and Xi-Le Zhao

Abstract—In this paper, we address the total variation (TV)-based nonlinear image restoration problems. In nonlinear image restoration problems, an original image is corrupted by a spatially-invariant blur, the build-in nonlinearity in imaging system, and the additive Gaussian white noise. We study the objective function consisting of the nonlinear least squares data-fitting term and the TV regularization term of the restored image. By making use of the structure of the objective function, an efficient alternating direction method of multipliers can be developed for solving the proposed model. The convergence of the numerical scheme is also studied. Numerical examples, including nonlinear image restoration and high-dynamic range imaging are reported to demonstrate the effectiveness of the proposed model and the efficiency of the proposed numerical scheme.

Index Terms—Nonlinearity, image restoration, total variation, alternating direction method of multipliers, high-dynamic range imaging.

I. INTRODUCTION

IMAGE restoration plays an important role in many applications of sciences and engineering such as medical and astronomical imaging, film restoration, image and video coding. The linear image formation model has been studied extensively for its simplicity. Indeed, the nonlinear image formation model provides an accurate description for a number of imaging systems [1], [2]. For example, the light intensity is nonlinearly translated into the pixel intensity in the photochemical and photoelectric systems [1], [15], [16]. The other examples contain camera response functions in

Manuscript received May 22, 2014; revised September 6, 2014; accepted November 3, 2014. Date of publication November 12, 2014; date of current version December 8, 2014. The work of M. K. Ng was supported in part by the Research Grants Council, General Research Fund, through the Hong Kong Baptist University (HKBU), Hong Kong, under Grant 202013, and in part by the HKBU Faculty Research Grant. The work of X.-L. Zhao was supported in part by 973 Program under Grant 2013CB329404, in part by the National Natural Science Foundation of China under Grant 61170311, Grant 61370147, Grant 61402082, in part by the Sichuan Province Science and Technology Research Project under Grant 2012GZX0080, and in part by the Fundamental Research Funds for the Central Universities under Grant ZYGX2013J106. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Javier Mateos.

C. Chen is with the Department of Mathematics, Hong Kong Baptist University, Hong Kong (e-mail: 12466220@life.hkbu.edu.hk).

M. K. Ng is with the Centre for Mathematical Imaging and Vision, Department of Mathematics, Hong Kong Baptist University, Hong Kong (e-mail: mng@math.hkbu.edu.hk).

X.-L. Zhao is with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 610051, China, and also with the Centre for Mathematical Imaging and Vision, Hong Kong Baptist University, Hong Kong (e-mail: xlzhao122003@163.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2014.2369953

image deblurring [12], high-dynamic-range imaging [14], and super-resolution image reconstruction [2], [6].

The nonlinear image degradation model for a spatially-invariant system is generally formulated as [1], [15], [20], and [22]:

$$g = s(Hf_{\text{true}}) + n. \quad (1)$$

Here $g \in \mathbb{R}^{m^2}$ is the observed image, the true image $f_{\text{true}} \in \mathbb{R}^{m^2}$ (corresponding to an m -by- m image) is corrupted by a component-wise nonlinearity $s(\cdot)$, a spatially-invariant blurring matrix $H \in \mathbb{R}^{m^2 \times m^2}$ and a noise vector $n \in \mathbb{R}^{m^2}$. The blurring matrix is a matrix of block circulant with circulant blocks (BCCB) when periodic boundary conditions are used. The blurring matrix is a matrix of block Toeplitz-plus-Hankel with Toeplitz-plus-Hankel blocks (BTHTHB) when Neumann boundary conditions are imposed [10]. In this work, we consider periodic boundary conditions.

The nonlinear image restoration model is more accurate but is often more computational intractable. Recently, great attentions have been paid to nonlinear image restoration problems. Tekalp and Pavlovic [15] proposed to transform the noisy and blurred image into the exposure domain using the inverse of the sensor nonlinearity, and derived a linear minimum mean square error deconvolution filter in the presence of multiplicative noise. In order to address the nonlinear least squares problem

$$\arg \min_f \frac{1}{2} \|s(Hf) - g\|_2^2, \quad (2)$$

Zervakis and Venetsanopoulos [22] used the steepest descent method. The convergence analysis of the nonlinear iterative algorithm were given in [22]. Zervakis and Venetsanopoulos [20] further considered the Gauss-Netwon (GN) algorithm for the nonlinear least squares problem (2). The GN method is based on the approximation of the nonlinearity involved in the objective function (2) with its first-order Taylor series expansion at f^k :

$$\arg \min_f \frac{1}{2} \|s(Hf^k) + D_s H(f - f^k) - g\|_2^2,$$

where f^k is the iterate at the k th iteration and D_s is a diagonal matrix with the diagonal vector being equal to the derivative of s at Hf^k . Thus the linear least squares problem can be solved and the iterate f^{k+1} at the $(k+1)$ th iteration is obtained by

$$f^{k+1} = f^k + (H^T D_s^2 H)^{-1} H^T D_s (g - s(Hf^k)),$$

where $H^T D_s^2 H$ is assumed to be invertible. The Levenberg-Marquardt modification is introduced to overcome the singularity of $H^T D_s^2 H$. The corresponding solution f^{k+1} is given by

$$f^{k+1} = f^k + (H^T D_s^2 H + \mu I)^{-1} H^T D_s(g - s(Hf^k)),$$

where μ is a positive constant. The main difficulty of the computational procedure is that the inverse of $H^T D_s^2 H + \mu I$ is required or the solution of the corresponding linear system is required. However, since the structure of the blurring matrix is destroyed by the diagonal matrix, the matrix $H^T D_s^2 H$ cannot be diagonalized by the fast Fourier transform. It is computational intractable for large-scale problems to find the inverse or solve the corresponding linear system. Zervakis and Venetsanopoulos [20] considered the pseudo-inverse approximation of H in the computational procedure:

$$\begin{aligned} f^k + H^T(HH^T + \mu I)^{-1}D_s^{-2}(HH^T + \mu I)^{-1}HH^TD_s \\ \times (g - s(Hf^k)), \end{aligned}$$

where the pseudo-inverse of H is approximated by $H^T(HH^T + \mu I)^{-1}$ as μ approaches zero. The benefit from this approximation is that the inverse of $(HH^T + \mu I)^{-1}$ can be efficiently computed via the fast Fourier transform. The resulting computational cost is of $O(m^2 \log m)$ at each iteration. The convergence properties of the GN method were also studied in [20].

It is interesting to note that there is a nice connection between the Levenberg-Marquardt modified solution of the nonlinear least squares problem (2) and the solution of the following Tikhonov regularized nonlinear least squares problem obtained by the Gauss-Netwon method

$$\arg \min_f \frac{1}{2} \|s(Hf) - g\|_2^2 + \mu \|f\|_2^2, \quad (3)$$

where μ is the regularization parameter. We remark there is no regularization term in the original nonlinear least squares problem (2). Recently, Lu et al. [9] considered the generalized Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm for solving the Tikhonov regularized nonlinear least squares problem (3).

In this work, we study the TV based nonlinear least squares problem as follows:

$$\arg \min_f E(f) := \frac{1}{2} \|s(Hf) - g\|_2^2 + \mu \sum_{i=1}^{m^2} \|D_i f\|_2, \quad (4)$$

subject to $a_1 \leq f \leq a_2$, where $D_i f \in \mathbb{R}^2$ is the discrete gradient of f at the i th pixel and the sum $\sum_{i=1}^{m^2} \|D_i f\|_2$ is the discrete total variation of f . The TV regularization [13] has proven to be successful for image restoration problem because of its ability to preserve sharp discontinuities (edges) in the restored images. Beside the TV regularizer, we consider the value range of f is between a_1 and a_2 . We remark that the existence of the minimizers of the proposed model (4) is guaranteed by Weierstrass extreme value theorem [25].

The main contribution of this work is to design an efficient algorithm for solving the TV based nonlinear image restoration model (4) to obtain the restored image. The main idea is to

split the original difficult optimization problem (4) into some easier subproblems under the ADM framework such that the nonlinearity, the blurring operator, and the TV regularization term can be tackled separately, and the resulting convergence is still guaranteed. The rest of this paper is structured as follows. In Section II, we develop the ADM method for solving the proposed model (4). Also the convergence of the proposed numerical scheme will be studied. In Section III, we provide some promising numerical results to demonstrate the effectiveness of the proposed model and the efficiency of the proposed numerical scheme. The numerical examples include the nonlinear image restoration and high dynamic range imaging. Finally, some concluding remarks are given in Section IV.

II. ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Because the TV regularization term in (4) is non-smooth, the existing methods for nonlinear image restoration problems cannot be extended to tackle the proposed model (4). In this section, we develop an efficient alternating direction method of multipliers to solve the TV based nonlinear image restoration model (4). The ADM method allows us to convert the original problem involving a nonlinear least squares term and a TV regularizer into a sequence of simpler problems: a decoupled minimization problem, a proximal mapping, and a structured linear system. The classical ADM method, which was proposed by Gabay and Mercier [7], is a practical version of the classical augmented Lagrangian method (ALM) for solving linearly constrained convex problems with a separate structure whose objective function is in the form of the sum of two individual functions without coupled variables. The philosophy behind the ADM method is to decompose the joint minimization into two easier subproblems such that the involved variables can be minimized separately in an alternating order; see [3] for a comprehensive review of the recent development of the ADM method.

We reformulate (4) as the following constrained optimization problem:

$$\begin{aligned} \arg \min_z \quad & \frac{1}{2} \|s(z) - g\|_2^2 + \mu \sum_{i=1}^{m^2} \|p_i\|_2 + \chi_{K_1}(u) + \chi_{K_2}(v), \\ \text{subject to} \quad & Hf = z, \quad f = u, \quad f = v, \quad D_i f = p_i, \end{aligned} \quad (5)$$

where $i = 1, 2, \dots, m^2$, and $\chi_{K_1}(u)$ and $\chi_{K_2}(v)$ are the indicator functions given by

$$\chi_{K_1}(u) = \begin{cases} 0, & \text{if } u - a_1 \geq 0, \\ \infty, & \text{otherwise,} \end{cases}$$

and

$$\chi_{K_2}(v) = \begin{cases} 0, & \text{if } v - a_2 \leq 0, \\ \infty, & \text{otherwise.} \end{cases}$$

Let $J_1(f) = 0$ and $J_2(y) = \frac{1}{2} \|s(z) - g\|_2^2 + \mu \sum_{i=1}^{m^2} \|p_i\|_2 + \chi_{K_1}(u) + \chi_{K_2}(v)$.

The constraints can be expressed as:

$$Bf + Cy := \begin{bmatrix} D \\ H \\ I \\ I \end{bmatrix} f - \begin{bmatrix} I & I & I \\ & I & I \end{bmatrix} \begin{bmatrix} p \\ z \\ u \\ v \end{bmatrix},$$

Algorithm 1 The Alternating Direction Method of Multipliers for Solving (5)**Input:** $g, f^0, \lambda^0, \mu, \beta, a_1$, and a_2 .**Output:** f^k .**while** not converged **do****Step 1.** Given f^{k-1} and λ_1^{k-1} , update p^k by using the shrinkage formula:

$$p_i^k = \max \left\{ \left\| D_i f^{k-1} + \frac{[\lambda_1^{k-1}]_i}{\beta} \right\|_2 - \frac{\mu}{\beta}, 0 \right\} \frac{D_i f^{k-1} + \frac{[\lambda_1^{k-1}]_i}{\beta}}{\left\| D_i f^{k-1} + \frac{[\lambda_1^{k-1}]_i}{\beta} \right\|_2}. \quad (7)$$

Step 2. Given f^{k-1} and λ_2^{k-1} , component-wisely update z^k by solving m^2 decoupled one-dimensional nonlinear equations using the Newton method:

$$\arg \min_z \frac{1}{2} \|s(z) - g\|_2^2 + \frac{\beta}{2} \left\| z - H f^{k-1} - \frac{\lambda_2^{k-1}}{\beta} \right\|_2^2. \quad (8)$$

Step 3. Given f^{k-1} , λ_3^{k-1} , and λ_4^{k-1} , component-wisely update u^k and v^k by projections

$$u^k = P_+ \left(f^{k-1} - a_1 + \frac{\lambda_3^{k-1}}{\beta} \right) + a_1, \quad (9)$$

$$v^k = P_- \left(f^{k-1} - a_2 + \frac{\lambda_4^{k-1}}{\beta} \right) + a_2, \quad (10)$$

where $P_+(x) = \max\{x, 0\}$ and $P_-(x) = \min\{x, 0\}$.**Step 4.** Given p^k, z^k, u^k, v^k , and λ^{k-1} , update f^k by solving the following linear system

$$(D^T D + H^T H + 2I)f = D^T(p^k - \frac{\lambda_1^{k-1}}{\beta}) + H^T(z^k - \frac{\lambda_2^{k-1}}{\beta}) + (u^k - \frac{\lambda_3^{k-1}}{\beta}) + (v^k - \frac{\lambda_4^{k-1}}{\beta}). \quad (11)$$

Step 5. Update λ^k by

$$\lambda^k = \lambda^{k-1} + \beta(Bf^k + Cy^k). \quad (12)$$

end while

where $D := [D_x; D_y]$ is a combination of the first-order finite difference operators D_x and D_y with respect to the horizontal direction and the vertical direction. D_x and D_y are m^2 -by- m^2 matrices constructed by stacking the first row and the second row of D_i , respectively. The optimization problem is in the form of the sum of two individual functions without crossed variables. The optimization problem fits the framework of the ADM method. Attaching the Lagrangian multiplier $\lambda = [\lambda_1^T, \lambda_2^T, \lambda_3^T, \lambda_4^T]^T$ to the linear constraints, the augmented Lagrangian function of (5) is given by

$$\begin{aligned} \mathcal{L}(f, y, \lambda) = & \frac{1}{2} \|s(z) - g\|_2^2 + \mu \sum_{i=1}^{m^2} \|p_i\|_2 + \chi_{K_1}(u) + \chi_{K_2}(v) \\ & + (\lambda, Bf + Cy) + \frac{\beta}{2} \|Bf + Cy\|_2^2, \end{aligned} \quad (6)$$

where $\beta > 0$ is the penalty parameter for the linear constraints to be satisfied.The ADM method performs minimization with respect to f and y alternatively, and then the update of λ follows. More precisely, the numerical scheme for solving (5) is described in Algorithm 1.In Step 1, we solve the p -subproblem by using the 2D shrinkage formula. In Step 3, we solve the u -subproblem and the v -subproblem by projecting onto

the constrained sets. In Step 4, we solve the f -subproblem by solving the structured linear system via the fast Fourier transform. It remains to solve the z -subproblem in Step 2. Before tackling the z -subproblem, we discuss the convexity of the z -subproblem (8). The following theorem states that the second quadratic penalty term in (8) contributes to convexify the z -subproblem. The detailed proof is provided in Appendix A.

Theorem 1: Assume the value range of g is between $[b_1, b_2]$, the nonlinear function s is convex or concave, s and s'' are monotone on $(0, \infty)$. Then there exists a scalar M such that the z -subproblem is strictly convex on the interval $[a_1, a_2] \subset (0, \infty)$ when β is greater than M .

For example, when s is the natural logarithm function $\ln(x)$ (see for instance [14], [22]) or s is the power function x^γ for some γ (see for instance [5], [8]), it satisfies the assumptions in Theorem 1. If the z -subproblem is strictly convex, the corresponding optimization problem (8) has an unique solution. In some cases, the closed-form solution of the z -subproblem is available. For example, the global minimizer can be determined by finding the roots of a cubic polynomial analytically when the nonlinear function is x^2 . However, in most cases, there is no closed-form solution of the z -subproblem. Note the optimization problem (8) can be decoupled into m^2 1D nonlinear equations, therefore the

solution can be calculated very efficiently by using the 1D Newton method. Since the estimated f^{k-1} might not be accurate enough, it is not necessary to calculate an exact solution of the z sub-problem at the k th iteration in Algorithm 1. For computationally efficiency, we implement an inexact version of Algorithm 1 in practice, which only runs several inner Newton iterations.

Thus the computational complexities of updating the variables p , z , u , and v in (7), (8), (9), and (10) are $O(m^2)$, $O(m^2 \log m)$, $O(m^2)$, and $O(m^2)$ operations, respectively. The structured linear system in (11) can be exactly and efficiently solved by using the fast Fourier transform in $O(m^2 \log m)$ operations. In summary, the total cost is of $O(m^2 \log m)$ at each iteration.

For the nonlinear least squares problem (2), the convergence properties of the steepest descent algorithm and Gauss-Newton algorithm were studied [20], [22]. Based on the global convergence theorem, the residual norm was proved to be monotone decreasing. For the Tikhonov regularized nonlinear least squares problem (3), the sequence generated by the BFGS method converges to a stationary point under some assumptions [9]. Since the TV regularization term is non-smooth, the previous convergence results might be difficult to be extended to the TV based nonlinear image restoration model (4). The global convergence can be guaranteed when the classical ADM method is applied to two block convex programming problems [3]. However, to best of our knowledge, there is no global convergence results in general for non-convex programming. The empirical evidences strongly suggest that Algorithm 1 has a very strong convergence behavior. Here we show that the sequence generated by Algorithm 1 converges to the following Karush-Kuhn-Tucker (KKT) point whenever it converges. A point $w := (f, p, z, u, v, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is a KKT point of problem (5) if it satisfies the following KKT conditions:

$$\begin{aligned} & D^T(p - \frac{\lambda_1}{\beta}) + H^T(z - \frac{\lambda_2}{\beta}) + (u - \frac{\lambda_3}{\beta}) + (v - \frac{\lambda_4}{\beta}) \\ & -(D^T D + H^T H + 2I)f = 0 \\ & s'(z)(s(z) - g) - \lambda_2 = 0 \\ & Df - p = 0 \\ & Hf - z = 0 \\ & f - u = 0 \\ & f - v = 0 \\ & \lambda_3 \leq 0 \leq (u - a_1), \quad \lambda_3 \circ (u - a_1) = 0 \\ & (v - a_2) \leq 0 \leq \lambda_4, \quad \lambda_4 \circ (v - a_2) = 0 \\ & 0 \in \mu \partial \|p_i\|_2 - [\lambda_1]_i \end{aligned} \quad (13)$$

where $i = 1, \dots, m^2$, and \circ refers to the Hadamard product.

Here we just state the main theorem. The proof is provided in Appendix B.

Theorem 2: Let $\{w^k\}_{k=1}^\infty$ be the sequence generated by Algorithm 1 that satisfies the condition

$$\lim_{k \rightarrow \infty} (w^{k+1} - w^k) = 0. \quad (14)$$



Fig. 1. The original (a) Cameraman, (b) Lena, (c) Parrot, and (d) Panda images.

Then any accumulation point of $\{w^k\}_{k=1}^\infty$ is a KKT point of problem (5). Consequently, any accumulation point of $\{f^k\}_{k=1}^\infty$ is a KKT point of problem (4).

Far from being satisfactory, the convergence result nevertheless provides some assurance on the behavior of the ADM method for solving nonlinear image restoration problems. The further theoretical refinements are certainly desirable.

III. EXPERIMENTAL RESULTS

In this section, we report numerical results for nonlinear image restoration and high-dynamic range imaging. The quality of the restored images is measured by the highest peak signal-to-noise ratio (PSNR)

$$\text{PSNR} = 10 \log_{10} \frac{\max_i \{[f_{\text{true}}]_i^2\}}{\frac{1}{m^2} \sum_{i=1}^{m^2} ([f]_i - [f_{\text{true}}]_i)^2},$$

where f_{true} and f are the true image and the restored image, respectively. Also, we evaluate the quality of the restored image by the structural similarity (SSIM) index, which has proven to be consistent with human eye perception; see [17] for more details. In all the methods, their corresponding regularization parameters are chosen by providing the highest PSNR value of the restored image. The stopping criterion of the numerical method is measured on the relative change of the successive solutions being less than a specified tolerance. All the tests are performed under Windows 7 and MATLAB Version 7.10 (R2010a) running on a desktop with an Intel Core2 Quad CPU at 2.66 GHz with 4 GB of memory.

A. Nonlinear Image Restoration

The performance of the nonlinear least squares (NLS) model and the TV based nonlinear least squares (TVNLS) model are compared for nonlinear image restoration. The NLS model is solved by the GN method, and the TVNLS model is solved by the ADM method. The testing image is the “Cameraman” image given in Fig. 1(a). The stopping tolerance

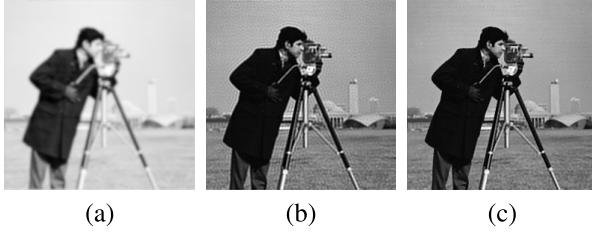


Fig. 2. (a) The observed image degraded by the natural logarithm nonlinearity, the restored images by using (b) the NLS model (PSNR: 26.95dB, SSIM: 0.8036, Iter: 1526, Time: 42.0 seconds, and $\mu = 5 \times 10^{-2}$), and (c) the TVNLS model (PSNR: 28.68dB, SSIM: 0.8848, Iter: 998, Time: 68.0 seconds, $\mu = 5 \times 10^{-7}$, and $\beta = 10^{-6}$).

is set to be 10^{-5} and the initial guess is set to be the inverse image $s^{-1}(g)$ unless otherwise specified, where s^{-1} is the component-wise inverse nonlinearity. The number of inner Newton iterations in Algorithm 1 is set to be 5.

In the first experiment, we evaluate the performance of the TVNLS model for recovering the nonlinear degraded images. In photochemical systems, the relation between optical density and the logarithm of incident intensity is widely assumed [15]. Here the logarithm function is used in regions where the response is linear with the logarithm of incident intensity. In photoelectronic systems, a power-law is employed in the relationship between incident light intensity and the output of the photoelectronic sensor [8]. The photoelectronic systems the function is raising to the power to the variable. We test two types of nonlinear functions including the natural logarithm function $\ln(x)$ and the power function x^2 . The “Cameraman” image is blurred by a Gaussian point spread function (MATLAB build-in function “fspecial” with hsize = 21 and sigma = 1.6). After the corresponding nonlinear response, 0.1% Gaussian white noise was further added to the degraded images.

In Fig. 2, we display the restoration results by using the NLS model and the TVNLS model, when the natural logarithm nonlinearity is considered. The corresponding PSNR values, SSIM values, iterations, CPU time, and parameters are listed in the caption of Fig. 2. We observe that the PSNR value and the SSIM value by using the TVNLS model are better than those by using the NLS model. Their corresponding percentage improvements are about 6.0% in PSNR and 9.2% in SSIM. Furthermore, we display the zoom-in “pillar” parts in Fig. 3 to visually illustrate the quality of the restored image by the NLS model and the TVNLS model. We see that the restored image obtained by the TVNLS model is better than that by the NLS model. On the other hand, we show the PSNR behavior of the NLS model and the TVNLS model with respect to the iteration number in Fig. 4. We observe in the earlier iterations that the increase of the PSNR value in the TVNLS model is slower than that in the NLS model. However, in the latter iterations of the NLS model, the PSNR value does not further increase. In contrast, the PSNR value is still increasing in the latter iterations of the TVNLS model. We note that the resulting PSNR value of the TVNLS model is higher than that of the NLS model.

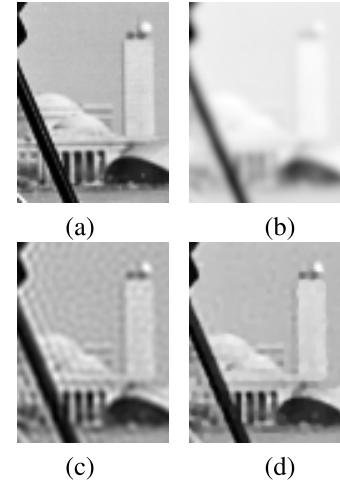


Fig. 3. The zoom-in “pillar” part of (a) the true image, (b) the observed image degraded by the natural logarithm nonlinearity, the restored images by using (c) the NLS model, and (d) the TVNLS model.

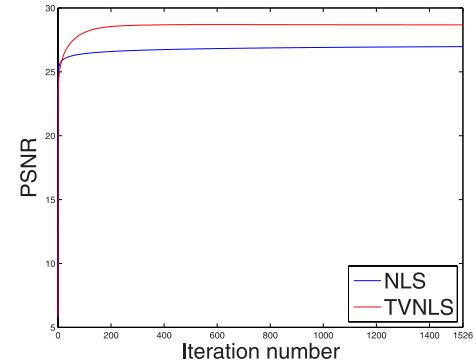


Fig. 4. The PSNRs of the NLS model and the TVNLS model with respect to the iteration number for the natural logarithm nonlinearity.

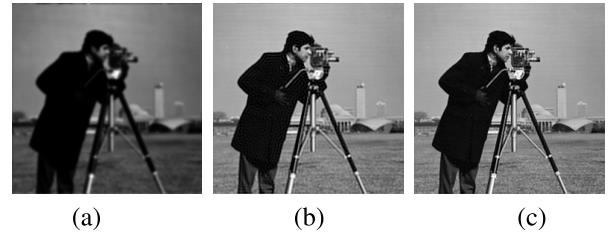


Fig. 5. (a) The observed image degraded by the power nonlinearity, the restored images by using (b) the NLS model (PSNR: 27.80 dB, SSIM: 0.8401, Iter: 1430, Time: 35.0 seconds, and $\mu = 2 \times 10^{-2}$), and (c) the TVNLS model (PSNR: 30.27 dB, SSIM: 0.9124, Iter: 1609, Time: 88.0 seconds, $\mu = 10$, and $\beta = 100$).

Next we present the restoration results in Figs. 5 and 6, when the power nonlinearity is used. The corresponding PSNR values, SSIM values, iterations, CPU time, and parameters are listed in the caption of Fig. 5. We note that the percentage improvements of the TVNLS model over the NLS model are about 8.2% in PSNR and 7.9% in SSIM. We see from the figures that the visual quality of the restored image by using the TVNLS model is better than that by using the NLS model. In Fig. 7, we display the PSNR behavior of the NLS model

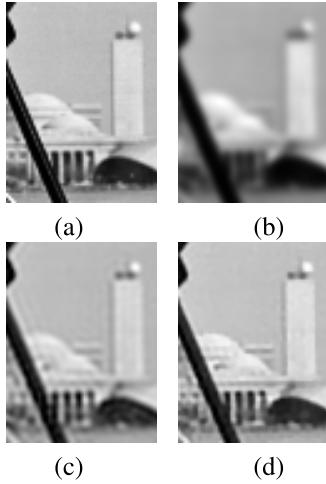


Fig. 6. The zoom-in “pillar” part of (a) the true image, (b) the observed image degraded by the power nonlinearity, the restored images by using (c) the NLS model, and (d) the TVNLS model.

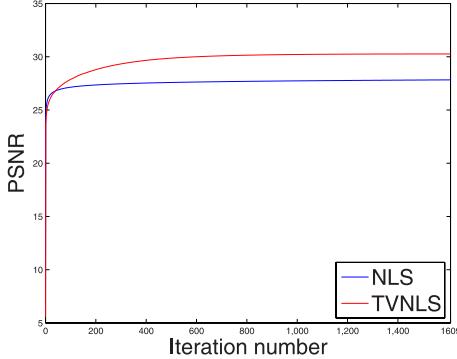


Fig. 7. The PSNRs of the NLS model and the TVNLS model with respect to the iteration number for the power nonlinearity.

and the TVNLS model with respect to the iteration number. The curve for the TVNLS model in Fig. 7 is similar to that in Fig. 4. The resulting PSNR value of the restored image by the TVNLS is higher than that by the NLS model.

Moreover, we make use of these two examples to numerically demonstrate the convergence of the ADM method for solving the TVNLS model. The convergence behaviors of the objective function value in log scale are displayed in Fig. 8 for different settings of regularization parameters μ and penalty parameters β . We observe from Fig. 8 that the objective function value is monotonically decreasing with respect to the iteration number for different parameter settings.

Finally, the performances of the NLS model and TVNLS model are evaluated for different blurs and noise levels. Here we consider Gaussian blur (MATLAB build-in function “fspecial” with `hsiz` = 21 and `sigma` = 1.6) and Moffat blur (MATLAB command: `psfMoffat([23 23],4,5)`, which is a function in the HNO package¹). And zero mean Gaussian white noise with standard deviation 10^{-3} and 10^{-2} is considered. The testing image is the Cameraman image given in Fig. 1(a). The stopping tolerance is set to be 10^{-5}

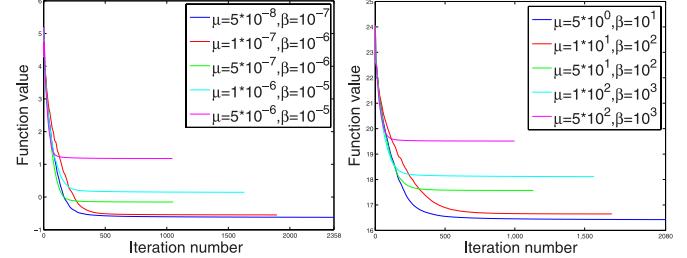


Fig. 8. The objective function value with respect to the iteration number (left: the natural logarithm nonlinearity, and right: the power nonlinearity).

and the initial guess is set to be the inverse image $s^{-1}(g)$. The numbers of inner Newton iterations in Algorithm 1 are set to be 5 and 20 for Gaussian blur and Moffat blur respectively. The restoration results are summarized in Table I for different blurs and noise levels. Similarly, we observe from Table I that the performance of the TVNLS model is superior to that of the NLS model in terms of PSNR values and SSIM values.

B. The Influence of Parameter Settings

In this subsection, we studied several parameter settings in the TVNLS model, namely, the number of inner Newton iterations, the penalty parameter, and the initial guess of the ADM method.

1) *The Number of Inner Newton Iterations:* We first investigate the performance of the TVNLS model with respect to the number of inner Newton iterations used in the ADM method. In Fig. 9, we use the examples in Fig. 2 and Fig. 5 to show the PSNRs of the restored “Cameraman” images and their required CPU time by using different numbers of inner Newton iterations. In these examples, we see that it is not necessary to run many inner Newton iterations in order to obtain high PSNR restored images. Indeed, the results suggest there is no improvement in PSNR value even more inner Newton iterations are employed, and five inner Newton iterations are sufficient for nonlinear image restoration with low computational time.

2) *The Penalty Parameter:* Next we study the influence of the penalty parameter β used in ADM method for solving the TVNLS model. In Fig. 10, we use the examples in Fig. 2 and Fig. 5 to show the PSNRs of the restored “Cameraman” images and their required CPU time by using different values of the penalty parameter. We see that the value of the penalty parameter indeed affects the performance of the TVNLS model. Also the best penalty parameters (among all the testing values) are different for the natural logarithm nonlinearity ($\beta = 10^{-6}$) and the power nonlinearity ($\beta = 10^2$). We note that a small penalty parameter value is used for the natural logarithm nonlinearity while a large penalty parameter value is used for the power nonlinearity.

3) *The Initial Guess:* Finally we test the sensitivity to the initial guess used in the ADM method for solving the TVNLS model. We consider the four testing images in Fig. 1. All testing images are blurred by a Gaussian point spread function (Matlab build-in function “fspecial” with `hsiz` = 21

¹<http://www2.imm.dtu.dk/~pch/HNO/>

TABLE I
THE RESTORATION RESULTS BY USING THE NLS MODEL AND THE TVNLS MODEL WITH DIFFERENT BLURS AND NOISE LEVELS

Nonlinearity	Blur	Standard Deviation	NLS				TVNLS			
			PSNR	SSIM	Iter	Time	PSNR	SSIM	Iter	Time
Logarithm	Gaussian	0.001	26.95	0.8036	1526	42.0	28.68	0.8848	998	43.0
		0.01	23.75	0.6012	1887	35.0	25.76	0.8120	1247	62.0
	Moffat	0.001	28.52	0.7987	1782	36.0	30.81	0.9039	903	48.0
		0.01	24.26	0.6201	1996	39.0	26.44	0.8225	1149	60.0
Power	Gaussian	0.001	27.80	0.8401	1430	35.0	30.27	0.9124	1609	80.0
		0.01	25.07	0.7451	1776	44.0	27.30	0.8619	1730	73.0
	Moffat	0.001	29.82	0.8599	1627	27.0	33.49	0.9304	1721	78.0
		0.01	26.12	0.7609	2677	45.0	28.88	0.8793	1541	71.0

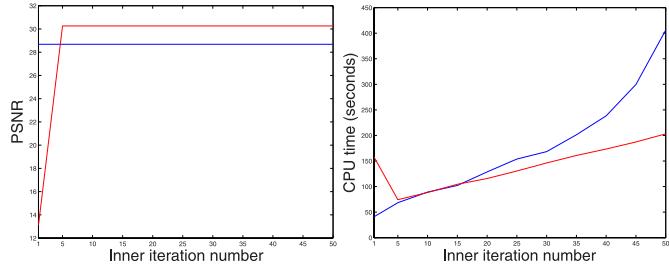


Fig. 9. The PSNR values and the CPU time with respect to the number of inner Newton iterations (blue: the natural logarithm nonlinearity, and red: the power nonlinearity).

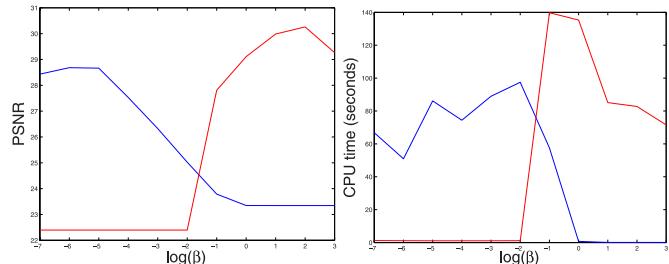


Fig. 10. The PSNR values and the CPU time with respect to the penalty parameter (blue: the natural logarithm nonlinearity, and red: the power nonlinearity).

and $\sigma = 1.6$). The blurred image is then degraded by the corresponding nonlinear response and 0.1% Gaussian white noise. We solve the TVNLS model by the ADM method with different initial guesses including the white image, the random image, the observed image g , and the inverse image $s^{-1}(g)$. In the previous experiment in Section III(A), we employ the inverse image as the initial guess and find that five inner Newton iterations are sufficient to obtain good restoration results. However, when we use the other initial guesses (the white image, the random image, and the observed image), more inner Newton iterations are required in order to obtain comparable restoration results. Here we set the number of inner Newton iterations to be 20 for all initial guesses for fair comparisons. The penalty parameters are set to be the same as those in the examples in Fig. 2 and Fig. 5. The restoration results for the four testing images are summarized in Table II. We see from the table that the results obtained by the ADM method in the TVNLS model are about the same for different initial guesses. However, the results obtained by the GN method in the NLS model are different for different

initial guesses. We note that the initial guesses of the white image, the observed image and the inverse image give more or less same restoration results. The initial guess of the random image produces poor restoration results. The GN method is very sensitive to the initial guess. It is clear from the table that the performance of the TVNLS model is significantly better than that of the NLS model.

C. High Dynamic Range Imaging

The low dynamic range (LDR) image responses in a nonlinear fashion with respect to the high dynamic range (HDR) radiance. The nonlinear response is formulated as

$$g = s(r), \quad (15)$$

where r is the true HDR radiance, g is the observed LDR image, and s is the camera response function which maps the HDR radiance r to the observed LDR image g . The closed-form nonlinear function is given by [14]

$$s(r) = c_1 \ln \left(1 + \frac{r}{c_2} \right), \quad (16)$$

where c_1 and c_2 are two positive constants in the camera setting. Here we apply the TVNLS model (4) to recover the true HDR radiance r from the noisy observed LDR image g . The TVNLS model is solved by the ADM method. Here several values of regularization parameters $\mu = i \times 10^j$ ($i = 1$ and 5 , and $j = 3, 4, 5$, and 6) are tested, and the recovered results with the best PSNR values are reported.

In this experiment, we test two 512×640 HDR images from the Stanford HDR dataset.² The value range of the first “office” HDR image is between 0.9929 and 5.6823×10^3 . The value range of the second “building” HDR image is between 13.4534 and 3.5893×10^4 . Here the constant c_2 is set to be 1 and the constant c_1 is set such that the value range of $s(r)$ is between 0 and 1. The observed LDR image is further corrupted by 1% Gaussian white noise. In the ADM method, we set the number of inner Newton iterations to be 5 and the initial guess to be the inverse image. For reference, the linearized majorize-minimize (MM) method [14] is developed to solve the above HDR radiance recover problem. The idea is to use the re-weighted least squares technique to tackle the non-smooth TV term and the linearized technique to tackle the nonlinear least squares data-fitting term. The main

²<http://white.stanford.edu/~brian/hdr/hdr.html>

TABLE II
THE RESTORATION RESULTS BY USING THE NLS MODEL AND THE TVNLS MODEL WITH DIFFERENT INITIAL GUESSES

Image	Nonlinearity	Model	White image		Random image		Observed image		Inverse image	
			PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Cameraman	Logarithm	NLS	26.94	0.8047	13.63	0.1811	26.95	0.8036	26.95	0.8036
		TVNLS	28.69	0.8848	28.70	0.8848	28.68	0.8848	28.68	0.8848
	Power	NLS	27.80	0.8401	13.82	0.1972	27.80	0.8401	27.80	0.8401
		TVNLS	30.30	0.9132	30.24	0.9126	30.41	0.9133	30.26	0.9129
Lena	Logarithm	NLS	28.91	0.8318	13.10	0.1355	28.91	0.8319	28.92	0.8309
		TVNLS	29.58	0.8697	29.48	0.8695	29.58	0.8697	29.58	0.8697
	Power	NLS	29.18	0.8849	13.30	0.1416	27.02	0.8516	29.20	0.8855
		TVNLS	30.77	0.9014	30.66	0.9001	30.75	0.9009	30.77	0.9014
Parrot	Logarithm	NLS	27.05	0.8492	14.00	0.2289	27.97	0.8707	27.99	0.8697
		TVNLS	29.43	0.9037	29.38	0.9034	29.43	0.9037	29.43	0.9037
	Power	NLS	29.06	0.8889	14.24	0.2322	24.36	0.7778	29.14	0.8894
		TVNLS	30.98	0.9219	30.69	0.9208	30.81	0.9211	30.98	0.9218
Panda	Logarithm	NLS	30.09	0.8176	14.30	0.1884	30.09	0.8177	30.10	0.8174
		TVNLS	31.10	0.8366	31.01	0.8364	31.11	0.8366	31.11	0.8366
	Power	NLS	30.62	0.8303	14.31	0.1552	27.30	0.7548	30.65	0.8308
		TVNLS	32.31	0.8609	32.18	0.8597	32.27	0.8604	32.31	0.8609



(a)

(b)



(c)

(d)

Fig. 11. (a) The tone mapped LDR image from the true HDR image, (b) the noisy observed LDR image, (c) the tone mapped LDR image from the recovered HDR image obtained by the MM method ($\mu = 10^{-4}$), and (d) the tone mapped LDR image from the recovered HDR image obtained by the ADM method ($\mu = 10^{-4}$, and $\beta = 4 \times 10^{-4}$).

computational bottleneck of the MM method is that many ill-conditioned systems are required to be solved at each iteration. In [14], conjugate gradient method is employed to solve the ill-conditioned linear systems. The stopping tolerance of the MM method and the ADM method is set to be 3×10^{-5} .

In Figs. 11-12, we show the recovered results by different methods. The recovered HDR images can be download from the website³ and checked by the HDR image viewer. For better visual effects, we show their tone mapped LDR images by using the fast bilateral filtering technique [24], which is one of the classic methods in the well-known HDR toolbox.⁴ In Fig. 11(a), it is the tone mapped LDR image from the



(a)

(b)



(c)

(d)

Fig. 12. (a) The tone mapped LDR image from the true HDR image, (b) the noisy observed LDR image, (c) the tone mapped LDR image from the recovered HDR image obtained by the MM method ($\mu = 5 \times 10^{-5}$), and (d) the tone mapped LDR image from the recovered HDR image obtained by the ADM method ($\mu = 5 \times 10^{-6}$, and $\beta = 3.5 \times 10^{-5}$).

original “office” HDR image. In Fig. 11(b), it is the noisy observed LDR image. In Fig. 11 (c)-(d), there are the tone mapped LDR images from the recovered HDR images by using the MM method and the ADM method. In the caption of Fig. 11, the corresponding values of the regularization parameter and the penalty parameter are listed. We note that the HDR image recovered by the ADM method is visually better than that by the MM method. In their tone mapped versions, we also see that the ADM method can provide better visual effects than the MM method. For better visual inspection, we also display “board” and “toy” zoom-in parts of the recovered images in Figs. 13-14. Indeed, the PSNR value of the recovered image by the ADM method (43.08 dB) is higher than that by the MM method (42.76 dB). Moreover, the computational time required by the ADM method (270 seconds) is less than that required by the MM method (389 seconds).

³<http://www.math.hkbu.edu.hk/~mng/hdr/hdr.html>

⁴<http://www.mathworks.com/matlabcentral/linkexchange/links/2792-the-hdr-toolbox>

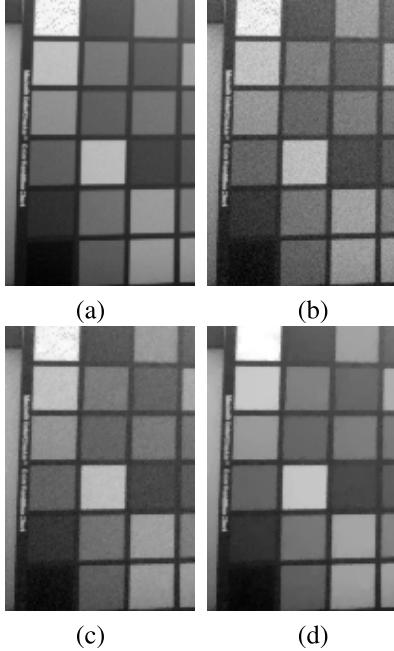


Fig. 13. The zoom-in ‘‘board’’ part of (a) the tone mapped LDR image from the true HDR image, (b) the noisy observed LDR image, (c) the tone mapped LDR image from the recovered HDR image obtained by the MM method, and (d) the tone mapped LDR image from the recovered HDR image obtained by the ADM method.

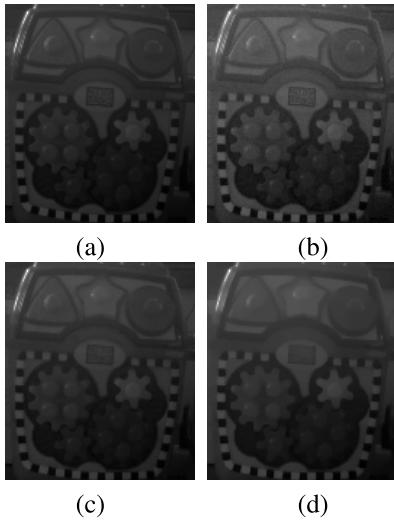


Fig. 14. The zoom-in ‘‘toy’’ part of (a) the tone mapped LDR image from the true HDR image, (b) the noisy observed LDR image, (c) the tone mapped LDR image from the recovered HDR image obtained by the MM method, and (d) the tone mapped LDR image from the recovered HDR image obtained by the ADM method.

In Fig. 12(a), it is the tone mapped LDR image from the original ‘‘building’’ HDR image. In Fig. 12(b), it is the noisy observed LDR image. In Fig. 12(c)–(d), there are the tone mapped LDR images from the recovered HDR images by using the MM method and the ADM method. Similarly, we observe that the PSNR value of the recovered image by the ADM method (60.44 dB) is higher than that by the MM method (60.43 dB). Moreover, the computational

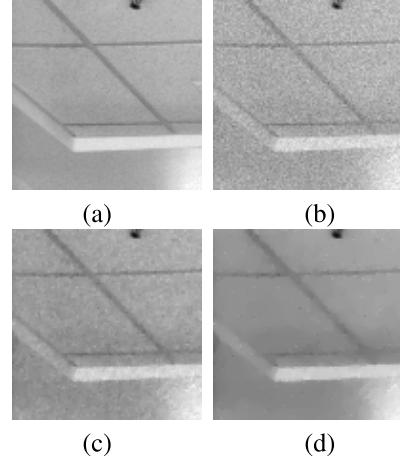


Fig. 15. The zoom-in ‘‘ceiling’’ part of (a) the tone mapped LDR image from the true HDR image, (b) the noisy observed LDR image, (c) the tone mapped LDR image from the recovered HDR image obtained by the MM method, and (d) the tone mapped LDR image from the recovered HDR image obtained by the ADM method.

time required by the ADM method (101 seconds) is less than that required by the MM method (413 seconds). For better visual inspection, we also display ‘‘ceiling’’ zoom-in part of the recovered images in Fig. 15.

IV. CONCLUDING REMARKS

In this work, we have proposed the TV based variational model to tackle nonlinear image restoration problems. In the model, we studied the functional by minimizing the nonlinear least squares data-fitting term and the TV regularization term. We have developed an efficient alternating direction method of multipliers to solve the proposed model. The convergence results show that the ADM method converges to a Karush-Kuhn-Tucker point whenever it converges. Some promising numerical examples including nonlinear image restoration and high dynamic range imaging are provided to illustrated the effectiveness of the proposed model and the efficiency of the proposed numerical scheme.

APPENDIX A CONVEXITY

Proof of Theorem 1

Proof: Let us define

$$h(z) := \frac{1}{2}(s(z) - g)^2 + \frac{\beta}{2}(z - d)^2.$$

We consider the case $s(z)$ and $s''(z)$ are increasing functions, the other three cases can be proved with the same arguments. If $s(z)$ is convex, it follows from the monotonicity of $s(z)$ and $s''(z)$ that

$$\begin{aligned} h''(z) &= s''(z)(s(z) - g) + (s'(z))^2 + \beta \\ &= s''(z) \left(\frac{\beta}{s''(z)} + s(z) - g \right) + (s'(z))^2 > 0, \end{aligned}$$

when $\beta > s''(a_2)(b_2 - s(a_1))$. On the other hand, if $s(z)$ is concave, it follows from the monotonicity of $s(z)$ and $s''(z)$

that

$$h''(z) = s''(z) \left(\frac{\beta}{s''(z)} + s(z) - g \right) + (s'(z))^2 > 0,$$

when $\beta > s''(a_1)(b_1 - s(a_2))$. If $s''(z) = 0$ for some z ,

$$h''(z) = (s'(z))^2 + \beta > 0.$$

Therefore, the convexity statement follows directly. \blacksquare

APPENDIX B CONVERGENCE

Proof of Theorem 2

Proof: We can rearrange the update formula in Algorithm 1 as

$$\begin{aligned} D^T(p^k - \frac{\lambda_1^{k-1}}{\beta}) + H^T(z^k - \frac{\lambda_2^{k-1}}{\beta}) + (u^k - \frac{\lambda_3^{k-1}}{\beta}) \\ + (v^k - \frac{\lambda_4^{k-1}}{\beta}) - (D^T D + H^T H + 2I)f^k = 0, \\ \beta(Hf^k - Hf^{k-1}) + (\lambda_2^k - \lambda_2^{k-1}) \\ - s'(z^k)(s(z^k) - g) - \lambda_2^k = 0, \\ \mathcal{P}_+(f^{k-1} - a_1 + \frac{\lambda_3^{k-1}}{\beta}) - (u^{k-1} - a_1) - (u^k - u^{k-1}) = 0, \\ \mathcal{P}_-(f^{k-1} - a_2 + \frac{\lambda_4^{k-1}}{\beta}) - (v^{k-1} - a_2) - (v^k - v^{k-1}) = 0, \\ \beta(Df^k - p^k) - (\lambda_1^k - \lambda_1^{k-1}) = 0, \\ \beta(Hf^k - z^k) - (\lambda_2^k - \lambda_2^{k-1}) = 0, \\ \beta(f^k - u^k) - (\lambda_3^k - \lambda_3^{k-1}) = 0, \\ \beta(f^k - v^k) - (\lambda_4^k - \lambda_4^{k-1}) = 0. \end{aligned}$$

Note $f^k = f^{k-1} + (f^k - f^{k-1})$, $u^k = u^{k-1} + (u^k - u^{k-1})$, $v^k = v^{k-1} + (v^k - v^{k-1})$, and $\lambda^k = \lambda^{k-1} + (\lambda^k - \lambda^{k-1})$, where the second terms vanish asymptotically. Letting k go to infinity, we obtain the following results

$$\begin{aligned} D^T(p^k - \frac{\lambda_1^k}{\beta}) + H^T(z^k - \frac{\lambda_2^k}{\beta}) + (u^k - \frac{\lambda_3^k}{\beta}) \\ + (v^k - \frac{\lambda_4^k}{\beta}) - (D^T D + H^T H + 2I)f^k \rightarrow 0, \\ \mathcal{P}_+(f^{k-1} - a_1 + \frac{\lambda_3^{k-1}}{\beta}) - (u^{k-1} - a_1) \rightarrow 0, \\ \mathcal{P}_-(f^{k-1} - a_2 + \frac{\lambda_4^{k-1}}{\beta}) - (v^{k-1} - a_2) \rightarrow 0, \\ s'(z^k)(s(z^k) - g) - \lambda_2^k \rightarrow 0, \\ Df^k - p^k \rightarrow 0, \\ Hf^k - z^k \rightarrow 0, \\ f^k - u^k \rightarrow 0, \\ f^k - v^k \rightarrow 0. \end{aligned}$$

Clearly, the KKT conditions (13) except the last three conditions are satisfied at any limit point $\hat{w} = (\hat{f}, \hat{p}, \hat{z}, \hat{u}, \hat{v}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4)$. We continue to verify the last three

conditions in the KKT conditions (13). We recall the subdifferential of $\|Ax\|_2$ [18] is defined by:

$$\partial \|Ax\|_2 = \begin{cases} \{A^T Ax / \|Ax\|_2\}, & \text{if } Ax \neq 0, \\ A^T h : \|h\|_2 \leq 1, & \text{otherwise.} \end{cases}$$

If $\|D_i \hat{f} + \frac{[\hat{\lambda}_1]_i}{\beta}\|_2 > \mu/\beta$, it follows from the shrinkage formula that \hat{p}_i is a non-zero vector and

$$\hat{p}_i = \left(\left\| D_i \hat{f} + \frac{[\hat{\lambda}_1]_i}{\beta} \right\|_2 - \frac{\mu}{\beta} \right) \frac{D_i \hat{f} + \frac{[\hat{\lambda}_1]_i}{\beta}}{\left\| D_i \hat{f} + \frac{[\hat{\lambda}_1]_i}{\beta} \right\|_2}.$$

Hence we have

$$\begin{aligned} \mu \partial \|\hat{p}_i\|_2 - [\hat{\lambda}_1]_i &= \mu \frac{\hat{p}_i}{\|\hat{p}_i\|_2} - [\hat{\lambda}_1]_i = \mu \frac{D_i \hat{f} + \frac{[\hat{\lambda}_1]_i}{\beta}}{\left\| D_i \hat{f} + \frac{[\hat{\lambda}_1]_i}{\beta} \right\|_2} \\ &\quad - (\beta D_i \hat{f} - \beta \hat{p}_i + [\hat{\lambda}_1]_i) = 0. \end{aligned}$$

If $\|D_i \hat{f} + \frac{[\hat{\lambda}_1]_i}{\beta}\|_2 \leq \mu/\beta$, we have $\hat{p}_i = D_i \hat{f} = 0$, which implies $[\hat{\lambda}_1]_i \leq \mu$. Combining $\|\hat{\lambda}_1\|_2 \leq \mu$ and $\mu \partial \|\hat{p}_i\|_2 = \mu h$, we obtain $0 \in \mu \partial \|\hat{p}_i\|_2 - [\hat{\lambda}_1]_i$, where h is any vector satisfying $\|h\|_2 \leq 1$. The dual feasibility and the complementary slackness remain to be verified. Because the non-negativity of $\hat{u} - a_1$ and the non-positivity $\hat{v} - a_2$ are guaranteed by the update formula in Algorithm 1, we only need to verify the non-positivity of $\hat{\lambda}_3$, the non-negativity of $\hat{\lambda}_4$, the complementarity between $\hat{\lambda}_3$ and $\hat{u} - a_1$, and the complementarity between $\hat{\lambda}_4$ and $\hat{v} - a_2$. From the update formula in Algorithm 1, we have

$$\hat{u} - a_1 = \mathcal{P}_+ \left(\hat{f} - a_1 + \frac{\hat{\lambda}_3}{\beta} \right), \quad (17a)$$

$$\hat{v} - a_2 = \mathcal{P}_- \left(\hat{f} - a_2 + \frac{\hat{\lambda}_4}{\beta} \right). \quad (17b)$$

If $[\hat{f} - a_1]_i = [\hat{u} - a_1]_i = 0$, we obtain $\mathcal{P}_+([\frac{\hat{\lambda}_3}{\beta}]_i) = 0$ and $[\hat{\lambda}_3]_i \leq 0$. On the other hand, if $[\hat{f} - a_1]_i = [\hat{u} - a_1]_i > 0$, then (17a) implies $[\hat{\lambda}_3]_i = 0$. The non-positivity of $\hat{\lambda}_3$ and the complementarity between $\hat{\lambda}_3$ and $\hat{u} - a_1$ are proved. The same argument can be applied to prove the non-negativity of $\hat{\lambda}_4$ and the complementarity between $\hat{\lambda}_4$ and $\hat{v} - a_2$. Therefore, the statement of $\{w^k\}$ for problem (5) is established. The statement of $\{f^k\}$ for problem (4) follows directly from the equivalence between problem (4) and problem (5). \blacksquare

ACKNOWLEDGMENTS

The authors would like to thank Dr. Wei Wang (Tongji University) for valuable discussions and Mr. Mo-Tong Qiao (Hong Kong Baptist University) for sharing the HDR toolbox and HDR test images.

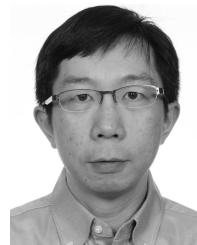
REFERENCES

- [1] H. Andrews and B. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1997.
- [2] B. K. Gunturk and X. Li, *Image Restoration: Fundamentals and Advances*. Boca Raton, FL, USA: CRC Press, 2012.

- [3] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
- [4] 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*. New York, NY, USA: Springer-Verlag, 2006, pp. 17–31.
- [5] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 787–794, 2006.
- [6] B. K. Gunturk, "Handling exposure time in multi-frame image restoration," in *Proc. IEEE ICASSP*, Mar. 2005, pp. 865–868.
- [7] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Comput. Math. Appl.*, vol. 2, no. 1, pp. 17–40, 1976.
- [8] B. R. Hunt, "Bayesian methods in nonlinear digital image restoration," *IEEE Trans. Comput.*, vol. 100, no. 3, pp. 219–229, Mar. 1977.
- [9] L.-Z. Lu, M. K. Ng, and F.-R. Lin, "Approximation BFGS methods for nonlinear image restoration," *J. Comput. Appl. Math.*, vol. 226, no. 1, pp. 84–91, 2009.
- [10] M. K. Ng, R. H. Chan, and W.-C. Tang, "A fast algorithm for deblurring models with Neumann boundary conditions," *SIAM J. Sci. Comput.*, vol. 21, no. 3, pp. 851–866, 1999.
- [11] M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," *SIAM J. Sci. Comput.*, vol. 32, no. 5, pp. 2710–2736, 2010.
- [12] S. Kim, Y.-W. Tai, S. J. Kim, M. S. Brown, and Y. Matsushita, "Nonlinear camera response functions and image deblurring," in *Proc. IEEE Conf. CVPR*, Jun. 2012, pp. 25–32.
- [13] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.
- [14] R. M. Rameshan, S. Chaudhuri, and R. Velmurugan, "High dynamic range imaging under noisy observations," in *Proc. 18th IEEE ICIP*, Sep. 2011, pp. 1333–1336.
- [15] A. M. Tekalp and G. Pavlović, "Restoration of scanned photographic images," in *Digital Image Restoration*, A. K. Katsaggelos, Ed. New York, NY, USA: Springer-Verlag, 1991, pp. 209–239.
- [16] H. J. Trussell and B. R. Hunt, "Improved methods of maximum a posteriori restoration," *IEEE Trans. Comput.*, vol. 28, no. 1, pp. 57–62, Jan. 1979.
- [17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [18] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM J. Imag. Sci.*, vol. 1, no. 3, pp. 248–272, 2008.
- [19] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, "An alternating direction algorithm for matrix completion with nonnegative factors," *Frontiers Math. China*, vol. 7, no. 2, pp. 365–384, 2012.
- [20] M. E. Zervakis and A. N. Venetsanopoulos, "Convergence properties of Gauss–Newton iterative algorithms in nonlinear image restoration," *Multidimensional Syst. Signal Process.*, vol. 2, no. 3, pp. 287–318, Aug. 1992.
- [21] M. E. Zervakis and A. N. Venetsanopoulos, "Iterative algorithms with fast-convergence rates in nonlinear image restoration," *Proc. SPIE*, vol. 1452, pp. 90–103, Jun. 1991.
- [22] M. E. Zervakis and A. N. Venetsanopoulos, "Iterative least squares estimators in nonlinear image restoration," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 927–945, Apr. 1992.
- [23] Y. Zhang, "An alternating direction algorithm for nonnegative matrix factorization," Dept. Comput. Appl. Math., Rice Univ., Houston, TX, USA, Tech. Rep. TR 10-03, 2010.
- [24] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 257–266, 2002.
- [25] D. Bertsekas, A. Nedic, and A. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Scientific, 2003.



Chuan Chen received the B.S. degree in mathematics from Sun Yat-sen University, Guangzhou, China, in 2012. He is currently pursuing the Ph.D. degree with the Department of Mathematics, Hong Kong Baptist University, Hong Kong. His main research interests include data mining, spectral graph, and numerical optimization.



Michael K. Ng is currently a Professor with the Department of Mathematics, Hong Kong Baptist University, Hong Kong. He received the B.Sc. and M.Phil. degrees from the University of Hong Kong, Hong Kong, in 1990 and 1992, respectively, and the Ph.D. degree from the Chinese University of Hong Kong, Hong Kong, in 1995. He was a Research Fellow with the Computer Sciences Laboratory, Australian National University, Canberra, ACT, Australia, from 1995 to 1997, and an Assistant/Associate Professor with the University of Hong Kong before joining Hong Kong Baptist University from 1997 to 2005. His research interests include bioinformatics, data mining, image processing, scientific computing, and data mining. He serves on the Editorial Boards of international journals. He is also the Principal Editor of the *Journal of Computational and Applied Mathematics*.



Xi-Le Zhao received the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2009 and 2012, respectively, where he is currently a Lecturer with the School of Mathematical Sciences. His main research interests include numerical linear algebra and numerical optimization.