Applied Numerical Mathematics ••• (••••) •••–•••

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

An effective alternating direction method of multipliers for color image restoration [☆]

Jianjun Zhang [a], James G. Nagy [b,*]

[a] Department of Mathematics, Shanghai University, Shanghai 200444, China
[b] Department of Mathematics, Emory University, Atlanta, GA 30322, USA

ARTICLE INFO

Article history:
Received 27 April 2020
Received in revised form 29 June 2020
Accepted 7 July 2020
Available online xxxx

Keywords:
Image restoration
Gaussian noise
Total variation
Regularization
Ill-posed problems
Matrix equation

ABSTRACT

Color image restoration is an ill-posed problem, and regularization is necessary. In this paper, we first formulate the color image restoration problem into a constrained minimization problem that minimizes a quadratic functional and subject to a constraint that the total variation of the color image is less than a given parameter δ. The advantages of the constrained minimization problem over the traditional unconstrained one is that the parameter δ has an obvious physical meaning and is easy to select. However solving the constrained minimization problem is generally more difficult than solving the corresponding unconstrained form. We propose an effective alternating direction method of multipliers for color image restoration by using the structure of the problem. We prove the convergence of the method in detail. Experimental results demonstrate that the proposed method is feasible and much more effective for color image restoration.

1. Introduction

Image restoration, and important task in image processing, is the process of reconstructing an image of an unknown scene from an observed distorted image, and it finds many applications in scientific computing and engineering.

Let x be a vector representation of the desired original image, and b be a vector representing the observed image. Then the forward model for image restoration is as follows,

$$b = \mathcal{A}x + e, \tag{1}$$

where $\mathcal{A}$ is a linear operator that describes the distortion process of the imaging system, and e represent additive noise, usually assumed to Gaussian. Image restoration is a broad term used for methods that attempt to recover x from the observed degraded image b. It is well-known that image restoration is a typical ill-posed problem [16,17]. In order to find a meaningful solution for image restoration, a regularization technique is applied to deal with the ill-posedness. One classical regularization technique is to restore the image by minimizing a functional that consists of a data-fitting term and a regularization term:

$$\min_x \left\{ \frac{1}{2}|\mathcal{A}x - b|^2 + \gamma\phi(x) \right\}, \tag{2}$$

☆ Work of J.G. Nagy was funded in part by The National Science Foundation under grant no. DMS-1819042.

* Corresponding author.
E-mail addresses: zhangjj@shu.edu.cn (J. Zhang), jnagy@emory.edu (J.G. Nagy).

https://doi.org/10.1016/j.apnum.2020.07.008
0168-9274/© 2020 IMACS. Published by Elsevier B.V. All rights reserved.

where the regularization term $\phi(x)$ is used to stabilize the solution by enforcing some prior constraints on x, and $\gamma > 0$ is a small regularization parameter that is used to balance the data-fitting term and the regularization term. Here and in the sequel, for any vectors $x, y \in \mathcal{R}^{mn}$, their inner product is defined as $\langle x, y\rangle = \sum_{i=1}^{mn} x_i y_i$, the 2-norm of x is defined as $|x| = \sqrt{\langle x, x\rangle}$. (We purposely use $|\cdot|$ to represent vector 2-norm instead of the standard $\|\cdot\|$ notation.)

Numerous kinds of regularization functionals have been studied; the classical and most well-known is Tikhonov regularization [26]. Image restoration based on Tikhonov regularization is popular because of its computational simplicity and efficiency [14,21]. The downside however is that it may oversmooth the computed solution. To overcome this weakness, edge-preserving regularization techniques have been proposed, including total variation (TV) regularization [24], half-quadratic regularization [3,22], high-order TV based regularization [9], total generalized variation regularization [6], nonlocal TV regularization [15] and structure tensor TV regularization [20].

Image restoration based on TV regularization requires solving the following minimization problem:

$$\min_x \left\{ \frac{1}{2}|\mathcal{A}x - b|^2 + \gamma \mathrm{TV}(x) \right\}, \tag{3}$$

where $\mathrm{TV}(x)$ is the TV norm of x.

Compared to the Tikhonov approach, TV regularization can preserve edges in the image. The price paid is that the TV regularization model (3) is more computationally difficult to solve, but some effective schemes have been proposed. For example, the time-marching method [24], the primal-dual Newton method [8], the variable splitting method [1], primal-dual methods [32], alternating minimization methods [18,27,28], and the alternating direction method of multipliers (ADMM) [5,30].

Aside from the numerical difficulty for the minimization of the TV regularization model, one has to tackle the difficulty of the selection of a suitable regularization parameter. If the regularization parameter is too large, the restored image is too smooth. If the regularization parameter is too small, the restored image will contain more oscillatory artifacts. Finding a suitable value for the regularization parameter is therefore crucial for image restoration.

To circumvent this difficulty, we can rewrite the optimization problem as an equivalent constrained minimization problem,

$$\min_x \frac{1}{2}|\mathcal{A}x - b|^2, \ \text{ subject to } \ \mathrm{TV}(x) \le \delta, \tag{4}$$

where δ is a regularization parameter that controls the smoothness of x.

Mathematically, problems (3) and (4) are equivalent and there is a correspondence between γ and δ. However, from computational point of view, these two problems are very different. For problem (3), the parameter γ does not have a physical meaning, and it is very difficult to determine a suitable value. On the other hand, the parameter δ in problem (4) has an obvious physical meaning that represents the smoothness of the image x. We can select δ by using information of b. The advantages of the constrained minimization (4) over unconstrained one (3) have been observed by many authors. See for example [2,11,23]. However, in general solving problem (4) is more difficult than solving problem (3).

In this paper we consider color image restoration, which we first formulate the problem as a constrained minimization problem that minimizes a quadratic functional subject to a constraint that the total variation of the color image is less than a given parameter δ. As previously mentioned, the advantages of the constrained minimization problem over the traditional unconstrained one is that the parameter δ has an obvious physical meaning and is easy to select. Then by exploiting structure inherent in the color image restoration model, we propose an effective alternating direction method of multipliers to solve the constrained minimization problem. We prove the convergence of the method in detail. Experimental results demonstrate that the proposed method is feasible and much more effective for color image restoration.

The contributions of this paper include: (i) We propose a new effective iteration method for color image restoration by using the structure of the considered problem. (ii) The proposed method is based on the minimization of a quadratic functional, subject to a constraint that the total variation of the color image is less than a given parameter δ, which has obvious meaning and is easy to tune. (iii) Convergence of the proposed method is established.

The rest of this paper is organized as follows. In Section 2, we present the forward model of color image restoration, and formulate the problem as a constrained minimization problem. Our proposed algorithm is presented in Section 3. Convergence analysis of the proposed method is given in Section 4. Experimental results are given in Section 5, and some conclusions are presented in Section 6.

2. Forward model and TV regularization

Pixels of a color image are encoded in three scalar values, namely, red, green and blue, and thus can be represented by a three-dimensional array $\mathcal{X}$ of size $m \times n \times 3$, in which the three channels of red, green and blue are respectively the three two-dimensional arrays $X_r \triangleq \mathcal{X}(:,:,1)$, $X_g \triangleq \mathcal{X}(:,:,2)$, and $X_b \triangleq \mathcal{X}(:,:,3)$. Accordingly, the forward model for color image restoration is much more complicated than grayscale image restoration, since blurs exist either within or across channels.

Please cite this article in press as: J. Zhang, J.G. Nagy, An effective alternating direction method of multipliers for color image restoration, Appl. Numer. Math. (2020), https://doi.org/10.1016/j.apnum.2020.07.008

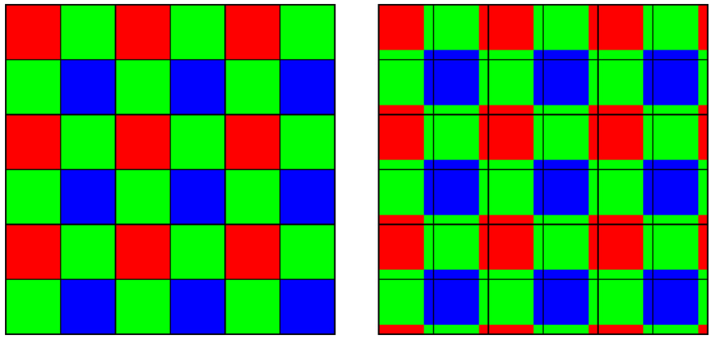

Fig. 1. Left: the Bayer pattern of sensors for recording color images on a single CCD; right: the situation that the Bayer pattern is shifted with respect to the CCD. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

In recording a color image, two strategies are often employed [17]. One strategy is splitting the incoming light in three separate color channels using color filters, and then recording these three color images on three different CCDs (charge-coupled devices). Another strategy is using a single CCD to record color pixels. To that end, a filter mask with a red/green/blue pattern, called a Bayer pattern, is placed on top of the CCD so that light in the different colors is recorded by different sensors on the same CCD. The Bayer pattern uses the property that the human eye is most sensitive to green, less to red and least to blue, and is shown in the left part of Fig. 1.

In either technique, color images recorded may suffer from color artifacts. For example, in the single-CCD approach, this occurs due to the Bayer pattern shift with respect to the CCD, as shown in the right part of Fig. 1. Due to the shift, a sensor meant to record green might be partially misaligned with the red. As a result, in addition to the within-channel blurring of each of the three color layers, there is also a cross-channel blurring among these layers. See [17] for details.

According to [17], assume that the optical blurring or the within-channel blurring takes place before color blurring which is spatially invariant. In addition, we assume that in all three channels the blurring is the same. To simplify the notation, we let $\boldsymbol{A}$ be the matrix that represents within-channel blurring, and define

$$\boldsymbol{A}_c = \begin{pmatrix} a_{rr} & a_{rg} & a_{rb} \\ a_{gr} & a_{gg} & a_{gb} \\ a_{br} & a_{bg} & a_{bb} \end{pmatrix}$$

to represent cross-channel blurring. Then the model for color blurring takes the following form:

$$\big(\mathrm{vec}(B_r), \mathrm{vec}(B_g), \mathrm{vec}(B_b)\big) = \big(\boldsymbol{A}\mathrm{vec}(X_r)\ \boldsymbol{A}\mathrm{vec}(X_g)\ \boldsymbol{A}\mathrm{vec}(X_b)\big)\, \boldsymbol{A}_c^T \tag{5}$$

where the three-dimensional array $\mathcal{B}$ of size $m \times n \times 3$ is the observed color image, and $B_r \triangleq \mathcal{B}(:,:,1)$, $B_g \triangleq \mathcal{B}(:,:,2)$, and $B_b \triangleq \mathcal{B}(:,:,3)$, and for any m-by-n matrix $M = (m_1, m_2, \cdots, m_n)$, $\mathrm{vec}(M) \triangleq (m_1^T, m_2^T, \cdots, m_n^T)^T$.

Let $x_r = \mathrm{vec}(X_r)$, $x_g = \mathrm{vec}(X_g)$, $x_b = \mathrm{vec}(X_b)$, $b_r = \mathrm{vec}(B_r)$, $b_g = \mathrm{vec}(B_g)$, $b_b = \mathrm{vec}(B_b)$, and set $\boldsymbol{X} = (x_r, x_g, x_b)$, $\boldsymbol{B} = (b_r, b_g, b_b)$. Then equation (5) can be written as the following matrix equation

$$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T.$$

In addition to color blurring, there is noise during the image formation process, and so the forward model for color image restoration then has the following form

$$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T + \boldsymbol{E}, \tag{6}$$

where we assume in this paper that $\boldsymbol{E}$ additive Gaussian white noise. This forward model for color image restoration was also studied in [4]. In order to stabilize the solution to equation (6), we rewrite it in the form (4). For this purpose, we need to define TV for a color image. There are different definitions of TV for color image, here we use multichannel TV [7,10,25,29,31].

For a vector $u \in \mathcal{R}^{mn}$, let u_{ij} represent u's $(i + (j-1)m)$th element. Let $\mathcal{I} \triangleq \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$. Then, assuming periodic boundary conditions, the discrete derivatives $\nabla_x u \in \mathcal{R}^{mn}$ of u in the horizontal and $\nabla_y u \in \mathcal{R}^{mn}$ in the vertical directions respectively, are defined by

$$(\nabla_x u)_{ij} = \begin{cases} u_{i,j+1} - u_{ij}, & 1 \le j < n, \\ u_{i1} - u_{in}, & j = n, \end{cases}$$

and

$$(\nabla_y u)_{ij} = \begin{cases} u_{i+1,j} - u_{ij}, & 1 \le i < n, \\ u_{1j} - u_{mj}, & i = m, \end{cases}$$

for $(i,j)\in\mathcal{I}$. For color image $\boldsymbol{X}$ written in the form $\boldsymbol{X}=(x_r,x_g,x_b)\in\mathcal{R}^{mn\times 3}$, we define the discrete gradient operator $\nabla:\mathcal{R}^{mn\times 3}\to\mathcal{R}^{mn\times 6}$ by $\nabla\boldsymbol{X}=(\nabla_x x_r,\nabla_y x_r,\nabla_x x_g,\nabla_y x_g,\nabla_x x_b,\nabla_y x_b)$. Then the discrete multichannel TV of $\boldsymbol{X}$, is defined by

$$\|\boldsymbol{X}\|_{MTV}:=\sum_{(i,j)\in\mathcal{I}}\sqrt{\sum_{k\in\{r,g,b\}}|(\nabla_x x_k)_{ij}|^2+|(\nabla_y x_k)_{ij}|^2}.$$

For any space $\mathcal{R}^{mn\times l}$ $(l\ge 2)$, we use the inner product $\langle\cdot,\cdot\rangle_F$ and norm $\|\cdot\|_F$, which are defined as follows: for $\boldsymbol{P}=(p^1,p^2,\cdots,p^l)\in\mathcal{R}^{mn\times l}$ and $\boldsymbol{Q}=(q^1,q^2,\cdots,q^l)\in\mathcal{R}^{mn\times l}$, $\langle\boldsymbol{P},\boldsymbol{Q}\rangle_F=\sum_{k=1}^{l}\langle p^k,q^k\rangle$ and $\|\boldsymbol{P}\|_F=\sqrt{\langle\boldsymbol{P},\boldsymbol{P}\rangle_F}$. Clearly $\|\boldsymbol{P}\|_F$ is the traditional Frobenius norm. Define $\|\boldsymbol{P}\|_1=\sum_{(i,j)\in\mathcal{I}}|\boldsymbol{P}_{ij}|$, where $|\boldsymbol{P}_{ij}|=|(p^1_{ij},p^2_{ij},\cdots,p^l_{ij})|=\sqrt{\sum_{k=1}^{l}(p^k_{ij})^2}$ for $(i,j)\in\mathcal{I}$. Using these notations, we have $\|\boldsymbol{X}\|_{MTV}=\|\nabla\boldsymbol{X}\|_1$, and we formulate the color image restoration problem as the following minimization problem.

$$\begin{cases}\min\dfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T-\boldsymbol{B}\|_F^2,\\ \text{subject to}\quad \|\nabla\boldsymbol{X}\|_1\le\delta,\end{cases}\tag{7}$$

or

$$\min\frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T-\boldsymbol{B}\|_F^2+\iota_{(\|\nabla\boldsymbol{X}\|_1\le\delta)}(\boldsymbol{X}),\tag{8}$$

where $\iota_\Omega(\omega)$ is the indicator function of Ω, i.e.,

$$\iota_\Omega(\omega)=\begin{cases}0, & \text{for }\ \omega\in\Omega,\\ +\infty, & \text{otherwise.}\end{cases}$$

3. Our method

In this section, we design an effective method for solving (8).

By introducing new variables $\boldsymbol{Y}$ and $\boldsymbol{Z}$, we can write the minimization problem (8) into

$$\begin{cases}\min\dfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T-\boldsymbol{B}\|_F^2+\iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z}),\\ \text{subject to}\quad \boldsymbol{X}=\boldsymbol{Y},\\ \qquad\qquad\quad\ \boldsymbol{Z}=\nabla\boldsymbol{Y}.\end{cases}\tag{9}$$

Now attaching the Lagrange multipliers $\boldsymbol{\Lambda_X}\in R^{mn\times 3}$ to the linear constraints $\boldsymbol{X}=\boldsymbol{Y}$, and $\boldsymbol{\Lambda_Z}\in R^{mn\times 6}$ to $\boldsymbol{Z}=\nabla\boldsymbol{Y}$, we get the augmented Lagrangian (AL) function of (9):

$$\begin{aligned}\mathcal{L}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{\Lambda_X},\boldsymbol{\Lambda_Z})=\ &\tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T-\boldsymbol{B}\|_F^2+\iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z})+\langle\boldsymbol{\Lambda_X},\boldsymbol{X}-\boldsymbol{Y}\rangle_F\\ &+\tfrac{\beta}{2}\|\boldsymbol{X}-\boldsymbol{Y}\|_F^2+\langle\boldsymbol{\Lambda_Z},\boldsymbol{Z}-\nabla\boldsymbol{Y}\rangle_F+\tfrac{\mu}{2}\|\boldsymbol{Z}-\nabla\boldsymbol{Y}\|_F^2,\end{aligned}$$

where $\beta>0,\mu>0$ are the penalty parameters for the violation of the constraint $\boldsymbol{X}=\boldsymbol{Y}$ and $\boldsymbol{Z}=\nabla\boldsymbol{Y}$. Hence, given initials $\boldsymbol{Y}^{(0)}$, $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(0)}$ and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(0)}$, ADMM for problem (9) computes a sequence of iterates $\boldsymbol{X}^{(1)}$, $\boldsymbol{Z}^{(1)}$, $\boldsymbol{Y}^{(1)}$, $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(1)}$, $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(1)}$, $\boldsymbol{X}^{(2)}$, $\boldsymbol{Z}^{(2)}$, $\boldsymbol{Y}^{(2)}$, $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(2)}$, $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(2)}$, $\cdots$, $\boldsymbol{X}^{(k)}$, $\boldsymbol{Z}^{(k)}$, $\boldsymbol{Y}^{(k)}$, $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}$ and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}$, $\cdots$, such that

$$\boldsymbol{X}^{(k+1)}=\underset{\boldsymbol{X}}{\operatorname{argmin}}\left\{\frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T-\boldsymbol{B}\|_F^2+\langle\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)},\boldsymbol{X}-\boldsymbol{Y}^{(k)}\rangle_F+\frac{\beta}{2}\|\boldsymbol{X}-\boldsymbol{Y}^{(k)}\|_F^2\right\},\tag{10}$$

$$\boldsymbol{Z}^{(k+1)}=\underset{\boldsymbol{Z}}{\operatorname{argmin}}\left\{\iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z})+\frac{\mu}{2}\|\boldsymbol{Z}-\nabla\boldsymbol{Y}^{(k)}\|_F^2+\langle\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)},\boldsymbol{Z}-\nabla\boldsymbol{Y}^{(k)}\rangle_F\right\},\tag{11}$$

$$\begin{aligned}\boldsymbol{Y}^{(k+1)}=\ \underset{\boldsymbol{Y}}{\operatorname{argmin}}&\left\{\frac{\beta}{2}\|\boldsymbol{X}^{(k+1)}-\boldsymbol{Y}\|_F^2+\langle\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)},\boldsymbol{X}^{(k+1)}-\boldsymbol{Y}\rangle_F\right.\\ &\left.+\ \tfrac{\mu}{2}\|\boldsymbol{Z}^{(k+1)}-\nabla\boldsymbol{Y}\|_F^2+\langle\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)},\boldsymbol{Z}^{(k+1)}-\nabla\boldsymbol{Y}\rangle_F\right\},\end{aligned}\tag{12}$$

and

$$\begin{aligned}\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k+1)}&=\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}+\beta(\boldsymbol{X}^{(k+1)}-\boldsymbol{Y}^{(k+1)}),\\ \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k+1)}&=\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}+\mu(\boldsymbol{Z}^{(k+1)}-\nabla\boldsymbol{Y}^{(k+1)}).\end{aligned}$$

Please cite this article in press as: J. Zhang, J.G. Nagy, An effective alternating direction method of multipliers for color image restoration, Appl. Numer. Math. (2020), https://doi.org/10.1016/j.apnum.2020.07.008

We now consider solving the above three subproblems.

For problem (10), by setting the gradient of

$$\frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \frac{\beta}{2}\|\boldsymbol{X} - \boldsymbol{Y}^{(k)}\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}, \boldsymbol{X} - \boldsymbol{Y}^{(k)}\rangle_F$$

with respect to $\boldsymbol{X}$ to zero, we have

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T\boldsymbol{A}_c + \beta\boldsymbol{X} = \boldsymbol{A}^T\boldsymbol{B}\boldsymbol{A}_c + \beta\boldsymbol{Y}^{(k)} - \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} \triangleq \boldsymbol{R}^{(k)}. \tag{13}$$

Let the singular value decomposition of matrix $\boldsymbol{A}_c$ be $\boldsymbol{A}_c = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$. Then (13) can be written as

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{X}\boldsymbol{V}\boldsymbol{\Sigma}^2\boldsymbol{V}^T + \beta\boldsymbol{X} = \boldsymbol{R}^{(k)},$$

or

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{X}\boldsymbol{V}\boldsymbol{\Sigma}^2 + \beta\boldsymbol{X}\boldsymbol{V} = \boldsymbol{R}^{(k)}\boldsymbol{V}. \tag{14}$$

Let $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$, $\boldsymbol{T} = \boldsymbol{X}\boldsymbol{V} = (t_1, t_2, t_3)$, and $\boldsymbol{G}^{(k)} = \boldsymbol{R}^{(k)}\boldsymbol{V} = (g_1^{(k)}, g_2^{(k)}, g_3^{(k)})$, then (14) can be written as

$$\begin{cases} (\sigma_1^2\boldsymbol{A}^T\boldsymbol{A} + \beta\boldsymbol{I})t_1 = g_1^{(k)}, \\ (\sigma_2^2\boldsymbol{A}^T\boldsymbol{A} + \beta\boldsymbol{I})t_2 = g_2^{(k)}, \\ (\sigma_3^2\boldsymbol{A}^T\boldsymbol{A} + \beta\boldsymbol{I})t_3 = g_3^{(k)}. \end{cases} \tag{15}$$

Under the periodic boundary condition for X_r, X_g and X_b, $\boldsymbol{A}^T\boldsymbol{A}$ is a block circulant with circulant blocks matrix. Therefore, the coefficient matrices on the left-hand sides of (15) can be diagonalized by 2D discrete Fourier transform $\mathcal{F}$. Let $\mathcal{F}(\boldsymbol{X})$ denote the discrete Fourier transform of $\boldsymbol{X}$. Let the symbol "$\circ$" denote the element-wise multiplication, and the division is taken element-wise. Let $g_i^{(k)} = \mathrm{vec}(G_i^{(k)})$, $i = 1, 2, 3$. Then we can write the solution to (15) as

$$t_i = \mathrm{vec}\left(\mathcal{F}^{-1}\left(\frac{\mathcal{F}(G_i^{(k)})}{\sigma_i^2\mathcal{F}(\boldsymbol{A})^* \circ \mathcal{F}(\boldsymbol{A}) + \beta\boldsymbol{1}}\right)\right), \text{ for } i = 1, 2, 3, \tag{16}$$

where the super script "*" denotes complex conjugacy and $\boldsymbol{1}$ is a matrix of the same size with $\mathcal{F}(\boldsymbol{A})$, with all its elements being 1. Accordingly, $\boldsymbol{X}^{(k+1)} = \boldsymbol{T}\boldsymbol{V}^T = [t_1, t_2, t_3]\boldsymbol{V}^T$.

The minimization problem (11) seems computationally challenging, but if we define $\boldsymbol{W}^{(k)} = \nabla\boldsymbol{Y}^{(k)} - \frac{\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}}{\mu}$, then (11) can be written as

$$\boldsymbol{Z}^{(k+1)} = \operatorname*{argmin}_{\boldsymbol{Z}}\left\{\iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z}) + \frac{\mu}{2}\|\boldsymbol{Z} - \boldsymbol{W}^{(k)}\|_F^2\right\}. \tag{17}$$

We note that $\boldsymbol{W}^{(k)} \in \mathcal{R}^{mn\times 6}$. Let $\boldsymbol{W}_i^{(k)}$ be the ith row vector of $\boldsymbol{W}^{(k)}$, and let $|\boldsymbol{W}_{(i)}^{(k)}|$ be the order statics of the norms $|\boldsymbol{W}_i^{(k)}|$ for $i = 1, 2, \cdots, mn$. That is, $|\boldsymbol{W}_{(1)}^{(k)}| \ge |\boldsymbol{W}_{(2)}^{(k)}| \ge \cdots \ge |\boldsymbol{W}_{(mn)}^{(k)}|$. Let I be the largest number such that

$$\sum_{i=1}^{I}|\boldsymbol{W}_{(i)}^{(k)}| - \delta < I|\boldsymbol{W}_{(I)}^{(k)}|,$$

and define $\lambda = \frac{1}{I}\left(\sum_{i=1}^{I}|\boldsymbol{W}_{(i)}^{(k)}| - \delta\right)$. Then problem (17) has a closed form solution:

$$\boldsymbol{Z}_i^{(k+1)} = \begin{cases} \max\left(\left(1 - \frac{\lambda}{|\boldsymbol{W}_i^{(k)}|}\right), 0\right)\boldsymbol{W}_i^{(k)}, & \text{if } \boldsymbol{W}_i^{(k)} \ne 0, \\ 0, & \text{otherwise}, \end{cases} \tag{18}$$

for $i = 1, 2, \cdots, mn$. The details are presented in the Appendix of this paper.

For problem (12), if we set the gradient of

$$\frac{\beta}{2}\|\boldsymbol{X}^{(k+1)} - \boldsymbol{Y}\|_F^2 + \langle\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}, \boldsymbol{X}^{(k+1)} - \boldsymbol{Y}\rangle_F + \frac{\mu}{2}\|\boldsymbol{Z}^{(k+1)} - \nabla\boldsymbol{Y}\|_F^2 + \langle\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \boldsymbol{Z}^{(k+1)} - \nabla\boldsymbol{Y}\rangle_F$$

with respect to $\boldsymbol{Y}$ to be zero, then the solution $\boldsymbol{Y}^{(k+1)}$ solves the following matrix equation:

$$(\mu\nabla^T\nabla + \beta I)\boldsymbol{Y} = \mu\nabla^T\boldsymbol{Z}^{(k+1)} + \nabla^T\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)} + (\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} + \beta\boldsymbol{X}^{(k+1)}). \tag{19}$$

So we can write the solution as follows:

$$\boldsymbol{Y}^{(k+1)} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(\mu\nabla^T \boldsymbol{Z}^{(k+1)} + \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)} + \left(\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} + \beta \boldsymbol{X}^{(k+1)}\right)\right)}{\mathcal{F}(\nabla_x)^* \circ \mathcal{F}(\nabla_x) + \mathcal{F}(\nabla_y)^* \circ \mathcal{F}(\nabla_y) + \beta \mathbf{1}}\right).$$

Based on the above discussion, we give the following algorithm.

Algorithm 1. ADMM for color image restoration.

1. Let $\boldsymbol{A}$, $\boldsymbol{A}_c$ and $\boldsymbol{B}$ be given. Let $\boldsymbol{A}_c = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ be singular value decomposition of $\boldsymbol{A}_c$. Given constants $\beta, \mu \in \mathcal{R}$, and given initial $\boldsymbol{Y}^{(0)} \in \mathcal{R}^{mn\times 3}$, $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(0)} \in \mathcal{R}^{mn\times 3}$, and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(0)} \in \mathcal{R}^{mn\times 6}$.
2. For $k = 0, 1, 2, \ldots,$
   - Compute $\boldsymbol{R}^{(k)} = \boldsymbol{A}^T \boldsymbol{B} \boldsymbol{A}_c + \beta \boldsymbol{Y}^{(k)} - \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}$.
   - Compute $\boldsymbol{G}^{(k)} = \boldsymbol{R}^{(k)} \boldsymbol{V} = (g_1^{(k)}, g_2^{(k)}, g_3^{(k)})$.
   - Compute $\boldsymbol{t}_i^{(k)} = \text{vec}\left(\mathcal{F}^{-1}\left(\frac{\mathcal{F}(G_i^{(k)})}{\sigma_i^2 \mathcal{F}(\boldsymbol{A})^* \circ \mathcal{F}(\boldsymbol{A}) + \beta \mathbf{1}}\right)\right)$, for $i = 1, 2, 3$.
   - Compute $\boldsymbol{X}^{(k+1)} = \boldsymbol{T}^{(k)} \boldsymbol{V}^T$, where $\boldsymbol{T}^{(k)} = (t_1^{(k)}, t_2^{(k)}, t_3^{(k)})$.
   - Compute $\boldsymbol{W}^{(k)} = \nabla \boldsymbol{Y}^{(k)} - \frac{\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}}{\mu}$.
   - Compute $\boldsymbol{Z}^{(k+1)}$ according to (18).
   - Compute $\boldsymbol{Y}^{(k+1)} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(\mu\nabla^T \boldsymbol{Z}^{(k+1)} + \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)} + \left(\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} + \beta \boldsymbol{X}^{(k+1)}\right)\right)}{\mathcal{F}(\nabla_x)^* \circ \mathcal{F}(\nabla_x) + \mathcal{F}(\nabla_y)^* \circ \mathcal{F}(\nabla_y) + \beta \mathbf{1}}\right)$.
   - $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k+1)} = \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} + \beta(\boldsymbol{X}_{\boldsymbol{X}}^{(k+1)} - \boldsymbol{Y}^{(k+1)})$.
   - $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k+1)} = \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)} + \mu(\boldsymbol{Z}^{(k+1)} - \nabla \boldsymbol{Y}^{(k+1)})$.

4. Convergence analysis

In this section, we establish the convergence theory of our proposed algorithm.

Definition 1. [13] Let Ω, Π be nonempty convex subset of $\mathcal{R}^N$ and $\mathcal{R}^K$. $(u^*, v^*) \in \Omega \times \Pi$ is called a saddle point of function ψ defined on $\Omega \times \Pi$, if for all $u \in \Omega$ and $v \in \Pi$

$$\psi(u^*, v) \le \psi(u^*, v^*) \le \psi(u, v^*). \tag{20}$$

Lemma 1. *[13] Let $\psi(u, v) = \psi_1(u, v) + \psi_2(u, v)$ be defined on $\Omega \times \Pi$ with*

$$\begin{aligned} &\forall u \in \Omega,\ \phi_2(v) = \psi_2(u, v) \text{ is a concave differentiable function on } \Pi, \\ &\forall v \in \Pi,\ \phi_1(u) = \psi_2(u, v) \text{ is a convex differentiable function on } \Omega. \end{aligned}$$

Then $(u^*, v^*) \in \Omega \times \Pi$ is a saddle point of function ψ if and only if

$$\begin{aligned} &\forall u \in \Omega,\ \psi_1(u, v^*) - \psi_1(u^*, v^*) + (u - u^*)^T \nabla_u \psi_2(u^*, v^*) \ge 0, \\ &\forall v \in \Pi,\ \psi_1(u^*, v^*) - \psi_1(u^*, v) - (v - v^*)^T \nabla_v \psi_2(u^*, v^*) \ge 0. \end{aligned} \tag{21}$$

Remark 1. From the proof of Lemma 1 in [13], we see that the right inequality of (20) is satisfied if and only if (21) is satisfied.

Lemma 2. *$\boldsymbol{X}^* \in \mathcal{R}^{mn\times 3}$ is a solution of (8) if and only if there exist $\boldsymbol{Y}^*$, $\boldsymbol{Z}^*$, $\boldsymbol{\Lambda}_{\boldsymbol{X}}^*$, and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^*$ such that*

$$\mathcal{L}(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{Z}^*, \boldsymbol{\Lambda}_{\boldsymbol{X}}, \boldsymbol{\Lambda}_{\boldsymbol{Z}}) \le \mathcal{L}(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{Z}^*, \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*) \le \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*). \tag{22}$$

Proof. ($\Leftarrow$) Suppose that $(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{Z}^*, \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*)$ is a solution of (22). Then from the left inequality of (22), we have that

$$\begin{aligned} &\tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \le \delta)}(\boldsymbol{Z}^*) + \langle \boldsymbol{\Lambda}_{\boldsymbol{X}}, \boldsymbol{X}^* - \boldsymbol{Y}^*\rangle_F \\ &+ \tfrac{\beta}{2}\|\boldsymbol{X}^* - \boldsymbol{Y}^*\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}, \boldsymbol{Z}^* - \nabla \boldsymbol{Y}^*\rangle_F + \tfrac{\mu}{2}\|\boldsymbol{Z}^* - \nabla \boldsymbol{Y}^*\|_F^2 \\ \le\ &\tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \le \delta)}(\boldsymbol{Z}^*) + \langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{X}^* - \boldsymbol{Y}^*\rangle_F \\ &+ \tfrac{\beta}{2}\|\boldsymbol{X}^* - \boldsymbol{Y}^*\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*, \boldsymbol{Z}^* - \nabla \boldsymbol{Y}^*\rangle_F + \tfrac{\mu}{2}\|\boldsymbol{Z}^* - \nabla \boldsymbol{Y}^*\|_F^2, \end{aligned}$$

from which we have

$$\langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^* - \boldsymbol{\Lambda}_{\boldsymbol{X}}, \boldsymbol{X}^* - \boldsymbol{Y}^* \rangle_F + \langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}^* - \boldsymbol{\Lambda}_{\boldsymbol{Z}}, \boldsymbol{Z}^* - \nabla \boldsymbol{Y}^* \rangle_F \geq 0.$$

Since $\boldsymbol{\Lambda}_{\boldsymbol{X}}$ and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}$ can take any elements in $\mathcal{R}^{mn\times 3}$ and $\mathcal{R}^{mn\times 6}$, we get $\boldsymbol{X}^* = \boldsymbol{Y}^*$, and $\boldsymbol{Z}^* = \nabla \boldsymbol{Y}^*$.

Now by using the right inequality of (22), we have that

$$\begin{aligned} & \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^*) + \langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{X}^* - \boldsymbol{Y}^* \rangle_F \\ & + \tfrac{\beta}{2}\|\boldsymbol{X}^* - \boldsymbol{Y}^*\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*, \boldsymbol{Z}^* - \nabla \boldsymbol{Y}^* \rangle_F + \tfrac{\mu}{2}\|\boldsymbol{Z}^* - \nabla \boldsymbol{Y}^*\|_F^2 \\ \leq \; & \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}) + \langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{X} - \boldsymbol{Y} \rangle_F \\ & + \tfrac{\beta}{2}\|\boldsymbol{X} - \boldsymbol{Y}\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*, \boldsymbol{Z} - \nabla \boldsymbol{Y} \rangle_F + \tfrac{\mu}{2}\|\boldsymbol{Z} - \nabla \boldsymbol{Y}\|_F^2, \end{aligned}$$

from which by taking $\boldsymbol{Y} = \boldsymbol{X}$, $\boldsymbol{Z} = \nabla \boldsymbol{Y} = \nabla \boldsymbol{X}$, we have

$$\frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\nabla \boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}) \geq \frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\nabla \boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}^*).$$

So $\boldsymbol{X}^*$ is a solution of (8).

($\Rightarrow$) Suppose $\boldsymbol{X}^*$ is a solution of (8), then $\forall\ \boldsymbol{X} \in \mathcal{R}^{mn\times 3}$,

$$\frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\nabla \boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}) \geq \frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\nabla \boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}^*),$$

and so

$$0 \in \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B})\boldsymbol{A}_c + \nabla^T \partial \iota_{(\|\nabla \boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}^*).$$

So there exist $\boldsymbol{\Lambda}_{\boldsymbol{X}}^*$ and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^*$ such that,

$$\boldsymbol{\Lambda}_{\boldsymbol{X}}^* = -\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B})\boldsymbol{A}_c,\ -\boldsymbol{\Lambda}_{\boldsymbol{Z}}^* \in \partial \iota_{(\|\nabla \boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}),\ \text{and}\ \boldsymbol{\Lambda}_{\boldsymbol{X}}^* + \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^* = 0. \tag{23}$$

By taking $\boldsymbol{Y}^* = \boldsymbol{X}^*$, $\boldsymbol{Z}^* = \nabla \boldsymbol{Y}^*$, we immediately obtain the left inequality of (22).

We now consider the function $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*)$. Since $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*)$ is convex on $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$, to prove the right inequality of (22), by Remark 1 we just need to prove that

$$\begin{aligned} & \left\langle \begin{pmatrix} \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B})\boldsymbol{A}_c + \beta(\boldsymbol{X}^* - \boldsymbol{Y}^*) + \boldsymbol{\Lambda}_{\boldsymbol{X}}^* \\ \beta(\boldsymbol{Y}^* - \boldsymbol{X}^*) - \boldsymbol{\Lambda}_{\boldsymbol{X}}^* + \mu \nabla^T(\nabla \boldsymbol{Y}^* - \boldsymbol{Z}^*) - \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^* \\ \mu(\boldsymbol{Z}^* - \nabla \boldsymbol{Y}^*) + \boldsymbol{\Lambda}_{\boldsymbol{Z}}^* \end{pmatrix}, \begin{pmatrix} \boldsymbol{X} - \boldsymbol{X}^* \\ \boldsymbol{Y} - \boldsymbol{Y}^* \\ \boldsymbol{Z} - \boldsymbol{Z}^* \end{pmatrix} \right\rangle_F \\ & + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}) - \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^*) \geq 0. \end{aligned}$$

From (23), and using $\boldsymbol{Y}^* = \boldsymbol{X}^*$, $\boldsymbol{Z}^* = \nabla \boldsymbol{Y}^*$, the above inequality is equivalent to

$$\iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}) - \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^*) - \langle -\boldsymbol{\Lambda}_{\boldsymbol{Z}}^*, \boldsymbol{Z} - \boldsymbol{Z}^* \rangle_F \geq 0,$$

which is true since $\iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z})$ is a convex function on $\boldsymbol{Z}$, and $-\boldsymbol{\Lambda}_{\boldsymbol{Z}}^* \in \partial \iota_{(\|\nabla \boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}^*)$ and $\boldsymbol{Z}^* = \nabla \boldsymbol{X}^*$. □

By using (22) and Remark 1, we have that, $(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{Z}^*, \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*)$ satisfies

$$\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B})\boldsymbol{A}_c + \beta(\boldsymbol{X}^* - \boldsymbol{Y}^*) + \boldsymbol{\Lambda}_{\boldsymbol{X}}^* = 0, \tag{24}$$

$$\beta(\boldsymbol{Y}^* - \boldsymbol{X}^*) - \boldsymbol{\Lambda}_{\boldsymbol{X}}^* + \mu \nabla^T(\nabla \boldsymbol{Y}^* - \boldsymbol{Z}^*) - \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^* = 0, \tag{25}$$

$$\iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}) - \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^*) + \langle \mu(\boldsymbol{Z}^* - \nabla \boldsymbol{Y}^*) + \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*, \boldsymbol{Z} - \boldsymbol{Z}^* \rangle_F \geq 0. \tag{26}$$

In the following, we present a convergence result for Algorithm (1).

Theorem 1. *Let $\boldsymbol{X}^* \in \mathcal{R}^{mn\times 3}$ be a solution of (8). Then the sequence $(\boldsymbol{X}^{(k)}, \boldsymbol{Z}^{(k)})$ generated by Algorithm 1 is a minimization sequence of (8). That is, it satisfies*

$$\lim_{k\to\infty} \left\{ \frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^{(k)}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^{(k)}) \right\} = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}^*).$$

Moreover, if $\boldsymbol{A}$ and $\boldsymbol{A}_c$ are nonsingular, then we have $\lim_{k\to\infty} \boldsymbol{X}^{(k)} = \boldsymbol{X}^*$, $\lim_{k\to\infty} \boldsymbol{Y}^{(k)} = \boldsymbol{X}^*$, and $\lim_{k\to\infty} \boldsymbol{Z}^{(k)} = \nabla \boldsymbol{X}^*$.

Please cite this article in press as: J. Zhang, J.G. Nagy, An effective alternating direction method of multipliers for color image restoration, Appl. Numer. Math. (2020), https://doi.org/10.1016/j.apnum.2020.07.008

Proof. Let $(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{Z}^*, \boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*)$ be a solution of (22), then from the proof of Lemma 2, we know that, $\boldsymbol{Y}^* = \boldsymbol{X}^*$, $\boldsymbol{Z}^* = \nabla \boldsymbol{Y}^*$. Let $\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)} = \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} - \boldsymbol{\Lambda}_{\boldsymbol{X}}^*$, $\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)} = \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)} - \boldsymbol{\Lambda}_{\boldsymbol{Z}}^*$, $\overline{\boldsymbol{X}}^{(k)} = \boldsymbol{X}^{(k)} - \boldsymbol{X}^*$, $\overline{\boldsymbol{Y}}^{(k)} = \boldsymbol{Y}^{(k)} - \boldsymbol{Y}^*$, and $\overline{\boldsymbol{Z}}^{(k)} = \boldsymbol{Z}^{(k)} - \boldsymbol{Z}^*$. From the updates $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k+1)}$ and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k+1)}$ of Algorithm 1, we have that $\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k+1)} = \overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)} + \beta\left(\overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k+1)}\right)$, and $\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k+1)} = \overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)} + \mu\left(\overline{\boldsymbol{Z}}^{(k+1)} - \nabla\overline{\boldsymbol{Y}}^{(k+1)}\right)$. These two equalities are equivalent to $\sqrt{\mu}\,\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k+1)} = \sqrt{\mu}\,\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)} + \sqrt{\mu}\beta\left(\overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k+1)}\right)$, and $\sqrt{\beta}\,\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k+1)} = \sqrt{\beta}\,\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)} + \sqrt{\beta}\mu\left(\overline{\boldsymbol{Z}}^{(k+1)} - \nabla\overline{\boldsymbol{Y}}^{(k+1)}\right)$, from which we can get

$$\mu\left(\|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k+1)}\|_F^2 - \|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)}\|_F^2\right) = 2\mu\beta\langle\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)}, \overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k+1)}\rangle_F + \mu\beta^2\|\overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k+1)}\|_F^2, \tag{27}$$

and

$$\beta\left(\|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k+1)}\|_F^2 - \|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}\|_F^2\right) = 2\mu\beta\langle\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}, \overline{\boldsymbol{Z}}^{(k+1)} - \nabla\overline{\boldsymbol{Y}}^{(k+1)}\rangle_F + \mu^2\beta\|\overline{\boldsymbol{Z}}^{(k+1)} - \nabla\overline{\boldsymbol{Y}}^{(k+1)}\|_F^2. \tag{28}$$

From (24) and (25), we get

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T\boldsymbol{A}_c + \beta\boldsymbol{X}^* = \boldsymbol{A}^T\boldsymbol{B}\boldsymbol{A}_c + \beta\boldsymbol{Y}^* - \boldsymbol{\Lambda}_{\boldsymbol{X}}^*,$$

and

$$(\mu\nabla^T\nabla + \beta I)\boldsymbol{Y}^* = \mu\nabla^T\boldsymbol{Z}^* + \nabla^T\boldsymbol{\Lambda}_{\boldsymbol{Z}}^* + \boldsymbol{\Lambda}_{\boldsymbol{X}}^* + \beta\boldsymbol{X}^*.$$

Notice that $\boldsymbol{X}^{(k+1)}$ is the solution of (13), and $\boldsymbol{Y}^{(k+1)}$ is the solution of (19), so we get that

$$\boldsymbol{A}^T\boldsymbol{A}\overline{\boldsymbol{X}}^{(k+1)}\boldsymbol{A}_c^T\boldsymbol{A}_c + \beta\overline{\boldsymbol{X}}^{(k+1)} = \beta\overline{\boldsymbol{Y}}^{(k)} - \overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)},$$

and

$$(\mu\nabla^T\nabla + \beta I)\overline{\boldsymbol{Y}}^{(k+1)} = \mu\nabla^T\overline{\boldsymbol{Z}}^{(k+1)} + \nabla^T\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)} + \overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)} + \beta\overline{\boldsymbol{X}}^{(k+1)}.$$

This then allows us to conclude that

$$\|\boldsymbol{A}\overline{\boldsymbol{X}}^{(k+1)}\boldsymbol{A}_c^T\|_F^2 + \beta\|\overline{\boldsymbol{X}}^{(k+1)}\|_F^2 = \beta\langle\overline{\boldsymbol{X}}^{(k+1)}, \overline{\boldsymbol{Y}}^{(k)}\rangle_F - \langle\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)}, \overline{\boldsymbol{X}}^{(k+1)}\rangle_F, \tag{29}$$

and

$$\begin{aligned}\mu\|\nabla\overline{\boldsymbol{Y}}^{(k+1)}\|_F^2 + \beta\|\overline{\boldsymbol{Y}}^{(k+1)}\|_F^2 = \quad & \beta\langle\overline{\boldsymbol{X}}^{(k+1)}, \overline{\boldsymbol{Y}}^{(k+1)}\rangle_F + \langle\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)}, \overline{\boldsymbol{Y}}^{(k+1)}\rangle_F \\ & + \mu\langle\nabla^T\overline{\boldsymbol{Z}}^{(k+1)}, \overline{\boldsymbol{Y}}^{(k+1)}\rangle_F + \langle\nabla^T\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}, \overline{\boldsymbol{Y}}^{(k+1)}\rangle_F.\end{aligned} \tag{30}$$

On the other hand, by (11) we get

$$\iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z}) - \iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z}^{(k+1)}) + \langle\mu\left(\boldsymbol{Z}^{(k+1)} - \nabla\boldsymbol{Y}^{(k)}\right) + \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \boldsymbol{Z} - \boldsymbol{Z}^{(k+1)}\rangle_F \ge 0. \tag{31}$$

Taking $\boldsymbol{Z} = \boldsymbol{Z}^{(k+1)}$ in (26), and $\boldsymbol{Z} = \boldsymbol{Z}^*$ in (31), and then adding these equations together, we have

$$\left\langle\overline{\boldsymbol{Z}}^{(k+1)}, \mu\left(\nabla\overline{\boldsymbol{Y}}^{(k)} - \overline{\boldsymbol{Z}}^{(k+1)} - \overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}\right)\right\rangle_F \ge 0,$$

from which we have

$$\mu\|\overline{\boldsymbol{Z}}^{(k+1)}\|_F^2 \le \langle\overline{\boldsymbol{Z}}^{(k+1)}, \mu\nabla\overline{\boldsymbol{Y}}^{(k)} - \overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}\rangle_F. \tag{32}$$

By adding (27), (28) and the results of (29), (30) and (32) by multiplying both sides of them with $2\beta\mu$, we obtain

$$\begin{aligned}& \mu\left(\|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k+1)}\|_F^2 - \|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)}\|_F^2\right) + \beta\left(\|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k+1)}\|_F^2 - \|\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}\|_F^2\right) \\ & + \beta^2\mu\left(\|\overline{\boldsymbol{Y}}^{(k+1)}\|^2 - \|\overline{\boldsymbol{Y}}^{(k)}\|_F^2\right) + \beta\mu^2\left(\|\nabla\overline{\boldsymbol{Y}}^{(k+1)}\|_F^2 - \|\nabla\overline{\boldsymbol{Y}}^{(k)}\|_F^2\right) \\ \le \quad & -\beta^2\mu\|\overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k)}\|_F^2 - \beta\mu^2\|\overline{\boldsymbol{Z}}^{(k+1)} - \nabla\overline{\boldsymbol{Y}}^{(k)}\|_F^2 - 2\beta\mu\|\boldsymbol{A}\overline{\boldsymbol{X}}^{(k+1)}\boldsymbol{A}_c^T\|_F^2.\end{aligned} \tag{33}$$

From (33), we have that, $\left\{\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)}\right\}$, $\left\{\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}\right\}$, $\left\{\overline{\boldsymbol{Y}}^{(k)}\right\}$ and $\left\{\nabla\overline{\boldsymbol{Y}}^{(k)}\right\}$ are convergent, and $\|\overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k)}\|_F$, $\|\overline{\boldsymbol{Z}}^{(k+1)} - \nabla\overline{\boldsymbol{Y}}^{(k)}\|_F$, and $\|\boldsymbol{A}\overline{\boldsymbol{X}}^{(k+1)}\boldsymbol{A}_c^T\|_F$ converge to zero.

Notice $\boldsymbol{X}^* = \boldsymbol{Y}^*$, $\boldsymbol{Z}^* = \nabla\boldsymbol{Y}^*$, from the right inequality of (22), we have that

$$\begin{aligned}& \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z}^*) \\ \le \quad & \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^{(k+1)}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z}^{(k+1)}) + \langle\boldsymbol{\Lambda}_{\boldsymbol{X}}^*, \boldsymbol{X}^{(k+1)} - \boldsymbol{Y}^{(k)}\rangle_F \\ & + \tfrac{\beta}{2}\|\boldsymbol{X}^{(k+1)} - \boldsymbol{Y}^{(k)}\|_F^2 + \langle\boldsymbol{\Lambda}_{\boldsymbol{Z}}^*, \boldsymbol{Z}^{(k+1)} - \nabla\boldsymbol{Y}^{(k)}\rangle_F + \tfrac{\mu}{2}\|\boldsymbol{Z}^{(k+1)} - \nabla\boldsymbol{Y}^{(k)}\|_F^2.\end{aligned} \tag{34}$$

Since

$$\boldsymbol{X}^{(k+1)} = \underset{\boldsymbol{X}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}, \boldsymbol{X} - \boldsymbol{Y}^{(k)} \rangle_F + \frac{\beta}{2}\|\boldsymbol{X} - \boldsymbol{Y}^{(k)}\|_F^2 \right\},$$

we have,

$$\frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 - \frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^{(k+1)}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \langle \beta\left(\boldsymbol{X}^{(k+1)} - \boldsymbol{Y}^{(k)}\right) + \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}, \boldsymbol{X} - \boldsymbol{X}^{(k+1)} \rangle_F \geq 0. \tag{35}$$

Since

$$\boldsymbol{Z}^{(k+1)} = \underset{\boldsymbol{Z}}{\operatorname{argmin}} \left\{ \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}) + \frac{\mu}{2}\|\boldsymbol{Z} - \nabla \boldsymbol{Y}^{(k)}\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \boldsymbol{Z} - \nabla \boldsymbol{Y}^{(k)} \rangle_F \right\},$$

we have,

$$\iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}) - \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^{(k+1)}) + \langle \mu\left(\boldsymbol{Z}^{(k+1)} - \nabla \boldsymbol{Y}^{(k)}\right) + \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \boldsymbol{Z} - \boldsymbol{Z}^{(k+1)} \rangle_F \geq 0. \tag{36}$$

Taking $\boldsymbol{X} = \boldsymbol{X}^*$ in (35) and $\boldsymbol{Z} = \boldsymbol{Z}^*$ in (36) we have

$$\begin{aligned} & \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^*) \\ \geq\ & \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^{(k+1)}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^{(k+1)}) \\ & + \langle \beta\left(\boldsymbol{X}^{(k+1)} - \boldsymbol{Y}^{(k)}\right) + \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}, \overline{\boldsymbol{X}}^{(k+1)} \rangle_F + \langle \mu\left(\boldsymbol{Z}^{(k+1)} - \nabla \boldsymbol{Y}^{(k)}\right) + \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \overline{\boldsymbol{Z}}^{(k+1)} \rangle_F. \end{aligned} \tag{37}$$

Since

$$\begin{aligned} \boldsymbol{Y}^{(k+1)} = \ & \underset{\boldsymbol{Y}}{\operatorname{argmin}} \left\{ \frac{\beta}{2}\|\boldsymbol{X}^{(k+1)} - \boldsymbol{Y}\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}, \boldsymbol{X}^{(k+1)} - \boldsymbol{Y} \rangle_F \right. \\ & \left. + \tfrac{\mu}{2}\|\boldsymbol{Z}^{(k+1)} - \nabla \boldsymbol{Y}\|_F^2 + \langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \boldsymbol{Z}^{(k+1)} - \nabla \boldsymbol{Y} \rangle_F \right\}, \end{aligned}$$

we have,

$$\beta\left(\boldsymbol{Y}^{(k+1)} - \boldsymbol{X}^{(k+1)}\right) - \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} + \mu \nabla^T\left(\nabla \boldsymbol{Y}^{(k+1)} - \boldsymbol{Z}^{(k+1)}\right) - \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)} = 0.$$

Therefore we have $\lim_{k\to\infty}(\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} + \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}) = 0$ because $\|\overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k)}\|_F$ and $\|\overline{\boldsymbol{Z}}^{(k+1)} - \nabla \overline{\boldsymbol{Y}}^{(k)}\|_F$ converge to zero, and $\left\{\overline{\boldsymbol{Y}}^{(k)}\right\}$ is convergent. Since $\overline{\boldsymbol{Y}}^{(k)}$, $\boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}$ and $\boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}$ are convergent, we therefore have $\lim_{k\to\infty}\left\langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)} + \nabla^T \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \overline{\boldsymbol{Y}}^{(k)} \right\rangle_F = 0$, or

$$\lim_{k\to\infty}\left(\left\langle \boldsymbol{\Lambda}_{\boldsymbol{X}}^{(k)}, \overline{\boldsymbol{Y}}^{(k)} \right\rangle_F + \left\langle \boldsymbol{\Lambda}_{\boldsymbol{Z}}^{(k)}, \nabla \overline{\boldsymbol{Y}}^{(k)} \right\rangle_F\right) = 0. \tag{38}$$

From the results that $\left\{\overline{\boldsymbol{\Lambda}}_{\boldsymbol{X}}^{(k)}\right\}$, $\left\{\overline{\boldsymbol{\Lambda}}_{\boldsymbol{Z}}^{(k)}\right\}$, $\left\{\overline{\boldsymbol{Y}}^{(k)}\right\}$ and $\left\{\nabla \overline{\boldsymbol{Y}}^{(k)}\right\}$ are convergent, and $\|\overline{\boldsymbol{X}}^{(k+1)} - \overline{\boldsymbol{Y}}^{(k)}\|_F$, $\|\overline{\boldsymbol{Z}}^{(k+1)} - \nabla \overline{\boldsymbol{Y}}^{(k)}\|_F$ converge to zero, and by using (38) we have from (34) and (37) that

$$\begin{aligned} & \lim_{k\to\infty}\left\{\frac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^{(k)}\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^{(k)})\right\} \\ & = \left\{\tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{Z}\|_1 \leq \delta)}(\boldsymbol{Z}^*)\right\} \\ & = \left\{\tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{X}^*\boldsymbol{A}_c^T - \boldsymbol{B}\|_F^2 + \iota_{(\|\boldsymbol{X}\|_1 \leq \delta)}(\boldsymbol{X}^*)\right\}. \end{aligned}$$

If $\boldsymbol{A}$ and $\boldsymbol{A}_c$ are nonsingular, then we have $\lim_{k\to\infty}\boldsymbol{X}^{(k)} = \boldsymbol{X}^*$, $\lim_{k\to\infty}\boldsymbol{Y}^{(k)} = \boldsymbol{Y}^* = \boldsymbol{X}^*$, and $\lim_{k\to\infty}\boldsymbol{Z}^{(k)} = \nabla \boldsymbol{Z}^* = \nabla \boldsymbol{X}^*$. □

5. Experimental results

In this section, we conduct two experiments to illustrate the effectiveness of our proposed method. In the first experiment, we study how the value of δ affects the quality of the image restoration. In the second experiment, we compare the restoration performance by using our proposed method with the well-known FTVd method [31]. FTVd is a very fast method; in each iteration, it uses six FFTs and one thresholding operation. For the convenience of discussion, we will use the notation ADMC for our proposed method.

We tested six typical images "House", "Lake", "Mandril", "Peppers', "Rose" and "Sunset" to investigate the performance of ADMC and FTVd. The six images are shown in Fig. 2.

We use the signal-to-noise ratio (SNR) [31]:

$$\mathrm{SNR} = 10\log_{10}\frac{\|\boldsymbol{X}_{\text{true}} - \mathrm{E}(\boldsymbol{X}_{\text{true}})\|_F^2}{\|\boldsymbol{X}_{\text{true}} - \boldsymbol{X}\|_F^2},$$

Fig. 2. Test images. From top row to bottom row: House, Lake, Mandril, Peppers, Rose and Sunset.

to measure the quality of the image restoration, where $\boldsymbol{X}_{\text{true}}$ and $\boldsymbol{X}$ denote the original image and the restored image respectively, and $\mathrm{E}(\boldsymbol{X}_{\text{true}})$ is the mean intensity value of $\boldsymbol{X}_{\text{true}}$.

All computations were performed in double precision using MATLAB 7.12 (R2011a) on an Intel(R) Core(TM) i7-2600 CPU @3.40 GHz, 4.00 GB RAM.

Experiment 1.

In this experiment, we investigate how the value of δ affects the quality of the image restoration. For this purpose, we first use (6) to generate a blurred and noisy image for two test images house and sunset, where we take $\boldsymbol{A} = G(7, 5)$, which is a Gaussian blur with square support size 7 and standard deviation 5, and $\boldsymbol{A}_c = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.25 & 0.5 & 0.25 \\ 0.15 & 0.1 & 0.75 \end{bmatrix}$. In each case we generate a noise free observation image $\boldsymbol{B}_{\text{nf}} = \boldsymbol{A}\boldsymbol{X}_{\text{true}}\boldsymbol{A}_c^T$. The blurred and noisy image $\boldsymbol{B}$ was generated by $\boldsymbol{B} = \boldsymbol{B}_{\text{nf}} + \boldsymbol{E}$, where $\boldsymbol{E}$ is a noise matrix whose entries are chosen from a normal distribution with mean 0 and variance 1, and scaled so that

$$\frac{\|\boldsymbol{E}\|_F}{\|\boldsymbol{B}_{\text{nf}}\|_F} = \eta, \tag{39}$$

and we use $\eta = 10^{-3}$ and $\eta = 10^{-2}$. Then we set δ to be $1.5\|\boldsymbol{B}\|_{MTV}$, $2\|\boldsymbol{B}\|_{MTV}$ and $2.5\|\boldsymbol{B}\|_{MTV}$. The results are shown in Figs. 3 to 6.

From the above figures, we see that the closer δ is to the TV of the true image, the better the obtained result is. We also see that, even if δ is not a good approximation to the TV of the true image, the obtained results are still acceptable.

Experiment 2.

In this experiment, we compare the results obtained by using our proposed method with that obtained by using FTVd. In order to provide a fair comparison, the experiments were done in the following way. For each test image, we use FTVd to restore the image, where the parameters in the algorithm are selected by trial and error to obtain best possible results. We then set δ to be the TV of above restored image, and finally we use our proposed method to restore the image. Tables 1 and 2 show the obtained results. From the tables, we see that our proposed method gets the highest SNR value for almost every test image.

Please cite this article in press as: J. Zhang, J.G. Nagy, An effective alternating direction method of multipliers for color image restoration, Appl. Numer. Math. (2020), https://doi.org/10.1016/j.apnum.2020.07.008

Fig. 3. From left to right: blurred and noise image (SNR = 8.98, $\eta = 0.001$), restored image ($\delta = 1.5\|\boldsymbol{B}\|_{MTV}$, SNR = 23.30), restored image ($\delta = 2\|\boldsymbol{B}\|_{MTV}$, SNR = 23.42), restored image ($\delta = 2.5\|\boldsymbol{B}\|_{MTV}$, SNR = 23.57).

Fig. 4. From left to right: blurred and noise image (SNR = 14.18, $\eta = 0.001$), restored image ($\delta = 1.5\|\boldsymbol{B}\|_{MTV}$, SNR = 25.12), restored image ($\delta = 2\|\boldsymbol{B}\|_{MTV}$, SNR = 25.45), restored image ($\delta = 2.5\|\boldsymbol{B}\|_{MTV}$, SNR = 25.50).

Fig. 5. From left to right: blurred and noise image (SNR = 8.95, $\eta = 0.01$), restored image ($\delta = 1.5\|\boldsymbol{B}\|_{MTV}$, SNR = 17.76), restored image ($\delta = 2.0\|\boldsymbol{B}\|_{MTV}$, SNR = 16.62), restored image ($\delta = 2.5\|\boldsymbol{B}\|_{MTV}$, SNR = 15.35).

Fig. 6. From left to right: blurred and noise image (SNR = 14.14, $\eta = 0.01$), restored image ($\delta = 1.5\|\boldsymbol{B}\|_{MTV}$, SNR = 19.91), restored image ($\delta = 2.0\|\boldsymbol{B}\|_{MTV}$, SNR = 19.31), restored image ($\delta = 2.5\|\boldsymbol{B}\|_{MTV}$, SNR = 18.43).

Table 1
SNR (db) for experiment with $\eta = 0.001$.

Algorithm	House	Lake	Mandril	Peppers	Rose	Sunset
ADMC	23.57	21.09	14.94	23.23	23.21	25.38
FTVd	23.51	21.07	14.95	23.20	23.17	25.33

Please cite this article in press as: J. Zhang, J.G. Nagy, An effective alternating direction method of multipliers for color image restoration, Appl. Numer. Math. (2020), https://doi.org/10.1016/j.apnum.2020.07.008

Table 2
SNR (db) for experiment with $\eta = 0.01$.

Algorithm	House	Lake	Mandril	Peppers	Rose	Sunset
ADMC	18.05	16.91	10.33	19.72	18.19	19.92
FTVd	17.99	16.88	10.30	19.73	18.20	19.84

6. Conclusions

We propose an effective alternating direction method of multipliers for color image restoration with TV regularization by formulating it as a constrained minimization problem and exploiting the structure to develop a very efficient algorithm. We prove convergence of the method in detail. Experimental results demonstrate that the proposed method is feasible and effective for color image restoration.

7. Appendix

In this section, we consider the minimization problem

$$\min_{\boldsymbol{Z}\in\mathcal{R}^{mn\times 6}} \left\{ \iota_{(\|\boldsymbol{Z}\|_1\le\delta)}(\boldsymbol{Z}) + \frac{\mu}{2}\|\boldsymbol{Z}-\boldsymbol{W}\|_F^2 \right\}. \tag{40}$$

This problem can be reformulated as

$$\min_{\boldsymbol{Z}\in\mathcal{R}^{mn\times 6}} \frac{1}{2}\|\boldsymbol{Z}-\boldsymbol{W}\|_F^2 \ \ s.t. \ \ \|\boldsymbol{Z}\|_1 \le \delta. \tag{41}$$

It is easy to see that if $\|\boldsymbol{W}\|_1 \le \delta$, then the solution is $\boldsymbol{Z}^* = \boldsymbol{W}$. Therefore, to solve the above problem, we may assume that $\|\boldsymbol{W}\|_1 > \delta$. In this case, since $\{\boldsymbol{Z} \,|\, \|\boldsymbol{Z}\|_1 \le \delta\}$ is a convex set, the optimal solution must be on the boundary of the constrained set and we therefore can replace the inequality constraint $\|\boldsymbol{Z}\|_1 \le \delta$ with equality constraint $\|\boldsymbol{Z}\|_1 = \delta$.

We now consider the following problem

$$\min_{\boldsymbol{Z}\in\mathcal{R}^{mn\times 6}} \frac{1}{2}\|\boldsymbol{Z}-\boldsymbol{W}\|_F^2 \ \ s.t. \ \ \|\boldsymbol{Z}\|_1 = \delta. \tag{42}$$

The Lagrangian function of the above problem is

$$L(\boldsymbol{X},\lambda) = \frac{1}{2}\|\boldsymbol{Z}-\boldsymbol{W}\|_F^2 + \lambda(\|\boldsymbol{Z}\|_1 - \delta), \tag{43}$$

where $\lambda \in \mathcal{R}$ is a Lagrangian multiplier. If $\boldsymbol{Z}$ is the minimizer of (42), then $0 \in \partial_{\boldsymbol{X}} L(\boldsymbol{Z},\lambda)$.

To characterize the optimal solution of the problem (42), we first give the following lemma.

Lemma 3. *[19] Let $x \in \mathcal{R}^l (l \ge 2)$, and $|x| = \sqrt{\sum_{i=1}^{l} x_i^2}$. Then the subdifferential of functional $f(x) = |x|$ is*

$$\partial f(x) = \begin{cases} \frac{x}{|x|}, & \text{if } x \neq 0, \\ \{h \in \mathcal{R}^l \,|\, |h| \le 1\}, & \text{otherwise}. \end{cases}$$

Denote $\boldsymbol{Z} = (\boldsymbol{Z}_1^T, \boldsymbol{Z}_2^T, \cdots, \boldsymbol{Z}_{mn}^T)^T$, where $\boldsymbol{Z}_i = (\boldsymbol{Z}_{i1}, \boldsymbol{Z}_{i2}, \cdots, \boldsymbol{Z}_{i6})$ is the ith row vector of $\boldsymbol{Z}$. Then using Lemma 3, we get for $i = 1, 2, \cdots, mn$,

$$\begin{cases} \boldsymbol{Z}_i - \boldsymbol{W}_i + \lambda \frac{\boldsymbol{Z}_i}{|\boldsymbol{Z}_i|} = 0, & \text{if } \boldsymbol{Z}_i \neq 0, \\ |\boldsymbol{W}_i| \le \lambda, & \text{otherwise}, \end{cases} \tag{44}$$

from which we see that $\boldsymbol{Z}_i$ is either parallel to $\boldsymbol{W}_i$ or $\boldsymbol{Z}_i = 0$ for $i = 1, 2, \cdots, mn$. On the other hand, since $\boldsymbol{Z}$ is the solution of (42), $\boldsymbol{Z}_i$ should have the same direction with $\boldsymbol{W}_i$ if $\boldsymbol{Z}_i \neq 0$.

Lemma 4. *Let $\boldsymbol{Z}$ be the optimal solution of the problem (42), then $\boldsymbol{Z}_i = \theta_i \boldsymbol{W}_i$, with $\theta_i \ge 0$.*

Proof. From (44) we have that $\boldsymbol{Z}_i = \theta_i \boldsymbol{W}_i$. The only thing we need to do is to show that $\theta_i \ge 0$ for all i. Suppose that there exists some i_0 such that $\theta_{i_0} < 0$. Then we define $\tilde{\boldsymbol{Z}} \in \mathcal{R}^{mn\times 6}$ with

$$\tilde{\boldsymbol{Z}}_i = \begin{cases} \boldsymbol{Z}_i, & \text{if } i \neq i_0, \\ -\boldsymbol{Z}_i, & \text{otherwise}. \end{cases}$$

Clearly $\|\tilde{\boldsymbol{Z}}\|_1 = \|\boldsymbol{Z}\|_1$, and is equal to δ. However, since $\theta_{i_0} < 0$,

$$\begin{aligned}
&\|\boldsymbol{Z} - \boldsymbol{W}\|_F^2 - \|\tilde{\boldsymbol{Z}} - \boldsymbol{W}\|_F^2 = |\boldsymbol{Z}_{i_0} - \boldsymbol{W}_{i_0}|^2 - |\tilde{\boldsymbol{Z}}_{i_0} - \boldsymbol{W}_{i_0}|^2 \\
&= |\theta_{i_0}\boldsymbol{W}_{i_0} - \boldsymbol{W}_{i_0}|^2 - |-\theta_{i_0}\boldsymbol{W}_{i_0} - \boldsymbol{W}_{i_0}|^2 \\
&= \left((\theta_{i_0} - 1)^2 - (-\theta_{i_0} - 1)^2\right)|\boldsymbol{W}_{i_0}|^2 > 0
\end{aligned}$$

which contradicts the assumption that $\boldsymbol{Z}$ is the optimal solution of problem (42). □

Based on Lemma 4, we have by using (44) that, if $\boldsymbol{Z}$ is the optimal solution of problem (42), then

$$\boldsymbol{Z}_i = \begin{cases} \max\left(\left(1 - \dfrac{\lambda}{|\boldsymbol{W}_i|}\right), 0\right)\boldsymbol{W}_i, & \text{if } \boldsymbol{W}_i \neq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{45}$$

Based on (45), we have the following lemma.

Lemma 5. *Let $\boldsymbol{Z}$ be the optimal solution of problem (42), and let i_1 and i_2 be two indices such that*

$$|\boldsymbol{W}_{i_2}| > |\boldsymbol{W}_{i_1}|.$$

If $|\boldsymbol{Z}_{i_2}| = 0$, then $|\boldsymbol{Z}_{i_1}| = 0$.

Proof. If $|\boldsymbol{Z}_{i_2}| = 0$, then from (45) we have $|\boldsymbol{W}_{i_2}| \le \lambda$. By the assumption we also have $|\boldsymbol{W}_{i_1}| \le \lambda$. Then using (45), we have $\boldsymbol{Z}_{i_1} = 0$, or equivalently $|\boldsymbol{Z}_{i_1}| = 0$. □

Using Lemma 5, and following [12], let $|\boldsymbol{Z}_{(i)}|$ be the order statics of the norms $|\boldsymbol{Z}_i|$. That is, $|\boldsymbol{Z}_{(1)}| \ge |\boldsymbol{Z}_{(2)}| \ge \cdots \ge |\boldsymbol{Z}_{(mn)}|$. Suppose that $|\boldsymbol{Z}_{(i)}| > 0$ for $i = 1, 2, \cdots, I$ and $|\boldsymbol{Z}_{(i)}| = 0$ for $i = I + 1, \cdots, mn$. Then we have

$$\begin{aligned}
&\|\boldsymbol{Z}\|_1 = \sum_{i=1}^{mn} |\boldsymbol{Z}_i| = \sum_{i=1}^{mn} |\boldsymbol{Z}_{(i)}| = \sum_{i=1}^{I} |\boldsymbol{Z}_{(i)}| \\
&= \sum_{i=1}^{I} |(\boldsymbol{W}_{(i)}| - \lambda) = \sum_{i=1}^{I} |\boldsymbol{W}_{(i)}| - \lambda I = \delta.
\end{aligned}$$

From which we have $\lambda = \frac{1}{I}\left(\sum_{i=1}^{I} |\boldsymbol{W}_{(i)}| - \delta\right)$. So if we find I, which is the largest number such that

$$1 - \frac{\frac{1}{I}\left(\sum_{i=1}^{I} |\boldsymbol{W}_{(i)}| - \delta\right)}{|\boldsymbol{W}_{(I)}|} > 0,$$

or

$$\sum_{i=1}^{I} |\boldsymbol{W}_{(i)}| - \delta < I|\boldsymbol{W}_{(I)}|,$$

we then obtain the optimal solution $\boldsymbol{Z}$ from (45).

References

[1] M.V. Afonso, J.M. Bioucas-Dias, M.A.T. Figueiredo, Fast image recovery using variable splitting and constrained optimization, IEEE Trans. Image Process. 19 (9) (2010) 2345–2356.

[2] M.V. Afonso, J.M. Bioucas-Dias, M.A.T. Figueiredo, An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems, IEEE Trans. Image Process. 20 (3) (2011) 681–695.

[3] Z.-Z. Bai, Y.-M. Huang, M.K. Ng, Block-triangular preconditioners for systems arising from edge-preserving image restoration, J. Comput. Math. 28 (6) (2010) 848–863.

[4] A.H. Bentbib, M. El Guide, K. Jbilou, E. Onunwor, L. Reichel, Solution methods for linear discrete ill-posed problems for color image restoration, BIT Numer. Math. 58 (3) (2018) 555–576.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2011) 1–122.

[6] K. Bredies, K. Kunisch, T. Pock, Total generalized variation, SIAM J. Imaging Sci. 3 (3) (2010) 492–526.

[7] A. Brook, R. Kimmel, N. Sochen, Variational restoration and edge detection for color images, J. Math. Imaging Vis. 18 (3) (2003) 247–268.

Please cite this article in press as: J. Zhang, J.G. Nagy, An effective alternating direction method of multipliers for color image restoration, Appl. Numer. Math. (2020), https://doi.org/10.1016/j.apnum.2020.07.008

[8] T.F. Chan, G.H. Golub, P. Mulet, A nonlinear primal-dual method for total variation-based image restoration, SIAM J. Sci. Comput. 20 (6) (1999) 1964–1977.
[9] T.F. Chan, A.L. Marquina, P. Mulet, High-order total variation-based image restoration, SIAM J. Sci. Comput. 22 (2) (2000) 503–516.
[10] T.F. Chan, S.H. Kang, J. Shen, Total variation denoising and enhancement color images based on the CB and HSV color models, J. Vis. Commun. Image Represent. 12 (4) (2001) 422–435.
[11] P.L. Combettes, J.C. Pesquet, Image restoration subject to a total variation constraint, IEEE Trans. Image Process. 13 (9) (2004) 1213–1222.
[12] J. Duchi, S. Shalev-Shwartz, Y. Singer, T. Chandra, Efficient projections onto the l_1-ball for learning in high dimensions, in: Proceedings of the 25th International Conference on Machine Learning, 2008.
[13] I. Ekeland, R. Temam, Convex Analysis and Variational Problems, Classics in Appl. Math., vol. 28, SIAM, Philadelphia, 1999.
[14] S. Gazzola, P. Novati, M.R. Russo, On Krylov projection methods and Tikhonov regularization, Electron. Trans. Numer. Anal. 44 (2015) 83–123.
[15] G. Gilboa, S. Osher, Nonlocal operators with applications to image processing, Multiscale Model. Simul. 7 (3) (2008) 1005–1028.
[16] P.C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, SIAM, Philadelphia, 1997.
[17] P.C. Hansen, J.G. Nagy, D.P. O'Leary, Deblurring Images: Matrices, Spectra, and Filtering, SIAM, Philadelphia, 2006.
[18] Y.-M. Huang, M.K. Ng, Y.-W. Wen, A fast total variation minimization method for image restoration, Multiscale Model. Simul. 7 (2) (2008) 774–795.
[19] J. Jahn, Introduction to the Theory of Nonlinear Optimization, third edition, Springer-Verlag, 2006.
[20] S. Lefkimmiatis, A. Roussos, P. Maragos, M. Unser, Structure tensor total variation, SIAM J. Imaging Sci. 8 (2) (2015) 1090–1122.
[21] J.G. Nagy, K. Palmer, L. Perrone, Iterative methods for image deblurring: a MATLAB object oriented approach, Numer. Algorithms 36 (1) (2004) 73–93.
[22] M.P. Nikolova, M.K. Ng, Analysis of half-quadratic minimization methods for signal and image recovery, SIAM J. Sci. Comput. 27 (3) (2005) 937–966.
[23] S. Ono, L_0 gradient projection, IEEE Trans. Image Process. 26 (4) (2017) 1554–1564.
[24] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.
[25] B. Tang, G. Sapiro, V. Caselles, Color image enhancement via chromaticity diffusion, IEEE Trans. Image Process. 10 (5) (2001) 701–707.
[26] A.N. Tikhonov, V.Y. Arsenin, Solutions of Ill-Posed Problems, Winston and Sons, Washington DC, 1977.
[27] Y.-L. Wang, J.-F. Yang, W.-T. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM J. Imaging Sci. 1 (3) (2008) 248–272.
[28] Y.-W. Wen, M.K. Ng, W.-K. Ching, Iterative algorithms based on decoupling of deblurring and denoising for image restoration, SIAM J. Sci. Comput. 30 (5) (2008) 2655–2674.
[29] Y.-W. Wen, M.K. Ng, Y.-M. Huang, Efficient total variation minimization methods for color image restoration, IEEE Trans. Image Process. 17 (11) (2008) 2081–2088.
[30] C.-L. Wu, X.-C. Tai, Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models, SIAM J. Imaging Sci. 3 (3) (2010) 300–339.
[31] J.-F. Yang, W.-T. Yin, Y. Zhang, Y.-L. Wang, A fast algorithm for edge-preserving variational multichannel image restoration, SIAM J. Imaging Sci. 2 (2) (2009) 569–592.
[32] X.-Q. Zhang, M. Burger, S. Osher, A unified primal-dual algorithm framework based on Bregman iteration, J. Sci. Comput. 46 (1) (2011) 20–46.

Please cite this article in press as: J. Zhang, J.G. Nagy, An effective alternating direction method of multipliers for color image restoration, Appl. Numer. Math. (2020), https://doi.org/10.1016/j.apnum.2020.07.008