Neurocomputing 170 (2015) 201–212

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Fast total-variation based image restoration based on derivative alternated direction optimization methods

Dongwei Ren [a], Hongzhi Zhang [a], David Zhang [b], Wangmeng Zuo [a,*]

[a] School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[b] Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 10 December 2013
Received in revised form
24 May 2014
Accepted 4 August 2014
Available online 15 June 2015

Keywords:
Total variation
Image restoration
Convex optimization
Alternating direction method of multipliers
Augmented Lagrangian method

ABSTRACT

The total variation (TV) model is one of the most successful methods for image restoration, as well as an ideal bed to develop optimization algorithms for solving sparse representation problems. Previous studies showed that derivative space formulation of the image restoration model is useful in improving the success rate in image recovery and kernel estimation performance in blind deconvolution. However, little attentions are paid on the model and algorithm for derivative space based image restoration. In this paper, we study the TV based image restoration (TVIR) by developing a novel derivative space-based reformulation together with an efficient derivative alternating direction method of multipliers (D-ADMM) algorithm. Thanks to the simplicity of the proposed derivative space reformulation, D-ADMM only requires four fast Fourier transform (FFT) operations per iteration, and is much more efficient than the other augmented Lagrangian methods. Numerical experiments show that, D-ADMM can obtain satisfactory restoration result and is much faster than the state-of-the-art TVIR algorithms.

1. Introduction

In many image processing applications, the image in hand is only a degraded observation $\mathbf{y}$ of the original image $\mathbf{x}$. In the linear degradation model, the procedure can be modeled as

$$\mathbf{y}=\mathbf{A}\mathbf{x}+\mathbf{e}, \tag{1}$$

where $\mathbf{A}$ is a linear operator and $\mathbf{e}$ is the additive Gaussian white noise (AGWN). Image restoration aims to estimate the clear image $\mathbf{x}$ from its degraded observation $\mathbf{y}$, and is well known as a typical linear inverse problem [1].

Since the linear operator $\mathbf{A}$ usually is ill-conditioned, the recovery of $\mathbf{x}$ from $\mathbf{y}$ is an ill-posed problem, and a typical image restoration models generally includes a fidelity term and a regularization term, resulting in the following optimization problem:

$$\min_{\mathbf{x}}\frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2+\tau R(\mathbf{x}), \tag{2}$$

where $R(\mathbf{x})$ is some regularizer on $\mathbf{x}$, and τ is the regularization parameter. By far, based on various models on image prior, a number of regularizers, e.g., total variation (TV) [2], gradient-based [3], wavelet-based [4], dictionary-based sparsity [5–7], and non-local models [8–10], have been developed for image restoration. Due to its simplicity and ability to preserve edges, the TV regularizer has been widely applied to various image restoration and recovery tasks, e.g., denoising [11,12], deconvolution [13–15], and compressed sensing (CS) [16,17].

The TV model is also an ideal bed to develop optimization algorithms for solving sparse representation problems. By far, a number of methods have been developed for TVIR. These algorithms, including split-Bregman [18], accelerated proximal gradient [19–21], and alternating direction method of multipliers [22–24], were applied to TVIR, and then were adopted for other image processing, computer vision, and machine learning tasks [3,25–29].

1.1. Related work

The TVIR model can be formulated as

$$\min_{\mathbf{x}}F(\mathbf{x})=\frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2+\tau\|\mathbf{D}\mathbf{x}\|, \tag{3}$$

where $\mathbf{D}=[\mathbf{D}_h^T,\mathbf{D}_v^T]^T$ is the discrete gradient operator, $\|\cdot\|$ denotes the norm in the gradient space (including both anisotropic and isotropic versions, and please refer to Section 2.1 for detailed definitions of TV regularizers), and τ is the regularization parameter.

The augmented Lagrangian methods (ALM) are one class of the most efficient among various TVIR algorithms. Because of the non-smoothness of the TV regularizer, variable splitting strategies usually are required in the ALM-based algorithms. By far, there are mainly two variable splitting strategies for ALM-based TVIR. In [23], an auxiliary variable $\mathbf{u}$ was introduced to substitute the

* Corresponding author. Tel.: +86 451 86412871.
E-mail address: cswmzuo@gmail.com (W. Zuo).

http://dx.doi.org/10.1016/j.neucom.2014.08.101
0925-2312/

variable $\mathbf{x}$ in the fidelity term, resulting in the following equality-constrained optimization problem,

$$\min_{\mathbf{x},\mathbf{u}} \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{y}\|_2^2+\tau\|\mathbf{D}\mathbf{x}\| \quad s.t.\ \mathbf{u}=\mathbf{x}. \tag{4}$$

By incorporating with the efficient TV-based denoising algorithm [12], Afonso et al. developed a split augmented Lagrangian shrinkage algorithm (SALSA) for TVIR. In [22,30], another variable splitting strategy was adopted in FTVd by introducing an auxiliary variable $\mathbf{d}$ to substitute $\mathbf{D}\mathbf{x}$ in the regularizer, resulting in the following equivalent formulation:

$$\min_{\mathbf{x},\mathbf{d}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2+\tau\|\mathbf{d}\| \quad s.t.\ \mathbf{d}=\mathbf{D}\mathbf{x}, \tag{5}$$

where $\mathbf{d}=[\mathbf{d}_h^T,\mathbf{d}_v^T]^T$ with $\mathbf{d}_h=\mathbf{D}_h\mathbf{x}$, $\mathbf{d}_v=\mathbf{D}_v\mathbf{x}$. With this formulation, each subproblem of the ALM algorithms can be efficiently solved, making FTVd the state-of-the-art TVIR methods in terms of computational efficiency.

Recently, derivative space formulation of the image restoration model had received considerable research interests and shown several unique advantages in compressed sensing [31,32], image restoration [33], and blind deconvolution [34]. In compressed sensing, Patel et al. [31] proposed a GradientRec approach which first used the compressed sensing (CS) algorithm to recover the gradient images and then reconstructed the original image from the gradient images. Because the gradient images are much more sparse, it had been shown in [31] that derivative space based GradientRec could obtain higher success rate in image recovery.

In image restoration, Michailovich [33] also introduced a variable $\mathbf{d}$ to substitute $\mathbf{D}\mathbf{x}$ in the regularizer. By assuming that the image $\mathbf{x}$ has zero mean value, a left inverse operator $\mathcal{U}(\mathbf{d})$ [35] can be employed to recover the original image from the derivative space, i.e., $\mathbf{x}=\mathcal{U}(\mathbf{d})$. Thus, the variable $\mathbf{x}$ can be removed from the model in Eq. (3), and TVIR can be formulated in the derivative space,

$$\min_{\mathbf{d}} \frac{1}{2}\|\mathbf{A}\mathcal{U}\{\mathbf{d}\}-\mathbf{y}\|^2+\tau\|\mathbf{d}\|. \tag{6}$$

Michailovich [33] proposed a TV-based iterative shrinkage (TVIS) algorithm for solving the model in Eq. (6).

In blind deconvolution, recent studies showed that better kernel estimation performance can generally be obtained in the derivative space than in the image space [34,36–38]. Cho and Lee [39] analyzed the condition numbers of the Hessians which indicated that the Hessian in the derivative space has a diagonally dominant structure and has a much smaller condition numbers than that in the image space.

Although previous studies had indicated the advantages of derivative space formulation, little attentions are paid on the proper modeling and efficient algorithms for derivative space based image restoration. For example, GradientRec only greedily solved the TV based CS problem and cannot guarantee the convergence to the solution of the original problem. The convergence rate of TVIS is $O(t^{-1})$, which is much slower than the state-of-the-art TVIR algorithms.

In this paper, we study the derivative space TVIR problem by proposing a novel derivative space - based reformulation together with an efficient derivative alternating direction method of multipliers (D-ADMM) algorithm. This work is an extension of [40], based on which we deduce an explicit formulation of TVIR in the derivative space and propose two ADMM-based algorithms to solve it efficiently. First, by analyzing the connections of image space and derivative space, we introduce an explicit equality constraint on the gradients $\mathbf{d}$, and suggest a novel derivative space based reformulation of TVIR. Compared with the formulation in [33], the proposed formulation is more concise and much easier to be solved. Then, we adopt the alternating direction method of multipliers (ADMM) algorithm to solve the constrained optimization problem, resulting in the proposed derivative-space ADMM (D-ADMM) algorithm. D-ADMM only requires four fast Fourier transform (FFT) operations per iteration, and is much more efficient than the other TVIR methods. Finally, experimental results show that D-ADMM can obtain satisfactory restoration results and is much faster than the state-of-the-art TVIR algorithms, e.g., FTVd and SALSA.

1.2. Organization

This paper is organized as follows. Section 2 introduces some background knowledge related to this paper. Section 3 presents the derivative space based reformulation of TVIR, and Section 4 describes the proposed D-ADMM algorithms. Section 5 provides the experimental results by comparing D-ADMM with the state-of-the-art methods. Finally, Section 5 ends this paper with some concluding remarks.

2. Preliminaries

In this section, we first introduce the discrete TV operators with periodic boundary conditions, then summarize the related proximal operators used in this paper, and finally, briefly review the ADMM algorithm.

2.1. The discrete TV operators

Analogous to [33], we assume that the image $\mathbf{x}$ should lie in the $\mathbb{R}^{m\times n}$ space $\mathbb{U}$ with zero mean value, i.e., $\mathbb{U}=\{\mathbf{x}\in\mathbb{R}^{m\times n}|\, mean(\mathbf{x})=0\}$. With the assumption of periodic boundary conditions, the gradient operator $\mathcal{D}$, also notated as ∇, is defined as

$$\begin{aligned}(\mathcal{D}_h\mathbf{x})_{k,l}&=\mathbf{x}_{k,l}-\mathbf{x}_{k,l-1} \quad \text{with } \mathbf{x}_{k,-1}=\mathbf{x}_{k,n-1}\\ (\mathcal{D}_v\mathbf{x})_{k,l}&=\mathbf{x}_{k,l}-\mathbf{x}_{k-1,l} \quad \text{with } \mathbf{x}_{-1,l}=\mathbf{x}_{m-1,l}\end{aligned} \tag{7}$$

where $k=0,1,2,\ldots,m-1$ and $l=0,1,2,\ldots,n-1$. Thus, the anisotropic TV [19,21] is defined as

$$\mathrm{TV}_a(\mathbf{x})=\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}\left(|(\mathcal{D}_h\mathbf{x})_{k,l}|+|(\mathcal{D}_v\mathbf{x})_{k,l}|\right). \tag{8}$$

The isotropic TV [19,21] is defined by

$$\mathrm{TV}_i(\mathbf{x})=\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}\sqrt{(\mathcal{D}_h\mathbf{x})_{k,l}^2+(\mathcal{D}_v\mathbf{x})_{k,l}^2}. \tag{9}$$

The adjoint operators $\mathcal{D}_h^*$ and $\mathcal{D}_v^*$ of $\mathcal{D}_h$ and $\mathcal{D}_v$ can be defined by

$$\begin{aligned}(\mathcal{D}_h^*\mathbf{x})_{k,l}&=\mathbf{x}_{k,l}-\mathbf{x}_{k,l+1} \quad \text{with } \mathbf{x}_{k,n}=\mathbf{x}_{k,0}\\ (\mathcal{D}_v^*\mathbf{x})_{k,l}&=\mathbf{x}_{k,l}-\mathbf{x}_{k+1,l} \quad \text{with } \mathbf{x}_{m,l}=\mathbf{x}_{0,l},\end{aligned} \tag{10}$$

respectively.

The images $\mathbf{x}$, $\mathbf{d}_h$ and $\mathbf{d}_v$ can be rearranged into the corresponding vectors, and vice versa. Thus, we use the same small bold notation to denote an image and its vectorization, and this should not cause ambiguity by referring to the context. Then the gradient operators $\mathcal{D}_h$ and $\mathcal{D}_v$ can be written as matrices $\mathbf{D}_h$ and $\mathbf{D}_v$ with $\mathbf{d}_h=\mathbf{D}_h\mathbf{x}$ and $\mathbf{d}_v=\mathbf{D}_v\mathbf{x}$, respectively, The corresponding adjoint operators $\mathcal{D}_h^*$ and $\mathcal{D}_v^*$ are associated with matrices $\mathbf{D}_h^T$ and $\mathbf{D}_v^T$, respectively.

2.2. Related proximal operators

Given a (nonsmooth) convex function $g(\mathbf{x})$ and a vector $\mathbf{z}$, the proximal operator with parameter λ of g is the function $\mathbf{prox}_{\lambda g}$