Neurocomputing 122 (2013) 245–257

Contents lists available at ScienceDirect

Neurocomputing

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

A sparse proximal Newton splitting method for constrained image deblurring

Han Pan*, Zhongliang Jing, Ming Lei, Rongli Liu, Bo Jin, Canlong Zhang

Aero, Information and control, Dongchuanglu, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 6 December 2012
Received in revised form
11 April 2013
Accepted 13 June 2013
Communicated by X. Gao
Available online 4 July 2013

Keywords:
Image deblurring
Optimization method
Total variation minimization
Proximal Newton splitting

ABSTRACT

It is a challenging task to recover a high quality image from the degraded images. This paper proposes a fast image deblurring algorithm. To deal with the limitations of the proximal Newton splitting scheme, a sparse framework is presented, which characterized by utilizing the sparse pattern of the approximated inverse Hessian matrix and relaxing the original assumption on the constant penalty parameter. The proposed framework provides a common update strategy by exploiting the second derivative information. To alleviate the difficulties introduced by the sub-problem of this framework, an approximate solution to the weighted norm based primal–dual problem is derived and studied. Moreover, its theoretical aspects are also investigated. Compared with the state-of-the-art methods in several numerical experiments, the proposed algorithm demonstrates the performance improvement and efficiency.

1. Introduction

Image restoration has been widely studied and discussed, first explored in the 1960s [1,2]. This problem is a classical inverse problem, which existed in various application domains, such as astronomical image processing, medical image reconstruction and microarray processing. In the past years, substantial amount of material concerned on this subject. The main task of image restoration is to recover the latent image under different distortion conditions and noisy scenario. These problems varied in the prior models, the regularization terms and the optimization methods. For simplify, a classic imaging model is widely adopted, which consists of three components, i.e. an observation image z, the latent image x and the additive noise η. This model can be formulated as follows:

$$z = K \otimes x + \eta, \tag{1}$$

where K denotes the convolution operator or point spread function (PSF) on the latent image x. More specifically, the operator is mainly assumed to be spatially invariant, which can be viewed as an approximate to the blurring process. And the noise, denoted by η, varies in different conditions include Gaussian, salt & pepper noise, and so on. Motivated by the specific properties of the imaging process of the particular hardware [3,4], some researchers proposed some efficient methods to restore a clear scene from a single image or multi-frame images. Unlike the conventional methods, some other schemes, such as the partial differential equation (PDE) with extreme learning machine [36] and the neural network (NN)-like methods [37,38], are presented to reconstruct the underlying image effectively.

In order to restore more edges or other obvious features of the latent image and reduce the presented noise, Rudin-Osher and Fatemi (ROF) [5] presented the total variation regularization term, which can be exhibited as follows:

$$F(x) = f(x) + g(x) = \min_x \frac{1}{2} \|K \otimes x - z\|_F^2 + \lambda TV(x), \tag{2}$$

where $\|\cdot\|_F$ is the *Frobenius*-norm, z denotes the degrade image; $f(x)$ stands for a convex function $\frac{1}{2}\|K \otimes x - z\|^2$; $g(x)$ stands for the regularization term or non-convex functions $\lambda TV(x)$; the parameter λ is a scalar positive parameter, which controls the regularization degree between the $f(x)$ and $g(x)$. This parameter can be automatically selected by some methods or rules such as generalized cross-validation (GCV), the discrepancy principle (DP), the L-curve [6] and the generalized Stein unbiased risk estimator (GSURE) method [36]. It can be noted that $F(x)$ is a non-smooth problem. The ROF's regularization term has two forms, i.e. the isotropic TV and the l_1-based isotropic TV (see details in Appendix A). It is important to note that the TV regularizer is non-smooth, which brings in some difficulties, such as non-differentiability of the TV term, huge storage requirement and heavy computation demand. To remedy these issues, various schemes were proposed to deal with different aspects of these issues. On the one hand, in order to overcome the drawbacks of the original ROF's TV model, some researchers proposed some improved total variation models include non-local and graph based

* Corresponding author. Tel.: +86 21 34206272.
E-mail addresses: hanpan@sjtu.edu.cn, panhamming@gmail.com (H. Pan), zljing@sjtu.edu.cn (Z. Jing), mlei@sjtu.edu.cn (M. Lei).

0925-2312/$ - see front matter Crown Copyright © 2013 Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.06.027

total variation models [7]. On the other hand, some optimization algorithms [6,8–10,29–31] were proposed to obtain a favorable solution, such as projected gradient method [6], proximal gradient method [8], a class of the Iterative Shrinkage/Thresholding Algorithms (ISTA) [9,29] and the SpaRSA (Sparse reconstruction by separable approximation) [10]. Among these methods, the families of the IST-like algorithms were examined to be simply and efficient minimization methods, which vary in the different strategies on the step size and the iterate variable.

1.1. Related works

Recently, an implicit optimization approach [8,9,11,12], named the proximal splitting or operator splitting scheme, is proved to be an efficient scheme, which can be categorized as first-order method. Its key procedure is to solve the sub-problem, i.e. the Moreau proximal minimization problem [11]. There are intensive research literatures on the proximal splitting scheme [14]. Among these works, the Forward–Backward splitting method [8] and its variants are widely applied include Backward–Backward algorithm [15], Beck–Teboulle proximal optimization method [8], Dykstra-like splitting methods [16] and Douglas–Rachford splitting method [17]. Some algorithms, extended from the proximal splitting scheme, have been developed to handle the real-world's problems directly, such as image restoration [8,9], compressive magnetic resonance (MR) image reconstruction [12], matrix completion [18] and machine learning [19,20]. Among the above mentioned schemes, a class of accelerated proximal splitting algorithms was developed [9]. In spite of these works, it can be noted that there is significantly less work to focus on the regularization parameter of the proximal term.

Another efficient way to handle this problem is the variable splitting technology [21,22], which is mainly coupled with the augmented Lagrangian optimization framework by converting the original problem into an equivalent constraint optimization problem. And this scheme has been applied to various application domains. For image restoration problems, an effective scheme [13,21,22,32,33] based on the classic alternating direction method of multipliers (ADMM) framework was proposed. A combined variable splitting and duality based image restoration method for the salt & pepper noise removal was proposed in [23]. For the multiple variables splitting case, Goldfarb [24] provided a novel solution. For the TV/l_1 and TV/l_2 minimization problems and its variants, Min Tao [32] exploited the original structure of the problems and proposed some effective algorithms. And this scheme also applied in some application domains, such as compressed sensing (CS) [25] and video restoration [33]. Although this scheme achieves a success in the real-world's application, its convergence rate sometimes depends on the choice or update of the penalty parameter.

1.2. The motivation and contribution

The motivation of this paper is to reduce the computational and dimension limitations of the frameworks in [26,27]. It is well known that the quasi-Newton-like methods suffer from the computational requirement and memory usage for obtaining the accurate representation of the true Hessian matrix as the dimension of the problem growing. In other words, when this framework applies to the large-scale problem, some issues come out, for instance the computation expense and the side effect of the approximation of the exact Hessian matrix. These shortcomings are a cue to develop an alternative approach in the hope that an approximation to this optimization scheme, which is computationally cheap without destroying the convergence properties of the original scheme. To handle these issues existed in the computability and implementation, a modification of the previous framework in [26,27] is proposed. As a result, an alternate approach for solving the resulting sub-problem (18) is presented.

The difference to the works in [8,9,21,22], presented for solving the sub-problem (18), can be summarized in three aspects. First, a substantive assumption on the penalty parameter σ in Eq. (7) is relaxed into a general representation, i.e. an iterate-variation and bound condition. The previous assumption is that the penalty parameter is chosen to be *Lipschitz* constant [8,9] or another constant μ in [21,22], which need to be estimated at the algorithm's beginning. In this paper, the proximal term is generalized by changing from $\frac{\sigma}{2}\|x-y\|_2^2$ into $\frac{1}{2}\|x-y\|_{\tilde{W}}^2$. The penalty parameter σ has an influence on the convergence process toward the optimal solution [40]. The proposed method can utilize the convergence behavior of the proximal splitting scheme [21,22]. Although the relaxation of the main assumption in [21,22] seems quite simple and natural, it will increase the practical numerical computation performance of the proposed method in term of the evaluation of the reconstruction performance significantly. Second, a common strategy for obtaining the suitable regularization matrix $\tilde{W}$ is given. Specially, a modified iterate-update penalty parameter method is provided. Compared to the works in [26,27], the proposed approach just partly exploits the second-derivative information, can be viewed as a balance between the computational performance and the memory requirement. Finally, a generalized framework, extended from the original proximal splitting schemes, is provided. In summary, although the pattern of the proposed algorithm and the methods in [21,22,26,27] is broadly similar, the main difference is the consideration of sparse pattern and the iterate-depend penalty or weighting matrix, which can deal with the overhead costs introduced by the size limitation in [26,27].

In this paper, a fast image deblurring algorithm, utilizing the sparse pattern of proximal Newton splitting framework, is proposed, which can be considered as an extension of the works in [26,27]. For resolving the weighted norms based Moreau proximal minimization problem, a sparse gradient projection method is presented to resolve the resulting sub-problem. Meanwhile, a modified symmetric rank 1 (SR1) approximation method is provided to determine the weighted penalty matrix adaptively. The main advantage of the proposed method is the introduction of the iteration-varying penalty or weighting matrix, which can control the progress of the estimated variable toward the optimal solution set dynamically. Specially, the presented method maintains the sparse structure of the weighting matrix and provides a unifying framework to the primary proximal splitting scheme. In other words, our method can generate appropriate penalty parameter quickly. In some degree, the proposed method can deal with a retrograde step toward feasibility. Another advantage of the proposed method also provides a mechanism for coping with the singularities when approximating the inverse Hessian matrix. At last, the theoretical and practical results of the proposed algorithm are presented. The extensive experiments against the state-of-art methods indicate that the presented method is more suitable for large scale optimization problems.

The paper is organized as follows. In Section 2, a briefly survey on related optimization for algorithm comparison is presented. In Section 3, the proposed method and its some theoretical results are exhibited. In Section 4, the numerical simulations and an extensive discussion are presented. Finally, the work is concluded in Section 5.

2. The related optimization algorithms

2.1. The basic notations

In this section, some notations in this paper are defined.

- The symbol $\|x\|_F$ denotes the *Frobenius* norm on the matrix space $x \in R^{m \times n}$, where m, n denote the size of matrix x.