

基于非凸的全变分和低秩混合正则化的图像去模糊模型和算法

孙 涛 李东升

(国防科技大学计算机学院 长沙 410073)

摘 要 非盲图像去模糊问题是从已知核的带噪声的线性卷积变换中恢复原始图像. 如果噪声是满足高斯分布的, 则可以直接使用最小二乘求解. 然而在大多数情况下, 去模糊问题都是高度病态的, 直接求解无法做到. 因此, 通常的做法是通过抽取原始图像的已知统计先验信息进行正则化来帮助求解问题. 两种常用的正则化是低秩和全变分. 早期的相关工作单独使用这两种正则化. 直到几年前, 人们才考虑将这两种正则化结合起来. 已有的结果表明, 混合正则化模型比单一模型具有更好的性能. 然而, 目前的混合正则化方法只是采用凸方法, 非凸的工作仍然是空白的. 考虑到非凸正则化在很多种情况下都比凸正则化的效果要好, 因此本文使用 $L_{1/2}$ 范数和 Schatten- $1/2$ 范数提出了一种新的非凸混合模型. 我们使用这两个非凸函数, 因为它们近端算子很容易计算. 这种非凸混合正则化模型本质上是一个非凸线性约束问题, 可以通过交替方向乘子法求解. 然而, 非凸性使得交替方向乘子法收敛十分困难. 因此, 我们转向求解原问题的惩罚问题. 将交替最小化方法应用于惩罚问题就可以得到提出的算法, 其中每个子步骤只涉及非常简单的计算. 由于惩罚参数很大时, 交替极小化算法速度会很慢, 为了加速算法, 针对惩罚参数我们使用了预热技术, 即选取很小的初值但是在迭代过程中不断将参数增大. 我们证明了该算法的收敛性. 数值实验验证了本文提出的模型和算法的有效性. 在非温和的假设下, 我们证明了算法的收敛性. 数值实验验证了本文提出的模型和算法的有效性.

关键词 低秩; 全变分; 图像去模糊; 非凸模型; 交替极小化

中图法分类号 TP311 **DOI号** 10.11897/SP.J.1016.2020.00643

Nonconvex Low-Rank and Total-Variation Regularized Model and Algorithm for Image Deblurring

SUN Tao LI Dong-Sheng

(College of Computer, National University of Defense Technology, Changsha 410073)

Abstract Non-blind image deblurring aims to reconstruct the original image for noised linear convolution transform with some known kernel. If the noise is Gaussian, one might use the least square minimization with the observed image and the kernel for this task. However, in most cases, the convolution operator makes the problem very ill-posed and hard to solve directly. To this end, regularizations, which are recruited to characterize known statistical priors about the original images, are then developed to help. Among them, two frequently used ones are the low-rank and total-variation regularizations. The earlier related works employed them separately, that is, just using one of them, not both. Until several years ago, people have considered combing these two regularizations together. Existing results show that the hybrid regularized model performs much better than the single one. However, the current composite regularization just uses the convex methodology. The

收稿日期: 2019-07-01; 在线出版日期: 2020-02-08. 本课题得到国家科技部重点研发项目(2018YFB0204300)和国家自然科学基金青年项目(61906200)资助. 孙涛(通信作者), 博士, 助理研究员, 主要研究领域为机器学习中的优化, 图像处理优化算法. E-mail: nudtsuntao@163.com. 李东升, 博士, 教授, 计算机学会(CCF)会员, 主要研究领域为分布式计算, 计算机网络, 云计算, 大数据.

nonconvex implementations are still missing. Considering nonconvex regularizations can beat convex ones in various cases, in this paper, we propose a novel nonconvex hybrid model in which, the $L_{1/2}$ and Schatten-1/2 norms are used. We use these two nonconvex functions due to that their proximal maps are easy to calculate even without convexity. Both $L_{1/2}$ and Schatten-1/2 norms enjoy closed-form proximal maps. The proposed nonconvex hybrid regularized model is naturally a nonconvex linearly constrained problem which can be solved by the alternating directions of multipliers. However, the nonconvexity breaks the theoretical guarantees. Thus, we turn to solve the penalty problem rather than the original form. The alternating minimization method applied to the penalty then yields the proposed algorithm, in which, each substep just involves very simple computations. In our algorithm, an important parameter is the penalty parameter. If it is infinity, the penalty is then identical to the original problem. But the large penalty parameter will make the algorithm iterate slowly. Thus, to improve the speed and narrow the penalty problem and the original one, for the penalty parameter, we use a warm-up technique, that is, increasing the penalty parameter in the iterations. The convergence of the algorithm is proved under very mild assumptions, which can be easily satisfied in applications. The numerical experiments are conducted on six natural test images. The performance of the proposed algorithm verifies the convergence theory. Comparisons with other algorithms demonstrate the efficiency of our algorithm.

Keywords low-rank; total variation; image deblurring; nonconvex model; alternating minimization

1 引 言

非盲的图像去模糊是图像处理领域中,尤其是自然图像处理中一类常见的问题.从数学角度该问题可以表示为:给定一个图像 $f^* \in \mathbb{R}^{m \times n}$ 和已知的模糊算子 B ,在观测到模糊且噪声污染的图像 $e = B(f^*) + b$ 后,通过数值算法有效地重构原始图像 f^* . 上式中 b 表示观测过程中的噪声,一般问题都假设噪声是满足高斯分布的.在本文中,我们也遵循这样的假设.该问题本质上是一个逆问题,最为直观的解决办法是最小二乘回归,也就是直接求解 $\min_f \|B(f) - e\|^2$.当然如果噪声不是满足高斯分布的,需要其他的回归方法,例如如果噪声是脉冲的,我们需要用鲁棒回归方法^[19-20],即极小化 $\|B(f) - e\|_1$.但是由于模糊算子 B 是病态的,对于高斯白噪声通过直接的最小二乘往往会得到无数个解,而这些解和 f^* 都相差很远.因此需要通过最小二乘问题加以正则化,即增加一项关于 f^* 的特征约束.由于图像自身的特征,人们普遍采用两种常见通用的正则化方法:低秩正则化^[1]和全变分正则化^[2].这两种方法是基于自然图像的两种不同的基本特征.在本文的第二章节,我们将详细地介绍这两种正则化以及它们的混合正则化.

由于混合正则化结合了两种正则化的性质,所以其效果往往会好于单个特征正则化的效果.但是现有的关于混合正则化的工作都只是停留在了凸方法的建模,混合正则化的发展也仅仅是使用了 L_1 范数和核范数.因此本文考虑非凸的低秩和全变分混合正则化模型.这是因为虽然 L_1 范数和核范数可以诱导稀疏和低秩特征,但是他们毕竟是零范数和秩范数的较为粗糙的近似.而且已有的结果也表明非凸函数相对于凸函数对于去模糊和低秩恢复的问题具有一定的优势^[3-4].所以我们考虑使用 $L_{1/2}$ 范数和 Schatten-1/2 范数正则化的非凸模型,以得到更好的去模糊效果.我们给出该非凸模型的求解算法及其算法的收敛性分析,我们也给出了在测试图像上新算法的数值结果以验证所提算法的有效性.

2 相关工作

由于相关工作有数学描述,因此在开始介绍相关工作之前,我们给出一些必要的数学符号.令 $X = (X_{i,j}) \in \mathbb{R}^{m \times n}$,本文中 $\|X\|$ 代表向量和矩阵的二范数或者 Frobenius 范数,即 $\|X\| := \sqrt{\sum_{i=1}^m \sum_{j=1}^n X_{i,j}^2}$. $\|\cdot\|_1$ 代表向量和矩阵的一范数,即 $\|X\|_1 :=$

$\sum_{i=1}^m \sum_{j=1}^n |X_{i,j}|$, 而 $L_{1/2}$ 范数的定义为 $\|X\|_{1/2} := (\sum_{i=1}^m \sum_{j=1}^n |X_{i,j}|^2)^{1/2}$, $\|X\|$ 的 Schatten-1/2 范数定义为 $\sum_i \sqrt{\sigma_i(X)}$, 其中 $\sigma_i(X)$ 代表矩阵 X 的第 i 个最大的特征值. 符号 \circ 代表矩阵元素相乘, 即 $(A \circ B)_{i,j} = A_{i,j} \cdot B_{i,j}$. 对于复数矩阵 X , X^* 代表 X 的共轭装置矩阵. 对 $r > 0$, 定义算子 \mathcal{S}_r 为 $r\sqrt{\cdot}$ 的近似点算子, $\mathcal{S}_r(t) \in \arg \min_m \{r\sqrt{|m|} + \frac{1}{2}(m-t)^2\}$. 若 X 为向量或矩阵, \mathcal{S}_r 即为作用到向量或矩阵的每个元素. Lu 等在 [10] 中证明了以下结论 $U \text{diag}[\mathcal{S}_r(\sigma_i(X))] V \in \arg \min_Y \{r \sum_i \sqrt{\sigma_i(Y)} + \frac{1}{2} \|Y - X\|^2\}$, 其中 U, V 是 X 的 SVD 分解的正交矩阵.

在低秩正则化部分,我们将给出直观的例子(自然图像的奇异值分布)来解释低秩正则化相关工作的合理性. 因此,我们选取了图像处理领域的六张代表性测试图片,这些图片也将用于我们的实验. 这六张测试图片(如图1所示)依次为:“Lena”、



图1 测试图像

“Barbara”、“Man”、“Cameraman”、“House”和“Pepper”.

2.1 低秩正则化

虽然自然图像本身往往都不是低秩的,但是一般都是主成分占优的. 也就是说在一般情况下,对于一个自然图像 f , 存在一个低秩矩阵 \hat{f} 使得 $\|f - \hat{f}\|$ 比较小. 一种常见的获得 \hat{f} 的办法就是直接对 f 做 SVD 分解, 然后取较大的奇异值对应的部分. 对于图1中的六张图像, 图2给出其奇异值分布, 可以看到这些图像的奇异值下降地都很快.

低秩正则化正是利用了这一特性, 最开始求解的问题数学表达式为

$$\min_f \left\{ \frac{1}{2} \|B(f) - e\|^2 + \lambda \cdot \text{rank}(f) \right\}, \quad (1)$$

其中 $\lambda > 0$ 是参数. 低秩性本质上是矩阵在谱域上的稀疏, 而矩阵的秩也可以看做是谱域上的零范数. 受压缩感知中的凸松弛方法启发, 人们也考虑问题(1)的凸近似模型, 其数学表达式为

$$\min_f \left\{ \frac{1}{2} \|B(f) - e\|^2 + \lambda \|f\|_* \right\}, \quad (2)$$

而 $\|f\|_*$ 则表示 f 的核范数, 也就是 f 所有奇异值之和. 相对于问题(1), 问题(2)的优势在于可以保证全局最优解同时在统计意义上更便于分析. 求解(2)的优化方法是向前向后分裂算法^[5]. 在后来的工作中, 人们发现由 Schatten- p 范数 ($0 < p < 1$) 正则化得到的效果往往比核范数要有优势^[4], 即如下问题

$$\min_f \left\{ \frac{1}{2} \|B(f) - e\|^2 + \lambda \sum_i \sigma_i(f)^p \right\}, \quad (3)$$

这是因为核范数是秩范数的一种相对粗糙的近似, 而 Schatten- p 范数更加接近秩范数. 对于一般的 $p \in (0, 1)$ 且 $p \neq 1/2, 2/3$, 求解(3)无法使用向前向后分裂算法, 这是因为此时 $\lambda \sum_i \sigma_i(f)^p$ 的近似点算子并不存在解析解. 这时使用的方法一般为重加权核范数极小化算法^[6-8,10]. 然而重加权核范数极小化算法并不是直接求解(3), 事实上他求解的是一个光滑化的 Schatten- p 范数, 即 $\sum_i (\sigma_i(f) + \epsilon)^p$ ($\epsilon > 0$ 且很小) 正则化的问题. 但 $p = 1/2, 2/3$ 时, $\sum_i \sigma_i(f)^p$ 有显式近似点表达. 特别当 $p = 1/2$ 时, 近似点可以有快速算法^[9-10].

2.2 全变分正则化

全变分正则化利用了自然图像是分片光滑的这一特性, 即相邻的像素值相减所得的矩阵是近似稀

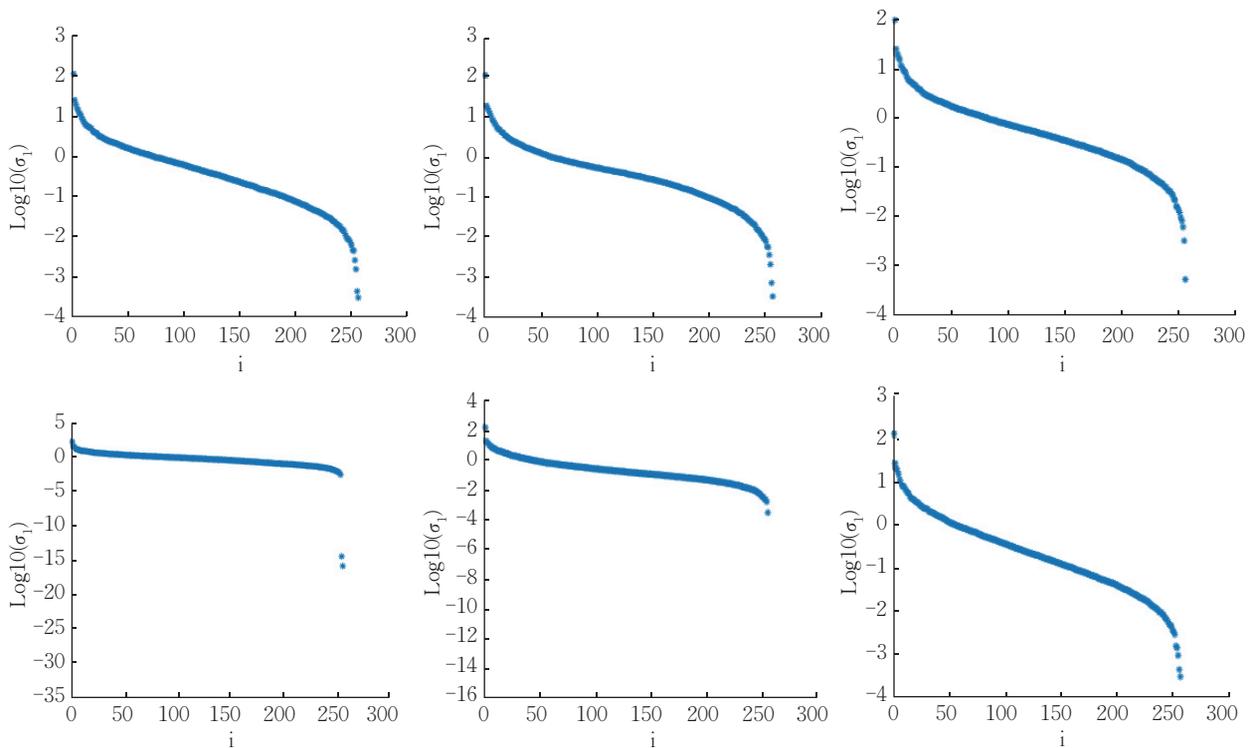


图2 奇异值分布

疏的. 全变分正则化即考虑 $D(X)$ 稀疏意义下的极小化. 凸的全变分正则化模型可以表达为

$$\min_f \left\{ \frac{1}{2} \|B(f) - e\|^2 + \eta \|D(f)\|_1 \right\}, \quad (4)$$

其中 $\eta > 0$ 是参数. 求解问题(4)的一般方法是ADMM^[11]或者ADM方法^[12]. 由于 p 范数往往有着更好的诱导稀疏性的特质, 很自然地, 也可以考虑 p 范数诱导稀疏性的模型, 即 $\|D(f)\|_p, 0 < p < 1$ 正则化^[3]

$$\min_f \left\{ \frac{1}{2} \|B(f) - e\|^2 + \eta \|D(f)\|_p \right\}. \quad (5)$$

对于问题(5), 求解的方法有非凸的ADMM^[13]和重加权ADMM算法^[14]. 但是由于理论上非凸ADMM和重加权ADMM收敛性的要求对涉及的矩阵较为苛刻使得去模糊问题不能满足该条件. 而且由于在非凸的ADMM中, 朗格朗日乘子取值很大, 这使得算法迭代地比较慢. 为了克服这两个缺点, 在[15]中, 作者提出了一种基于参数预热技术的交替极小化算法.

2.3 混合正则化

全变分正则化和低秩正则化是基于图像两种不同的基本特征: 一个是图像的分片光滑性质, 即图像的横向和纵向的差分 $D(X)$ 十分稀疏; 另一个则是图像的低秩性, 即图像往往可以被一个低秩矩阵

逼近地很好. 图像去模糊的前期工作都是使用单个性质, 其组合使用是由He等人在2016年提出的^[16]在2016年提出的. 由于使用了两种不同的特征, 其效果要好于单个正则化的效果. 在[16]中, 作者提出如下混合正则化模型

$$\min_f \left\{ \frac{1}{2} \|B(f) - e\|^2 + \lambda \|f\|_* + \eta \|D(f)\|_1 \right\}. \quad (6)$$

混合正则化(6)结合了低秩和全变分正则化, 其去模糊效果往往要好于单个特征正则化模型的效果. 该模型只使用了凸函数, 本文旨在建立非凸的混合正则化模型以及设计相应的算法.

3 非凸混合正则化模型和优化算法

由于 $L_{1/2}$ 范数和Schatten-1/2核范数的近似点算子都有快速算法, 这使得我们有足够的理由考虑使用 $L_{1/2}$ 范数和Schatten-1/2核范数替代(6)中的 L_1 范数和核范数. 我们考虑以下正则化问题

$$\min_f \left\{ \frac{1}{2} \|B(f) - e\|^2 + \lambda \sum_i \sqrt{\sigma_i(f)} + \eta \|D(f)\|_{\frac{1}{2}} \right\}, \quad (7)$$

其中 $\eta, \lambda > 0$ 是正则化参数. 一个直观的求解(7)的方法是ADMM算法. 但是由于我们求解的是非凸

问题,在这种情况下,ADMM算法的收敛性较为苛刻且参数确定较为困难.事实上,直接使用ADMM很难收敛.因此在本文中,我们考虑一种针对(7)的罚子法.通过引用两个新变量 y 和 z 以及一个很大的惩罚参数 $\rho > 0$,我们考虑以下非约束的优化模型

$$\min_{f,y,z} \left\{ \ell_{\rho}(f,y,z) := \frac{1}{2} \|B(f) - e\|^2 + \lambda \sum_i \sqrt{\sigma_i(y)} + \eta \|z\|_{\frac{1}{2}} + \frac{\rho}{2} \|f - y\|^2 + \frac{\rho}{2} \|D(f) - z\|^2 \right\} \quad (8)$$

对于问题(8)我们使用交替极小化算法,也就是在每一步迭代固定其他两个变量而极小化其中的一个.但是如果 ρ 过大,会使得迭代过程很慢.因此我们考虑一种逐步增大的技巧,也就是选取较小的 ρ ,但是在迭代过程中不断增大 ρ ,具体来说可以选择一个 $\alpha > 1$ 在每步迭代中 $\rho \leftarrow \alpha\rho$ 但是始终使 ρ 置于一个固定的大的正数值 $\bar{\rho}$ 之下.该技巧广泛应用很多工作中^[15,17].如果直接使用交替极小化算法,其在第 $k+1$ 次迭代的具体的数学迭代表式为

$$\begin{cases} f^{k+1} \in \arg \min_f \ell_{\rho^k}(f, y^k, z^k), \\ y^{k+1} \in \arg \min_y \ell_{\rho^k}(f^{k+1}, y, z^k), \\ z^{k+1} \in \arg \min_z \ell_{\rho^k}(f^{k+1}, y^{k+1}, z), \\ \rho^{k+1} = \min \{ \alpha\rho^k, \bar{\rho} \}. \end{cases}$$

然而这个算法无法保证充分下降,因此我们考虑对 y^{k+1} 和 z^{k+1} 使用以下交替极小化算法格式

$$\begin{cases} y^{k+1} \in \arg \min_y \left\{ \ell_{\rho^k}(f^{k+1}, y, z^k) + \frac{\delta\rho^k}{2} \|y - y^k\|^2 \right\}, \\ z^{k+1} \in \arg \min_z \left\{ \ell_{\rho^k}(f^{k+1}, y^{k+1}, z) + \frac{\delta\rho^k}{2} \|z - z^k\|^2 \right\}, \end{cases}$$

其中 $\delta > 0$ 是参数.我们给出 $k+1$ 步的各个变量的具体迭代格式.令 $D_{i,j} = (D_{i,j}^{(1)}, D_{i,j}^{(2)}) \in \mathbb{R}^2$ 为像素在 (i,j) 位置的变分映射,其数学定义为

$$D_{i,j}(X) := \begin{cases} (X_{i+1,j} - X_{i,j}, X_{i,j+1} - X_{i,j}), & \text{if } 1 \leq i \leq n-1, 1 \leq j \leq n-1, \\ (0, X_{i,j+1} - X_{i,j}), & \text{if } i = n, 1 \leq j \leq n-1, \\ (X_{i+1,j} - X_{i,j}, 0), & \text{if } 1 \leq i \leq n-1, j = 0, \\ (0, 0), & \text{if } i = n, j = n. \end{cases}$$

再定义 $D^{(1)} := (D_{i,j}^{(1)})$, $D^{(2)} := (D_{i,j}^{(2)})$ 和 $D := \begin{pmatrix} D^{(1)} \\ D^{(2)} \end{pmatrix}$.求解 f^{k+1} 本质是解如下的线性方程

$$\begin{aligned} & ((D^{(1)})^{\top} D^{(1)} + (D^{(2)})^{\top} D^{(2)} + B^{\top} B + \rho^k \cdot \text{Id}) f^{k+1} \\ & = (D^{(1)})^{\top} z^k + (D^{(2)})^{\top} z^k + B^{\top} e + \rho^k y^k. \end{aligned} \quad (9)$$

考虑到(9)中涉及的矩阵具有明显的结构信息,我们考虑使用[12]中对于类似线性方程组的求解算法:快速傅里叶变换.对(9)两边做快速傅里叶变换,由傅里叶卷积定理可以马上得到

$$f^{k+1} = \mathcal{F}^{-1} \left[\begin{aligned} & (\mathcal{F}(D^{(1)})^{\circ} \mathcal{F}(z^k) + \mathcal{F}(D^{(2)})^{\circ} \mathcal{F}(z^k) + \\ & \mathcal{F}(B)^{\circ} \mathcal{F}(e) + \rho^k \mathcal{F}(y^k)) \\ & / (\mathcal{F}(D^{(1)})^{\circ} \mathcal{F}(D^{(1)}) + \mathcal{F}(D^{(2)})^{\circ} \mathcal{F}(D^{(2)}) + \\ & \mathcal{F}(B)^{\circ} \mathcal{F}(B) + \rho^k \cdot \mathbf{1}^{\top} \mathbf{1}) \end{aligned} \right] \quad (10)$$

需要指出的是上式中的除也指的是元素相除.我们再考虑求解 y^{k+1} ,通过化简可以得到

$$y^{k+1} \in \arg \min_y \left\{ \frac{\lambda}{(1+\delta)\rho^k} \sum_i \sqrt{\sigma_i(y)} + \frac{1}{2} \|y - \frac{f^{k+1} + \delta y^k}{1+\delta}\|^2 \right\}.$$

因此求解 y^{k+1} 实际就是 $(f^{k+1} + \delta y^k)/(1+\delta)$ 在 $\lambda \sum_i \sqrt{\sigma_i(y)}/(1+\delta)\rho^k$ 的近似点算子下的映射,也就是

$$y^{k+1} = U^k \text{diag} \left[\mathcal{S}_{\frac{\lambda}{(1+\delta)\rho^k}} \left(\frac{f^{k+1} + \delta y^k}{1+\delta} \right) \right] V^k,$$

其中 U^k, V^k 是 $(f^{k+1} + \delta y^k)/(1+\delta)$ 的SVD分解的正交矩阵.类似地,我们可以得到对于 z^{k+1} 的迭代公式

$$z^{k+1} = \mathcal{S}_{\frac{\eta}{(1+\delta)\rho^k}} \left(\frac{1}{1+\delta} [D^{(1)}(f^k), D^{(2)}(f^k)] + \frac{\delta}{1+\delta} z^k \right).$$

4 非凸收敛性分析

在本节中,我们给出算法的收敛性分析.我们首先介绍一下函数次微分的概念^[18].令 $J: \mathbb{R}^N \rightarrow (-\infty, +\infty]$ 为合适的闭函数,即 $J(x) > -\infty$ 且集合 $\{(x,t) | t \geq f(x)\}$ 是闭集.给定一个 $x \in \text{dom}(J)$, J 在 x 处的Frechet次微分记为 $\hat{\partial}J(x)$,其定义为 $u \in \mathbb{R}^N$ 且满足 $\liminf_{y \neq x, y \rightarrow x} \frac{J(y) - J(x) - \langle u, y - x \rangle}{\|y - x\|} \geq 0$.当 $x \notin \text{dom}(J)$,我们记为 $\hat{\partial}J(x) = \emptyset$. J 在 $x \in \mathbb{R}^N$ 次梯度记为 $\partial J(x)$,是由以下过程定义的 $\partial J(x) := \{u \in \mathbb{R}^N : \exists x^k \rightarrow x, J(x^k) \rightarrow J(x), u^k \in \hat{\partial}J(x^k) \rightarrow u, k \rightarrow \infty\}$.

若某个点 x 极小化函数 J ,称 x 为 J 的临界点,且有 $0 \in \partial J(x)$,反之不然.由定义也可以知道对于任

意 x ,次微分 $\partial J(x)$ 是闭集合.我们再回顾一个关于强凸函数的性质,假设 h 是强凸的,若 $x^* \in \arg \min_x h(x)$ 那么

$$h(x) - h(x^*) \geq \frac{\nu}{2} \|x - x^*\|^2. \quad (11)$$

定理. 如果 $k > \ln \frac{\bar{\rho}}{\rho_0} / \ln \alpha + 1 =: \bar{K}$,那么

$$\begin{aligned} \ell_{\bar{\rho}}(f^k, y^k, z^k) - \ell_{\bar{\rho}}(f^{k+1}, y^{k+1}, z^{k+1}) &\geq \frac{\bar{\rho}}{2} \|f^{k+1} - f^k\|^2 \\ &+ \frac{\delta \bar{\rho}}{2} \|y^{k+1} - y^k\|^2 + \frac{\delta \bar{\rho}}{2} \|z^{k+1} - z^k\|^2. \end{aligned} \quad (12)$$

对于点列 $(f^k, y^k, z^k)_{k \geq 0}$ 的任意聚点,那么 (f^*, y^*, z^*) 也是函数 $\ell_{\bar{\rho}}(f, y, z)$ 的临界点.

证明. 容易发现,当 $k > \ln \frac{\bar{\rho}}{\rho_0} / \ln \alpha + 1$,我们有 $\rho^k = \bar{\rho}$.注意到 $\ell_{\bar{\rho}}(f, y^k, z^k)$ 相对于 f 是 $\bar{\rho}$ 强凸函数,由(11)以及 $f^{k+1} \in \arg \min_f \ell_{\bar{\rho}}(f, y^k, z^k)$,那么

$$\ell_{\bar{\rho}}(f^k, y^k, z^k) - \ell_{\bar{\rho}}(f^{k+1}, y^k, z^k) \geq \frac{\bar{\rho}}{2} \|f^{k+1} - f^k\|^2. \quad (13)$$

再由算法格式直接可以得到

$$\begin{aligned} \ell_{\bar{\rho}}(f^{k+1}, y^k, z^k) &\geq \ell_{\bar{\rho}}(f^{k+1}, y^{k+1}, z^k) + \frac{\delta \bar{\rho}}{2} \|y^{k+1} - y^k\|^2, \\ \ell_{\bar{\rho}}(f^{k+1}, y^{k+1}, z^k) &\geq \ell_{\bar{\rho}}(f^{k+1}, y^{k+1}, z^{k+1}) + \frac{\delta \bar{\rho}}{2} \|z^{k+1} - z^k\|^2. \end{aligned} \quad (14)$$

结合(13)和(14),我们即可得到(12).注意到 $\ell_{\bar{\rho}}(f, y, z) \geq 0$,我们可以得到

$$\begin{aligned} \sum_{i=\bar{K}}^k \left\{ \frac{\bar{\rho}}{2} \|f^{i+1} - f^i\|^2 + \frac{\delta \bar{\rho}}{2} \|y^{i+1} - y^i\|^2 + \frac{\delta \bar{\rho}}{2} \|z^{i+1} - z^i\|^2 \right\} &\leq \ell_{\bar{\rho}}(f^0, y^0, z^0) - \ell_{\bar{\rho}}(f^{k+1}, y^{k+1}, z^{k+1}) \\ &\leq \ell_{\bar{\rho}}(f^0, y^0, z^0). \end{aligned}$$

令 $k \rightarrow +\infty$,在上式中我们就可以得到

$$\begin{aligned} \sum_k \frac{\bar{\rho}}{2} \|f^{k+1} - f^k\|^2 + \frac{\delta \bar{\rho}}{2} \|y^{k+1} - y^k\|^2 \\ + \frac{\delta \bar{\rho}}{2} \|z^{k+1} - z^k\|^2 < +\infty. \end{aligned}$$

因此,我们可以得出以下结论

$$\lim_k \|f^{k+1} - f^k\| = \lim_k \|y^{k+1} - y^k\| = \lim_k \|z^{k+1} - z^k\| = 0. \quad (15)$$

令 $(f^{k_j}, y^{k_j}, z^{k_j}) \rightarrow (f^*, y^*, z^*)$,由(15),也就有 $(f^{k_j+1}, y^{k_j+1}, z^{k_j+1}) \rightarrow (f^*, y^*, z^*)$.当 $k \geq \bar{K}$ 时,在 k_j 次的迭代中,根据次微分的性质容易发现有

$$\begin{cases} 0 \in \partial_f \ell_{\bar{\rho}}(f^{k_j+1}, y^{k_j}, z^{k_j}), \\ 0 \in \partial_y \ell_{\bar{\rho}}(f^{k_j+1}, y^{k_j+1}, z^{k_j}) + \delta \rho (y^{k_j+1} - y^{k_j}), \\ 0 \in \partial_z \ell_{\bar{\rho}}(f^{k_j+1}, y^{k_j+1}, z^{k_j+1}) + \delta \rho (z^{k_j+1} - z^{k_j}). \end{cases}$$

令 $j \rightarrow +\infty$,我们就能推出

$$\begin{cases} 0 \in \partial_f \ell_{\bar{\rho}}(f^*, y^*, z^*), \\ 0 \in \partial_y \ell_{\bar{\rho}}(f^*, y^*, z^*), \\ 0 \in \partial_z \ell_{\bar{\rho}}(f^*, y^*, z^*), \end{cases}$$

也就是说 (f^*, y^*, z^*) 也是函数 $\ell_{\bar{\rho}}(f, y, z)$ 的临界点.

由定理1,我们可以看出, ρ_0 和 α 的选取不会对算法最终的结果造成影响.在算法中,对算法性能有影响的参数是 $\delta, \bar{\rho}$.

5 数值实验

在本节中,我们给出算法的相关实验.所有的实验都是MATLAB程序编写的,我们使用的是Intel CPU @2.6 GHz的笔记本计算机.我们使用峰值信噪比(PSNR)来表征图像去模糊的效果.本部分包括两部分:第一部分是算法自身的表现,包括算法敛性和参数 δ 和 $\bar{\rho}$ 对算法的影响.第二部分是算法和已有的经典算法的性能比较.

5.1 算法自身表现

在这个测试中,噪声水平设为 10^{-7} .其他参数设置为 $\delta = 10^{-12}$, $\rho^0 = 0.1$, $\alpha = 3$, $\eta = \lambda = 10^{-7}$.最大迭代步数设为100.图3给出了算法在高斯模糊算子作用于“Lena”,“Man”,“Cameraman”,以及运动模糊算子作用于“Barbara”,“House”,“Pepper”的去模糊表现.可以看到PSNR最终都趋于平稳.容易发现,算法在最开始的几步迭代过程中非常不稳定.这是由于我们用的是惩罚算子的方法.因此目标函数和算法极小化的函数有所不同,而且在计算过程中我们不断增大惩罚算子,这使得目标函数一直在变化.因此导致了收敛曲线下降的情况.其实对于恒定惩罚算子的情况,曲线一直是下降的.

对于 $\delta = 1, 10^{-3}, 10^{-6}, \bar{\rho} = 10^7, 10^9$ 的情形,图4给出算法在高斯模糊算子下的去模糊效果比较.从左至右,从上至下依次为 $\delta = 1, \bar{\rho} = 10^7$, $\delta = 1, \bar{\rho} = 10^9$, $\delta = 10^{-3}, \bar{\rho} = 10^7$, $\delta = 10^{-3}, \bar{\rho} = 10^9$, $\delta = 10^{-6}, \bar{\rho} = 10^7$ 和 $\delta = 10^{-6}, \bar{\rho} = 10^9$.

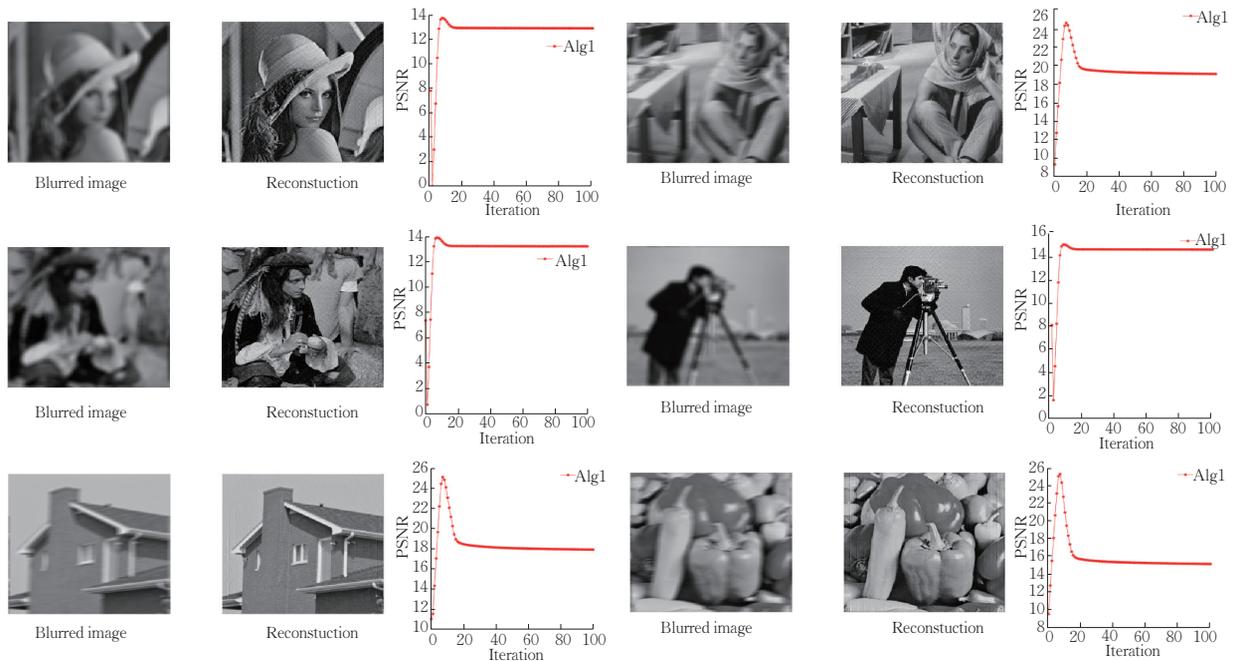


图3 算法性能

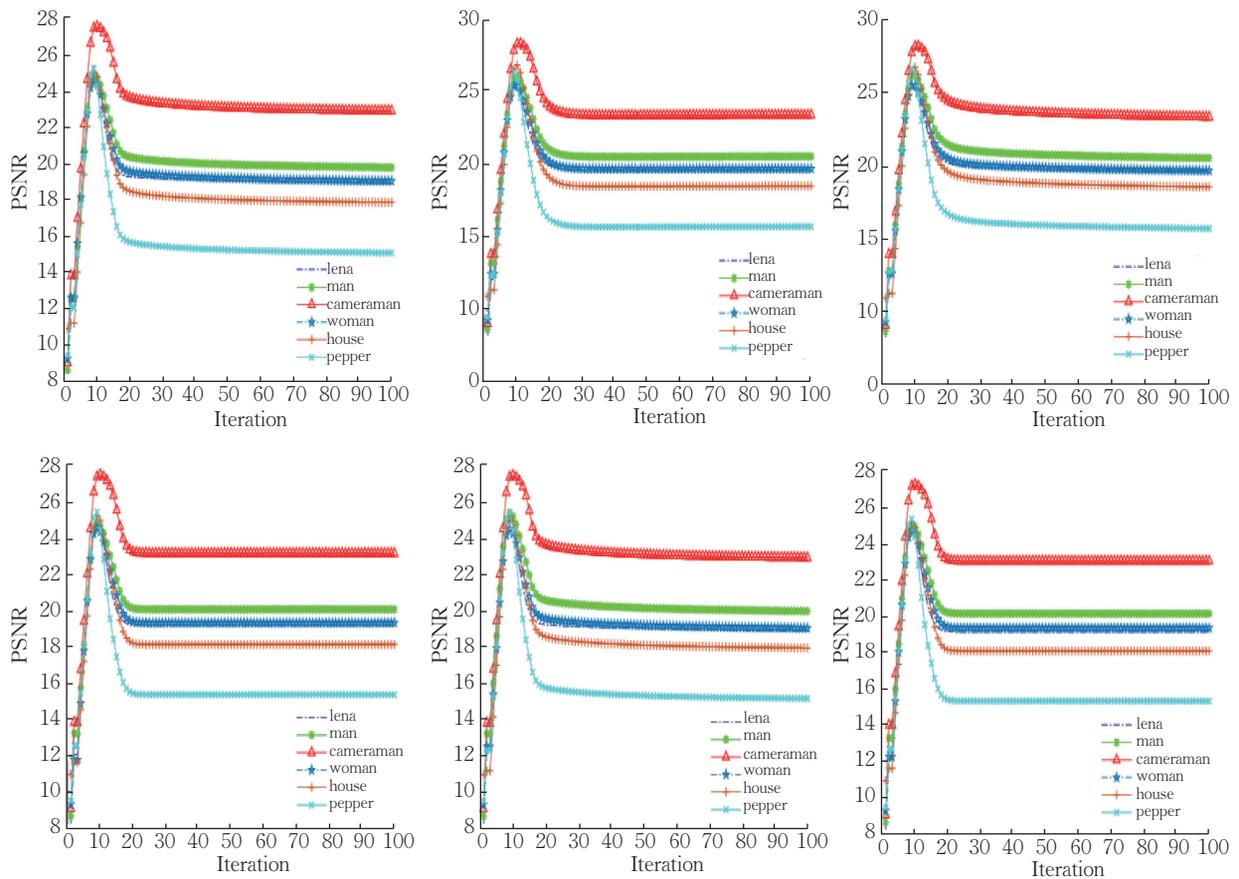


图4 不同 δ 和 $\bar{\rho}$ 下的算法表现

5.2 和其他算法的比较

我们考虑的对比如算法有四个,分别是 TV1^[12], TVLR^[16], TV05^[23]. 另外我们考虑重加权的低秩方

法^[24]. 但是单单只用低秩正则化对于去模糊的效果并不好,因此我们增加了 TV 项得到新的方法称之为 LR05TV. 对于混合正则化模型,我们使用 $\eta =$

$\lambda = 10^{-7}$. 在所有算法中,我们使用相同的 $\rho_0, \bar{\rho}$ 和 α . 迭代步数设为100. 运动模糊核是由MATLAB代码为 `fspecial('motion', 17, 3)`, 而高斯模糊算子则为 `fspecial('gaussian', 17, 3)`. 图5给出了在六幅图不同算法性能比较. 图6给出不同算法对于不同测试图象PSNR随迭代步数变化的趋势图. 在表1中,我们给出各个算法对于不同测试图像最后输出的结果和所需的迭代步数. 从表1中可以看到,本文算法相对于已有的经典算法都可以稳定地提高1~3点的dB. 本文给出的算法收敛时输出的结果PSNR要高于其他算法,从而说明了本文算法表现地更优秀.

6 讨 论

我们给出的算法可以用于解决盲去模糊问题的问题. 考虑将文献[25]中提出的经典的混合化盲去模糊模型作如下改变得到

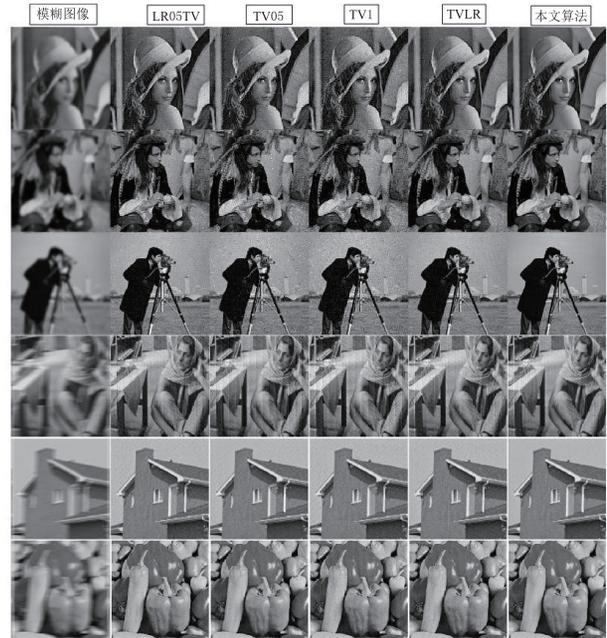


图5 不同算法性能比较

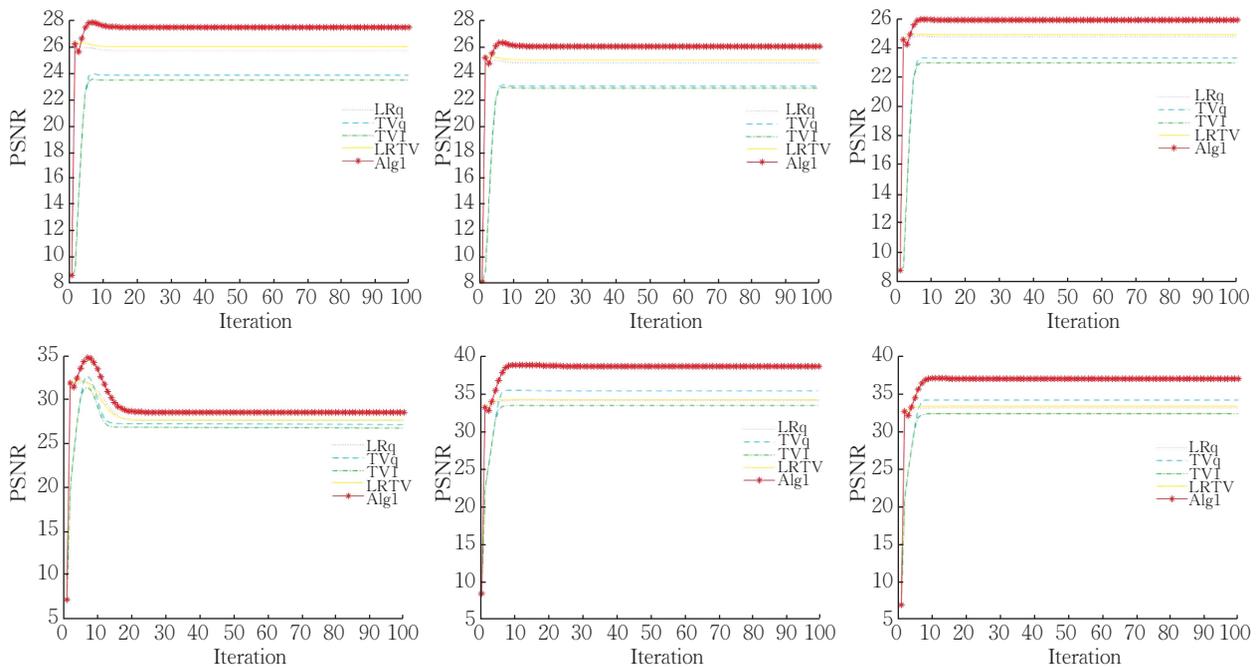


图6 不同算法对于不同测试图象PSNR随迭代步数变化的趋势

表1 不同算法在不同测试图片去模糊问题中收敛所需迭代步数(Iter)以及最后的PSNR(dB)

测试图片	LRq		TVq		TV1		LRTV		Alg1	
	Iter	PSNR								
Lena	23	30.4	29	30.1	23	29.5	22	30.4	24	31.8
Barbara	19	27.8	34	27.3	29	26.9	18	27.8	19	28.8
Man	25	30.7	29	30.8	25	30.2	24	30.8	26	32.2
Cameraman	26	32.2	21	33.0	19	31.8	25	32.2	27	34.3
House	6	33.9	7	35.1	7	33.2	5	34.0	19	38.4
Pepper	6	32.8	7	33.8	7	32.1	5	33.0	8	36.7

$$\min_{B,f} \left\{ \frac{1}{2} \|B(f) - e\|^2 + \lambda \sum_i \sqrt{\sigma_i(f)} + \eta \|D(f)\|_{\frac{1}{2}} \right\}, \text{ s.t. } K_B \geq 0, \|K_B\|_1 = 1.$$

其中 K_B 表示 B 的卷积核, 即 $B(f) = K_B * f$. 求解该问题可以通过 B 和 f 的交替极小得以解决. 即从初始化的 B^0, f^0 出发, 在第 k 步作如下迭代

$$\left\{ \begin{array}{l} f^{k+1} = \arg \min_f \left\{ \frac{1}{2} \|B^k(f) - e\|^2 + \lambda \sum_i \sqrt{\sigma_i(f)} + \eta \|D(f)\|_{\frac{1}{2}} \right\} \\ B^{k+1} = \arg \min \{ \|B(f^{k+1}) - e\|^2 \}, \text{ s.t. } K_B \geq 0, \\ \|K_B\|_1 = 1. \end{array} \right.$$

我们的算法可以用于求解第一个子问题.

我们使用的先验正则化项是有可能和深度学习结合的. 在论文[21]以及[22]中, 作者考虑了一种非线性反应扩散的深度学习框架用于图像去噪. 不同于以往的优化模型使用固定的正则化函数, 该框架运用学习的方法在一族函数中寻找 φ 来代替 l_1 范数. 从而全变分的正则项就是 $\varphi(D(f))$. 这要比 $L_{1/2}$ 范数更为一般. 但是这系列的工作都没有考虑使用低秩先验. 我们可以考虑在其模型上增加低秩的正则项. 但是这种直接的添加会增加其他计算上的复杂性. 如何有效地结合低秩正则化将留到下一步研究.

7 结 论

在本文中, 我们提出一种新的非凸的低秩和全变分的混合去模糊模型. 该模型使用了 $L_{1/2}$ 范数和 Schatten- $1/2$ 范数来表征图像的两种不同特征. 对于提出了模型我们使用交替极小化算法并给出了算法的收敛性分析. 在数值实验中, 该算法的性能表现优于已有的经典算法从而验证了算法的有效性.

参 考 文 献

- [1] Candes E J, Li X, Ma Y, et al. Robust principal component analysis? *Journal of the ACM (JACM)*, 2011, 58(3): 11. 2
- [2] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 1992, 60(1-4): 259-268
- [3] Hintermüller M, Wu T. Nonconvex TVq-models in image restoration: Analysis and a trust-region regularization-based superlinearly convergent solver. *SIAM Journal on Imaging Sciences*, 2013, 6(3): 1385-1415
- [4] Nie F, Huang H, Ding C. Low-rank matrix recovery via efficient Schatten p-norm minimization//*Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*. Toronto, Canada, 2012: 655-661
- [5] Combettes P L, Wajs V R. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 2005, 4(4): 1168-1200
- [6] Lu C, Tang J, Yan S, Lin Z. Generalized nonconvex nonsmooth low-rank minimization//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Ohio, USA, 2014: 4130-4137
- [7] Lu Z, Zhang Y. Schatten-p quasi-norm regularized matrix optimization via iterative reweighted singular value minimization. *arXiv preprint arXiv:1401.0869*, 2015
- [8] Sun T, Jiang H, Cheng L. Convergence of proximal iteratively reweighted nuclear norm algorithm for image processing. *IEEE Transactions on Image Processing*, 2017, 26(12): 5632-5644
- [9] Xu Z, Chang X, Xu F, et al. $L_{1/2}$ regularization: A thresholding representation theory and a fast solver. *IEEE Transactions on neural networks and learning systems*, 2012, 23(7): 1013-1027
- [10] Lu C, Zhu C, Xu C, et al. Generalized singular value thresholding//*Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*. Austin, USA, 2015: 1805-1811
- [11] Gabay D, Mercier B. A dual algorithm for the solution of non linear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 1976, 2(1): 17-40
- [12] Wang Y, Yang J, Yin W, et al. A new alternating minimization algorithm for total variation image reconstruction. *SIAM Journal on Imaging Sciences*, 2008, 1(3): 248-272.
- [13] Wang Y, Yin W, Zeng J. Global convergence of ADMM in nonconvex nonsmooth optimization. *Journal of Scientific Computing*, 2019, 78(1): 29-63
- [14] Sun T, Jiang H, Cheng L, et al. Iteratively linearized reweighted alternating direction method of multipliers for a class of nonconvex problems. *IEEE Transactions on Signal Processing*, 2018, 66(20): 5380-5391
- [15] Sun T, Li D, Jiang H, et al. Iteratively reweighted penalty alternating minimization methods with continuation for image deblurring.//*Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Brighton, United Kingdom, 2019: 3757-3761
- [16] He W, Zhang H, Zhang L, et al. Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration. *IEEE transactions on geoscience and remote sensing*, 2016, 54(1): 178-188
- [17] Padcharoen A, Kitkuan D, Kumam P, et al. Accelerated alternating minimization algorithm for Poisson noisy image

- recovery. *Inverse Problems in Science and Engineering*, 2020, 27(1): 1-26
- [18] Mordukhovich B S. *Variational analysis and generalized differentiation I: Basic theory [M]*. New York: Springer Science & Business Media, 2006
- [19] Huber P J. Robust regression: asymptotics, conjectures and Monte Carlo. *The Annals of Statistics*, 1973, 1(5): 799-821.
- [20] Sun T, Cheng L. Reweighted fast iterative shrinkage thresholding algorithm with restarts for ℓ_1 - ℓ_1 minimisation. *IET Signal Processing*, 2016, 10(1): 28-36.
- [21] Chen Y, Yu W, Pock T. On learning optimized reaction diffusion processes for effective image restoration// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA, 2015: 5261-5269
- [22] Qiao P, Dou Y, Feng W, et al. Learning non-local image diffusion for image denoising// *Proceedings of the 25th ACM international Conference on Multimedia*. Mountain View, USA, 2017: 1847-1855
- [23] Hintermüller M, Wu T. Nonconvex TV-q-models in image restoration: analysis and a trust-region regularization-based superlinearly convergent solver. *SIAM Journal on Imaging Sciences*, 2013, 6(3): 1385-1415
- [24] Gu S, Xie Q, Meng D, et al. Weighted nuclear norm minimization and its applications to low level vision. *International journal of computer vision*, 2017, 121(2): 183-208
- [25] Chan T F, Wong C K. Total variation blind deconvolution. *IEEE transactions on Image Processing*, 1998, 7(3): 370-375



Sun Tao, Ph. D. assistant professor, His current research interests include optimization for machine learning and image processing.

Li Dong-Sheng, Ph. D. professor. His current research interests include distributed computing, computer network, cloud computing, and big data research.

Background

This paper studies the optimization methods for the image processing. Specially, we consider the math methods for image deblurring. Basically, it can be regarded as an intersect of applied mathematics and computer science.

The current research in this area of image deblurring focuses on the convex and nonconvex TV model. Recently, with low-rank property, the composite TV+LR model is proposed. With the extra prior, the reconstructions are much

better.

However, the current TV+LR model only uses the convex function. Due to the nonconvex functions always perform better in the image processing compared with the convex ones, we proposed the nonconvex TV+LR model, which has never been proposed before. And the numerical results show that our model can work much better than the convex TV+LR.