

# 全变分图像复原的数值方法

邱欣欣

2014 年 8 月 21 日

## 1 引言

基于全变分的图像复原模型最初由 Rudin, Osher, and Fatemi(ROF) 在 [30] 中提出。此后, 它成为图像复原领域最流行和有效的模型之一。现在, 它不仅能处理基本的图像去噪问题, 还广泛应用于其他图像复原问题, 如去模糊, 盲复原, 和图像修补等, 这得益于 TV 良好的边缘保持性能。此后, 人们对于 TV 开展了进一步的研究: 如何构建更多复杂的变体以更好的重建图像, 设计更快的算法以求解最优化问题, 以及发现更多 TV 的应用领域。本文主要介绍全变分模型的基本概念及相关的数值方法。

## 2 基于全变分的数学模型

### 2.1 基本定义

在 Rudin, Osher, and Fatemi 的工作 [30] 中, 一幅图像  $u$  的 TV:  $\Omega \rightarrow \mathbb{R}$  被定义为

$$\int_{\Omega} |\nabla u| dx$$

其中  $|\nabla u| = \sqrt{u_x^2 + u_y^2}$ ,  $\Omega \in \mathbb{R}^2$  是一个有界的开集。

## 2.2 ROF 模型

ROF 模型处理的是下列的问题，给定一个噪声图像  $f: \Omega \rightarrow \mathbb{R}$ ，我们的目的是得到干净的图像  $u$ 。为了解决这个问题，Rudin, Osher 和 Fatemi 提出了下列的最小化问题：

$$\begin{aligned} & \min \int_{\Omega} |\nabla u| dx \\ \text{subject to } & \int_{\Omega} u dx dy = \int_{\Omega} f dx dy \quad \int_{\Omega} \frac{1}{2} (u - f)^2 dx dy = \sigma^2, \quad \sigma > 0 \end{aligned}$$

但通常我们考虑的是更一般的形式

$$\min_u \frac{1}{2} \int_{\Omega} (Ku - f)^2 dx + \lambda \int_{\Omega} |\nabla u| \quad (1)$$

在上式中， $f$  是观察到的图像， $u$  是去噪图像， $K$  是模糊算子，当  $K = I$  时，上述问题转化成去噪问题。TV 范数在平的区域内是不可微的，此时  $|\nabla u| = 0$ ，因此直接求解上述的最优化过程是困难的。由此人们提出了一系列的改进算法，如对偶/原始对偶方法，Bregman 迭代方法，图割方法 (graph cut methods) 等。需要注意的是，这些方法用于不同形式的 TV 公式，不同的性质和目的，不同的普适性，因而不能简单的进行对比。

## 2.3 BV( $\Omega$ )

Bounded variation (BV) 空间是一个极小化 ROF 模型的理想的函数空间，因为 BV 提供解的正则化并且允许不连续性（边界）。许多其他空间例如 Sobolev 空间  $W^{1,1}$  不允许边界。在定义 BV 空间之前，我们声明 TV 的定义如下

$$\int_{\Omega} |\nabla u| = \sup \left\{ \int_{\Omega} u \nabla \cdot p \, dx : p \in C_c^1(\Omega, \mathbb{R}^n), |p(x)| \leq 1 \, \forall x \in \Omega \right\}$$

其中  $u \in L^1(\Omega)$ ， $\Omega \subseteq \mathbb{R}^n$  是一个有界空集。因此 BV 空间定义为

$\{u \in L^1(\Omega) | \int_{\Omega} |\nabla u| < \infty\}$ 。现在 BV 空间函数相当于有界 TV 半范数的  $L^1$  函数。

同时，我们得到下列的等式形式

$$\int_{\Omega} |\nabla u| = \max_{p \in C_c^1(\Omega), |p| \leq 1} \int_{\Omega} \nabla u \cdot p = \max_{|p| \leq 1} \int_{\Omega} -u \nabla \cdot p \quad (2)$$

其中， $p : \Omega \rightarrow \mathbb{R}^2$ 。

注：以下介绍的方法如没有特殊说明，只适用于图像去噪与图像去模糊问题。

### 3 对偶和原始对偶方法

为了克服直接求解 TV 问题的困难，许多研究致力于通过求解公式 (1) 的变形形式来间接求解。其中，基于对偶与原始对偶的方法在实践中取得了良好的表现，下面介绍的是几个成功的对偶和原始对偶方法。

#### 3.1 Chan–Golub–Mulet 原始对偶方法 (Chan–Golub–Mulet’s Primal-Dual Method)

在早期的原始对偶方法研究中，Carter[6] 提出了 TV 去噪模型的一个 Fenchel 对偶公式，并研究了原始对偶的内点法，原始对偶松弛法。Chan, Golub, 和 Mulet (CGM) 在 [11] 中提出了一个原始对偶系统：由一个原始变量  $u$  和一个 Fenchel 对偶变量  $p$  组成的方程组。它是目前求解 TV 模型最有效的方法之一。它的过程是首先求解公式 (1) 的欧拉 - 拉格朗日方程：

$$K^T(Ku - f) - \lambda \nabla \cdot \left( \frac{\nabla u}{\sqrt{|\nabla u|^2 + \epsilon}} \right) = 0$$

引入一个辅助变量  $p$ ,

$$p = \left( \frac{\nabla u}{\sqrt{|\nabla u|^2 + \epsilon}} \right)$$

因此, 得到

$$\begin{aligned} p\sqrt{|\nabla u|^2 + \epsilon} &= \nabla u \\ K^T(Ku - f) - \lambda \operatorname{div} p &= 0 \end{aligned} \tag{3}$$

文中采用了牛顿法来求解方程组 (3)。

事实上, 我们可以把方程 (1) 看做一个原始问题, 它有相应的对偶问题, 而这两个问题的共同的解由方程组 (3) 给出, 也就是说通过求解方程组 (3) 可以求得 (1) 的解  $u$  和其对偶问题的解  $p$ , 具体的解释如下。

对 TV 的对偶式

$$\sup_{|p| \leq 1} \int_{\Omega} u \nabla \cdot p \, dx \, dy$$

其中使

$$p(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$$

满足限制条件, 那么对所有  $x \in \Omega, \nabla u(x) \neq 0$ ,

$$p|\nabla u| - \nabla u = 0$$

这就是方程组 (3) 中的第二个方程。则方程 (1) 可以写成

$$\min_u \sup_{\|p\|_{\infty} \leq 1} \Phi(u, p) \tag{4}$$

$$\Phi(u, p) = \int_{\Omega} \left( -\lambda u \nabla \cdot p + \frac{1}{2} (Ku - f)^2 \right) dx \, dy$$

根据对偶理论，我们得到 (4) 的解  $(u^*, p^*)$

$$\min_u \sup_{\|p\|_\infty \leq 1} \Phi(u, p) = \Phi(u^*, p^*) = \sup_{\|p\|_\infty \leq 1} \min_u \Phi(u, p)$$

其中， $p^*$  是下列对偶问题的解

$$\sup_{\|p\|_\infty \leq 1} \Psi(p)$$

$$\Psi(p) = \min_u \Phi(u, p)$$

由上式我们得出这个极小化问题必须满足下列等式

$$-\lambda \operatorname{div} \cdot p + K^*(Ku - f) = 0$$

这就是方程组 (3) 的第一个方程。

## 3.2 Chambolle 对偶方法 (Chambolle's Dual Method)

### 3.2.1 Notation

ROF 模型在 BV 空间作为一个无限维最优化问题，为了能够数值的求解它，必须离散化这个问题，这个方法首先由 Rudin 等在 [30] 中提出：“最优化然后离散”，将方程离散成一个标准的有限差分格式来求解。这种方法效果很好，并且数值解能收敛于一个稳定的程度并和连续的 ROF 模型的收敛结果趋于一致。

Chambolle 算法中需要用到离散化的方法，以下是相关的注释。

$X$  表示欧几里得空间  $\mathbb{R}^{N \times N}$ ， $p$  是映射  $p: \{1, \dots, N\} \times \{1, \dots, N\} \rightarrow \mathbb{R}^2$ ， $Y = X \times X$ 。

$$J_d(u) = \|\nabla u\|_Y = \sum_{1 \leq i, j \leq N} |(\nabla u)_{i,j}|$$

对于  $p \in Y$ ,  $u \in X$ , 利用对偶公式 (2) 我们有

$$J_d(u) := \max_{p \in \nu} \langle u, \text{div} p \rangle$$

其中,  $\nu = \{p \in Y : |p_{i,j}|^2 - 1 \leq 0, \forall i, j \in \{1, \dots, N\}\}$ .

### 3.2.2 Chambolle's algorithm

由 Chambolle 在 [7] 中提到了一个纯对偶方法, 它的 ROF 目标函数只写成对偶变量的形式。在文章中, 他提出求解下列问题的对偶式 (与 (1) 形式略有不同)

$$\min_{u \in X} J_d(u) + \frac{1}{2\lambda} \|u - f\|_X^2 \quad (5)$$

其中  $\|\cdot\|_X$  代表欧几里得范数,  $f \in X$ , 上式可化为

$$\begin{aligned} \min_{u \in X} J_d(u) + \frac{1}{2\lambda} \|u - f\|_X^2 &= \min_{u \in X} \max_{p \in \nu} \langle u, \text{div} p \rangle + \frac{1}{2\lambda} \|u - f\|_X^2 \\ &= \max_{p \in \nu} \min_{u \in X} \langle u, \text{div} p \rangle + \frac{1}{2\lambda} \|u - f\|_X^2 \end{aligned}$$

显式的求解  $u$  的极小值, 我们得到  $u = f - \lambda \text{div} p$ , 然后

$$\begin{aligned} \max_{p \in \nu} \min_{u \in X} \langle u, \text{div} p \rangle + \frac{1}{2\lambda} \|u - f\|_X^2 &= \max_{p \in \nu} \langle f, \text{div} p \rangle - \frac{\lambda}{2} \|\text{div} p\|_X^2 \\ &= -\frac{\lambda}{2} \min_{p \in \nu} \left( \left\| \text{div} p - \frac{f}{\lambda} \right\|_X^2 - \left\| \frac{f}{\lambda} \right\|_X^2 \right) \end{aligned}$$

因此如果  $p^*$  是下式的解

$$\min_{p \in \nu} \left\| \text{div} p - \frac{f}{\lambda} \right\|_X^2 \quad (6)$$

那么  $u = f - \lambda \text{div} p^*$  是 (5) 式的解。从而把求原始问题 (5) 的解变为求其对偶式 (6) 的解, 降低了求解难度, 其具体求解 (6) 的过程在《Mathematical Problems in Image Processing》第三章第二节有具体的阐述。

### 3.3 原始对偶有效集方法 (Primal-Dual Active-Set Method)

Hintermüller 和 Kunisch 在 [21] 中考虑到用 Fenchel 对偶方法去解决一个约束二次对偶问题并且得到了一种十分有效的有效集方法来处理这种约束。这种方法把变量分为有效集合和无效集合, 因此能够根据它们的性质区别对待。

#### 3.3.1 有效和无效的概念

一个不等式约束  $a^T x \geq \beta$  在点  $x$  被称作是有效的 (active) 如果约束满足等式  $a^T x = \beta$ , 一个不等式约束被称作是无效的 (inactive) 如果约束严格满足  $a^T x > \beta$  或违反约束条件使  $a^T x < \beta$ 。形式为  $a^T x = \beta$  的等式约束总是被认为是有效的 [25]。

PDAS 的优点在于如果在最优解处的有效集和无效集的不等式约束已知并作为先验知识, 可以通过解一个更简单的等式约束二次规划问题得到最终解。在算法中, 有效集方法做出对有效和无效约束在最优解的预测, 这种预测通常是错误的, 因此需要不断地试验并且改变当前的预测以得到最终的有效和无效变量。每一次迭代中需要解一个由有效集约束定义的 KKT 系统, 通过不断迭代当有相当数量的约束是有效的时候, 整个方程系统规模已经变的很小, 可以较容易求解。

#### 3.3.2 具体过程

考虑到下列的一般性的二次问题

$$\min_{y, y \leq \psi} \frac{1}{2} \langle y, Ay \rangle - \langle f, y \rangle$$

其中  $\psi$  是  $\mathbb{R}^n$  中的向量，其 KKT 条件为

$$\begin{aligned} Ay + \lambda &= f \\ \lambda \odot (\psi - y) &= 0 \\ \lambda &\geq 0 \\ \psi - y &\geq 0 \end{aligned} \tag{7}$$

$\lambda$  是拉格朗日乘子向量， $\odot$  代表分量乘积。前面提到需要对最终的有效与无效变量进行预测，这个预测是通过比较  $\lambda$  和  $\psi - y$  对 0 的相近度判断的。如果  $\psi - y$  趋近于 0 的速度是  $\lambda$  的  $c$  倍，那么这个变量被预测为有效的。

将 KKT 系统简写成下列的式子:

$$\begin{aligned} Ay^{k+1} + \lambda^{k+1} &= f \\ C(y, \lambda) &= 0 \end{aligned}$$

其中  $C(y, \lambda) = \min(\lambda, c(\psi - y))$ ,  $c$  为任意正常数，得到目前的原始对偶对  $(y, \lambda)$ 。

PDAS 算法如下:

1. 初始化  $y^0$ ,  $\lambda^0$ , 设  $k = 0$ 。
2. 设  $I^k = \{i : \lambda_i^k - c(\psi - y^k)_i \leq 0\}$ , 并且  $A^k = \{i : \lambda_i^k - c(\psi - y^k)_i > 0\}$ 。
3. 求解

$$\begin{aligned} Ay^{k+1} + \lambda^{k+1} &= f \\ y^{k+1} &= \psi \quad \text{on } A^k \\ \lambda^{k+1} &= 0 \quad \text{on } I^k \end{aligned}$$

4. 停止, 或设  $k = k + 1$  并且返回步骤 2。



在 [23] 中, Krishnan 等考虑到带边界约束的 TV 去模糊问题, 并提出了结合 cgm 与 pdas 算法的非负约束 cgm 算法。其中处理图像  $u$  和它的对偶  $p$  使用的是 cgm 方法, 而在  $u$  的边界约束则用 pdas 算法处理。结果的最优化条件被证明是半平滑的。

注: 除了以上三种方法以外, 下面的方法只是做一简略的介绍。

### 3.4 原始对偶混合梯度方法 (Primal-Dual Hybrid Gradient Method)

原始方法和对偶方法有各自的优点和相应的在数值方面需要克服的困难, 因此 Zhu 和 Chan 在 [32] 处提出了原始对偶混合算法 (PDHG) 算法, 可以实现轮流在原始形式和对偶形式中转换。

这种方法基于原始对偶公式:

$$G(u, p) := \frac{1}{2} \int_{\Omega} (Ku - f)^2 dx + \lambda \int_{\Omega} u \operatorname{div} p dx \rightarrow \inf_u \sup_{|p| \leq 1}$$

这个公式有两个子问题:

$$\sup_{|p| \leq 1} G(u, p) \quad \inf_u G(u, p)$$

PDHG 方法对每个子问题仅使用一步的梯度下降法, 基本原理是当两个变量没有一个是最优的时, 对每个子问题进行迭代都无法得到有用的信息, 直到收敛。通过一个初始猜想  $u^0$ , 不断重复下面的两个步骤:

$$\begin{aligned} p^{k+1} &= p_{|p| \leq 1}(p^k - \tau_k \nabla u^k) \\ u^{k+1} &= u^k - \theta_k (K^T(Ku^k - f) + \lambda \operatorname{div} p^{k+1}) \end{aligned}$$

$P_{|p| \leq 1}$  是到可行集  $p: |p| \leq 1$  的投影, 可以选择合适的步长  $\tau_k$  和  $\theta_k$  来

优化结果，一些相应的步长策略在 [7] 中提出。

在文中的结果表明这个简单的算法比 Split Bregman 迭代要快，也快于 Chambolle 的半隐式对偶算法。关于 PDHG 算法与其他算法的联系在文章 [16] 中有所描述。

### 3.5 半平滑牛顿法 (Semi-Smooth Newton's Method)

Ng 等 [27] 基于 Chambolle 的算法提出了一种半平滑牛顿算法。由 Section 3.2 我们可以得到其 KKT 条件如下：

$$\begin{aligned} -(\nabla(\lambda \operatorname{div} \mathbf{p} - \mathbf{g}))_{i,j} + \mu_{i,j} p_{i,j} &= 0 \\ |p_{i,j}|^2 - 1 &\leq 0 \\ \mu_{i,j} &\geq 0 \\ \mu_{i,j} (|p_{i,j}|^2 - 1) &= 0 \end{aligned}$$

我们了解到，如果约束满足  $a \geq 0, b \geq 0$  并且  $ab = 0$ ，那么它们能够统一成下列的等式约束：

$$\phi(a, b) := \sqrt{a^2 + b^2} - a - b = 0$$

其中  $\phi$  被称作 Fisher-Burmeister 函数，因此上述的 KKT 系统能被写为：

$$\begin{aligned} -(\nabla(\lambda \operatorname{div} \mathbf{p} - \mathbf{g}))_{i,j} + \eta_{i,j} p_{i,j} &= 0 \\ \phi(\mu_{i,j}, 1 - |p_{i,j}|^2) &= 0 \end{aligned}$$

Ng 等观察到这个系统是半平滑的，因此提出了使用半平滑牛顿法求解这个系统。在这个方法里，如果系统的雅可比矩阵由于缺乏足够的平滑度，不能定义为经典的形式，那么可以用相近点的广义的雅可比矩阵代替。如

果要求解的系统至少是半平滑的，并且其收敛处的广义雅可比矩阵满足某些可逆性条件，那么这个方法的收敛是超线性的。其具体的求解过程见论文 [9]。

注：在 Hintermuller 等 [20] 的论文中提到 PDAS 方法是一种特殊的半平滑牛顿法。

### 3.6 总结

以上提到的对偶或原始对偶方法，它们都满足原始对偶的相关理论，即在符合某些条件下 (KKT 条件)，原始最优化问题和对偶最优化问题可以取得相同的最优解，从而可以将难解的原始问题转化到相对易解的对偶问题或原始对偶问题上去。如 CGM 方法中将求解方程 (1) 准化为求解原始对偶方程组 (2)，Chambolle 对偶方法中将求解方程 (5) 转化为求解对偶式 (6)，PDAS 方法将二次规划问题转化为原始对偶系统 (7) 等，降低了求解困难，最终实现了对全变分一般化模型的求解。

## 4 Bregman 迭代

### 4.1 原始 Bregman 迭代 (Original Bregman Iteration)

Bregman 迭代是由 Osher 等 [29] 为实现 TV 去噪提出的，它也被推广以解决许多凸的逆问题。在每一步，在前一步被去除的信号又被加回来，这是为了降低 ROF 模型中对比度的丢失。

令  $f_0 = f$ ，下列步骤不断重复：

1. 令

$$u_{j+1} = \arg_u \min \left\{ \frac{1}{2} \int_{\Omega} (u - f_j)^2 dx + \lambda \int_{\Omega} |\nabla u| \right\}$$

2. 令  $f_{j+1} = f_j + (f - u_{j+1})$ .

Bregman 距离被定义为

$$D_J^p(u, v) = J(u) - J(v) - \langle p, u - v \rangle$$

$p$  是  $J$  的次梯度的元素, 在 TV 去噪情况下,  $J(u) = \lambda \int_{\Omega} |\nabla u|$ 。那么 Bregman 迭代变为

1. 令

$$u_{j+1} = \arg_u \min \left\{ \frac{1}{2} \int_{\Omega} (u - f)^2 dx + D_J^{p_j}(u, u_j) \right\}$$

2. 令  $f_{j+1} = f_j + (f - u_{j+1})$ .

3. 令  $p_{j+1} = f_{j+1} - f$ .

在迭代中, 只要  $L_2$  距离比噪声分量的量级大,  $u_j$  和干净图像的 Bregman 距离是单调递减的。Bregman 迭代的一个有趣的特征是, 在离散的情况下, 如果把  $\|u - f\|^2$  项换为  $\|Au - f\|^2$ , 其中  $Au = f$  是欠定的, 那么我们能得到下列的基追踪问题 (Basis pursuit problem) :

$$\min_u \{J(u) | Au = f\}$$

当  $\|Au - f\|^2$  用在目标中代替  $\|u - f\|^2$ , Bregman 迭代变为

1. 令

$$u_{j+1} = \arg_u \min \left\{ \frac{1}{2} \int_{\Omega} (Au - f)^2 dx + D_J^{p_j}(u, u_j) \right\}$$

2. 令  $f_{j+1} = f_j + (f - Au_{j+1})$ .

3. 令  $p_{j+1} = A^T(f_{j+1} - f)$ .

## 4.2 分割 Bregman 迭代 (Split Bregman Iteration)

最近, Goldstein 和 Osher[19] 提出了分割 Bregman 迭代, 能够有效的解决 ROF 问题。其主要思想是引入一个新的变量使 TV 极小化问题变为一个  $L^1$  极小化问题, 以便能用 Bregman 迭代有效的求解。其中, 将变量

$q = \nabla u$  引入目标函数中，然后用罚函数法求解以加强约束  $q = \nabla u$ 。包含罚函数的目标变为

$$G(u, q) = \frac{\alpha}{2} \int_{\Omega} |q - \nabla u|^2 dx + \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \lambda \int_{\Omega} |q| dx$$

当用  $y$  代替  $(u, q)$ ，那么上述的目标能被定义为

$$\min_y \left\{ \frac{\alpha}{2} \int_{\Omega} |Ay|^2 dx + J(y) \right\}$$

其中，

$$\begin{aligned} Ay &= q - \nabla u \\ J(y) &= \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \lambda \int_{\Omega} |q| dx \end{aligned}$$

当  $\alpha \rightarrow \infty$ ，这就是基追踪问题 (Basis pursuit problem)，事实上，即使  $\alpha$  是固定有限的，当使用 Bregman 迭代时，它收敛于问题

$$\min_y \left\{ J(y) | Ay = 0 \right\}$$

因此约束  $q = \nabla u$  在收敛处满足。

## 5 图割法 (Graph Cut Methods)

Graph Cut[8], [15] 方法近来在解决各式变分问题方面有较大的发展，它能快速的解决许多实际的问题甚至得到一些非凸问题的全局最优解。由于 Graph Cut 问题是组合学的，目标函数不得不是离散的。因此，给定的

$m \times n$  的图像  $f$  是一个从  $\mathbb{Z}_m \times \mathbb{Z}_n$  到  $\mathbb{Z}_k$  的函数。ROF 问题因此变为

$$F(u) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n (u_{i,j} - f_{i,j})^2 + \lambda \|u\|_{TV} \rightarrow \min_{u: \mathbb{Z}_m \times \mathbb{Z}_n \rightarrow \mathbb{Z}_k}$$

其中  $\|u\|_{TV}$  是离散 TV 形式。这个问题可以转化成下列的 Graph Cut 问题:

$$\min_{\nu^1, \nu^2, \dots, \nu^{K-2}} \sum_{k=0}^{K-2} F^k(\nu^k)$$

然后进一步求解, 关于具体的求解过程, 可以参考 [8], [15].

## 6 二次规划 (Quadratic Programming)

离散的各向异性 TV 是一个分段的线性函数, Fu 等在 [17] 中指出通过引入一些辅助变量, 可以将 TV 转换成一个有着线性约束的线性函数。

Fu 等考虑的目标函数为

$$F(u) = \frac{1}{2} \|Ku - f\|^2 + \lambda \sum_{i,j} |u_{i+1,j} - u_{i,j}| + |u_{i,j+1} - u_{i,j}|$$

也可以写作

$$F(u) = \frac{1}{2} \|Ku - f\|^2 + \lambda \|Ru\|_1$$

其中  $R$  是一个  $2mn \times mn$  矩阵, 如果使用原始各向同性 TV, 那就不能写做这个形式。

[17] 中采用的方法是令  $\nu = Ru$  然后把它分割成正的和负的部分:  $\nu^+ = \max(\nu, 0)$ ,  $\nu^- = \max(-\nu, 0)$ 。目标函数可以写作

$$G(u, \nu^+, \nu^-) = \frac{1}{2} \|Ku - f\|^2 + \lambda (1^T \nu^+ + 1^T \nu^-)$$

这是一个二次函数，但是添加了一些线性约束：

$$\begin{aligned} Ru &= \nu^+ - \nu^- \\ \nu^+, -\nu^- &\geq 0 \end{aligned}$$

现在，可以用标准原始对偶内点法来求解这个问题。这里的对偶指的是线性约束的拉格朗日乘数。

## 7 二阶锥规划 (Second-Order Cone Programming)

Goldfarb 和 Yin 在 [18] 中提出了可以用于各向同性 TV 形式的二阶锥规划 (SOCP) 方法，他们考虑的是下述的 ROF 问题：

$$\min_u \|u\|_{TV} \quad \text{subject to } \|u - f\| \leq \sigma$$

其中  $\sigma$  是假设已知的噪声的标准差。

让  $w_{i,j}^x = u_{i+1,j} - u_{i,j}$ ,  $w_{i,j}^y = u_{i,j+1} - u_{i,j}$ , TV 变为

$$\sum_{i,j} \sqrt{(w_{i,j}^x)^2 + (w_{i,j}^y)^2}$$

引入变量  $\nu = f - u$ ,  $t$ , 和约束

$$(w_{i,j}^x)^2 + (w_{i,j}^y)^2 \leq t_{i,j}^2$$

则 TV 最小化问题变为

$$\begin{aligned}
& \min \sum_{i,j} t_{i,j} \\
& s.t. \quad u + \nu = f \\
& \quad \quad w_{i,j}^x = u_{i+1,j} - u_{i,j} \\
& \quad \quad w_{i,j}^y = u_{i,j+1} - u_{i,j} \\
& \quad \quad (\sigma, \nu) \in cone^{mn+1} \\
& \quad \quad (t_{i,j}, w_{i,j}^x, w_{i,j}^y) \in cone^3
\end{aligned}$$

这里  $cone^n$  是  $\mathbb{R}^n$  的二阶锥:

$$\{x \in \mathbb{R}^n : \|(x_2, x_3, \dots, x_n)\| \leq x_1\}$$

最优解满足

$$t_{i,j}^2 = (w_{i,j}^x)^2 + (w_{i,j}^y)^2$$

因此

$$\sum_{i,j} t_{i,j} = \sum_{i,j} \sqrt{(w_{i,j}^x)^2 + (w_{i,j}^y)^2} = \|u\|_{TV}$$

SOCP 方法可以用内点法求解, 但 Goldfarb 和 Yin 提出了一个域分解方法, 通过将规模较大的规划问题分解成几个子问题来有效求解。

## 8 优化最小化 (Majorization-Minimization)

Majorization-Minimization(MM) 方法是求解最优化的一种方法, 它的主要思想是在方法的每一步, 用一个简单的函数 (叫做代理函数) 将目标函数代替, 这样它的最小化问题便易于求解, 最终结果给出原始问题的一个较小的目标值。在很多情况下, 甚至能将多维问题转化成一系列一维问



题。最近，一些作者 [3] 应用这种方法求解去模糊问题。求解 MM 方法通常采用预处理共轭梯度法，以及余弦变换，多重网格法等。其具体解释求解过程参考 [3]。

## 9 分裂法 (Splitting Methods)

分裂法也是一种可以求解 TV 去模糊问题的方法，它的思想是将原始问题分离成去模糊过程和 TV 正则化过程再分开求解。分裂法一般基于下列形式的目标函数

$$F(u) = J_1(u) + J_2(Au)$$

$A$  为线性算子， $J_1$  和  $J_2$  的最小化是分开进行的。对下列的离散 ROF 模型：

$$F(u) = \frac{1}{2}\|Ku - f\|^2 + \lambda\|\nabla u\|_1$$

Huang[22]，Bresson 和 Chan[4] 考虑到下列的分裂

$$G(u, v) = \frac{1}{2}\|Ku - f\|^2 + \frac{\alpha}{2}\|u - v\|^2 + \lambda\|\nabla u\|_1$$

最小化  $u$  可得

$$(K^T K + \alpha I)u = K^T f + \alpha v$$

这个式子可以用快速傅里叶变换求解，只要模糊矩阵  $K$  能被对角化成一个快速转换矩阵。最小化  $v$  是一个 ROF 去噪问题，可以用上面提到的去噪方法求解，[22]，[4] 使用的是 Chambolle 对偶算法。TV 去噪问题比去模糊问题更容易求解，一些基于 Graph cut 的方法不能直接应用于去模糊问题，但是通过分裂方法，我们可以应用 Graph cut 方法求解去噪问题。分裂法通常速度较快，并且可以应用于其他图像处理问题如图像分割等。

## 10 $L^1$ Fitting

正如前面提到的，ROF 模型的一个缺点是对比度的丢失。一个简单的修正 ROF 模型以补偿对比度丢失的方法是将变分公式中保真项的  $L^2$  范数替换为  $L^1$  范数。

$$\int_{\Omega} |\nabla u| + \lambda \int_{\Omega} |u - f| dx \quad (8)$$

Alliney[1] 研究了公式 (8) 处理一维信号的离散形式，Nikolova[28] 研究了高维形式。已经证明在处理某些特殊形式的噪声（脉冲噪声）如椒盐噪声时，这种方法要比标准 ROF 形式更加高效。Chan 和 Esedoglu 在 [9] 中研究了连续形式的解。

(8) 具有对比度不变性：如果  $f(x)$  是给定的图像， $u(x)$  是 (6) 的解，那么如果  $cf(x)$  是给定的图像， $cu(x)$  是 (6) 的解。

除了有更好的对比度保持性能，模型 (6) 被证明对处理”形状 (shapes)”的去噪问题有较好的效果。

## 11 TV 在图像盲复原、图像修复、图像分割中的应用

除了上述所说 TV 在图像去噪，图像去模糊方面的应用，TV 还应用于图像盲复原、图像修复的研究。其中，Chan 与 Wong 在 [14] 提出了下列 TV 盲复原模型

$$\min_{u, k \in BV} \|k * u - f\|_2^2 + \lambda_u \|u\|_{TV} + \lambda_k \|k\|_{TV}$$

其中  $f$  是观察到的图像， $k$  是点扩散函数 (point spread function)， $u$  是干净的图像， $*$  代表卷积算子。

Masnou 和 Morel[26], Bertalmio 等 [2], Chan and Shen[12] 提出了一些图像修补方法与模型。

TV 极小化问题也应用于图像分割问题, Chan 和 Vese[13] 提出了一个基于水平集的方法结合 Mumford–Shah 模型的方法来求解分割问题, Chan 等 [10], Krishnan 在 [24] 中考虑了离散的二相 Mumford–Shah 模型来求解。

此外, 全变分还成功应用在图像缩放, 颜色加强, 图像融合, 图像分解等方面。在《Level Set and PDE Based Reconstruction Methods in Imaging》一书中由 Martin Burger 和 Stanley Osher 撰写的《A Guide to the TV Zoo》一节中有详细的说明 [5]。

注: 本文档主要参考了《Handbook of Mathematical Methods in Imaging》[31]23、24 章的内容。

## 参考文献

- [1] S Alliney. A property of the minimum vectors of a regularizing functional defined by means of the absolute norm. *IEEE Trans. Signal Process*, 1997.
- [2] M Bertalmio, G Sapiro, V Caselles, and C Ballester. Image inpainting. *Proc. 27th Ann. Conf. on Comput. Graph. Interactive Tech*, 2000.
- [3] J Bioucas-Dias, M Figueiredo, and R Nowak. Total variation-based image deconvolution: a majorization-minimization approach. *ICASSP*, 2006.
- [4] X Bresson and T Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Probl Imaging*, 2008.

- [5] Martin Burger and Stanley Osher. *A Guide to the TV Zoo*. Springer International Publishing, 2013.
- [6] J Carter. Dual methods for total variation-based image restoration. *Ph.D. thesis*, 2001.
- [7] A Chambolle. An algorithm for total variation minimization and applications. *J Math Imaging Vis*, 2004.
- [8] A. Chambolle and J Darbon. On total variation minimization and surface evolution using parametric maximum flows. *Int J Comput Vis*, 1997.
- [9] T Chan and S Esedoglu. Aspects of total variation regularized  $l^1$  function approximation. *UCLA CAM Report*, 2004.
- [10] T. Chan, S. Esedoglu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM J Appl Math*, 2006.
- [11] T Chan, G Golub, and P Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J Sci Comp*, 1999.
- [12] T Chan and J Shen. Mathematical models of local non-texture inpaintings. *SIAM J. Appl. Math*, 2001.
- [13] T Chan and L Vese. Active contours without edges. *IEEE Trans Image Process*, 2001.
- [14] T Chan and C Wong. Total variation blind deconvolution. *IEEE Trans. Image Process*, 1998.

- [15] J Darbon and M Sigelle. Image restoration with discrete constrained total variation part i: Fast and exact optimization. *J Math Imaging Vis*, 2006.
- [16] E Esser, X Zhang, and T Chan. A general framework for a class of first order primal-dual algorithms for tv minimization. *UCLA CAM Report*, 2009.
- [17] H Fu, M Ng, M Nikolova, and J Barlow. Efficient minimization methods of mixed  $l^2 - l^1$  and  $l^1 - l^1$  norms for image restoration. *SIAM J Sci Comput*, 2006.
- [18] D Goldfarb and W Yin. Second-order cone programming methods for total variation based image restoration. *SIAM J Sci Comput*, 2005.
- [19] T Goldstein and S Osher. The split bregman method for  $l^1$ -regularization problems. *SIAM J Imaging Sci*, 2009.
- [20] M Hintermüller, K Ito, and K Kunisch. The primal-dual active set strategy as a semismooth newton’s method. *SIAM J Optim*, 2003.
- [21] M Hintermüller and G Stadler. A primal-dual algorithm for tv-based inf-convolution-type image restoration. *SIAM J Sci Comput*, 2006.
- [22] Y Huang, M Ng, and Y Wen. A fast total variation minimization method for image restoration. *Multiscale Model Simulat*, 2008.
- [23] D Krishnan, P Lin, and A Yip. A primal-dual active-set method for non-negativity constrained total variation deblurring problems. *IEEE Trans Image Process*, 2007.

- [24] D Krishnan, Q Pham, and A Yip. A primal dual active set algorithm for bilaterally constrained total variation deblurring and piecewise constant mumford-shah segmentation problems. *Adv Comput Math*, 2009.
- [25] CM Maes. A regularized active-set method for sparse convex quadratic programming. *Ph.D. thesis*, 2010.
- [26] S Masnou and J Morel. Level-lines based disocclusion. *Proc. 5th IEEE Int. Conf. on Image Process*, 1998.
- [27] M Ng, L Qi, Y Tang, and Y Huang. On semismooth newton’s methods for total variation minimization. *J Math Imaging Vis*, 2007.
- [28] M Nikolova. Minimizers of cost-functions involving nonsmooth data-fidelity terms. *SIAM J. Numer. Anal.*, 2002.
- [29] S Osher, M Burger, D Goldfarb, J Xu, and W Yin. An iterative regularization method for total variation based image restoration. *Multiscale Model Simulat*, 2005.
- [30] L Rudin, S Osher, and E Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 1992.
- [31] Otmar Scherzer, editor. *Handbook of Mathematical Methods in Imaging*. Springer Science+Business Media, LLC, 2011.
- [32] M. Zhu and T. Chan. An efficient primal-dual hybrid gradient algorithm for total variation image restoration. *UCLA CAM Report*, 2008.