

混合全变差正则化的全波形反演

吴法选, 何清龙*

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2022年7月23日; 录用日期: 2022年8月23日; 发布日期: 2022年9月1日

摘要

全波形反演(FWI)是一个高度非线性和不适定的数学物理反问题。全变差(TV)正则化方法具有保持解的不连续的性质, 然而其对具有一定倾斜角度(如分段线性)的区域会呈现阶梯状伪影。二阶TV (TV2)正则化方法可以减弱阶梯现象, 同时很好地保留反演的边缘信息。但二阶TV正则化方法与传统TV正则化相比, 其需要更多的计算量。因此, 本文基于TV和TV2正则化方法, 提出了混合的全变差正则化方法(HTV)。基于Marmousi2模型和Sigsbee模型进行数值实验, 数值结果表明相对于TV和TV2正则化方法, HTV正则化方法在反演精度方面具有较好的计算表现。

关键词

全变差正则化, 二阶全变差正则化, 混合正则化, 全波形反演

Full Waveform Inversion with Compound Total Variation Regularization

Faxuan Wu, Qinglong He*

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Received: Jul. 23rd, 2022; accepted: Aug. 23rd, 2022; published: Sep. 1st, 2022

Abstract

Full waveform inversion (FWI) is a highly nonlinear and ill-posed mathematical physics inverse problem. Total variation (TV) regularization method has the property of preserving the discontinuity of the solution. However, it leads to the stair-casing artifacts for regions with certain skew angles (e.g., piecewise linearity). The second-order TV (TV2) regularization method can attenuate the staircase phenomenon while preserving the edge information of the inversion resolution well.

*通讯作者。

However, compared with the conventional TV regularization method, the TV2 regularization method requires large computation costs. Therefore, we combine the advantages of TV and TV2 regularization and propose a hybrid regularization (HTV) method. Numerical experiments based on the Marmousi2 model and the Sigsbee model are conducted, and the numerical results show that the HTV regularization method has better computational performance in terms of inversion accuracy compared with the TV and TV2 regularization methods.

Keywords

Total Variation Regularization, Second-Order Total Variation Regularization, Compound Total Variation Regularization, Full Waveform Inversion

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

全波形反演(Full waveform inversion, 简称为 FWI)是一种能有效地利用地震波场运动学和动力学信息的反演方法, 具有高分辨率的成像特性。理论上, 求解 FWI 问题可以转化为求解如下偏微分方程约束的极小化问题[1] [2]

$$\min f(m) = \|Ru(m) - d_{obs}\|_2^2, \quad (1)$$

其中, $\|\cdot\|_2$ 表示 L_2 范数, m 是反演参数(如速度, 密度等), R 代表线性投影算子, 它用于提取接收器位置的波场, d_{obs} 表示观测数据, 合成波场 u 满足方程

$$F(u(m), m) = 0, \quad (2)$$

方程(2)可以是频率域或时间域声波方程、弹性波方程等。从数值计算角度, 本文只考虑 2D 频率域声波方程, 且假设声波方程中密度恒定, 模型参数为压力波速度 $m = v_p$, 因此, 方程(2)即为著名的 Helmholtz 方程

$$F(u(m), u) := \Delta u + \frac{\omega^2}{m^2} u - q_s(x, x_s) = 0, \quad (3)$$

其中, ω 是角频率, v_p 是声波速度, $q_s(x, x_s)$ 是震源项。

FWI 通常是一个不适定问题, 需要引入适当的正则化方法才能获得较好的反演结果。目标函数(1)引入正则化惩罚项变为

$$f(m) = \|Ru(m) - d_{obs}\|_2^2 + \lambda J(m), \quad (4)$$

其中, $J(m)$ 是惩罚项(正则化项), 其具有添加先验信息以限制模型空间的作用以及稳定极小化过程的作用, $\lambda > 0$ 为正则化参数, 其作用是平衡正则化项和数据拟合项, 取值取决于具体的问题。常见的正则化方法有基于 L_2 -norm 的正则化, 如 Tikhonov 正则化[3] [4], 该方法具有获取过度光滑的解的作用; 基于 L_1 -norm 的正则化, 如全变差(TV)正则化[5] [6] [7], 其可以很好地保留解的边缘信息。

TV 正则化具有保留重构解的不连续性信息的优点。针对 FWI 反演问题, 地层介质模型参数(声波速度)往往具有不连续性且在幅度上通常存在较大差异。因此, 基于 TV 正则化的反演方法广泛应用于求解

FWI 反演问题[8] [9] [10]。然而, 从全变差函数的定义可知, **TV** 正则化方法只能提取 x 方向和 z 方向的不连续边界信息, 因此对于具有一定倾斜角度的不连续解往往将其重构为分片常数, 称为著名的阶梯现象[10]。为了克服传统 **TV** 正则化的不足, 基于二阶偏导数信息的二阶 **TV** 正则化方法逐渐受到了学者的关注。二阶正则化方法不仅可以减弱一阶 **TV** 重构结果的阶梯现象, 而且还能很好地保留解的边缘信息[11] [12] [13] [14]。然而, 与传统的 **TV** 正则化方法相比, 在极小化的过程中, 由于二阶 **TV** 正则化方法需要计算更多的一阶导数或二阶导数信息, 导致计算量急剧增加[15]。本文结合一阶 **TV** 正则化和二阶 **TV** 正则化方法各自的优点, 提出了混合形式的正则化方法(**HTV**)。给出了二阶 **TV** 正则化方法的梯度计算方法, 并基于 2D 的 **Marmousi2** 模型和 **Sigsbee** 模型进行数值实验, 验证 **HTV** 方法的有效性。

2. 方法

2.1. 一阶 TV 正则化方法

TV 正则化(一阶 **TV**)能够有效提高反演的稳定性和重构的分辨率, 且可以很好地保留解的边缘信息。1992 年, **Rudin** 首次提出了全变差正则化方法[6]

$$J_v = \int_{\Omega} |\nabla m| dx dz, \quad (5)$$

其中, Ω 表示积分区域, $\nabla m = (\partial m / \partial x, \partial m / \partial z)^T$, $\partial m / \partial x, \partial m / \partial z$ 分别表示模型参数关于 x -和 z -方向的偏导数。在(5)式中, 由于函数 $g(x) = |x|$ 在原点处不可导, 因此通常利用函数 $h(x) = \sqrt{x^2 + \beta_v}$ 近似 $g(x)$, β_v 表示光滑参数, 通常取固定值, 本研究中取为 $\beta_v = 1.0 \times 10^{-4}$ 。于是全变分正则化惩罚项变为

$$J_v = \int_{\Omega} \sqrt{\left(\frac{\partial m}{\partial x}\right)^2 + \left(\frac{\partial m}{\partial z}\right)^2 + \beta_v} dx dz, \quad (6)$$

为了对 **TV** 惩罚项进行离散, 并引入记号

$$m_{i,j}^x = \frac{m_{i+1,j} - m_{i,j}}{\Delta x}, m_{i,j}^z = \frac{m_{i,j+1} - m_{i,j}}{\Delta z}, \quad (7)$$

其中, Δx 和 Δz 分别表示 x -和 z -方向的空间离散步长。于是, 一阶 **TV** 正则化项的向前差分离散形式为

$$J_v = \sum_{i=1}^{N_x-1} \sum_{j=1}^{N_z-1} \sqrt{\left(m_{i,j}^x\right)^2 + \left(m_{i,j}^z\right)^2 + \beta_v} \Delta x \Delta z, \quad (8)$$

其中, N_x, N_z 分别表示 x -和 z -方向的空间采样点数量。

由于一阶 **TV** 惩罚项具有保持解不连续的性质, 其在图像处理 and **FWI** 中得到了广泛的应用。然而, 对于具有一定倾角几何特征的重构问题, 一阶 **TV** 正则化方法常常产生阶梯现象。

2.2. 二阶 TV 正则化方法

基于二阶导数信息的二阶 **TV** 惩罚项(**TV2**)具有获取参数的某些几何信息(不连续性、尖点等)的特性, 因此, 能够提高重构结果的分辨率。**TV2** 惩罚项定义为[12] [15]

$$J_{v2}(m) = \int_{\Omega} \sqrt{\left(\frac{\partial^2 m}{\partial x^2}\right)^2 + \beta_{v2}} + \sqrt{\left(\frac{\partial^2 m}{\partial z^2}\right)^2 + \beta_{v2}} dx dz, \quad (9)$$

其中, Ω 表示积分区域, $\partial^2 m / \partial x^2, \partial^2 m / \partial z^2$ 分别表示模型参数关于 x -和 z -方向二阶偏导数, β_{v2} 表示光滑参数, 取为 $\beta_{v2} = 1.0 \times 10^{-4}$ 。

类似地, 为了对 TV2 惩罚项进行离散, 引入如下变量

$$m_{i,j}^{xx} = \frac{m_{i+1,j} - 2m_{i,j} + m_{i-1,j}}{\Delta x}, \quad (10)$$

$$m_{i,j}^{zz} = \frac{m_{i,j+1} - 2m_{i,j} + m_{i,j-1}}{\Delta z}, \quad (11)$$

其中, $\Delta x, \Delta z$ ($\Delta x = \Delta z$) 分别表示 x -和 z -方向的空间离散步长。于是 TV2 正则化惩罚项的中心差分离散格式为

$$J_{tv2} = \sum_{i=1}^{N_x-1} \sum_{j=1}^{N_z-1} \sqrt{(m_{i,j}^{xx})^2 + \beta_{tv2}} + \sqrt{(m_{i,j}^{zz})^2 + \beta_{tv2}} \Delta x \Delta z, \quad (12)$$

在数值计算中, TV2 能够很好地保留几何上具有一定倾角的解的不连续信息; 然而, 相对于 TV 正则化方法, 由于 TV2 的高阶导数的高阶光滑特性使其对水平和垂直方向上的不连续信息具有光滑性, 且 TV2 会引入额外的计算量, 不利用于求解大规模反演问题。

2.3. 混合正则化方法

为了进一步提高算法的重构分辨率和计算效率, 基于 TV 和 TV2 正则化方法各自的特性, 本文提出了混合 TV 正则化方法和 TV2 正则化方法的正则化方法 (HTV)。为了调节不同方向上的不连续性信息的权重, 采用 λ 和 μ 两个正则化参数, 以控制 TV 和 TV2 在目标泛函中的权重。则 HTV 正则化的 FWI 目标函数为

$$f_{hiv}(m) = \|Ru(m) - d_{obs}\|_2^2 + \lambda J_{tv}(m) + \mu J_{tv2}(m), \quad (13)$$

需要注意的是, 当 λ 为 0 时, (12)式变为 TV2 正则化方法的目标函数; 当 μ 为 0 时, (12)式为 TV 正则化方法的目标函数。

3. 目标函数的优化方法及梯度计算

3.1. 目标函数优化方法

在极小化目标函数(12)时, 常见的优化方法有 Newton 型方法(如 Gauss Newton, Truncated Newton 等)和梯度型方法(如最速下降法, 共轭梯度法)等[16] [17] [18]。一般地, Newton 型方法需要计算 Hessian 矩阵及其逆矩阵, 需要大量的计算量和存储量。而梯度型算法, 尤其是非线性共轭梯度算法(CG), 其具有迭代格式简单、储存量小等优点, 因此常用于求解大规模优化问题。

对于 FWI 反演问题, 采用 CG 方法的参数更新格式为

$$m_{k+1} = m_k + \alpha_k d_k \quad (14)$$

其中, α_k 为步长, d_k 表示下降方向, d_k 的迭代格式为

$$d_0 = -g_0, d_k = -g_k + \beta_{k-1} d_{k-1}, \quad (15)$$

其中, $g_k = \nabla f_{hiv}(m_k)$ 表示 $f_{hiv}(m)$ 在 m_k 处的梯度, β_{k-1} 是共轭系数。当利用 CG 方法求解大规模优化问题时, 当前迭代步的梯度 g_k 与之前的下降方向 $\{d_i\}, i = 0, 1, \dots, k-1$ 满足正交性时, 才能保证 CG 方法的快速收敛性。然而, 针对大规模高度非线性的 FWI 问题, 该正交性往往很难得到保证。本文采用一种高效且带预处理的 CG 方法(CG-DESCENT)极小化目标函数(12)式[16]。

当正交性不满足时, CG-DESCENT 方法通过在一个子空间上求解一个子优化问题对该正交性进行恢复, 因此, 其具有较高的收敛速度和计算效率。因此, 带预处理的 CG-DESCENT 方法由三个不同的迭

代组成: 一个是标准 CG 迭代, 其共轭系数 β_k 为:

$$\beta_k = \left(y_k - \eta_k \frac{d_k \|y_k\|^2}{d_k^T y_k} \right) \frac{g_{k+1}}{d_k^T y_k}, \quad (16)$$

其中, $y_k = g_{k+1} - g_k$, $\eta_k > \frac{1}{4}$, 在理论上, $\eta_k = 1$ 是最佳选择; 第二个是子空间迭代, 当正交性失去时, 其通过 L-BFGS 方法在子空间上求解一个子优化问题来恢复正交性以提高算法的计算效率; 第三个是预处理, 当子空间迭代结束回到原空间时, 将子空间得到的 Hessian 矩阵作为预处理矩阵来加速算法的收敛速度(见参考文献[16])。

3.2. 梯度计算

利用 CG-DESCENT 求解 FWI 反演问题的一个关键因素时目标函数梯度的计算。首先考虑数据误差项的梯度计算, 对(1)式中 $f(m)$ 的两端关于未知参数 m 进行微分, 并利用微分的链式规则得

$$\begin{aligned} \nabla f(m) &= \left(\frac{\partial Ru(m)}{\partial m} \right)^* (Ru(m) - d_{obs}) \\ &= \left(\frac{\partial u(m)}{\partial m} \right)^* R^* (Ru(m) - d_{obs}), \end{aligned} \quad (17)$$

在上式中, $*$ 表示随算子的伴随, $\partial u(m)/\partial m$ 是 Jacobian 矩阵。若对(17)式采用有限差分法直接计算 Jacobian 矩阵, 则需要计算同参数数量相同的波场模拟, 而这对于大规模 FWI 反演问题是不现实的, 且随最计算规模的增加也不允许保存该大规模的 Jacobian 矩阵。为了避免计算 Jacobian 矩阵, 本文采用一阶伴随方法计算数据残差项的梯度[17][18]

$$\nabla f(m) = \left(\frac{\partial F(u(m), m)}{\partial m} \right)^* u^\dagger, \quad (18)$$

其中, u^\dagger 满足一阶伴随方程, 即

$$\left(\frac{\partial F(u(m), m)}{\partial u(m)} \right)^* u^\dagger = -R^* (Ru(m) - d_{obs}). \quad (19)$$

从(18)和(19)式可以看出, 计算数据残差项的梯度只需要求解两个正演问题, 一个是计算正演波场 u , 另一个是计算伴随波场 u^\dagger 。因此, 利用伴随方法计算梯度是非常有效的。本文在频率域求解 FWI 问题时, 每次迭代只取数据残差项梯度的实部。

为了求解 TV 的梯度, 根据式(7)和(8), TV 关于模型参数 $m_{i,j}$ 的梯度表示为

$$\begin{aligned} \frac{\partial J_{tv}(m)}{\partial m_{i,j}} &= \frac{-m_{x1} - m_{z1}}{\sqrt{(m_{x1})^2 + (m_{z1})^2 + \beta_{tv}^2}} \\ &+ \frac{m_{x2}}{\sqrt{(m_{x2})^2 + (m_{z2})^2 + \beta_{tv}^2}} \\ &+ \frac{m_{x3}}{\sqrt{(m_{x3})^2 + (m_{z3})^2 + \beta_{tv}^2}}, \end{aligned} \quad (20)$$

其中,

$$\begin{aligned}
 m_{x1} &= \frac{m_{i+1,j} - m_{i,j}}{\Delta x}, m_{z1} = \frac{m_{i,j+1} - m_{i,j}}{\Delta z}, \\
 m_{x2} &= \frac{m_{i,j} - m_{i-1,j}}{\Delta x}, m_{z2} = \frac{m_{i-1,j+1} - m_{i-1,j}}{\Delta z}, \\
 m_{x3} &= \frac{m_{i+1,j-1} - m_{i,j-1}}{\Delta x}, m_{z3} = \frac{m_{i,j} - m_{i,j-1}}{\Delta z}.
 \end{aligned} \tag{21}$$

由式(10), (11)及(12), TV2 关于模型参数 $m_{i,j}$ 的梯度表示为

$$\begin{aligned}
 \frac{\partial J_{tv2}(m)}{\partial m_{i,j}} &= \frac{-2m_{xx1}}{\sqrt{(m_{xx1})^2 + \beta_{tv2}^2}} + \frac{-2m_{zz1}}{\sqrt{(m_{zz1})^2 + \beta_{tv2}^2}} \\
 &+ \frac{m_{xx2}}{\sqrt{(m_{xx2})^2 + \beta_{tv2}^2}} + \frac{m_{xx3}}{\sqrt{(m_{xx3})^2 + \beta_{tv2}^2}} \\
 &+ \frac{m_{zz2}}{\sqrt{(m_{zz2})^2 + \beta_{tv2}^2}} + \frac{m_{zz3}}{\sqrt{(m_{zz3})^2 + \beta_{tv2}^2}},
 \end{aligned} \tag{22}$$

其中,

$$\begin{aligned}
 m_{xx1} &= \frac{m_{i+1,j} - 2m_{i,j} + m_{i-1,j}}{(\Delta x)^2}, m_{zz1} = \frac{m_{i,j+1} - 2m_{i,j} + m_{i,j-1}}{(\Delta z)^2}, \\
 m_{xx2} &= \frac{m_{i,j} - 2m_{i-1,j} + m_{i-2,j}}{(\Delta x)^2}, m_{xx3} = \frac{m_{i+2,j} - 2m_{i+1,j} + m_{i,j}}{(\Delta x)^2}, \\
 m_{zz2} &= \frac{m_{i,j} - 2m_{i,j-1} + m_{i,j-2}}{(\Delta z)^2}, m_{zz3} = \frac{m_{i,j+2} - 2m_{i,j+1} + m_{i,j}}{(\Delta z)^2}.
 \end{aligned} \tag{23}$$

式(23)适用于所有的内点, 即对 $2 \leq i \leq N_x - 2$ 和 $2 \leq j \leq N_z - 2$, 空间采样点的边界为 $m_{i,0} = 0$, $m_{0,j} = 0$, $m_{N_x-1,j} = 0$, $m_{i,N_z-1} = 0$ 。

4. 数值实验

为了数值求解波动方程(3), 首先采用混合网格有限差分法进行离散, 形成线性方程组[19]。在离散过程中, 引入完全匹配层(PML)吸收边界条件以减少来自边界反射波场对波场的影响[20]。基于 MPI 和 OpenMP 的混合并行框架求解该大规模稀疏线性方程组[21]。

正则化参数控制目标函数的数据拟合项和正则化项之间的平衡。因此, 正则化参数控制正则化解的特性的一个重要量, 应谨慎选取这些参数。正则化参数越大, 对应的正则化项占总目标函数比例越大, 在相同优化方法和相同迭代步的条件下, 所需要的计算量越大。在反演过程中, 正则化参数 λ, μ 取为固定值, 对于 TV 和 TV2 正则化方法, 取 $\lambda = 1.0 \times 10^{-4}$, $\mu = 1.0 \times 10^{-4}$, 对于 HTV 正则化方法, 取 $\lambda = 1.0 \times 10^{-4}$, $\mu = 1.0 \times 10^{-5}$, 该正则化参数通过实验获得, 即根据 TV 和 TV2 反演结果选取该参数。基于 Marmousi2 模型和 Sigsbee 模型进行数值实验以验证 HTV 正则化方法的有效性, 并分别与 TV 和 TV2 正则化方法进行对比以验证 HTV 的高效性。为了定量评估方法的计算表现, 引入峰值信噪比指数(PSNR)指标[22]

$$\text{PSNR}(m^t, m^r) = 20.0 \times \log_{10} \left(\frac{\text{MAX}(m^t)}{\sqrt{\text{MSE}(m^t, m^r)}} \right), \tag{24}$$

其中,

$$\text{MAX}(m') = \max_{i,j}(m'_{i,j}), \quad (25)$$

$$\text{MSE}(m', m^r) = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (m'_{i,j}, m^r_{i,j}), \quad (26)$$

其中, m' 和 m^r 分别表示真实模型和重构模型。

4.1. Marmousi2 模型

在地震成像问题中, Marmousi2 模型是用来测试地震成像方法的基准模型之一。原始 Marmousi2 模型的模型参数为: 深 3.5 km, 宽 17 km, 模型由 199 个地质层和一个 450 m 的深水层构成。为了减弱 FWI 问题的非线性对算法的影响, 在本文的实验中保持上表面的水层参数不更新。

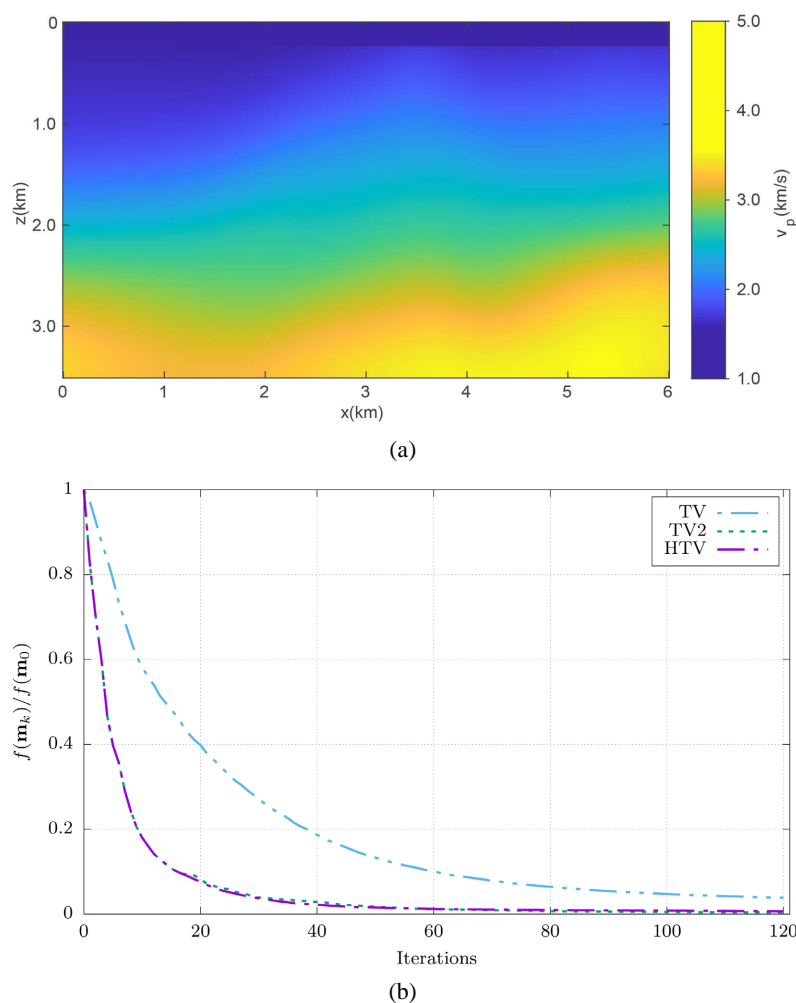


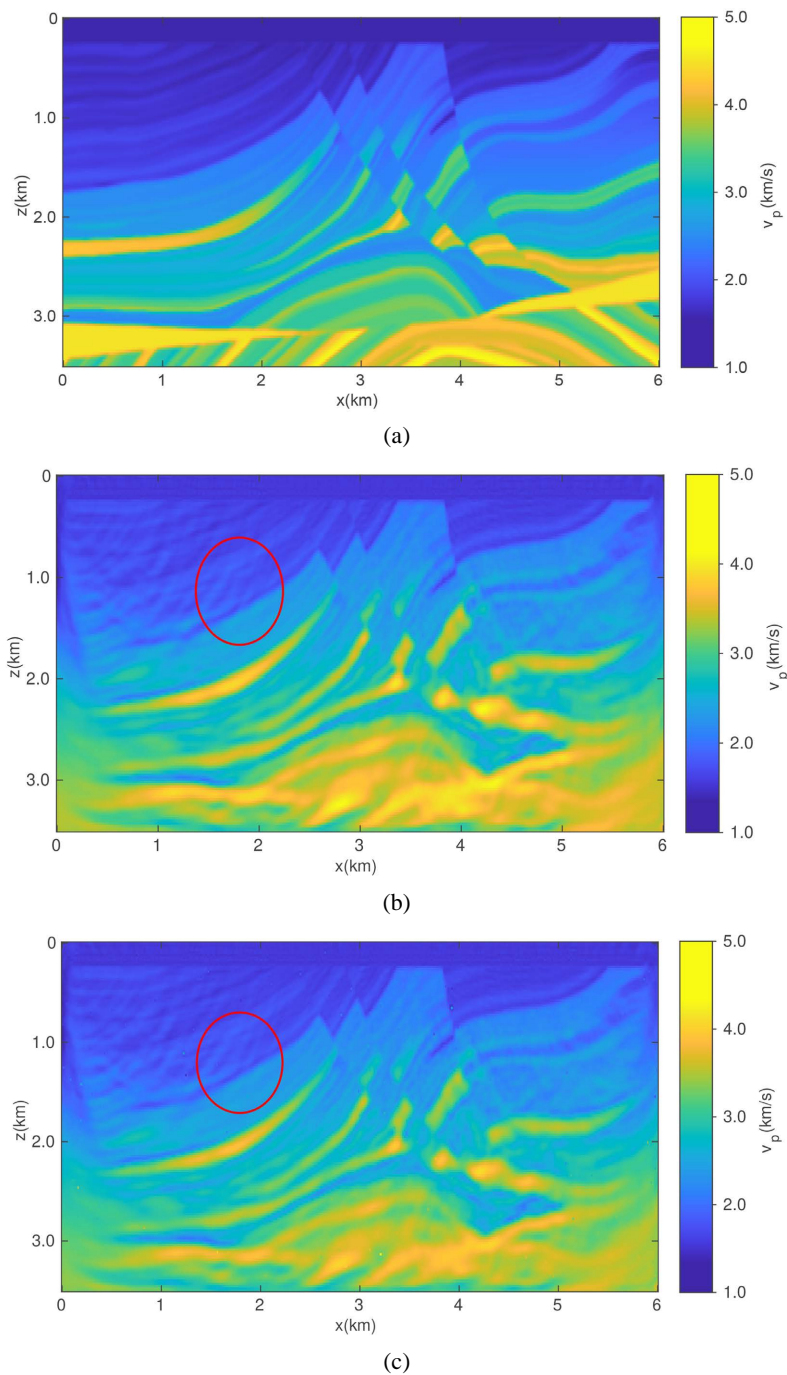
Figure 1. (a) The initial guess of the Marmousi2 model; (b) The convergence curve for three regularization methods

图 1. (a) Marmousi2 初始速度模型; (b) 三种正则化方法的收敛曲线

从数值计算角度考虑, 在本实验中对原始模型进行重采样获得反演的真实模型。重采样的 Marmousi2 模型的模型大小为 481×146 , 空间采样间隔为 $\Delta x = \Delta z = 20$ m, 在反演区域四周均设置 10 个 PML 网格。因此, 总的计算规模为 $N_x \times N_z = 501 \times 166$ 。图 1(a)所示为重采样的 Marmousi2 真实模型, 图 2(a)为相应

的初始模型, 该初始模型由高斯滤波器对真实模型进行平滑化获得。

在反演过程中, 115 个震源和 460 个接收器均匀分布在模型的上表面。震源和接收器的间隔分别为 80 m 和 20 m。反演过程中所使用的频率为 3.0 Hz, 4.5 Hz, 6.0 Hz, 7.5 Hz, 9.0 Hz, 10.5 Hz 和 12.0 Hz。对不同的正则化方法, 均采用从低频到高频的逐次反演策略。每个频率的最大迭代次数设置为 120 次。如图 1(b)所示为频率为 3.0 Hz 时, TV、TV2 和 HTV 正则化方法的收敛曲线。图 1(b)表明三种正则化方法均收敛, 且相对于 TV 正则化方法, TV2 和 HTV 的收敛速度明显快于 TV 正则化方法; TV2 和 HTV 的收敛速度相当。



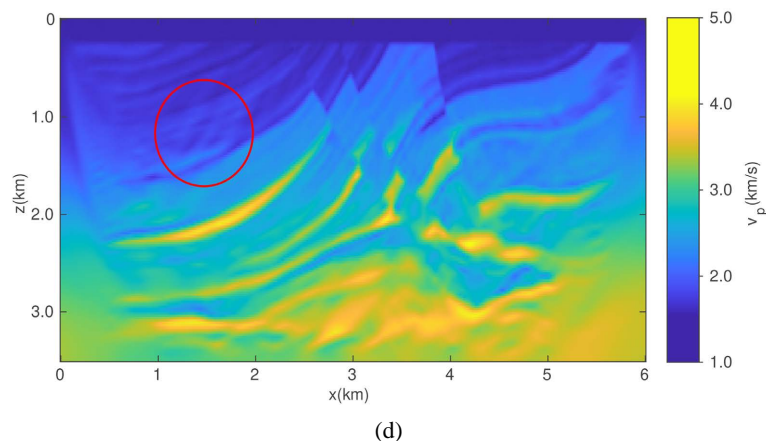


Figure 2. (a) The true velocity of the Marmousi2 model; (b) The reconstructed results of TV method; (c) The reconstructed results of TV2 method; and (d) The reconstructed results of HTV method

图 2. (a) Marmousi2 真实速度模型; (b) 表示 TV 方法的反演结果; (c) 表示 TV2 方法的反演结果; (d) 表示 HTV 方法的最终反演结果

图 2 所示为三种方法每个频率进行 120 次迭代后的最终的重构结果。图 2(b)为 TV 正则化方法的反演结果, 图 2(c)为 TV2 正则化方法的反演结果, 图 2(d)为 HTV 正则化方法的反演结果。从图 2 可以看出, 三种正则化方法均可较好地重构 Marmousi2 模型。从图 2 中还可以看出, 当不连续边界较小且有一定的倾斜时, TV 正则化方法的重构结果为阶梯形状, 即产生与真实结果不相符合的虚假边界。TV2 和 HTV 能够捕捉到斜面区域的不连续性, 对具有倾斜的部分的重构表现比 TV 方法更好(如图 2 中椭圆区域)。

图 3 所示为模型在不同水平位置(3.84 km, 5.84 km 和 7.0 km)处的垂直剖面。图 3 再次说明了三种正则化方法均能较好地重构 Marmousi2 模型。从图 2 和图 3 中可以看出每种反演方法的重构分辨率随着深度的增加而降低, 该现象是由于震源的有限照明所引起的。

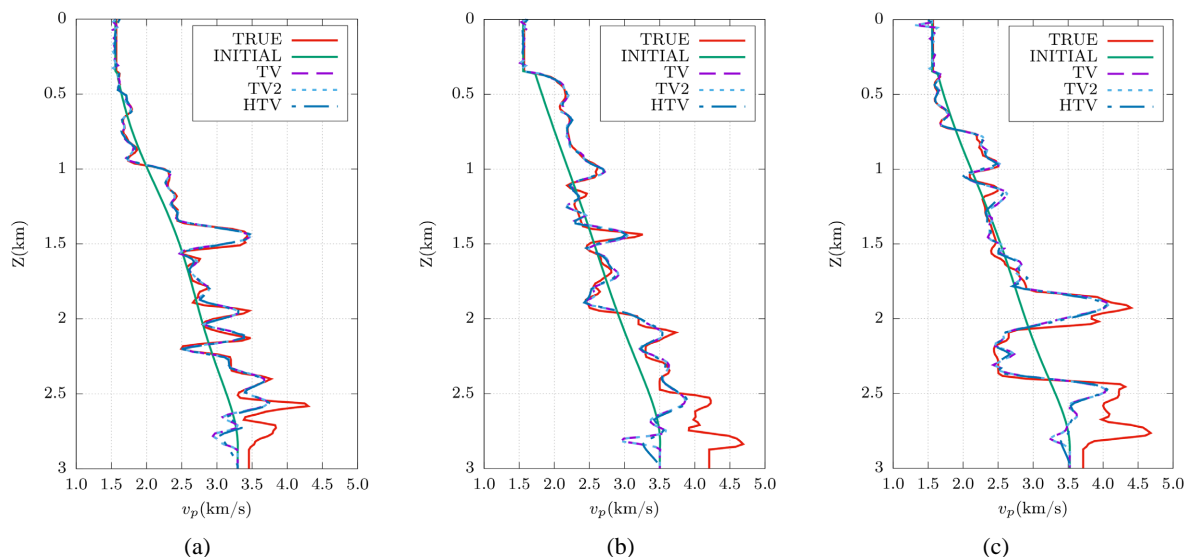


Figure 3. Vertical profiles for Marmousi2 model, TRUE-true model, INITIAL-background model, TV-TV regularization method, TV2-TV2 regularization method, HTV-hybridize TV and TV2: (a) at horizontal position of 3.84 km; (b) at horizontal position of 5.84 km; (c) at horizontal position of 7.0 km

图 3. Marmousi2 模型的垂直剖面图: TRUE-真实模型、INITIAL-初始模型、TV-TV 正则化、TV2-TV2 正则化、HTV-混合 TV 和 TV2。(a) 在 3.84 km 的水平位置, (b) 在 5.84 km 的水平位置, (c) 在 7.0 km 的水平位置

表 1 为三种正则化方法的目标函数计算次数(NF), 梯度计算次数(NG), 峰值信噪比(PSNR)以及 $f(m_k)/f(m_0)$ 的统计量对比。从表 1 可以看出, TV2 和 HTV 在峰值信噪比指数方面要优于 TV 正则化方法, HTV 正则化方法在目标函数估计次数与梯度计算次数方面优于 TV2 正则化方法, 说明了混合 HTV 在在计算效率上优于 TV2 方法。

Table 1. Comparison between TV, TV2 and HTV regularization methods in the terms of the number of function estimations (NF), the number of gradient estimations (NG), and the peak signal-to-noise ratio (PSNR)

表 1. TV, TV2 和 MTV 正则化方法在目标函数估计次数(NF), 梯度估计次数(NG)以及峰值信噪比(PSNR)的对比

正则化方法	NF	NG	PSNR
TV	1833	616	21.01
TV2	4590	1811	21.21
HTV	2105	750	21.52

4.2. Sigsbee 模型

Sigsbee 模型顶部构造分界面存在着强烈的速度间断, 随着深度的加深以及背景介质速度的增加, 会产生多散射场问题, 给问题的求解提出了挑战。本节利用重采样的 Sigsbee 模型, 其大小为 391×144 , 空间采样间隔为 $\Delta x = \Delta z = 15 \text{ m}$ 。为了减弱边界波场反射对重构结果的影响, 在矩形区域边界添加 10 个网格的 PML 吸收层。因此, 计算区域网格大小为 411×166 。图 4(a)所示为对真实模型进行 Gaussian 光滑化所得到的初始模型, 重采样的 Sigsbee 真实模型如图 5(a)所示。

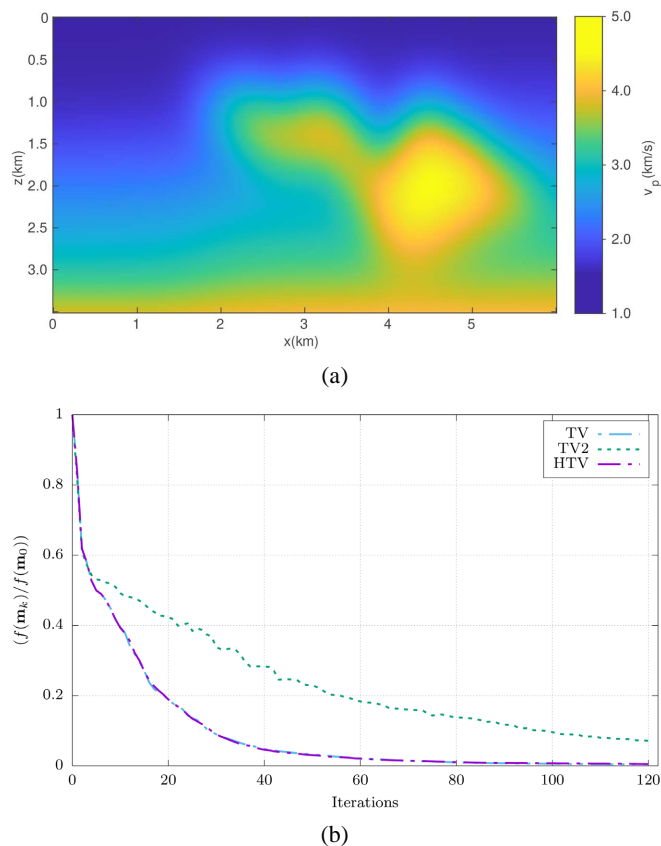
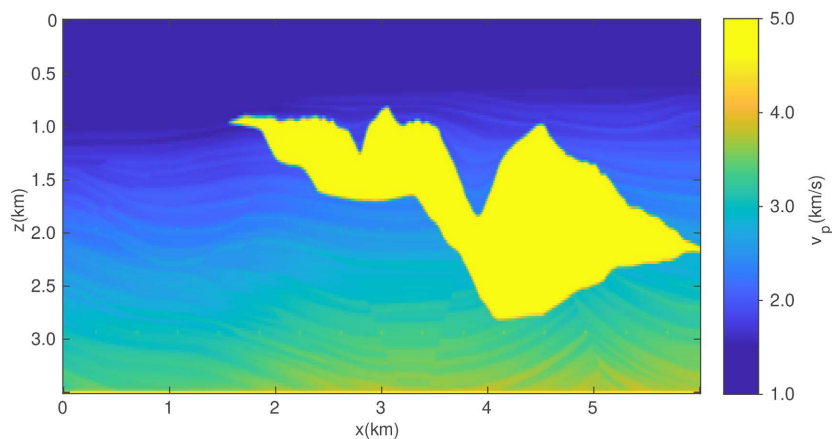
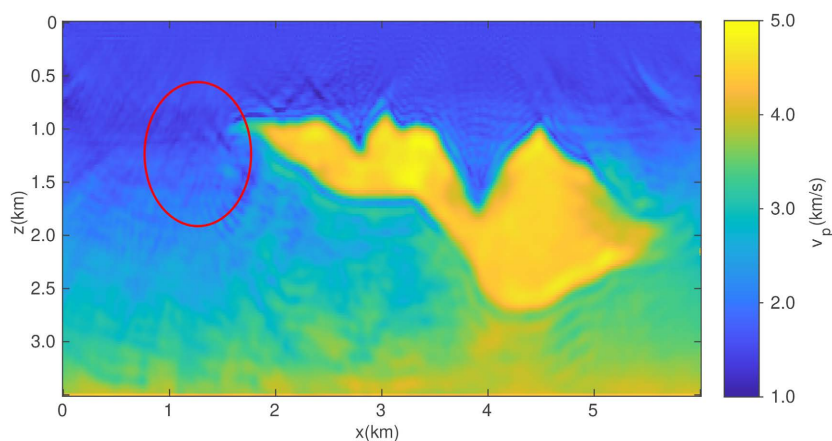


Figure 4. (a) The initial guess of the Sigsbee model; (b) The convergence curve for three regularization methods
图 4. (a) Sigsbee 初始速度模型; (b) 三种正则化方法的收敛曲线

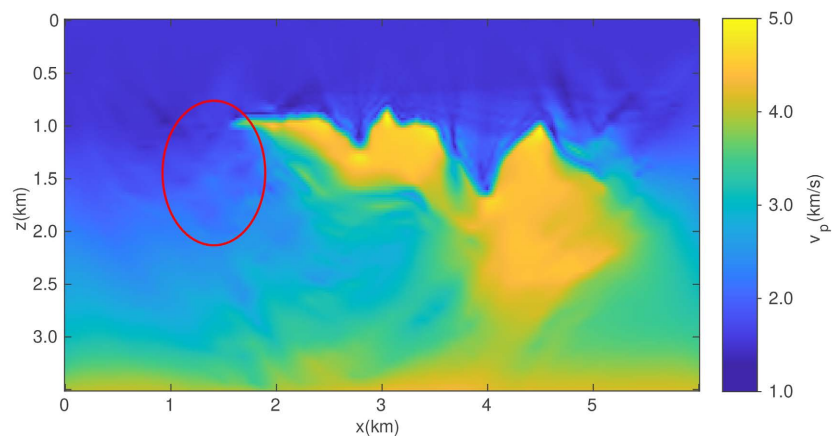
在反演过程中, 地表观测系统由 97 个震源和 390 个接收器组成, 震源和接收器均匀分布在计算区域的上表面。21 个离散频率从 2.5 Hz 到 17.0 Hz 进行采样, 采样间隔为 0.725 Hz。本实验也采用从低频到高频的反演策略逐次反演, 每个频率的最大迭代步数为 120 步。如图 4(b)所示为频率为 2.5 Hz 时, 三种正则化方法的目标函数 $f(m_k)/f(m_0)$ 相对于迭代步的变化曲线。图 4(b)表明对 Sigsbee 模型, 三种正则化方法均收敛, 且 TV 正则化方法的收敛速度与 HTV 正则化方法的收敛曲线基本一致。



(a)



(b)



(c)

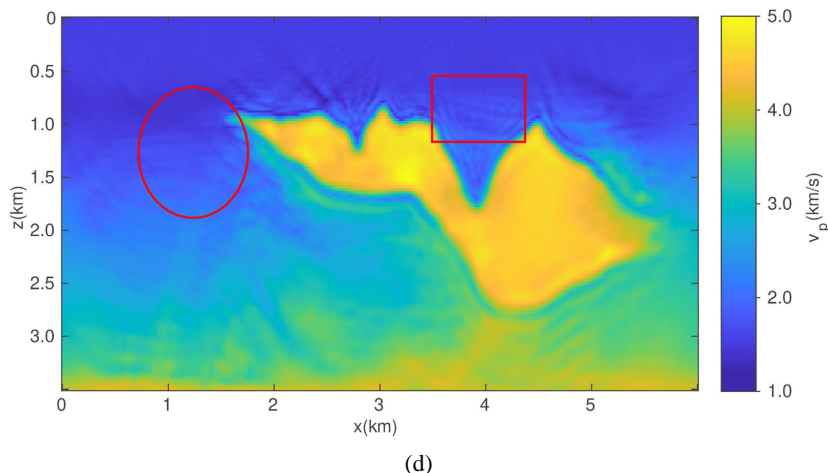


Figure 5. (a) The true velocity of the Sigsbee model; (b) The reconstructed results of TV method; (c) The reconstructed results of TV2 method; and (d) The reconstructed results of HTV method

图 5. (a) Sigsbee 真实速度模型; (b) 表示 TV 方法的反演结果; (c) 表示 TV2 方法的反演结果; (d) 表示 HTV 方法的最终反演结果

图 5(b)~(d)所示分别为 TV, TV2 和 HTV 正则化方法的最终反演结果。图 6(a)为速度模型在 $x = 2.49$ km 处的垂直剖面图, 图 6(b)为速度模型在 $x = 2.88$ km 处的垂直剖面图, 图 6(c)为速度模型在 $x = 4.5$ km 处的垂直剖面图。图 5 和图 6 说明三种正则化方法均可较好地重构 Sigsbee 模型。其次, TV2 正则化方法相比于 TV 方法, 在浅层的反演精度更高(如图 5 中椭圆区域所示)。与 TV 和 TV2 相比, HTV 正则化方法能够准确的反演出高速比部分的 Sigsbee 模型(如图 5 中椭圆和矩形区域所示), 说明 HTV 正则化方法是有效的。

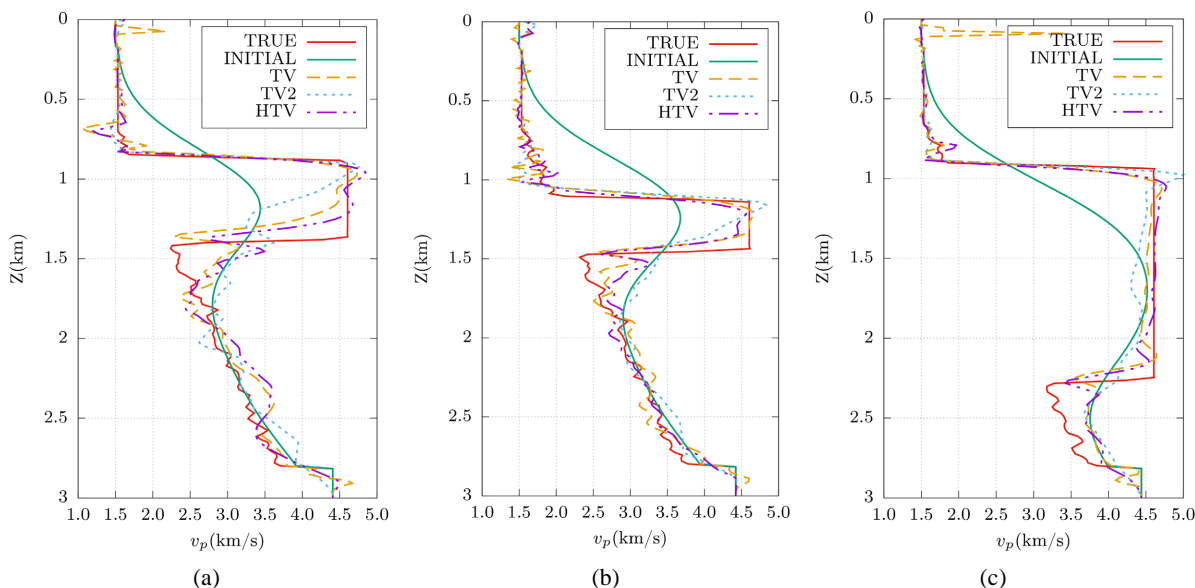


Figure 6. Vertical profiles for Sigsbee model: TRUE-true model, INITIAL-background model, TV-TV regularization method, TV2-TV2 regularization method, HTV-hybridize TV and TV2: (a) at horizontal position of 3.84 km; (b) at horizontal position of 5.84 km; (c) at horizontal position of 7.0 km

图 6. Sigsbee 模型的垂直剖面: TRUE-真实模型、INITIAL-初始模型、TV-TV 正则化、TV2-TV2 正则化、HTV-混合 TV 和 TV2。(a) 在 3.84 km 的水平位置; (b) 在 5.84 km 的水平位置; (c) 在 7.0 km 的水平位置

为了说明 HTV 正则化方法在计算效率和重构分辨率方面的优势, 对三种正则化方法的目标函数计算次数(NF), 梯度计算次数(NG), 峰值信噪比(PSNR)进行对比(见表 2)。从表 2 中可以看出, 相对于经典的 TV 和 TV2 正则化方法, HTV 正则化方法具有较高的计算效率和重构精度。

Table 2. Comparison between TV, TV2 and HTV regularization methods in the terms of the number of function estimations (NF), the number of gradient estimations (NG), and the peak signal-to-noise ratio (PSNR)

表 2. TV, TV2 和 MTV 正则化方法在目标函数估计次数(NF), 梯度估计次数(NG)以及峰值信噪比(PSNR)的对比

正则化方法	NF	NG	PSNR
TV	2433	906	22.02
TV2	2657	1018	22.67
HTV	2280	838	23.04

5. 结论与展望

全波形反演是一个大规模不适定问题, 需要引入正则化方法以克服 FWI 问题的不适定性。本文基于一阶和二阶 TV 正则化方法, 提出了一种混合的正则化方法(HTV)。HTV 正则化方法能够克服传统 TV 正则化方法的阶梯现象, 从而提高反演成像的分辨率。此外, 引入了共轭梯度方法(CG-DESCENT)求解本文提出的正则化 FWI 问题。该算法可以恢复梯度和先前下降方向之间的正交性, 提高算法的计算效率。基于 Marmousi2 模型和 Sigsbee 模型, 从数值角度验证了 HTV 正则化方法的收敛性和高效性。数值结果表明, 相对于 TV 和 TV2 正则化方法, HTV 方法能够保留分段常数区域和倾斜区域的边界信息。

尽管数值实验表明, HTV 正则化 FWI 方法相较于传统的方法重构精度要高, 但仍有一些问题需要进一步研究:

1) 本文的正则化参数 λ 和 μ 设置为常数, 并且效果良好。但是为了保持数据残差项和正则化项之间的平衡, 正则化参数需要随着迭代次数的增减而动态变化。

2) 由于震源的有限照明, 重构分辨率随着深度的增加而降低。然而这种现象可以通过使用适当的预处理或高阶优化方法来缓解。

因此, 未来工作将系统地研究正则化参数的选取方法同时减少迭代的次数。此外, 研究高阶优化方法(如 Truncated-Newton, Gauss-Newton 方法)和预处理方法来实现混合 TV 正则化 FWI, 也将是下一步的工作。

基金项目

本文获得国家自然科学基金(资助号: 11801111)、贵州省科技计划项目([2019]1122)。

参考文献

- [1] Tarantola, A. and Valette, B. (1982) Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion. *Reviews of Geophysics*, **20**, 219-232. <https://doi.org/10.1029/RG020i002p00219>
- [2] Tarantola, A. (1984) Inversion of Seismic Reflection Data in the Acoustic Approximation. *Geophysics*, **49**, 1259-1266. <https://doi.org/10.1190/1.1441754>
- [3] Tikhonov, A.N. (1963) Regularization of Incorrectly Posed Problems. *Soviet Mathematics Doklady*, **4**, 1624-1627.
- [4] Aghamiry, H.S., Gholami, A. and Operto, S. (2019) Compound Regularization of Full-Waveform Inversion for Imaging Piecewise Media. *IEEE Transactions on Geoscience and Remote Sensing*, **58**, 1192-1204. <https://doi.org/10.1109/TGRS.2019.2944464>
- [5] Anagaw, A.Y. and Sacchi, M.D. (2011) Full Waveform Inversion with Total Variation Regularization. *Recovery-CSPG CSEG CWLS Convention 2011*, Alberta, 9-11 May 2011, 1-4.

-
- [6] Royden, H.L. and Fitzpatrick, P. (1988) Real Analysis. Vol. 32, Macmillan, New York.
- [7] Aghamiry, H.S., Gholami, A. and Operto, S. (2019) Implementing Bound Constraints and Total-Variation Regularization in Extended Full-Waveform Inversion with the Alternating Direction Method of Multiplier: Application to Large Contrast Media. *Geophysical Journal International*, **218**, 855-872. <https://doi.org/10.1093/gji/ggz189>
- [8] Li, J., Bai, L. and Liu, H. (2021) Numerical Verification of Full Waveform Inversion for the Chang'E-5 Lunar Regolith Penetrating Array Radar. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1-10. <https://doi.org/10.1109/TGRS.2021.3098104>
- [9] Anagaw, A.Y. and Sacchi, M.D. (2012) Edge-Preserving Seismic Imaging Using the Total Variation Method. *Journal of Geophysics and Engineering*, **9**, 138-146. <https://doi.org/10.1088/1742-2132/9/2/138>
- [10] Anagaw, A. and Sacchi, M.D. (2022) A Regularization by Denoising (RED) Scheme for 3-D FWI Model Updates in Large-Contrast Media. *Geophysical Journal International*, **229**, 814-827. <https://doi.org/10.1093/gji/ggab505>
- [11] Du, Z., Liu, D., Wu, G., Cai, J., Yu, X. and Hu, G. (2021) A High-Order Total-Variation Regularisation Method for Full-Waveform Inversion. *Journal of Geophysics and Engineering*, **18**, 241-252. <https://doi.org/10.1093/jge/gxab010>
- [12] Steidl, G. (2006) A Note on the Dual Treatment of Higher-Order Regularization Functionals. *Computing*, **76**, 135-148. <https://doi.org/10.1007/s00607-005-0129-z>
- [13] Bredies, K. and Holler, M. (2020) Higher-Order Total Variation Approaches and Generalisations. *Inverse Problems*, **36**, Article ID: 123001. <https://doi.org/10.1088/1361-6420/ab8f80>
- [14] Lv, X.G., Song, Y.Z., Wang, S.X. and Le, J. (2013) Image Restoration with a High-Order Total Variation Minimization Method. *Applied Mathematical Modelling*, **37**, 8210-8224. <https://doi.org/10.1016/j.apm.2013.03.028>
- [15] Yuan, J., Schnörr, C. and Steidl, G. (2009) Total-Variation Based Piecewise Affine Regularization. *International Conference on Scale Space and Variational Methods in Computer Vision*, Voss, 1-5 June 2009, 552-564. https://doi.org/10.1007/978-3-642-02256-2_46
- [16] Hager, W.W. and Zhang, H. (2013) The Limited Memory Conjugate Gradient Method. *SIAM Journal on Optimization*, **23**, 2150-2168. <https://doi.org/10.1137/120898097>
- [17] He, Q. and Wang, Y. (2020) Inexact Newton-Type Methods Based on Lanczos Orthonormal Method and Application for Full Waveform Inversion. *Inverse Problems*, **36**, Article ID: 115007. <https://doi.org/10.1088/1361-6420/abb8ea>
- [18] Yan, X., He, Q. and Wang, Y. (2022) Truncated Trust Region Method for Nonlinear Inverse Problems and Application in Full-Waveform Inversion. *Journal of Computational and Applied Mathematics*, **404**, Article ID: 113896. <https://doi.org/10.1016/j.cam.2021.113896>
- [19] Hustedt, B., Operto, S. and Virieux, J. (2004) Mixed-Grid and Staggered-Grid Finite-Difference Methods for Frequency-Domain Acoustic Wave Modelling. *Geophysical Journal International*, **157**, 1269-1296. <https://doi.org/10.1111/j.1365-246X.2004.02289.x>
- [20] Berenger, J.P. (1994) A Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *Journal of Computational Physics*, **114**, 185-200. <https://doi.org/10.1006/jcph.1994.1159>
- [21] Ghysels, P., Li, X.S., Rouet, F.H., Williams, S. and Napov, A. (2016) An Efficient Multicore Implementation of a Novel HSS-Structured Multifrontal Solver Using Randomized Sampling. *SIAM Journal on Scientific Computing*, **38**, S358-S384. <https://doi.org/10.1137/15M1010117>
- [22] Yoo, J.C. and Ahn, C.W. (2012) Image Matching Using Peak Signal-to-Noise Ratio-Based Occlusion Detection. *IET image processing*, **6**, 483-495. <https://doi.org/10.1049/iet-ipr.2011.0025>