Total Variation Image Restoration Using the Primal-Dual Hybrid Gradient Algorithm

Shayan Ravari

Abstract

This document presents the theory, mathematical reasoning, and application of the Primal-Dual Hybrid Gradient (PDHG) algorithm in performing Total Variation (TV) image restoration. We delve into the Rudin-Osher-Fatemi (ROF) model for image denoising, discussing both its primal and dual formulations, and derive the PDHG algorithm by leveraging the saddle-point structure of the optimization problem. The implementation of the algorithm is demonstrated using Python, and results are presented to showcase its effectiveness in denoising images while preserving important features such as edges. Through detailed explanations and code examples, this document aims to provide a comprehensive understanding of how the PDHG algorithm can be applied to real-world image processing tasks.

Contents

1 Introduction

Image denoising is a fundamental problem in image processing and computer vision, where the goal is to recover a clean image from a noisy observation. Noise can enter the imaging process through various sources, such as sensor imperfections, environmental conditions, or transmission errors. Effective denoising is crucial for enhancing image quality and is a prerequisite for higher-level tasks like segmentation, object detection, and recognition. An ideal denoising algorithm should not only remove noise but also preserve essential features of the image, particularly edges and textures. Traditional methods like Gaussian smoothing tend to blur edges, leading to loss of important structural information. To address this issue, advanced techniques that promote edge preservation have been developed.

Total Variation (TV) regularization, introduced by Rudin, Osher, and Fatemi in 1992 [1], has become a cornerstone in the field of image denoising. The TV norm measures the integral of the absolute gradient of the image, promoting piecewise constant solutions. This property effectively reduces noise while maintaining sharp edges, making TV-based models highly effective for image restoration tasks. This document explores the use of the Primal-Dual Hybrid Gradient (PDHG) algorithm to efficiently solve the TV denoising problem formulated by the ROF model. We provide a detailed mathematical formulation, derive the algorithm using primal-dual concepts, and demonstrate its application through implementation. The aim is to offer a comprehensive guide that bridges the theoretical foundations with practical execution.

2 Mathematical Formulation

2.1 The ROF Model

The Rudin-Osher-Fatemi (ROF) model formulates the image denoising problem as an optimization task that balances fidelity to the observed data with smoothness imposed by the total variation regularization. Given a noisy image $f\colon \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ represents the image domain (typically a rectangular grid), the objective is to find a denoised image u that minimizes the total variation while remaining close to the observed data.

2.1.1 Constrained ROF Model

The constrained ROF model is expressed as:

$$\min_u \int_\Omega |\nabla u| \, \mathrm{d}x \quad \text{subject to} \quad \|u - f\|_2^2 \leq \sigma^2 |\Omega|, \tag{1}$$

where $|\nabla u|$ denotes the total variation of u, representing the integral of the gradient magnitude over the image domain. The term $\|u - f\|_2^2$ is the squared L^2 norm measuring the discrepancy between the denoised image u and the observed image f. The parameter σ^2 is an estimate of the noise variance in the observed image, and $|\Omega|$ is the area (number of pixels) of the image domain. This formulation seeks the smoothest possible image (minimal total variation) that is still consistent with the observed data within a specified noise level. This model was first proposed by Rudin, Osher, and Fatemi [1].

2.1.2 Unconstrained ROF Model

By introducing a Lagrange multiplier $\lambda > 0$, the constrained problem can be reformulated into an unconstrained optimization problem:

$$\min_{u} \int_{\Omega} |\nabla u| \, \mathrm{d}x + \frac{\lambda}{2} \|u - f\|_2^2. \tag{2}$$

In this formulation, the first term $\int_{\Omega} |\nabla u| \, \mathrm{d}x$ promotes smoothness by penalizing large gradients, effectively reducing noise while preserving edges. The second term $\frac{\lambda}{2}\|u - f\|_2^2$ ensures fidelity to the observed data, preventing excessive deviation from the input image. The parameter λ controls the balance between the regularization and fidelity terms. A larger λ places more emphasis on fitting the observed data, potentially retaining more noise, while a smaller λ emphasizes smoothness, which may over-smooth the image.

2.2 Total Variation Norm

The total variation (TV) of a function $u \in BV(\Omega)$ (functions of bounded variation) is defined as:

$$\int_{\Omega} |\nabla u| \, \mathrm{d}x = \int_{\Omega} \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2} \, \mathrm{d}x, \tag{3}$$

where ∇u is the gradient of u, and the integral is taken over the entire image domain Ω. The TV norm has the desirable property of being sensitive to edges (locations of high gradient magnitude) while being less sensitive to small fluctuations due to noise. This makes it particularly suitable for image denoising tasks where edge preservation is important.

3 Discretization

To implement the PDHG algorithm numerically, we discretize the continuous image domain and operators.

3.1 Image Representation

We represent the image domain Ω as a discrete grid of size $n \times n$, resulting in $N = n^2$ pixels. The image u is represented as a two-dimensional array of pixel values $u_{i,j}$, where (i, j) indexes the pixel location with $1 \leq i, j \leq n$. The observed noisy image f is similarly represented as $f_{i,j}$.

3.2 Discrete Gradient Operator

The discrete gradient ∇u at pixel (i, j) is approximated using forward finite differences:

$$\begin{aligned} (\nabla u)_{i,j}^x &= \begin{cases} u_{i+1,j} - u_{i,j}, & \text{if } i < n, \\ 0, & \text{if } i = n, \end{cases} \\ (\nabla u)_{i,j}^y &= \begin{cases} u_{i,j+1} - u_{i,j}, & \text{if } j < n, \\ 0, & \text{if } j = n. \end{cases} \end{aligned} \tag{4}$$

Here, $(\nabla u)^x_{i,j}$ approximates the partial derivative with respect to x at pixel (i, j), and $(\nabla u)^y_{i,j}$ approximates the partial derivative with respect to y. Boundary conditions are handled by assigning a zero gradient at the image edges, assuming Neumann boundary conditions.

3.3 Discrete Divergence Operator

The divergence $\nabla \cdot p$ of a vector field $p = (p^x, p^y)$ is approximated using backward finite differences:

$$(\nabla \cdot p)_{i,j} = \begin{cases} p^x_{i,j} - p^x_{i-1,j}, & \text{if } 1 < i \leq n, \\ p^x_{1,j}, & \text{if } i = 1, \end{cases} + \begin{cases} p^y_{i,j} - p^y_{i,j-1}, & \text{if } 1 < j \leq n, \\ p^y_{i,1}, & \text{if } j = 1. \end{cases} \tag{5}$$

In this formulation, $(\nabla \cdot p)_{i,j}$ computes the net flux exiting the pixel (i, j). The divergence operator is the negative adjoint of the gradient operator under appropriate boundary conditions. Proper handling of boundary conditions is crucial for ensuring that the discrete operators approximate their continuous counterparts accurately.

4 Primal-Dual Formulation

The optimization problem posed by the ROF model involves a non-smooth term due to the presence of the TV norm. To efficiently solve this problem, we employ a primal-dual formulation that leverages convex optimization techniques.

4.1 Dual Problem

The total variation norm can be expressed using its dual representation, which introduces a dual variable p:

$$\int_\Omega |\nabla u| \, \mathrm{d}x = \max_{\|p\| \leq 1} \int_\Omega \nabla u \cdot p \, \mathrm{d}x. \tag{6}$$

Using this dual representation, the ROF model can be reformulated as finding the saddle point of the Lagrangian:

$$\min_u \max_{\|p\| \leq 1} \int_\Omega \nabla u \cdot p \, \mathrm{d}x + \frac{\lambda}{2} \|u - f\|_2^2. \tag{7}$$

This primal-dual formulation allows us to handle the non-smoothness of the TV norm more effectively. Primal-dual algorithms have been developed to solve such saddle-point problems efficiently [2, 4].

4.2 Primal-Dual Hybrid Gradient Algorithm

The PDHG algorithm is designed to solve convex-concave saddle-point problems of the form:

$$\min_u \max_p \mathcal{L}(u, p) = G(u) + \langle Ku, p \rangle - F^*(p), \tag{8}$$

where $G(u)$ is a convex function in u, $F^*(p)$ is the convex conjugate of a convex function $F(p)$, K is a linear operator, and $\langle \cdot, \cdot \rangle$ denotes the inner product. For the ROF model, we identify $G(u) = \frac{\lambda}{2}\|u - f\|_2^2$, $F^*(p) = \delta_{\mathcal{P}}(p)$ where $\delta_{\mathcal{P}}(p)$ is the indicator function of the set $\mathcal{P} = \{p \,|\, \|p_{i,j}\|_2 \leq 1\}$, and $K = \nabla$, the gradient operator. The PDHG algorithm, introduced by Zhu and Chan [4], alternates between gradient ascent steps in the dual variable p and gradient descent steps in the primal variable u, with appropriate projections to handle constraints.

5 PDHG Algorithm

5.1 Algorithm Steps

At each iteration k, the PDHG algorithm performs the following updates:

1. **Dual Update:**

$$p^{k+1} = \mathrm{Proj}_{\mathcal{P}}\left(p^k + \tau\lambda\nabla u^k\right), \tag{9}$$

where $\mathrm{Proj}_{\mathcal{P}}$ projects p onto the constraint set $\mathcal{P} = \{p \,|\, \|p_{i,j}\|_2 \leq 1\}$, and $\tau > 0$ is the dual step size controlling the magnitude of the update in p.

2. **Primal Update:**

$$u^{k+1} = (1-\theta)u^k + \theta\left(f + \frac{1}{\lambda}\nabla \cdot p^{k+1}\right), \tag{10}$$

where $\theta \in (0, 1]$ is the primal step size determining the weight of the new update. The term $f + \frac{1}{\lambda}\nabla \cdot p^{k+1}$ combines the data fidelity and the influence of the dual variable.

5.2 Projection Onto the Unit Ball

The projection onto the set $\mathcal{P}$ is performed component-wise for each pixel (i, j):

$$p_{i,j} \leftarrow \frac{p_{i,j}}{\max(1, \|p_{i,j}\|_2)}. \tag{11}$$

This operation ensures that the constraint $\|p_{i,j}\|_2 \leq 1$ is satisfied at every pixel, which is essential for the convergence and stability of the algorithm.

5.3 Convergence Conditions

For the PDHG algorithm to converge, the step sizes τ and θ must satisfy certain conditions related to the operator norm of ∇:

$$\tau\theta\lambda^2\|\nabla\|^2 < 1. \tag{12}$$

In practice, since the operator norm $\|\nabla\|$ can be difficult to compute exactly, it is common to select τ and θ such that $\tau\theta \leq \frac{1}{4}$, which generally ensures convergence for images with normalized pixel values.

6 Conclusion

The Primal-Dual Hybrid Gradient algorithm provides an efficient and effective method for solving the Total Variation denoising problem posed by the ROF model. By leveraging the primal-dual formulation, the algorithm addresses the non-smoothness of the TV norm and efficiently handles the associated constraints through projection. The implementation demonstrates that the PDHG algorithm can restore images corrupted with noise while preserving essential features such as edges and textures. This balance between noise reduction and detail preservation is a key advantage of TV-based denoising methods. Moreover, the algorithm's simplicity and reliance on basic operations make it suitable for large-scale image processing tasks. Its flexibility allows for adaptation to other inverse problems in imaging, such as deblurring and compressed sensing.

References

[1] Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4), 259-268.

[2] Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2), 89-97.

[3] Goldstein, T., & Osher, S. (2009). The split Bregman method for L1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2), 323-343.

[4] Zhu, M., & Chan, T. (2008). An efficient primal-dual hybrid gradient algorithm for total variation image restoration. *UCLA CAM Report*, 34.