MATIMYÁS MATEMATIKA
ISSN 0115-6926

Journal of the Mathematical Society of the Philippines
Vol. 30 Nos. 2-3 (2007) pp. 103-112

A Primal-Dual Island Search Method for Total Variation-Based Image Restoration

Marrick Neri
Institute of Mathematics
University of the Philippines Diliman
Quezon City 1101, Philippines
email: marrick@math.upd.edu.ph

Michael Hintermüller
University of Graz
Graz, Austria
email: michael.hintermueller@uni-graz.at

Abstract

For the denoising of blocky images, based on the nonsmooth TV-model, a new method combining the algorithm introduced by Chambolle [1] and a primal-dual active set strategy due to Hintermüller and Stadler [7] is introduced. The active set technique aims at decomposing the image into edges (active sets) and flat zones (inactive sets, which we call islands). The data reconstruction on the islands is based on an averaging technique reflecting the statistics of the noise. Utilizing the primal and a corresponding dual variable in the island detection stabilizes the algorithm. The dual update is based on the strategies in [1] and [7]. Numerical tests showing that the method is highly efficient in removing the noise, in restoring edges, and in the reconstruction of flat image features.

1 Introduction

Given an observed noisy image d in a domain Ω, we want to get the best reconstruction $s \in \mathrm{L}^2(\Omega)$ of the original clean image. To date there are many ways of doing so, the more successful of which use variational or partial differential equation based approaches. Variational models are efficient in removing high oscillations which are often associated to noise. The most popular variational example is the Rudin-Osher-Fatemi (ROF) total variation based image denoising model [9]. Specifically, the authors in [9] proposed solving the minimization problem

$$\min \quad \int_\Omega |\nabla s| \, dx \tag{1}$$

$$\text{subject to } s + v = d, \tag{2}$$

$$\int_\Omega |v|^2 \, dx \leq \sigma^2, \tag{3}$$

where $|\cdot|$ denotes the Euclidean norm, v is Gaussian noise, and σ^2 is the noise variance. When the variance is undetermined, ROF proposed minimizing the unconstrained problem

$$\int_\Omega |s-d|^2\ dx + \alpha \int_\Omega |\nabla s|\ dx$$

which yields the same solution as rof$_o forasuitablyselectedLagrangemultiplier\alpha$ [2]. The ROF model is effective in preserving edges as it allows discontinuities in the reconstruction. Currently there are various approaches to solve the ROF, some of them we cite here. In [9] the authors used a delayed time marching scheme. A primal-dual method was proposed in [3], while a fixed-point algorithm to solve the dual of the ROF was introduced in [1]. Active set methods that exploit the primal-dual structure of ROF were used in [6] and [7]. A second-order cone programming method was presented in [4]. In this research, we develop an algorithm that utilizes an active set method and that exploits the statistics of the noise.

In this paper, we introduce a primal-dual active set method that exploits the statistics of the noise in image subdomains where the norm of the gradient is some small value. For images with piecewise constant regions, the method is shown to be effective and fast in both restoring flatness features and sharp edges. However, the updates on the primal variable is not strictly a Newton update, and thus convergence with respect to Karush-Kuhn-Tucker residuals is not obtained.

In the ensuing discussions, Ω is a simply connected domain in R^2 with Lipschitz continuous boundary $\partial\Omega$. In the discrete setting, Ω is the $n \times n$ pixel-square. We denote by ∇ the (distributional) gradient operator. The divergence operator div is the adjoint of ∇, and $\triangle$ denotes the Laplace operator.

For a vector $v := \begin{bmatrix} v_x \\ v_y \end{bmatrix} \in R^{2N}$, with $N = n^2$, the Euclidean norm $|v| \in R^N$ is given by $|v|_i = \left(v_{x_i}^2 + v_{y_i}^2\right)^{\frac{1}{2}}$ for $i = 1, \ldots, N$. Given $v, w \in R^m$, we also define the following operations: $q = v/w$ implies $q_i = v_i/w_i$, and $q = v \star w$ implies $q_i = v_i w_i$ for $i = 1, \ldots, m$.

On a closed subset P in Ω, we denote the subvector u_P of u those whose indices are in P. Given a vector $w \in R^m$, we denote by $D(w)$ the $m \times m$ diagonal matrix with diagonal entries w_i, $i = 1, \ldots, m$. The vector of ones in R^m is represented by 1_m, and I_m is the $m \times m$ identity matrix.

The motivation for this work is the assumption that in Ω, the value of pixels in a closed and flat subdomain, say $\mathcal{I}$, is constant. This makes $|(\nabla u)_i| = 0$ for almost every $i \in \mathcal{I}$, whereas on the boundary of $\mathcal{I}$, $|(\nabla u)_i|$ can be large. Assuming that the image is corrupted with additive white noise, we can apply an averaging scheme to denoise flat portions of the image, and then use standard denoising methods on the boundary set. This approach can greatly lessen the amount of variable pixels to solve while recovering sharpness in edges and flat features.

2 The Problem

A popular and well researched model for image denoising is the Rudin-Osher-Fatemi (ROF) total variation model:

$$\min_{s \in BV(\Omega)} \tfrac{1}{2}\Omega|s-d|^2\ dx + \alpha\Omega|\nabla s|\ dx, \tag{ROF}$$

where the space of bounded variations BV is defined by

$$BV(\Omega) := \{f \in L^1 \ : \ \int_\Omega |\nabla f| < \infty\}.$$

The first term in ROF is the data fidelity term which is responsible for recovering the features of the image. The second term is the total variation (TV) term which aids in preserving edges and removing high oscillations - usually typical with noise- in the image [8, 10]. The ROF is highly effective in image denoising; however, it exhibits the staircase effect, i.e., smooth curves are approximated by piecewise constant functionals. Since the TV term is also nondifferentiable at zero, many approaches applied a regularization of the TV term by a small positive parameter ε, i.e., by replacing $|\nabla s|$ with $\sqrt{|\nabla s|^2 + \varepsilon}$ (e.g. see [9, 8, 10]).

An alternative approach to solving the ROF is to solve its dual form. The dual of the ROF is a differentiable quadratic problem involving simple constraints. Applying the Fenchel duality theorem (cf. [7]) on ROF, we obtain the dual problem

$$\sup_{\substack{p \in L^2(\Omega^2), \\ |p| \leq \alpha \text{ a.e. in } \Omega}} \left\{-\tfrac{1}{2}\text{div } p + d + \tfrac{1}{2}d\right\}. \tag{DROF}$$

The solutions s^* and p^* of ROF and DROF respectively are characterized by the optimality conditions [7]

$$\text{div } p^* = s^* - d, \tag{4}$$

$$\left.\begin{array}{rl} |\nabla s^*| \star p^* - \alpha \nabla s^* = 0 & \text{if} \quad |p^*| = \alpha \\ \nabla s^* = 0 & \text{if} \quad |p^*| < \alpha \end{array}\right\} \quad \text{in } L^2. \tag{5}$$

Despite the smoothness of the DROF, its solution is not unique because the kernel of the divergence operator is nonempty. In [1] Chambolle proposed a fixed point method to solve the DROF. The method obtains a dual solution that determines the unique solution to the ROF.

3 On Chambolle's algorithm

Let Ω_h be the discrete domain Ω, and $Y = \Omega_h \times \Omega_h$. The discrete gradient operator is given by the matrix $\nabla_h = \begin{bmatrix} D_x \\ D_y \end{bmatrix}$, where D_x and D_y are respectively the horizontal and vertical forward difference operators. For the discrete divergence operator we use $\text{div}_h = -\nabla_h^\top$, and $\triangle_h$ denotes the discrete Laplacian obtained using the standard five-point stencil with Dirichlet boundary conditions. We now review the results in [1].

A solution of ROF can be simply given by $s^* = d + \pi_{\alpha\bar{K}}(d)$, where $\pi_{\alpha\bar{K}}$ is the nonlinear projection and $\bar{K}$ is the closure of the set

$$K = \left\{\text{div } \psi \ : \ \psi \in C_0^1(\Omega; R^2), |\psi(x)| \leq 1 \ \forall x \in \Omega\right\}.$$

Chambolle developed a fixed point method that computes this projection in dimension 2 by way of solving the problem

$$\min_{\substack{p \in Y, \\ |p|^2 \leq 1 \text{ a.e. in } X}} \quad \alpha \text{div}_h \ p + d \tag{6}$$

.

Karush-Kuhn-Tucker conditions yield the existence of a Lagrange multiplier λ for the constraint in dualC. The corresponding Euler-Lagrange equation is

$$-(\nabla(\alpha \mathrm{div}_h\ p+d))+\lambda \star p=0, \tag{7}$$

where $\lambda \in R^{2N}$ and nonnegative.

Complementary conditions on λ and p state that whenever $|p_i|=1$, then $\lambda_i>0$, and when $|p_i|<1$, $\lambda_i=0$. For both cases,

$$\lambda_i=|(\nabla(\alpha \mathrm{div}_h\ p+d))_i|.$$

In the algorithm proposed in [1] (we label as algorithm C) the p-update is $p^{k+1}=\begin{bmatrix} p_x^{k+1} \\ p_y^{k+1} \end{bmatrix}$ where

$$p_l^{k+1}=\frac{p_l^k+\tau D_l(\mathrm{div}_h\ p^k+d/\alpha)}{1+\tau|\nabla_h(\mathrm{div}\ p^k+d/\alpha)|},\ l=x,y. \tag{8}$$

We get the resulting theorem on the convergence property of the algorithm.
Theorem (Theorem 3.1, [1])

Let $\tau \leq 1/8$. Then $\alpha div_h\ p^k$ converges to $\pi_{\alpha \bar{K}}(d)$ as $k \to \infty$.

Setting $s^k=d+\alpha \mathrm{div}_h\ p^k$ and $\beta=\alpha/\tau$, the p-update can be expressed as

$$p_l^{k+1}=\frac{\beta p_l^k+D_l s^k}{\beta+|(\nabla_h s^k)|},\quad l=x,y. \tag{9}$$

With this representation by s, the Lagrange multiplier λ can be expressed as $\lambda=|\nabla_h s|$.

Numerical implementations of algorithm C show it to be efficient in denoising images and restoring edges. It runs fast but it also requires many iterations to converge.

In [7] Hintermüller and Stadler presented a primal-dual active set method that solves a regularized version of DROF. The method was shown to converge superlinearly to the unique solution of the predual of the version.

4 On Hintermüller and Stadler's algorithm

A regularization of the dual problem DROF by a positive parameter γ was proposed in [7]:

$$\sup_{\substack{p\ \in\ L^2(\Omega^2),\\ |p|\ \leq\ \alpha \text{ a.e. in } \Omega}} \left\{-\tfrac{1}{2}\mathrm{div}\ p+d+\tfrac{1}{2}d\right\}-\frac{\gamma^{s-1}}{s\alpha^{s-1}}p_{L^s}^s, \tag{10}$$

where $1<s\leq 2$. The objective functional in rdual is L^s-uniformly concave, and thus a unique solution p^* exists for every fixed γ. We show the results in [7] for the case $s=2$.

By the Fenchel duality theorem, the dual of rdual is given by

$$\min_{s\in H_0^1(\Omega)} \frac{1}{2}\Omega|s-d|^2\ dx+\alpha\Omega\Phi_\gamma(\nabla s)\ dx, \tag{11}$$

where for $v \in \mathbf{L}^2$, $\Phi_\gamma(v)(x):=\begin{cases} |v(x)|-\frac{1}{2}\gamma & \text{if} \quad |v(x)|\geq\gamma, \\ \frac{1}{2\gamma}|v(x)|^2 & \text{if} \quad |v(x)|<\gamma. \end{cases}$

The solutions p^*_γ and s^*_γ of problems rdual and rprimal, respectively, satisfy

$$s^*_\gamma - \operatorname{div}\, p^*_\gamma = d \text{ in } H^{-1}(\Omega), \tag{12}$$

$$\left.\begin{array}{rl} |\nabla s^*_\gamma| \star p^*_\gamma - \alpha \nabla s^*_\gamma = 0 & \text{if} \quad |p^*_\gamma| = \alpha, \\ \gamma p^*_\gamma - \alpha \nabla s^*_\gamma = 0 & \text{if} \quad |p^*_\gamma| < \alpha \end{array}\right\} \text{ in } \mathbf{L}^2(\Omega). \tag{13}$$

Remark. These conditions mirror almost exactly the conditions in rofgr1 and rofgr2, differing only in the added term γp^*_γ in kkt2. The equations in kkt2 can be joined into one equation

$$\max(\gamma, |\nabla s^*_\gamma|) \star p^*_\gamma - \alpha \nabla s^*_\gamma = 0. \tag{14}$$

The results in generalized differentiability and semismoothness of the max operator and the (Euclidean) ℓ^2-norm (cf. [5]) allow the use of a Newton step to the discretized forms of kkt1 and maxkkt at the approximations s^k and p^k:

$$\begin{pmatrix} I_N & -\operatorname{div}_h \\ G\nabla & D(\tilde{m}^k) \end{pmatrix} \begin{pmatrix} \delta_s \\ \delta_p \end{pmatrix} = \begin{pmatrix} -s^k + \operatorname{div}_h p^k + d \\ \alpha \nabla_h s^k - D(\tilde{m}^k) p^k \end{pmatrix}, \tag{15}$$

where

$$\begin{array}{rcl} G & = & \left(-\alpha I + \chi_{A_{k+1}} D(p^k) J(\nabla_h s^k)\right), \\ \tilde{m}^k & = & \max\left(\gamma 1_{2N}, \eta(\nabla_h s^k)\right) \in R^{2N}, \end{array}$$

and the mapping $\eta : R^{2N} \to R^{2N}$ is defined by

$$(\eta(v))_i = |v|_i \quad \text{for } v \in R^{2N},\ i = 1, \ldots, 2N.$$

In G, we have $\chi_{A_{k+1}} = D(t^k) \in R^{2N \times 2N}$ with

$$t^k_i := \begin{cases} 1 & \text{if } \eta(\nabla_h s^k)_i \geq \gamma, \\ 0 & \text{else.} \end{cases}$$

The variable t^k determines where an index i belongs: to the active set if $t^k_i = 1$; otherwise, to the inactive set. The matrix J denotes the Jacobian of η, i.e., for $\nabla_h s = (D_x s, D_y s)^\top \in R^{2N}$,

$$J(\nabla_h s) = (D(\eta(\nabla_h s)))^{-1} \begin{pmatrix} D(D_x s) & D(D_y s) \\ D(D_x s) & D(D_y s) \end{pmatrix}$$

Since $\tilde{m}^k_i > 0$, for $i = 1, \ldots, 2N$, the matrix $D(\tilde{m}^k)$ is invertible. Solving for δ_p, we get

$$\delta_p = \alpha D(\tilde{m}^k)^{-1} \nabla_h s^k_\gamma - p^k - D(\tilde{m}^k)^{-1} G \nabla_h \delta_s. \tag{16}$$

Substituting δ_p to the first equation in Newtonstep, we get

$$H_k \delta_s = f_k, \tag{17}$$

where the matrix H_k and the right hand side f_k are defined as

$$\begin{array}{rcl} H_k & := & I_N + \nabla_h^\top D(\tilde{m}^k)^{-1} \left(\alpha I_{2N} - \chi_{A_{k+1}} D(p^k) J(\nabla_h s^k)\right) \nabla_h, \\ f_k & := & -s^k + d - \alpha \nabla_h^\top D(\tilde{m}^k)^{-1} \nabla_h s^k. \end{array}$$

The updates follow:

$$s^{k+1} = s^k + \delta_s \quad \text{and} \quad p^{k+1} = p^k + \delta_p. \tag{18}$$

It was shown in [7] that although the matrix H_k is in general not symmetric, in the solution (s^*_γ, p^*_γ) it is symmetric. In case H_k is not positive definite, p^k is projected into the feasible region, e.g. by replacing p_i^k by $\alpha \max\left\{\alpha, 1/|p_i^k|\right\} p_i^k$ for each i wherein $|p_i| > \alpha$. The projection results in a modified system matrix H_k^+ which is positive definite. Assuming that $s^k \to s^*$ and $p^k \to p^*$, the modified system matrices H_k^+ converges to H_k as $k \to \infty$ (see Lemma 3.5, [7]).

In [5] it was shown that primal-dual active set methods are equivalent to the semismooth Newton method. In [7] the authors introduced a primal-dual active set algorithm (we label as Algorithm *HS*) that applies the above updates on (s, p). The following result was established.

Theorem (Theorem 3.6, [7])

Assuming the feasibility of p^k for every $k \in N$, the iterates (s^k, p^k) in Algorithm HS converge superlinearly to (s^*, p^*), provided that (s^0, p^0) is sufficiently close to (s^*, p^*). Moreover, this convergence is global.

In terms of number of iterations, algorithm *HS* converges faster than algorithm *C*. However, *HS* requires solving the full system du, which can be memory-expensive. The implementation in [7] made use of preconditioned conjugate gradient method with incomplete Cholesky factorization to speed up the solution process.

For a simple image such as a binary image with blocky structures, it may be faster to recover the original image if we know exactly the edge set and the closed domains bounded by the edges. We could simply assign a single value to all pixels in a closed region. The problem is simplified to determining the closed regions and the edges. Consequently, the problem would require fewer variables to solve compared to solving the full system.

We introduce a primal-dual active set approach on the ROF that detects the edge set, utilizes the statistics of the noise in the closed domains, and reconstructs boundaries and piecewise-constant features of the image. The proposed algorithm is observed to be most effective in reconstructing images with blocky features. We call the proposed algorithm as the primal-dual island method.

5 The primal-dual island (PDI) method

Usually, an image is made up of flat portions and closed edges. In a flat portion, the value of the pixels is constant, thus the gradient is zero. We will call the edges the boundaries and the flat portions *islands*.

Definition Let $\mathcal{I}$ be a subset of Ω such that $|(\nabla s)_i| = 0$ for all $i \in \mathcal{I}$. Then $\mathcal{I}$ is called an island in Ω.

Since in an island the gradient is zero, we will call the set of islands in Ω the inactive set I. The set of boundaries enclosing the islands will be the active set A. These assignments for I and A agree with the corresponding definitions in [7]. Note that for every $\mathcal{I}_r \subset I$, $r = 1, \ldots, T$, such that $I_r \cap I_j = \emptyset$ for $r \neq j$, we have $I = I_1 \cup \cdots \cup I_T$.

Given an original image d^o we add to it Gaussian noise ξ^o to obtain the observed image $d := d^o + \xi^o$. We assume that over an island with a sufficient number of pixels, the mean of the noise is some small positive number ε. Thus a good approximation of the pixels in the island is the mean of that island. Clearly, averaging removes high oscillations in an island and recovers flatness.

In the proposed PDI method, we decompose the components of the solution s into A and I. The main feature is the s-update in I. Over $I_r \subset I$ the update for s is the mean value $\bar{d}_{I_r}$, i.e.,

$$\forall i \in I_r,\ s_i^k = \bar{d}_{I_r} := \frac{1}{|I_r|_c} \sum_{i \in I_r} d_i = \frac{1}{|I_r|_c} \sum_{i \in I_r} d_i^o + \varepsilon_i, \tag{19}$$

where we define $|\cdot|_c$ to be the set cardinality. The PDI comes cheap to implement since the total number of variables to solve equals $T + |A|_c$.

We hybridize algorithms C and HS to adopt the PDI method. On the active set we use the standard updates of these algorithms. The averaging scheme on I can give a better TV-term value, but possibly not a better value in the data fidelity term. Since the update method is not entirely a gradient method or a Newton method, convergence properties associated with the latter methods cannot be expected from PDI at its current state. Convergence of PDI in the sense of a minimum KKT residual may not be achieved. Here we use as KKT residual $KKT_{res} = (\|F_1\|^2 + \|F_2\|^2)^{\frac{1}{2}}$, where F_1 and F_2 are the left hand sides, respectively, of kkt1 and maxkkt.

5.1 Algorithms PDI-C and PDI-HS

We use the observation in [6] that the multiplier $|\nabla_h s|$ is efficient in detecting edges. Let $\gamma > 0$. In equation EulerC the Lagrange multiplier is $\lambda_i^k = |(\nabla_h s^k)_i|,\ i = 1, \ldots, 2N$, where we set $s^k = d + \alpha \mathrm{div}_h\, p^k$. Note that $\lambda_i = \lambda_{i+N}$ and $|p_i| = |p_{i+N}|$. For $\gamma > 0$ we apply the active set determination in [5]:

$$A^{k+1} = \left\{ i \in \{1, \ldots, N\} \ :\ \lambda_i^k + \gamma(|p_i^k| - 1) > 0 \right\} \quad \text{and} \quad I^{k+1} = \Omega \setminus A^{k+1}. \tag{20}$$

We modify algorithm C so that it will utilize the PDI method. The resulting algorithm is PDI-C:

Algorithm PDI-C

1. *Initialize* $p^0 = 0$ *and* $s^0 = d$. *Set* $k = 0$.
2. *Determine the active and inactive sets* A^{k+1} *and* I^{k+1} *respectively, by* AIC.
3. *Update* p *according to* newp1.
4. *Update* s*: in* I^{k+1}*, use* uI*; and in* A^{k+1}, $s_{A^{k+1}} = (d + \alpha\, div_h\ p^k)_{A^{k+1}}$.
5. *Stop, or set* $k = k + 1$ *and go to step 2.*

The regularized dual problem rdual was solved in [7] by a primal-dual active set method. In this method, the active and inactive sets were determined as

$$A^{k+1} = \{ i \in \{1, \ldots, N\} \ :\ |(\nabla_h s)_i| > \gamma \} \quad \text{and} \quad I^{k+1} = \Omega \setminus A^{k+1}. \tag{21}$$

Implementing the PDI method on algorithm HS, we get the following algorithm:

Algorithm PDI-HS

1. *Initialize* $p^0 = 0$ *and* $s^0 = d$. *Set* $k = 0$.
2. *Determine the active and inactive sets* A^{k+1} *and* I^{k+1}*, respectively, using* AIHS

3. Update s: in I^{k+1}, use uI*; and in A^{k+1}, $s^{k+1}_{A^{k+1}} = s^k_{A^{k+1}} + \delta^k_{s_{A^{k+1}}}$,.*
4. Update p according to upHS*.*
5. Stop, or set $k = k + 1$ and go to step 2.

In step 3, $\delta^k_{s_{A^{k+1}}}$ is obtained by solving in du the subsystem of equations

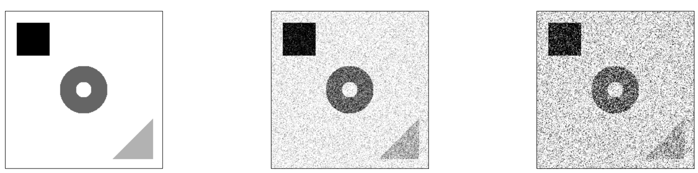
$$H_{AA}\delta^k_{s_{A^{k+1}}} = f_A - H_{AI}\delta^k_{s_{I^{k+1}}}.$$

By H_{AA} we mean the submatrix of H with row indices and column indices in A. The matrix H_{AI} is defined similarly. In the inactive set, the increment $\delta^k_{s_{I^{k+1}}} = s^{k+1}_{I^{k+1}} - s^k_{I^{k+1}}$.

6 Numerical computations

We present here numerical implementations of PDI-C and PDI-HS on a 255×255 image (figure 2(a)) with noise standard deviation $n_l = 0.2$ and $n_l = 0.4$ (figures 2(b), 2(c)) and with corresponding TV-regularization parameters $\alpha = 0.4$ and $\alpha = 0.7$ respectively. The algorithms terminate when the desired KKT residual accuracy is reached or upon some other stopping criterion, e.g. a specified time limit.

Figure 1: Image data



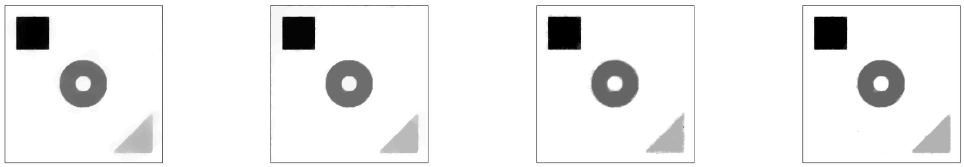
(a) True Image (b) With noise $\eta_0 = 0.2$ (c) With noise $\eta_0 = 0.4$

Results on the noisy images show that the PDI method can recover almost accurately edges and flat portions. For instance in Figure 2 both algorithms C and HS tend to lessen contrast, particularly in the hole of the annulus, while the PDI restores better contrast. At the higher noise level 40% the loss of contrast in the solutions of C and HS is more evident, especially along the edges of the triangle and the annulus (figures 4(a), 4(b)).

The PDI method gives a small residual to the complementarity system maxkkt, showing that the averaging update is very good for recovering flat portions. However, the method yields larger residual for kkt1. Due to the higher KKT residual, the image reconstruction of PDI - though visually could be better than the results of algorithms C and HS - may not be the minimizer of ROF.

We observe that the PDI method works better with algorithm HS. In both noise levels, the PDI-HS reconstructions approximate better the original image within fewer iterations. Tables 1 and 2 show some statistics of the implementations. The entry *image residual* denotes the distance of the reconstruction from the clean image, which in natural cases may not be known.

Figure 2: Reconstructions for image with 20% noise.

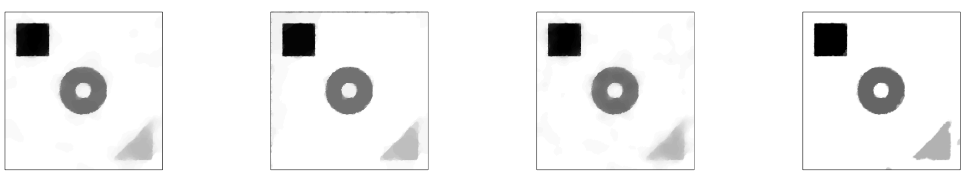


(a) C (b) HS (c) PDI-C (d) PDI-HS

Table 1: $\alpha = 0.4$, noise $n_l = 0.2$

	γ	iterations	KKT_{res}	image residual	function cost	time
C	1.375×10^{-3}	497	0.1632	5.9308	0.7795	85.3125
PDIC	1.375×10^{-3}	41	19.2501	6.9209	0.8742	185.3438
HS	5.42×10^{-5}	14	2.6×10^{-7}	6.2528	1.5178	46.5781
PDIHS	3.75×10^{-2}	11	52.6347	4.6782	0.7279	85.2031

Figure 3: Reconstructions for image with 40% noise



(a) C (b) HS (c) PDI-C (d) PDI-HS

Table 2: $\alpha = 0.7, n_l = 0.4$

	γ	iterations	KKT_{res}	image residual.	cost	time
C	7.9×10^{-3}	875	0.1685	9.2777	1.4045	131.3594
PDIC	7.9×10^{-5}	376	0.7613	10.0393	1.5514	185.4063
HS	5.42×10^{-5}	14	3.7×10^{-7}	10.3948	2.6754	49.1719
PDIHS	3.83×10^{-2}	30	103.2406	8.9188	1.2478	131.3125

Conclusion

Although the PDI method does not converge in the smooth sense, its use of the active set method and its averaging scheme makes it a robust method in restoring edges and blocky features. Numerical results show that the PDI method outperforms standard methods (algorithms C and HS) in that it yields solutions which are good reconstructions of images with piecewise-constant features.

References

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

[2] A. Chambolle and P. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997.

[3] T. Chan, G. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM Journal on Scientific Computing*, 20:1964–1977, 1999.

[4] D. Goldfarb and W. Yin. Second-order cone programming methods for total variation-based image restoration. *SIAM Journal on Scientific Computing*, 27(2):622–645, 2005.

[5] M. Hintermüller, K. Ito, and K. Kunisch. A primal-dual active set strategy as a semismooth newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2003.

[6] M. Hintermüller and K. Kunisch. Total bounded variation regularization as a bilaterally constrained optimization problem. *SIAM Journal on Applied Mathematics*, 64(4):1311–1333, 2004.

[7] M. Hintermüller and G. Stadler. An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. *SIAM Journal on Scientific Computing*, 28(1):1–23, 2006.

[8] T. Kärkkäinen, K. Majava, and M. Mäkelä. Comparison of formulations and solution methods for image restoration problems. *Inverse Problems*, 17:1977–1995, 2001.

[9] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[10] C. Vogel. *Computational methods for inverse problems*. SIAM Frontiers in Applied Mathematics., 2002.