Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman

Pascal Getreuer

published • 2012-05-19
reference • Pascal Getreuer, *Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman*, Image Processing On Line, 2012. http://dx.doi.org/10.5201/ipol.2012.g-tvd

- Pascal Getreuer pascal.getreuer@yale.edu, Yale University

Communicated by Jean-Michel Morel morel@cmla.ens-cachan.fr, CMLA, ENS Cachan
Edited by Pascal Getreuer pascal.getreuer@yale.edu, Yale University

Overview

Denoising is the problem of removing noise from an image. The most commonly studied case is with additive white Gaussian noise (AWGN), where the observed noisy image f is related to the underlying true image u by

$$f = u + \eta,$$

and η is at each point in space independently and identically distributed as a zero-mean Gaussian random variable.

Total variation (TV) regularization is a technique that was originally developed for AWGN image denoising by Rudin, Osher, and Fatemi [9]. The TV regularization technique has since been applied to a multitude of other imaging problems, see for example Chan and Shen's book [20]. We focus here on the split Bregman algorithm of Goldstein and Osher [31] for TV-regularized denoising.

TV Regularization

Rudin, Osher, and Fatemi [9] proposed to estimate the denoised image u as the solution of a minimization problem,

$$\operatorname*{arg\,min}_{u \in BV(\Omega)} \|u\|_{\mathrm{TV}(\Omega)} + \frac{\lambda}{2} \int_\Omega \big(f(x) - u(x)\big)^2 \, dx,$$

where λ is a positive parameter. This problem is referred to as the Rudin-Osher-Fatemi or ROF problem. Denoising is performed as an infinite-dimensional minimization problem, where the search space is all bounded variation (BV) images. A function u is in $BV(\Omega)$ if it is integrable and there exists a Radon measure Du such that

$$\int_\Omega u(x) \operatorname{div} \vec{g}(x) \, dx = -\int_\Omega \langle \vec{g}, Du(x) \rangle \quad \text{for all } \vec{g} \in C_c^1(\Omega, \mathbb{R}^2)^2.$$

This measure Du is the distributional gradient of u. When u is smooth, $Du(x) = \nabla u(x)\, dx$. The *total variation (TV) seminorm* of u is

$$\|u\|_{\mathrm{TV}(\Omega)} := \int_\Omega |Du| := \sup\Big\{ \int_\Omega u \operatorname{div} \vec{g}\, dx : \vec{g} \in C_c^1(\Omega, \mathbb{R}^2)^2, \sqrt{g_1^2 + g_2^2} \le 1 \Big\}.$$

When u is smooth, TV is equivalently the integral of its gradient magnitude,

$$\|u\|_{\mathrm{TV}(\Omega)} = \int_\Omega |\nabla u|\, dx.$$

The TV term in the minimization discourages the solution from having oscillations, yet it does allow the solution to have discontinuities. The second term encourages the solution to be close to the observed image f. By this combination, the minimization finds a denoised image. If $f \in L^2$, the minimizer of the ROF problem exists and is unique and is stable in L^2 with respect to perturbations in f [24].

From a Bayesian point of view, this formulation is a maximum a posteriori estimate using a TV prior. From the AWGN noise model, the conditional probability density $p(f|u)$ is

$$\begin{aligned} p\big(f(x)|u\big) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{1}{2\sigma^2} \int_\Omega \big(f(x) - u(x)\big)^2\, dx\Big) \\ -\log p\big(f|u\big) &= const + \frac{1}{2\sigma^2} \int_\Omega (f-u)^2\, dx, \end{aligned}$$

where σ is the noise variance. The maximum a posteriori estimate is

$$\begin{aligned} u &= \arg\max_u p(u|f) \\ &= \arg\max_u p(u) p(f|u) \\ &= \arg\min_u -\log p(u) - \log p(f|u) \\ &= \arg\min_u -\log p(u) + \frac{1}{2\sigma^2} \int_\Omega (f-u)^2\, dx. \end{aligned}$$

The $-\log p(u)$ term is the *prior* on u, an *a priori* assumption on the likelihood of a solution u. With total variation regularization, the selected prior is

$$-\log p(u) = \mu \|u\|_{\mathrm{TV}(\Omega)},$$

where μ is a positive parameter controlling the regularization strength. A larger value of μ places more emphasis on the prior, leading to a more regular solution. The ROF problem is equivalent to the maximum a posteriori estimate with $1/\lambda = \mu\sigma^2$,

$$\arg\min_u \mu \|u\|_{\mathrm{TV}(\Omega)} + \frac{1}{2\sigma^2} \int_\Omega \big(f(x) - u(x)\big)^2\, dx.$$

Through this connection with maximum a posteriori estimation, TV-regularized denoising can be extended to other noise models. Alliney [12] and Chan and Esedoḡlu [19] developed TV denoising for Laplace noise (L^1 data fidelity),

$$\arg\min_u \|u\|_{\mathrm{TV}(\Omega)} + \lambda \int_\Omega \big|f(x) - u(x)\big|\, dx,$$

Le, Chartrand, and Asaki [25] developed TV denoising for Poisson noise,

$$\arg\min_u \|u\|_{\mathrm{TV}(\Omega)} + \lambda \int_\Omega \big(u(x) - f(x)\log u(x)\big)\,dx.$$

TV denoising has been similarly extended to multiplicative noise [17],[45] and Rician noise [42].

Furthermore, these models can be extended to use a spatially varying λ (see for example [23]) to impose a locally adapted regularization strength at different points of space,

$$\arg\min_u \|u\|_{\mathrm{TV}(\Omega)} + \frac{1}{2}\int_\Omega \lambda(x)\big(f(x) - u(x)\big)^2\,dx.$$

TV-based inpainting [20] is an interesting special case where λ is set to zero over some region of space. For x where $\lambda(x) = 0$, the observed value $f(x)$ is ignored and $u(x)$ is only influenced by the $||u||_{\mathrm{TV}}$ term.

The choice of noise model can significantly affect the denoising results. For better results, the noise model should agree with the actual noise distribution in the image.

Algorithms

For numerical solution of the minimization problem, several approaches for implementing the TV seminorm have been proposed in the literature. TV is most often discretized by

$$\|u\|_{\mathrm{TV}(\Omega)} \approx \sum_{i,j} \sqrt{(\nabla_x u)_{i,j}^2 + (\nabla_y u)_{i,j}^2},$$

where ∇_x and ∇_y are discretizations of the horizontal and vertical derivatives. A difficulty with TV is that it has a derivative singularity when u is locally constant. To avoid this, some algorithms regularize TV by introducing a small parameter $\varepsilon > 0$ within the square root,

$$\sum_{i,j} \sqrt{\epsilon^2 + (\nabla_x u)_{i,j}^2 + (\nabla_y u)_{i,j}^2}.$$

Let ∇_x^+, ∇_x^-, ∇_y^+, ∇_y^- denote the forward (+) and backward (−) finite differences in the x and y directions and let $m(a,b)$ denote the minmod operator

$$m(a,b) = \left(\frac{\operatorname{sign} a + \operatorname{sign} b}{2}\right)\min(|a|,|b|).$$

Several ways to discretize the derivatives are

- One-sided difference $(\nabla_x u)^2 = (\nabla_x^+ u)^2$
- Central difference $(\nabla_x u)^2 = \big((\nabla_x^+ u + \nabla_x^- u)/2\big)^2$
- Geometric average $(\nabla_x u)^2 = \big((\nabla_x^+ u)^2 + (\nabla_x^- u)^2\big)/2$
- Minmod $(\nabla_x u)^2 = m(\nabla_x^+ u, \nabla_x^- u)^2$
- Upwind discretization [39] $(\nabla_x u)^2 = \big(\max(\nabla_x^+ u, 0)^2 + \max(\nabla_x^- u, 0)^2\big)/2$

Central differences are undesirable for TV discretization because they miss thin structures. The central difference at (i,j) does not depend on $u_{i,j}$:

$$\frac{\nabla_x^+ u_{i,j} + \nabla_x^- u_{i,j}}{2} = \frac{(u_{i+1,j} - u_{i,j}) + (u_{i,j} - u_{i-1,j})}{2} = \frac{u_{i+1,j} - u_{i-1,j}}{2}.$$

Therefore, if u has a one-sample wide structure like $u_{0,j} = 1$ and $u_{i,j} = 0$ for all $i \neq 0$, the variation at $(0,j)$ estimated by central differences is zero. To avoid this problem, one-sided differences can be used, however, they are not symmetric. The geometric average, minmod, and upwind estimates listed above regain symmetry by combining the forward and backward one-sided differences, though at the cost that then the derivatives are nonlinear. Another concern is whether a TV discretization is consistent, that is, whether the discrete TV converges to the true TV as the grid resolution becomes infinitely fine, see for example Wang and Lucier [40].

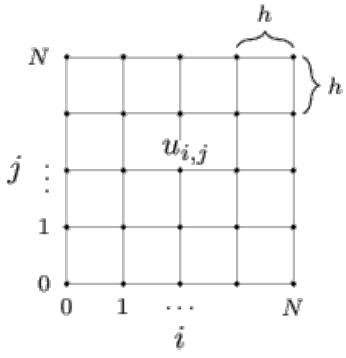
Another twist is that some algorithms substitute TV with the anisotropic TV,

$$\sum_{i,j} \big(|(\nabla_x u)_{i,j}| + |(\nabla_y u)_{i,j}| \big).$$

The usual TV is invariant to rotation of the domain, but anisotropic TV is not. However, it allows for other approaches that do not apply with the usual TV, for example Hochbaum's exact algorithm [16] and graph-cuts [30].

As first proposed by Rudin, Osher, and Fatemi in [9], an alternative to discretizing the minimization problem directly is to discretize its gradient descent PDE. Through calculus of variations, the gradient descent PDE of the minimization is

$$\begin{cases} \partial_t u = \operatorname{div} \frac{\nabla u}{|\nabla u|} + \lambda(f - u), \\ \nu \cdot \nabla u = 0 \quad \text{on } \partial\Omega. \end{cases}$$



The sampling grid.

Since the problem is convex, the steady state solution of the gradient descent is the minimizer of the problem. Therefore, the minimizer can be obtained numerically by evolving a finite difference approximation of this PDE. An explicit scheme for this was developed in [9]. Let $u_{i,j}$ denote samples on a grid, $u_{i,j} := u(ih,jh)$, $i, j = 0,1,\ldots N$, $Nh = 1$. Gradient descent is performed by iterating

Algorithm 1.

$$u_{i,j}^{n+1} = u_{i,j}^n + dt\Big[\nabla_x^-\Big(\frac{\nabla_x^+ u_{i,j}^n}{\sqrt{(\nabla_x^+ u_{i,j}^n)^2 + (m(\nabla_y^+ u_{i,j}^n, \nabla_y^- u_{i,j}^n))^2}}\Big) + \nabla_y^-\Big(\frac{\nabla_y^+ u_{i,j}^n}{\sqrt{(\nabla_y^+ u_{i,j}^n)^2 + (m(\nabla_x^+ u_{i,j}^n, \nabla_y^- u_{i,j}^n))^2}}\Big)\Big] + dt\lambda(f_{i,j} - u_{i,j}^n), \quad i,j = 1,\ldots,N-1,$$

$$u_{0,j}^n = u_{1,j}^n, \quad u_{N,j}^n = u_{N-1,j}^n, \quad u_{i,0}^n = u_{i,1}^n, \quad u_{i,N}^n = u_{i,N-1}^n, \quad i,j = 0,\ldots,N,$$

where dt is a small positive timestep parameter. The discretization is symmetric through a balance of forward and backward differences. In the divisions, notice that the numerator is always smaller in magnitude than the denominator. In the special case that the denominator is

zero (where u is locally constant), the quotient is evaluated as 0/0 = 0. The second line imposes the zero Neumann boundary condition.

Instead of evolving the gradient descent, another approach taken for example with the digital TV filter [14] is to solve for the steady state directly:

$$0 = \operatorname{div} \frac{\nabla u}{|\nabla u|} + \lambda(f - u).$$

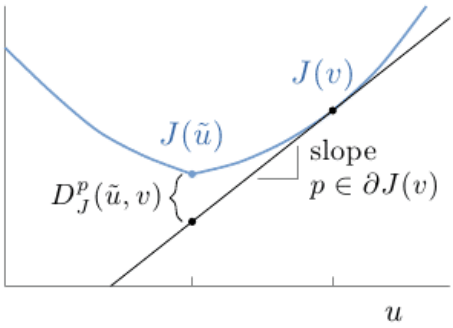
Many other algorithms for TV denoising have been developed, especially for the Gaussian noise model, and this continues to be an active area of research. Numerous algorithms have been proposed to solve the TV denoising minimization, too many to list them all here. To name a few, there are algorithms based on duality [18],[36],[44], Newton-based methods [27], graph cuts [30], and frame shrinkage [38]. Most recent methods employ operator splitting [6],[26],[29],[31],[33],[34],[41],[43], particularly the split Bregman algorithm discussed in the next few sections.

Bregman Iteration

Bregman iteration [1],[22],[28],[35] is a technique for solving constrained convex minimization problems of the form

$$\arg\min_u J(u) \quad \text{subject to} \quad H(u) = 0$$

where J and H are (possibly nondifferentiable) convex functionals on defined on a Hilbert space. We assume there exists u minimizing H for which $H(u) = 0$ and $J(u) < \infty$. The key idea is the *Bregman distance*.



The Bregman distance $D_J^p(\tilde{u}, v)$.

The Bregman distance is defined as

$$D_J^p(u, v) := J(u) - J(v) - \langle p, u - v \rangle, \quad p \in \partial J(v).$$

Bregman distance compares the value $J(u)$ with the tangent plane $J(v) + \langle p, u - v \rangle$. The figure above illustrates the distance in one dimension. The horizontal axis denotes u, the blue curve denotes $J(u)$, and the black line is the tangent plane $J(v) + \langle p, u - v \rangle$. Here, ∂J is the *subdifferential* of J [10], which is defined as

$$\partial J(v) := \{p : J(u) \geq J(v) + \langle p, u - v \rangle \ \forall u\}.$$

Bregman distance is not a distance in the usual sense because it is not symmetric. However, it

does satisfy other distance-like properties following from the definition of the distance and the convexity of J [22]:

- $D_J^p(v, v) = 0$
- $D_J^p(u, v) \geq 0$
- $D_J^p(u, v) + D_J^{\tilde{p}}(v, \tilde{v}) - D_J^{\tilde{p}}(u, \tilde{v}) = \langle p - \tilde{p}, v - u \rangle$.

Given a starting point u^0 and parameter $\gamma > 0$, the Bregman iteration algorithm is formally

Algorithm 2.

$$u^{k+1} = \arg\min_u D_J^{p^k}(u, u^k) + \gamma H(u), \quad p^k \in \partial J(u^k).$$

Existence of the solutions u^{k+1} is nontrivial if the search space is infinite dimensional. This is studied in [22], with particular attention to the case where J is total variation.

Because u^{k+1} minimizes $D_J^{p^k}(u, u^k) + \gamma H(u)$,

$$D_J^{p^k}(u^{k+1}, u^k) + \gamma H(u^{k+1}) \leq D_J^{p^k}(u^k, u^k) + \gamma H(u^k),$$

so the iteration has the property

$$\gamma H(u^{k+1}) \leq D_J^{p^k}(u^{k+1}, u^k) + \gamma H(u^{k+1}) \leq D_J^{p^k}(u^k, u^k) + \gamma H(u^k) = \gamma H(u^k),$$

s o $H(u^k)$ decreases monotonically. Some stronger convergence results under additional assumptions will be discussed shortly.

We will consider here the case when H is differentiable. In this case the sub-differential of H is its gradient ∇H, and the sub-differential of the Lagrangian is given by

$$\partial_u \big(J(u) - J(u^k) - \langle p^k, u - u^k \rangle + \gamma H(u) \big) = \partial J - p^k + \gamma \nabla H.$$

Since u^{k+1} minimizes $D_J^{p^k}(u, u^k) + \gamma H(u)$, the optimality condition is then

$$\begin{aligned} & 0 \in \partial J(u^{k+1}) - p^k + \gamma \nabla H(u^{k+1}) \\ \Leftrightarrow \quad & p^k - \gamma \nabla H(u^{k+1}) \in \partial J(u^{k+1}). \end{aligned}$$

Therefore, $p^{k+1} \in \partial J(u^{k+1})$ can be selected as $p^{k+1} = p^k - \gamma \nabla H(u^{k+1})$. Bregman iteration with this rule is

Algorithm 3.

$p^0 \in \partial J(u^0)$
for $k = 0, 1, \ldots$
$\quad u^{k+1} = \arg\min_u D_J^{p^k}(u, u^k) + \gamma H(u)$
$\quad p^{k+1} = p^k - \gamma \nabla H(u^{k+1})$

Suppose that H is differentiable and that the solutions u^{k+1} exist and are obtained by

Algorithm 3, then the following convergence results hold [22]: for any $\tilde{u}$ such that $H(\tilde{u}) = 0$ and $J(\tilde{u}) < \infty$,

$$D^{p^{k+1}}(\tilde{u}, u^{k+1}) \leq D^{p^k}(\tilde{u}, u^k)$$

and

$$H(u^k) \leq \frac{J(\tilde{u})}{\gamma k}.$$

Particularly, (u^k) is a minimizing sequence of H.

A remarkable feature of Bregman iteration is that the limiting solution satisfies the constraint $H(u) = 0$ exactly for any positive value of the parameter γ. The value of γ does, however, affect the convergence speed and numerical conditioning of the minimization problems, so γ should be selected according to these considerations.

A case of practical importance, including our application to TV denoising, is where u is in $\mathbb{R}^n$ with linear equality constraints. Let A be a matrix and set

$$H(u) = \tfrac{1}{2}\|Au - f\|_2^2,$$

then Bregman iteration simplifies [28],[31] to Algorithm 4. Furthermore, when the constraints are linear, Bregman iteration is equivalent [28] to the augmented Lagrangian method (also known as the method of multipliers) introduced by Hestenes [2] and Powell [3].

Algorithm 4.

$u^0 \in \mathbb{R}^n, b^0 = 0$
for $k = 0, 1, \ldots$
$\quad u^{k+1} = \arg\min_u J(u) + \frac{\gamma}{2}\|Au - f + b^k\|_2^2$
$\quad b^{k+1} = b^k + Au^{k+1} - f$

The subgradients p^k are represented by the auxiliary variables b^k, which are added within the quadratic penalty term. Jia, Zhao, and Zhao [35] proved that the above iteration converges to the solution of the constrained minimization problem for TV denoising for both the isotropic or anisotropic TV discretization.

Discrete Derivatives

We describe here a methodology for discrete derivatives and boundary handling of uniformly sampled functions. These discrete derivatives will be used in the denoising algorithm.

In one dimension, let $(f_n), n \in \mathbb{Z}$, denote uniformly-spaced samples of a bounded function f. We define the discrete first derivative of f as the forward difference

$$\partial f_n := f_{n+1} - f_n.$$

In two dimensions, the discrete gradient of $u_{i,j}, (i,j) \in \mathbb{Z}^2$, is defined as applying ∂ separately along the x and y dimensions,

$$\nabla u_{i,j} := \begin{pmatrix} \partial_x u_{i,j} \\ \partial_y u_{i,j} \end{pmatrix} = \begin{pmatrix} u_{i+1,j} - u_{i,j} \\ u_{i,j+1} - u_{i,j} \end{pmatrix}.$$

In analogy to the standard notation for continuous partial derivatives, the subscript on ∂ denotes along which dimension the difference is applied.

Note that the negative adjoint $-\partial^*$ is the backward difference,

$$\begin{aligned} \sum_{n\in\mathbb{Z}} (\partial^* f_n) g_n &:= \sum_{n\in\mathbb{Z}} f_n (\partial g_n) \\ &= \sum_{n\in\mathbb{Z}} f_n (g_{n+1} - g_n) \\ &= \sum_{n\in\mathbb{Z}} (f_{n-1} - f_n) g_n, \quad \forall g \in \ell^1, \\ \Rightarrow -\partial^* f_n &= f_n - f_{n-1}. \end{aligned}$$

We define discrete divergence through the relationship div $:= -\nabla^*$. For a vector field $\vec{v}_{i,j} = (v^x_{i,j}, v^y_{i,j})^T$,

$$\begin{aligned} \operatorname{div} \vec{v}_{i,j} &:= -\nabla^* v_{i,j} \\ &= -\partial^*_x v^x_{i,j} - \partial^*_y v^y_{i,j} \\ &= v^x_{i,j} - v^x_{i-1,j} + v^y_{i,j} - v^y_{i,j-1}. \end{aligned}$$

The discrete Laplacian follows from the relationship $\Delta :=$ div ∇,

$$\begin{aligned} \Delta u_{i,j} &:= \operatorname{div} \nabla u_{i,j} \\ &= -\partial^*_x \partial_x u_{i,j} - \partial^*_y \partial_y u_{i,j} \\ &= -4u_{i,j} + u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}. \end{aligned}$$

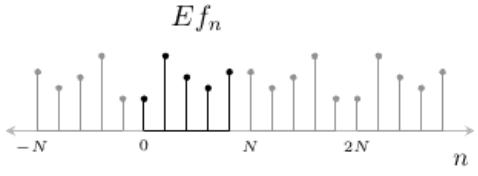
We now address the issue of boundary handling. Above, we defined discrete derivatives assuming samples f_n are available for all integer n. On a finite-length signal $f_0, f_1, \ldots, f_{N-1}$, the forward differences can be computed directly for n in the interior,

$$\partial f_n = f_{n+1} - f_n, \quad n = 0, \ldots, N-2.$$

However, the forward difference at the right endpoint $n = N - 1$ would require the unknown sample f_N. Special handling is needed on the boundaries.

Define the half-sample symmetric extension Ef,

$$Ef_n = \begin{cases} f_n & \text{if } n = 0, \ldots, N-1, \\ Ef_{-1-n} & \text{if } n < 0, \\ Ef_{N-1-n} & \text{if } n \geq N. \end{cases}$$



The definition is recursive since multiple reflections may be needed to obtain an index between 0 and $N - 1$. We also consider the tensor product of this extension applied to an $N \times N$ image $u_{i,j}$, $i = 0, \ldots, N-1, j = 0, \ldots, N-1$. We define the discrete derivative of finite-length f as the forward difference of Ef,

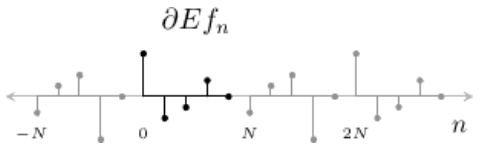
$$\partial f_n := Ef_{n+1} - Ef_n = \begin{cases} f_{n+1} - f_n & \text{if } n = 0, \dots, N-2, \\ 0 & \text{if } n = -1 \text{ or } N-1. \end{cases}$$

This discrete derivative may be viewed as an $N \times N$ matrix multiplication,

$$\begin{pmatrix} \partial f_0 \\ \partial f_1 \\ \vdots \\ \partial f_{N-2} \\ \partial f_{N-1} \end{pmatrix} = \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-2} \\ f_{N-1} \end{pmatrix}.$$

Noting that Ef is $2N$-periodic, the discrete gradient may also be viewed as a cyclic convolution of the reflected signal $(f_0, \ \dots, f_{N-1}, f_{N-1}, \ \dots, f_0)$ with the filter $h_{-1} = 1$, $h_0 = -1$, and h zero otherwise.

We define the discrete gradient of an $N \times N$ image u as $\nabla u = (\partial_x u, \partial_y u)^{\mathrm{T}}$. Due to the symmetric extension, ∂Ef is (whole-sample) anti-symmetric about the points $n = N - 1 \pmod N$:



Let g be such an anti-symmetric signal. Then $-\partial^* g$ is

$$\begin{pmatrix} -\partial^* g_0 \\ -\partial^* g_1 \\ \vdots \\ -\partial^* g_{N-2} \\ -\partial^* g_{N-1} \end{pmatrix} = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 0 \end{pmatrix} \begin{pmatrix} g_0 \\ g_1 \\ \vdots \\ g_{N-2} \\ g_{N-1} \end{pmatrix}.$$

This is the negative transpose of the matrix above for ∂. To explain the endpoints, note that g_{-1} and g_{N-1} are zero by the anti-symmetric property, which implies

$$\begin{aligned} -\partial^* g_0 &= g_0 - g_{-1} &&= g_0, \\ -\partial^* g_{N-1} &= g_{N-1} - g_{N-2} &&= -g_{N-2}. \end{aligned}$$

Similarly in two dimensions, we define the discrete divergence of an $N \times N$ vector field $v = (v^x, v^y)^{\mathrm{T}}$ as $\operatorname{div} \vec{v} = -\partial_x^* v^x - \partial_y^* v^y$.

Finally, second differences are obtained as

$$-\partial^*\partial f_n := -\partial^*\partial E f_n = \begin{cases} f_1 - f_0 & \text{if } n = 0, \\ f_{n+1} - 2f_n + f_{n-1} & \text{if } n = 1, \ldots, N-2, \\ f_{N-2} - f_{N-1} & \text{if } n = N-1, \end{cases}$$

$$\begin{pmatrix} -\partial^*\partial f_0 \\ -\partial^*\partial f_1 \\ \vdots \\ -\partial^*\partial f_{N-2} \\ -\partial^*\partial f_{N-1} \end{pmatrix} = \begin{pmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-2} \\ f_{N-1} \end{pmatrix}.$$

We define the discrete Laplacian of an $N \times N$ image as $\Delta u = -\partial_x^*\partial_x u - \partial_y^*\partial_y u$. In the image interior, this is the 5-point Laplacian $-4u_{i,j} + u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}$.

Split Bregman for Gaussian Noise

Here we focus on the split Bregman algorithm of Goldstein and Osher [31]. Split Bregman is a flexible algorithm for solving nondifferentiable convex minimization problems, and it is especially efficient for problems with L^1 or TV regularization. Goldstein and Osher [31] discuss in particular its application to TV-regularized Gaussian denoising. It is easy to extend to other noise models (described in a later section) and related problems like TV-regularized deblurring and inpainting [37],[46],[47].

Total variation is approximated by summing the vector magnitude $|\nabla u_{i,j}|$ over all pixels,

$$\|u\|_{\mathrm{TV}(\Omega)} \approx \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |\nabla u_{i,j}|,$$

where ∇u is the discrete gradient developed in the previous section. The split Bregman idea is to apply operator splitting and then use Bregman iteration to solve the resulting constrained minimization problem:

$$\begin{aligned} &\arg\min_{d,u} \sum_{i,j} |d_{i,j}| + \frac{\lambda}{2} \sum_{i,j} (f_{i,j} - u_{i,j})^2 \\ &\text{subject to } d = \nabla u. \end{aligned}$$

By introducing d, the first and second terms are not directly interacting. The split problem is solved using Bregman iteration as in Algorithm 4. In each iteration of the Bregman method, following problem is solved:

$$\arg\min_{d,u} \sum_{i,j} |d_{i,j}| + \frac{\lambda}{2} \sum_{i,j} (f_{i,j} - u_{i,j})^2 + \frac{\gamma}{2} \sum_{i,j} |d_{i,j} - \nabla u_{i,j} - b_{i,j}|^2$$

where b is a variable related to the Bregman iteration algorithm and the penalty parameter γ is a positive constant. Goldstein and Osher proposed to solve this problem by an alternating direction method [4],[5],[7], in each step minimizing either d or u while keeping the other variable fixed. Esser [32] showed that for linear constraints, split Bregman with this alternating direction method is equivalent to the alternating direction method of multipliers, which was introduced by Glowinski and Marocco [4] and Gabay and Mercier [5].

d subproblem. With u fixed, the d subproblem is

$$\arg\min_{d} \sum_{i,j} |d_{i,j}| + \frac{\gamma}{2} \sum_{i,j} |d_{i,j} - \nabla u_{i,j} - b_{i,j}|^2.$$

This problem decouples over space and has a closed-form solution as a vectorial shrinkage,

$$d_{i,j} = \frac{\nabla u_{i,j} + b_{i,j}}{|\nabla u_{i,j} + b_{i,j}|} \max\{|\nabla u_{i,j} + b_{i,j}| - 1/\gamma, 0\}.$$

u subproblem. With d fixed, the u subproblem is

$$\arg\min_{u} \frac{\lambda}{2} \sum_{i,j} (u_{i,j} - f_{i,j})^2 + \frac{\gamma}{2} \sum_{i,j} |\nabla u_{i,j} - d_{i,j} + b_{i,j}|^2.$$

The optimal u satisfies a discrete screened Poisson equation,

$$\begin{aligned} \lambda(u - f) + \gamma \nabla^*(\nabla u - d + b) &= 0 \\ \lambda u - \gamma \Delta u &= \lambda f - \gamma \operatorname{div}(d - b), \end{aligned}$$

where $\operatorname{div} := -\nabla^*$ and $\Delta := \operatorname{div} \nabla$ are the discrete divergence and discrete Laplacian developed in the previous section.

The optimality equation may be solved for u in the Fourier or DCT domain or by iterative matrix techniques. In this work, we follow Goldstein and Osher's suggestion [31] to approximate the solution to this equation with one sweep of Gauss-Seidel per Bregman iteration. The subproblem is solved once for each Bregman iteration, so the combined effect of the sweeps over multiple iterations solves the subproblem accurately.

Updating b.

We enforce the constraint $d = \nabla u$ by applying Algorithm 4 with $H(u) = ½||\nabla u - d||^2$. The auxiliary variable b is initialized to zero and updated after each Bregman iteration as

$$b^{k+1} = b^k + \nabla u - d.$$

Selecting the penalty parameter γ.

As we discussed previously, Bregman iteration ensures that the limiting solution satisfies the constraint $H(u) = 0$ exactly for any positive value γ. Therefore, a good choice of γ is where both d and u subproblems converge quickly and are numerically well-conditioned.

In the d subproblem, the solution d is equal to $(\nabla u + b)$ after shrinking its vector magnitude by $1/\gamma$. This effect is more dramatic when γ is small.

The u subproblem behaves oppositely. The updated u is found by solving

$$\lambda u - \gamma \Delta u = \lambda f - \gamma \operatorname{div}(d - b).$$

The effect of the subproblem increases when γ is larger because the Δu term creates stronger spatial interaction and d has more influence on the solution. However, the conditioning also

worsens as γ increases and is ill-conditioned in the limit $\gamma \to \infty$.

Therefore, γ should be neither extremely larger nor small for good convergence. In the examples, we fix $\gamma = 5$. We have found that the algorithm is fairly insensitive to the exact value of γ.

The overall algorithm is

Algorithm 5.

```
Initialize u = 0, d = 0, b = 0
while ||u_cur - u_prev||_2 > Tol
   Solve the d subproblem
   Solve the u subproblem
   b = b + ∇u - d
```

where the solutions of the subproblems are as developed above. In the xth subproblem ($x = d$ or u), the solution is computed using the current values of the other variables and overwrites the previous value for x. Convergence is tested by the mean square difference between the current and previous iterate of u. In the implementation, the default parameter values are $Tol = ||f||_2/1000$ and $\gamma = 5$.

A similar algorithm to split Bregman is the FTVd algorithm by Wang, Yang, Yin, and Zhang [29]. In FTVd, operator splitting and alternating direction minimization is applied in the same way as in split Bregman. But instead of using Bregman iteration, FTVd enforces the constraint by gradually increasing the penalty parameter γ in a continuation scheme. The downside of continuation schemes is that penalty parameter may need to become very large to satisfy the constraint accurately, which degrades the numerical conditioning and convergence speed. Bregman iteration avoids these problems because γ stays fixed. On the other hand, FTVd is advantageous in TV-regularized deconvolution, where it requires one fewer FFT transform per iteration than split Bregman.

For color images, the vectorial TV (VTV) is used in place of TV,

$$\|u\|_{\mathrm{VTV}(\Omega)} := \int_\Omega \Big(\sum_{i \in \mathrm{channels}} |\nabla u_i(x)|^2 \Big)^{1/2} dx.$$

The grayscale algorithm extends directly to VTV-regularized denoising.

Tuning λ

The choice of the parameter λ affects how much the image is regularized, balancing between removing the noise and preserving the signal content. Parameter tuning can generally be approached as a meta-optimization where λ is selected to optimize some criterion of the denoising result. A straightforward method for parameter tuning is the discrepancy principle: λ is selected to match the noise variance σ^2. For TV denoising, the discrepancy principle suggests to solve a constrained form of the ROF problem

$$\arg\min_u \|u\|_{\mathrm{TV}(\Omega)} \quad \text{subject to} \int_\Omega \big(f(x) - u(x)\big)^2\, dx = \sigma^2 |\Omega|.$$

The discrepancy principle has an observed tendency to overestimate the mean squared error optimal choice of λ and slightly over-smoothing the solution, see for example [8]. We nevertheless follow it here as a simple automatic selection of the parameter.

Let $\langle f \rangle$ denote the mean value of f. We assume that the variance of f is at least as large as the noise level

$$\int_\Omega \big(f(x) - \langle f \rangle\big)^2\, dx \geq \sigma^2 |\Omega|,$$

which is likely to be true since f is supposed to have additive noise of variance σ^2. Under this condition, the problem is equivalent to the unconstrained minimization

$$\arg\min_u \|u\|_{\mathrm{TV}(\Omega)} + \frac{\lambda}{2} \int_\Omega \big(f(x) - u(x)\big)^2\, dx,$$

with λ as the Lagrangian multiplier for the constraint. There exists a unique value of λ for which the minimizers of the two problems are the same. Unfortunately, the relationship between σ and λ is indirect; there is no closed-form formula to obtain the value of λ corresponding to a particular σ.

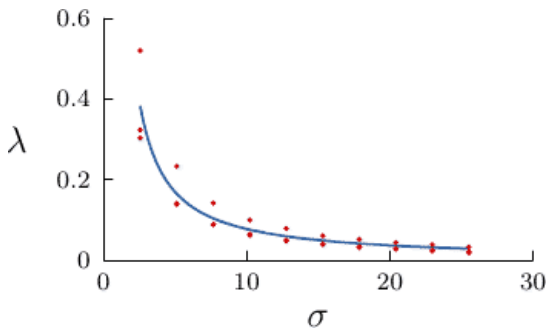
While there are some algorithms that can solve the constrained problem directly with σ, most algorithms solve the unconstrained version with λ. To find a value of λ so that $||f-u||_2{}^2$ is approximately σ^2, an effective algorithm proposed by Chambolle [18] is

Algorithm 6.

```
Iterate
   u = argmin_u ||u||_TV + λ/2 ||f - u||_2^2
   λ = λ ||f - u||_2/σ
```

The sequence of λ produced by this iteration is proven to converge monotonically to the unique λ such that $||f-u||_2{}^2 = \sigma^2$. We initialize the iteration with the following empirical estimate of λ,

$$\lambda_0 = \frac{0.7079}{\sigma} + \frac{0.6849}{\sigma^2},$$



Red: optimal λ values for three images. Blue: empirical estimate.

where σ is the noise standard deviation relative to the intensity range [0,255]. The iteration solves the unconstrained problem with the current estimate of λ and then updates λ according to $||f-u||_2$. To speed up the minimizations, note that the u computed in one iteration can be used

as the initial guess in the following iteration.

The iteration converges quickly for most images and noise levels. We perform five iterations to tune λ, which is sufficiently accurate so that $||f-u||_2$ is usually within 10% of σ.

Split Bregman for Laplace and Poisson Noise

For a general noise model, TV-regularized denoising takes the form

$$\arg\min_u \|u\|_{\mathrm{TV}(\Omega)} + \lambda \int_\Omega F\big(u(x), f(x)\big)\, dx,$$

where F specifies the noise model,

$$F\big(u(x), f(x)\big) = \begin{cases} |u(x) - f(x)| & \text{Laplace noise,} \\ u(x) - f(x)\log u(x) & \text{Poisson noise.} \end{cases}$$

The split Bregman algorithm may be applied if the problem is convex, which is the case with the Laplace and Poisson noise models. As developed in [37], a splitting with two auxiliary variables can be used to separate F from the derivative terms,

$$\begin{aligned} &\arg\min_{d,z,u} \sum_{i,j} |d_{i,j}| + \lambda \sum_{i,j} F(z_{i,j}, f_{i,j}) \\ &\text{subject to } d = \nabla u,\ z = u. \end{aligned}$$

In each iteration of the Bregman method, following problem is solved:

$$\begin{aligned} &\arg\min_{d,z,u} \sum_{i,j} |d_{i,j}| + \lambda \sum_{i,j} F(z_{i,j}, f_{i,j}) \\ &\quad + \frac{\gamma_1}{2} \sum_{i,j} |d_{i,j} - \nabla u_{i,j} - b^1_{i,j}|^2\, dx + \frac{\gamma_2}{2} \sum_{i,j} (z_{i,j} - u_{i,j} - b^2_{i,j})^2 \end{aligned}$$

where b^1 and b^2 are variables related to the Bregman iteration. As in the Gaussian case, this problem is solved by minimizing one variable at a time with the other two fixed.

d subproblem. With z and u fixed, the d subproblem is the same as before

$$d_{i,j} = \frac{\nabla u_{i,j} + b^1_{i,j}}{|\nabla u_{i,j} + b^1_{i,j}|} \max\big\{|\nabla u_{i,j} + b^1_{i,j}| - 1/\gamma_1, 0\big\}.$$

z subproblem. With d and u fixed, the z subproblem is

$$\arg\min_z \lambda \sum_{i,j} F(z_{i,j}, f_{i,j}) + \frac{\gamma_2}{2} \sum_{i,j} (z_{i,j} - u_{i,j} - b^2_{i,j})^2.$$

The solution decouples over i,j and the optimal z satisfies

$$\lambda \partial_z F(z, f) + \gamma_2 (z - u - b^2) = 0.$$

For the Laplace noise model with $F(z,f) = |z - f|$, the solution is

$$z_{i,j} = f_{i,j} + \operatorname{sign} s_{i,j} \max\{|s_{i,j}| - \tfrac{\lambda}{\gamma_2}, 0\},$$
$$s = u - f + b^2.$$

For the Poisson noise model with $F(z,f) = z - f \log z$, the solution is

$$z_{i,j} = s_{i,j}/2 + \sqrt{(s_{i,j}/2)^2 + \tfrac{\lambda}{\gamma_2} f_{i,j}},$$
$$s = u - \tfrac{\lambda}{\gamma_2} + b^2.$$

u subproblem. With d and z fixed, the u subproblem is

$$\arg\min_u \frac{\gamma_1}{2} \sum_{i,j} |\nabla u_{i,j} - d_{i,j} + b^1_{i,j}|^2 + \frac{\gamma_2}{2} \sum_{i,j} (u_{i,j} - z_{i,j} + b^2_{i,j})^2.$$

The optimal u satisfies

$$\gamma_2 u - \gamma_1 \Delta u = \gamma_2 (z - b^2) - \gamma_1 \operatorname{div}(d - b^1),$$

which as before is approximated by one sweep of Gauss-Seidel iteration.

The overall algorithm is

Algorithm 7.

```
Initialize u = 0, z = 0, b2 = 0, d = 0, b1 = 0
while ||u_cur - u_prev||_2 > Tol
   Solve the d subproblem
   Solve the u subproblem
   Solve the z subproblem
   b1 = b1 + ∇u - d
   b2 = b2 + u - z
```

In the implementation, the default parameter values are $Tol = ||f||_2/1000$, $\gamma_1 = 5$, $\gamma_2 = 8$.

As with Gaussian denoising, λ can be selected according to the discrepancy principle to match the noise standard deviation. While there is no theoretical guarantee of convergence in this case, we find that iterations similar to Algorithm 6 also work with Laplace and Poisson noise.

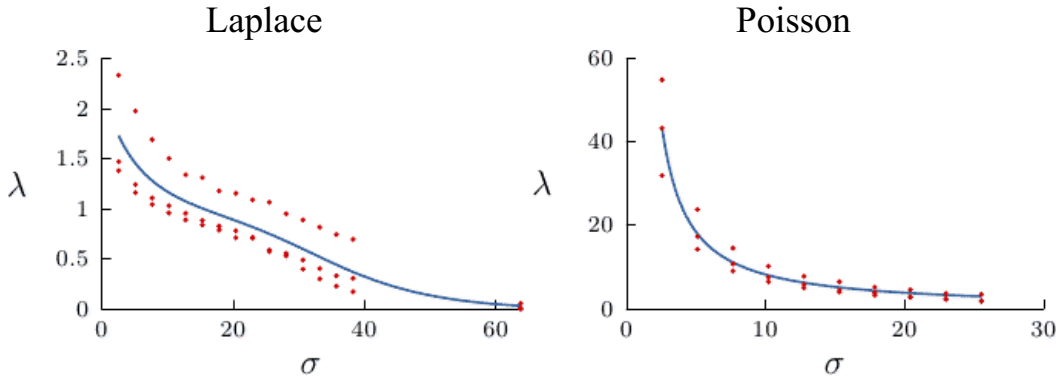
λ tuning for Laplace noise

```
λ = (-270.5 σ + 21572) / (σ^3 - 52.07 σ^2 + 1063 σ + 9677)
Iterate
   u = argmin_u ||u||_TV + λ ||f - u||_1
   λ = λ sqrt(||f - u||_2/σ)
```

λ tuning for Poisson noise

```
λ = 72.39/σ + 97.67/σ^2
Iterate
   u = argmin_u ||u||_TV + λ ∫ ((u - f)log u)
   λ = λ ||f - u||_2/σ
```

Empirical estimates are used to initialize λ. For Laplace noise, the sequence of λ tends to oscillate, so a square root is included in the update formula to dampen the oscillations.



Red: optimal λ values for three images. Blue: empirical estimates.

Implementation

This software is distributed under the terms of the simplified BSD license.

- source code zip tar.gz
- online documentation

Please see the `readme.txt` file inside or the online documentation for details.

Future software releases and updates will be posted at http://dev.ipol.im/~getreuer/code.

Examples

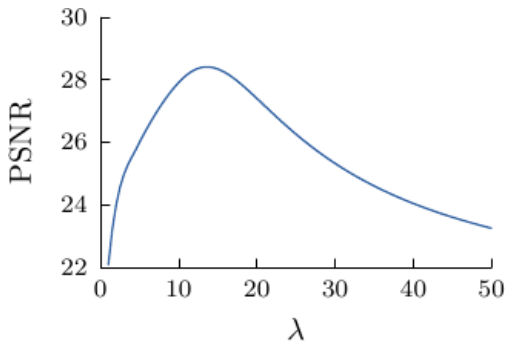
The first example demonstrates how for TV-regularized Gaussian denoising the value of λ influences the result. A smaller value of λ implies stronger denoising. When λ is very small, the image becomes cartoon-like with sharp jumps between nearly flat regions. The λ parameter needs to be balanced to remove noise without losing too much signal content.

Input f (PSNR 20.15) $\lambda = 5$ (PSNR 26.00) $\lambda = 10$ (PSNR 27.87)

$\lambda = 20$ (PSNR 27.34) $\lambda = 40$ (PSNR 24.01)

TV-regularized denoising with increasing values of λ.

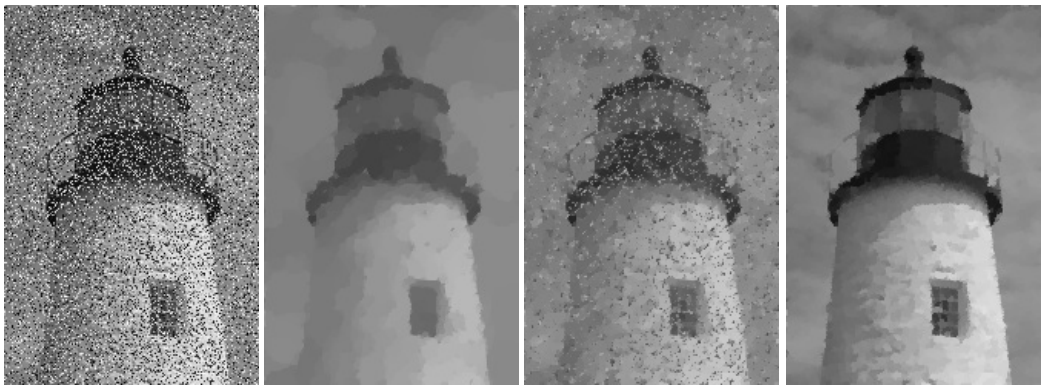
The plot shows the PSNR vs. λ for the previous example. The optimal λ in this case is about 13.4.



PSNR vs. λ for the previous example.

To illustrate the importance of the noise model, the image in this example has been corrupted with impulsive noise. The Gaussian noise model works poorly on impulsive noise: λ must be very small to remove all the noise, but this also removes much of the signal content. Better results are obtained with the Laplace noise model, which better approximates the distribution of impulsive noise.

Input f (PSNR 13.26) | Gaussian, $\lambda = 4$ (PSNR 20.28) | Gaussian, $\lambda = 8$ (PSNR 19.70) | Laplace, $\lambda = 1.25$ (PSNR 25.85)



The Laplace model is more effective for removing impulsive noise.

The next example demonstrates VTV-regularized Gaussian denoising on a color image.

Exact

Input f (PSNR 18.57) Denoised u with $\lambda = 7$ (PSNR 28.24)

A problem with TV regularization is a loss of contrast. Suppose that f has value h within a disk or radius r and is 0 outside,

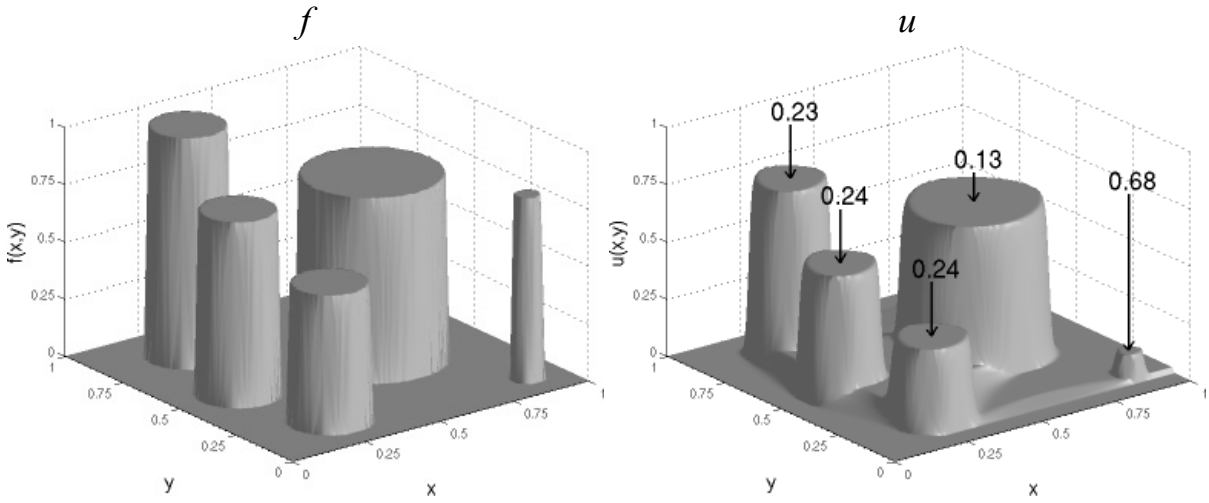
$$f(x) = \begin{cases} h & \text{if } |x| \le r, \\ 0 & \text{otherwise.} \end{cases}$$

Then if $\Omega = \mathbb{R}^2$, Meyer [15] showed that TV-regularized Gaussian denoising decreases the value within the disk by $2/(\lambda r)$,

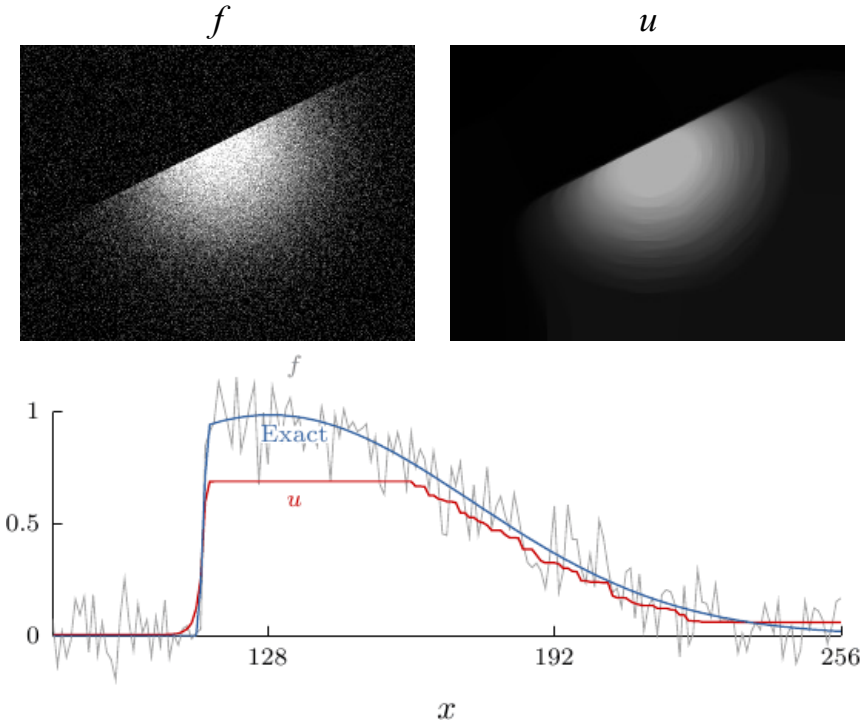
$$u(x) = \begin{cases} \max\{0, h - \frac{2}{\lambda r}\} & \text{if } |x| \le r, \\ 0 & \text{otherwise.} \end{cases}$$

Note that although the contrast is diminished, the edge of the circle is maintained exactly. Strong and Chan [11] made a similar analytic investigation under the assumption that the edges do not move, and showed that the behavior is similar when Ω is compact and for shapes other than disks.

The figure below verifies the diminishing contrast numerically. The image contains three disks of radius 0.11 with different heights, one large disk of radius 0.2, and a small disk of radius 0.04. The solution shows the decrease in value of each disk. For the three disks of radius 0.11, the decrease is almost the same, despite their differing heights, and the decrease is smaller for the large disk and larger for the small disk.



Noted by Nikolova [13], another problem with TV regularization is the "staircase artifact," a tendency to produce small flat regions with artificial edges. This effect is demonstrated below. The exact image has one jump along the center and is otherwise smooth. The plot shows a cross section of the images to visualize the stair steps. The loss of contrast effect is also visible in the peak of the dome.



Chan et al. [21] discuss solutions that have been developed to reduce the loss of contrast and

staircasing artifacts.

References

1. L.M. Bregman, "The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex optimization," USSR Computational Mathematics and Mathematical Physics, 7:200–217, 1967.
2. M. R. Hestenes, "Multiplier and Gradient Methods," Journal of Optimization Theory and Applications, vol. 4, pp. 303–320, 1969.
3. M. J. D. Powell, "A Method for nonlinear Constraints in Minimization Problems," in Optimization, R. Fletcher, ed., Academic Press, New York, pp. 283–298, 1969.
4. R. Glowinski, A. Marrocco, "Sur l'approximation par éléments finis d'ordre un, et la résolution par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires," Rev. Francaise dAut. Inf. Rech. Oper., R-2, pp. 41–76, 1975.
5. D. Gabay, B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," Computers & Mathematics with Applications, vol. 2, no. 1, pp. 17–40, 1976.
6. P.-L. Lions, B. Mercier, "Splitting Algorithms for the Sum of Two Nonlinear Operators," SIAM J. Numerical Analysis, vol. 16, pp. 964–979, 1979.
7. D. P. Bertsekas, "Constrained Optimization and Lagrange Multiplier Methods," Academic Press, New York, 1982.
8. N.P. Galatsanos, A.K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," IEEE Transactions on Image Processing, vol. 1, no. 3, pp. 322–336, 1992.
9. L.I. Rudin, S. Osher, E. Fatemi, "Nonlinear total variation based noise removal algorithms." Physica D, vol. 60, pp. 259–268, 1992.
10. J.-B. Hiriart-Urruty, C. Lemaréchal, "Convex Analysis and Minimization Algorithms I, II" (two parts), vol. 305–306 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, 1993. ISBN: 3540568506 and 3642081622.
11. D. M. Strong, T. F. Chan, "Exact Solutions to Total Variation Regularization Problems," UCLA CAM Report 96-41, 1996.
12. S. Alliney, "A property of the minimum vectors of a regularizing functional defined by means of the absolute norm," IEEE Trans. on Signal Processing, vol. 45, no. 4, pp. 913–917, 1997.
13. M. Nikolova, "Local Strong Homogeneity of a Regularized Estimator," SIAM Journal on Applied Mathematics, vol. 61, no. 2, pp. 633–658, 2000.
14. T.F. Chan, S. Osher, J. Shen. "The digital TV filter and nonlinear denoising," IEEE Transactions on Image Processing, vol. 10, no. 2, pp. 231–241, 2001.
15. Y. Meyer, "Oscillating Patterns in Image Processing and Nonlinear Evolution Equations," University Lecture Series, vol. 22, AMS Providence, RI, 2001. ISBN: 0821829203.
16. D. S. Hochbaum, "An efficient algorithm for image segmentation, Markov random fields and related problems," Journal of the ACM, vol. 48, no. 4, pp. 686–701, 2001.
17. L. Rudin, P.-L. Lions, S. Osher, "Multiplicative Denoising and Deblurring: Theory and Algorithms," Geometric Level Set Methods in Imaging, Vision, and Graphics, Part III, pp. 103–119, 2003.
18. A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," JMIV, 20(1-2), 2004.
19. T.F. Chan, S. Esedoḡlu. "Aspects of total variation regularized L^1 function

approximation." SIAM Journal on Applied Mathematics, vol. 65, no. 5, pp. 1817–1837, 2005.
20. T.F. Chan, J. Shen. Image Processing and Analysis: Variational, PDE, wavelet, and Stochastic Methods. SIAM, 2005. ISBN: 089871589X.
21. T.F. Chan, S. Esedoglu, F. Park, A.M. Yip, "Recent developments in total variation image restoration," Mathematical Models of Computer Vision, Springer Verlag, 2005.
22. S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin, "An Iterative Regularization Method for Total Variation-Based Image Restoration," Multiscale Model. Simul., vol. 4, no. 2, pp. 460–489, 2005.
23. G. Gilboa, N. Sochen, Y. Y. Zeevi, "Variational denoising of partly textured images by spatially varying constraints," IEEE Trans. Image Processing, vol. 15, no. 8, pp. 2281–2289, 2006.
24. A. Haddad. "Stability in a Class of Variational Methods." Appl. Comput. Harmon. Anal. 23, pp. 57–73, 2007.
25. T. Le, R. Chartrand, T. Asaki. "A Variational Approach to Constructing Images Corrupted by Poisson Noise," JMIV, vol. 27(3), pp. 257–263, 2007.
26. C. Chaux, P. L. Combettes, J.-C. Pesquet, V. R. Wajs, "A variational formulation for frame-based inverse problems," Inverse Problems, vol. 23, no. 4, 2007.
27. M. K. Ng, L. Qi, Y.-F. Yang, Y.-M. Huang, "On Semismooth Newton's Methods for Total Variation Minimization," Journal of Mathematical Imaging and Vision, vol. 27, no. 3, pp. 265–276, 2007.
28. W. Yin, S. Osher, D. Goldfarb, J. Darbon. "Bregman Iterative Algorithms for l_1 Minimization with Applications to Compressed Sensing." SIAM Journal on Imaging Sciences, vol. 1, no. 1, pp. 143–168, 2008.
29. Y. Wang, J. Yang, W. Yin, Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," SIAM J. Imaging Sci., vol. 1, no. 3, pp. 248–272, 2008.
30. A. Chambolle, J. Darbon. "On Total Variation Minimization and Surface Evolution using Parametric Maximum Flows," IJCV, vol. 84, no. 3, pp. 288–307, 2009.
31. T. Goldstein, S. Osher, "The Split Bregman Method for L1 Regularized Problems," SIAM J. Imaging Sci., vol. 2, no. 2, pp. 323–343, 2009.
32. E. Esser, "Applications of Lagrangian-Based Alternating Direction Methods and Connections to Split Bregman," UCLA CAM Report 09-21, 2009.
33. A. Beck, M. Teboulle, "Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems," IEEE Transactions on Image Processing, vol. 18, no. 11, pp. 2419–2434, 2009.
34. T. Pock, A. Chambolle, D. Cremers, H. Bischof, "A convex relaxation approach for computing minimal partitions," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 810–817, 2009.
35. R.Q. Jia, H. Zhao, W. Zhao, "Convergence analysis of the Bregman method for the variational model of image denoising," Applied and Computational Harmonic Analysis, vol. 27, no. 3, pp. 367–379, 2009.
36. M. Zhu, S.J. Wright, T.F. Chan, "Duality-based algorithms for total-variation-regularized image restoration," J. Comput. Optim. Appl., vol. 47, no. 3, pp. 377–400, 2010.
37. P. Getreuer, tvreg v2: Variational Imaging Methods for Denoising, Deconvolution, Inpainting, and Segmentation, 2010.
38. S. Setzer. "Operator Splittings, Bregman Methods and Frame Shrinkage in Image Processing," International Journal of Computer Vision, vol. 92, no 3, pp. 265–280, 2011.
39. A. Chambolle, S. E. Levine, B. J. Lucier, "An Upwind Finite-Difference Method for Total

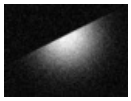
Variation–Based Image Smoothing,“ SIAM J. Imaging Sci., vol. 4, pp. 277–299, 2011.
40. J. Wang, B. J. Lucier, “Error bounds for finite-difference methods for Rudin-Osher-Fatemi image smoothing,“ SIAM J. Numer. Anal., vol. 49, no. 2, pp. 845–868, 2011.
41. S. Becker, J. Bobin, E. J. Candès, “NESTA: A Fast and Accurate First-Order Method for Sparse Recovery,” SIAM J. Imaging Sci., vol. 4, no. 1, pp. 1–39, 2011.
42. P. Getreuer, M. Tong, L. Vese, “A Variational Model for the Restoration of MR Images Corrupted by Blur and Rician Noise,” Proc. ISVC, Part I, LNCS 6938, 2011.
43. M. Tao, X. M. Yuan, “An inexact parallel splitting augmented Lagrangian methods for monotone variational inequalities with separable structures,” Computational Optimization and Applications, pp. 1–23, 2011.
44. C. Couprie, L. Grady, H. Talbot, L. Najman, “Combinatorial Continuous Max Flow,” SIAM J. Imaging Sci. 4, pp. 905–930, 2011.
45. Y. Dong, T. Zeng, “A Convex Variational Model for Restoring Blurred Images with Multiplicative Noise,” UCLA CAM Report 12–23, 2012.
46. P. Getreuer, “Total Variation Deconvolution using Split Bregman,” Image Processing On Line, submitted.
47. P. Getreuer, “Total Variation Inpainting using Split Bregman,” Image Processing On Line, submitted.

This material is based upon work supported by the National Science Foundation under Award No. DMS-1004694. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Work partially supported by the MISS project of Centre National d'Etudes Spatiales, the Office of Naval Research under grant N00014-97-1-0839 and by the European Research Council, advanced grant “Twelve labours.”

image credits

 Kodak Image Suite, image 19

 Kurt Getreuer

 Pascal Getreuer