Nonlinear Inverse Scale Space Methods for Image Restoration

Martin Burger[1], Stanley Osher[2], Jinjun Xu[2] and Guy Gilboa[2]

[1] Industrial Mathematics Institute, Johannes Kepler University,
Altenbergerstr. 69, A 4040 Linz, Austria.
martin.burger@jku.at
[2] Department of Mathematics, UCLA, Los Angeles, CA 90095, USA.
{sjo, jjxu , gilboa}@math.ucla.edu

Abstract. In this paper we generalize the iterated refinement method, introduced by the authors in [8], to a time-continuous inverse scale-space formulation. The iterated refinement procedure yields a sequence of convex variational problems, evolving toward the noisy image.
The inverse scale space method arises as a limit for a penalization parameter tending to zero, while the number of iteration steps tends to infinity. For the limiting flow, similar properties as for the iterated refinement procedure hold. Specifically, when a discrepancy principle is used as the stopping criterion, the error between the reconstruction and the noise-free image decreases until termination, even if only the noisy image is available and a bound on the variance of the noise is known.
The inverse flow is computed directly for one-dimensional signals, yielding high quality restorations. In higher spatial dimensions, we introduce a relaxation technique using two evolution equations. These equations allow accurate, efficient and straightforward implementation.

1 Introduction

The processing of noisy images is a central task in mathematical imaging. Over the last decades, a variety of methods have been proposed ranging from filtering methods to variational approaches to techniques based on the solution of partial differential equations. Since the noise in images is usually expected to be a small scale feature, particular attention has been paid to methods separating scales, in particular those smoothing small scale features faster than large scale ones, so-called *scale space methods*.

Scale space methods are obtained for example by nonlinear diffusion filters [9] of the form

$$\frac{\partial u}{\partial t} = \mathrm{div}(\gamma(|\nabla u|^2)\nabla u), \tag{1}$$

in $\Omega \times \mathbb{R}_+$ with $u(x,0) = f(x)$, where $f : \Omega \to \mathbb{R}$ denotes the given image intensity (Ω being a bounded open subset in $\mathbb{R}^2$) and $u : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ the flow of smoothed images. The diffusion coefficient involves a positive and monotone function γ. For such methods it can be shown that small scales are smoothed

faster than large ones, so if the method is stopped at a suitable final time, we may expect that noise is smoothed while large-scale features are preserved to some extent. Diffusion filters can be related to regularization theory (cf. [13]) with certain regularization functionals, but foundations of choosing optimal stopping times are still missing.

Recently, *inverse scale space methods* have been introduced in [12], which are based on a different paradigm. Instead of starting with the noisy image and gradually smoothing it, inverse scale space methods start with the image $u(x,0) = 0$ and approach the noisy image f as time increases, with large scales converging faster than small ones. Thus, if the method is stopped at a suitable time, large scale features may already be incorporated into the reconstruction, while small scale features (including the "noise") are still missing. The inverse scale space method can also be related to regularization theory, in particular iterated Tikhonov regularization (cf. [6, 12]) with the same regularization functionals as for diffusion filters. The construction of inverse scale space methods in [12] worked well for quadratic regularization functionals, but did not yield convincing results for other important functionals, in particular for the total variation. In this paper we present a different version of constructing inverse scale space methods as the limit of an iterated refinement procedure previously introduced by the authors (cf. [8]) and demonstrate its applicability to image restoration. With the new approach we are able to perform inverse scale space methods even for the total variation functional, and, in contrast to diffusion filters, we obtain a simple stopping criterion for the methods.

2 Iterated Refinement

In [8], an iterated refinement procedure for total variation restoration was introduced, motivated by the variational problem

$$u = \operatorname*{argmin}_{u \in BV(\Omega)} \left\{ |u|_{BV} + \frac{\lambda}{2} \|f - u\|_{L^2}^2 \right\} \tag{2}$$

for some scale parameter $\lambda > 0$, where $BV(\Omega)$ denotes the space of functions with bounded variation on Ω, equipped with BV seminorm which is formally given by

$$|u|_{BV} = \int_{\Omega} |\nabla u|,$$

also referred to as the *total variation* (TV) of u. This is the ROF model, introduced to the field of image restoration in [11].

In [8] the authors showed that an iterative procedure (which turned out to be equivalent to Bregman's relaxation method, cf. [1], and proximal point algorithms, cf. [3]) could be used to improve the quality of regularized solutions to inverse problems, based on regularization functionals as in (2). Given a convex functional $J(u)$, e.g., $J(u) = |u|_{BV}$, the *iterated refinement method* defines a sequence $\{u_k\}$ by:

– Set $u_0 = 0,\ p_0 = 0$;
– Given u_{k-1} and p_{k-1},
  • step 1: compute $u_k = \arg\min_u Q_k(u)$ with

$$Q_k : u \longmapsto J(u) - J(u_{k-1}) - \langle p_{k-1}, u - u_{k-1}\rangle + \frac{\lambda}{2}\|f - u\|^2_{L^2}, \qquad (3)$$

  where $\langle\cdot,\cdot\rangle$ denotes the usual duality product;
  • step 2: update the dual variable $p_k = p_{k-1} + \lambda(f - u_k)$.
– Increase k by 1 and continue.

The quantity p_k is identified with $\partial J(u_k)$, as discussed below. This procedure improves the quality of reconstruction for many problems with discontinuous solutions, e.g., deblurring and denoising of images (cf. [7, 8]) when $\|f - u\|^2_{L^2}$ is replaced by an appropriate fitting term for individual examples.

Note that the regularization term used in the first step is a so-called generalized *Bregman distance* between u and u_{k-1}, defined as follows,

$$D(u, v) = J(u) - J(v) - \langle u - v, p\rangle,\ p \in \partial J(v),$$

where $\partial J(v)$ is the subgradient of the convex functional $J(v)$. Note that the subgradient may contain more than one element if the functional J is not continuously differentiable, so that the distance would depend on the specific choice of the subgradient. However, we shall suppress the dependence on the subgradient in the notation below. Note that for strictly convex functionals the subgradient contains at most one element and $D(u, v)$ is a scalar distance, that is strictly positive for $u \neq v$. We can then rewrite the functional Q_k minimized in each iteration step as: $Q_k(u) = D(u, u_{k-1}) + \frac{\lambda}{2}\|f - u\|^2_{L^2}$. The Bregman distance and the associated iteration were not used in this fashion previously, but they have been rather employed to minimize functions $H(u, f)$ where H is a (usually complicated) convex function of u having a unique minimum (cf. e.g., [3]).

It was shown in [8] that the iterated refinement method yields a well-defined sequence of minimizers u_k and subgradients $p_k \in \partial J(u_k)$. Moreover, it was proved that the sequence $\{u_k\}$ satisfies $\|u_k - f\|^2_{L^2} \leq \|u_{k-1} - f\|^2_{L^2}$ and if $f \in BV(\Omega)$, then $\|u_k - f\|^2_{L^2} \leq \frac{J(f)}{k}$, i.e., u_k converges monotonically to f in $L^2(\Omega)$ with a rate of $\frac{1}{\sqrt{k}}$. Of course, this convergence result does not give particular information on the behavior of the method as a denoising method, in particular for the typical case of a noisy image f.

The key denoising result obtained in [8] is as follows: for $g \in BV(\Omega)$ we have

$$D(g, u_k) < D(g, u_{k-1}) \quad \text{if } \|f - u_k\|_{L^2} \geq \tau\|g - f\|_{L^2} \qquad (4)$$

for any $\tau > 1$. Thus, the distance between a restored image u_k and a possible exact image g is decreasing until the L^2-distance of f and u_k is larger than the L^2-distance of f and g. This result can be used to construct a stopping rule for our iterative procedure. If we have an estimate of the variance of the noise, i.e., $f = g + n, \|n\|_{L^2} = \sigma$, where $g \in BV(\Omega)$ is the noise-free image and n is the

noise, then we can stop the iteration at the first k for which $\|f - u_{k+1}\|_{L_2} < \tau\sigma$. The choice of τ allows some freedom to apply the stopping rule also in the case when we only know an upper bound for σ.

It is interesting to note that for TV-based denoising where $J(u) = |u|_{BV}$, the sequence $\{u_k\}$ has the following interpretation (cf. [8]):

- Let $u_1 = \arg\min J(u) + \frac{\lambda}{2}\|f - u\|_{L^2}^2$;
- Define $v_1 = f - u_1$;
- Then inductively for $k \geq 2$, let

$$u_k = \arg\min \left\{ J(u) + \frac{\lambda}{2}\|f + v_{k-1} - u\|_{L^2}^2 \right\}$$

and $f + v_{k-1} = u_k + v_k$.

In other words, we *add* the "small scales" v_{k-1} back to f and perform ROF minimization with f replaced by $f + v_{k-1}$ and decompose this function into "large scales" (u_k) plus "small scales" (v_k). This interpretation already yields a multiscale interpretation of the method, since the "small scales" are somehow doubled in each step and so their larger parts can be incorporated into the large scale part after the next iteration. A related procedure involving the ROF model using Tikhonov-Morozov rather than Bregman iteration which multiplies λ by two in each step yields a multiscale method suggested in [6] and analyzed in [14].

3 Inverse Scale Space Methods

In the following we generalize the concept of inverse scale space theory introduced in [6, 12] in the context of Tikhonov regularization for the case

$$J(u) = \frac{1}{2}\int_\Omega |\nabla u|^2. \tag{5}$$

We shall derive general inverse scale space methods as a limit of the iterated refinement procedure for $\lambda \to 0$, concentrating in particular on the functional

$$J(u) = \int_\Omega \sqrt{|\nabla u|^2 + \epsilon^2}. \tag{6}$$

Recall that for a special $\lambda > 0$ the iterative refinement procedure constructs sequences u_k^λ of primal and p_k^λ of dual variables such that $u_0^\lambda = p_0^\lambda = 0$,

$$\begin{aligned} u_k^\lambda &= \arg\min_{u \in BV(\Omega)} \left\{ D(u, u_{k-1}^\lambda) + \frac{\lambda}{2}\|f - u\|_{L^2}^2 \right\} \\ p_k^\lambda &= \partial J(u_k^\lambda) = J'(u_k^\lambda). \end{aligned}$$

From the Euler-Lagrange equation $J'(u_k^\lambda) - p_{k-1}^\lambda + \lambda(u_k^\lambda - f) = 0$ and $p_k^\lambda = J'(u_k^\lambda)$ we are led to the relation:

$$\frac{p_k^\lambda - p_{k-1}^\lambda}{\lambda} = f - u_k, \quad k = 1, 2, \ldots$$

for the updates. We now reinterpret $\lambda = \Delta t$ as a time step and the difference quotient on the left-hand side as an approximation of a time derivative. Setting $t_k = k\Delta t$, $p^{\Delta t}(t_k) = p_k^{\Delta t}$, and $u^{\Delta t}(t_k) = u_k^{\Delta t}$, we have $p_{k-1}^{\Delta t} = p^{\Delta t}(t_{k-1}) = p^{\Delta t}(t_k - \Delta t)$ and

$$\frac{p^{\Delta t}(t_k) - p^{\Delta t}(t_k - \Delta t)}{\Delta t} = f - u^{\Delta t}(t_k).$$

For $\Delta t \downarrow 0$, dropping subindex k we arrive at the differential equation

$$\frac{\partial p}{\partial t}(t) = f - u(t), \qquad p(t) = J'(u(t)), \tag{7}$$

with initial values given by $u(0) = p(0) = 0$. We assume $\int_\Omega f = 0$.

If the flow $u(t)$ according to (7) exists and is well behaved (which can be shown under reasonable assumptions on the functional J, in particular for total variation, cf. [2]), it is an inverse scale space method in the sense of [6]. This means that the flow starts at $u(0) = 0$ and incorporates finer and finer scales (with the concept of scale depending on the functional J) finally converging again to the image f as $t \to \infty$, i.e. $\lim_{t\to\infty} u(t) = f$. Through (7) the image $u(t)$ flows from the smoothest possible image ($u(0) = 0$) to the noisy image f. Our goal is to use the flow to denoise the image, and therefore we shall use a finite stopping time for the flow. As we shall see below, we can use a simple stopping criterion related to the fitting term $\|u(t) - f\|^2$ only.

3.1 Behaviour for Quadratic Regularization

We start by briefly reviewing the results obtained in [6] for (5). In this case we obtain from the variation of the functional J the equation $p = -\Delta u$ in Ω with boundary condition $\frac{\partial u}{\partial n} = 0$ on $\partial\Omega$. Given p, $\int_\Omega p = 0$, there exists a unique solution $u = -\Delta^{-1}p$ (since $\int_\Omega u = \int_\Omega f = 0$).

A simple manipulation (and the fact that $\frac{\partial f}{\partial t} = 0$) leads us to the equation $\frac{\partial}{\partial t}(u-f) = \Delta^{-1}(u-f) = -A(u-f)$, with the notation $A := -(\Delta)^{-1}$. Thus, the function $v = u - f$ satisfies a simple linear ordinary differential equation in the function space, whose solution is given by $u(t) - f = v(t) = e^{-tA}v(0) = -e^{-tA}f$. It is well-known that A is a positive definite operator and thus, $e^{-tA}f$ decays to zero. As a consequence, the difference $u(t) - f = -e^{-tA}f$ decays exponentially as $t \to \infty$.

3.2 General Convex Regularization

We consider the case of general convex functionals on the digital image space, i.e., $J : \mathbb{R}^N \to \mathbb{R}$. If J is continuously differentiable, we can compute the implicitly defined $u = u(p)$ as the solution of $J'(u(p)) = p$. Note that if J is smooth and strictly convex, the Jacobian of J' given by J'' is positive definite, and hence, the existence of a solution is guaranteed under a standard condition like $J(0) = 0$ by the inverse function theorem.

A possibility to invert the equation for u is the use of the the dual functional (cf. [4]), defined by $J^*(p) := \sup_u \{\langle u, p\rangle - J(u)\}$. Then one can easily show that $p = \partial_u J(u)$ is equivalent to $u = \partial_p J^*(p)$ and we obtain an explicit relation for $u(p)$ provided we can compute the dual functional J^*.

Under the above conditions, we can obtain some important estimates for the inverse scale space flow (7) associated to J. We start by computing the time-derivative of the fitting functional and the (partial) time derivative of u:

$$\frac{1}{2}\frac{d}{dt}\|u(t) - f\|_{L^2}^2 = \langle u(t) - f, \partial_t u(t)\rangle$$
$$\partial_t u(t) = \frac{d}{dt}(\partial_p J^*(p(t))) = H^*(p(t))\partial_t p(t) = -H^*(p(t))(u(t) - f),$$

where we used the notation $H^* = \partial_{pp}^2 J^*$ for the Hessian of the dual functional. If J^* is strictly convex, then there exists a constant $a > 0$ such that $H^*(q) \geq a$ for all $q \in \mathbb{R}$. Hence, combining the above estimates we deduce

$$\frac{1}{2}\frac{d}{dt}\|u - f\|_{L^2}^2 \leq -\langle u(t) - f, H^*(p(t))(u(t) - f)\rangle \leq -a\|u - f\|_{L^2}^2$$

and from a standard ordinary differential equation argument we deduce

$$\|u(t) - f\|_{L^2} \leq e^{-a(t-s)}\|u(s) - f\|_{L^2} \leq e^{-at}\|f\|_{L^2}$$

if $t > s$. Thus, as $t \to \infty$ we obtain convergence $u(t) \to f$ with exponential decay of the error in the L^2-norm.

Note that for the above L^2-estimates, we do not need severe assumptions on f, so that the estimate holds for a clean image as well as for a noisy version used in the algorithm. If we assume that f is a clean image and $J(f) < \infty$, then we can also obtain a decay estimate on the error in the Bregman distance via

$$\begin{aligned}\frac{d}{dt}D(f, u(t)) &= \frac{d}{dt}\big[J(f) - J(u(t)) - \langle f - u(t), p(t)\rangle\big] \\ &= -\langle f - u(t), \partial_t p(t)\rangle = -\|u(t) - f\|_{L^2}^2 \leq -\|f\|^2 e^{-2at}.\end{aligned}$$

In fact

$$D(f, u(t)) - D(f, u(s)) \leq \frac{1}{2a}\|f\|_{L^2}^2[e^{-2at} - e^{-2as}] < 0,$$
$$D(f, u(t)) \leq \frac{1}{2a}\|f\|_{L^2}^2[e^{-2at} - 1] + J(f).$$

All results so far give information about the convergence of u to the clean image f (with finite value $J(f)$) only. In a more practical situation, f is the noisy version of an image g to be restored, and we might even have $J(f) = \infty$, while $J(g) < \infty$. In this case we can derive a similar estimate as follows:

$$\begin{aligned}\frac{d}{dt}D(g, u) &= \langle -\partial_t p(t), g - u(t)\rangle = -\langle f - u(t), g - u(t)\rangle \\ &= -\|f - u(t)\|_{L^2}^2 - \langle f - u(t), g - f\rangle \leq -\frac{\|f - u(t)\|_{L^2}^2}{2} + \frac{\|f - g\|_{L^2}^2}{2}.\end{aligned}$$

The last term on the right-hand side is negative if $\|f-u(t)\|_{L^2} > \|f-g\|_{L^2}$. This means that $u(t)$ approaches any "noise free" image g in the sense of Bregman distance, as long as the residual (the L^2 difference between $u(t)$ and f) is larger than the difference between the noisy image f and g. The left-hand side, namely the residual $\|f-u(t)\|_{L^2}$ can be monitored during the iteration, it only involves the known noisy image f and the computed restoration $u(t)$. The right-hand side is not known for the "real" image g to be restored, since g itself is unknown. However, in typical imaging situations, an estimate for the noise variance is known, which yields a bound of the form $\|f-g\|_{L^2} \leq \sigma$. The above estimate guarantees that the distance $D(g,u)$ is decreasing at least as long as $\|f-u(t)\|_{L^2} > \sigma$, and one could terminate the inverse scale space flow for the minimal t_* such that $\|f-u(t_*)\|_{L^2} = \sigma$. This stopping criterion is well-known in the theory of iterative regularization of inverse problems as the so-called *discrepancy principle* (cf. [5, 10] for a detailed discussion).

4 Direct Solution for Regularized Total Variation in 1D

In the following we discuss the numerical solution of (7) in 1D. We recall here that $p(t) \in \partial J(u(t))$ and $u \in \partial J^*(p)$. For the (nondifferentiable) total variation functional we only have (multivalued) subgradients instead of derivatives and therefore we shall instead consider the regularized total variation $J(u) = \int \sqrt{|\nabla u|^2 + \epsilon^2}$, with

$$\partial J(u(t)) = -\mathrm{div}\left(\frac{\nabla u}{\sqrt{|\nabla u|^2+\epsilon^2}}\right) = p(t). \tag{8}$$

Note that since $\partial J(u+c) = \partial J(u)$, the solution of (8) is not unique if we take the standard assumption that u satisfies homogeneous Neumann boundary condition. In this case, the solvability condition is $\int p(x,t)\,dx = 0$ for all t and the conservation of mean value gives an additional property implying uniqueness, namely $\int u\ dx \equiv \int f\ dx = 0$.

For a fixed time t, we have to solve

$$-\left(\frac{u_x}{\sqrt{u_x^2+\epsilon^2}}\right)_x = p \ \text{ in } D=(a,b), \quad \int_a^b u\,dx = 0, \tag{9}$$

If we denote $q := \dfrac{u_x}{\sqrt{u_x^2+\epsilon^2}}$, then

$$q(x,t) = -\int_a^x p(s,t)\,ds = \int_x^b p(s,t)\,ds \tag{10}$$

and hence, $u_x = \epsilon \dfrac{q}{\sqrt{1-q^2}}$. Therefore,

$$u(x,t) = \epsilon \int_a^x \frac{q(y,t)}{\sqrt{1-q^2(y,t)}}\,dy + C \tag{11}$$

where C is a constant chosen to normalize $\int_a^b u(x)\,dx = 0$. We mention that the same formula for u can be obtained by duality arguments, since J^* can be explicitly calculated in spatial dimension one.

5 Relaxed Inverse Scale Space Flow

In order to implement the process in any dimension we resort to a new kind of approximation.

Consider the following coupled equations:

$$\begin{aligned} u_t &= -p(u) + \lambda(f + v - u), \\ v_t &= \alpha(f - u), \end{aligned} \tag{12}$$

where $u|_{t=0} = v|_{t=0} = 0$ and $\alpha > 0$ is a constant. These equations can be viewed as a time-continuous interpretation of the discrete iterated refinements procedure.

It is easy to see that the steady state of these equations ($u_t = 0, v_t = 0$) is: $u = f$, $v = \frac{p(f)}{\lambda}$. We still would like to show that for any $f \in BV$, the solutions converge to this steady state. We will do the analogue of this only for the simple linear case below. Our numerical evidence indicates that this is indeed true for the regularized TV flow in one and two dimensions.

By taking the time derivative of the first equation in (12) and substituting for v_t by using the second equation, we can view this process as a single, second order in time, evolution:

$$u_{tt} + (\lambda + p_u)u_t + \alpha\lambda u = \alpha\lambda f, \tag{13}$$

where $u|_{t=0} = 0, u_t|_{t=0} = \lambda f$. Here we assume $p(0) = 0$. We now analyze the linear case, where $-p(u) = \Delta u$. Rewriting the flow in the frequency domain ξ, by taking the Fourier transform, the characteristic equation is $r^2 + (\lambda + |\xi|^2)r + \alpha\lambda = 0$, with the solutions

$$r_\pm = \frac{-(\lambda + |\xi|^2) \pm \sqrt{(\lambda + |\xi|^2)^2 - 4\alpha\lambda}}{2}. \tag{14}$$

Using the Taylor approximation $\sqrt{1+x} \approx 1 + \frac{x}{2}$, $x \ll 1$, one can approximate (for frequencies for which $|\xi|^4 \gg \alpha\lambda$)

$$r_\pm \approx \frac{-(\lambda + |\xi|^2)(1 \pm (1 - \frac{2\alpha\lambda}{(\lambda+|\xi|^2)^2}))}{2}, \tag{15}$$

obtaining two roots with different characteristic behavior: $r_+ \approx -(\lambda + |\xi|^2)$, $r_- \approx \frac{-\alpha\lambda}{\lambda+|\xi|^2}$. The Fourier transform of the solution is

$$U(\xi) = (c_+ e^{r_+ t} + c_- e^{r_- t} + 1)F(\xi) \tag{16}$$

where $c_+ = \dfrac{\lambda + r_-}{r_+ - r_-}$, $c_- = \dfrac{\lambda + r_+}{r_- - r_+}$.

We observe that the first part, containing r_+, corresponds to a Gaussian convolution, which decays very quickly with time. The approximate second part, containing r_-, corresponds to the inverse scale-space solution (with some scaling in time) which we actually want to solve. Our numerical results indicate that this kind of behavior extends to the nonlinear process.

From (14) we see that for both parts to have decaying exponential solutions (real valued $r_\pm$) we should require $\alpha \leq \frac{\lambda}{4}$. In the numerical experiments below we set $\alpha = \frac{\lambda}{4}$.

The relaxed inverse scale space flow has about the same complexity as the standard gradient descent to steady state approach of ROF. Moreover, for the linear case, where $J(u) = \frac{1}{2}\int |\nabla u|^2$, as shown in [12], we obtain a step size estimation of $O(1)$ for the direct solution of inverse scale space flow.

6 Results

In this section we present some numerical examples. We solve the 1D problems by the direct approach discussed above and the relaxed inverse flow in order to test and compare their behavior. Motivated by the accordance of one-dimensional results, we only used the computationally cheaper relaxed flow for two-dimensional tests. In all experiments we use a uniform spatial grid of size $h = 1$, a standard assumption in imaging problems.

Example 1: We first consider a 1D denoising problem. Figure 1 shows the clean signal g, the noisy signal f, the noise n ($\sigma = \|n\|_{L^2} = 10 \approx 24\%\|g\|_2$) and the solutions recovered from ROF, direct inverse flow and relaxed inverse flow, respectively. The typical signal loss can be observed in the result of ROF, and as expected the loss is much smaller in the results of the inverse scale flows. This confirms the intuition that the inverse TV flows yields better restorations than the original ROF model.

The regularization parameter was chosen as $\epsilon = 1.5$ and for all three experiments, and all restored images u satisfy $\|f - u\|_{L^2} \approx \sigma$. The reason we use a relatively large ϵ here is that the direct solver method is sensitive to numerical error. Moreover, we used the parameter $\lambda = 0.02$ for ROF, time step $\Delta t = 5 \times 10^{-9}$ for the inverse TV flow, and $\lambda = 0.001$, $\Delta t = 0.1$ for the relaxed inverse TV flow. However, the difference of Δt and t in these two inverse flow experiments are only due to the different scaling, the number of time steps until the stopping time $t^* = \min\{t : \|f - u(t)\|_2 \leq \sigma\}$ was reached, is about the same.

Example 2: Now we turn to the denoising of 2D images. In this example we consider an image with different scales and shapes and corrupted by Gaussian noise, which is shown in Figure 2. $SNR(f) = 7.4, \sigma = \|f - g\|_2 = 40$. Figure 3 shows the results from ROF, iterated TV refinement (Bregman ROF), and relaxed inverse TV flow, row-by-row respectively. In each row we display the resulting restoration u and its corresponding residual $w = f - u$ and part of w. One observes that for the ROF model, visible signal is still contained in w (e.g. the small blocks and grids) and it's much smaller in the other two models. This

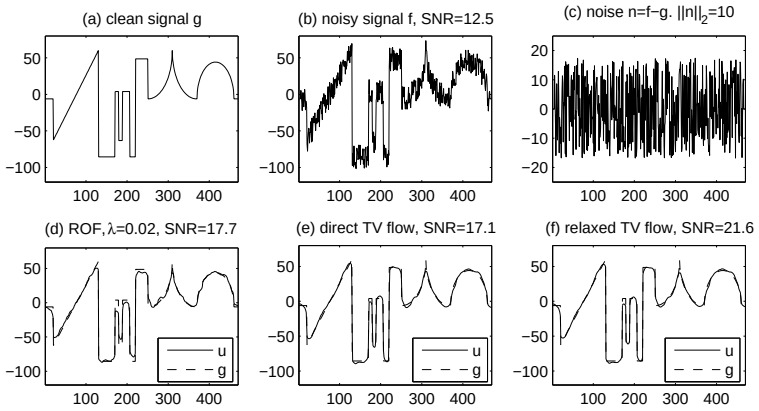

Fig. 1. 1D denoising. (a): clean g; (b): noisy f; (c): noise n; (d)-(f): u recovered from ROF, direct inverse flow and relaxed inverse flow.

is also quantified by the signal-to-noise ratios $SNR(u) = 9.9, 11.8, 12.5$ obtained for these three experiments respectively.

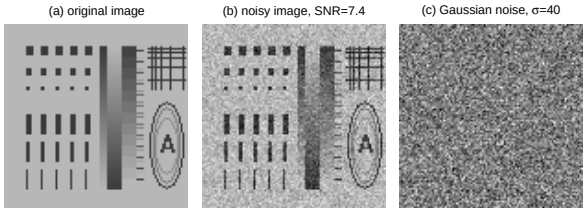

Fig. 2. 2D shape image. (a): original image; (b): noisy image; (c): Gaussian noise.

Example 3: In this final example we denoise a real satellite image with the same methods as used in example 2. Figure 4 shows the data and results. Here we have $SNR(f) = 6.3$ and $SNR(u) = 14.3, 15.1, 14.8$ for ROF, Bregman ROF and inverse TV flow, respectively. Again one can see some visible signal in w (such as the long antenna) for the ROF model, but less signal for the other two experiments.

Acknowledgement

The work of M.B. has been supported by the Austrian National Science Foundation FWF through Project SFB F 013/08. The work of S.O., J.X. and G.G. has been supported by the NIH through grant U54RR021813, and the NSF through

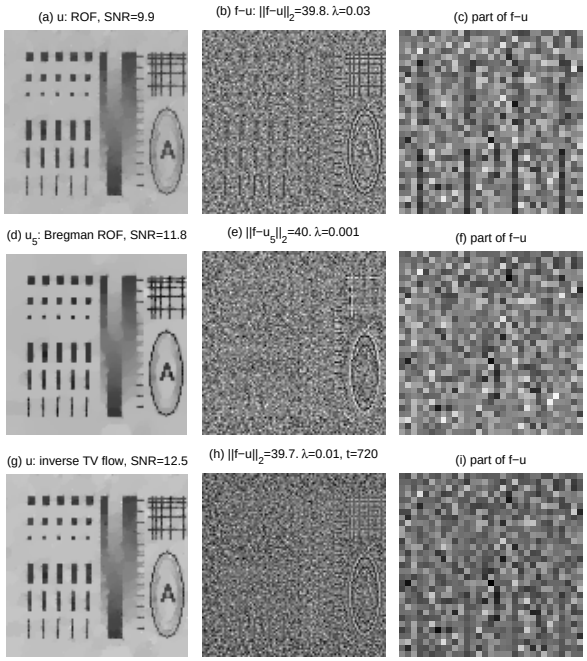

Fig. 3. 2D denoising on shape image. row-by-row: denoised u, residual $w = f - u$ and part of w from ROF, Bregman ROF and relaxed inverse TV flow.

grants DMS-0312222, ACI-0321917 and DMI-0327077. The authors would like to thank Mark Green and Barry Merriman (both UCLA) for useful and stimulating discussions.

References

1. L. M. Bregman. The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Comp. Math. and Math. Phys.*, 7:200–217, 1967.
2. M. Burger, D. Goldfarb, S. Osher, J. Xu, and W. Yin. Inverse total variation flow. in preparation.
3. G. Chen and M. Teboulle. Convergence analysis of a proximal-like minimization algorithm using bregman functions. *SIAM J. Optim.*, 3:538–543, 1993.
4. I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North-Holland Publishers, Amsterdam, 1976.
5. H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
6. C. Groetsch and O. Scherzer. Nonstationary iterated Tikhonov-Morozov method and third order differential equations for the evaluation of unbounded operators. *Math. Methods Appl. Sci.*, 23:1287–1300, 2000.

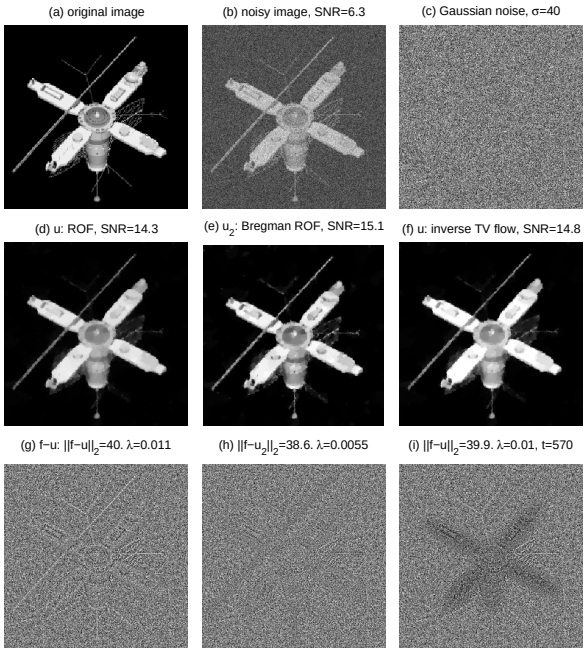

Fig. 4. 2D denoising on satellite image. first row: original, noisy image and noise; second row and third row: u and $w = f - u$ from ROF, Bregman ROF and relaxed inverse TV flow, column-by-column.

7. L. He, A. Marquina, and S. Osher. Blind deconvolution using TV regularization and Bregman iteration. *Int. J. of Imaging Systems and Technology*, 5:74–83, 2005.
8. S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation based image restoration. *Multiscale Model. and Simul.*, 4:460–489, 2005.
9. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. PAMI*, 12(7):629–639, 1990.
10. R. Plato. The discrepancy principle for iterative and parametric methods to solve linear ill-posed problems. *Numer. Math.*, 75:99–120, 1996.
11. L. I. Rudin, S. J. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992.
12. O. Scherzer and C. Groetsch. Inverse scale space theory for inverse problems. In M.Kerckhove, editor, *Scale-Space and Morphology in Computer Vision, Lecture Notes in Comput. Sci. 2106*, pages 317–325. Springer, New York, 2001.
13. O. Scherzer and J. Weickert. Relations between regularization and diffusion filtering. *J. Math. Imaging and Vision*, 12:43–63, 2000.
14. E. Tadmor, S. Nezzar, and L. Vese. A multiscale image representation using hierarchical (BV, L^2) decompositions. *Multiscale Model. Simul.*, 2:554–579, 2004.