UCLA
UCLA Previously Published Works

Title
Image Recovery via Nonlocal Operators

Permalink
https://escholarship.org/uc/item/9fk7q75p

Journal
Journal of Scientific Computing, 42(2)

ISSN
1573-7691

Authors
Lou, Yifei
Zhang, Xiaoqun
Osher, Stanley
et al.

Publication Date
2010-02-01

DOI
10.1007/s10915-009-9320-2

Peer reviewed

Image Recovery via Nonlocal Operators

**Yifei Lou · Xiaoqun Zhang · Stanley Osher ·
Andrea Bertozzi**

Abstract This paper considers two nonlocal regularizations for image recovery, which exploit the spatial interactions in images. We get superior results using preprocessed data as input for the weighted functionals. Applications discussed include image deconvolution and tomographic reconstruction. The numerical results show our method outperforms some previous ones.

Keywords Nonlocal methods · Inverse problem · Deconvolution · Tomography ·
Variational model

1 Introduction

We consider a general model:

$$f = \mathcal{K}u + n, \tag{1}$$

where f is the observed data, $\mathcal{K}$ is a continuous linear operator, u is the original image and n is white noise. Image recovery is an inverse problem, which refers to the estimation of u given the data f and continuous operator $\mathcal{K}$. For example, when the operator $\mathcal{K}$ is formulated as a convolution with a shift invariant kernel, this inverse problem becomes

Y. Lou (✉) · X. Zhang · S. Osher · A. Bertozzi
Mathematics Department, UCLA, Box 951555, Los Angeles, CA 90095-1555, USA
e-mail: yflou@math.ucla.edu

X. Zhang
e-mail: xqzhang@math.ucla.edu

S. Osher
e-mail: sjo@math.ucla.edu

A. Bertozzi
e-mail: bertozzi@math.ucla.edu

image deconvolution. Another example occurs in the context of tomographic reconstruction in which $\mathcal{K}$ represents an attenuated Radon transform and the data f are observed as photon counts [2]. These are the two examples considered here.

It is standard to approach the inverse problem by the method of regularization. For example, Tikhonov regularization [24] can be formulated as a one-step filter via Fourier transform, in the case of image deconvolution. However, this procedure tends to amplify the high frequencies where noise is conspicuous. Another popular regularization is total variation (TV). Due to its virtue of preserving edges, it is widely used in the many applications of image processing, such as blind deconvolution [7, 15, 20], inpainting [6] and super-resolution [21]. However, these TV-based methods do not preserve the fine structures, details and textures.

Buades *et al.* [3] introduce the nonlocal means filter to image denoising, which takes advantage of similar patches in the same image. The similarity is represented by a weight function, which induces a weighted averaging filter. The application of the nonlocal means filter to different image processing tasks is an active research area [4, 12, 22]. However, it remains challenging to extend previous nonlocal models to general inverse problems successfully. For example, the same authors use neighborhood filters to stabilize the inverse heat equation [4], following the direction in the PDE community to formulate Gaussian blurring as diffusion [16, 17, 19]. But this method does not work very well, since observed data and original images do not necessarily have same similarity distribution and structures. In this paper, we propose a key step for nonlocal methods to work effectively in the general case, which is to use a crude solution of the inverse problem to construct the weight. This solution can be obtained by any fast image restoration method, e.g. Tikhonov regularization [24] for image deconvolution and Filtered Back Projection (FBP) [2] for tomography. This step amounts to preprocessing the data to exploit the useful spatial information in image. It is more important for tomography, since the observed data and the image lie in different spaces. Along with the preprocessing the data to build the weight, we show that a nonlocal extension of total variation provides an effective mechanism for the inverse problem.

The paper is organized as follows: we review the related approaches in Sect. 2. A mathematic framework for nonlocal operators is studied in Sect. 3. In Sect. 4, we extend the TV-based model to a nonlocal setting with the applications to deconvolution and tomographic reconstruction. Section 5 contains numerical experiments.

2 Prior Work

2.1 Linear Models

First we review two common examples of linear regularizations to solve the image recovery problem. One is called Tikhonov regularization with identity [24], which is to solve the following minimization problem:

$$u = \arg\min \int (f - \mathcal{K}u)^2 + \mu \int u^2, \qquad (2)$$

where $\mu > 0$ is a regularizing parameter. The variational formulation gives us $\mathcal{K}^*(\mathcal{K}u - f) + \mu u = 0$, where $\mathcal{K}^*$ is the adjoint of $\mathcal{K}$. An alternative regularization [1] is to replace the L^2 norm in the last term in (2) with H^1 semi-norm. The second approach is called Wiener filter for image deconvolution:

$$u = \arg\min \int (f - \mathcal{K}u)^2 + \nu \int |\nabla u|^2. \qquad (3)$$

Similarly, the minimizer u is a solution of the Euler-Lagrange equation $\mathcal{K}^*(\mathcal{K}u - f) - \nu \Delta u = 0$, where Δu is the Laplacian of u.

These linear methods are very simple and fast to implement. For example, when $\mathcal{K}$ is a convolution operator, it only involves one step of calculation, *e.g.*, the fast Fourier transform. In general, the minimizer can be obtained by an efficient conjugate gradient algorithm. However, in order to compensate for the noise, a large regularizing parameter μ or ν is necessary, which tends to smear out the edges.

2.2 TV Regularization

Total variation [23] was proposed as regularizing functional for image restoration by Rudin, Osher and Fatemi, so-called ROF model,

$$u = \arg\min_u \int |\nabla u| + \frac{\lambda}{2} \int (f - \mathcal{K}u)^2, \tag{4}$$

where $\lambda > 0$ is Lagrange multiplier. This yields a nonlinear model. One way of solving this minimization problem is to evolve a PDE, whose steady state corresponds to a minima of the energy (4),

$$\begin{cases} u_t = \nabla \frac{\nabla u}{|\nabla u|} + \lambda \mathcal{K}^*(f - \mathcal{K}u) & \text{in } \Omega, \\ \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega. \end{cases} \tag{5}$$

Note that when $\mathcal{K}$ is the identity operator, we obtain the problem of image denoising. There are many fast methods to minimize (4) such as [5, 8, 13, 14].

2.3 Nonlocal Methods

The nonlocal methods in image processing are generalized from the Yaroslavsky filter [25] and patch-based methods, originally proposed for texture synthesis [10]. The idea is to restore an unknown pixel using other similar pixels. The resemblance is regarded in terms of a patch centered at each pixel, not just the intensity of the pixel itself. In order to denoise a pixel, it is better to average the nearby pixels with similar structures (patches). This idea is generalized to a famous neighborhood denoising filter, the Non Local Means (NLM) by Buades *et al.* in [3]:

$$NL_v(x) := \frac{1}{C(x)} \int_\Omega w_v(x, y)v(y)dy, \tag{6}$$

where v is the reference image, the weight function $w_v(x, y)$ and the normalizing factor $C(x)$ have the form:

$$w_v(x, y) = \exp\left(-\frac{(G_a * |v(x + \cdot) - v(y + \cdot)|^2)(0)}{h^2}\right), \tag{7}$$

$$C(x) = \int_\Omega w_v(x, y)dy, \tag{8}$$

where G_a is the Gaussian kernel with standard deviation a and h is a filtering parameter. In general h corresponds to the noise level; usually we set it to be the standard deviation of the noise. The weights are significant only if the window around y looks like the corresponding window around x. Thus the self-similarity is used to reduce the noise. The authors [3]

consider using the noisy data f as the reference image to construct the weight and by this weighed averaging, the structures, *e.g.* edges, are reinforced, while the noise gets cancelled out.

The application to a general inverse problem is not direct since observed data and original images do not necessarily have same similarity distribution and structures. In the case of deconvolution, based on the hypothesis that the deblurred image must maintain the same similarities as the blurry image, Buades *et al.* proposed a nonlocal deblurring model in [4]:

$$u = \arg\min_u \int (u - NL_f(u))^2 + \frac{\lambda}{2} \int (f - \mathcal{K}u)^2,$$

where NL_f is defined in (6) with the reference to be the blurry noisy data, *i.e.* $v = f$. This hypothesis has a strong limitation since it forces the deblurred image to maintain the same coherence as the given blurred one.

The nonlocal means can be regarded as one step of fixed point iteration to solve the optimality conditions of the general functional [18],

$$J(u) := \int_{\Omega \times \Omega} g\left(\frac{|u(x) - u(y)|^2}{h^2}\right) w(|x - y|)dxdy.$$

where g is a general functional. Inspired by graph theory [26], Gilboa and Osher [11] propose a particular functional g, which turns out to be a linear regularization for denoising and can be applied to supervised segmentation as well; they formalize a systematic and coherent framework [12] for nonlocal operators.

3 Nonlocal Framework

3.1 Nonlocal Operators

Here we review the work from [12]. Let $\Omega \subset \mathbf{R}^2$, $x, y \in \Omega$, $w(x, y)$ be a weight function. Furthermore $w(x, y)$ is assumed to be nonnegative and symmetric. We define

- nonlocal gradient $\nabla_w u : \Omega \to \Omega \times \Omega$

$$(\nabla_w u)(x, y) := (u(y) - u(x))\sqrt{w(x, y)}.$$

- nonlocal divergence $\operatorname{div}_w \vec{v} : \Omega \times \Omega \to \Omega$

$$(\operatorname{div}_w \vec{v})(x) := \int_\Omega (v(x, y) - v(y, x))\sqrt{w(x, y)}dy.$$

We design two types of regularization functionals based on the nonlocal operators.

$$J_{NL/H^1}(u) = \frac{1}{4} \int |\nabla_w u|^2 = \iint_{\Omega \times \Omega} (u(x) - u(y))^2 w(x, y)dxdy, \tag{9}$$

$$J_{NL/TV}(u) = \int |\nabla_w u| = \int_\Omega \sqrt{\int_\Omega (u(x) - u(y))^2 w(x, y)dy}dx. \tag{10}$$

Note that the functional in (9) is analogous to the standard H^1 semi-norm, so it is denoted as NL/H^1; similarly the one in (10) is denoted as NL/TV.

We calculate the Euler-Lagrange of the functionals above, thus obtaining

$$L_{NL/H^1} u = - \int_\Omega (u(y) - u(x)) w(x, y) dy, \tag{11}$$

$$L_{NL/TV} u = - \int_\Omega (u(y) - u(x)) w(x, y) \left[\frac{1}{|\nabla_w u(x)|} + \frac{1}{|\nabla_w u(y)|} \right] dy. \tag{12}$$

3.2 Computing the Weights

In this paper we use the weight function as (7), which is the same formula in [3] and [4], except that we use a preprocessed image as the reference to construct the weight. Another difference is that we do not require normalization of the weight since we formulate a minimization problem. In practice, in order to improve computational time and storage efficiency, we only compute the weights in a semi-local searching window for each pixel. The details of the preprocessing are given in the next section, since it varies from different applications.

4 Two Applications

4.1 Image Deconvolution

We use the nonlocal regularization functionals in (9) and (10) to perform deconvolution, by incorporating the convolution into the data fidelity. In this case, the general model $\mathcal{K}u$ amounts to the convolution between the image u and a circular shift-invariant blur kernel g. Furthermore, periodic boundary conditions for the convolution are assumed, thus the fast Fourier transform can be applied to realize the convolution.

We define the total energy as

$$E(u) = J(u) + \frac{\lambda}{2} \int (g * u - f)^2, \tag{13}$$

We use gradient descent to update the solution by the Euler-Lagrange of (13)

$$u_t = -Lu + \lambda \tilde{g} * (f - g * u), \tag{14}$$

where the operator L is the corresponding gradient flow with respect to the functional J, that is either (11) or (12), and $\tilde{g}$ is the adjoint of g.

A better way to construct the weight function (7) is to use a preprocessed image instead of the blurry and noisy data f. Ideally the preprocessed image should be sharper than the blurry image and easy to compute. For example, we may use the linear model mentioned in Sect. 2.1. Although these methods suffer from amplifying the noise frequencies, the nonlocal framework can automatically handle its side-effects. Since the noise statistics of the preprocessed image is changed, we should alter the parameter h in the weight function to compensate for the noise accordingly. We choose Tikhonov regularization to be the preprocessor considering that it produces sharper results than Wiener filter, which is to minimize the H^1 semi-norm.

The minimizer of the linear model (2) can be obtained in the frequency domain (ξ, η):

$$\hat{u}(\xi, \eta) = \frac{\hat{g}^*(\xi, \eta) \hat{f}(\xi, \eta)}{|\hat{g}|^2(\xi, \eta) + \mu}, \tag{15}$$

where $\wedge$ denotes the Fourier transform and $*$ is the conjugate operator.

In the case where we know some statistics of additive noise, *e.g.* the standard deviation of noise, a proper parameter μ is chosen such that the restored image u in (15) satisfies

$$\iint (f - g * u)^2 dxdy = \iint n^2 dxdy = \sigma^2.$$

By Parseval's theorem, this is equivalent to finding μ satisfying

$$\sigma^2 = \iint (f - g * u)^2 dxdy = \iint |\hat{f} - \hat{g}\hat{u}|^2 d\xi d\eta = \iint \frac{\mu \hat{f}(\xi, \eta)}{|\hat{g}|^2 + \mu} d\xi d\eta.$$

Notice that the right-hand side is monotonically increasing function in μ, hence there exist a unique solution μ, which can be determined via bisection.

Since the above procedure describes a linear operator applied to f, we can denote Inv_g. It follows from the image degradation model (1) that

$$\tilde{f} = \mathrm{Inv}_g \circ f = \mathrm{Inv}_g \circ (g * u + n).$$

Note that after preprocessing the data, the noise has L^2 norm

$$\|\mathrm{Inv}_g \circ n\|_{L^2} \leqslant \|\mathrm{Inv}_g\|_{L^2} \|n\|_{L^2} = \sigma \|\mathrm{Inv}_g\|_{L^2}.$$

Hence a good choice is $h = \sigma \|\mathrm{Inv}_g\|_{L^2}$.

4.2 Tomographic Reconstruction

We provide another example of the nonlocal functional model in the context of Computerized Tomography (CT). In a simplified parallel tomographic problem, an observed body slice is modelled as a two-dimensional distribution $(x, y) \mapsto u(x, y)$ of the X-ray attenuation constant, and a line integral called projection $p(r, \theta)$ represents the total attenuation of a beam of X-rays parameterized by (r, θ), where $r \in \mathbb{R}$ is the signed perpendicular distance from the line to the origin and $\theta \in [0, \pi)$ is the angle between the perpendicular vector and the x-axis. As shown in Fig. 1, a projection obtained by illuminating the object along the line is given by

$$Ru : p(r, \theta) = \iint_{-\infty}^{\infty} u(x, y)\delta(x \cos \theta + y \sin \theta - r)dxdy,$$

where δ denotes the Dirac delta-distribution. The linear operator $R : u(x, y) \mapsto p(r, \theta)$ is called Radon transform. The tomographic reconstruction problem is to estimate the distribution $u(x, y)$ from a finite number of measured line integrals $p(r, \theta)$. The standard reconstruction algorithm in clinical applications is the Filtered Back Projection, which is a direct discretization of an analytical formula for the inverse of the Radon transform. As the name indicates, this method can be decomposed into two steps: (1) the one-dimensional projection data along each orientation θ is applied via a *ramp filter*; (2) each pixel of the image u is obtained by summing on those filtered projections passing through this pixel. In fact, the last step is actually an adjoint transform of the Radon transform [2], which is called *back projection*. In the presence of noise, this problem becomes difficult since the inverse of the Radon transform is unbounded. Traditionally, a low-pass filter has to be applied to

Fig. 1 Radon transform in $\mathbb{R}^2$
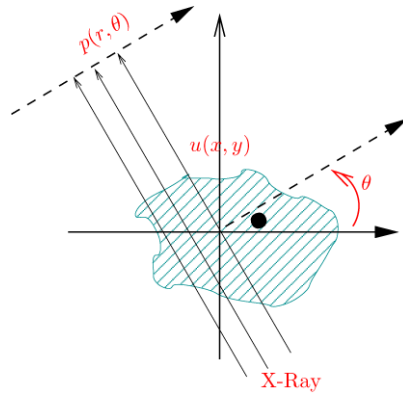
compensate the noise amplification by ramp filter. However this step smooths edges, where important structures are located.

Instead we formulate this problem as a variational problem

$$\min J(u) + \frac{\lambda}{2} \|Ru - p\|_2^2,$$

where R denotes a discrete Radon transform (often a sparse matrix) and p represents the measured projection data, and the functional J can be chosen from (9) and (10). The solution is found iteratively using a gradient algorithm.

The weight function $w(x, y)$ is needed to measure the similarity of image features between pixels x and y. Since the observed projection data are in the Radon domain, which do not directly reflect the similarity information of the image, we need a more accurate image to compute the weight. We consider the FBP image as a crude and fast solution to build the weight. Moreover, we choose the filter parameter h to be the estimated noise variance in the filtered back projection image. Here we use a wavelet based noise estimation model introduced by Donoho *et al.* [9] in 1994. More precisely, we first apply a wavelet transform on the filtered back projection image, then the noise variance of this image is estimated by using the medial absolute deviation of the wavelet coefficients at its finest scale, *i.e.*:

$$\hat{\sigma}^2 = \frac{\text{median}(|y_i|)}{0.6745},$$

where y_i are wavelet coefficients in the finest subband of the filtered back projection image. Note that an alternative choice is to compute the norm of an discrete inverse of the Radon transform.

4.3 Algorithm

Algorithm 1 summarizes the steps of the proposed methods for two applications: image deconvolution and tomographic reconstruction. Note that it can be generalized to other applications.

5 Experiments

In this section we present some numerical results using the nonlocal functionals: NL/H^1 and NL/TV. We compare these to traditional methods, such as Tikhonov regularization [24],

---

Algorithm 1 Image Recovery via Nonlocal Operators

Input: Data f and the degradation model $\mathcal{K}$.

Preprocessing: Compute u^0 using Tikhonov regularization for image deconvolution or filtered back projection for tomographic reconstruction.

Weight computation: For each pixel x, the weight $w(x, y)$ is computed for every y belonging to a searching window of size 21×21 centered at x, according to (7) for $v = u^0$. The choice of h was discussed in the previous section.

Minimization

$$E(u) = J(u) + \frac{\lambda}{2} \int (f - \mathcal{K}u)^2$$

Choose a fixed λ, and set $e = 1e - 5$, $dt = 0.1$, $k = 0$ and *maxIter* $= 500$. Use the preprocessed data u^0 as the initial guess for the iteration.

while $\|u^{k+1} - u^k\| > e$ and $k < maxIter$ **do**

   Gradient descent

$$u^{k+1} = u^k - dt \left(Lu^k + \lambda \mathcal{K}^*(\mathcal{K}u^k - f) \right)$$

   where L is either (11) or (12).

   $k = k + 1$.

end while

Output: final solution u^k.

Comment: We run the minimization for a wide range of λ and choose the best result.

Table 1 The statistics of blur and noise we add to the images

Image	Blur kernel	Gaussian noise
Shape	Gaussian with $\sigma_b = 2$	$\sigma_n = 10$
Barbara	Gaussian with $\sigma_b = 1$	$\sigma_n = 5$
Cameraman	9×9 box average	$\sigma_n = 3$

ROF [23] and the NL-means deblurring model [4] by Buades *et al*. As for the nonlocal weight, we compute it from either the noisy blurry image or the preprocessed image.

We use signal-to-noise ratio (SNR) as a means of judging performance

$$SNR(u, f) = 20 \log_{10} \left\{ \frac{\|f - \bar{f}\|_{L^2}}{\|f - u\|_{L^2}} \right\},$$

where u is the original image, f is the recovery and $\bar{f}$ is its mean value.

5.1 Image Deconvolution

We test all the methods on three images: a synthetic image, Barbara and Cameraman with various kinds of blur and noise as listed in Table 1. The synthetic one is referred to as Shape since it contains geometric features. The Cameraman image is high contrast and has many edges, while the image of Barbara contains more textures.

We use Tikhonov regularization to preprocess the data with the method parameter μ determined by the standard deviation of the noise. Then we use this preprocessed data to compute the weight function with filter parameter h chosen to compensate the amplified

Table 2 SNR for different methods

Image	Shape	Barbara	Cameraman
Noisy input	11.6213	10.0363	8.4600
Tikhonov [1]	13.1970	10.5680	11.2940
ROF [23]	14.2192	11.4257	12.3448
NLM deblur [4]	16.3289	12.0903	12.4269
NL/H^1	16.8084	11.7952	12.4599
NL/TV	17.5251	12.0066	12.7495
Tikhonov + [4]	17.5269	**12.7369**	13.5099
Tikhonov + NL/H^1	19.6982	12.1182	13.5185
Tikhonov + NL/TV	**20.9401**	12.4616	**13.6194**

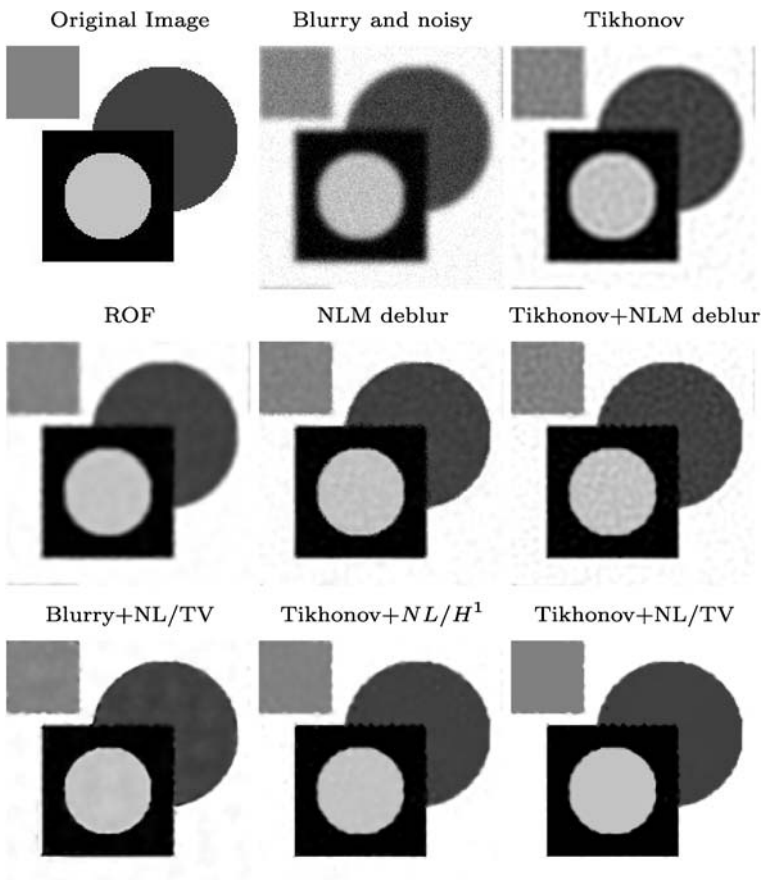**Fig. 2** A 150×150 image with Gaussian blur $\sigma_b = 2$ and Gaussian noise $\sigma_n = 10$

noise. Please refer to Sect. 4.1 for details. Once we have the weight, we solve the minimizer of (13) with either NL/H^1 or NL/TV as regularization via gradient descent. For each
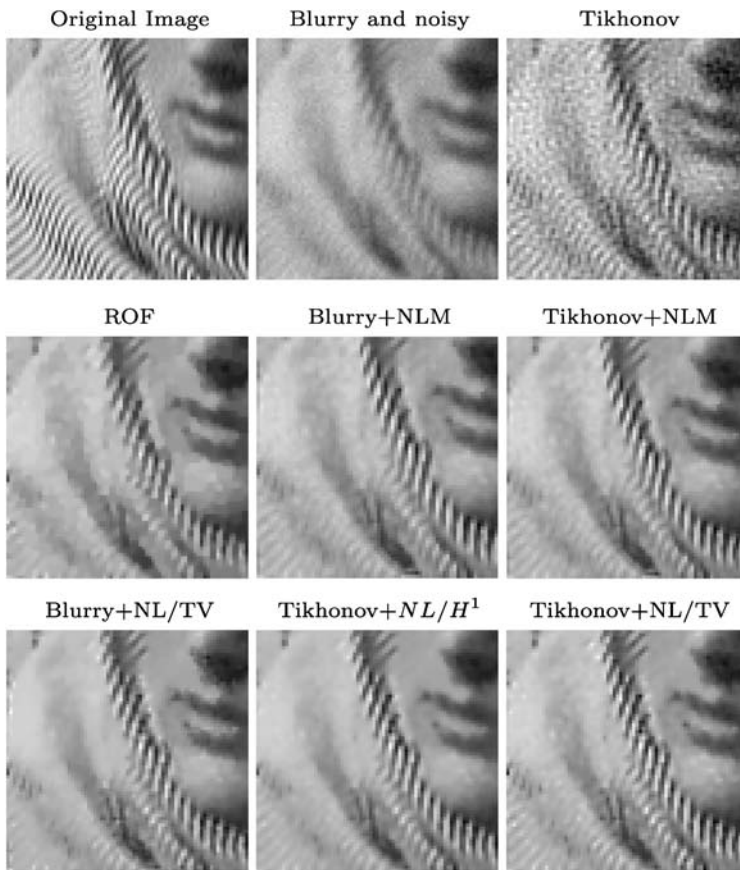
Fig. 3 A 200×200 image with Gaussian blur $\sigma_b = 1$ and Gaussian noise $\sigma_n = 5$ cropped to 75×75 pixels

method, we present our result with an optimal parameter λ chosen from a series of values with wide range.

Regardless of which way we compute the nonlocal weights, visually NL/H^1 returns a smoother image than NL/TV does. We list the SNR of all the methods in Table 2. For Figs. 2 and 4, the best reconstruction is the one that combines a preprocessor by Tikhonov filter with NL/TV regularization functional both in terms of SNR and direct inspection as well. Although Tikhonov + NLM deblur achieves the highest SNR in Barbara example, Tikhonov + NL/TV restores most textures especially in the bottom left corner as zoomed in Fig. 3.

For computational time, it takes about 2.9 seconds for a dual core desktop with 2.99 GHz processor and 1.99 GB memory to construct the weight function of a 256×256 image in MATLAB. Once we have constructed the weight, the iteration of NL/H^1 and NL/TV is comparable to ROF in speed (NL/H^1 is more efficient than NL/TV). The computation speed depends on the number of iterations. In general, it takes around 200 seconds for 500 steps iteration. For some examples, the energy converges very quickly. Also, code could be optimized.
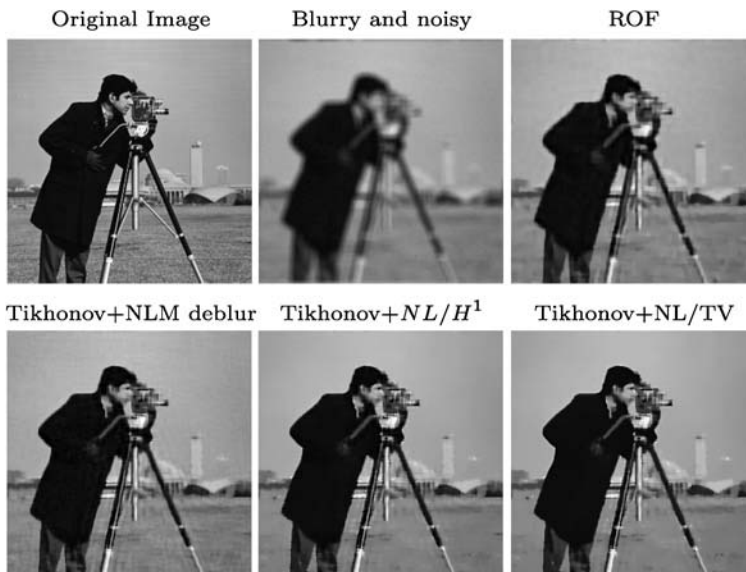
Fig. 4 A 256×256 image with box average kernel 9×9 and Gaussian noise $\sigma_n = 3$

5.2 Tomographic Reconstruction

We have tried the algorithms on a synthetic Shepp-Logan head phantom and a real brain image, for which projection data have been computed using a Radon transform matrix.

The size of each image is 128×128. The number of projection angles is chosen to be the size of the image and there are $\sqrt{2} * 128 \approx 185$ X-ray parallel beams for each angle in order to cover the whole image domain. Therefore the sizes of projection data for both images are 185×128. Then we add white noise of $\sigma_n = 100$ to the projection data of both images. Since the projection data lies in a different space than the image does, [4] can not be directly applied in this case. Therefore, the weight function for the nonlocal methods is estimated from the image reconstructed by FBP, which is also used as the initial guess for all the variational models. As we discussed above, we compute the weight function $w(x, y)$ from an initial filtered back projection image and estimate noise level to determine the value of parameter h in the weight.

Figures 5 and 6 show the simulation results in the presence of Gaussian white noise in the projection of a phantom and a brain image. The efficiency of FBP is very poor. *TV*-regularization performs well but the details of the images are slightly blurred. Similar to deconvolution, NL/H^1 gives a visually smoother image than NL/TV does, but NL/TV archives a slightly better SNR. Furthermore, NL/H^1 gives smoother results but details are well-preserved. For both images, nonlocal functionals outperform the traditional methods.

6 Conclusion

Nonlocal functionals have recently been introduced for the case of image denoising. We have extended their utilities to more general inverse problems. We discuss two applications of the general model: image deconvolution and tomographic reconstruction. We found nonlocal
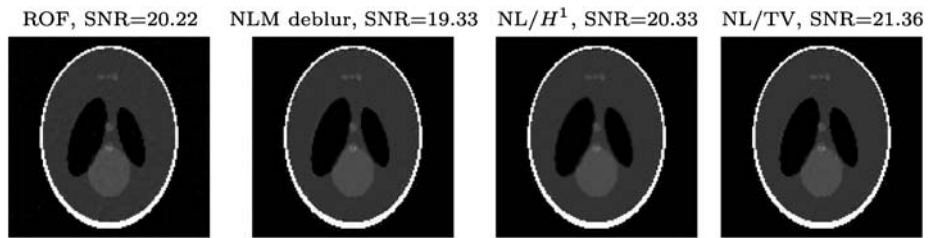
ROF, SNR=20.22   NLM deblur, SNR=19.33   NL/H^1, SNR=20.33   NL/TV, SNR=21.36

Fig. 5 Results of reconstruction from noisy projection data with SNR = 28.1db

Original Image   FBP, SNR=9.68   ROF, SNR=14.77

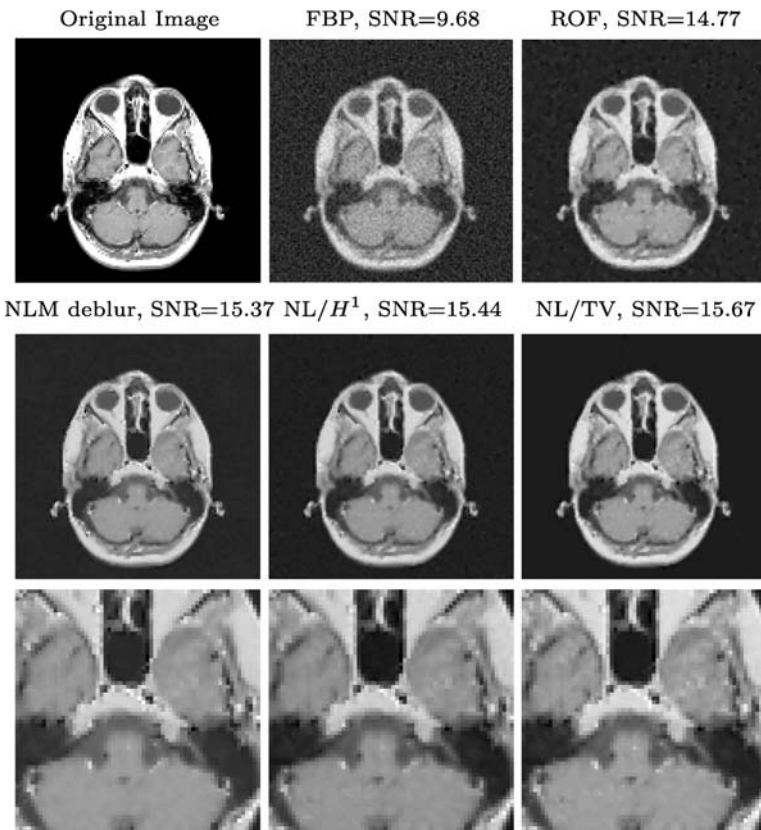NLM deblur, SNR=15.37 NL/H^1, SNR=15.44   NL/TV, SNR=15.67

Fig. 6 Results of reconstruction from noisy projection data with SNR = 26.04db. On the *last row* is the enlarged central part of each reconstruction

functionals to give superior results, provided that the data is preprocessed using older simple techniques to construct the nonlocal weight functional.

References

1. Andrews, H.C., Hunt, B.R.: Digital Image Restoration. Prentice-Hall, Englewood Cliffs (1977)
2. Avinash, C., Malcolm, S.: Principles of Computerized Tomographic Imaging. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2001)
3. Buades, A., Coll, B., Morel, J.-M.: A review of image denoising algorithms, with a new one. Multiscale Model. Simul. **4**(2), 490–530 (2005)
4. Buades, A., Coll, B., Morel, J.M.: Image enhancement by non-local reverse heat equation. Technical Report 22, CMLA (2006)
5. Chambolle, A.: An algorithm for total variation minimization and applications. J. Math. Imaging Vis. **20**, 89–97 (2004)
6. Chan, T.F., Shen, J.: Mathematical models for local non-texture inpainting. J. Appl. Math. **62**(3), 1019–1043 (2002)
7. Chan, T.F., Wong, C.K.: Total variation blind deconvolution. IEEE Trans. Image Process. **7**(3), 370–375 (1998)
8. Darbon, J., Osher, S.: Fast discrete optimization for sparse approximations and deconvolution. Technical report, UCLA (2007, to appear)
9. Donoho, D., Johnstone, J.: Ideal spatial adaption via wavelet shrinkage. Biometrika **81**(3), 425–455 (1994)
10. Efros, A.A., Leung, T.K.: Texture synthesis by non-parameteric sampling. In: ICCV, vol. 2, pp. 1033–1038 (1999)
11. Gilboa, G., Osher, S.: Nonlocal linear image regularization and supervised segmentation. Multiscale Model. Simul. **6**(2), 595–630 (2007)
12. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. Multiscale Model. Simul. **7**(3), 1005–1028 (2008)
13. Goldfarb, D., Yin, W.: Second-order cone programming methods for total variation based image restoration. SIAM J. Sci. Comput. **27**(2), 622–645 (2004)
14. Goldstein, T., Osher, S.: The split Bregman method for L1-regularized problems. SIAM J. Imaging Sci. 2(2), 323–343
15. He, L., Marquina, A., Osher, S.: Blind deconvolution using TV regularization and Bregman iteration. Int. J. Imaging Syst. Technol. **15**(1), 74–83 (2005)
16. Hummel, R.A., Kimia, B., Zucker, S.W.: Gaussian blur and the heat equation: forward and inverse solutions. In: IEEE Computer Vision and Pattern Recognition, vol. 2, pp. 416–421 (1983)
17. Hummel, R.A., Kimia, B., Zucker, S.W.: Deblurring Gaussian blur. In: Computer Vision, Graphics, and Image Processing, vol. 38, pp. 66–80 (1987)
18. Kindermann, S., Osher, S., Jones, P.W.: Deblurring and denoising of images by nonlocal functionals. Multiscale Model. Simul. **4**(4), 1091–1115 (2005)
19. Kramer, H.P., Bruckner, J.B.: Iterations of a non-linear transformation for enhancement of digital images. In: Pattern Recognition, vol. 7, pp. 53–58 (1975)
20. Marquina, A.: Nonlinear inverse scale space methods for total variation blind deconvolution. SIAM J. Imaging Sci. **2**(1), 64–83 (2009)
21. Marquina, A., Osher, S.: Image super-resolution by TV-regularization and Bregman iteration. J. Sci. Comput. **37**(3), 367–382 (2008)
22. Peyré, G., Bougleux, S., Cohen, L.: Non-local regularization of inverse problems. In: ECCV '08: European Conference on Computer Vision, pp. 57–68. Springer, Berlin (2008)
23. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)
24. Tikhonov, A., Arsenin, V.: Solution of Ill-Posed Problems. Wiley, New York (1977)
25. Yaroslavsky, L.P.: Digital Picture Processing, an Introduction. Springer, Berlin (1985)
26. Zhou, D., Scholkopf, B.: A regularization framework for learning from graph data. In: ICML Workshop on Stat. Relational Learning and Its Connections to Other Fields (2004)