Compressed sensing by inverse scale space and curvelet thresholding

Jianwei Ma

School of Aerospace, Tsinghua University, Beijing 100084, China

Abstract

Compressed sensing provides a new sampling theory for data acquisition, which says that compressible signals can be exactly reconstructed from highly incomplete sets of linear measurements. It is significant to many applications, e.g., medical imaging and remote sensing, especially for measurements limited by physical and physiological constraints, or extremely expensive. In this paper we proposed a recovery algorithm from a view of reaction-diffusion equations, by applying curvelet thresholding in inverse scale space flows. Numerical experiments in medical CT and aerospace remote sensing show its good performances for recovery of detailed features from incomplete and inaccurate measurements, in comparison with some existing methods.

Index Terms

Inverse problem, compressed sensing, image recovery, incomplete measurements, geometric wavelets, aerospace remote sensing.

I. Introduction

Compressed sensing (CS) or compressive sampling [7–10, 18] is a new direction in the fields of signal processing. It says that we can recover an compressible signal u from a very small set of incomplete measurements f. The CS needs far fewer measurements than conventional measurements that are limited by Shannon/Nyquist sampling theorem: the

sampling rate must be at least twice the maximum frequency of signal. If we let $\Phi \in \mathcal{C}^{K,N}$, $K << N$ be a CS matrix or an imaging lens in optical architecture, the encoding can be described as

$$f = \Phi u + \epsilon. \tag{1}$$

Here ϵ denotes possible measurement noises. It seems hopeless solving the ill-posed underdetermined equations since the rows are much fewer than the columns in Φ. Most time, the x is compressible by a transform Ψ, thus we have $f = \Phi\Psi\vartheta + \epsilon$. The CS told us that we can accurately recover the sparse coefficients ϑ thus u by solving a constrained optimal problem.

$$\min \|\vartheta\|_{l_1}, \;\; subject\, to \; \|\Phi\Psi\vartheta - f\|_{l_2} \leq \epsilon. \tag{2}$$

Promising performances of CS theorems have been demonstrated in a few potential applications including compressive imaging, wireless sensing, analog-to-information conversion, biosensing for DNA microarrays, optical architecture, and surface metrology (see e.g., [1, 3, 11, 27, 28, 31, 32] and http://www.dsp.ece.rice.edu/cs/). To make the CS work well, one must handle successfully two goals: constructing of CS matrices and designing of nonlinear recover algorithms. To construct the CS matrices, Candès et al. [7, 10] proposed a sufficient condition, named restricted isometry property (RIP). Roughly speaking, the CS matrix should be noise-like incoherent in the sparse transform domain. The greater incoherence, the smaller number of measurements needed. Frequent-used measurements are random matrices which are incoherent to most transforms. So far, a few special CS measurement matrices have been presented, including sparse 0/1 random matrices [2], Toeplitz block matrices [40], structurally random matrices and scrambled block Hadamard ensemble [23], etc. How to build determined measurement matrices that satisfy a modified RIP was considered by DeVore [17]. Very recently, Romberg [38] presented a new framework named random convolution measurement, i.e., convolution with a random pulse waveform followed by random

time domain subsampling. Elad [20] proposed to design optimal adjustable measurement matrices by using an average measure of the so-called mutual-coherence of the effective dictionary. On the other hand, a few recovery algorithms have been proposed in the last couple of years. Typically, e.g., linear programming [7], gradient projection sparse reconstruction [22], orthogonal matching pursuit (OMP) [43], stagewise OMP [19], and iterative thresholding (IT) [4, 15, 16, 31, 32, 36]. Here we concern ourselves to the nonlinear recovery algorithm from a new way: partial differential equation (PDE) method.

In fact, the PDE method is one of classical filtering methods in signal processing. It has been successfully explored for denoising and enhancement of signals. Unlike tools from computational harmonic analysis (CHA) (e.g., wavelets) that suffer from pseudo-Gibbs artifacts and shift/rotation variance, applications of PDE (e.g., anisotropic nonlinear diffusion) are almost free from the lacks of CHA. Recent researches are increasingly focusing on the relations and combinations of both, see e.g., [29,30,41] among lots of literature. The essential idea of PDE is scale space theory. For a nonlinear diffusion

$$\frac{\partial u(x,t)}{\partial t} = \nabla \cdot (g(|\nabla u(x,t)|)\, \nabla u(x,t)) \tag{3}$$

with the given noisy image as initial condition $u_0 = u(x,0) = f(x)$ and periodic boundary conditions. Here the time t acts as a scale parameter for filtering, which is known to lead to a stronger smoothing of u with increasing t. A various t produces scale space. The diffusivity g controls the smoothing process by admitting strong diffusion if the gradient ∇u is small (possibly caused by small-scale feature) and slowing down the smoothing for large gradients.

Recently, an inverse scale space (ISS) was introduced by Scherzer and Groetsch [39] for ill-posed inverse problems. Unlike traditional forward diffusions that start with the observed noisy image and gradually smooth it, the ISS methods start with an arbitrary image and approach the observation image f as time increases. It generates a data representation in

inverse scale space from a smooth u_0 to detailed f at a continuum of scales t. If we stop at a suitable time, detailed features are still missing while large scale features have been reconstructed.

So far, only little literature on ISS has been published. Burger et al. [5] provided a so-called relaxed inverse flow for numerical computation of the ISS method. Lie and Nordbotten [26] further proved that the relaxed inverse flow is convergent for convex regularization functionals. Xu and Osher [45] applied the nonlinear inverse scale space method to wavelet domain denoising, in which the authors replaced the gradient ∇u by wavelet coefficients thus the original PDE is uncoupled. When we finished the draft, we make aware of references [24,46] that respectively involve Bregman iteration and ISS method for CS. As declared in [24] and [5], the ISS method is more efficient than the Bregman iteration in terms of computation cost and performance. In the following section 2 we will show the close relation between the Bregman iteration and ISS. However, both methods in [24,46] did not consider the constraint of sparse transform, and also can not recover well the detailed textural components. In this paper we apply a curvelet-combined ISS method for recovery of CS from incomplete and inaccurate measurements. In numerical experiments, we show the superior performance of our method to reconstruct the detailed features, in comparison to ISS method, wavelet-TV reconstruction, and iterative curvelet thresholding without the combination of ISS.

II. Curvelet-combined inverse scale space for CS

A. Curvelet transform

Curvelet transform [12,13] is a new geometric multiscale transform. As wavelet transform owns good performance to represent isotropic point singularities, curvelet transform allows an optimal sparse representation of objects with C^2 singularities. The needle-shape elements of this transform own very high directional sensitivity and anisotropy. For a smooth object

f with discontinuities along smooth curves, the best m-term approximation $\tilde{f}_m$ by curvelet thresholding obeys $\|f - \tilde{f}_m\|_2^2 \leq Cm^{-2} (\log m)^3$, while for wavelets the decay rate is only m^{-1}.

Unlike wavelets, indexed by two parameters, a system of curvelets is indexed by three parameters. Let μ be the collection of triple index (j, l, k), where j, l and $k = (k_1, k_2)$ are respectively scale, orientation and translation parameters. The curvelets are defined as functions of $x \in \mathbb{R}^2$ by

$$\psi_\mu (x) = \psi_j (R_{\theta_J} (x - k_\delta)) . \tag{4}$$

In the above, ψ is a waveform oscillatory in the horizontal direction and bell-shaped (nonoscillatory) along the vertical direction. R_{θ_J} is a rotation matrix of angle $\theta_J = 2\pi \cdot 2^{-\lfloor j/2 \rfloor} \cdot l$, $J = (j, l)$ indexing the scale/angle, with $\lfloor \cdot \rfloor$ denoting the integer part, while the translation parameter is given by $k_\delta = R_{\theta_J}^{-1} \left(k_1 \cdot 2^{-j}, k_2 \cdot 2^{-j/2} \right)$. The curvelet elements are obtained by anisotropic dilations, rotations and translations of a collection of unit scale oscillatory blobs. For a function f, the curvelet coefficients are given by

$$\vartheta_\mu = \langle f, \psi_\mu \rangle = \int_{\mathbb{R}^2} f (x) \, \bar{\psi}_\mu (x) \, dx, \tag{5}$$

and it can be directly evaluated by frequency-window partitions in the second-generation version of discrete curvelet transform [13]. The forward and inverse curvelet transform have the computational cost of $\mathcal{O}(N^2 \log N)$ for an $(N \times N)$ image. We refer readers to [12, 13, 30, 33, 42] for details of curvelet transform and its applications.

The curvelet thresholding function can be defined,

$$S_\sigma(f) = \sum_\mu \chi(\vartheta_\mu(f)) \, \psi_\mu, \tag{6}$$

where χ can be taken as a soft thresholding function defined for a fixed threshold $\sigma > 0$,

$$\chi_s(x) = \begin{cases} x - \sigma, & x \geq \sigma, \\ 0, & |x| < \sigma, \\ x + \sigma & x \leq -\sigma, \end{cases}$$

or a hard thresholding function

$$\chi_h(x) = \begin{cases} x, & |x| \geq \sigma, \\ 0, & |x| < \sigma, \end{cases}$$

B. Curvelet-based inverse scale space

In this section, we first briefly review the theorem of ISS and a relaxed ISS scheme for practical computation, then we apply the ISS scheme combining with iterative curvelet thresholding for CS.

Considering additive fidelity term, we can rewrite a general diffusion form of Eq. (3) as

$$\partial_t u = -\partial_u J(u) + \lambda \partial_u H(f, u), \tag{7}$$

or in regularization form

$$u = \min_u \{J(u) + \lambda H(f, u)\}. \tag{8}$$

Here $J(u)$ is a convex regularization functional, e.g., $J(u) = |u|_{BV} = \int_\Omega |\nabla u|$ for total variation regularization/diffusion. $H(f, u)$ is a convex fidelity functional, e.g., $\frac{1}{2}\|f - u\|_{L^2}$.

We further deduce a so-called iterated refinement method (IRM) or Bregman iteration [34] as

$$u_k = \min_u \{D(u, u_{k-1}) + \lambda H(f, u)\}, \tag{9}$$

where D is a Bregman distance function defined by

$$D(u, v) = J(u) - J(v) - \langle u - v, p \rangle \quad p \in \partial_u J(u). \tag{10}$$

Omitting constant parts that are not relevant to the minimization from Eq. (10), we obtained

$$u_k = \min_u \{J(u) - \langle u, p(u_{k-1})\rangle + \lambda H(f,u)\}. \tag{11}$$

By deducing the Euler-Lagrange equation of (11) and assuming $\lambda \to 0$, we obtain a constrained PDE for inverse scale space

$$\partial_t p = -\partial_u H(f,u), \;\; p \in \partial_u J(u). \tag{12}$$

with initial conditions $u|_{t=0} = 0, p|_{t=0} = 0$.

The ISS method can provide more accurate results than the IRM because we can compute the stopping time more accurately in the ISS due to the continuous evolution. The IRM can be interpreted as an implicit Euler discretization of the ISS. But in comparison with IRM, we have a much faster approximate algorithm, so-called relaxed ISS flow for the computation of ISS [5, 6]:

$$\begin{aligned} \partial_t u &= -p + \lambda(-\partial_u H(f,u) + v), \\ \partial_t v &= -\alpha \partial_u H(f,u), \end{aligned} \tag{13}$$

with initial conditions $u|_{t=0} = 0, v|_{t=0} = 0$.

In particular, taking e.g., H=$\frac{1}{2}\|f - \Phi u\|_{L^2} = \frac{1}{2}\|f - \Phi\Psi\vartheta\|_{L^2}$, we have the following ISS scheme for CS.

$$\begin{aligned} \partial_t u &= \nabla \cdot (g(|\nabla(S * u)|)\, \nabla u) + \lambda(\Phi^T(f - \Phi u) + v), \\ \partial_t v &= \alpha \Phi^T (f - \Phi u). \end{aligned} \tag{14}$$

The first term on the right side of $\partial_t u$ is a diffusion with a certain regularization S to smooth u, while the second term is a fidelity to sharp u and make it approach the observations f. Here the S is an optional smoothing operator or pre-filtering regularization,

which acts as a preprocessing to reduce the influence of noise during the diffusion process. The classical S is taken as Gaussian-filtering regularization [14]. Other possibilities, e.g., time-delay regularization, wavelet and curvelet regularization can be also considered [30,37]. The frequently applied diffusivity is total variation (TV) diffusivity $g(x) := 1/x$, and Perona-Malik (PM) diffusivity [35] $g(x) := 1/(1 + x^2/\gamma^2)$ with a suitable chosen contrast parameter γ. This model is essentially a reaction-diffusion equation.

In the CS theorem a sparse transform Ψ can be seen as a prior knowledge. For instance, we can take Ψ as curvelet transform if the measurement objects consist of curve singularities. There are at least two ways to take advantage of the prior knowledge in our model. The first way is that one computes the equations in coefficient ϑ domain by taking advantage of the sparsity of the instant field u in the transform domain. But, following this way, one has to learn knowledge on multiscale sparse computing of PDEs, and has to deal with the boundary condition in curvelet domain, which is a challenge problem in curvelet-based solving of PDEs. The second way is that one applies thresholding $S_\sigma(x)$ for the fidelity term, e.g., $S_\sigma(\Phi^T(f - \Phi u))$ in Eq. (14). In order to make it work robustly, in practice, we can directly apply the thresholding as $S_\sigma(u + \Phi^T(f - \Phi u))$ instead of the residual parts. Using the forward difference with step τ for discretization of time variable, we propose a ISS-IT scheme

$$u^{k+1} = S_\sigma(u^k + \tau\lambda\Phi^T(f - \Phi u^k)) + \tau\nabla \cdot (g(|\nabla(S * u^k)|)\nabla u^k) + \tau\lambda v^k, \tag{15}$$

$$v^{k+1} = \alpha\Phi^T(f - \Phi u^k). \tag{16}$$

The first term of the right side in Eq. (15) is a thresholding term, which is related to weighted IT methods [4, 16, 31, 36]. Readers can check the difference easily between these IT methods and our method. Our model combines the diffusion to suppress the sharpened artifacts resulted from the fidelity. The model is also motivated by our recent reaction-

diffusion equation [37] and iterative curvelet thresholding method [31]. We refer readers to [16, 37] for analysis of related mathematical properties.

Now we give a remark for the regularization operator S in the above inverse diffusion equations. In spite of having many desirable properties, the classic PM model is a notoriously ill-posed problem [14,21,25]. Another drawback of the PM diffusion is its sensitivity to noise. Noise often introduces very large oscillations of the gradient ∇u, therefore the gradient-based model possibly misjudges true edges and heavy noise leading to undesirable diffusion in regions where there is no true edge. In [14], a classical Gaussian regularization is proposed by using Gaussian pre-filtering. However, Gaussian filtering blurs edges and finer textures. This behavior seems somewhat against the purpose of the PM equation (i.e. sharpen the edges). In [37], we suggested to take the regularization operator as the curvelet-shrinkage. In this paper, we apply the second-generation discrete curvelet transform [12,13] for the digital filtering and thresholding in the ISS-IT scheme. The curvelet pre-processing can effectively remove the noise while preserving the edges well, leading to better discontinuity-preserving diffusion.

C. Numerical procedure

We apply a four-pixel scheme [44] for spatial discretization of the inverse diffusion equation (15). As recommend by one of the reviewers, we provide the numerical procedure of computing the divergence term $\nabla \cdot (g(|\nabla(S * u)|)\nabla u)$. For simplify, we use $\tilde{u}$ to denote the $S * u$ below.

The divergence expression can be decomposed by means of two orthonormal basis vectors x_1, x_2,

$$\nabla \cdot \left(g(|\nabla \tilde{u}|)\, \nabla u\right) = \partial_{x_1}(g(|\nabla \tilde{u}|)\, \partial_{x_1} u) + \partial_{x_2}(g(|\nabla \tilde{u}|)\, \partial_{x_2} u).$$

Choosing $x_1 = (1, 0)$, $x_2 = (0, 1)$, and replacing the derivatives by finite differences, we

obtain the discrete scheme

$$\begin{aligned}\nabla\cdot\left(g(|\nabla\tilde{u}|)\,\nabla u\right) = {} & g(|\tilde{u}^k_{i+1,j}-\tilde{u}^k_{i,j}|)\,(u^k_{i+1,j}-u^k_{i,j})\\ & -g(|\tilde{u}^k_{i,j}-\tilde{u}^k_{i-1,j}|)\,(u^k_{i,j}-u^k_{i-1,j})\\ & +g(|\tilde{u}^k_{i,j+1}-\tilde{u}^k_{i,j}|)\,(u^k_{i,j+1}-u^k_{i,j})\\ & -g(|\tilde{u}^k_{i,j}-\tilde{u}^k_{i,j-1}|)\,(u^k_{i,j}-u^k_{i,j-1}).\end{aligned}$$

Here, $u^k_{i,j}$ denotes the sampled values of u^k, i.e., $u^k_{i,j}=u^k(i,j)$ for the suitable scaled image. Analogously, choosing the diagonal directions $x_1=\frac{1}{\sqrt{2}}(1,1)$, $x_2=\frac{1}{\sqrt{2}}(1,-1)$ we find

$$\begin{aligned}\nabla\cdot\left(g(|\nabla\tilde{u}|)\,\nabla u\right) = {} & \frac{1}{2}(g(\frac{|\tilde{u}^k_{i+1,j+1}-\tilde{u}^k_{i,j}|}{\sqrt{2}})(u^k_{i+1,j+1}-u^k_{i,j})\\ & -g(\frac{|\tilde{u}^k_{i,j}-\tilde{u}^k_{i-1,j-1}|}{\sqrt{2}})(u^k_{i,j}-u^k_{i-1,j-1})\\ & +g(\frac{|\tilde{u}^k_{i+1,j-1}-\tilde{u}^k_{i,j}|}{\sqrt{2}})(u^k_{i+1,j-1}-u^k_{i,j})\\ & -g(\frac{|\tilde{u}^k_{i-1,j+1}-\tilde{u}^k_{i,j}|}{\sqrt{2}})(u^k_{i-1,j+1}-u^k_{i,j})).\end{aligned}$$

Averaging the two equations leads to a four-direction finite differences for the divergence term,

$$\sum\nolimits^{1}_{\substack{r,s=-1\\(r,s)\neq(0,0)}}\frac{g(\sqrt{2}^{1-|r|-|s|}(|\tilde{u}_{i+r,j+s}-\tilde{u}_{i,j}|))(u_{i+r,j+s}-u_{i,j})}{r^2+s^2}.$$

It should be stressed that the computing is not applied in curvelet-coefficient space, but by a spacial discretization. This means that each time this term is computed, the curvelet thresholding is used for u before the diffusion operator is applied. This does not seems to be very efficient in terms of computational cost. But in practices, one only needs to use the curvelet regularization S in the starting several iterations to get a satisfying result. As mentioned in our previous work [30, 37], there is a way to compute the diffusion operator in curvelet space by replacing gradient term by curvelet coefficients, i.e., use $g(|\vartheta_\mu|)$ instead of

$g(|\nabla S u|)$. This strategy can reduce the computational cost, but creates artifacts to some extent. There is big room for future work.

III. Experiments

Figure 1 shows the good performance of our proposed method in comparison with other methods. Fig. 1 (a) is an original Shepp-Logan phantom in CT imaging. We assume that the image is unknown and any pixel information can not be used in our recovery methods. We want to recover the unknown image from a few random measurements. The only prior knowledge that we can use is its sparsity in curvelet domain. The measurement matrix Φ that we consider in this paper is pseudo-random undersampling measurements in Fourier frequency domain. Fig. 1 (b) shows an example of Φ, i.e., a various-density point mask with 25% undersampling factor [27, 28]. Fig. 1 (c) is the reconstruction by zero-filling with density compensation (zf-w/dc) suggested in [27]. This result is used as an initial value in the following iterative methods. Fig. 1 (d) is recovered by the $wavelet-TV$ method proposed by Lustig et al. [27], i.e., solving a minimization of $\|\vartheta\|_{l_1} + \lambda_w TV(u)$ subject to $\|\Phi\Psi\vartheta - f\| \leq \epsilon$. Here the ϑ denotes wavelet coefficients. This method is implemented based on a nonlinear projection gradient technique instead of iterative thresholding. The Daubechies DB4 wavelet transform with three-level decomposition and the weight $\lambda_w = 0.01$ are used in this case. Fig. 1 (e) is recovered by iterative curvelet thresholding proposed by the author in [31], with decreasing hard threshold value $\sigma = \sigma_0(1 - k/k_{number})$, where initial threshold value is $\sigma_0 = 0.04$, k is the index of iterations, and $k_{number} = 120$ is the number of total iterations. Fig. 1 (f) is recovered by ISS with TV diffusivity, in which the detailed features are lost. Fig. 1 (g) is the result recovered by our ISS-curvelet method with TV diffusivity. Here we use parameters $\tau = 0.002, \alpha = 100, \lambda = 500$, iteration $k_{number} = 120$ times, and hard thresholding with decreasing threshold value $\sigma = \sigma_0(1 - k/k_{number})$, $\sigma_0 = 0.04$. Fig. 1 (h) is

recovered by the ISS-curvelet method with PM diffusivity, using parameters $\tau = 0.02, \alpha = 10, \lambda = 50, \gamma = 0.002$, the same iterations and thresholding. For contrast-enhancement cases, the PM diffusion often displays better performances. It should be noted that the best experiential parameter setting has been taken for all methods.

Figure 2 shows the removed components by the above methods. (a)-(d) are the components removed by $wavelet - TV$, iterative curvelet thresholding, ISS, and our ISS-curvelet with TV diffusivity, respectively. It can be seen that the wavelet and TV methods kick out the edge features too much while the proposed method recovers the detail feature very well, and also achieves a higher SNR (signal-to-noise ratio) in comparison with other methods. Furthermore, Figure 3 shows the various of SNR and $l_1 - norm$ recovery errors as the iteration increases, in term of the proposed method. Study of the rate of change will be helpful for adaptive choose of parameters in future researches.

We also tested that the proposed method is robust for measurement noises and various numbers of measurements. More iterations are needed for bigger noises and fewer measurements. Quantitative speaking, using the above ISS-curvelet with the same parameters for noise case (ϵ is produced by 0.01 times random numbers), we got similar clear image but with SNR=40.88 dB. And using 150 iterations and keeping other parameters unchanged, we have SNR=44.82 dB.

Compressed sensing based methods, e.g. single-pixel camera, have been applied successfully on satellite remote sensing in [32]. It is very significant in this field to reduce the number of sensors, imaging time, storage space and power consumption of on-board cameras in satellites, and reduce the data to be transmitted back to earth [32]. In figure 4, we apply the proposed ISS-IT method for recovery of compressed remote sensing. Since we consider the random Fourier measurement in these experiments, in fast, this is a multi-pixel but

one-times (MPOT) imaging camera, as described in [32]. Fig. 4 (a) is an original unknown cloud image with 512×512 size, considering measurement noise (0.005 times random numbers) and the same 25% sampling strategy as above. Fig. 4 (b) and (c) are respectively the zero-filling reconstruction and $wavelet - TV$ reconstruction. Fig. 4 (d) is the proposed ISS-IT reconstruction with 50 iterations. In this case, the elapsed computation time of the $wavelet - TV$ and ISS-IT is 2382.13 seconds and 2710.37 seconds using a laptop with Intel Pentium processor 1.86 GHz and 512 MB memory, respectively.

Our method achieves higher visual quality and SNR, because it takes advantage of the merits of iterative curvelet thresholding and inverse anisotropic diffusion. Of course, the combining algorithm needs to pay more computational costs than simple thresholding. However, in the CS-based remote sensing, this computational costs can be done in a rather cheap way by digital computers on ground, instead of on-board chips of camera in satellites or probes. In practice, there are two way to improve the computational speed: 1) The number of iterations can be reduced than respective iterative methods, due to the curvelet thresholding and anisotropic can benefit from each other in the ISS-IT framework, especially for textural images with noises. 2) The regularization S can be applied only for the first few iterations.

IV. Conclusion

This paper is concerned with a computational method for compressed sensing based on inverse scale space and curvelet thresholding. A combination of several existing methodologies (inverse scale space flow, nonlinear diffusion, curvelet thresholding) is used to build a new iterative algorithm to reconstruct sparse representations from fewer measurements. The motivation of this method is that we apply a diffusion to suppress the sharpened artifacts resulted from the fidelity in a reaction-diffusion system. In comparison with some existing

methods, the proposed method can recover the detailed features much better. Applications on medical CI imaging and satellite remote sensing show great potential to reduce the data acquisition.

The current version can be improved at least from two aspects in future work. 1) The number of parameters in this method: τ, λ, σ, and k_{number} probably can be decreased by finding out a good combination. 2) To improve the iterative computational expenses, and do more detailed comparisons to the alternatives in terms of computational speed.

Acknowledgments: The author would like to thank financial support from NSFC Grant No. 40704019, TBRF (JC2007030), and PetroChina Innovation Fund (060511-1-1).

References

[1] R. Baraniuk, A lecture on compressive sensing, *IEEE Signal Processing Magazine* **24** (4), 118-121 (2007).

[2] R. Berinde, P. Indyk, Sparse recovery using sparse random matrices, Tech. Report of MIT, (2008).

[3] J. Bobin, J. Starck, R. Ottensamer, Compressed sensing in astronomy, *IEEE J. Selected Topics in Signal Process.*, submitted, (2008).

[4] T. Blumensath, M. Davies, Iterative thresholding for sparse approximations, *J. Fourier Anal. Appl.*, to appear, (2008).

[5] M. Burger, G. Gilboa, S. Osher, J. Xu, Nonlinear inverse scale space methods, *Comm. Math. Sci.* **4** (1), 179-212 (2006).

[6] M. Burger, K. Frick, S. Osher, O. Scherzer, Inverse total variation flow, *Multiscale Model. Simul.* **6**, 366-386 (2007).

[7] E. Candès, T. Tao, Decoding by linear programming, *IEEE Trans. Infor. Theory* **51**, 4203-4215 (2005).

[8] E. Candès, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate information, *Commun. Pure Appl. Math.* **59**, 1207-1233 (2005).

[9] E. Candès, T. Tao, Near optimal signal recovery from random projections: Universal encoding straegies, *IEEE Trans. Inform. Theory* **52** (12), 5406-5425 (2006).

[10] E. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory* **52** (2), 489-590 (2006).

[11] E. Candès, Compressive sampling, *Int. Congress of Mathematicians*, vol. 3, pp. 1433-1452, Madrid, Spain, 2006.

[12] E. Candès, D. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise singularities, *Comm. Pure Appl. Math.* **57**, 219-266 (2004).

[13] E. Candès, L. Demanet, D. Donoho, L. Ying, Fast discrete curvelet transforms, *Multiscale Model. Simul.* **5** (3), 8610-889 (2006).

[14] F. Catté, P. Lions, J. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM J. Numer. Anal.* **32**, 1895-1909 (1992).

[15] P. Combettes, V. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.* **4**, 1168-1200 (2005).

[16] I. Daubechies, M. De Friese, C. De Mol, An iterative thresholding algorithm for linear inverse problems wiht a sparsity constriant, *Commun. Pure Appl. Math.* **57**, 1413-1457 (2004).

[17] R. A. DeVore, Deterministic constructions of compressed sensing matrices, *J. of Complexity* **23**, 918-925 (2007).

[18] D. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory* **52** (4), 1289-1306 (2006).

[19] D. Donoho, Y. Tsaig, I. Drori, J. Starck, Sparse solution of underdetermined linear

equatins by stagewise Orthogonal Matching Pursuit, *IEEE Trans. Inform. Theory*, to appear, (2008).

[20] M. Elad, Optimized projections for compressed sensing, *IEEE Trans. Signal Process.* **55** (12), 5695-5702 (2007).

[21] S. Esedoglu, An analysis of the Perona-Malik scheme, *Comm. Pure Appl. Math.* **54**, 1442-1487 (2001).

[22] M. Figueiredo, R. Nowak, S. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, *IEEE J. Select Topic in Signal Process.* **1** (4), 586-597 (2007).

[23] L. Gan, T. Do, T. Tran, Fast compressive imaging using scrambled block Hadamard ensemble, preprint, (2008).

[24] L. He, T. Chang, S. Osher, T. Fang, P. Speier, MR image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods, UCLA CAM Report 06-35, (2006).

[25] S. Kichenassamy, The perona-malik paradox, *SIAM J. Appl. Math.* **57** (5), 1328-1342 (1997).

[26] J. Lie, J. Nordbotten, inverse scale spaces for nonlinear filtering, *J. Math. Imaging Vis.* **27** (1),41-50 (2007).

[27] M. Lustig, D. Donoho, J. Pauly, Sparse MRI: the application of compressed sensing for rapid MR imaging, *Magnetic Resonance in Medicine* **58** (6), 1182-1195 (2007).

[28] M. Lustig, D. Donoho, J. Santos, J. Pauly, Compressed sensing MRI, *IEEE Signal Process. Magazine* **25** (2), 72-82 (2008).

[29] J. Ma, M. Fenn, Combined complex ridgelet shrinkage and total variation minimization, *SIAM J. Sci. Comput.* **28** (3), 984-1000 (2006).

[30] J. Ma, G. Plonka, Combined curvelet shrinage and nonlinear anisotropic diffusion, *IEEE Trans. Image Process.* **16** (9), 2198-2206 (2007).

[31] J. Ma, Compressed sensing for surface metrology, *IEEE Trans. Insturm. Measure.*, revision submitted, (2008).

[32] J. Ma, F.-X. Le Dimet, Deblurring from highly incomplete measurements for remote sensing, *IEEE Trans. Geosci. Remote Sensing*, to appear, (2008).

[33] J. Ma, A. Antoniadis, F.-X. Le Dimet, Curvelet-based snake for multiscale detection and tracking of geophysical fluids, *IEEE Trans. Geosci. Remote Sensing* **44** (12), 3626-3638 (2006).

[34] S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Ying, An iterative regularization method for total variation baased image restoration, *Multiscale Model. Simul.* **4**, 460-489 (2005).

[35] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.* **12**, 629-639 (1990).

[36] G. Peyré, Best basis compressed sensing, *Proc. of SSVM07*, pp.80-91, June 2007.

[37] G. Plonka, J. Ma, Nonlinear regularized reaction-diffusion filters for denoising of images with textures, *IEEE Trans. Image Process.* **17** (8), 1283-1294 (2008).

[38] J. Romberg, Compressive sensing by random convolution, *SIAM J. Image Science*, submitted, 2008.

[39] O. Scherzer, C. Groetsch, Inverse scale space theory for inverse problems, In M. Kerckhove (Ed.), *Scale-space 2001*, LNSC 2106, pp. 317-325 Springer, New York, 2001.

[40] F. Sebert, Y. Zou, L. Ying, Toeplitz block matrices in compressed sensing and their applications in imaging, *Int. Conf. Tech. Appl. in Biomedicine*, 30-31 May 2008, pp. 47-50.

[41] G. Steidl, J. Weickert, T. Brox, P. Mrázek, M. Welk, On the equivalence of soft wavelet

shrinkage, total variation diffusion, total variation regularization, and sides, *SIAM J. Numer. Anal.* **42** (2), 686-713 (2004).

[42] J. Starck, E. Candès, D. Donoho, The curvelet transform for image denoising, *IEEE Trans. Image Process.* **11**, 670-684 (2002).

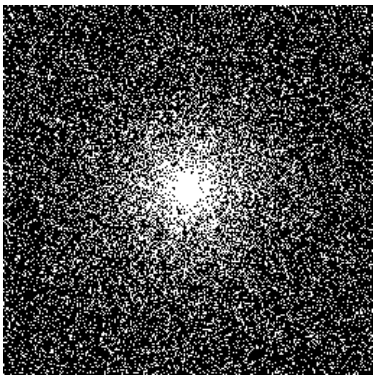
[43] J. Tropp, A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. Inform. Theory* **53** (12), 4655-4666 (2008).

[44] M. Welk, G. Steidl, J. Weickert: A four–pixel scheme for singular differential equations. In R. Kimmel et al. (Eds.), *Scale-space and PDE methods in computer vision.* LNCS, Springer, Berlin, pp. 610–621, 2005.

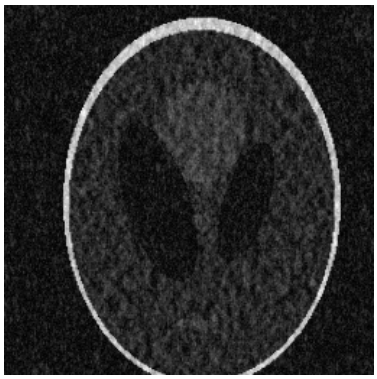
[45] J. Xu, S. Osher, Iterative regularization and nonlinear inverse scale space applied to wavelet based denoising, *IEEE Trans. Image Process.* **16** (2), 534-544 (2007).

[46] W. Yin, S. Osher, D. Goldfarb, J. Darbon, Bregman iterative algorithms for l_1-minimization with applications to compressed sensing, *SIAM J. Imaging Sciences* **1** (1), 143-168 (2008).

(a)



(b)

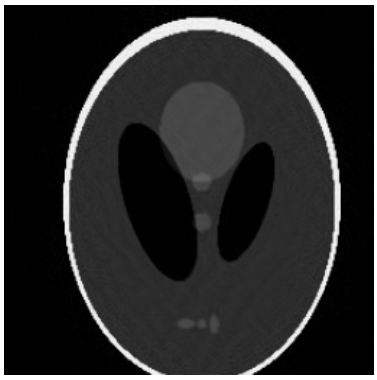


(c)

SNR = 35.28 dB

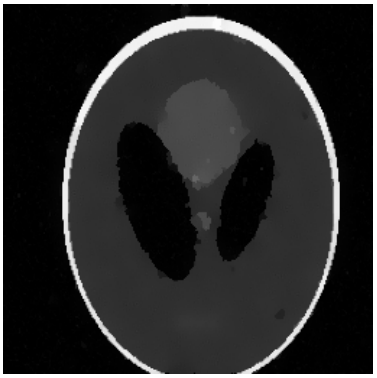
(d)

SNR = 38.11 dB



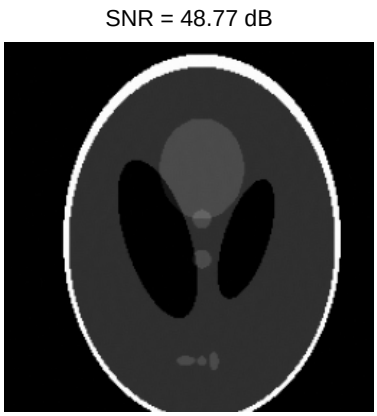
(e)

SNR = 27.81 dB

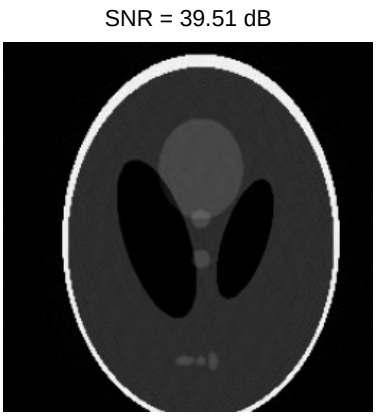


(f)

Fig. 1. Compressed sensing for a Shepp-Logan phantom using different methods. (a) original unknown image. (b) 25% pseudo-random undersampling mask in frequency domain. (c) recovery by zero-filling reconstruction. (d) by wavelet-TV reconstruction. (e) by iterative curvelet thresholding. (f) by ISS method with TV diffusivity. (g) by our ISS-curvelet method with TV diffusivity. (h) by our ISS-curvelet method with PM diffusivity.



(g)



(h)

Figure 1: continue.

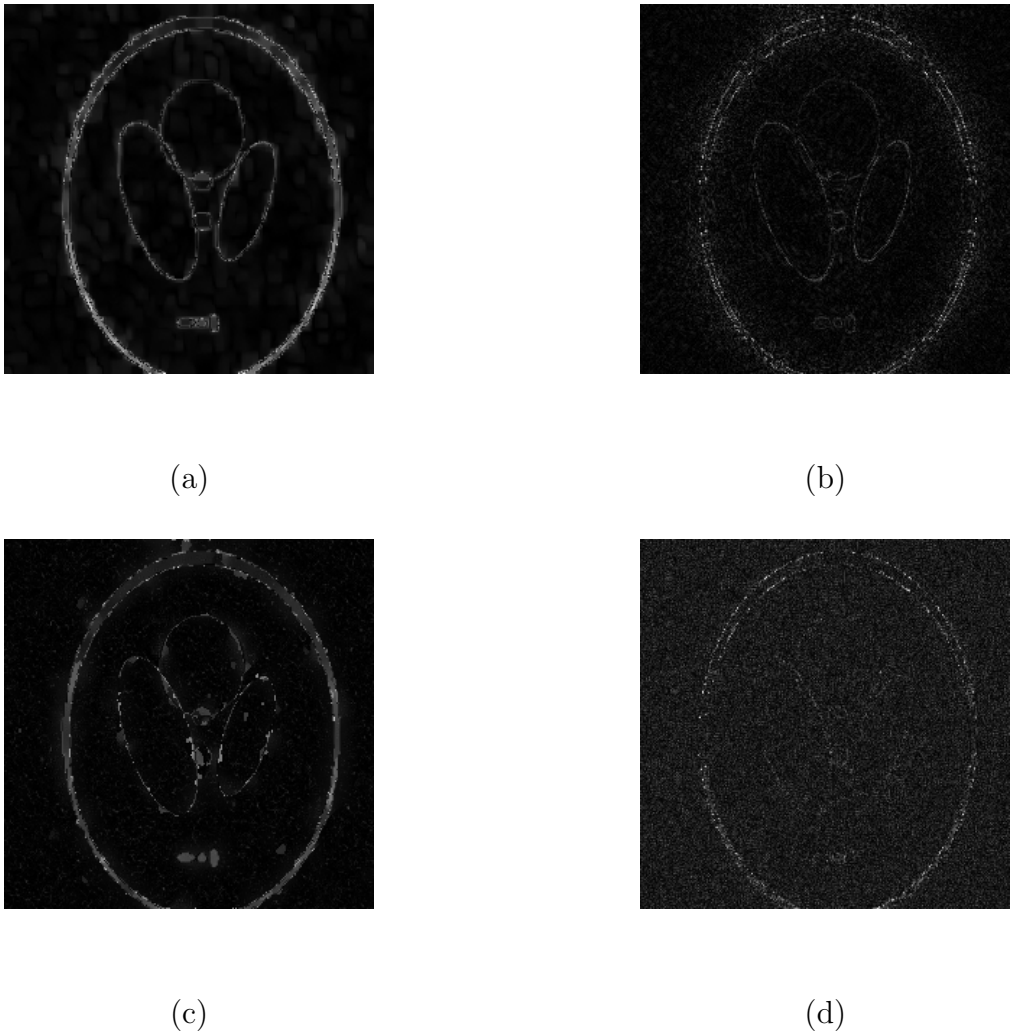
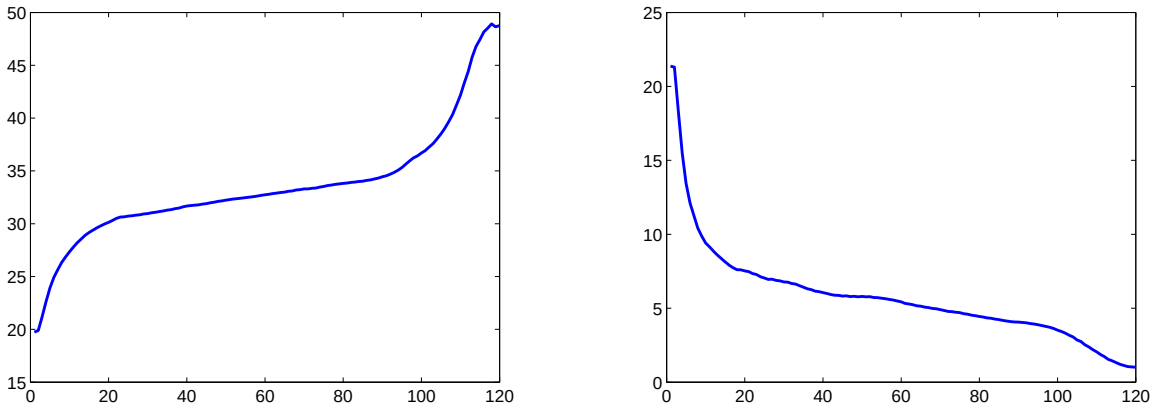


Fig. 2. Recovery error by different reconstruction methods. (a) by wavelet-TV method. (b) by iterative curvelet thresholding. (c) by ISS. (d) by ISS-curvelet with TV diffusivity.



(a) (b)

Fig. 3. The SNR vs. iterations (a) and recovery error vs. iterations (b) using our ISS-curvelet method.

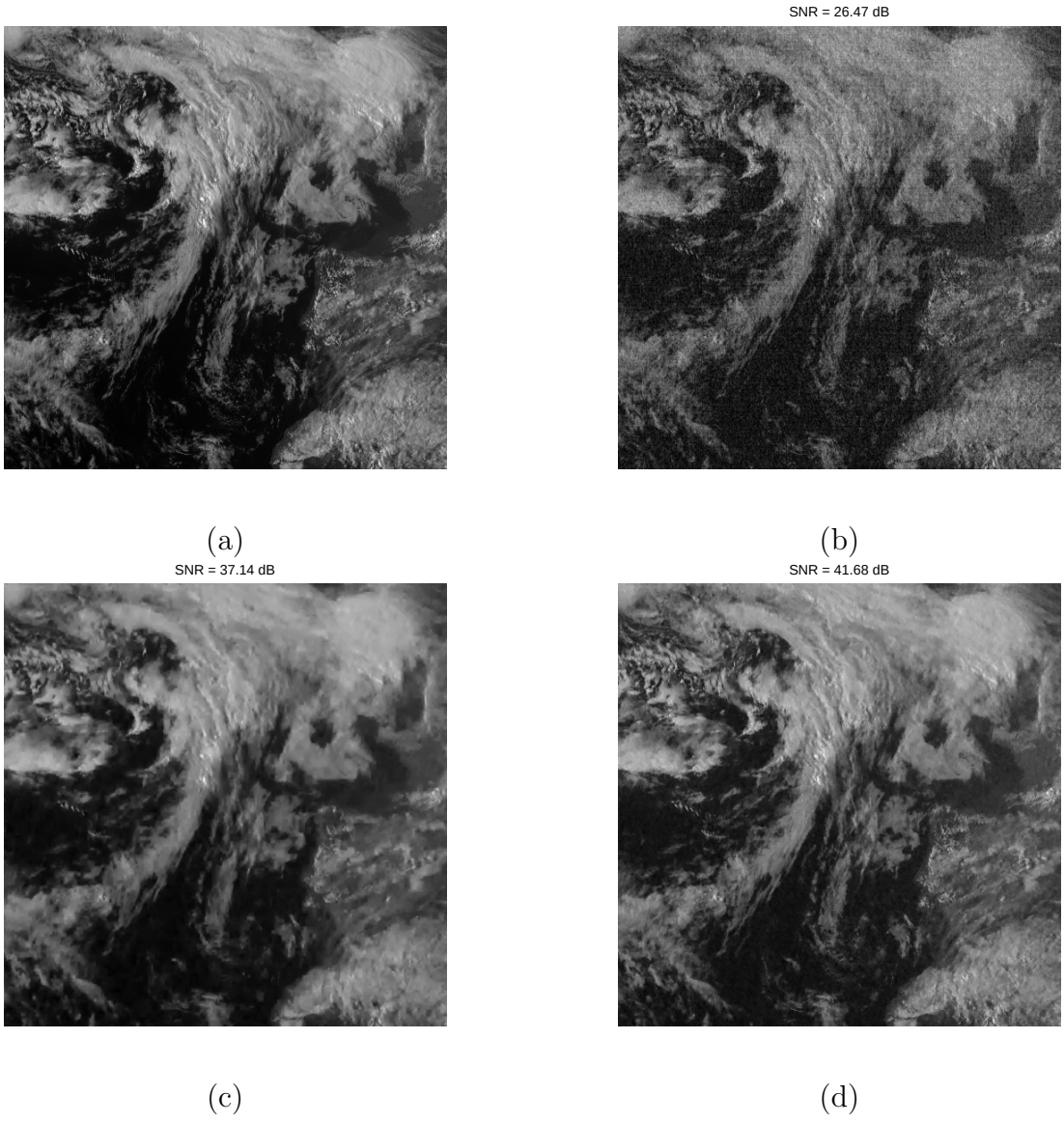


Fig. 4. CS for satellite remote sensing. (a) original image. (b) zero-filling reconstruction (SNR=26.47 dB). (c) $wavelet - TV$ reconstruction (SNR=37.14 dB). (d) our ISS-curvelet reconstruction (SNR=41.68 dB).