RESEARCH **Open Access**

A parameter-adaptive iterative regularization model for image denoising

Wenshu Li[1*], Chao Zhao[1], Qiegen Liu[2], Qingjiang Shi[1] and Shen Xu[1]

Abstract

In this article, an iterative regularization model (IRM) with adaptive parameter is addressed. IRM has gained a lot of attentions. But constant scale parameter becomes very sensitive for the fast convergence. It becomes very important to optimize the scale parameter adaptively. Therefore, we introduce a novel IRM with varying scale parameter because of the fact that when the scale parameter is smaller, the number of the iteration will enhance by IRM. A method to estimate a scale parameter is proposed according to the trend of the scale parameter. And the theoretical justification for this approach can be inferred. Numerical experiments show that the proposed methods with varying scale parameter can efficiently remove noise, reduce the number of iteration, and well preserve the details of images.

Keywords: Iterative regularization, Total variation, Variational methods, Image denoising

Introduction

During the last decade, in spite of the sophistication of the recently proposed methods, some algorithms have not yet attained a desirable level of applicability for image denoising, which is still a challenge at the crossing of functional analysis and statistics. The relations between variational regularization method and wavelet shrinkage have become one of the most active areas of research [1-5].

In this article, we are motivated by the following classical denoising problem of image degraded by additive white Gaussian noise. Given a noisy image $f(x, y): \Omega \rightarrow$, where Ω is a bounded open subset of σ^2, we want to obtain a decomposition equation:

$$f(x, y) = g(x, y) + n(x, y) \qquad (1)$$

where $g(x,y)$ is the true image and $n(x,y)$ is the noise with $(x, y) \in \Omega$ and $n(x, y)(0, \sigma^2)$

The most classical variational model is

$$u = \arg \min_{u \in \mathrm{BV}(\Omega)} \left\{ J(u) + \lambda \|f - u\|_2^2 \right\} \qquad (2)$$

or its corresponding constrained version

$$u = \arg \min_{u \in \mathrm{BV}(\Omega)} J(u) . \mathrm{s.t} .. \|f - u\|_2^2 = \sigma^2 \qquad (2a)$$

For some scale parameter $\lambda > 0$, where $\mathrm{BV}(\Omega)$ denotes the space of functions with bounded variation on Ω, $\bullet_2$ is L_2 norm. $J(u)$ is the regularization item and $\|f - u\|_2^2$ is the fitting item. λ is chosen to balance inconsistency (first term) and the deviation (second term) from the noise image $f(x,y)$ and depends on the noise norm σ. Therefore, a mass of researchers are concentrated on the regularization item $J(u)$. The total variation model of Rudin–Osher–Fatemi (ROF) for image denoising is considered to be the better denoising model. But, there were two serious issues about the ROF model [6-11]. First, it was very complicated to compute the solutions of the optimization problems induced by the variational method. Second, it was difficult to extract textures from images by using the ROF model. For the first issue, Goldstein and Osher recently introduced the split Bregman method for $L1$ regularized problems. The Bregman method gave rise to very efficient algorithms for solutions of the ROF model. Meyer [12] did some very interesting analysis by characterizing textures which he defines as "highly oscillatory patterns in image processing" as elements of the dual space of $\mathrm{BV}(\Omega)$. An iterative regularization model (IRM) [13], which replaces the regularization term by a

* Correspondence: Charlie@zstu.edu.cn
[1]Lab of Intelligence Detection and System, School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China
Full list of author information is available at the end of the article

generalized Bregman distance [14,15], was proposed. This model is formulated as

$$u_{k+1} = \arg \min_{u \in \mathrm{BV}(\Omega)} \left\{ J(u) + \frac{\lambda}{2} \|f + v_k - u\|_2^2 \right\} \qquad (3a)$$

$$v_{k+1} = v_k + f - u_{k+1} \qquad (3b)$$

Large λ corresponds to very little noise removal, and hence $u(x,y)$ is quickly close to $f(x,y)$ and the quality of image denoising is not effective. Small λ yields an over-smoothed $u(x,y)$ and the iterated times will be enhanced. In spite of the sophistication of the recently proposed methods, most algorithms have not yet attained a desirable level of applicability [15-18].

In this article, we proposed a new denoising method with varying scale parameter where the regularization item is $J(u) = \iint_{\Omega} |\nabla u| dx dy$. We deduce a method to gain the scale parameter from the iterative regularization. Finally, some numerical examples are presented and show that our method improves the quality of the image denoising and reduces the optimal number of iterations.

The remainder of this article is organized as follows. In "IRM" section, we mainly review IRM and its some attributes. The proposed method is introduced in "IRM with varying scale parameter" section; the experimental results of our method are given in "Result and discussion" section. This article is summarized in "Conclusion" section.

IRM
IRM makes use of some signals in the removed residual part for these denoising algorithms [19,20]. For $p \in \partial J(v)$, we define the non-negative quantity

$$D^p(u,v) \equiv D_J^p(u,v) \equiv J(u) - J(v) - <p, u - v> \qquad (4)$$

Then, the equivalent representation of Equation (3) is

$$u_{k+1} = \arg \min_{u \in \mathrm{BV}(\Omega)} D^{p_k}(u, u_k) + \frac{\lambda}{2} \|f - u\|_2^2 \} \qquad (5a)$$

$$p_{k+1} = p_k + f - u_{k+1} \qquad (5b)$$

where $u_0 = 0$ and $D_J^{p_k}(u, u_k)$ are the Bregman distance between u and u_k. As the optimal number of iteration k increases, u is close to the noisy image f. The scale parameter λ tunes the weight between the regularization and fidelity terms. The iterated refinement method yields a well-defined sequence of minimizers $\{u_k\}$ which satisfies $\|u_k - f\|_2^2 \leq \|u_{k-1} - f\|_2^2$ and if $f \in \mathrm{BV}(\Omega)$, then $\|u_k -$

$f\|_2^2 \leq \frac{J(f)}{k}$, i.e., u_k converges monotonically to f in $L^2(\Omega)$ with a rate of $\frac{1}{\sqrt{k}}$. For $g \in \mathrm{BV}(\Omega)$ and $\gamma > 1$, we have $D(g, u_k) \leq D(g, u_{k-1})$ subject to $\|u_k - f\|_2 \geq \gamma \|g - f\|_2$.

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 [13].

It should be stressed that the Bregman-based methodology, in the last few years, has made rapid development due to the tireless efforts of Osher and collaborators [18,21-23]. A key breakthrough among is that, with adequate initializations, the Bregman method equals to the augmented Lagrangian algorithm [7,22]. Furthermore, many efficient algorithms are proposed to enable fast implementation [21,24,25].

IRM with varying scale parameter
We know that for IRM the bigger the scale parameter λ is, the smaller the number of iteration is to the stop criterion, but u is quickly close to the noise image f, the quality of the image denoising is not ideal. When the scale parameter λ is smaller, the number of the iteration will enhance. Therefore, it is important to choose an optimal value λ.

Varying scale parameter
For that, let $J(u) = \iint_{\Omega} |\nabla u| dx dy$, differentiating both sides with respect to u for Equation (3a) we have

$$\nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right) + \lambda(f + v_k - u) = 0 \qquad (6)$$

Multiplying Equation (6) by $\nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right)$ and integrating over x and y, we get

$$\iint_{\Omega} \nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right) \nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right) dx dy$$
$$+ \lambda \iint_{\Omega} \nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right) (f + v_k - u) dx dy = 0 \qquad (7)$$

Then, we have the following equation

$$\lambda = \frac{\iint_{\Omega} \nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right) \nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right) dx dy}{\iint_{\Omega} \nabla \cdot \left(\frac{1}{|\nabla u|} \nabla u \right) (u - f - v_k) dx dy} \qquad (8)$$

In numerical implementation, we use accordingly λ_{k+1} denotes λ in Equation (8). Applying the proposed scale parameter to IRM with initial values $u^0 = 0, v^0 = 0$, we obtain different scale parameters λ_{k+1} for different iterations. Equation (3) should be written as

$$\tilde{\lambda}_{k+1} = \frac{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|}\nabla u_k\right)\nabla \cdot \left(\frac{1}{|\nabla u_k|}\nabla u_k\right)dxdy}{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|}\nabla u_k\right)(u_k - f - v_k)dxdy}$$

(9a)

$$u_{k+1} = \arg \min_{u \in BV(\Omega)} \left\{|\nabla u| + \lambda_{k+1}\|f + v_k - u\|_2^2\right\} \quad (9b)$$

$$v_{k+1} = v_k + (f - u_{k+1}) \quad (9c)$$

This gives us an adaptive value λ_{k+1}, which appears to converge as $k \to \infty$. The theoretical justification for this approach comes from Appendices 1 and 2.

Initial scale parameter

By the numerical experiment, we discover that the quality of image denoising is not ideal when initial values $u_0 = 0, v_0 = 0$. For example, if the initial condition holds, there is a question that Equation (9a) will be divided by zero.

If we randomly give an initial scale parameter value λ_0, we calculate λ_k by

$$\tilde{\lambda}_k = \frac{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|}\nabla u_k\right)\nabla \cdot \left(\frac{1}{|\nabla u_k|}\nabla u_k\right)dxdy}{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|}\nabla u_k\right)(u_k - f - v_k)dxdy}$$

(10)

after the iterations are taken some steps. We gain the sequence vector $\{\lambda_k\}$ and find that λ_k has some properties as follows:

(a) the sequence vector $\{\lambda_k\}$ is monotonically decreasing as the number of iteration k increases (see Figure 1c);

(b) as the number of iteration k increases, the sequence vector $\{\lambda_k\}$ will at first decrease, and then increase closely to λ_0 (see Figure 1b);

(c) the sequence vector $\{\lambda_k\}$ is monotonically increasing as the number of iteration k increases (see Figure 1a).

Therefore, we can obtain the initial value of varying scale parameter by the trend of the sequence vector $\{\lambda_k\}$ as follows:

(1) If the sequence vector $\{\lambda_k\}$ is monotonically decreasing at first as the number of iteration k increases, we consider that the random selected λ_0 is contented with the property of (a) or (b). Then, the initial scale parameter λ_1 of our proposed method is equal to λ_k^-. Usually, k is equal to 3.

(2) If the sequence vector $\{\lambda_k\}$ is monotonically increasing as the number of iteration k increases, the random selected λ_0 is contented with the property of (c). Then, the initial scale parameter of our proposed method $\lambda_1 = \lambda_1^-$ or $\lambda_1 = \lambda^0/p$ with the constant $p > 1$. Usually, $p = 2$.

In Figure 1, as the example of 'Barbara' image, the trends of the sequence vector $\{\lambda_k\}$ are gained when the scale parameter λ_0 is 8.33, 4.34, and 0.013, respectively.

IRM framework with varying scale parameter

According to the above two sections, our general iterative regularization procedure can be formulated as follows.

$$\begin{cases} u_{j+1} = \arg \min_{u \in BV(\Omega)} \left\{J(u) + \frac{\lambda 0}{2}\|f + v_j - u\|_2^2\right\} \\ v_{j+1} = v_j + f - u_{j+1} \end{cases} \quad (11)$$

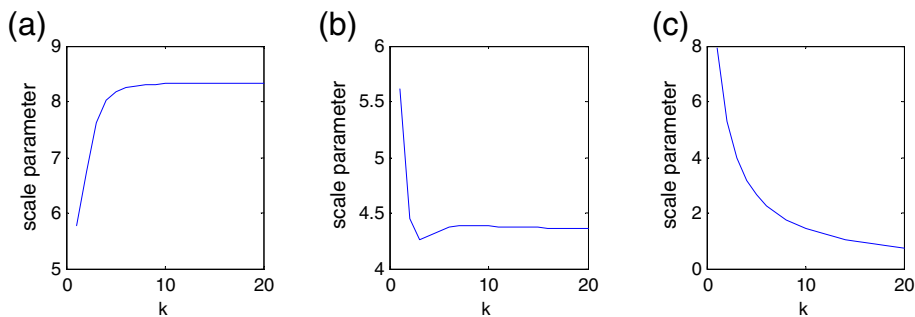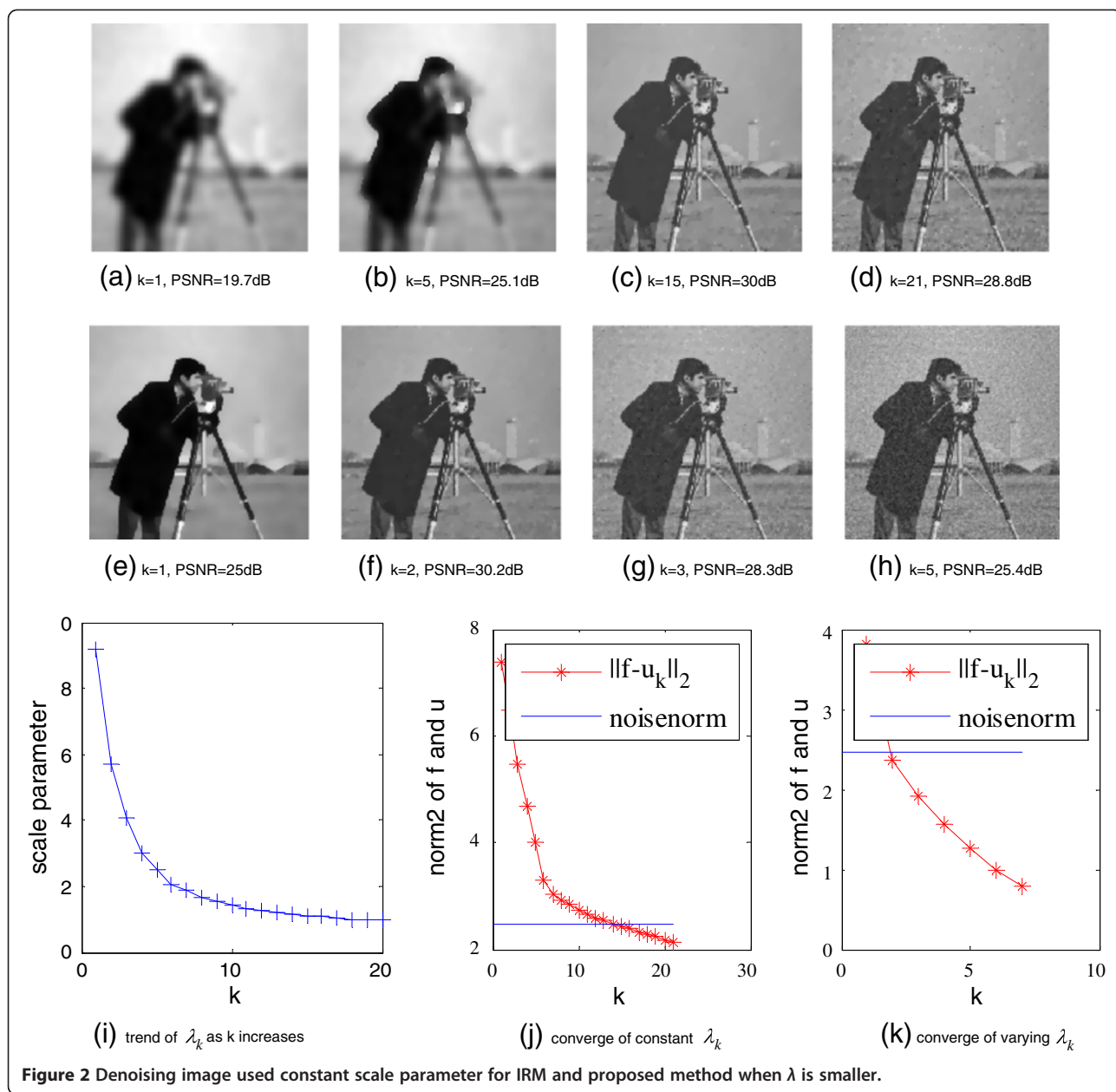Step 1: We randomly select λ_0. Let $u_0 = 0, v_0 = 0$ and $j = 0, 1, 2\ldots$

Figure 1 Trends of the scale parameter λ_k change as the number of iteration k increases.

Figure 2 Denoising image used constant scale parameter for IRM and proposed method when λ is smaller.

(1) According to Equations (11) and (10), we calculate u_{j+1}, v_{j+1}, and λ_j by the number of iteration j. Generally $j = 2$.

(2) We observe the trend of the sequence vector $\{\lambda_k\}$. According to the properties of "Initial scale parameter" section, we get the initial value λ_1 of our proposed method.

Step 2: Let $u_0 = 0$, $v_0 = 0$ and $k = 1, 2 \ldots$

(1) According to Equation (9) and the initial value λ_1, we calculate u_{k+1}, v_{k+1}, and λ_{k+1}.

(2) We get image u_k and stop the iteration when $\|f - u\|_k \leq \sigma$ (as the stopping criterion).

Result and discussion

All solutions to the variational problem were obtained using gradient descent in a standard fashion [21-28]. Now, we use Chambolle Algorithm [8]. The only non-trivial difficulty comes when $|\nabla u| \approx 0$. We fix this, as is usual, by perturbing $J(u) = \iint_{\Omega} |\nabla u| dxdy$ to $J_{\varepsilon}(u) = \iint_{\Omega} \sqrt{|\nabla u|^2 + \varepsilon^2} dxdy$, where ε is a small positive number.
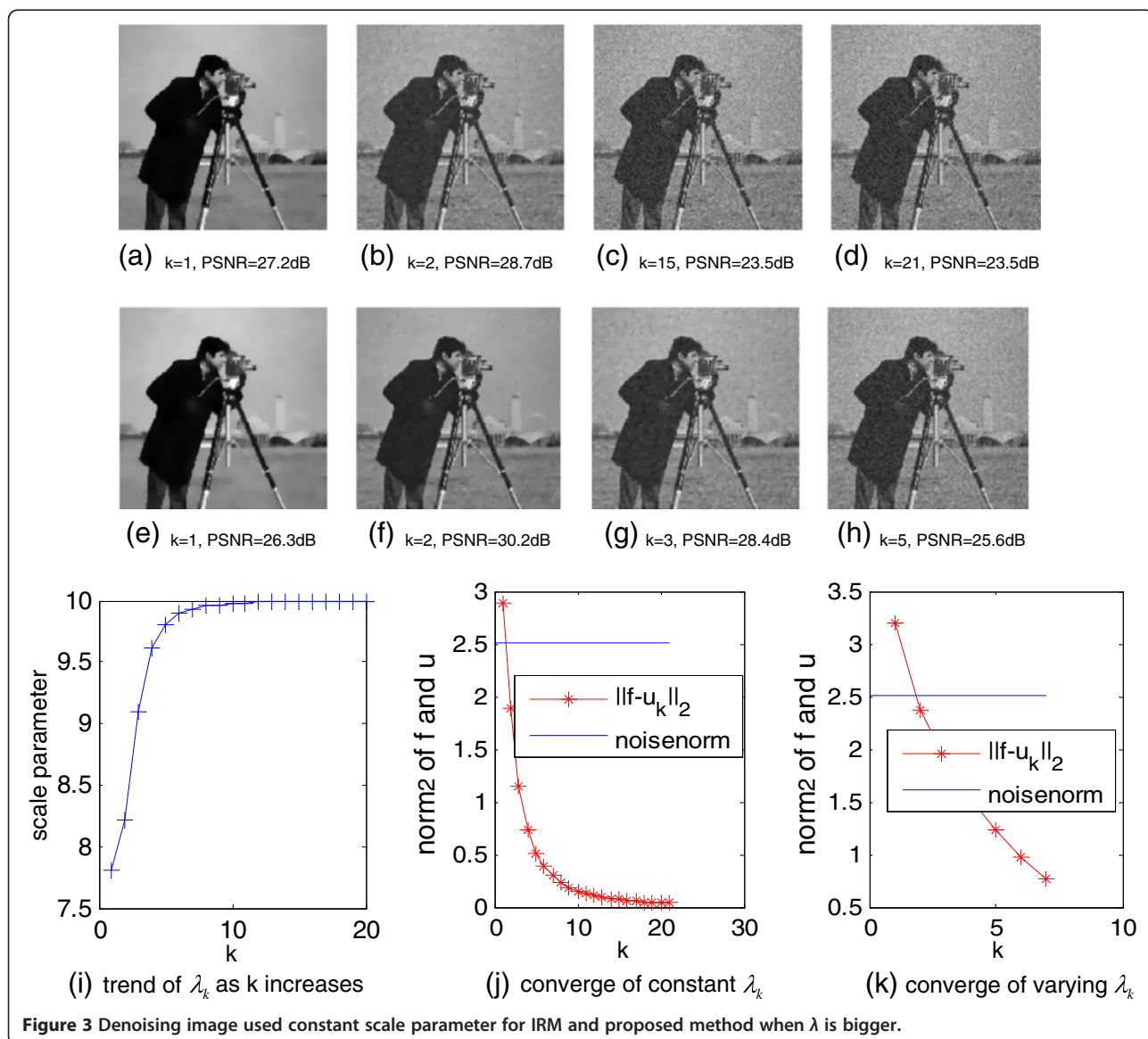
To be extent, the 'stair-casing' effect of this method can be decreased. In our calculations, we too $k = 10^{-12}$; the step of iteration unit τ for Chambolle Algorithm is 0.2. Without loss of generality, the performance of the denoising algorithms is measured in terms of peak signal-to-noise-ratio (PSNR) [29], which can be defined as follows

$$\text{PSNR} = 10 * \log_{10} \left(\frac{255^2}{\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} (f_{mn} - u_{mn})^2} \right) \tag{12}$$

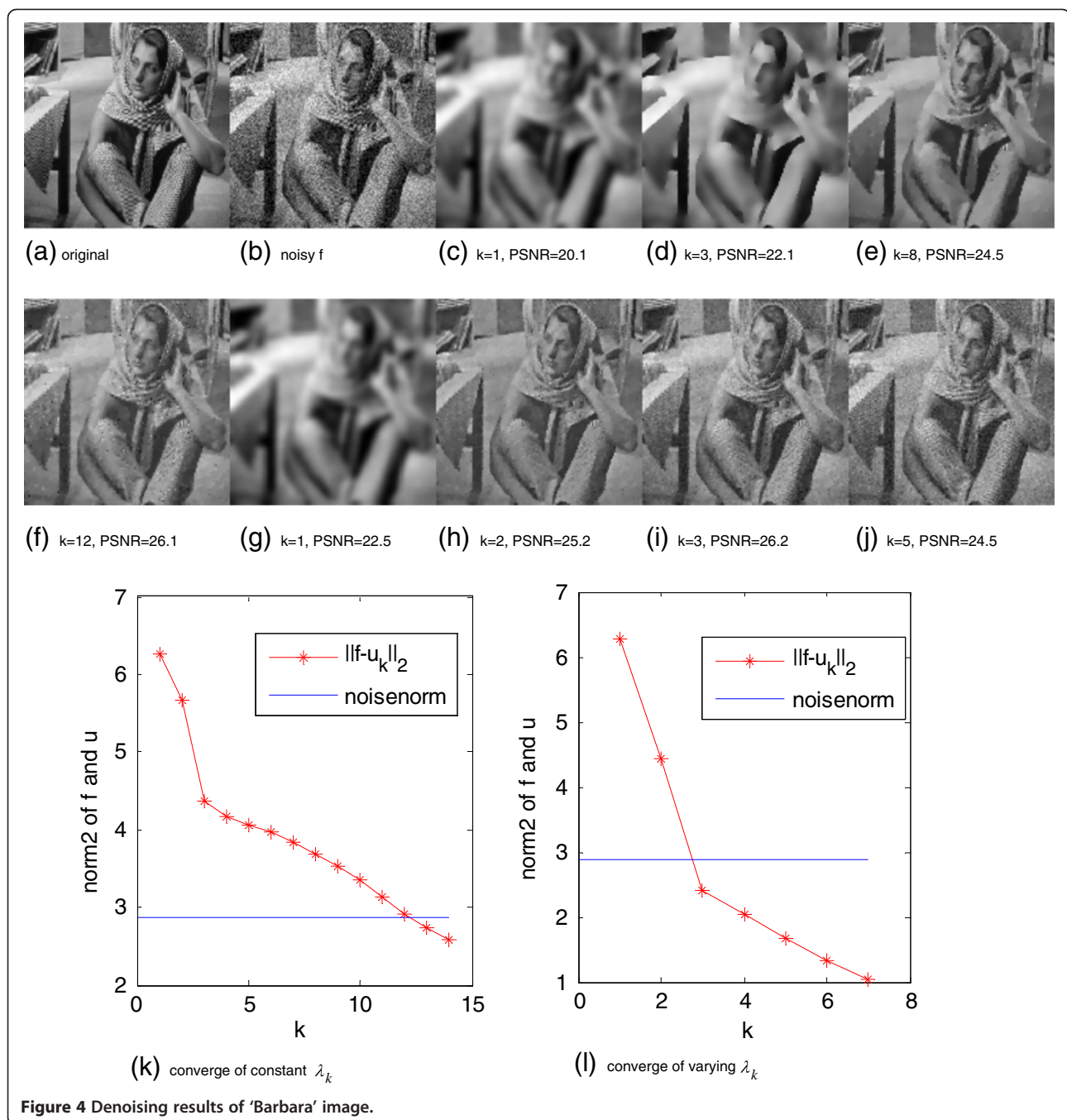where f is the original image and u is the denoising image.

Convergence analysis

Figures 2 and 3 show the results with constant and varying scale parameter of IRM with 'cameramen' image added Gauss noise $\sigma = 20$ when λ is smaller and bigger, respectively. In Figure 2, the first row results show that more iteration steps are required to stop criterion with smaller scale parameter $\lambda_0 = 0.67$; the second row results show that our proposed methods require less iterations to get the optimal denoising results. At first, we used constant scale parameter $\lambda_0 = 0.67$ to iterate three times and got a sequence vector $\{\lambda_k\}$ decreased in the first image of the third row. According to Equation (11), we got the initial value $\lambda_1 = \lambda_2 = 5.74$. The last two plots (*i*) and (*j*) show that $\|f - u\|_{k2}$ decreases monotonically with the iteration, first dropping below σ at the optimal iterate $k = 12$ and 2, respectively. It shows that our proposed

(a) k=1, PSNR=27.2dB (b) k=2, PSNR=28.7dB (c) k=15, PSNR=23.5dB (d) k=21, PSNR=23.5dB

(e) k=1, PSNR=26.3dB (f) k=2, PSNR=30.2dB (g) k=3, PSNR=28.4dB (h) k=5, PSNR=25.6dB

(i) trend of λ_k as k increases (j) converge of constant λ_k (k) converge of varying λ_k

Figure 3 Denoising image used constant scale parameter for IRM and proposed method when λ is bigger.
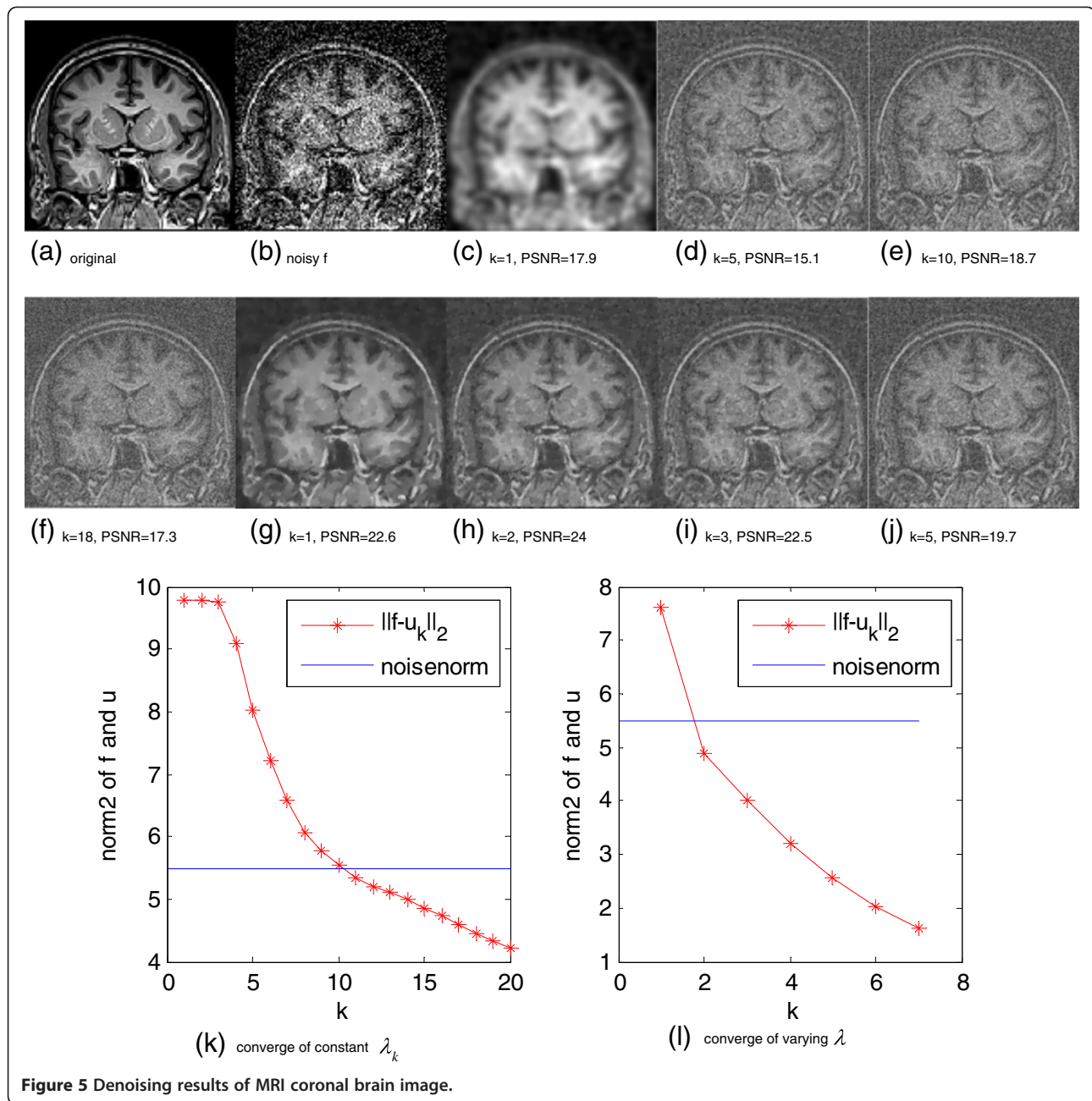
method converge faster than IRM with the constant scale parameter. In Figure 3, as can be seen, with large scale parameter $\lambda_0 = 10$ the original IRM convergences to the noisy image f quickly, and only one iterative needed to reach the stop criterion. Obviously, the denoising result is not satisfied. However, promising result is obtained by our varying scale parameter strategy where the initial value $\lambda_1 = \lambda_2 = 5.74$ according to Equation (10).

Preserved textures analysis

Figure 4 shows the denoising results of 'Barbara' image with Gauss white noise σ = 25.5. The constant scale parameter λ for IRM is 1. Compared with the constant scale parameter for IRM in Figure 4c–f, our proposed method can preserved more textures in Figure 4g–j. The last two plots (k) and (l) show that $\|f - u\|_{k2}$ decreases monotonically with the iterations, first dropping below σ at the optimal iterate $k = 12$ and 2, respectively. It shows that

(a) original   (b) noisy f   (c) k=1, PSNR=20.1   (d) k=3, PSNR=22.1   (e) k=8, PSNR=24.5

(f) k=12, PSNR=26.1   (g) k=1, PSNR=22.5   (h) k=2, PSNR=25.2   (i) k=3, PSNR=26.2   (j) k=5, PSNR=24.5

(k) converge of constant λ_k

(l) converge of varying λ_k

Figure 4 Denoising results of 'Barbara' image.

Figure 5 Denoising results of MRI coronal brain image.

our proposed method converges faster than IRM with the constant scale parameter.

Denoising analysis for MRI coronal brain
The denoising results of MRI coronal brain image with Gauss white noise $\sigma = 53.83$ are shown in Figure 5. The constant scale parameter λ for IRM is 1. Compared with the constant scale parameter for IRM in Figure 5c–f, our proposed method can preserve more textures in Figure 5g–j. The last two plots (k) and (l) show that $\|f - u\|_{k2}$ decreases monotonically with the iterations, first dropping below σ at the optimal iterate $k = 13$ and 3,

respectively. It shows that our proposed method converges faster than IRM with the constant scale parameter and has more texture details in the denoised image.

Computational cost analysis
We have made a comparison in terms of computational time (see Tables 1 and 2) by MATLAB 7.1, which is used

Table 1 Computer time of formula (9)

	Computing (9a)	Computing the sub-problem (9b)
Computational time	0.026 s	1.281 s

Table 2 Computer time of our method and the conventional IRM

	Our method	The conventional IRM
Computational time	3.136 s	13.708 s

on a PC equipped with AMD 2.31 GHz CPU and 3 GB RAM. In fact, it should be noted that the term "Fast" is relatively for computing (9a) when compared with computing the sub-problem (9b), i.e., in this article even we use the efficient Chambolle Algorithm to solve the sub-problem (9b) and set the inner iterations is 40. To take the results of "Preserved textures analysis" section for denoising 'Barbara' image (256 × 256) as an example, the average time of computing (9b) once is 1.281 s, while that of (9a) once is only 0.026 s. Moreover, combined with fact that the outer iterations of the conventional IRM is 13 while our adaptive IRM is 3, the whole computational time of the conventional IRM is 13.708 s while that of our method is 3.136 s. Therefore, we know that our adaptive scheme is really faster than the conventional one.

In addition, we have made a comparison between wavelet + wiener, curvelets include hard threshold, soft threshold, and block threshold regulation algorithm and our method. Denoising results of Lena image are shown that our algorithm improves PSNR than traditional method in Table 3.

Conclusion

A novel IRM with adaptive scale parameter is proposed in order to decrease the sensitivity of constant scale parameter, optimize the scale parameter adaptively in the IRM, and attain a desirable level of applicability for image denoising. We replace the classic regularization item and deduce the equation of the adaptive scale parameter, because we know that the scale parameter is smaller, the number of the iteration will enhance by IRM. Then, the rule of varying scale parameter by the trend of the sequence vector is attained. A new iterative scale parameter λ is obtained according to the trend of the sequence vector. In general, we can get the initial scale parameter λ just using three steps of iteration. We have seen with practical examples that our proposed method can reduce

Table 3 A variety of image denoising algorithms are compared for Lena image

Gauss white noise σ	10	15	20	25	30
Wavelet + wiener	30.65	28.06	26.94	25.48	24.39
Hard threshold of curvelet	31.67	29.58	28.47	26.43	24.97
Soft threshold of curvelet	31.99	30.76	29.78	28.79	26.34
Block threshold of curvelet	32.76	32.33	31.09	29.57	28.08
Our method	33.97	33.68	32.59	30.14	29.64

the number of iterations. Thus, a fast and robust method is got.

Appendix 1

A constrained problem is defined as

$$\begin{cases} \min F(X) \\ s.t. N^T X - b \geq 0 \end{cases} \tag{13}$$

We know that the constraint equation is $N^T X - b = 0$. The basic assumption is that X lies in the subspace tangent to the active constraints, i.e., $X_{i+1} = X_i + \alpha S$, where S is the direction with the most negative directional derivative and α is the iterative step length, both X_i and X_{i+1} satisfy the constraint equations. Therefore, we obtain

$$N^T S = 0 \tag{14}$$

If we want the steepest descent direction satisfying Equation (14), we can pose the problem as

$$\begin{cases} \min S^T \nabla F \\ s.t. N^T S = 0 \text{ and } S^T S = 1 \end{cases} \tag{15}$$

The Euler–Lagrange equation (15) has the formulation

$$L(S, \lambda, \mu) = S^T \nabla F - S^T N \lambda - \mu(S^T S - 1) \tag{16}$$

The derivative of L with respect to S is

$$\frac{\partial L}{\partial S} = \nabla F - N\lambda - 2\mu S = 0 \tag{17}$$

Recall that $N^T S = 0$ in Equation (14) and multiplying Equation (17) by N^T, we get

$$N^T \nabla F - N^T N \lambda = 0 \tag{18}$$

Therefore, we get the value

$$\lambda = \frac{N^T \nabla F}{N^T N} \tag{19}$$

So, the proposition holds.

Appendix 2

In this article, we want to solve the question

$$u_{k+1} = \arg \min_{u \in \text{BV}(\Omega)} \left\{ \iint_\Omega |\nabla u| dx dy + \frac{\lambda}{2} \|f + v_k - u\|_2^2 \right\} \tag{20}$$

It is equivalent to

$$u_{k+1} = \arg\min_{u\varepsilon BV(\Omega)} \left\{ \lambda_1 \iint_\Omega |\nabla u| dxdy + \frac{1}{2} \|f + v_k - u\|_2^2 \right\} \tag{21}$$

where $\lambda_1 = \frac{1}{\lambda}$. Then, Equation (21) is rewritten to

$$\begin{cases} \min_{u\varepsilon BV(\Omega)} \frac{1}{2} \|f + v_k - u\|_2^2 \\ \text{s.t.} \iint_\Omega |\nabla u| dxdy \geq 0 \end{cases} \tag{22}$$

Since $\iint_\Omega |\nabla u| dxdy = \|\nabla u_1\| \approx \sqrt{\|\nabla u\|_1^2 + \varepsilon} = < \frac{\nabla^* \nabla u}{\sqrt{\|\nabla u\|_1^2 + \varepsilon}}, u >$, we let $\frac{\nabla^* \nabla u}{\sqrt{\nabla u_1^2 + \varepsilon}}$ be approximated by $\frac{\nabla^* \nabla u_k}{\sqrt{\|\nabla u_k\|_1^2 + \varepsilon}}$, and let $u = X, b = 0, N = \frac{\nabla^* \nabla u_k}{\sqrt{\|\nabla u_k\|_1^2 + \varepsilon}}$ and $(u) = \frac{1}{2} \|f + v_k - u\|_2^2$, we have $\iint_\Omega |\nabla u| dxdy == N^T u \geq 0$, then according to Appendix 1, we obtain

$$\lambda_1 = \frac{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|} \nabla u^k\right)(u_k - f - v_k) dxdy}{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|} \nabla u^k\right) \nabla \cdot \left(\frac{1}{|\nabla u_k|} \nabla u_k\right) dxdy} \tag{23}$$

Therefore, we obtain the parameter

$$\lambda = \frac{1}{\lambda_1} = \frac{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|} \nabla u_k\right) \nabla \cdot \left(\frac{1}{|\nabla u_k|} \nabla u_k\right) dxdy}{\iint_\Omega \nabla \cdot \left(\frac{1}{|\nabla u_k|} \nabla u_k\right)(u_k - f - v_k) dxdy} \tag{24}$$

So, the proposition holds.

Author details
[1]Lab of Intelligence Detection and System, School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China. [2]School of Life Science and Technology, Shanghai Jiao Tong University, Shanghai 200240, China.

References
1. L Rudin, S Osher, E Fatemi, Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)
2. R Devore, B Jawerth, B Lucier, Image compression through wavelet transform coding. IEEE Trans. Inf. Theory **38**, 719–746 (1992)
3. DL Donoho, De-noising by soft-threshold. IEEE Trans. Inf. Theory **41**(3), 613–627 (1995)
4. K Dabov, A Foi, V Katkovnik, K Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. **16**, 2080–2095 (2007)
5. R Yan, L Shao, SD Cvetkovic, J Klijn, Improved nonlocal means based on pre-classification and invariant block matching. IEEE/OSA J. Disp. Technol. **8**(4), 212–218 (2012)
6. S Osher, A Solè, L Vese, Image decomposition and restoration using total variation minimization and the H-1 norm. SIAM J. Multiscale Model. Simulat. **1**(3), 349–370 (2003)
7. M Burger, G Gilboa, S Osher, JJ Xu, Nonlinear inverse scale space methods. Commun. Math. Sci. **4**(1), 175–208 (2006)
8. A Chambolle, R DeVore, NY Lee, B Lucier, Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. IEEE Trans. Image Process. **7**(3), 319–335 (1998)
9. A Chambolle, BJ Lucier, Interpreting translation-invariant wavelet shrinkage as a new image smoothing scale space. IEEE Trans. Image Process. **10**, 993–1000 (2001)
10. G Steidl, J Weickert, T Brox, P MraZek, 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)
11. 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)
12. Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations* (American Mathematical Society, Boston, 2002)
13. S Osher, M Burger, D Goldfarb, J Xu, W Yin, An iterative regularization method for total variation based image restoration. Multiscale Model. Simulat. **4**, 460–489 (2005)
14. L Bregman, The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. U.S.S.R. Comput. Math. Math. Phys. **7**, 200–217 (1967)
15. BB Hao, M Li, XC Feng, Wavelet iterative regularization for image restoration with varying scale parameter. Signal Process. Image Commun. **23**(6), 433–441 (2008)
16. B Liu, K King, M Steckner, J Xie, J Sheng, L Ying, Regularized sensitivity encoding (SENSE) reconstruction using Bregman iterations. Magn. Reson. Med. **61**(1), 145–152 (2009)
17. A Chambolle, An algorithm for total variation minimization and applications. J. Math. Imag. Vis. **20**, 89–97 (2004)
18. L Vese, S Osher, Modeling textures with total variation minimization and oscillating patterns in image processing. J. Sci. Comput. **19**, 553–572 (2003)
19. T Weissman, E Ordentlich, G Seroussi, S Verdu, M Weinberger, Universal discrete denoising: known channel. IEEE Trans. Inf. Theory **51**, 5–28 (2005)
20. P Perona, J Malik, Scale space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**, 629–639 (1990)
21. T Goldstein, S Osher, The split Bregman method for L1-regularized problems. SIAM J. Imag. Sci. **2**(2), 323–343 (2009)
22. W Yin, S Osher, D Goldfarb, J Darbon, Bregman iterative algorithms for l1-minimization with applications to compressed sensing. SIAM J. Imag. Sci. **1**, 142–168 (2009)
23. L He, TC Chang, S Osher, T Fang, P Speier, *MR image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods*, 2006. Technical Report CAM Report 06–35, UCLA
24. E Esser, *Applications of Lagrangian-based alternating direction methods and connections to split Bregman*, 2009. CAM Report 09–31, UCLA
25. Y Wang, J Yang, W Yin, Y Zhang, A new alternating minimization algorithm for total variation image reconstruction. SIAM J. Imag. Sci. **1**(3), 248–272 (2008)
26. J Ma, Compressed sensing by inverse scale space and curvelet thresholding. Appl. Math. Comput. **206**(2), 980–988 (2008)

27. J Ma, DF Le, Deblurring from highly incomplete measurements for remote sensing. IEEE Trans. Geosci. Remote Sensing (GeoRS) **3**(47), 792–802 (2009)
28. G Gilboa, N Sochen, YY Zeevi, Texture preserving variational denoising using an adaptive fidelity term, in *Proc VLSM Nice, France* **1**, 2003, pp. 137–144
29. L Shao, H Zhang, G de Haan, An overview and performance evaluation of classification-based least squares trained filters. IEEE Trans. Image Process. **17**, 1772–178 (2008)