Regularized Sensitivity Encoding (SENSE) Reconstruction Using Bregman Iterations

Bo Liu, Kevin King, Michael Steckner, Jun Xie, Jinhua Sheng, and Leslie Ying

In parallel imaging, the signal-to-noise ratio (SNR) of sensitivity encoding (SENSE) reconstruction is usually degraded by the ill-conditioning problem, which becomes especially serious at large acceleration factors. Existing regularization methods have been shown to alleviate the problem. However, they usually suffer from image artifacts at high acceleration factors due to the large data inconsistency resulting from heavy regularization. In this paper, we propose Bregman iteration for SENSE regularization. Unlike the existing regularization methods where the regularization function is fixed, the method adaptively updates the regularization function using the Bregman distance at different iterations, such that the iteration gradually removes the aliasing artifacts and recovers fine structures before the noise finally comes back. With a discrepancy principle as the stopping criterion, our results demonstrate that the reconstructed image using Bregman iteration preserves both sharp edges lost in Tikhonov regularization and fine structures missed in total variation (TV) regularization, while reducing more noise and aliasing artifacts. Magn Reson Med 61: 145–152, 2009. © 2008 Wiley-Liss, Inc.

Key words: SENSE; parallel imaging; total variation regularization; Bregman iteration; compressed sensing

Parallel magnetic resonance imaging (MRI), as a fast imaging method, uses an array of RF receiver surface coils to acquire multiple sets of undersampled k-space data simultaneously. Over the past few years, a number of parallel imaging techniques have been proposed for image reconstruction from these undersampled data. These methods include image-domain methods such as sensitivity encoding (SENSE) (1) and partially parallel imaging with localized sensitivities (PILS) (2); k-space methods such simultaneous acquisition of spatial harmonics (SMASH) (3), generalized autocalibrating partially parallel acquisitions (GRAPPA) (4), and parallel MRI with adaptive radius in k-space (PARS) (5); and hybrid methods such as sensitivity profiles from an array of coils for encoding and reconstruction in parallel (SPACE-RIP) (6). Several practical difficulties prevent parallel MRI from fully achieving maximal speeds (7). A common issue with parallel imaging is signal-to-noise ratio (SNR) degradation. An intrinsic SNR reduction is inevitable due to the reduced number of data samples. In SENSE, the reconstruction problem is formulated as solving a set of linear equations

\[ Ef = d, \]

where \( d \) is the vector formed from all k-space data acquired in all channels, \( f \) is the unknown vector defining the desired full field of view (FOV) image to be computed (both with a lexicographical column ordering of the two-dimensional (2D) array components), and \( E \) is the encoding matrix. These equations can be very ill-conditioned depending on the coil configurations and sampling trajectories, which further deteriorates SNR dramatically in the reconstructed image, especially when a large acceleration factor is used. The problem has been partially addressed by optimizing coil geometry (6–11), optimizing sampling trajectory (12,13), or introducing regularizations (14–19).

Among these techniques, regularization does not need modifications in hardware or data acquisition. It usually requires some additional prior information in reconstruction, but some iterative reconstruction methods without explicit regularization terms are also known to have inherent regularization capability when stopped properly (20–22). Tikhonov regularization has been widely investigated primarily due to the existence of a closed-form solution. In the general Tikhonov framework, the SENSE reconstruction with arbitrary trajectories is given by (17–19):

\[ f_{\text{reg}} = \arg \min_f \| d - Ef \|_2^2 + \lambda \| A^H A f - f \|_2^2, \]

where \( \| \cdot \|_2 \) is the \( L_2 \) norm, \( A \) is a positive semidefinite matrix, and the regularization parameter \( \lambda \) is a positive constant chosen to balance the data inconsistency (first term) and the deviation (second term) from the prior image \( f \), known as the regularization image. A closed-form solution for \( f_{\text{reg}} \) exists and is given by

\[ f_{\text{reg}} = f + (E^H \psi^{-1} E + \lambda A^H A)^{-1} E^H \psi^{-1} (d - Ef), \]

where \( \psi \) is the noise correlation matrix. In previous studies (17–19), the matrix \( A \) was assumed to be the identity matrix; the regularization parameter \( \lambda \) was set heuristically as a constant for the entire image, or automatically in a spatial dependent fashion using the L-curve method (17); and the regularization image \( f \) was set to zero or a low-resolution image was generated from the Nyquist-sampled autocalibration data acquired near the center of the k-space (23,24). The regularization image, usually of poor quality (e.g., low resolution), introduces bias in Tikhonov regularization, which is imparted as image blurring, or residual aliasing artifacts at high reduction factors (17).
Some edge-preserving regularization methods have recently been proposed for SENSE, such as total variation (TV) regularization (25,26), EPGRAM (16), and MAP-SENSE (27). These methods replace the smoothness prior in the second term of Eq. [2] by an edge-preserving prior that imposes a relationship between neighboring pixels, and the reconstruction usually requires iterative algorithms. A drawback of these methods is the possible loss of texture due to the piecewise smooth assumption (16).

Recently, an iterative regularization technique based on Bregman distance was developed in image processing for denoising (28). The Bregman iteration method improves on the TV regularization by gradually recovering the fine-scale structures that are usually lost in the TV regularization due to the piecewise smooth constraint, and has found some success in conventional MR image reconstruction with reduced radial samples (29,30).

In this work, we investigate SENSE regularization using Bregman iteration to address the issues with the existing SENSE regularization methods. Unlike the existing methods, where the regularization term is fixed, the proposed method iteratively updates the regularization function based on the Bregman distance between the reconstructions from two consecutive iterations such that the regularization effect is adaptively reduced. The method is equivalent to TV regularization at the first iteration, but it gradually recovers more fine structures as the iteration continues, and eventually converges to the basic SENSE reconstruction without regularization when the suppressed noise comes back. Our phantom and in vivo results demonstrate that when stopped properly, the Bregman iteration method is better than the competing regularization methods in suppressing noise, maintaining fine details, and reducing aliasing artifacts.

**THEORY**

The method starts with the TV regularization for SENSE (25,31), given by

$$ f_{\text{eq}} = \arg \min_f \| d - Ef \|^2 + \lambda |f|_{\text{TV}}. \tag{4} $$

where the regularization term is called the TV norm of an image, which is defined as a function of the image gradient:

$$ |f|_{\text{TV}} = \sum \sqrt{\left| \nabla_x f \right|^2 + \left| \nabla_y f \right|^2}, \tag{5} $$

where $f$ denotes the 2D image, $\nabla_x$ and $\nabla_y$ denote the gradient along $x$ and $y$, respectively, and $| \cdot |$ denotes the complex modulus. The TV prior is based on the fact that the highly oscillatory noise usually increases the TV of an image (31). TV regularization is known to be edge-preserving, but may result in loss of texture due to the assumption that the image TV is small (30). The detail textures usually have large local variations, which lead to its increased TV value. As a result, the minimization of Eq. [4] in TV regularization reduces fine structures at the same time it suppresses noise in reconstruction.

To better preserve fine structures in reconstruction, we next use an iterative procedure, called Bregman iteration (32,33), to improve upon TV regularization. Specifically, instead of stopping at the solution to Eq. [4] (denoted as $f_1$ for the first iteration), we adaptively refine TV regularization by updating the regularization function iteratively. At the $k$th iteration, the image is reconstructed by (28)

$$ f_k = \arg \min_{f} \| d - Ef \|^2 + \lambda D(f, f_{k-1}), \text{ for } k>1, \tag{6} $$

where $D(f, f_k)$ is the Bregman distance between $f$ and $f_k$ associated with the TV norm, defined as (32,33)

$$ D(f, f_k) = |f|_{\text{TV}} - |f_k|_{\text{TV}} - \langle f - f_k, \partial(|f_k|_{\text{TV}}) \rangle. \tag{7} $$

where $\langle \cdot, \cdot \rangle$ denotes the inner product, and $\partial(|f|_{\text{TV}})$ is an element of the subgradient of the TV norm at point $f_k$. The Bregman distance associated with the TV norm is graphically represented in Fig. 1. As a measure of the convexity of the TV norm, it indicates the increase in $|f|_{\text{TV}}$ over $|f_k|_{\text{TV}}$ above linear growth with slope $\partial(|f_k|_{\text{TV}})$. Since the TV norm is a convex function, $D(f, f_k)$ is also convex in $f$ for each $f_k$, and thus the optimization of Eq. [6] has a unique solution.

Using TV regularization with Bregman iteration, it has been shown (28) that the sequence $E_k$ monotonically converges to the acquired noisy data $d$ in the $L_2$ sense, i.e., the reconstruction $f_k$ approaches the unregularized basic SENSE reconstruction. More importantly, for $\lambda$ sufficiently large, the Bregman distance between the reconstruction $f_k$ and the true, noise-free image $f$ also monotonically de-
creases, when the iteration number \( k \) is less than \( \tilde{k} \), where \( \tilde{k} \) is an integer that satisfies

\[
\tilde{k} = \max_{k \in \mathbb{N}} \|d - Ef_k\| \geq \tau \|d - Ef_0\|. \tag{8}
\]

for any constant \( \tau > 1 \). When the iteration further continues with \( k \) greater than \( \tilde{k} \), the reconstruction deviates away from the true, noise-free image as it converges to the unregularized SENSE reconstruction. Due to this semiconvergence property of Bregman iteration, a stopping criterion is needed to obtain a reconstruction closest to the clean, true image before the noise comes back. The trade-off between data consistency (residual aliasing artifacts reduction) and regularization (SNR improvement) is controlled by the stopping criteria. Based on Eq. [8], we use a discrepancy principle (28) that stops the iterative procedure when the data inconsistency \( \|d - Ef_k\| \) is reduced to below the measurement noise level for the first time. The noise level is estimated by \( \|d_y\|_2 \), where \( d_y \) denotes the multichannel k-space data from a dummy scan sampled at the same locations for \( d \). In contrast to Tikhonov and TV regularizations, where the choice of \( \lambda \) heavily affects the final reconstruction quality, the \( \lambda \) in Bregman iteration has little effect on image quality as long as it is sufficiently large, as demonstrated in the Results section. It has been proved mathematically (28) that for \( \lambda \) sufficiently large, the images will have more and more fine structures than the TV regularized reconstruction as the iteration continues. When stopped according to the discrepancy principle, a larger \( \lambda \) will increase the number of iterations needed to reach the stopping criterion, where the reconstruction stays almost the same. On the other hand, if \( \lambda \) is too small, the TV regularization weighting will be too small and the reconstruction will be too close to the conventional SENSE to observe the semiconvergence.

### MATERIALS AND METHODS

To simplify implementation of the minimization in Eq. [6], we use its equivalent representation (28):

\[
f_k = \arg \min_{f} \left\{ \|d + v_{k-1} - Ef_k\|_2^2 + \lambda \|f\|_T \right\}, \tag{9}
\]

where \( v_k = v_{k-1} + d - Ef_k \) for \( k > 0 \), \( v_0 = 0 \). It is seen that the above minimization in each Bregman iteration is the same as the TV regularization formulation (Eq. [4]) except that \( d \) is replaced by \( d + v_{k-1} \) in the first term. It becomes TV regularization when the iteration index is \( k = 1 \). The standard algorithms for the variation problem with the Rudin-Osher-Fatemi (ROF) model, such as time-marching (31), fixed-point (34), and primal-dual algorithms (35), can be used to solve the minimization problem in Eq. [9]. Here we employ the iterative time-marching method, a detailed description of which is given in the Appendix. All code was written in MATLAB (The MathWorks, Natick, MA, USA).

In our experiments, two data sets were acquired. In the phantom experiment, a set of data was collected on a Hitachi Airis Elite (Kashiwa, Chiba, Japan) 0.3T permanent magnet scanner with a four-channel head coil and a single-slice spin-echo sequence (TE/TR = 40/1000 ms, 8.4 kHz bandwidth [bw], matrix size = 256 \times 256, FOV = 220 mm²). In the human experiment, a set of in vivo brain data was acquired on a 3T commercial scanner (GE Healthcare, Waukesha, WI, USA) and eight-channel head coil (Invivo, Gainesville, FL, USA) with a 2D TI-weighted spin-echo protocol (axial plane, TE/TR = 11/700 ms, 22-cm FOV, 10 slices, matrix size = 256 \times 256). Informed consent was obtained from the volunteer in accordance with the institutional review board policy.

To demonstrate its performance, we first applied the proposed method to a set of simulated noise-free data generated from the phantom experiment. To simulate the noise-free case, both the sum-of-square (SoS) reconstruction and the coil sensitivities were first estimated from the actual data set, and then their product was Fourier transformed to generate k-space data. One out of every three phase encodings was kept to simulate a reduction factor of 3.

We then applied the method to the experimental data. For both the phantom and brain data sets, the full k-space data were acquired and their SoS reconstructions were used as the reference for comparison. One out of every three and four phase encodings was then manually kept to simulate reduction factors of 3 and 4 for the phantom and in vivo data, respectively. The central 32 fully sampled phase encodings were truncated by a cosine taper window (36) and used to generate a set of low-resolution images. These images provide estimated coil sensitivity profiles after normalization by their SoS reconstruction.

In reconstruction, basic SENSE, Tikhonov regularization, TV regularization (first Bregman iteration), and Bregman iteration were used to generate the final images based on the reduced data. The error image was calculated as the difference between the reconstruction and the reference image. Each Bregman iteration included 30 time-marching steps. The discrepancy principle was used as the stopping rule for Bregman iterations. The number of Bregman iterations used for the reconstruction was displayed on the top-right corner of the image. In Tikhonov regularization, the low-resolution SoS reconstruction was used as the regularization image, and the regularization parameter was automatically chosen by the L-curve method (17).

### RESULTS

The results for the simulated data are shown in Fig. 2. It can be seen that in the absence of noise and with exact knowledge of coil sensitivities, basic SENSE is able to reconstruct the original image perfectly. In contrast, the TV-regularized image has residual aliasing artifacts due to data inconsistency, which is clearly seen in the error images on the bottom of Fig. 2. As Bregman iterations continue beyond TV regularization, the aliasing artifacts disappear, and the reconstruction eventually converges to the true image as well as the basic SENSE reconstruction. Because there is no noise in this case, the reconstructions keep getting closer to the original image as the number of iterations increases, as shown in the convergence curve for the normalized mean squared error (NMSPE) (37) in Fig. 3. In both TV and Bregman, \( \lambda = 8 \) was used.
The results for the phantom experimental data are shown in Fig. 4; \( \lambda = 80 \) was used for the Bregman iteration. For a reduction factor of 3 with four channels, the basic SENSE reconstruction has high noise due to the ill-conditioning problem. All regularization methods are able to suppress noise with different degrees. However, both Tikhonov and TV regularizations result in aliasing artifacts because at relatively large reduction factors, the regularization term has to be weighted heavily for noise reduction, and hence the data inconsistency is well beyond noise level. In contrast, the Bregman iteration is able to remove the aliasing artifacts almost completely as it approaches the data consistency gradually, which is better seen in the error images on the bottom of Fig. 4. Because the phantom image mostly satisfies the piecewise smooth constraint in TV regularization, the Bregman iteration does not bring in much additional detail. In this case, the Bregman iteration is used primarily to remove the residual aliasing artifacts in TV regularization.

The reconstructions from the eight-channel in vivo data with a reduction factor of 4 are shown in Fig. 5. TV regularization is able to suppress more noise than Tikhonov regularization, but both result in residual aliasing artifacts. In addition, some high-intensity details are lost due to the piecewise smooth constraint in TV regularization, which makes the image look blurry. In comparison, the Bregman iteration is able to remove the aliasing artifacts and also bring back the fine structures in the original image. The improvement of Bregman iteration over TV regularization in reduction of artifacts is clearly seen in the error images on the bottom of Fig. 5. According to the discrepancy principle, the method stops at seven Bregman iterations. The parameter \( \lambda = 12.5 \) was used. To demonstrate how Bregman iterations gradually remove the artifacts, then bring in details, and finally bring back noise, we show in Fig. 6 a sequence of images with different numbers of Bregman iterations, and in Fig. 7 the corresponding error images for \( \lambda = 12.5 \). For actual data, the Bregman iterations also converge to the basic SENSE reconstruction. We also compare in Fig. 8 the NMSE as a function of the number of Bregman iterations when \( \lambda = 12.5 \) and \( \lambda = 25 \) are used in Eq. [6]. It is shown that Bregman iterations have a semiconvergence behavior, and the convergence speed is affected by the regularization parameter \( \lambda \). Large \( \lambda \) converges slowly and small \( \lambda \) converges rapidly, but the lowest NMSEs of the two curves are almost the same. The final reconstructions based on the discrepancy principle are also the same visually (so only \( \lambda = 12.5 \) is shown in Fig. 6). When \( \lambda \) becomes too small, the NMSE of reconstruction diverges and the noise immediately comes back at the second Bregman iteration before the aliasing artifacts are completely removed. It is also worth noting that the discrepancy principle does not necessarily give the least NMSE, but usually the most satisfying reconstruction visually. For example, for \( \lambda = 12.5 \), although four iterations give the least image error, there are still visible aliasing artifacts, as better seen in the error images in Fig. 7. In comparison, the reconstruction after seven iterations determined by the stopping rule gives fewer artifacts with a slight increase in noise.

In all results, each Bregman iteration, composed of 30 time-marching steps, took 3.4 s on an HP XW8400 work-
station with 2.3 GHz CPU and 4 GB RAM, which about doubles the running time for basic SENSE.

**DISCUSSION**

In summary, our results demonstrate the following properties of Bregman iteration for SENSE regularization. First, the Bregman iteration converges to the basic SENSE reconstruction as the iteration number goes to infinity. Second, as $\lambda$ becomes sufficiently large, the Bregman iteration first removes the aliasing artifacts, then recovers the fine structures, and finally recovers noise. Therefore, when stopped properly, Bregman iteration is able to remove these artifacts while still keeping the noise low. Third, the reconstruction barely depends on the value of the regularization parameter $\lambda$ as long as it is larger than a threshold. The reconstruction converges most rapidly at this threshold. As $\lambda$ increases, the convergence speed is reduced, and the convergence curve is flatter at the minimum, as seen in Fig. 8. It suggests that the recovery of fine structures before noise would be slower to observe and the reconstruction is less sensitive to inaccurate noise levels used for the stopping rule. Therefore, a practical approach is to choose a rather large $\lambda$ (e.g., 100) to guarantee reconstruction quality for a wide range of image contents and noise levels (0.3T phantom and 3T brain images). On the other hand, we have found heuristically that doubling $\lambda$ would approximately half the convergence rate and thus double the reconstruction time. This prolonged reconstruction time may not be a concern as computers become faster. If it is a concern, optimization of $\lambda$ for fast convergence using iteration-adaptive $\lambda$ may be of interest for future study. Finally, the computational complexity of each Bregman iteration is sufficiently low for practical use. The benefit of the Bregman iteration method is especially important at large reduction factors when Tikhonov and TV regularizations fail to reconstruct high-quality images even at the best compromise between data consistency and

![FIG. 4. Phantom images reconstructed from a set of actual four-channel data with a reduction factor of 3. The error images for TV and Bregman are shown on the bottom row. Basic SENSE reconstruction has large noise. Tikhonov and TV regularizations introduce aliasing artifacts when suppressing noise. Bregman iteration is able to remove these artifacts while still keeping the noise low.](image)

![FIG. 5. Brain images reconstructed from a set of actual eight-channel data with a reduction factor of 4. The error images for TV and Bregman with respect to the reference are shown on the bottom row. The acceleration was along the anterior-to-posterior direction. Similarly to the phantom results, Bregman iteration reduces the aliasing artifacts in Tikhonov and TV regularizations. In addition, Bregman iteration reconstruction has more fine structures and noise than TV-regularized reconstruction which is seen to be blocky.](image)
noise reduction. In this case, the aliasing artifacts are already apparent (suggesting too-heavy regularization) while the noise is still large (suggesting insufficient regularization) in conventional regularized reconstructions. In addition, although the results presented here are based on reconstructions using data from Cartesian trajectories, the proposed Bregman iteration method should be easily extended to non-Cartesian trajectories.

As a regularization method for SNR enhancement, it is very interesting to see how much the proposed method improves on SNR with different parameters. The NMSE in the Results provides a combined metric for image quality since it includes contributions from both noise and artifacts. A pure SNR metric such as the g-factor map (1) is also interesting. In Tikhonov regularization, there is a closed-form expression for the g-factor map to characterize the spatially-dependent noise amplification (17), which suggests that the larger the regularization parameter $\lambda$, the better is the SNR achieved. However, with an iterative method, such a g-factor map is difficult to calculate analytically. Here we calculate empirically the noise amplification. Specifically, we replace the measured data by pure, independently, identically distributed complex Gaussian noise and reconstruct the image using different reconstruction methods. The reconstruction represents the spatial distribution of noise amplification, and is thus named a pseudo g-factor map. Although slight noise correlation exists between coils, we found its effect to be low on the g-factor map of basic SENSE, and therefore the correlation is ignored in generating the pseudo g-factor map. Using the sensitivity maps from the in vivo brain data, Fig. 9 shows the pseudo g-factor map in comparison with the g-factor map for basic SENSE. The pseudo g-factor map is seen to be close to the g-factor map in representing the spatial distribution of noise. Figure 10 shows the pseudo g-factor map for different numbers of Bregman iterations in parallel with Fig. 6. Similar to the conclusion for Tikhonov
regularization (17), the larger the $\lambda$, the lower the noise is in TV regularization. For fixed $\lambda = 12.5$, the noise increases with the number of Bregman iterations and eventually approaches that of basic SENSE. At seven iterations, the noise level is still low and quite uniform spatially, which demonstrates the effectiveness of the proposed method in noise suppression.

The definition of Bregman distance used for Bregman iteration in this paper is based on the TV norm. Bregman iterations are simple to extend to other norms, such as the $L_1$ norm of wavelet transform. Different norms might bring an additional benefit for Bregman iteration and should also be an interesting topic for future study.

Bregman iteration can also find applications in compressed sensing (38,39). In compressed sensing with encoding matrix $E$ and sparse transformation $W$, the signal is reconstructed from measurement $d$ by the constrained nonlinear convex program:

$$\text{Minimize} \|Wf\|_1, \text{ subject to } Ef = d. \quad [10]$$

Equation [10] can be solved robustly using interior point methods (40) but with high computational complexity. It has been proven (41) that Bregman iteration also approaches the solution to Eq. [10] with much less computational complexity. In addition, when stopped according to the discrepancy principle, Bregman iteration is also an alternative method for solving the following program for compressed sensing in the case of measurement noise:

$$\text{Minimize} \|Wf\|_1, \text{ subject to } \|Ef - d\|_2 < \epsilon \quad [11]$$

Although they did not call it compressed sensing, Chang et al. (29) used Bregman iteration for compressed sensing with success. Bregman iteration may also be applied to other recent work on compressed sensing in MRI (42–44).

**APPENDIX**

In the time-marching algorithm (31), the Euler-Lagrange equation for the cost function in Eq. [9] is first obtained:

$$0 = -\lambda \nabla \cdot \left( \frac{\nabla f}{\|\nabla f\|} \right) - 2E�(d + v_{k-1} - Ef). \quad [A1]$$

We then establish the following time-dependent partial differential equation:

$$\delta = \lambda \nabla \cdot \left( \frac{\nabla f}{\|\nabla f\|} \right) + 2E�(d + v_{k-1} - Ef), \quad [A2]$$

which updates the solution iteratively according to

$$f_{k+1} = f_k + \Delta t \cdot \left( \lambda \nabla \cdot \left( \frac{\nabla f_k}{\|\nabla f_k\|} \right) + 2E�(d + v_{k-1} - Ef_k) \right). \quad [A3]$$

where $w$ is the index for the number of time-marching steps, $f_0$ is the initial guess for the desired image, and $\Delta t$ is the step size. The solution is guaranteed to converge independently of the initial image if $\Delta t$ satisfies the Courant-Friedrichs-Lewy condition (31). Here we use the low-resolution SoS reconstruction instead of the zero image to speed up convergence. The step size was chosen to be $\Delta t = 1$.

![FIG. 10. Pseudo g-factor maps for different numbers of Bregman iterations. The noise level is low and uniform spatially with a small iteration number. It increases and eventually approaches that of SENSE as the iteration number increases.](image)
REFERENCES


