An Augmented Lagrangian Method for Regularized MRI Reconstruction Using SENSE

S. Ramani[1], and J. A. Fessler[1]
[1]EECS Department, University of Michigan, Ann Arbor, Michigan, United States

Introduction: Standard reconstruction methods for SENSitivity Encoding (SENSE) [1, 2] suffer from SNR degradation due to *k*-space undersampling and instability arising from correlation in sensitivity maps. Nonquadratic regularization can be used to improve reconstruction quality but demands computation intensive nonlinear optimization. We present a new algorithm for (nonquadratic) regularized reconstruction from sensitivity-encoded data—SENSE-reconstruction—using the augmented Lagrangian (AL) formalism that constitutes a powerful framework for solving large-scale constrained optimization problems. Based on numerical experiments with *in-vivo* human brain data, we demonstrate that the proposed AL algorithm converges faster compared to general-purpose algorithms such as nonlinear conjugate gradient (NCG) (that has been used for compressed sensing MRI [3]) and state-of-the-art Monotone Fast Iterative Shrinkage-Thresholding Algorithm (MFISTA) [4].

Problem Formulation: We employ a combination of TV and l_1-regularization and formulate regularized SENSE-reconstruction as **P0:** $\mathbf{x}_{recon} = arg\ min_{\mathbf{x}}\ 1/2\ \|\mathbf{d} - \mathbf{FSx}\|^2 + \lambda_1\ \|\mathbf{Wx}\|_1 + \lambda_2\ \mathrm{TV}\{\mathbf{x}\}$, where $\mathbf{d}$ represents undersampled data from a coil-array, $\mathbf{F}$ and $\mathbf{S}$ represent the Fourier encoding and sensitivity-map matrices, respectively, $\mathbf{W}$ represents 2-levels of undecimated Haar-wavelet transform (excluding "scaling" coefficients), and $\mathrm{TV}\{\mathbf{x}\} = \sum_i (\sum_n |[\mathbf{R}_n\ \mathbf{x}]_i|^2)^{1/2}$ is composed of matrices $\mathbf{R}_n$ that correspond to finite-differences along the n-th spatial dimension of the image, n=1,2,…d, $[\mathbf{x}]_i$ denotes the i-th element of $\mathbf{x}$, and λ_1, and λ_2 are regularization parameters. The above formulation can be used to deal with correlated noise after applying a suitable noise-decorrelation procedure [2].

Method: We write **P0** as an equivalent constrained optimization problem **P1:** $arg\ min_{\mathbf{u}}\ f(\mathbf{u})$ subject to $\mathbf{u}_0 = \mathbf{Sx}$, $\mathbf{u}_1 = \mathbf{Ru}_2$, and $\mathbf{u}_2 = \mathbf{x}$, where $\mathbf{R} = [\mathbf{W}^{\mathrm{T}}\ \mathbf{R}_{11}{}^{\mathrm{T}} \ldots \mathbf{R}_{1d}{}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{u}_1 = [\mathbf{u}_{10}{}^{\mathrm{H}}\ \mathbf{u}_{11}{}^{\mathrm{H}} \ldots \mathbf{u}_{1d}{}^{\mathrm{H}}]^{\mathrm{H}}$ whose components $\{\mathbf{u}_{1n}\}$ correspond to the block-rows of $\mathbf{R}$, $\mathbf{u} = [\mathbf{u}_0\ \mathbf{u}_1\ \mathbf{u}_2\ \mathbf{x}]$, and $f(\mathbf{u}) = 1/2\ \|\mathbf{d} - \mathbf{Fu}_0\|^2 + \lambda_1\ \|\mathbf{u}_{10}\|_1 + \lambda_2 \sum_i (\sum_n |[\mathbf{u}_{1n}]_i|^2)^{1/2}$. This type of constrained reformulation is different from those in [5, 6]. We construct an AL function [7] for **P1** using a Lagrange multiplier $\boldsymbol{\eta} = [\boldsymbol{\eta}_0\ \boldsymbol{\eta}_1\ \boldsymbol{\eta}_2]$ as $L(\mathbf{u}, \boldsymbol{\eta}, \mu) = f(\mathbf{u}) + \mu/2\ \|\mathbf{u}_0 - \mathbf{Sx} - \boldsymbol{\eta}_0\|^2 + \mu\nu_1/2\ \|\mathbf{u}_1 - \mathbf{Ru}_2 - \boldsymbol{\eta}_1\|^2 + \mu\nu_2/2\ \|\mathbf{u}_2 - \mathbf{x} - \boldsymbol{\eta}_2\|^2$, where $\mu, \nu_1, \nu_2 > 0$ are scalars that do not affect the solution of **P1** (and **P0**). The traditional AL scheme [7] for solving **P1** alternates between jointly minimizing L with respect to $\mathbf{u} = [\mathbf{u}_0\ \mathbf{u}_1\ \mathbf{u}_2\ \mathbf{x}]$ for a fixed $\boldsymbol{\eta}$ and updating $\boldsymbol{\eta}$. The joint-minimization step can be computationally expensive, so we apply an *alternating minimization* scheme (that minimizes L alternatively with respect to $\mathbf{u}_0$, $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{x}$ one at a time while holding the others at their most recent values). This simplifies the minimization leading to the following **AL** algorithm for solving **P1** (and **P0**).

AL algorithm:

Select $\mathbf{u}^{(0)}$, $\boldsymbol{\eta}^{(0)}$ *and* $\mu, \nu_1, \nu_2 > 0$; *set iteration number* $m = 0$;

Repeat until stop criterion is met;

1. $\mathbf{u}_0{}^{(m+1)} = arg\ min_{\mathbf{u}_0}\ 1/2\ \|\mathbf{d} - \mathbf{Fu}_0\|^2 + \mu/2\ \|\mathbf{u}_0 - \mathbf{Sx}^{(m)} - \boldsymbol{\eta}_0{}^{(m)}\|^2$;
2. $\mathbf{u}_1{}^{(m+1)} = arg\ min_{\mathbf{u}_1}\ \lambda_1\ \|\mathbf{u}_{10}\|_1 + \lambda_2 \sum_i (\sum_n |[\mathbf{u}_{1n}]_i|^2)^{1/2} + \mu\nu_1/2\ \|\mathbf{u}_1 - \mathbf{Ru}_2{}^{(m)} - \boldsymbol{\eta}_1{}^{(m)}\|^2$;
3. $\mathbf{u}_2{}^{(m+1)} = arg\ min_{\mathbf{u}_2}\ \nu_1/2\ \|\mathbf{u}_1{}^{(m+1)} - \mathbf{Ru}_2 - \boldsymbol{\eta}_1{}^{(m)}\|^2 + \nu_2/2\ \|\mathbf{u}_2 - \mathbf{x}^{(m)} - \boldsymbol{\eta}_2{}^{(m)}\|^2$;
4. $\mathbf{x}^{(m+1)} = arg\ min_{\mathbf{x}}\ 1/2\ \|\mathbf{u}_0{}^{(m+1)} - \mathbf{Sx} - \boldsymbol{\eta}_0{}^{(m)}\|^2 + \nu_2/2\ \|\mathbf{u}_2{}^{(m+1)} - \mathbf{x} - \boldsymbol{\eta}_2{}^{(m)}\|^2$;
5. $\boldsymbol{\eta}_0{}^{(m+1)} = \boldsymbol{\eta}_0{}^{(m)} - (\mathbf{u}_0{}^{(m+1)} - \mathbf{Sx}^{(m+1)})$;
6. $\boldsymbol{\eta}_1{}^{(m+1)} = \boldsymbol{\eta}_1{}^{(m)} - (\mathbf{u}_1{}^{(m+1)} - \mathbf{Ru}_2{}^{(m+1)})$;
7. $\boldsymbol{\eta}_2{}^{(m+1)} = \boldsymbol{\eta}_2{}^{(m)} - (\mathbf{u}_2{}^{(m+1)} - \mathbf{x}^{(m+1)})$;
8. $m = m + 1$;

Step 2 of the **AL** algorithm is a denoising problem (that can further be decoupled in terms of the individual elements of $\mathbf{u}_1$) whose solution is obtained using shrinkage operations [5]. Steps 1, 3, and 4 involve quadratic costs and therefore have closed-form updates with simple matrix-inverses that can be computed efficiently. We adjust μ, ν_1, and ν_2 based on easily computed condition numbers of these matrix-inverses for fast convergence of the proposed **AL** algorithm.

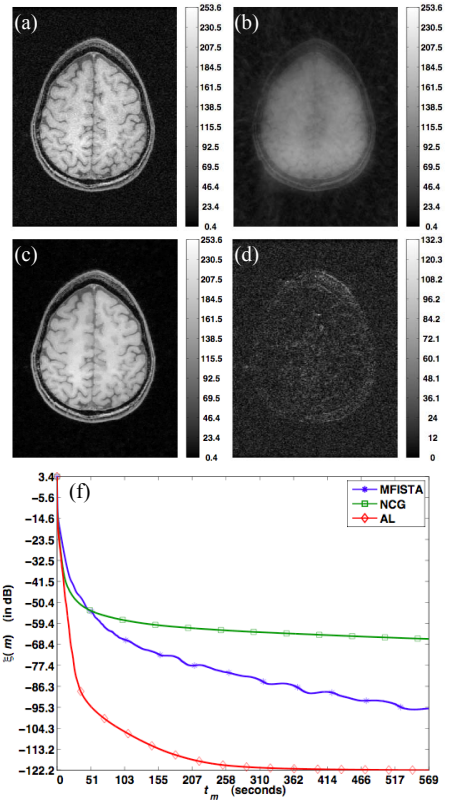

Fig. 1 *Experiment with a 2-D slice of a in-vivo human brain data-set: (a) Body-coil image; (b) Zero-filled iFFT reconstruction; (c) Regularized SENSE-reconstruction* $\mathbf{x}^*$*; (d) Absolute difference between (a) and (c); (e) Poisson-disk-based undersampling pattern (on a Cartesian grid); (f) Plot of* $\xi(m)$ *versus* t_m *for* **AL***, NCG and MFISTA.*

Results: We used a 3-D *in-vivo* human brain data-set (256x144 uniformly-spaced samples in the phase-encode plane, and 128 samples along the read-out direction) acquired from a GE 3T scanner (T_R = 25 ms, T_E = 5.172 ms, and voxel-size = 1x1.35x1 mm^3), with a 8-channel head-coil. Figure 1a shows one slice of the iFFT-reconstruction of fully-sampled data collected with a body-coil for comparison. To estimate the sensitivity maps $\mathbf{S}$, we separately optimized a quadratic-regularized least-squares criterion that encouraged smooth maps that "closely" fit a low-resolution version of the body-coil image in Figure 1a to low-resolution head-coil images, all obtained from iFFT-reconstruction of corresponding central 32x32 phase-encodes. We estimated the inverse of noise covariance matrix from data collected during a dummy scan where only the static magnetic-field (and no RF excitations) was applied and carried out noise-decorrelation of data as described in [2]. We then performed regularized SENSE-reconstruction of the 2-D slice (corresponding to Figure 1a) from undersampled phase-encodes where we applied the Poisson-disk-based undersampling pattern (with a reduction factor of ≈ 6) in Figure 1e in the phase-encode plane. We obtained the reconstruction $\mathbf{x}^*$ (that represents a solution of **P0**) numerically by running thousands of iterations of MFISTA. Figure 1c shows $\mathbf{x}^*$ where aliasing artifacts and noise have been suppressed considerably compared to the conventional (square-root of sum-of-squares) zero-filled iFFT-reconstruction in Figure 1b. We compared the proposed **AL** algorithm to NCG and MFISTA in terms of speed of convergence to the reconstruction $\mathbf{x}^*$ by computing the normalized l_2-distance $\xi(m)$ between the estimate $\mathbf{x}^{(m)}$ and $\mathbf{x}^*$ as $\xi(m) = 20\ \log_{10}(\|\mathbf{x}^{(m)} - \mathbf{x}^*\|_2 / \|\mathbf{x}^*\|_2)$. We evaluated $\xi(m)$ for all algorithms and plotted it as a function of algorithm run-time t_m (time elapsed from start until iteration m) in Figure 1f. It is seen that **AL** converges faster than NCG and MFISTA. We obtained promising results (not shown) similar to that in Figure 1 for experiments with several slices of the above human brain data-set and with various synthetic data-sets and a real breast-phantom data-set.

Conclusions: The augmented Lagrangian (AL) formalism combined with alternating minimization can be used for solving large-scale unconstrained optimization problems effectively by posing them as equivalent constrained tasks. We adopted such an approach for regularized SENSE-reconstruction and developed an **AL** algorithm that converges faster than NCG and MFISTA, is simple to implement and can also be easily extended to handle (combinations of) several nonquadratic regularization criteria.

Acknowledgements: This work was supported by the Swiss National Science Foundation under fellowship PBELP2-125446 and in part by the National Institutes of Health under grant P01 CA87634. The authors would like to thank Dr. Jon-Fredrik Nielsen, University of Michigan, for providing the *in-vivo* human brain data-set used in the experiments.

References: [1] Pruessmann et al, 1999, *MRM*, 42: 952-962; [2] Pruessmann et al, 2001, *MRM*, 46: 638-651; [3] Lustig et al, 2007, *MRM,* 58: 1182-1195; [4] Beck et al, 2009, *IEEE TIP*, 18: 2419-2434; [5] Goldstein et al, 2009, *SIAM J. Img. Sci.*, 2: 323-343; [6] Afonso et al, 2010, *IEEE TIP*, 19: 2345-2356; [7] Bertsekas, 1976, *Automatica*, 12: 133-145.