Adaptive regularization in compressed sensing using the discrepancy principle

K. F. King[1], L. Marinelli[2], and C. J. Hardy[2]
[1]Applied Science Lab, GE Healthcare, Waukesha, WI, United States, [2]GE Global Research, Niskayuna, New York, United States

Introduction

Compressed sensing (1) can be combined with image space parallel imaging (SENSE) (2). The image m is typically reconstructed by minimizing an objective function $J(m) = \|Em - y\|_2^2 + \lambda\|\Psi m\|_1$ where y is the measured (multicoil) k-space data, E is the coil encoding operator that includes both the B1 receive field for each coil and the Fourier operator, Ψ is a transform that produces a compressible function when operating on the image m, and the constant λ is adjusted to balance data fidelity and artifact reduction (first and second terms in J). If λ is chosen too small, uncorrected aliasing remains in the image. If λ is chosen too large, the image appears blurry. Many methods have been developed to automatically determine the optimal λ (3). Most are problematic for MRI data because either the assumptions are inapplicable or the computational burden is too high. Here we focus on a simple method called the discrepancy principle that chooses λ based on the size of the first term in J, also called the discrepancy term. A version of this method has been used for GRAPPA regularization (4).

Methods

Let the terms in J be denoted $J1 = \|\Psi m\|_1$ and $J2 = \|Em - y\|_2^2$. A small λ gives small J2, making the reconstructed image consistent with the measured k-space data. In particular, the noise in Em will be consistent with the noise in y, resulting in a noisy image. Conversely, a larger λ gives smaller J1, reducing aliasing, and resulting in a relatively larger J2 with noise in Em that is lower than the noise in y, giving a denoised image. The discrepancy principle chooses λ large enough to give J2 about the same as σ^2, the variance of y. This gives some denoising because the noise in Em is less than the noise in y. The noise variance σ^2 can be easily measured using a short acquisition without gradients or RF. A simple method for finding the optimal λ starts with a λ that is too large. The image that minimizes J is obtained and J2 is computed. The process terminates if $J2 < \alpha\sigma^2$ where $\alpha \sim 1$, otherwise λ is reduced and the process repeated using the final m from the previous iteration as the initial guess for the next one. One complication is that the coil sensitivity component of E must be estimated from a small number of Nyquist samples near the center of k-space. The low spatial resolution approximation for E can cause Em to differ from y. A slightly better method is to stop reducing λ when $J2 < \varepsilon + \alpha\sigma^2$ where ε is called the model error. The model error can be estimated by assuming that $\varepsilon = J2$ when $\lambda = 0$, solving for m with $\lambda = 0$, and then computing J2.

A 3T commercial scanner (GE Healthcare, Waukesha, WI) was used to scan a volunteer using both a single channel transmit/receive quadrature head coil and an 8-channel receive-only head coil with a T1-weighted 2D spin echo pulse sequence. Fully sampled data were randomly undersampled in the ky direction to simulate compressed sensing acquisition. The image was reconstructed by minimizing J using a conjugate gradient (CG) algorithm implemented in MATLAB (The Mathworks, Natick, MA). To estimate ε a fixed total reduction factor of three was used while varying the number of Nyquist lines. The single channel and 8-channel model errors ε were compared for varying numbers of Nyquist lines. The single channel coil should have very small ε because of the uniform coil sensitivity, and hence differences in ε between the two coils should result from the low spatial resolution estimation of E. To test the discrepancy principle, an optimal λ was first empirically determined. The starting λ was roughly 4 times this value, and λ was reduced by $\sqrt{2}$ after each solution for m. The solution for m was found using a maximum of 16 CG iterations for each value of λ. A 2D Daubechies-4 wavelet was used for Ψ.

Results

For the single channel coil, with $\lambda = 0$, J2 was essentially zero after a single CG iteration, independent of the number of Nyquist lines, indicating that ε is negligible for this case as expected. $J2/\sigma^2$ for the 8-channel coil is shown in Fig. 1 as a function of the number of Nyquist lines. For 8 Nyquist lines, J2 is comparable to σ^2 but decreases as expected as the number of Nyquist lines increases, indicating that coil sensitivity spatial resolution is one source of ε. The k-space undersampling pattern shown in Fig. 2 (3-fold reduction, 32 Nyquist lines) was used to evaluate the discrepancy principle. The 8-channel coil images are shown with $\lambda = 0$ (Fig. 3), empirical optimal $\lambda = 1500$ (Fig. 4), and discrepancy principle with initial $\lambda = 6000$ and final $\lambda = 1500$ (Fig. 5). For the discrepancy principle reconstruction, $\varepsilon = 0$ and $\alpha = 1$ were used to find the final λ. Although $\varepsilon = 0$ and $\alpha = 1$ gave acceptable results here, in general these would not be the optimal values.

Conclusions

The discrepancy principle is a method for finding optimal regularization parameters for SENSE parallel imaging combined with compressed sensing. A similar approach has been used for GRAPPA parallel imaging. One source of model error is the low spatial resolution estimation of coil sensitivity. A sufficient number of Nyquist samples for the coil estimate results in negligible model error. The method currently only works for a single sparsifying transform (single λ). This is probably adequate for most cases, however some images may benefit from two sparsifying transforms. The method also takes several times longer than the reconstruction time for a single λ. This can be reduced by optimizing the starting λ, the factor for reducing λ after each solution, and the number of CG steps per solution.

References

1. Lustig M, Donoho D, Pauly JM, Magn Reson Med 2007;58:1182.
2. King K, Proc. ISMRM 2008; p. 1488.
3. Kilmer ME, O'Leary DP, SIAM J Matrix Anal Appl 2001;22:1204.
4. Qu P, Wang C, Shen GX, J Magn Reson Imaging 2006;24:248.

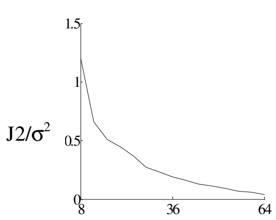

Figure 1. Effect of Nyquist lines

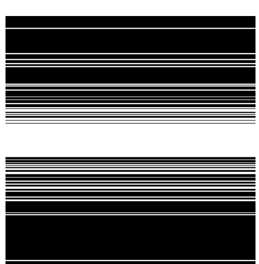

Figure 2. k-Space sampling

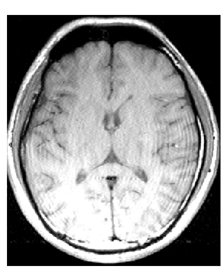

Figure 3. $\lambda = 0$

Figure 4. Empirical λ

Figure 5. Adaptive λ