**Proceedings of the 28th IEEE**
**EMBS Annual International Conference**
**New York City, USA, Aug 30-Sept 3, 2006**



# *Statistical Aspects of Parallel Imaging Reconstruction*

Ashish Raj, Bryan Kressler, Gurmeet Singh, Ramin Zabih, Yi Wang

*Abstract*— **A statistical interpretation of existing parallel magnetic resonance imaging methods reveals that th eunderlying noise model is of additive independent Gaussian noise. In reality MR imaging processes suffer from a variety of noise, errors and other uncertainties. A careful statistical analysis of these uncertainties can potentially allow significant improvement of the reconstruction process. In this paper we present such an analysis and describe a few very recent approaches to handle these statistical models. We show examples of simulation and in vivo reconstructed data which demonstrate the potential of the statistical approach.**

## I. INTRODUCTION

The standard image reconstruction algorithm for parallel imaging can be regarded as a least-squares fitting of the estimated image $\underline{x}$ to the measured data $\underline{y}$. These parallel imaging techniques include, *inter alia*, SENSE [1-4], SMASH [5-6] and GRAPPA [7-8]. Good comparative reviews were presented in [9]. While all of these methods are mathematically similar, SENSE is the exact reconstruction method [10] and will be the focus of this paper. These schemes use multiple coils to reconstruct (unfold) the unaliased image from under-sampled data in Fourier- or k-space. Successful unfolding relies on receiver diversity - each coil "sees" a slightly different image, since their spatial sensitivity profiles are different. Noise is a major source of uncertainties ubiquitous in all MR measurements. Effects of noise in MR signals are accounted for by a noise correlation matrix in the standard matrix inversion formulation of parallel imaging reconstruction. The noise correlation matrix is determined by the electronic coupling of coil elements and the imaging objects. One consequence of noise in parallel imaging is amplified noise in the final reconstructed images. Regularization incorporating prior information may be used to improve signal-to-noise ratio in image reconstruction [11-13].

Another source of uncertainties in parallel imaging reconstruction is error or noise in the estimation of coil sensitivity map. This type of error fundamentally alters the mathematics of the image reconstruction. Here we present general statistical formulations of image reconstruction that account for various uncertainties.

## II. MAXIMUM LIKELIHOOD (ML) FORMULATION FOR PARALLEL IMAGING

For simplicity, we assume noise is independent identically distributed (iid) Gaussian noise (after diagonalizing and normalizing the noise correlation matrix). To illustrate the statistical formulation, let us consider the case of error in sensitivity map. The standard parallel imaging reconstruction problem assumes a perfect encoding matrix E (product of sensitivity map and Fourier encoding factor) and maximizes the fitting of image x to the observed data y:

$$y = Ex + n,$$

If the noise process n is independent Gaussian, the probability of observing data y given a certain underlying image x is there fore given by the likelihood function

$$L(x|y) \sim \exp(-\tfrac{1}{2} \|y-Ex\|^2).$$

Under the Maximum Likelihood (ML) formulation, the best estimate of the image x is the one that maximizes the above conditional probability. Under the above iid Gaussian model, this amounts to a least squares minimization

$$\underline{x} = \arg\min_x \|y-Ex\|^2, \qquad (1a)$$

$$\underline{x} = (E^H E)^{-1} E^H y. \qquad (1b)$$

Eq.1a simply states an estimated image that agrees with observed data with the least-squares error, and Eq.1b is a closed form solution. This is the classic SENSE approach.

Unfortunately a typical parallel imaging process suffers from many other sources of noise and errors that are not accurately captured by the above additive model. Indeed, a major source of errors in SENSE occurs when it is difficult to obtain artifact-free sensitivity maps. In practice coil sensitivities are subject to noise processes since they are obtained from MR data, whether via division by a body coil image, or from low-frequency calibration lines. In addition, the encoding and decoding sensitivities are not identical in practice due to physiological motion, misalignment of coils between scans, etc.

We have addressed this issue by extending the classic least squares problem by including an error term $\Delta E$ in the sensitivity map,

$$y = (E+\Delta E)x + n. \qquad (2)$$

Here y represents measurements, n noise. The likelihood L(x|y) of image x fitting for a given observed data y is determined by the total effective noise, which may be

assumed to be characterized by iid Gaussian for simplicity. Let the total noise be g(x) = y-Ex,

$$R \equiv \langle g(x)g(x)^H \rangle = \sigma_n^2 I + \sigma_s^2 B. \quad (3)$$

Here the first term accounts for the standard signal noise, and the second term accounts for the sensitivity map errors. Then a solution to problem stated in Eq.2 is :

$$L(x|y) \sim \exp(-\tfrac{1}{2}(y-Ex)^H R^{-1}(y-Ex));$$

$$\underline{x} \sim \arg\min_x (y-Ex)^H R^{-1}(y-Ex). \quad (4)$$

In general Eq.4 does not have a closed form solution. In the case of Cartesian sampling, the computation for solution to Eq.4 can be efficiently found using Newton iteration. Preliminary data from simulation and imaging experiments demonstrate that this maximum likelihood solution provide much less noisy artifacts than the standard SENSE reconstruction.

The computation may become intractable for the problem stated in Eq.2 for a general arbitrary sensitivity map error $\Delta E$. The Gaussian noise assumption comes from the process of estimating sensitivity from pre-scanned MR signal that is contaminated with Gaussian noise. The lack of spatial resolution in the estimated sensitivity map and the spatial misregistration due to motion occurred between prescan for sensitivity map and scan for image data may not fit well into the Gaussian model. Interestingly even in these cases, the solution in Eq.4 can provide image quality improvement.

## III. Tikhonov regularization, Edge preserving Regularization, and graph cut solution

### A. Tikhonov regularization

The inverse problem of Eq.1 becomes progressively ill-posed with increasing acceleration, leading to unacceptable noise amplification. The Tikhonov-type regularization can be used to stabilize inversion and reduce noise in the reconstructed image [11,13]. The corresponding modification to the likelihood of image x fitting observed data y is modified as,

$$L(x|y) \sim \exp(-\tfrac{1}{2}(\|y-Ex\|^2 + \lambda\|A(x-x_r)\|^2));$$

$$x_{reg} = \arg\min_x \{\|y-Ex\|^2 + \lambda\|A(x-x_r)\|^2\}. \quad (5a)$$

Compared to Eq.1, the additional second penalizes non-smooth solutions through an appropriate matrix **A** and a prior reference image $\mathbf{x}_r$. Its closed-form solution is

$$\underline{x}_{reg} = x_r + (E^H E + \lambda^2 A^H A)^{-1} E^H (y-Ex_r). \quad (5b)$$

The standard choice is A = 1, the identical matrix. If there is no reference image ($x_r$ = 0), then Eq.5 computes the minimum norm solution. In general, the Tikhonov regularization imposes global smoothness in the reconstructed image, accordingly minimizing noise effects. The degree of noise reduction by regularization is controlled by the parameter $\lambda$. A small $\lambda$ corresponds to a high data fidelity and a small noise reduction. A high $\lambda$ corresponds to a low data fidelity (possible residual aliasing artifacts) and a substantial noise reduction. This fundamental aliasing/noise limit may be overcome by generalizing the regularization term and incorporating prior knowledge of targeted images.

### B. General image prior, Bayesian reconstruction, and graph cut solution

The requirement of reference image in Eq.5 may be too restrictive. We have extended the regularization term in Eq.5 to a more general form characterized by a probability function Pr(x) that is determined by image property parameters (cost function G(x)),

$$Pr(x) \sim \exp(-G(x)). \quad (6)$$

This probability is incorporated into the likelihood function for estimating image x according to the Bayesian approach:

$$L_{prior}(x|y) = L(x|y)\, Pr(x). \quad (7)$$

In the absence of a prior (G=0), Eq.6 reduces to the maximum likelihood formulation L(x|y) as in Eqs. 1&4. The Tikhonov regularization in Eq.5 be interpreted at $G(x) = \lambda\|A(x-x_r)\|^2$. In general, the term G(x) can succinctly express both the traditional smoothness assumptions as well as more complicated but powerful priors modeled by Markov Random Fields (MRFs) [14].

The solution to Eq.7 is the known method of maximum a posteriori (MAP) estimation, which is the maximum likelihood (ML) estimation constrained to a prior distribution over the image to estimate or a regularization of the ML estimation. The solution is formally expressed as

$$x_{MAP} = \arg\min_x \{\|y-Ex\|^2 + G(x)\}. \quad (8)$$

Conventional image priors impose spatial smoothness assuming that intensities vary smoothly across the entire image. Such an assumption may cause excessive edge blurring and become inappropriate for most medical image data with discontinuities at organ or pathology boundaries. An edge preserving prior G(x) would be very desirable.

A function imposing piecewise smoothness rather than global smoothness would be a natural choice for the cost function G(x):

$$G(x) = \sum_{p,q} V(x_p, x_q), \quad (9)$$

where the summation is over a spatial neighborhood, usually the 8 connected neighboring points. The separation cost $V(x_p,x_q)$ characterizes piecewise smoothness by weighting the intensities in the specified neighborhood. A useful choice for V is the truncated linear function,

$$V(x_p,x_q) = \lambda \min(|x_p-x_q|, K), \quad (10)$$

which treats neighboring intensity differences within the threshold K as noise and penalizes the difference. The threshold K allows large differences at organ edges to go unpenalized. Note that in Eq.10, the linear intensity distance

is used to greater intensity variation than that of the widely used $L_2$ norm.

In general, the computation for Eq.8 requires exponential number of steps and becomes unaffordable in practice. Here we outline a solution for the case of 1D Cartesian parallel imaging, Eq.8 reduces to minimizing the following objective energy function:

$$a^2 + \sum b(p)x_p^2 - 2\sum c(p)x_p + 2\sum_{Na} d(p,q)x_p\, x_q + \lambda \sum_{Ns} \min(|x_p - x_q|, K), \quad (11)$$

which can be solved using the fast graph cuts method [15-16]. The pixel intensity $x_p$ is represented by a finite discrete set of labels L. The minimization is iterated over individual labels $\alpha$ belonging to L ($\alpha$ -expansion move algorithm [15]), which reduces the integer problem into a succession of binary problems. The binary problems can be solved using the minimum cut over a graph representation of the image, as described in [15-16]. There exist a number of fast and easy methods to compute this minimum cut, via the maximum flow algorithm. Label substitution may be used to condition the energy function for graph cuts.

## IV. Results

### A. Maximum Likelihood estimation under sensitivity errors

Figure 1 shows the reconstruction results obtained by different methods on a torso data set under 3x acceleration and 8 coils. This data set suffered from motion mis-registration, ringing in sensitivity maps and non-uniform sensitivity noise due to the division by the sum-of-squares image. SENSE and ML-SENSE results are compared with the Total Least Squares (TLS) method which was proposed earlier as a possible way to handle sensitivity errors [17]. Clearly both SENSE and TLS are inadequate under sensitivity errors, but ML-SENSE is more tolerant to these errors.

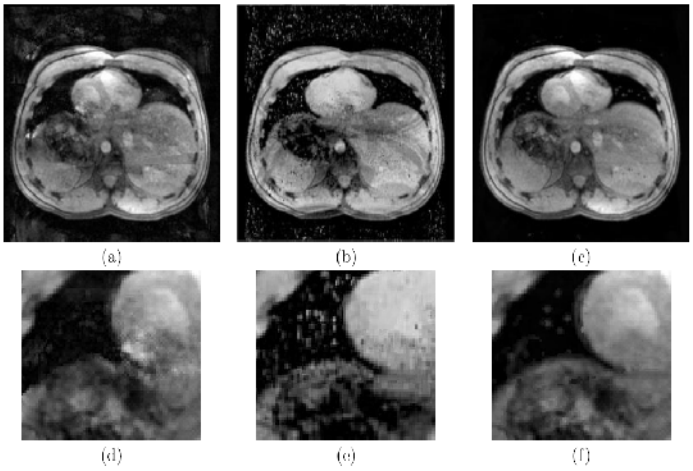


Figure 1: Reconstruction results of torso scan, 8 coils, 3 times acceleration: (a) Standard SENSE with regularization, (b) TLS, (c) ML-SENSE, (d) Zooming into (a), (e) Zooming into (b), and (f) zooming into (c). Note the distortions at the heart-stomach boundary and strip artifact across the liver in SENSE output.

### B. Regularization and Graph Cut Results

Figure 2 shows reconstruction results obtained from a brain scan using 8 coils and accelerated 4 times. High field strength (4 Tesla) structural MRI brain data was obtained using a whole body scanner (Bruker/Siemens Germany) equipped with a standard birdcage, 8 channel phased-array transmit/receive head-coil localized cylindrically around the S-I axis. Volumetric T1-weighted images (1 x 1 x 1 $mm^3$ resolution) were acquired using a Magnetization-Prepared RApid Gradient Echo (MPRAGE) sequence with TI/TR = 950/2300 ms timing and a flip angle of 8°. From the figure, the Graph Cut method appears to produce visually observable improvement compared to SENSE regularized with two choices of the regularization parameter.

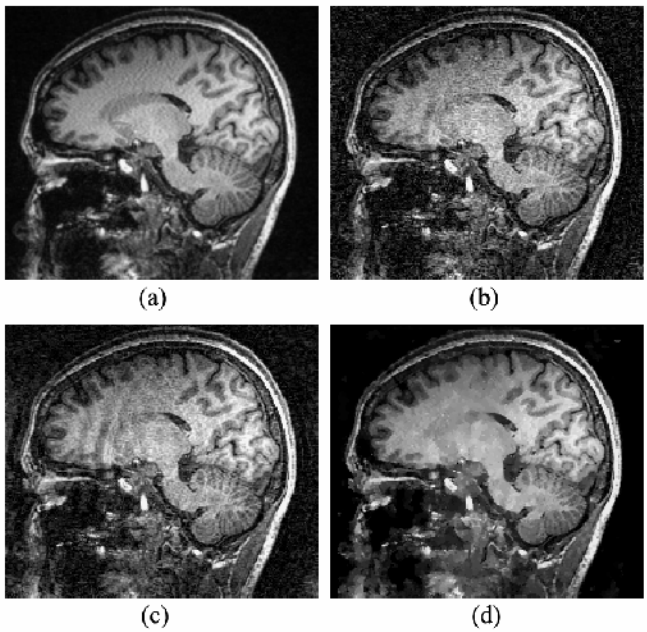


Figure 2: In vivo brain result with 4 times acceleration using 8 coils. Views were acquired vertically. (a) Reference image (b) SENSE regularized with $\lambda = 0.08$ (c) SENSE regularized with $\lambda = 0.16$, (d) Graph Cut reconstruction

Table 1 shows the quantitative comparison for a number of data sets we evaluated. SNR and g-values were obtained by the technique described in [18]. There appears to be a substantial improvement in both these quantities using Graph Cuts compared to conventional SENSE.

Table 2: Mean reconstructed SNR and mean g-factor values of SENSE and Graph Cut.

| | | Reg. SENSE | Graph Cut |
|---|---|---|---|

| | R | mean SNR | Mean g | Mean SNR | Mean g |
|---|---|---|---|---|---|
| Brain A | 4 | 8 | 4.6 | 23 | 1.7 |
| Brain B | 5 | 10 | 3.5 | 17 | 2.2 |
| Cardiac A | 3 | 20 | 2.3 | 33 | 1.5 |
| Cardiac B | 3 | 15 | 3.3 | 36 | 1.4 |

## V. Discussion and Conclusion

The experimental data shown above indicates the promise of improved performance and visual quality by a statistical approach to parallel imaging. This is a direct result of better error models as well as better prior models. The use of a powerful edge-preserving prior model in the Graph Cut method has produced visible improvement. The technique makes use of some advanced combinatorial optimization algorithms borrowed from graph theory. These combinatorial methods are needed because the Bayesian estimation task associated with our edge-preserving priors is generally too hard for traditional optimization methods like conjugate gradients, Newton methods, etc.

The preliminary data available from both the ML-SENSE and Graph Cut approaches is promising, but much more in-depth empirical evaluation must be performed. An exhaustive clinical evaluation is needed to further characterize the method and identify specific clinical applications – this work is ongoing.

## References


[1] Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: Sensitivity Encoding For Fast MRI. Magnetic Resonance in Medicine 1999; 42(5): 952-962.

[2] Pruessmann KP, Weiger M, Boernert P, Boesiger P. Advances In Sensitivity Encoding With Arbitrary K-Space Trajectories. Magnetic Resonance in Medicine 2001; 46(4):638--651.

[3] Weiger M, Pruessmann KP, Boesiger P. 2D SENSE For Faster 3D MRI. Magnetic Resonance Materials in Biology, Physics and Medicine 2002; 14(1):10-19.

[4] Van den Brink J, Watanabe Y, Kuhl CK, Chung T, Muthupillai R, Cauteren MV, Yamada K, Dymarkowski S, Bogaert J, Maki JH, Matos C, Casselman JW, Hoogeveen RM. Implications Of SENSE MR In Routine Clinical Practice. European Journal of Radiology 2003; 46(1):3-27.

[5] Sodickson DK, Manning WJ. Simultaneous Acquisition Of Spatial Harmonics (SMASH): Fast Imaging With Radiofrequency Coil Arrays. Magnetic Resonance in Medicine 1997; 38(4): 591-603.

[6] Sodickson DK, McKenzie CA, Ohliger MA, Yeh EN, Price MD. Recent Advances In Image Reconstruction, Coil Sensitivity Calibration, And Coil Array Design For SMASH And Generalized Parallel MRI. Magnetic Resonance Materials in Biology, Physics and Medicine 2002; 13(3): 158-63.

[7] Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA). Magnetic Resonance in Medicine 2002; 47(6): 1202-10.

[8] Wintersperger BJ, Nikolaou K, Dietrich O, Rieber J, Nittka M, Reiser MF, Schoenberg SO. Single Breath-Hold Real-Time Cine MR Imaging: Improved Temporal Resolution Using Generalized Autocalibrating Partially Parallel Acquisition (GRAPPA) Algorithm. European Journal of Radiology 2003; 13(8): 1931-6.

[9] Blaimer M, Breuer F, Mueller M, Heidemann RM, Griswold MA, Jakob PM. SMASH, SENSE, PILS, GRAPPA: How To Choose The Optimal Method. Top. Magnetic Resonance Imaging 2004; 15(4): 223-36.

[10] Wang Y. Description Of Parallel Imaging In MRI Using Multiple Coils. Magnetic Resonance in Medicine 2000; 44(3): 495-9.

[11] Lin F, Kwang K, BelliveauJ, Wald L. Parallel Imaging Reconstruction Using Automatic Regularization. Magnetic Resonance in Medicine 2004; 51(3): 559-67.

[12] Bammer R, Auer M, Keeling SL, Augustin M, Stables LA, Prokesch RW, Stollberger R, Moseley ME, Fazekas F. Diffusion Tensor Imaging Using Single-Shot SENSE-EPI. Magnetic Resonance in Medicine 2002; 48(1): 128-136.

[13] Ying L, Xu D, Liang ZP. On Tikhonov Regularization For Image Reconstruction In Parallel MRI. Proc IEEE EMBS 2004: 1056-9.

[14] Li S. *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, 1995.

[15] Boykov Y, Veksler O, Zabih R. Fast Approximate Energy Minimization Via Graph Cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 2001; 23(11): 1222-39.

[16] Kolmogorov V, Zabih R. What Energy Functions Can Be Minimized Via Graph Cuts? IEEE Transactions on Pattern Analysis and Machine Intelligence 2004; 26(2): 147-59.

[17] Liang ZP, Bammer R, Ji J, Pelc N, Glover G. Making Better SENSE: Wavelet Denoising, Tikhonov Regularization, And Total Least Squares. Proc. ISMRM 2002: 2388.

[18] Raj A, Singh G, Zabih R, Kressler B, Wang Y, Schuff N, Weiner M. Bayesian Parallel Imaging with Edge-Preserving Priors, Magnetic Resonance in Medicine 2006, in print.