WAVELET-REGULARIZED RECONSTRUCTION FOR RAPID MRI

M. Guerquin-Kern[1], D. Van De Ville[1], C. Vonesch[1], J.-C. Baritaux[1], K. P. Pruessmann[2] and M. Unser[1]

[1]Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne
[2] Institute for Biomedical Engineering, University and ETH Zurich, CH-8092 Zürich

ABSTRACT

We propose a reconstruction scheme adapted to MRI that takes advantage of a sparsity constraint in the wavelet domain. We show that artifacts are significantly reduced compared to conventional reconstruction methods. Our approach is also competitive with Total Variation regularization both in terms of MSE and computation time. We show that ℓ^1 regularization allows partial recovery of the missing k-space regions. We also present a multi-level version that significantly reduces the computational cost.

Index Terms— MRI, non-linear reconstruction, sparsity, wavelets, compressed sensing, thresholded Landweber, multi-level strategy

1. INTRODUCTION

Magnetic Resonance Imaging (MRI) scanners extract measurements that correspond to the Fourier transform of the object under investigation. When these so-called k-space samples are located on a Cartesian grid, one can rely on the inverse Fourier Transform for fast image reconstruction. When the sampling is non-Cartesian (e.g., spiral or radial sampling), a common approach is to regrid the data. However, interpolation in k-space leads to reconstruction artifacts that are difficult to characterize. Moreover, rapid MRI comes with undersampled k-space trajectories and noisy measures. Therefore, state-of-the-art methods tackle the reconstruction as an inverse problem with a proper regularization to make it well-conditioned. The most commonly-used algorithms in the field are linear and are linked to quadratic regularization. Unfortunately, in the case of severe undersampling, they suffer from noise propagation, blurring, ringing or aliasing errors. These reconstruction artifacts are similar to those encountered in image deconvolution where the research focus during the past decade has shifted towards non-linear optimization; in particular, ℓ^1 wavelet regularization which has been found to outperform traditional linear algorithms (e.g. Wiener filtering) [1].

this work was supported by the Swiss National Competence Center in Biomedical Imaging (NCCBI).

The theory of Compressed Sensing (CS) establishes a strong connection between ℓ^1 regularization and sparsity and is recently attracting a lot of attention in MRI [2,3]. The great majority of CS reconstruction algorithms consider a finite-difference sparsifying transform which is formally equivalent to Total Variation (TV) regularization. Surprisingly, ℓ^1 wavelet regularization has received little attention in MRI so far, with the notable exception of the work of Liu [4] which uses a differentiable proxy of the regularization functional to facilitate the optimization. In principle, it is possible to solve the ℓ^1 wavelet-regularized MRI reconstruction problem exactly using an adaptation of Daubechies et al.'s Thresholded Landweber (TL) algorithm [1, 5]. A potential difficulty that needs to be overcome is that the procedure converges slowly when the forward model is poorly conditioned.

In this work, we develop a version of the TL algorithm that is specifically tailored to the structure of the MRI problem. Based on simulations, we demonstrate high-quality reconstructions compared to state-of-the-art algorithms. We also show that our non-linear algorithm is capable of recovering missing k-space regions which is not the case for quadratic regularization. Finally, we propose a modified version of the algorithm that deploys a multi-level strategy to speed-up the reconstruction.

2. MRI AS AN INVERSE PROBLEM

2.1. Forward model

2.1.1. Physics

In this work, we consider a single receiving coil with homogeneous sensitivity. The corresponding model for the complex time-varying MR signal is

$$\mathrm{m}(\mathrm{t}) \propto \int_{\mathbb{R}^2} \rho(\mathbf{r}) \mathrm{e}^{2\mathrm{j}\pi\langle \mathbf{k}(\mathrm{t}), \mathbf{r}\rangle} \, \mathrm{d}\mathbf{r}, \tag{1}$$

where ρ is the unknown proton density map to be recovered and $\mathbf{k}(\mathrm{t})$ denotes the so-called k-space trajectory. The latter is implemented by time-varying magnetic gradient fields.

2.1.2. Discretization

In order to perform a numerical reconstruction, we must provide a discretized version of the forward model (1). Time is sampled at N instants resulting in the k-space samples $\{\mathbf{k}_n\}$ and the concatenated measurement vector $\mathbf{m} = (m_1, \ldots, m_N)$.

The signal to be reconstructed is represented as a linear combination of basis functions that are shifted versions of a generator φ on a finite Cartesian 2-D grid C_s:

$$\rho(\mathbf{r}) = \sum_{\mathbf{p} \in C_s} c[\mathbf{p}] \varphi(\mathbf{r} - \mathbf{p}). \tag{2}$$

The signal is thereby parametrized by a set of M coefficients $\{c[\mathbf{p}]\}$, represented as a vector $\mathbf{c}$.

Then, (1) can be rewritten as follows:

$$m_n = \widehat{\varphi}(-2\pi \mathbf{k}_n) \sum_{\mathbf{p} \in C_s} c[\mathbf{p}] e^{2j\pi \langle \mathbf{k}_n, \mathbf{p} \rangle} + b_n. \tag{3}$$

The term b_n is introduced to represent both the measurement noise and model mismatch. This model is linear; thus there exists a $N \times M$ matrix $\mathbf{E}$ such that:

$$\mathbf{m} = \mathbf{E}\mathbf{c} + \mathbf{b}. \tag{4}$$

2.2. Variational formulation

2.2.1. General framework

The solution $\widetilde{\mathbf{c}}$ is defined as the minimizer of a cost function that involves two terms: the data fidelity $\mathcal{F}$ and the regularization $\mathcal{R}$ that favors solutions according to given prior knowledge. This is summarized as

$$\widetilde{\mathbf{c}} = \arg \min_{\mathbf{x}} \mathcal{F}(\mathbf{m} - \mathbf{E}\mathbf{x}) + \lambda \mathcal{R}(\mathbf{x}), \tag{5}$$

where the tuning parameter λ balances the effects of the two terms. $\mathcal{F}$ is chosen as the square of the ℓ^2-norm : $\|\mathbf{m} - \mathbf{E}\mathbf{x}\|_{\ell^2}^2$, which is justified when the noise is Gaussian. The ill-conditioning, inherent to undersampled trajectories, imposes the choice of an adequate regularization term $\mathcal{R}$.

Standard Tikhonov regularization, corresponding to a quadratic regularization term $\|\mathbf{R}\mathbf{x}\|_{\ell^2}^2$, leads to the formal linear solution $(\mathbf{E}^{\mathrm{H}}\mathbf{E} + \lambda \mathbf{R}^{\mathrm{H}}\mathbf{R})^{-1}\mathbf{E}^{\mathrm{H}}\mathbf{m}$ (the superscript $^{\mathrm{H}}$ denotes the Hermitian adjoint). Such linear solutions are quite tractable both theoretically and numerically. However, quadratic cost functionals are of limited use for measuring the plausibility of natural images.

TV reconstruction is related to the ℓ^1-norm of the modulus of the gradient and is an optimal regularization for piecewise-constant solutions.

2.2.2. Wavelet regularization

The underlying idea of wavelet regularization is that natural images tend to be sparse in the wavelet domain. We hence would like to favor, among all the possible candidates, a solution that has only few significant (non-zero) wavelet coefficients.

Based on the property that a small ℓ^1-norm promotes sparsity, this solution is defined as:

$$\boldsymbol{W}\widetilde{\mathbf{c}} = \arg \min_{\boldsymbol{w}} \left\| \mathbf{m} - \mathbf{E}\boldsymbol{W}^{-1}\boldsymbol{w} \right\|_{\ell^2}^2 + \lambda \|\boldsymbol{w}\|_{\ell^1}, \tag{6}$$

where $\boldsymbol{W}$ and $\boldsymbol{W}^{-1}$ are the wavelet decomposition and synthesis matrices, respectively.

3. ITERATIVE MINIMIZATION

3.1. Principle of the TL algorithm

The important point for comprehension is that the solution of the simpler wavelet denoising problem ($\mathbf{E}$ is the identity matrix and $\boldsymbol{W}$ is orthonormal) is a single-step thresholding: $\widetilde{\mathbf{c}} = \boldsymbol{W}^{-1}\mathcal{T}_\lambda(\boldsymbol{W}\mathbf{m})$, as was first observed by Chambolle [6]. Daubechies et al.'s algorithm [1, 5] can then be explained by iteratively bounding the initial reconstruction problem by a simpler denoising problem. Specifically, at iteration step n, one defines the auxiliary variable $z_n = \widetilde{\mathbf{c}}_n + \alpha^{-1}(\mathbf{a} - \mathbf{E}^{\mathrm{H}}\mathbf{E}\widetilde{\mathbf{c}}_n)$ with $\mathbf{a} = \mathbf{E}^{\mathrm{H}}\mathbf{m}$ where the wavelet vector $\widetilde{\mathbf{c}}_n$ specifies the current estimate of the solution. One then considers the upper bound on data fidelity

$$\|\mathbf{m} - \mathbf{E}\mathbf{x}\|_{\ell^2}^2 \leqslant \mathrm{const}(\mathbf{x}) + \alpha \|z_n - \mathbf{x}\|_{\ell^2}^2$$

which is valid provided that α is greater than the largest eigenvalue of $\mathbf{E}^{\mathrm{H}}\mathbf{E}$. The equality holds for $\mathbf{x} = \widetilde{\mathbf{c}}_n$.

The minimization is then performed by:

Initialization of $\widetilde{\mathbf{c}}_0$;
repeat
  $z_n \leftarrow \widetilde{\mathbf{c}}_n + \tau(\mathbf{a} - \mathbf{E}^{\mathrm{H}}\mathbf{E}\widetilde{\mathbf{c}}_n)$ (Landweber updating);
  $\widetilde{\mathbf{c}}_{n+1} \leftarrow \boldsymbol{W}^{-1}\mathcal{T}_{\lambda\tau}(\boldsymbol{W}z_n)$ (wavelet thresholding);
  $n \leftarrow n + 1$;
until *convergence* ;

3.2. Structure of the problem

Considering (3), we rewrite the measurements as:

$$m_n = \widehat{\varphi}(-2\pi \mathbf{k}_n)\mathbf{h}_n{}^{\mathrm{H}}\mathbf{c} + b_n, \tag{7}$$

where $\mathbf{h}_n$ denotes $[e^{-2j\pi\langle \mathbf{k}_n, \mathbf{p}_1 \rangle} \ldots e^{-2j\pi\langle \mathbf{k}_n, \mathbf{p}_M \rangle}]^{\mathrm{T}}$ and $\{\mathbf{p}_i\}_{i \in [1..M]}$ is the index scanning the grid C_s.

The encoding matrix is then decomposed as:

$$\mathbf{E} = \mathbf{D}_{\widehat{\varphi}}\mathbf{H}, \tag{8}$$

with $\mathbf{H} = [\mathbf{h}_1 \ldots \mathbf{h}_{N_s}]^{\mathrm{T}}$ and $\mathbf{D}_{\widehat{\varphi}} = \mathrm{diag}\,(\widehat{\varphi}(-2\pi \mathbf{k}_n))$.

A fundamental point for our algorithm is that the multiplication with the matrix $\mathsf{E}^{\mathrm{H}}\mathsf{E}$ corresponds to a 2-D convolution (Block-Toeplitz matrix). Indeed, by defining the kernel $\mathsf{G}[\mathbf{p}] = \sum_{n=1}^{N} e^{-2j\pi\langle \mathbf{k}_n\,,\,\mathbf{p}\rangle} |\widehat{\varphi}(-2\pi\mathbf{k}_n)|^2$ and using (8), we get:

$$y[\mathbf{p}] = (\mathsf{E}^{\mathrm{H}}\mathsf{E}x)[\mathbf{p}] = \sum_{\mathbf{q}\in C_s} \mathsf{G}[\mathbf{p}-\mathbf{q}]x[\mathbf{q}]. \qquad (9)$$

Note that the kernel G has a support twice as large as C_s in each dimension.

The matrix-vector multiplication with $\mathsf{E}^{\mathrm{H}}\mathsf{E}$ corresponds to the most computer-intensive part of Algorithm 3.1. Based on the above property, we implement this operation by a pointwise multiplication in the frequency domain using FFTs on a grid twice as large as C_s [7]. The advantage is that this computation is exact while it avoids the use of regridding.

3.3. Bound estimation

In Section 3.1, we mentioned that the parameter $\alpha = \tau^{-1}$ must be chosen greater than the largest eigenvalue of $\mathsf{E}^{\mathrm{H}}\mathsf{E}$. We use a standard power method to estimate this value:

Initialization of ν_0 as a random vector;
repeat
  $\alpha_n \leftarrow \|\mathsf{E}\nu_n\|^2 / \|\nu_n\|^2$ (eigenvalue);
  $\nu_{n+1} \leftarrow \mathsf{E}^{\mathrm{H}}\mathsf{E}\nu_n / \left\|\mathsf{E}^{\mathrm{H}}\mathsf{E}\nu_n\right\|$ (eigenvector);
  $n \leftarrow n+1$;
until *convergence* ;

Here too, we avoid the use of regridding by rewriting $\|\mathsf{E}\nu\|^2 = \langle \nu\,,\,\mathsf{E}^{\mathrm{H}}\mathsf{E}\nu\rangle$.

3.4. Performance

3.4.1. Description of the experiments

We simulated MRI experiments with two original images: a 1024×1024 version of the Shepp-Logan brain phantom and a 512×512 T2 clinical image. We considered variable-density spiral trajectories as described in [8]; the parameters are: $\alpha = 2$, 8 interleaves, undersampling factor $R = 4$, supporting 256×256 and respectively 192×192 matrices for reconstruction. We added white Gaussian complex noise with Signal-to-Error-Ratio 20 dB and 30 dB to the simulated data. All algorithms were coded with MATLAB 7.7 (The MathWorks Inc. 2008, Natick) on a 8-cores machine with 4 GB RAM, running MacOs 10.5. To allow for an objective comparison of methods, we systematically selected the regularization parameter that provided the best SER performance (Oracle solution).

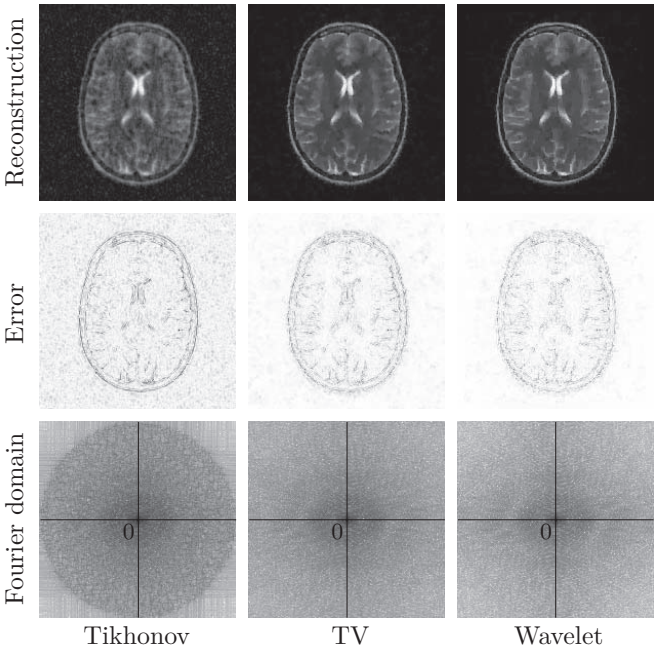

Fig. 1. Comparison of reconstruction results for the T2 image. Errors: same intensity scale and dark points representing high intensities. Fourier: log of modulus.

3.4.2. Results

For the linear reconstruction, we used a Conjugate Gradient (CG) loop, with a tolerance fixed to $1e-8$. The TV reconstruction was implemented using the Iterative Re-weighted Least Square (IRLS) method [9] with 10 outer iterations. For the linear solver, we also applied CG with a tolerance $1e-8$. For the wavelet reconstruction, we chose the Haar basis, which is the simplest and fastest wavelet transform. We considered 3 decomposition levels and used cycle-spinning as in [1] to avoid blocking artifacts. We took the zero image as an initialization. The step size τ was chosen the largest possible to speed up reconstruction. We stopped the algorithm when convergence was reached by monitoring the Signal-to-Error-Ratio (SER) and the cost function. We show the reconstructed images in Fig. 1. Corresponding SERs and computation times are given in Table 1.

The simulations show that wavelet regularization is competitive with TV for the Shepp-Logan brain phantom which is piecewise-constant. It performs even better (observed SER improvement: 0.7 dB) for the T2-MR image. Wavelet regularization appears to restore more details. In Fig. 1, we can note that high frequency components—missing in the spiral sampling—are partly recovered by both non-linear approaches. This extrapolation is made possible by the ℓ^1 regularization constraints that induce some correlations in the Fourier domain.

Image (noise)	Shepp-Logan (20 dB)			T2 image (30 dB)		
Regularization	ℓ^2	**TV**	**Wavelets**	ℓ^2	**TV**	**Wavelets**
λ_{opt}	2.7	0.3	0.25	0.3	50	60
SER_{opt} (dB)	8.02	12.08	**12.19**	9.76	13.13	**13.88**
duration (s)	**7**	146	117	**4**	46	268

Table 1. Comparison of reconstruction quality in terms of SER for different types of regularization.

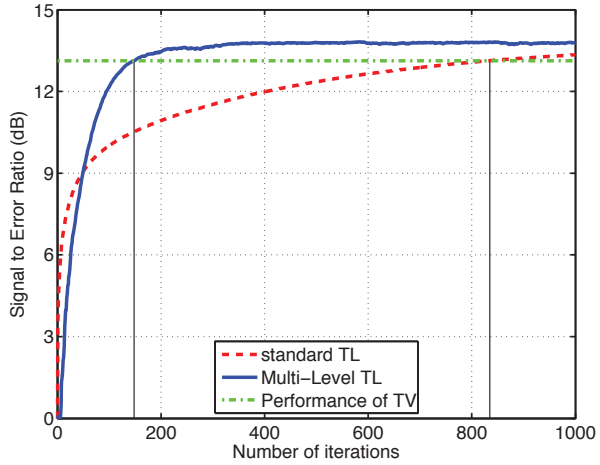

Fig. 2. Speed acceleration of the Multi-Level TL algorithm.

4. MULTI-LEVEL IMPLEMENTATION

To speed up the computation, we propose to formulate the problem at different resolutions and to use a "coarse-to-fine"[1] iterative optimization strategy. Our scheme is inspired by the work of Vonesch et al. in the context of deconvolution [10]. The key idea is to take advantage the multi-level representation to iterate at different scales while adapting the majorization parameters (α) to the current subband.

To implement the method efficiently we use the fact that the updating matrix $\mathbf{E}^{\mathrm{H}}\mathbf{E}$, at every resolution scale, is equivalent to a convolution on a grid twice as large. The delicate point of the algorithm is the suitable handling and propagation of boundary conditions.

We compared the number of iterations required to reach the TV quality for the standard and the multi-level implementation. Results are shown Fig. 2. Because of the chosen multi-level strategy, the computational costs per iteration are comparable. We observe a speed up by a factor of 5.5.

5. CONCLUSION

We successfully adapted and implemented wavelet-regularized reconstruction for MRI. Our experimental results suggest that it is competitive with TV regularization both in terms of MSE and computation time. It also appears that the prior corresponding to ℓ^1 regularization may be better adapted to MR images than the classical ℓ^2-term. We have observed that wavelet reconstruction usually outperforms TV for images that contain textured areas and/or many small details.

The current limitation of basic iterative wavelet reconstruction methods is their slow convergence for poorly-conditioned problems. We have shown how to improve the situation by using subband-adaptive step sizes together with a coarse-to-fine updating strategy. An important issue that still needs to be addressed is the optimal adjustment of the regularization parameter for real-data experiments.

[1] At every iteration each subband is updated once, from the coarsest to the finest scale.

6. REFERENCES

[1] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Sig. Proc.*, vol. 12, no. 8, pp. 906–916, August 2003.

[2] M. Lustig, D.L.. Donoho, and J.M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, pp. 1182–1195, 2007.

[3] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed Sensing in dynamic MRI," *Magn. Reson. Med.*, vol. 59, no. 2, pp. 365–373, Feb 2008.

[4] B. Liu, E. Abdelsalam, J. Sheng, and L. Ying, "Improved spiral sense reconstruction using a multiscale wavelet model," in *Proc. IEEE ISBI 2008*, 2008, pp. 1505–1508.

[5] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, August 2004.

[6] A. Chambolle, R. A. DeVore, N.-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Trans. Sig. Proc.*, vol. 7, no. 3, pp. 319–335, March 1998.

[7] F. Wajer and K. P. Pruessmann, "Major speedup of reconstruction for sensitivity encoding with arbitrary trajectories," *Proc., ISMRM, 8th Annual Meeting, Glasgow*, 2001.

[8] D.-H. Kim, E. Adalsteinsson, and D. M. Spielman, "Simple analytic variable density spiral design," *Magn Reson. in Med.*, vol. 50, no. 1, pp. 214–219, 2003.

[9] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for signal processing*, Prentice Hall, 2000.

[10] C. Vonesch and M. Unser, "A fast multilevel algorithm for wavelet-regularized image restoration," *IEEE Trans. Imag. Proc.*, vol. 18, no. 3, pp. 509–523, March 2009.