AN EFFICIENT AUXILIARY VARIABLE METHOD FOR QUANTIFICATION OF SPIN DENSITY, R_2^* DECAY AND FIELD INHOMOGENEITY MAPS IN MAGNETIC RESONANCE IMAGING

Chenxi Hu, Student Member, IEEE, Stanley J. Reeves, Senior Member, IEEE

Magnetic Resonance Imaging Research Center, Auburn University
560 Devall Drive, Auburn, AL 36830
czh0032@auburn.edu, reevesj@auburn.edu

ABSTRACT

Quantification of spin density, R_2^* decay and off-resonance frequency maps is very important in some applications of magnetic resonance imaging (MRI). To reconstruct these parameter maps, a time-varying model such as mono-exponentials must be used to represent the signal from each voxel. When only a single-shot trajectory is adopted, the underlying reconstruction problem is significantly nonlinear and therefore requires an iterative algorithm. The regularized trust region method previously proposed to address this problem is stable but lacks speed. In this paper, we propose a novel auxiliary variable method that is very efficient in solving the underlying optimization problem. This method introduces an auxiliary variable in the spatial-temporal domain that separates the data fidelity term and the structure fidelity term. The algorithm then alternately optimizes the data fidelity and the structure fidelity to reach the solution. The data fidelity optimization has a closed-form solution and can be solved very efficiently. The structure fidelity optimization fits the exponential model with the auxiliary variable and can also be rapidly computed. Some preliminary comparisons between the auxiliary variable method and the trust region method show that the new method is 10 times faster than the trust region method at a reasonable reconstruction precision.

Index Terms— Auxiliary variable method, variable splitting, image reconstruction, MRI

1. INTRODUCTION

In many applications of MRI, the signal at each voxel does not remain constant through the imaging time frame. A typical time-domain model treats each voxel as a decaying exponential with unknown amplitude (spin density), R_2^* decay, and off-resonance frequency [1, 2]. Joint estimation of these parameter maps can dramatically reduce the artifacts caused by T_2^* relaxation, field inhomogeneities, and susceptibility differences. Furthermore, the quantification of these parameter maps itself is important in some MR applications such as BOLD functional MRI. However, through the years a robust and fast solution remains challenging due to the associated nonlinear, ill-conditioned, and large scale cost function.

The most commonly used method for joint reconstruction of the spin density, R_2^*, and off-resonance frequency maps is the conjugate gradient (CG) method [3, 4]. While the method works well with an educated initialization of the variable sets, the method experiences large difficulties to converge rapidly when the initialization is poor. Recently, the trust region (TR) method has been applied to the joint reconstruction problem and outperforms CG in both stability and speed [5]. However, this method does not leverage the structure of the forward operator, and its speed gain mainly comes from its stability. In this work, we propose a novel auxiliary variable (AV) method to solve the joint reconstruction problem. The AV method introduces an auxiliary variable set to separate the original data fidelity term into two terms. Both the two terms exploit the special structure of the relative forward operator and therefore can be solved efficiently. The speed of this method is therefore much higher than the CG and TR method which need to evaluate the nonlinear function and gradient in each iteration. A number of researchers have exploited this strategy to implement nonquadratic regularization and other constraints more efficiently ([6, 7, 8, 9, 10]).

The auxiliary variable is introduced by a penalty term in the cost function that links the unconstrained form of the time-dependent image (the auxiliary variable) to the nonlinear function of time adopted for each voxel. Each iteration consists of two steps—minimization with respect to the time-varying image to be estimated and minimization with respect to the parameters of the time-domain models. When the augmented cost function is minimized with respect to the unconstrained image (the auxiliary variable), the constraint term is not directly enforced but only the proximity of the image to the time model through the penalty term. When the augmented cost function is minimized with respect to the time model parameters, this step does not require estimating the full time-varying image but only enforces proximity to it through the penalty term. As a result, each iteration is greatly simplified.

2. THEORY

2.1. The general objective function

The k-space data is modeled by the following discretized equation:

$$s_l = \sum_{n=0}^{N-1} m_n e^{z_n t_l} e^{-2\pi\iota[\mathbf{k}_l \cdot \mathbf{r}_n]} + \varepsilon_l,\ l = 1, 2, \cdots L \quad (1)$$

where s_l denotes the k-space data, $t_l, \vec{k}_l, \varepsilon_l,\ l = 1, 2, \cdots, L$ denote the L samples of time, the underlying trajectory, and the noise. $\vec{r} := \text{vec}(\mathbf{r}_n)$, $\vec{m} := \text{vec}(m_n)$, and $\vec{z} := \text{vec}(z_n)$ denote the spatial coordinates, the spin density map, and the complex frequency which takes the R_2^* decay and the off-resonance frequency maps as its real and imaginary part, respectively. When a single-shot trajectory is adopted, direct inverse Fourier transform is not workable because the k-space at each time point is very sparsely sampled. Furthermore, the forward operator in (1) is nonlinear and ill-conditioned and hence calls for a regularized iterative method. Because the noise in (1) is Gaussian, a natural way to formulate the cost function is the least-squares approach:

$$\|s_l - \sum_{n=0}^{N-1} m_n e^{z_n t_l} e^{-2\pi\iota[\mathbf{k}_l \cdot \mathbf{r}_n]}\|^2 \quad (2)$$

With regularization, (2) becomes

$$\|s_l - \sum_{n} m_n e^{z_n t_l} e^{-2\pi\iota[\mathbf{k}_l \cdot \mathbf{r}_n]}\|^2 + \lambda_1 \|R\vec{m}\|^2 + \lambda_2 \|R\vec{z}\|^2 \quad (3)$$

where R is the regularization matrix. In this work, we use the first-order difference operator for all regularization methods.

2.2. The auxiliary variable method

2.2.1. Formulating the cost function for the AV method

Let's divide the signal acquisition time frame into multiple segments. Each segment is sufficiently short in time so that all samplings within this segment can be modeled as happening at a single time point. Suppose there are I such time segments. We call the spatial image at the ith segment the ith frame, and the Fourier transform of the ith frame the ith k-frame. The ith frame $\vec{\rho}_i(\vec{m}, \vec{z})$ and its relation to the ith k-frame $\vec{S}_i$ can be represented by

$$\vec{S}_i = \mathcal{F}\{\vec{\rho}_i(\vec{m}, \vec{z})\} \quad (4)$$

where $\mathcal{F}$ represents the spatial Fourier transform. If all samples are acquired on Cartesian grids of each k-frame, then the ith k-frame samples are determined by partial rows of the Fourier matrix prescribed by the trajectory. The cost function (3) can be rewritten by

$$\|\vec{S} - DF\vec{\rho}(\vec{m}, \vec{z})\|^2 + \lambda_1 \|R\vec{m}\|^2 + \lambda_2 \|R\vec{z}\|^2 \quad (5)$$

where D represents the row-wise downsampling of the Fourier matrix F, and $\vec{S} := \text{vec}(\vec{S}_i)$. Due to the nonlinear relationship between $\vec{\rho}$ and $(\vec{m}, \vec{z})$, the estimation process must include the constraint implied by the nonlinear model, making the computational complexity large. A method to decouple the solution step from the nonlinear model constraint is to introduce an auxiliary variable u in the spatial-temporal domain to decouple $\vec{S}$ and $\vec{\rho}$. Then (5) becomes

$$\begin{aligned} &\|\vec{S} - DF\vec{u}\|^2 + \alpha\|u - \vec{\rho}(\vec{m}, \vec{z})\|^2 \\ &+\lambda_1 \|R\vec{m}\|^2 + \lambda_2 \|R\vec{z}\|^2 \end{aligned} \quad (6)$$

where the relation between the variable and the data in the first term is linear and highly structured (Fourier transform) and the operator in the second term is decoupled in the spatial domain. We then minimize function (6) in an alternating two-step fashion: in the first step we minimize w.r.t. $\vec{u}$ given an estimate for $\vec{\rho}(\vec{m}, \vec{z})$, and in the second step we minimize w.r.t. $\vec{m}$ and $\vec{z}$ given the estimate of $\vec{u}$ from the first step. Solving in this fashion can maximally leverage the spatial structure of the operator in the first and the second term.

Although the cost function in (6) is not ill-conditioned w.r.t. $\vec{m}$ and $\vec{z}$ due to the regularization, it still may be ill-conditioned w.r.t. $\vec{u}$ especially when α is small. Furthermore, for large α the solution may be relatively close to $\vec{\rho}(\vec{m}, \vec{z})$ but without maintaining the smoothness, oscillatory, and decay characteristics of that term. We therefore add a temporal-spatial regularization term to the cost function (6):

$$\begin{aligned} &\|\vec{S} - DF\vec{u}\|^2 + \beta\|RE\vec{u}\|^2 \\ &+\alpha\|\vec{u} - \vec{\rho}(\vec{m}, \vec{z})\|^2 + \lambda_1 \|R\vec{m}\|^2 + \lambda_2 \|R\vec{z}\|^2 \end{aligned} \quad (7)$$

where E represents a first-order difference operator in time. The composite operator RE thus represents a temporal-spatial regularization operator to make $\vec{u}$ smooth in both spatial and temporal domains.

2.2.2. Minimizing the cost function

The cost function associated with the AV method is minimized in a two-step iterative manner. In the first step of each iteration, one minimizes

$$\|\vec{S} - DF\vec{u}\|^2 + \beta\|RE\vec{u}\|^2 + \alpha\|\vec{u} - \vec{\rho}(\vec{m}, \vec{z})\|^2 \quad (8)$$

w.r.t. u given $\vec{\rho}$. Notice this is a quadratic function. Also, since R is a convolution in the spatial domain, R is diagonalizable through F. The solution to (8) is given below

$$\begin{aligned} \hat{\vec{u}} &= (F^H D^H DF + \beta(RE)^H RE + \alpha I)^{-1}(F^H D^H S + \alpha\vec{\rho}) \\ &= (F^H D^H DF + \beta F^H X^H XF + \alpha F^H F)^{-1}(F^H D^H S + \alpha\vec{\rho}) \\ &= F^H (D^H D + \beta X^H X + \alpha I)^{-1} F(F^H D^H S + \alpha\vec{\rho}) \\ &= F^H (D^H D + \beta X^H X + \alpha I)^{-1}(D^H S + \alpha F\vec{\rho}) \end{aligned} \quad (9)$$

where X is a bidiagonal block matrix that satisfies $RE = F^H X F$. $D^H D + \beta X^H X + \alpha I$ is a tridiagonal block matrix where each block is a diagonal matrix. (9) can be solved in $O(n)$ complexity by using the fast Fourier transform (FFT) and the tridiagonal matrix algorithm (the Thomas algorithm).

In the second step of each iteration, one minimizes

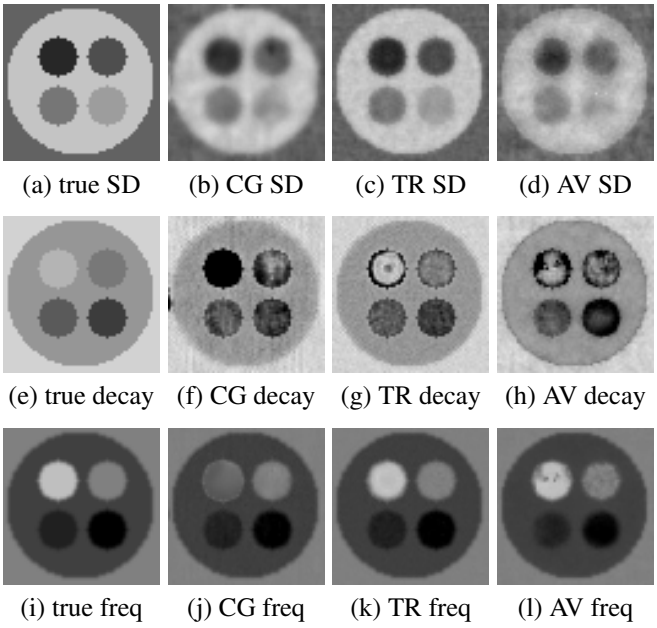
$$\|\vec{u} - \vec{m}e^{\vec{z}t}\|^2 + \frac{\lambda_1}{\alpha}\|R\vec{m}\|^2 + \frac{\lambda_2}{\alpha}\|R\vec{z}\|^2 \qquad (10)$$

where $\vec{u}$ is given by step 1. Notice this function is highly decoupled in the spatial domain and therefore is much easier to solve than the problem of minimizing (3). In this work, we minimize it in two steps, where we first find an educated initialization to the variables by curve fitting in the logarithmic domain and then run a preconditioned nonlinear conjugate gradient (PCG) algorithm to find the true minimum. In the first sub-step, we find $|\vec{m}|$ and $\mathrm{Re}\{\vec{z}\}$ by fitting the logarithm of $|\vec{u}|$ along the time axis and find $\angle\vec{m}$ and $\mathrm{Im}\{\vec{z}\}$ by fitting the unwrapped phase of $\vec{u}$. Notice the resulting matrices are all very easy to invert. PCG in the second sub-step uses a diagonal preconditioner.

The AV method recursively runs the two steps. Three parameters are associated with the method: α, β, and λ. Generally, we adopt a schedule such that α increases, β reduces, and λ reduces when the iteration grows so that the solution of (7) approaches that of (2). However, for practical purpose, the parameters β and λ do not reduce to zero to maintain the regularization effects.

3. RESULTS

We compared the three methods—CG, TR, and AV—based on a synthetic cylindrical phantom (Fig 1 left column), which contains four small cylindrical objects that have different spin density, R_2^* decay, and off-resonance frequency maps. The size of the image is 64×64. The ranges of the R_2^* decay and off-resonance frequency within the phantom are $[-50, -10]$ sec^{-1} and $[-200, 100]$ Hz. The simulations were run using MATLAB on an Intel i7-4700MQ quad-core processor. The AV program uses only a single core while the TR and CG use four cores with a parallel implementation. The signal-to-noise ratio (SNR) is defined by the norm of the signal divided by the standard deviation of added noise. In this simulation, white Gaussian noise is added with SNR = 10, which is a considerable amount of noise for testing algorithms. The precision of the reconstruction results is evaluated by means of normalized mean-square error (NMSE), which is defined by the norm of the reconstruction error divided by the norm of the ground truth. The trajectory used in the data synthesis is the rosette trajectory. Rosette is a single-shot non-Cartesian trajectory that has been remarked by many authors to be optimal in terms of the encoding capability over the complex frequency map [1, 4]. Since the AV method in our current development requires a Cartesian trajectory, the used rosette trajectory is rounded to nearby grid locations for preliminary algorithm justifications. 128 k-frames are used and this number needs to be carefully chosen to avoid aliasing in the reconstruction of the off-resonance frequency. For an 80 ms-long trajectory, each k-frame in the model covers less than 1 ms, causing only a small quantification error. We only consider the single-coil reconstruction in this paper.



(a) true SD (b) CG SD (c) TR SD (d) AV SD

(e) true decay (f) CG decay (g) TR decay (h) AV decay

(i) true freq (j) CG freq (k) TR freq (l) AV freq

Fig. 1. Comparison between CG, TR, and AV on the reconstruction of the spin density (SD), R_2^* decay, and the off-resonance frequency maps

Fig 1 shows the reconstruction results from CG, TR, and AV as compared with the ground truth. Fig 2 shows the convergence profiles associated with the three methods. Clearly, TR has more precise estimates than the other two methods. AV has a better precision than CG overall and is close to TR. In terms of the speed, AV is much faster than both TR and CG. The speed gain of AV over TR is more than 10 fold and over CG is more than 20 fold. This large speed gain makes AV a suitable method for initialization of variables. In a simulation which is not shown, TR converges in 2 minutes with initialization given by the AV method and generally converges around 18 minutes with poor initialization.

4. DISCUSSIONS AND CONCLUSIONS

The simulation shows that the AV method is much faster than both the CG and TR method. The speed gain has two sources: first, the good utilization of the structure allows for efficient matching of the data; and second, the constraint on $\|\vec{u} - \vec{\rho}\|$ allows a large region for the underlying $\vec{z}$ to vary in step 1, causing a rapid convergence in early phase of the iteration.

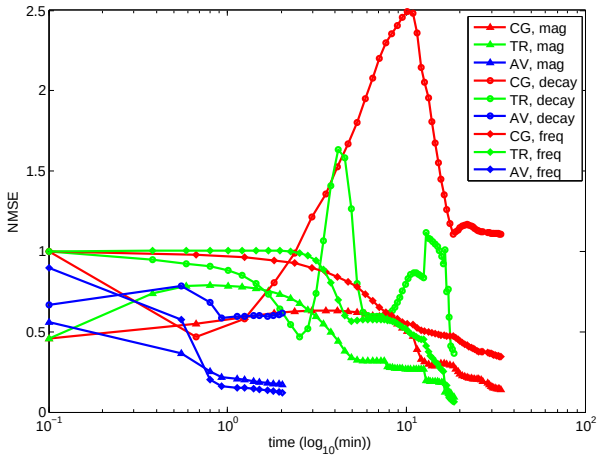


Fig. 2. Convergence profiles of CG, TR, and AV on the reconstruction of the spin density (mag), R_2^* decay, and the off-resonance frequency maps

The simulation shows that the AV method leads to a worse reconstruction than TR. This may be caused by the slow convergence of the AV method when the residual becomes very small. There are many reasons which may explain the slow convergence. For example, the AV method is very similar to the alternating projection method. Therefore, it may also have the convergence property of the alternating projection method, which can be arbitrarily slow. However, to thoroughly understand the convergence behavior of the AV method, further analysis is required. Although the precision of the AV method appears to be limited, the AV method is still important since it provides a very good initialization in a very short time. After that, other methods such as the TR method can be used to refine the result. This two-step procedure largely reduces the computation time compared to the situation when TR utilizes a poor initialization.

The method proposed in this paper works only for Cartesian sampling. For non-Cartesian sampling, directly gridding the k-space samples may cause a large rounding error. Therefore, modifications to step 1 are required for non-Cartesian trajectories, and this topic is still under investigation. In addition, the AV method has many parameters, and this characteristic is similar to the TR method. Trial-and-error is currently used to choose these parameters, but a more intelligent method is highly desired.

In conclusion, we have presented a novel auxiliary variable method which can rapidly solve the nonlinear large-scale optimization problem associated with joint reconstruction of the spin density, R_2^* decay, and off-resonance frequency maps. The auxiliary variable method exploits the structure between the images and the data, making the method remarkably efficient. The speed of the auxiliary variable method using a single core is roughly 10 times faster than the trust region method and 20 times faster than the conjugate gradient method, both of which run on four cores in parallel. The precision of the auxiliary variable method is reasonable under strong noise. This characteristic makes the method a desired initialization method for the joint reconstruction problem.

5. REFERENCES

[1] Donald B Twieg, "Parsing local signal evolution directly from a single-shot MRI signal: A new approach for fMRI," *Magnetic Resonance in Medicine*, vol. 50, no. 5, pp. 1043–1052, 2003.

[2] Jeffrey Fessler, "Model-based image reconstruction for MRI," *Signal Processing Magazine, IEEE*, vol. 27, no. 4, pp. 81–89, 2010.

[3] K.T. Block, M. Uecker, and J. Frahm, "Model-based iterative reconstruction for radial fast spin-echo MRI," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 11, pp. 1759–1769, Nov 2009.

[4] Donald B Twieg and Stanley J Reeves, "Basic properties of SS-PARSE parameter estimates," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 5, pp. 1156–1172, 2010.

[5] Chenxi Hu and Stanley J Reeves, "A regularized trust region method for joint reconstruction of spin magnitude, T_2^* decay, and off-resonance field map," 2014, IEEE International Conference on Imaging Processing.

[6] S. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," vol. 4, no. 7, pp. 932–946, 1995.

[7] Pierre Charbonnier, Laure Blanc-Feraud, Gilles Aubert, and Michel Barlaud, "Deterministic edge-preserving regularization in computed imaging," vol. 6, no. 2, pp. 298–311, 1997.

[8] J. Idier, "Convex half-quadratic criteria and interacting auxiliary variables for image restoration," *Image Processing, IEEE Transactions on*, vol. 10, no. 7, pp. 1001–1009, Jul 2001.

[9] Wangmeng Zuo and Zhouchen Lin, "A generalized accelerated proximal gradient approach for total-variation-based image restoration," *Image Processing, IEEE Transactions on*, vol. 20, no. 10, pp. 2748–2759, Oct 2011.

[10] S. Ramani and J.A. Fessler, "Parallel mr image reconstruction using augmented lagrangian methods," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 3, pp. 694–706, March 2011.