A KERNEL APPROACH TO PARALLEL MRI RECONSTRUCTION

Yuchou Chang[1], Dong Liang[1], Leslie Ying[1]

[1]Department of Electrical Engineering and Computer Science, University of Wisconsin-Milwaukee, Milwaukee, WI, USA

ABSTRACT

GRAPPA has been widely used as a *k*-space-based parallel MRI reconstruction technique. It linearly combines the acquired *k*-space signals to estimate the missing *k*-space signals where the coefficients are obtained by linear regression using auto-calibration signals. At high acceleration factors, GRAPPA reconstruction can suffer from a high level of noise even with a large number of auto-calibration signals. In this work, we improve the GRAPPA model using a kernel approach. Specifically, the acquired *k*-space data are mapped through a nonlinear transform to a high-dimensional space and then linearly combined to estimate the missing *k*-space data. A polynomial kernel is investigated in this work. Experimental results using phantom and *in vivo* data demonstrate that the proposed kernel GRAPPA method can significantly improve the reconstruction quality over the existing methods.

Index Terms— Parallel MRI, GRAPPA, Kernel method, Nonlinear filtering

1. INTRODUCTION

Generalized autocalibrating partially parallel acquisitions (GRAPPA) [1] has been widely used for reconstructing MR images from reduced acquisitions with multiple receivers. When the net acceleration factor is high, GRAPPA reconstruction can suffer from aliasing artifacts and noise amplifications. Methods have been developed in recent years to improve GRAPPA using localized coil calibration and variable density sampling [2], multicolumn multiline interpolation [3], regularization [4,5], iteratively reweighted least-squares [6], high-pass filtering [7], cross-validation [8,9], iterative optimization [10], virtual coil using conjugate symmetric signals [11], multi-slice weighting [12], or infinite pulse response (IIR) filtering [13], etc.

The conventional GRAPPA methods [1,3] reconstruct the missing *k*-space data by a linear combination of the acquired data, where the coefficients for combination are estimated using some additionally acquired auto-calibration signal (ACS) lines. Huang et al [14] analyzed two kinds of errors in GRAPPA reconstruction: truncation error and inversion error. Nana et al [8,9] extended the analysis and used more general terms: model error and noise-related error. While the first kind of error mainly originates from a limited number of ACS lines and data truncation, the second kind of errors generates from noise in the measured data and noise-induced error in estimating the coefficients for linear combination. Some methods use regularization [4,5] or iterative reweighted least-squares [6] to reduce the noise-induced error in the estimated coefficients. However, they are not widely accepted partially due to limited improvement or high computational complexity.

In this paper, we propose a kernel method to improve the conventional GRAPPA model and reduce the reconstruction error. We define a polynomial kernel function for the nonlinear mapping and the reconstruction problem of the missing *k*-space data is thereby formulated as a nonlinear combination of the acquired *k*-space data. Experimental results demonstrate that the proposed method outperforms the conventional GRAPPA, regularized GRAPPA, and the iterative reweighted least-squares methods in suppressing the spatial-varying noise.

2. BACKGROUND

The GRAPPA reconstruction can be represented as

$$S_j\left(k_y + r\Delta k_y, k_x\right) = \sum_{l=1}^{L}\sum_{b=B_1}^{B_2}\sum_{h=H_1}^{H_2} w_{j,r}\left(l,b,h\right) \times S_l\left(k_y + bR\Delta k_y, k_x + h\Delta k_x\right) \quad (1)$$

where the unacquired *k*-space signal S_j on the left-hand side is obtained by a linear combination of the acquired *k*-space signals on the right-hand side. Here w denotes the coefficient set, R represents the reduction factor, j is the target coil, l counts all coils, b and h transverse the acquired neighboring *k*-space data in k_y and k_x directions respectively, and the variables k_x and k_y represent the coordinates along the frequency- and phase-encoding directions, respectively. The formulation of GRAPPA can be simplified as a matrix equation

$$\mathbf{b}_{M\times 1} = \mathbf{A}_{M\times K}\mathbf{x}_{T\times K}\ , \quad (2)$$

where $\mathbf{A}$ represents the matrix comprised of the acquired data, $\mathbf{b}$ denotes the vector of the missing data, and $\mathbf{x}$ represents the coefficients.

In general, the coefficients depend on the coil sensitivities and are not known a priori. In GRAPPA, some auto-calibration data are acquired and used as the vector $\mathbf{b}$ to estimate the coefficient vector $\mathbf{x}$. The least-squares method is commonly used to calculate the coefficients:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\| \mathbf{b} - \mathbf{A}\mathbf{x} \right\|^2 \ . \quad (3)$$

When the matrix **A** is ill-conditioned at high reduction factors, the noise can be greatly amplified in the estimated coefficients. Regularization methods [4,5] have been used to solve for coefficients using

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\| \mathbf{b} - \mathbf{A}\mathbf{x} \right\|^2 + \lambda R(\mathbf{x}) \ , \quad (4)$$

where $R(\mathbf{x})$ is a regularization function (e.g., $R(\mathbf{x})=\|\mathbf{x}\|_2$ in Tikhonov) and λ is regularization parameter. It also has been noted that "outliers" due to noise and low sensitivity can lead to large error in the least-squares method. Iterative reweighted least-squares method [6] iteratively assigns and adjusts weights for the acquired data. "Outliers" are given less weights or removed in the final estimation so that the fitting accuracy and reconstruction quality is improved.

3. PROPOSED METHOD

The above GRAPPA formulation models the reconstruction as a standard linear regression and prediction problem. In practice, this linear model is not necessarily valid due to the noise in all measurements. The kernel method has recently been studied and shown to outperform linear regression and prediction [15,16,17]. The idea of kernel method is to transform the data nonlinearly to a higher dimensional space such that linear combination in the new space can approximate a broader class of nonlinear functions. Due to the improved accuracy in the regression model, the kernel method can improve the regression and prediction accuracy.

To transform the least square estimation into a kernel method algorithm, we apply a nonlinear mapping over the acquired *k*-space data $S_l\left(k_y + bR\Delta k_y, k_x + h\Delta k_x\right)$. Under such a mapping, we will need to solve the following linear system:

$$\mathbf{b} = \Phi(\mathbf{A})\mathbf{x} \ , \quad (5)$$

where $\Phi(\mathbf{A}) = \left[\Phi(\mathbf{a}_1), \Phi(\mathbf{a}_2), \cdots, \Phi(\mathbf{a}_M)\right]^T$ is a $M \times N_K$ matrix, the superscript T represents matrix transpose, M *is* the number of acquired ACS data, N_K is the dimension in the reproducing kernel Hilbert space (RKHS) [18] which is usually much higher than M, and $\mathbf{a}_i$'s are row vectors of the matrix **A**.

A kernel is a continuous, symmetric and positive-definite function, and is related to the mapping Φ in that

$$\kappa(\mathbf{a}_1, \mathbf{a}_2) = <\Phi(\mathbf{a}_1), \Phi(\mathbf{a}_2)>, \quad \forall \mathbf{a}_1, \mathbf{a}_2 \in \mathbf{A} \ , \quad (6)$$

where <,> represents the inner product. The generally used kernel includes polynomial kernel [19] and Gaussian kernel [20].

It is seen that the regression process to find **x** can still be solved by a linear algorithm:

$$\hat{\mathbf{x}} = \left(\left(\Phi(\mathbf{A})\right)^H \left(\Phi(\mathbf{A})\right) \right)^{-1} \left(\Phi(\mathbf{A})\right)^H \mathbf{b} \ . \quad (7)$$

Once the coefficients are estimated in Eq. (7), they are used in the prediction process to reconstruct the missing data in outer *k*-space, like the conventional GRAPPA does.

Since polynomials are widely used to approximate smooth unknown functions, we choose a polynomial kernel for Φ mapping. A polynomial kernel takes the following form

$$\kappa\left(\mathbf{a}_i, \mathbf{a}_j\right) = \left(\gamma \mathbf{a}_i^T \mathbf{a}_j + r\right)^d \ , \quad (8)$$

where γ and r are scalars and d is the degree of the polynomial. The polynomial kernel is the inner product between two vectors $\Phi(\mathbf{a}_i)$ and $\Phi(\mathbf{a}_j)$. If $\gamma = r = 1$ and $d = 2$, $\Phi(\mathbf{a})$ is given by [21]

$$\Phi(\mathbf{a}) = \left[1, \sqrt{2}a_1, \ldots, \sqrt{2}a_K, a_1^2, \ldots, a_K^2, \sqrt{2}a_1a_2, \ldots, \sqrt{2}a_i a_j, \ldots\right]^T \ , \quad (9)$$

where $a_1, a_2, \ldots, a_K$ are components of the vector **a**. We can see that the vector also includes the second-order terms. However, all components of $\Phi(\mathbf{a})$ cannot be incorporated in calculating the coefficients due to both large computation cost and over-fitting problem [22]. We keep the first-order components $1, \sqrt{2}\mathbf{a}_1, \ldots, \sqrt{2}\mathbf{a}_K$, and randomly and sparsely choose the second-order components to construct a hypothesis space

$$\tilde{\Phi}(\mathbf{a}) = \left[1, \sqrt{2}a_1, \sqrt{2}a_2, \cdots, \sqrt{2}a_K, a_{i_1}a_{j_1}, a_{i_2}a_{j_2}, \cdots\right], \quad (10)$$

where i, j are randomly chosen from 1, 2,…,K, such that the size of $\Phi(\mathbf{a})$ is equal to the desired dimension N_K. The proposed kernel GRAPPA method is thereby formulated as

$$\begin{aligned} S_j\left(k_y + r\Delta k_y, k_x\right) &= w_{j,r}^{(0)} \times 1 + \\ &\sum_{l=1}^{L} \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2} w_{j,r}^{(1)}(l,b,h) \times S_l\left(k_y + bR\Delta k_y, k_x + h\Delta k_x\right) \\ &+ \sum_{n=1}^{N} w_{j,r}^{(2)}(n) \times P(n) \times Q(n), \end{aligned} \quad (11)$$

where $P(n)$ and $Q(n)$ are randomly chosen from $S_l\left(k_y + bR\Delta k_y, k_x + h\Delta k_x\right)$ with $l = 1, \cdots, L; \quad b = B_1, \cdots, B_2;$ and $h = H_1, \cdots, H_2$, and N is the number of randomly selected second-order terms.

The above kernel formulation represents a more general model for GRAPPA, which includes the conventional GRAPPA as a special case. It is seen that the second part of kernel GRAPPA in Eq. (11) is equivalent to the conventional GRAPPA, which mainly captures the linear relationship between the missing and acquired signals in absent of noise and approximations. The first and third parts of the Eq. (11) can be used to characterize other nonlinear effects in practice such that noise and approximation errors are suppressed.

The performance of the proposed kernel GRAPPA method depends on the dimension of the RKHS, or the number of second-order terms N in Eq. (11). If the dimension is too low, prediction is inaccurate because the kernel space is not complex enough to accurately describe

the true relationship between the missing and acquired data, and thus the reconstructed image still suffers from large errors. On the other hand, if the dimension is too high, the model overfits the calibration data but poorly represents the missing data, thus leading to aliasing artifacts in reconstruction. Similar to conventional GAPPA where the number of blocks has to be optimized [8], the dimension of RKHS also needs to be chosen carefully to achieve a good reconstruction quality. As analyzed and demonstrated in our experiments, reconstruction result is generally good when N is 3-12 times of the number of the first-order terms. Within this range, the reconstruction is not sensitive to the changes in N.

4. EXPERIMENTAL RESULTS

The proposed method was tested on both phantom and *in vivo* data sets. All data sets were acquired in full and then manually undersampled to simulate the accelerated acquisition. The sum-of-squares (SoS) reconstruction from the fully sampled data of all channels was used as the reference image for comparison. The reconstructions from reduced data using the conventional GRAPPA, Tikhonov regularized GRAPPA [4], GRAPPA with iterative reweighted least-squares (IRLS) [6], and the proposed kernel GRAPPA method were compared visually and also in terms of normalized mean-squared error (NMSE). In all methods, the number of blocks (size of b) takes 4 and the number of columns (size of h) takes 15 for the phantom data and 7 for the brain data. An outer reduction factor (ORF) of 5 with 56 ACS lines was used for the 8-channel phantom data and an ORF of 4 with 48 ACS lines for the 4-channel brain data. For the proposed kernel GRAPPA method, N was chosen to be 4 and 7 times of the number of the first-order terms for the phantom and brain data, respectively.

Figure 1 shows the reconstructions of the phantom. It is seen that the conventional GRAPPA suffers from large noise at high accelerations. Tikhonov regularization and IRLS can both improve the signal-to-noise ratio (SNR) to some extent but at the cost of aliasing artifacts. The proposed kernel GRAPPA method suppresses most noise without introducing additional artifacts.

Figure 2 shows the reconstruction results from a set of *in vivo* brain data. A small region is zoomed to show more details. It is seen that the reconstruction using the proposed method achieves a quality superior to all other methods. The proposed method effectively removes the spatial-varying noise in the conventional GRAPPA reconstruction without the artifacts in Tikhonov regularization and IRLS methods. Furthermore, the proposed method also preserves the resolution of the original image without blurring. Quantitative measure via NMSE also shows that the proposed method outperforms other methods.

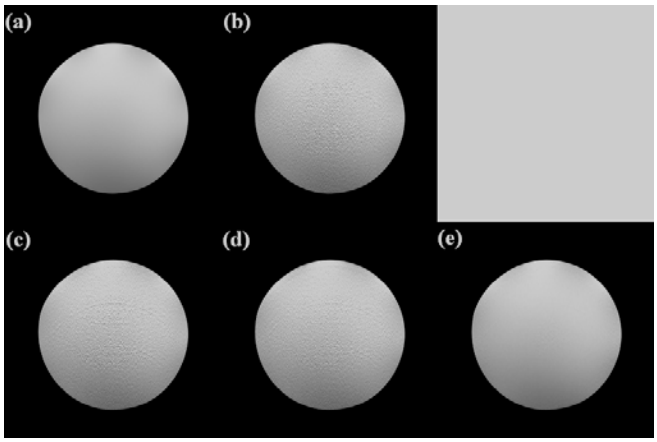

Fig.1 With (a) the SoS reconstruction of the phantom image as the reference, we compare (b) conventional GRAPPA, (c) Tikhonov regularization, (d) IRLS reconstruction, and (e) the proposed method when ORF = 5, ACS = 56.

In the proposed method, the number of second-order terms N needs to be tuned to achieve optimal reconstruction quality. We study the effects of N in reconstruction quality by changing it from one time of the number of the first-order terms to larger times. When the value of N is increasing, noise is gradually suppressed, but aliasing artifacts gradually appear. Quantitative measure via NMSE also illustrates that too small or too large N deteriorates reconstruction quality as shown in Fig. 3. There is an optimal range for the value of N at the bottom of the U shape where both noise and aliasing artifacts are suppressed. The rather flat bottom of the U shape indicates that the proposed method is insensitive to the choice of N within a certain range. Because the value of N directly affects the computational complexity, smaller N's are preferred within this range. It is also worth noting that the performance is insensitive to the choice of second-order terms (e.g., random or fixed) as long as the same number of terms is used. The computation time of the proposed method is about 3-5 times of the conventional GRAPPA and Tikhonov-regularized GRAPPA, while IRLS is the most time consuming among all. Furthermore, Tikhonov regularization can also be easily incorporated into the proposed reconstruction method in high-dimensional RKHS.

5. CONCLUSION

In this paper, we propose a novel kernel-based k-space reconstruction algorithm for parallel MRI. The method maps the data onto a higher dimensional space through a nonlinear transformation such that the nonlinear model can characterize the relationship between the acquired and missing data more accurately. The experimental results demonstrate that the proposed method is superior to the conventional GRAPPA and other improved GRAPPA methods in suppressing both noise and artifacts.

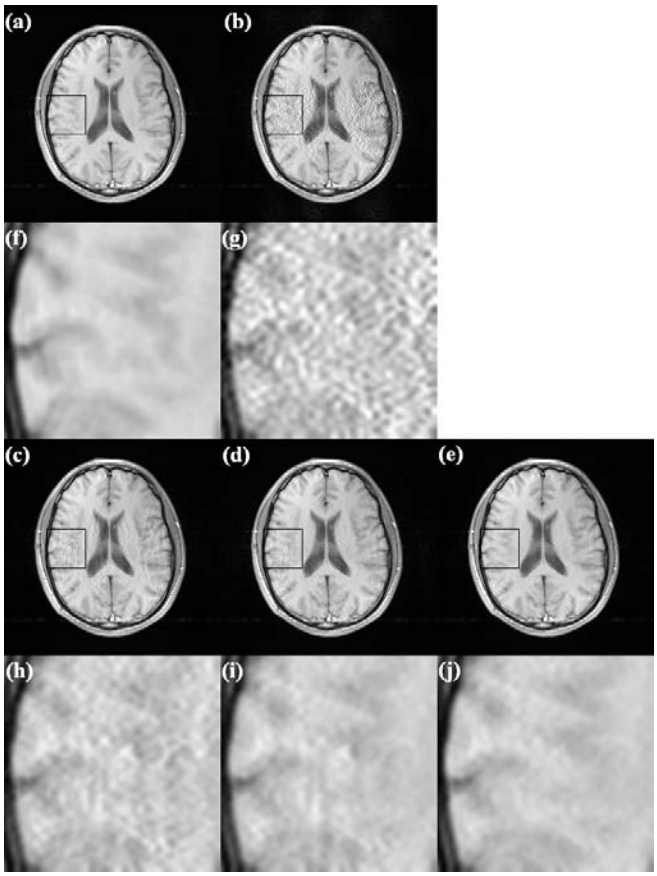

Fig.2 With (a) the SoS reconstruction of the 4-chanel brain image as the reference, we compare (b) conventional GRAPPA, (c) Tikhonov regularization, (d) IRLS reconstruction, and (e) the proposed method for ORF = 4, ACS = 48. Subfigures (f) – (j) show zoomed-in portions of (a) – (e) respectively.

6. ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation CBET-0731226 and a grant from The Lynde and Harry Bradley Foundation.

7. REFERENCES

[1] M.A. Griswold, P.M. Jakob, R.M. Heidemann, N. Mathias, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, pp. 1202-1210, Jun. 2002.

[2] J. Park, Q. Zhang, V. Jellus, O. Simonetti, and D. Li, "Artifact and noise suppression in GRAPPA imaging using improved k-space coil calibration and variable density sampling," *Magn. Reson. Med.*, vol. 53, pp. 189-193, Jan. 2005.

[3] Z. Wang, J. Wang, and J.A. Detre, "Improved data reconstruction method for GRAPPA," *Magn. Reson. Med.*, vol. 54, pp. 738-742, Jul. 2005.

[4] P. Qu, C. Wang, and G.X. Shen, "Discrepancy-based adaptive regularization for GRAPPA reconstruction," *J. Magn. Reson. Imaging*, vol. 24, pp. 248-255, Jul. 2006.

[5] F.H. Lin, "Prior-regularized GRAPPA reconstruction," In *Proc. of 14th ISMRM Scientific Meeting*, (Seattle, WA), 3656, May 2006.

[6] D. Huo, and D.L. Wilson, "Robust GRAPPA reconstruction and its evaluation with the perceptual difference model," *J. Magn. Reson. Imaging*, vol. 27, pp. 1412-1420, Jun. 2008.

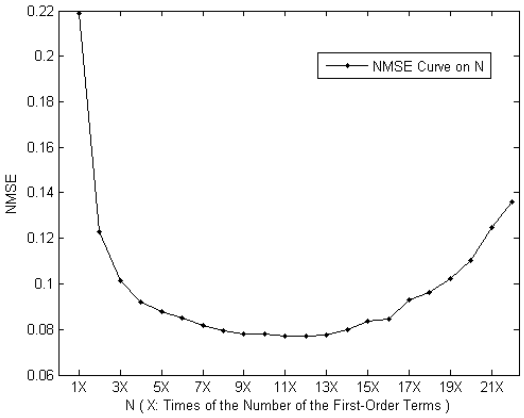

Fig.3 NMSE curves of the proposed method as a function of different *N* values for the brain data.

[7] F. Huang, Y. Li, S. Vijayakumar, S. Hertel, and G.R. Duensing, "High-pass GRAPPA: an image support reduction technique for improved partially parallel imaging," *Magn. Reson. Med.*, vol.59, pp. 642-649, 2008.

[8] R. Nana, T. Zhao, K. Heberlein, S.M. LaConte, and X. Hu, "Cross-validation-based kernel support selection for improved GRAPPA reconstruction," *Magn. Reson. Med.*, vol. 59, pp. 819-825, Apr. 2008.

[9] R. Nana and X. Hu, "Data consistency criterion for selecting parameters for *k*-space-based reconstruction in parallel imaging," *Magn. Reson. Imaging*, vol. 28, pp. 119-128, Jan. 2010.

[10] T. Zhao and X. Hu, "Iterative GRAPPA (iGRAPPA) for improved parallel imaging reconstruction," *Magn. Reson. Med.*, vol. 59, pp. 903-907, Apr. 2008.

[11] M. Blaimer, M. Gutberlet, P. Kellman, F.A. Breuer, H. Kostler, and M.A. Griswold, "Virtual coil concept for improved parallel MRI employing conjugate symmetric signals," *Magn. Reson. Med.*, vol. 61, pp. 93-102, Jan. 2009.

[12] M. Honal, S. Bauer, U. Ludwig, and J. Leupold, "Increasing efficiency of parallel imaging for 2D multislice acquisitions," *Magn. Reson. Med.*, vol. 61, pp. 1459-1470, Jun. 2009.

[13] Z. Chen, J. Zhang, R. Yang, P. Kellman, L.A. Johnston, and G.F. Egan, "IIR GRAPPA for Parallel MR Image Reconstruction," *Magn. Reson. Med.*, vol. 63, pp. 502-509, Feb. 2010.

[14] F. Huang and G.R. Duensing, "A theoretical analysis of errors in GRAPPA," In *Proc. of 14th ISMRM Scientific Meeting*, (Seattle, WA), 2468, May 2006.

[15] P. Tomaso and S. Steve, "The mathematics of learning: dealing with data," *Notices of the American Mathematical Society*, vol. 50, pp. 537-544, May, 2003.

[16] W. Liu, P.P. Pokharel, and J.C. Principe, "The kernel least-mean-square algorithm," *IEEE Trans. on Signal Processing*, vol. 56, pp. 543-554, Feb. 2008.

[17] G. Camps-Valls, J.L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel methods in bioengineering, signal and image processing," *Idea Group Publishing*, London, UK, Jan. 2007.

[18] A. Berlinet and C. Thomas-Agnan,*Reproducing kernel Hilbert spaces in probability and statistics*, *Springer-Verlag*, New York, Dec. 2003.

[19] T. Frieb and R. Harrison, "A kernel-based adaline," In *Proc. Eur. Symp Artificial Neural Networks (ESANN)*, (Bruges, Belgium), pp. 245-250, Apr. 1999.

[20] K. Kim, M. O. Franz, and B. Scholkopf, "Iterative kernel principal component analysis for image modeling," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 27, pp. 1351-1366, Sep. 2005.

[21] Y.W. Chang, C.J. Hsieh, K.W. Chang, M. Ringgaard, and C.J. Lin, "Training and testing low-degree polynomial data mappings via linear SVM," *J. of Mach. Learning Research*, vol. 11, pp. 1471-1490, Apr. 2010.

[22] T. Zhang, "Learning bounds for kernel regression using effective data dimensionality," *Neural Computation*, vol. 17, pp. 2077-2098, 2005.