SCIENTIFIC REP🌼RTS

OPEN

Improved Parallel Magnertic Resonance Imaging reconstruction with Complex Proximal Support Vector Regression

Lin Xu[1], Qian Zheng[2] & Tao Jiang[1]

Generalized Auto-calibrating Partially Parallel Acquisitions (GRAPPA) has been widely used to reduce imaging time in Magnetic Resonance Imaging. GRAPPA synthesizes missing data by using a linear interpolation of neighboring measured data over all coils, thus accuracy of the interpolation weights fitting to the auto-calibrating signal data is crucial for the GRAPPA reconstruction. Conventional GRAPPA algorithms fitting the interpolation weights with a least squares solution are sensitive to interpolation window size. MKGRAPPA that estimates the interpolation weights with support vector machine reduces the sensitivity of the k-space reconstruction to interpolation window size, whereas it is computationally expensive. In this study, a robust GRAPPA reconstruction method is proposed that applies an extended proximal support vector regression (PSVR) to fit the interpolation weights with wavelet kernel mapping. Experimental results on *in vivo* MRI data show that the proposed PSVR-GRAPPA method visually improves overall quality compared to conventional GRAPPA methods, while it has faster reconstruction speed compared to MKGRAPPA.

Parallel imaging technology using multi-channel phased array coils is widely used to speed up Magnetic Resonance Imaging (MRI) scanning by acquiring only a fraction of k-space data. Alias-free image can be reconstructed from incomplete k-space data with various reconstruction algorithms, e.g., sensitivity encoding (SENSE), simultaneous acquisition of spatial harmonics (SMASH), generalized auto-calibrating partially parallel acquisition (GRAPPA) and their derivatives[1–7]. These reconstruction methods can be generally classified into image domain methods and k-space methods. The image domain methods[1,7] unfold aliased signals in image with accurate coil sensitivity information. The k-space methods linearly join adjacent acquired data over all coils to interpolate the value of missing data where interpolation weights are estimated by fitting to a small number of additional acquired auto-calibration signals (ACS). The k-space methods are superior in cases where determination of the coil sensitivity is difficult, whereas conventional k-space methods suffer from low signal-noise-rate (SNR) of reconstructed image. This pitfall was resolved by GRAPPA that reconstructs individual coil image separately and then combines all coil images in a sum-of-squares (SOS) fashion[8].

GRAPPA reconstruction accuracy strongly depends on the selection of interpolation window. Inadequate interpolation windows could cause either under-fitting or over-fitting of interpolation weights, which can lead to serious noise and artifacts in the reconstructed image. Methods in[9–11] defined various kinds of error metrics to evaluate all potential interpolation windows and chose the interpolation window with minimal error to reconstruct image, whereas such methods just focus on the reconstruction error in calibration region. Nana *et al.*[12] utilized the shift-invariance of interpolation weights in k-space to approximate the predication error by calculating the difference between acquired signals and their estimates obtained based on the interpolation of the missing data. Although this method can determine the optimal interpolation window, time-exhausted reconstruction of all acquired k-space data for each possible interpolation window limits its application on large number of coils data.

Although the optimal interpolation window can be determined with aforementioned methods under low reduce factors, the conventional GRAPPA reconstructions using linear interpolation perform poorly at high

[1]Control Engineering College, Chengdu University of Information Technology, Chengdu, China. [2]Zhengzhou University of Light Industry, Zhengzhou, China. Correspondence and requests for materials should be addressed to T.J. (email: mripet@163.com)
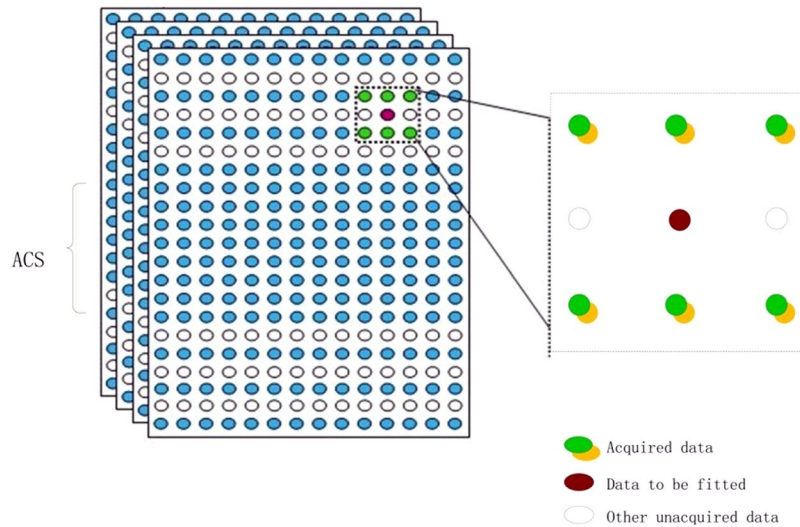
Figure 1. Schematic description of GRAPPA.

acceleration situations aroused by under-fitting of interpolation weights. Chang *et al.* proposed a nonlinear GRAPPA (NLGRAPPA) reconstruction[13], which uses a truncated polynomial kernel to map the *k*-space data onto a higher dimension feature space. Despite providing a more accurate fitting on the spatial correlation of coils compared with linear GRAPPA, the nonlinear transformation makes NLGRAPPA reconstruction more sensitive to interpolation windows. KerNL[14] introduced kernel tricks to represent the general nonlinear relationship between acquired and missing *k*-space data that improves both image quality and computation efficiency at high reduction factors. More recently, MKGRAPPA[15] reformulated estimating the optimal interpolation weights as a multi-kernel learning problem that can reduce the sensitivity of the *k*-space reconstruction to interpolation windows. Nevertheless, the MKGRAPPA is computationally expensive due to the involved semi-infinite linear programming.

In this study, we theoretically analyze errors in GRAPPA reconstruction from the perspective of structure risk minimization (SRM) principle[16] that can achieve favorable compromise between fitting error and complexity of fitting function for regression problem. We thereby develop a robust GRAPPA reconstruction method that estimates the interpolation weights by proximal support vector machine (PSVR)[17]. PSVR is a advantaged version of support vector machine (SVM)[16] developed from the SRM principle. PSVR has looser constraints than does SVM, with comparable performance and much lower computational cost. Experimental results of *in vivo* brain imaging are provided to demonstrate the performance of the proposed method named "PSVR-GRAPPA" compared with the GRAPPA and MKGRAPPA methods.

The rest of this paper is organized as follow: Section 2 reviews the theory of GRAPPA algorithm and SRM, then describes the details of proposed method which applies an extended PSVR to fit the interpolation weights with wavelet kernel mapping; Section 3 applies the proposed method *in vivo* experiments; Section 4 gives the conclusions and future work to be done.

Methods

Review of GRAPPA.   GRAPPA estimates the missing *k*-space data in each coil by a linear combination of its nearby acquired data over all coils, which can be mathematically represented as

$$sl(kx, ky + m\Delta ky) = \sum_{j=1}^{L} \sum_{h=H_1}^{H_2} \sum_{t=N_1}^{N_2} w_{l,m}(j, h, t)s_j(kx + h \cdot \Delta kx, ky + t \cdot R \cdot \Delta ky)$$
(1)

where *s* represents the *k*-space signal and *w* denotes the interpolation weights, (k_x, k_y) are *k*-space coordinates along the phase encoding (PE) and frequency encoding (FE) directions, respectively. Δk_x and Δk_y are sampling intervals along the PE and FE directions, respectively. In Eq. (1), *j* and *l* represent coil indexes, *L* is the number of coils in the array, *m* is the offset from the acquired data, N_1 and N_2 define the lower and upper bounds of the interpolation window along the PE direction, H_1 and H_2 define the left and right bounds of the interpolation window along the FE direction, respectively.

As shown in Fig. 1, the central region of *k*-space is fully sampled to calibrate the interpolation weights *w*. A linear system with knowing input and output is constructed as follows,

$$y = Aw + \varepsilon$$
(2)

where *y* is a vector constituted of all target data, and *A* is an $\ell*M$ matrix composed of the interpolation source data supposing that there are *M* interpolation source data points in the interpolation window and ℓ target data points to be interpolated.
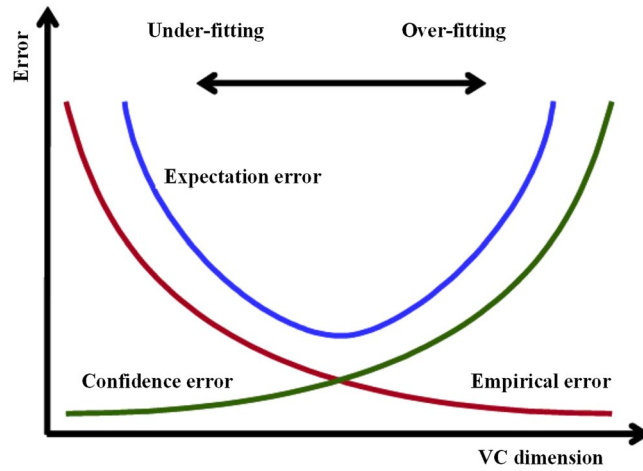
Figure 2. The expectation error is the sum of the empirical error and the confidence error. The empirical error is decreased with the VC dimension, while the confidence error is increased.

SRM model of GRAPPA.

In general, conventional GRAPPA reconstruction methods utilize the least squares method to estimate the interpolation coefficients w, which minimizes the following empirical error,

$$R_{emp}(w) = \sum_{i=1}^{\ell}\left(y_i - \sum_{k=1}^{M}w_k a_i^k\right)^2 = \left\|y_i - \sum_{k=1}^{M}w_k a_i^k\right\|_2 \tag{3}$$

Equation (3) indicates that $R_{emp}(w)$ decreases with the increasing of interpolation window size, and $R_{emp}(w)$ gets close to zero when M equals to l. Nonetheless, total error would not be minimal as interpolation window is oversized. Previous studies[11,12] indicate that the GRAPPA reconstruction error would increase as the interpolation window size increases if the latter reaches a threshold. Therefore, both interpolation weights w and interpolation window size M need to be optimally determined to obtain a perfect reconstruction.

GRAPPA reconstruction can be viewed as a supervised learning problem. According to the capacity concept of Vapnik-Chervonenkis (VC) theory[16], the total error of supervised learning problem can be described as expectation error $R(w)$, which can be represented as follows,

$$R(w) \leq R_{emp}(w) + \Phi(h/\ell) \tag{4}$$

where $R_{emp}(w)$ is the empirical error decreased as h increases, and $\Phi(h/\ell)$ is confidence error. h is VC dimension indicating the complexity of interpolation weights. The dimension of interpolation weights w is a good estimation of VC dimension for linear interpolation weights, thus a larger interpolation window indicates a higher VC dimension. Equation (4) shows that increasing the interpolation window size can reduce the empirical error while increase the confidence error. As shown in Fig. 2, the expectation error is an "U"-shape function of VC dimension. Consequently, the total expectation error will increase when the interpolation window size increases after a certain value.

Actually, the least squares solution only minimizes $R_{emp}(w)$ but not $R(w)$, which is the optimal solution in the sense that $\ell \rightarrow +\infty$, i.e., $\Phi(h/\ell) \rightarrow 0$. Without considering the confidence error, the interpolation weights fitted with the least squares method have poor generalization ability that the fitted model adapts properly to new data.

Proposed PSVR-GRAPPA.

It is well-known that superior generalization performance can be obtained from SVM and more importantly, the performance does not depend on the dimensionality of the input data[16]. Therefore, the SVM could effectively address the aforementioned problem in the GRAPPA reconstruction. Unfortunately, the original SVM that solves the constrained quadratic programming (QP) problem with linear constraints has a time complexity $O(N^3)$[18]. PSVR is an alternate formulation of SVM regression that dramatically reduces the computation. Consequently, a complex PSVR for GRAPPA is proposed to reduce the sensitivity of interpolation window size in GRAPPA reconstruction.

Complex PSVR for GRAPPA.

GRAPPA employs an unbiased interpolation weights to approximate the relation of interpolation source and target data, whereas recent work[13] indicated that the bias drastically increases as the noise increases at high accelerations. For this reason, we added a bias variable "b" into the cost function to improve fitting accuracy. Hence, interpolation weights can be solved with the following optimization problem:

$$J = \min \frac{1}{2}(w^H w + b^H b) + \frac{C}{2}\sum_{k=1}^{l}\eta_k^{\ H}\eta_k$$
$$\text{s.t } y_k - w^H\varphi(x_k) - b = \eta_k, \ k = 1, ..., l \tag{5}$$

where w and b are coefficients of interpolation weights, C is a punishment factor to balance the training error rate and the complexity of the model. η_κ is a slack variable for residual term. The superscript H denotes the operation of Hermitian Transpose. $\varphi(x_\kappa)$ is a vector containing the source data to interpolate target data y_κ. In Eq. (5), the $w^H w$ is an estimate of VC dimension of interpolation weights and $\eta_\kappa{}^H \eta_\kappa$ represents individual empirical error, thereby minimizing the sum of such two terms can obtain an optimal solution with minimal expectation error. Theoretically, the minimal expectation error can give the best reconstruction quality given the training samples and interpolation window size. However, it is worth to note that inappropriate determination parameter would discount the performance of minimal expectation error. A detailed parameters optimization will be given in the below subsection.

Considering that the k-space data are complex value while the original PSVR was proposed at real space, we define the operation

$$L(\varepsilon_n) = L(\Re\{\varepsilon_n\}) + L(\Im\{\varepsilon_n\}) \tag{6}$$

where $\Re\{\}$ and $\Im\{\}$ denote the operations of taking the real part and imaginary part of complex data, respectively. Consequently, problem Eq. (5) can be stated as:

$$
\begin{aligned}
J = \ & \min \ \frac{1}{2}(w^H w + b^H b) + \frac{C}{2}\sum_{k=1}^{l}(\xi_k{}^2 + \zeta_k{}^2) \\
& \text{s.t } \Re(y_k - w^H\varphi(x_k) - b) = \xi_k, k = 1, ..., l \\
& \Im(y_k - w^H\varphi(x_k) - b) = \zeta_k, k = 1, ..., l
\end{aligned}
\tag{7}
$$

where slack variables ξ, ζ are introduced for both real and imaginary residual terms. The dual problem of Eq. (7) is obtained by introducing Lagrange multipliers α and β:

$$
\begin{aligned}
L = \ & \min \ \frac{1}{2}(w^H w + b^H b) + \frac{C}{2}\sum_{k=1}^{l}(\xi_k{}^2 + \zeta_k{}^2) \\
& + \sum_{k=1}^{l}\alpha_k\Re[w^H\varphi(x_k) + b + \xi_k - y_k] \\
& + \sum_{k=1}^{l}\beta_k\Im[w^H\varphi(x_k) + b + \xi_k - y_k]
\end{aligned}
\tag{8}
$$

From the Karush-Kuhn-Tucker (KKT) conditions for optimality, we find:

$$
\begin{aligned}
\frac{\partial L}{\partial w} &= 0 \rightarrow w = -\sum_{k=1}^{l}\alpha_k\Re(\varphi(x_k)) - \sum_{k=1}^{l}\beta_k\Im(\varphi(x_k)) \\
\frac{\partial L}{\partial b} &= 0 \rightarrow \sum_{k=1}^{l}(\alpha_k + j \cdot \beta_k) = -b \\
\frac{\partial L}{\partial \xi} &= 0 \rightarrow \alpha_k = C \cdot \xi_k, k = 1, ..., l \\
\frac{\partial L}{\partial \zeta} &= 0 \rightarrow \beta_k = C \cdot \zeta_k, k = 1, ..., l \\
\frac{\partial L}{\partial \alpha} &= 0 \rightarrow \Re[w^H\varphi(x_k) + b + \xi_k - y_k] = 0 \\
\frac{\partial L}{\partial \beta} &= 0 \rightarrow \Im[w^H\varphi(x_k) + b + \xi_k - y_k] = 0
\end{aligned}
\tag{9}
$$

After elimination of the variables w, ξ and ζ, one can derive the following linear system:

$$
\begin{bmatrix} \Omega_r + C^{-1}I - \overrightarrow{1}_v^T \overrightarrow{1}_v & 0 \\ 0 & \Omega_i + C^{-1}I - \overrightarrow{1}_u^T \overrightarrow{1}_u \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \Re(y) \\ \Im(y) \end{bmatrix}
\tag{10}
$$

where I is an identity matrix, $\overrightarrow{1} = [1; ...; 1]^T$, Ω_r and Ω_i are Gramm matrices wherein $\Omega_{kl} = K(x_k, x_l)$, $k = 1, ..., l$, and $K(x_k, x_l)$ is a kernel function that will be described in the below subsection.

Lagrange multipliers α and β can be solved in Eq. (10) with ACS data, then other missing data can be predicted as follows,

$$
\begin{aligned}
f(m) = \ & w^H\varphi(m, x) + b = \sum_{k=1}^{l}\alpha_k K(\Re(m'), \Re(x_k)) \\
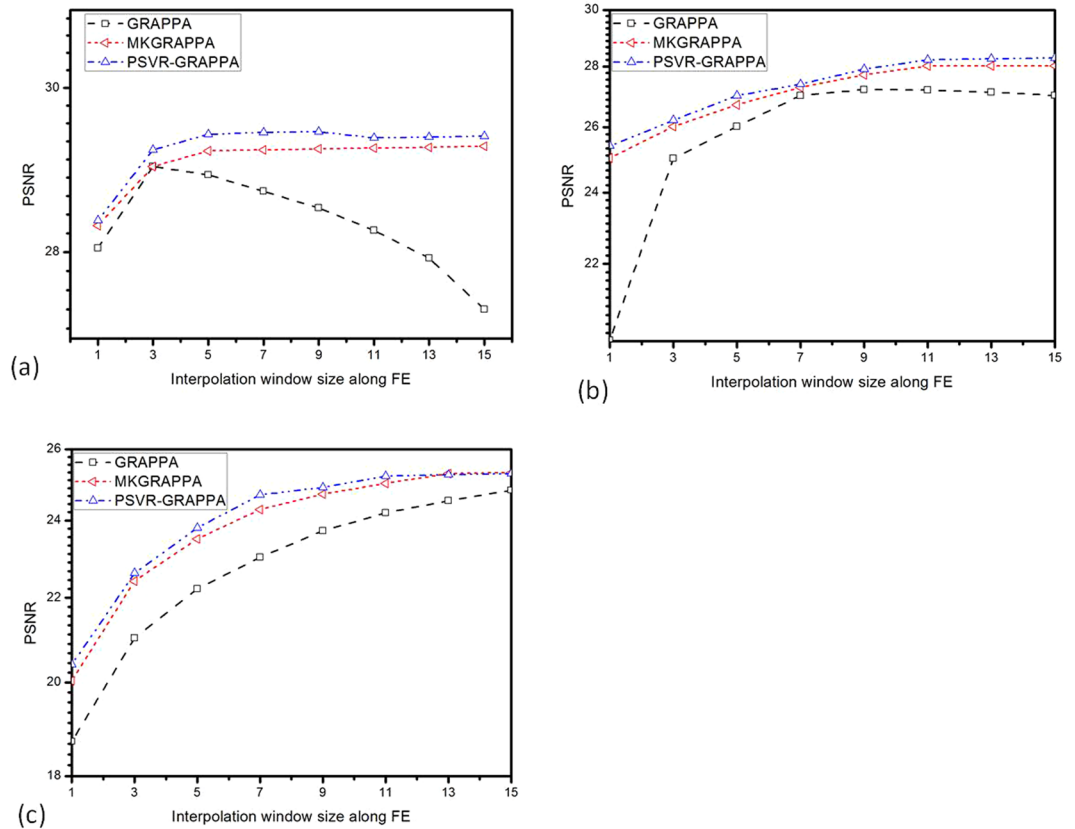& + j*\sum_{k=1}^{l}\beta_k K(\Im(m'), \Im(x_k)) + b
\end{aligned}
\tag{11}
$$

Figure 3. Plots of PSNR of brain data vs. interpolation window size along the FE direction for different sampling conditions: (**a**) $R = 2$, ACS lines $= 12$; (**b**) $R = 3$, ACS lines $= 24$; and (**c**) $R = 4$, ACS lines $= 40$.

where vector x consists of interpolation source data in ACS, and m' denotes the surrounding acquired data of missing data m, j is the square root of -1. The bias term b can be computed as follows,

$$b = -\sum_{k=1}^{l}(\alpha_k + j \cdot \beta_k)$$

Experments and Results

Data acquisition.   The proposed method was evaluated on two sets of *in vivo* human brain data (axial and sagittal) acquired using an SE pulse sequence (TE/TR $= 14/400$ ms, 33.3 kHz bandwidth, 256×256 pixels, FOV $= 240*240$ mm^2) on a 1.5 T scanner (Siemens Healthcare, Erlangen, Germany) with an 8-channel head coils. The data were fully sampled and later decimated in the PE direction by various factors to mimic parallel imaging acquisition procedure. All *in vivo* data were collected from healthy adult human volunteers with written informed consent from the participants in accordance with policies of the Institutional Review Board(IRB). All experimental protocols were approved by the IRB of University of Electronic Science and Technology of China (Chengdu, China).

Parameters optimization.   The punishment factor C needs to be optimally tuned in the proposed method. Although global or local optimization techniques such as genetic or gradient algorithms can obtain an approximately optimal parameter, such time-consuming methods may not be proper in the applications of real-time requirement. In this work, a heuristics selection algorithm[19] is utilized for the determining of punishment factor C:

$$C = \max(|\overline{y} + 3\sigma|, |\overline{y} - 3\sigma|) \tag{12}$$

where $\overline{y}$ and σ are mean and standard deviation of interpolation target k-space data in ACS region, respectively. Consequently, the parameter C was adaptively computed for each reconstruction task.

The kernel mapping is the basic characteristic for SVM. One can flexibly choose appropriate kernels for various applications. Nonetheless, it is intractable to choose the best kernel in practical application. Therefore, the performances of common used kernels were investigated with a set of sagittal data in this study. GRAPPA and PSVR-GRAPPA with different kernels were compared under the same sampling condition: R is 5 and 40 ACS lines were used to calibrate the interpolation weights. These kernels include:
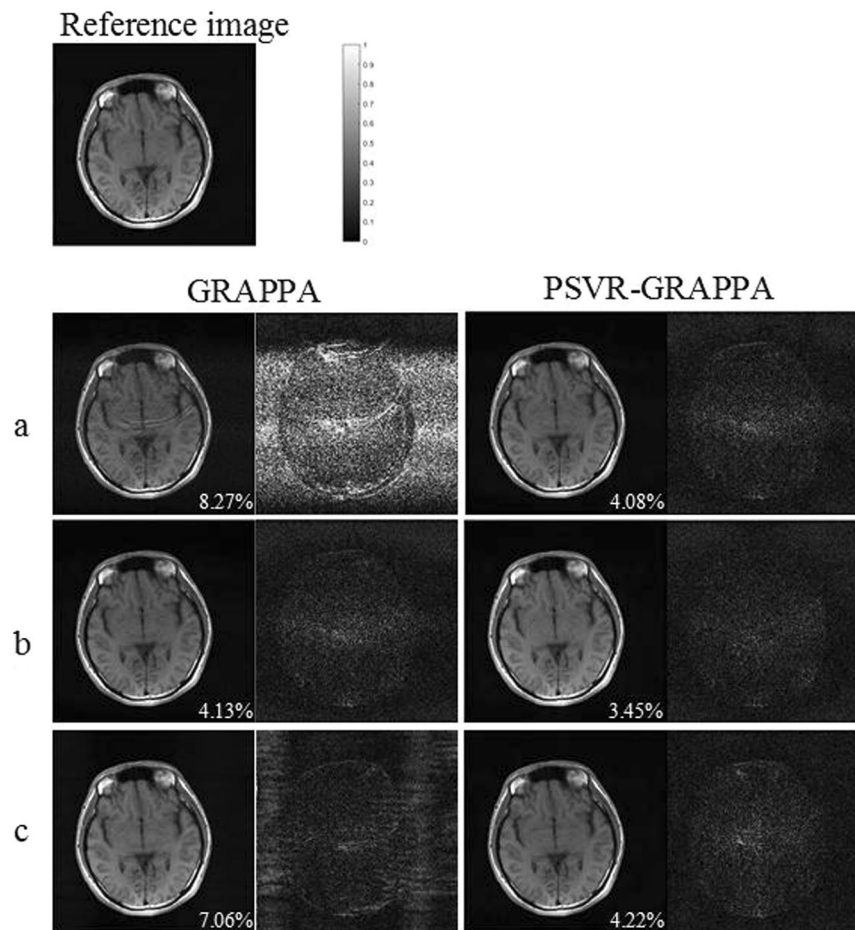
Reference image

GRAPPA · PSVR-GRAPPA

Figure 4. Axial brain images reconstructed by GRAPPA(left) and PSVR-GRAPPA (right) under sampling parameters: $R = 3$ with 24 ACS lines. Both methods were reconstructed in the entire k-space with interpolation windows: 4*3(**a**), 2*7 (**b**) and 4*15 (**c**). Nonaccelerated image is used as reference (top). To the right of each reconstructed image its absolute error with the nonaccelerated image is shown.

Polynomial: $K(x, y) = (x \cdot y + \theta)^d$, where θ and d are positive integers to define the constant term and order, respectively.

RBF: $K(x, y) = \exp(-\|x - y\|^2/\gamma^2)$, where γ is a parameter to turn the width of RBF kernel.

Wavelet kernel: $K(x, y) = \prod_{i=1}^{N} h\left(\frac{x_i - y_i}{q}\right)$ where $h(x) = \cos(1.75x)\exp\left(-\frac{x^2}{2}\right)$.

The polynomial kernel with 2-order and $\theta = 1$, the width of RBF kernel γ was chosen followed the work[18]: $\gamma = \frac{d_{\max}}{\sqrt{2l}}$, where d_{max} is the maximum distances between all training data. Based upon a number of empirical comparisons, we found that the following equation works well for estimating a desirable scale factor for wavelet kernel:

$$q = d_{\max} \log l \tag{13}$$

Image reconstruction. Reconstructions were performed on a desktop personal computer using MATLAB (The Mathworks, Natick, MA, USA). A quantitative assessment of reconstruction performance was performed by computing the NMSE and PSNR of reconstructed image:

$$NMSE = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|I_{i,j}^{recon} - I_{i,j}^{ref}\right|^2 / \sum_{i=1}^{N}\sum_{j=1}^{N}\left|I_{i,j}^{ref}\right|^2}$$

$$PSNR = 10 \times \log(255 \times 255/NMSE)$$

where $I_{i,j}^{ref}$ is the result of non-accelerated scenario, i and j are the indices of image pixel along PE and FE directions, respectively.

The effect of interpolation window size on reconstruction quality was investigated with the sagittal data. As reduce factor R increased from 2 to 4, the images were reconstructed by GRAPPA and PSVR-GRAPPA with varying interpolation windows, respectively. The interpolation window size along the PE direction was fixed to 2, and
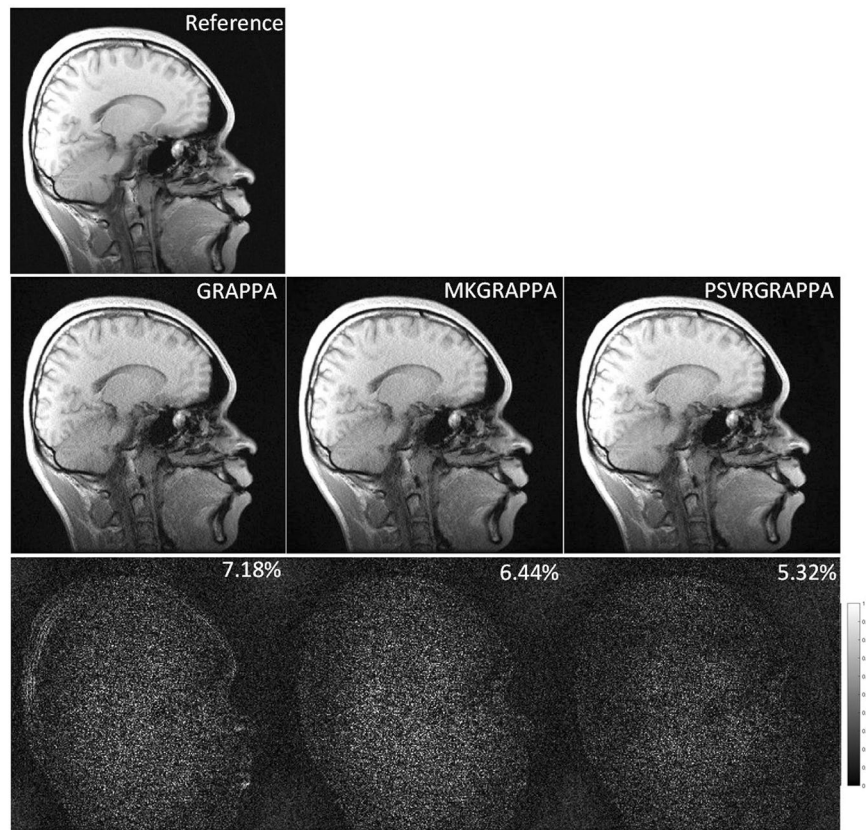
Figure 5. Sagittal brain images reconstructed by GRAPPA(left), MKGRAPPA(middle), and PSVR-GRAPPA (right) under sampling parameters: $R = 4$ with 40 ACS lines. All methods were reconstructed in the entire k-space with an interpolation window of 2*9. Nonaccelerated image is used as reference (top). To the bottom of each reconstructed image its absolute error with the nonaccelerated image is shown.

that along the FE direction was varied from 1 to 15. ACS lines were adopted with 12, 24 and 40 for R of 2, 3 and 4, respectively. To qualitatively analysis the effect of interpolation window size on reconstruction quality, the axial brain dataset was reconstructed by GRAPPA and PSVR-GRAPPA with three kinds of interpolation windows: 4*3, 2*7, 4*15. The sampling parameters: $R = 3$, ACS lines = 24.

The reconstruction quality of PSVR-GRAPPA with wavelet kernel mapping was compared against those of GRAPPA and MKGRAPPA with the sagittal brain dataset. All methods reconstructed in the entire k-space with an interpolation window size of 2*9. The sampling parameters: $R = 4$, ACS lines = 40. The full k-space data were reconstructed by 2-D Inverse Fast Fourier Transform (IFFT) for each coil, and then combined all coil images in the SOS method to give the reference image.

Results.   Figure 3 presents plots of variation of PSNR as increasing of interpolation window size in GRAPPA, MKGRAPPA and PSVR-GRAPPA at R of 2 (Fig. 3a), 3 (Fig. 3b) and 4 (Fig. 3c). Black lines represent the results of GRAPPA, red lines represent the results of GRAPPA, and blue lines indicate the results of PSVR-GRAPPA. Totally PSVR-GRAPPA reconstructed images have highest PSNRs among the three methods. All methods have low PSNRs when the interpolation window size is small regardless to the R. This phenomenon can be interpreted with the under-fitting of interpolation weights aroused by the small size of interpolation window. Therefore, the PSNR can be increased by increasing the interpolation window size. In Fig. 3a,b, the increasing of the interpolation window will decrease PSNR once the interpolation window size reaches a certain threshold, which is caused by the over-fitting of interpolation weights. At high accelerations (Fig. 3c), the linear model cannot catch sufficiently the complexity of spatial information of coils in k-space since the interpolation target data are far away from the interpolation source data, which results an under-fitting of interpolation weights. As a result, the fitting error that decreases as VC dimension h increases grows to be the dominating error. Hence, the PSNRs of all three methods increased with the increasing of the interpolation window size in Fig. 3c. Nonetheless, the PSNRs of MKGRAPPA and the proposed method increase faster than that of GRAPPA because that the nonlinear kernel has a higher h than linear model does. Although Fig. 3 shows that the reconstruction performance of all methods can be affected with interpolation window size along FE, MKGRAPPA and PSVR-GRAPPA based on the structural risk minimization theory are less sensitive to the changes of interpolation window than GRAPPA.

Figure 4 shows axial brain images reconstructed by GRAPPA and PSVR-GRAPPA with interpolation windows: 4*3(Fig. 4a), 2*7(Fig. 4b) and 4*15(Fig. 4c). The absolute error image is show on the right of the reconstruction result, respectively. Serious aliasing artifacts and amplified noise can be seen in the image reconstructed by GRAPPA
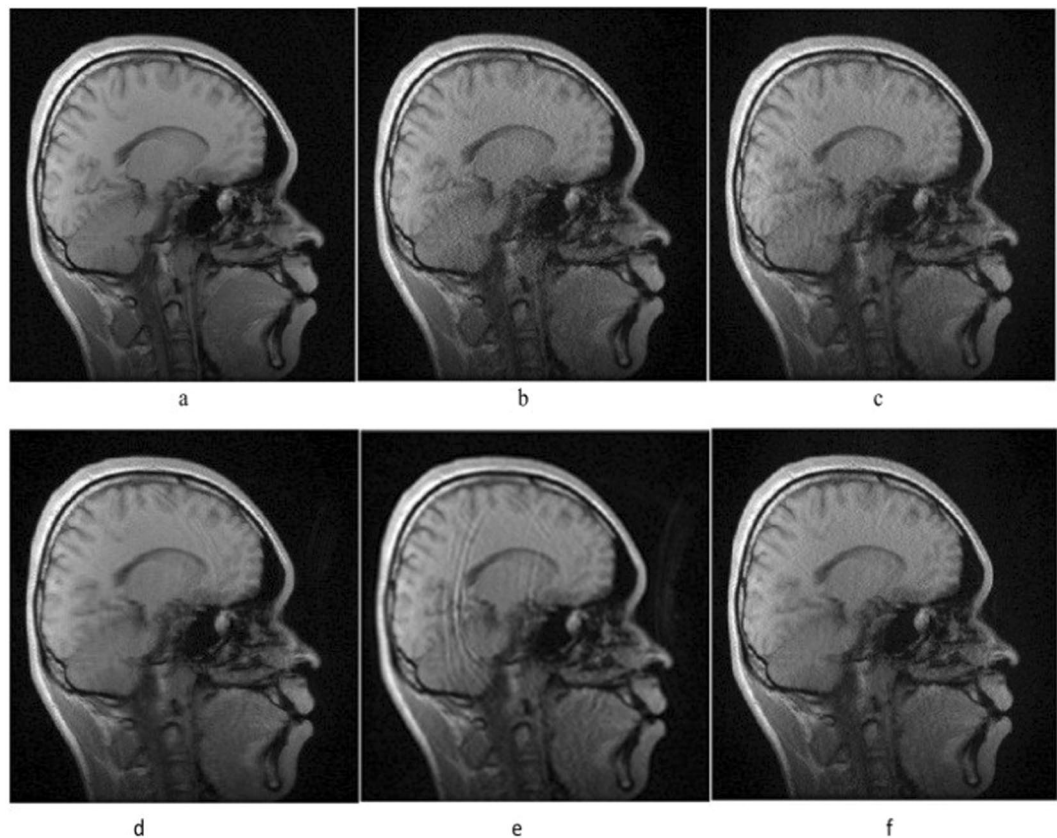
Figure 6. Sagittal brain images reconstructed by IFFT with full *k*-space data (**a**), GRAPPA method (**b**), and PSVR-GRAPPA with linear kernel (**c**), Polynomial kernel (**d**), RBF(**e**), wavelet kernel (**f**).

	$R=3$ ACS $=24$	$R=4$ ACS $=40$	$R=5$ ACS $=60$
GRAPPA	11	41	245
NLGRAPPA	17	103	348
MKGRAPPA	14	50	279

Table 1. Comparison of time consumption between GRAPPA, MKGRAPPA and PSVR-GRAPPA(sec).

with the interpolation window of 4*3, which results from the under-fitting of interpolation weights. Aliasing artifacts and amplified noise are significantly alleviated in the image reconstructed by GRAPPA with the interpolation window of 2*7. The image given by GRAPPA with the interpolation window of 4*15 has less noise at the cost of increased artifacts arising from the over-fitting of interpolation weights due to the over-sized interpolation window. Images reconstructed by PSVR-GRAPPA has no obvious aliasing artifacts and noise. Furthermore, less difference can be found among images reconstructed by PSVR-GRAPPA with such three kinds of interpolation windows. Quantitatively, PSVR-GRAPPA consistently yields smaller NMSEs than GRAPPA.

Figure 5 compares GRAPPA, MKGRAPPA, and PSVR- GRAPPA with an *R* of 4 using the sagittal brain dataset. The images reconstructed by GRAPPA and MKGRAPPA contain noticeable aliasing artifacts, which are significantly alleviated by PSVR-GRAPPA. Quantitatively, the NMSE of the PSVR-GRAPPA reconstruction is 5.32%, which is lower than those of GRAPPA (7.18%) and MKGRAPPA (6.44%) reconstructions.

Table 1 lists the time consumption of three methods *in vivo* studies. MKGRAPPA has the largest time consumption owing to the semi-infinite linear programming. Nevertheless, the proposed PSVR-GRAPPA is slightly slower than GRAPPA.

Figure 6 shows images reconstructed by different kernels. Figure 6a is the reference image reconstructed with full *k*-space data, other result images are reconstructed by GRAPPA method (Fig. 6b), and PSVR-GRAPPA with linear kernel (Fig. 6c), polynomial kernel (Fig. 6d), RBF kernel (Fig. 6e), wavelet kernel (Fig. 6f), respectively. The image reconstructed by linear kernel has similar noise level as GRAPPA reconstructed image, the noise is significantly reduced in the image resulted by polynomial kernel at the cost of blurs and artifacts, and the artifacts is severe in the image of RBF kernel reconstructed image. The noise is suppressed in the image reconstructed by wavelet kernel without visible artifacts, which looks closer to the reference image (Fig. 6a). In theory, wavelet kernel can approximate arbitrary functions that can give better outcomes than other kernels[20].

Conclusions

This study introduces a novel GRAPPA reconstruction method based on proximal support vector regression. The proposed PSVR-GRAPPA method has the advantage of adaptively balancing under- fitting and over-fitting of interpolation weights without significant increase in computational load. The *in vivo* experiments have demonstrated the proposed approach can significantly reduce the noise and aliasing artifacts in GRAPPA reconstruction and the reconstruction performance is insensitive to the interpolation window configuration. Compared with MKGRAPPA, although PSVR-GRAPPA gives implicit improvement on image quality, the latter has less computational load than the former.

References

1. Pruessmann, K. P., Weiger, M., Scheidegger, M. B. & Boesiger, P. SENSE: sensitivity encoding for fast MRI. *Magn. Reson. Med.* **42**, 952–962 (1999).
2. Sodickson, D. K. & Manning, W. J. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn. Reson. Med.* **38**, 591–603 (1997).
3. Jakob, P. M., Griswold, M. A., Edelman, R. R. & Sodickson, D. K. AUTO-SMASH: a self-calibrating technique for SMASH imaging. SiMultaneous Acquisition of Spatial Harmonics. *MAGMA* **7**, 42–54 (1998).
4. Heidemann, R. M., Griswold, M. A., Haase, A. & Jakob, P. M. VD-AUTO-SMASH imaging. *Magn. Reson. Med.* **45**, 1066–1074, https://doi.org/10.1002/mrm.1141 (2001).
5. Bydder, M., Larkman, D. J. & Hajnal, J. V. Generalized SMASH imaging. *Magn. Reson. Med.* **47**, 160–170, https://doi.org/10.1002/mrm.10044 (2002).
6. Griswold, M. A. *et al.* Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* **47**, 1202–1210, https://doi.org/10.1002/mrm.10171 (2002).
7. Ying, L. & Sheng, J. Joint image reconstruction and sensitivity estimation in SENSE (JSENSE). *Magn. Reson. Med.* **57**, 1196–1202, https://doi.org/10.1002/mrm.21245 (2007).
8. Blaimer, M. *et al.* SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method. *Top. Magn. Reson. Imaging* **15**, 223–236, doi: 00002142-200408000-00002 (2004).
9. Qu, P., Shen, G. X., Wang, C., Wu, B. & Yuan, J. Tailored utilization of acquired k-space points for GRAPPA reconstruction. *J. Magn. Reson.* **174**, 60–67 (2005).
10. Samsonov, A. A. On optimality of parallel MRI reconstruction in k-space. *Magn. Reson. Med.* **59**, 156–164, https://doi.org/10.1002/mrm.21466 (2008).
11. Nana, R., Zhao, T., Heberlein, K., LaConte, S. M. & Hu, X. Cross-validation-based kernel support selection for improved GRAPPA reconstruction. *Magn. Reson. Med.* **59**, 819–825, https://doi.org/10.1002/mrm.21535 (2008).
12. Nana, R. & Hu, X. Data consistency criterion for selecting parameters for k-space-based reconstruction in parallel imaging. *Magn. Reson. Imaging* **28**, 119–128, https://doi.org/10.1016/j.mri.2009.05.047 (2010).
13. Chang, Y., Liang, D. & Ying, L. Nonlinear GRAPPA: A kernel approach to parallel MRI reconstruction. *Magn. Reson. Med.* **68**, 730–740, https://doi.org/10.1002/mrm.23279 (2012).
14. Lyu, J., Zhou, Y., Nakarmi, U., Shi, C. & Ying, L. In *International Society of Magnetic Resonance in Medicine Scientific Meeting (ISMRM)*, pp. 2130, Toronto, ON, Canada (2015).
15. Xu, L., Feng, Y., Liu, X., Kang, L. & Chen, W. Robust GRAPPA reconstruction using sparse multi-kernel learning with least squares support vector regression. *Magn Reson Imaging* **32**, 91–101 (2014).
16. Vapnik, V. N. *Statistical Learning Theory*. (Wiley, 1998).
17. Glenn, F. & Mangasarian, O. L. In *Provost and Srikant, editors,Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 77–86 (ACM, 2001a).
18. IvorW, T., James, T. K. & Cheung, P.-M. In *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (AISTAT'05)*.
19. Cherkassky, V. & Ma, Y. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw.* **17**, 113–126, https://doi.org/10.1016/s0893-6080(03)00169-2 (2004).
20. Li Zhang, W. Z. & Jiao, L. Wavelet Support Vector Machine. *IEEE Transactions On Systems, Man, And Cybernetics—PART B: Cybernetics* **34**, 34–39 (2004).

Acknowledgements

Author Contributions

Lin Xu and Tao Jiang designed experiments; Lin Xu and Qian Zheng carried out experiments and analyzed experimental results. Lin Xu, Qian Zheng and Tao Jiang wrote the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.