Published in final edited form as:

Magn Reson Med. 2018 November ; 80(5): 1871–1881. doi:10.1002/mrm.27192.

Machine Learning RF Shimming: Prediction by Iteratively Projected Ridge Regression

Julianna D Ianni^{1,2}, Zhipeng Cao^{1,2}, and William A Grissom^{1,2,3,4}

¹Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, TN, United States

²Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, United States

³Department of Radiology, Vanderbilt University, Nashville, TN, United States

⁴Department of Electrical Engineering, Vanderbilt University, Nashville, TN, United States

Abstract

Purpose—To obviate online slice-by-slice RF shim optimization and reduce B_1^+ mapping requirements for patient-specific RF shimming in high-field MRI.

Theory and Methods—RF Shim Prediction by Iteratively Projected Ridge Regression (PIPRR) predicts patient-specific, SAR-efficient RF shims with a machine learning approach that merges learning with training shim design. To evaluate it, a set of B_1^+ maps was simulated for 100 human heads for a 24-element coil at 7T. Features were derived from tissue masks and the DC Fourier coefficients of the coils' B_1^+ maps in each slice, which were used for kernelized ridge regression prediction of SAR-efficient RF shim weights. Predicted shims were compared to directly designed shims, circularly polarized mode, and nearest-neighbor shims predicted using the same features.

Results—PIPRR predictions had 87% and 13% lower B_1^+ coefficients of variation compared to circularly polarized mode and nearest-neighbor shims, respectively, and achieved homogeneity and SAR similar to that of directly designed shims. Predictions were calculated in 4.92 ms on average.

Conclusion—PIPRR predicted uniform, SAR-efficient RF shims, and could save a large amount of B_1^+ mapping and computation time in RF-shimmed ultra-high field MRI.

Keywords

RF shimming; inhomogeneity correction; RF prediction; supervised learning; machine learning; tailored RF

Introduction

Despite the advantages afforded by a high signal-to-noise ratio, magnetic resonance imaging (MRI) at ultra-high field strengths (7 Tesla (T) and higher) has faced several barriers to widespread adoption. Among these is the increasing spatial inhomogeneity and subject dependence of transmit radiofrequency (RF) (B_1^+) fields with increasing field strength, resulting from an RF wavelength on the order of the size of the imaged object [1, 2]. The resulting signal non-uniformity creates spatially-varying contrast, can obscure underlying pathology, and complicates quantitative imaging [3, 4].

Patient-tailored RF shimming with multiple transmit channels and coils is a highly flexible approach to mitigating B_1^+ inhomogeneity, in which each coil in an array is driven with a unique amplitude and phase to achieve a more uniform combined field across a slice or volume [5, 6]. The amplitudes and phases can be tuned for each subject and scan geometry. However, this currently requires the B_1^+ fields of each coil to be measured in each subject.

Furthermore, since only the amplitude of the combined B_1^+ field needs to be uniform, optimization of the weights requires solution of a non-convex magnitude-least squares (MLS) problem [7]. Several advanced optimization approaches have been proposed to robustly solve this problem [8], but these require considerable computation, and a globally optimal solution cannot be guaranteed in a short computation time, especially for a large number of coils and slices. Furthermore, the fastest methods for full B_1^+ mapping currently require approximately 16 s per coil to scan a whole-brain volume at 3–4 mm isotropic resolution [9]. Mapping the 24 coils used in this work would require 6.5 minutes of scanning, after which RF shim calculation would require approximately 30 seconds per slice. For a whole-brain volume then, slice-by-slice mapping and RF shim calculation requires 10–15 minutes while the patient lies in the scanner, or approximately 25% of a total scan duration.

As an alternative to patient-specific B_1^+ mapping and RF pulse design, Gras et al. [10, 11] have proposed the concept of “universal” multidimensional parallel transmission pulses, which are jointly optimized over a large set of patient brains such that a single pre-optimized pulse can be applied to any adult brain. However, this approach trades homogeneity for broad applicability, since by nature solutions cannot be both universal *and* tailored. Mirfin et al. explored using a neural network for tailored spokes pulse prediction [12], but this approach was not successful in interpolating over the space of MLS problem solutions, which can have very different phase profiles. This resulted in an inability to produce homogeneous flip angle profiles.

In this work, a machine learning method is presented and tested in simulation for the instantaneous prediction of patient-tailored, SAR-efficient RF shims. The method, called RF Shim Prediction by Iteratively Projected Ridge Regression (PIPRR), avoids the weakness of Mirfin et al.’s approach by merging the design of the training shims with learning. This makes it possible to interpolate over the training shims using kernelized ridge regression.

Simulation results show that PIPRR can predict SAR-efficient RF shim weights with negligible online compute times and minimal B_1^+ map data (only the DC Fourier coefficients of the B_1^+ maps were used here). In the following, we describe the PIPRR method, characterize it in terms of the amount of training data required, robustness to noise, and required features, and validate it against directly designed shims as well as shims predicted by nearest-neighbor using the same features. Preliminary reports of this work were presented in Refs. [13–15].

Theory

Magnitude Least-Squares RF Shimming

Patient-tailored RF shimming is typically posed as a magnitude least squares optimization problem [7, 16–20]:

$$\hat{\mathbf{b}} = \underset{\mathbf{b}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{m} - |\mathbf{A}\mathbf{b}| \|^2_{\mathbf{W}} + \frac{\lambda}{2} \mathbf{R}(\mathbf{b}), \quad (1)$$

$$= \underset{\mathbf{b}, \phi}{\operatorname{argmin}} \frac{1}{2} \left\| \operatorname{diag} \left(e^{i\phi_i} \right) \mathbf{m} - \mathbf{A}\mathbf{b} \right\|_{\mathbf{W}}^2 + \frac{\lambda}{2} \mathbf{R}(\mathbf{b}), \quad (2)$$

where the vector $\mathbf{b}$ contains the complex-valued transmit RF weights for each coil, $\mathbf{m}$ is a vector containing the desired excitation pattern at all spatial locations ($\mathbf{m}$ is typically a vector of ones), ϕ is the desired target phase pattern which replaces the absolute value around $\mathbf{A}\mathbf{b}$ and is jointly optimized with the RF weights, i indexes spatial locations, $\mathbf{A}$ is a matrix containing the B_1^+ maps for each coil at each spatial location, which are concatenated in the column dimension, the diagonal matrix $\mathbf{W}$ contains spatially-dependent weights that select samples within a tissue mask, λ is a regularization parameter, and $\mathbf{R}$ is a quadratic regularization function that can be used, e.g., to regularize SAR or RF power. This work focuses on slice-by-slice RF shimming, in which a unique shim vector $\mathbf{b}$ is determined for each slice in an imaged volume. A typical iterative approach to solve Equation 2 alternates between updating the RF shim weights $\mathbf{b}$ while holding the target phase ϕ fixed, and updating the target phase ϕ for fixed shim weights $\mathbf{b}$ by setting it equal to the phase of $\mathbf{A}\mathbf{b}$. The shim weights can be updated using a standard solver for regularized least-squares problems, such as regularized pseudoinverse or the conjugate gradients (CG) algorithm.

Kernelized Ridge Regression Prediction of RF Shims

Instead of designing RF shims by solving the problem in Equation 2, we will apply kernelized ridge regression (KRR) to predict the RF weights from a length- N_{feat} complex vector of slice-specific features $\mathbf{f}$, which includes a bias entry. The RF shim weight $\hat{b}_j$ for coil j predicted by KRR can be written as:

$$\hat{b}_j = \mathbf{p}_j^T \mathbf{f}, \quad (3)$$

where the length- N_{feat} complex-valued feature weight vector $\mathbf{p}_j$ relates each feature in $\mathbf{f}$ to coil j 's shim weight. The entries of $\mathbf{f}$ are normalized by the mean and standard deviation of each feature across the training slices. The method is kernelized because we include first order cross-products of features in $\mathbf{f}$, as will be described later. KRR learns the feature weights contained in the $\mathbf{p}_j$ vectors by fitting a regularized linear model to a set of training shim weights for each coil, as:

$$\hat{\mathbf{p}}_j = \underset{\mathbf{p}_j}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{b}_{j,train} - \mathbf{F} \mathbf{p}_j\|^2 + \frac{\beta}{2} \|\mathbf{p}_j\|^2, \quad (4)$$

where the length- N_{train} vector $\mathbf{b}_{j,train}$ contains coil j 's shim weights for all the training slices, $\mathbf{F}$ is an $N_{train} \times N_{feat}$ matrix of features for each training slice, and β is the ridge regression regularization parameter. Each column of $\mathbf{F}$ is normalized to have zero mean and standard deviation one. Given this model, one might directly use a set of training shims that are solutions to Equation 2 to solve Equation 4 for the feature weights. However, as will be shown later, this results in very poor shims due to large variations in shimmed phase between MLS problem solutions. In the following, we describe the RF Shim PIPRR algorithm, which solves this problem by merging the design of a set of training shims for KRR with learning the coil-specific feature weights.

RF Shim PIPRR Algorithm

In order to derive a set of training shims that can be fitted with low errors to each slice's features using KRR, we propose to merge the design of the training shim weights for KRR (the $\mathbf{b}_{j,train}$ vectors) with learning the coil-specific feature weights (the $\mathbf{p}_j$ vectors). This is realized by alternately taking a few iterations towards the solution to Equation 2, and projecting the current training shim weights $\hat{\mathbf{b}}_{j,train}$ onto the space of shim weights that are predictable by KRR, as:

$$\hat{\mathbf{b}}_{j,train} \leftarrow \mathbf{F}(\mathbf{F}^H \mathbf{F} + \beta \mathbf{I})^{-1} \mathbf{F}^H \hat{\mathbf{b}}_{j,train}, \quad (5)$$

for each coil, where H denotes Hermitian transpose. Note that the regularized projector matrix $\mathbf{F}(\mathbf{F}^H \mathbf{F} + \beta \mathbf{I})^{-1} \mathbf{F}^H$ does not change and is calculated once, before the iterations start. The steps of the training algorithm are summarized as follows:

PIPRR: Training shim design and weight learning

- 1: Set target phase to circularly polarized (CP) mode phase for each training slice.
- 2: **repeat**

PIPRR: Training shim design and weight learning

- 3: Update training shims to minimize Equation 2 for each training slice individually, using a few CG iterations (Equation 2).
 - 4: Project training shims onto the set that is predictable by KRR (Equation 5).
 - 5: Update target phase to that produced by current shims for each training slice.
 - 6: **until** change in cost (Equation 2 summed across training slices) < tolerance
 - 7: Calculate final feature weights for each coil by solving Equation 4, using regularized pseudo-inverse.
-

Overall, this procedure corresponds to an alternating minimization over a non-convex set (the set of solutions to Equation 2) and a convex set (the vector space spanned by the feature projector matrix) [21].

Methods

Electromagnetic Simulations and Features

An in silico dataset of B_1^+ maps was generated by electromagnetic simulation of a 24-element 7T loop coil array with shielding in 100 human head models, using XFDTD (Remcom Inc., State College, PA, USA). The coils were driven by current sources and simulated in isolation. They were arranged in 3 rows of 8, as shown in Figure 1. The total height of the array was 20.5 cm, and the diameter was 30 cm. The subjects were generated by magnifying 50 Ella and 50 Duke head phantoms from the Virtual Family [22] in three dimensions, where the sizes in each dimension were drawn at random from population-based normal distributions [23]. Magnification factors ranged from 0.93x to 1.10x, 0.87x to 1.08x, and 0.92x to 1.21x in the left-right, anteroposterior, and craniocaudal dimensions, respectively. This corresponds to head sizes ranging from 13.5–16.5 cm left-right and 17.7 to 21.9 cm anterior-posterior. Measured from the base of the cerebellum, the head sizes ranged from 12 to 17 cm in the superior-inferior dimension. Figures 2a and b show the range of independent scaling factors of the original Duke and Ella models in each dimension. The resolution of the simulated maps was 5 mm isotropic, with no slice gap. The dataset contained 31 axial slices for each subject, centered in the middle of the coil; slices containing fewer than 20 voxels were discarded. Local 10-gram SAR matrices were calculated for each head using the adaptive spherical mask-based method of Ref. [?], and virtual observation point (VOP) SAR matrices were calculated from them using a parameter u equal to 20% of the maximum spectral norm [24]. Each head had between 122 and 303 VOP matrices. To define a regularization term in Equation 2 which discouraged shim solutions with high global or local SAR, the VOP matrices were summed with a global SAR matrix which was scaled up by a factor of 3.1 to reflect a $3.1\times$ tighter limit on global head SAR (max 3.2 W/kg) versus local SAR (max 10 W/kg). With this construction, the regularization term had the form: $R(\mathbf{b}) \triangleq \mathbf{b}^H \mathbf{C} \mathbf{b}$, where $\mathbf{C}$ is the sum of the VOPs and the scaled global SAR matrices. For comparison with RF Shim PIPRR, directly-designed shims were computed for every slice of every subject by solving Equation 2 using 100 random target phase initializations, which were obtained by generating 100 random sets of shim weights. The result that minimized the cost function of Equation 2 was taken as the ‘Direct Design’ (DD) shim for each slice.

A vector of 560 features was calculated for each simulated brain slice. This included the DC Fourier coefficient of each coil's B_1^+ map (24 features), the mean x and y coordinates of voxels in the tissue mask (i.e., the mask centroid's coordinates) (2 features), standard deviations of the within-mask x and y coordinates (2 features), slice z-position (1 feature), the central 3 Fourier shape descriptors [25] of the slice mask (3 features), and all 1st-order cross-terms of these features (528 features). A bias term was also included, and the masks were the same as those contained in the $\mathbf{W}$ matrices for shim design. Fourier shape descriptors were computed as the Fourier transform of the ordered set of complex coordinates representing the outermost outline of the brain mask [25].

Algorithm Implementation

PIPRR was implemented in MATLAB R2015a (The Mathworks, Natick, MA, USA) using Vander-bilt University's Advanced Computing Center for Research and Education (ACCRES) cluster to parallelize computation across folds (described below). The RF shim weight updates used 3 CG iterations, and the SAR regularization parameter λ was 0.01. λ was set based on DD shims to the maximum power of ten that did not increase the MLS cost by more than 1% on average. The target excitation pattern vectors $\mathbf{m}$ were set to one inside the tissue masks for each slice. The PIPRR stopping criterion and the KRR regularization parameter β were set as described below. Code for PIPRR and the 100 subjects' data can be downloaded from <https://github.com/wgrissom/piprr>.

Experiments

Comparison to other RF shim designs—To evaluate the PIPRR algorithm, a 10-fold random cross-validation was performed across all the simulated heads. Each fold comprised a unique division of the heads into training and test sets, with 90 of the heads used for training and 10 used for testing. Slices from the same head were not split between the training and test sets. Shims were predicted for slices in the test sets using the learned feature weights from the training sets. One fold of data was used to determine the KRR regularization parameter β and a stopping tolerance for PIPRR, to prevent overfitting in the other folds. PIPRR was run for several values of β in increments of powers of 10, and the value of β that reached the smallest test shim cost (the sum of the cost in Equation 2 across all test slices) over the PIPRR iterations ($\beta = 1$) was used for all other runs. The tolerance was set equal to the maximum of 10^{-6} and the consecutive difference in training shim costs (the sum of the cost in Equation 2 across all training slices) immediately before the $\beta = 1$ run's test shim costs started increasing. That fold was then omitted from all results, and the other folds used the determined stopping tolerance. PIPRR shims were compared to the coil array's CP mode, directly designed shims, and shims obtained by nearest-neighbor (NN), in terms of shimmed B_1^+ coefficients of variation (CoV; B_1^+ standard deviation divided by mean B_1^+) and the calculated SAR regularization terms $\mathbf{R}(\mathbf{b})$. The CP mode was established by driving each coil with a phase shift according to its angular position. The NN shim predictions used the same features as PIPRR, and were generated by directly selecting the shim from the training set whose features were closest to a test slice's features, in terms of mean-squared differences.

Required training data and features, and noise sensitivity—Four additional experiments were performed to characterize the PIPRR algorithm, in terms of required training data, sensitivity to noise, importance of different feature classes, and the amount of B_1^+ mapping data required. First, to characterize the amount of training data required by the algorithm, PIPRR training was repeated using 10 to 90 randomly-selected training heads from one fold, and the shimmed test B_1^+ CoVs and SAR penalties were calculated for each number of training heads. The effect of additive noise in the test features was studied by adding independent Gaussian noise over a range of signal-to-noise ratios (SNRs) to the test B_1^+ maps in each fold, calculating the corresponding SNRs of the B_1^+ map DC Fourier coefficients, and then adding noise to the remaining features to achieve the same SNRs. A total of 10 B_1^+ map SNR levels between 10 and 100 were simulated, and the shimmed B_1^+ CoVs were calculated for each level. The importance of each feature class was determined by calculating the norm of the set of weights assigned to each feature in PIPRR's final KRR weight learning step, while varying the regularization parameter β over several orders of magnitude. The weight norms were combined across folds, and the feature classes were ranked according to their norms. The effect of each feature class on the quality of test shim predictions was then calculated by dropping all feature classes from the model and then individually re-incorporating them in their order of importance. Finally, the effect of further reducing the required B_1^+ mapping data for PIPRR was characterized by leaving out coils' B_1^+ DC Fourier coefficients from the feature sets for each fold, but still predicting shims for all 24 coils. The number of retained coil coefficients ranged from 24 down to zero, with approximately equidistant spacing of retained coils.

Results

Comparison to other RF shim designs

Figure 3 shows three-plane shimmed B_1^+ patterns for all the slices in a representative Duke head model; additional slices from this volume are shown in Figure S1. PIPRR shims were similar in shape and error level to the DD shims, while NN shims contained larger errors and significant discontinuities between some adjacent axial slices, as indicated by the arrows. To illustrate the overall quality of the shims and the kinds of errors that arose in each case, Figure 4 shows the best-, median-, and worst-case shimmed B_1^+ patterns across all slices and folds in terms of B_1^+ coefficient of variation (CoV) for CP mode, NN, DD, and PIPRR training and test shims. Figure 5a shows box plots of B_1^+ CoVs for each method, across all test slices. Mean $|B_1^+|$ CoVs were 0.213 for CP mode, 0.021 for DD shims, 0.041 for NN shims, 0.389 for KRR shims, and 0.031 for PIPRR-predicted training and test shims. Thus, DD shims were the most uniform, followed by the PIPRR shims. There is also little difference between the PIPRR training shims and the DD shims, indicating that the projection step did not significantly degrade the quality of the training shims. NN produced relatively uniform shims on average, but also produced outliers with very poor shims, as illustrated by the worst case NN pattern in Figure 4. Conventional KRR applied to the DD

shims failed because of the large variation in phase between the training shims, resulting in high CoVs and holes in the B_1^+ patterns in Figure 4. Figure 5b shows the SAR penalty terms for each test slice, normalized to the mean of the DD SAR penalties. The mean SAR penalty term, normalized to that of the DD shims, was 8.78 a.u. for CP mode, 0.98 a.u. for NN, 1.40 a.u. for KRR, 0.89 a.u. for PIPRR training shims and 0.88 a.u. for PIPRR test shims. The PIPRR SAR penalty values are similar to or lower than the DD values, and are lowest among the predicted shims. Supporting Figure S2 plots separated global and maximum local SAR values for all test slices (normalized to the mean of the DD global and maximum local SARs), which reflect similar relationships between the different shimming methods. Supporting Figures S3 and S4 further show whole-head local 10-gram SAR maps of DD, NN and PIPRR shims, for Duke and Ella model slices for which the three designs had similar B_1^+ CoV's. The time required to train PIPRR was 5 days (without parallelization across training slices), but the average time to predict the shims for a test slice was 4.92 ms (4.86 ms for feature calculation and 0.06 ms for prediction via multiplication with the feature weights), on a 2.8 GHz desktop PC. In comparison, computing the DD shims took approximately 30 seconds per slice on the same PC.

Required training data and features, and noise sensitivity

Figure 6 shows box plots of one fold's test set B_1^+ CoVs when PIPRR is run using a variable number of randomly-chosen heads from the training set. A drop-off in inhomogeneous outliers and 75th percentiles occurs between 50 and 60 heads. Supporting Figure S5 shows that the SAR penalties varied in a similar manner with the number of training heads. Similar drop-offs in inhomogeneity and SAR outliers were observed for the other folds, which indicates that 50–60 heads were required to sufficiently sample each dimension of the head size distributions, without completely missing some dimensions of head size scaling. Inhomogeneity and SAR of predicted shims were comparable when 60 heads were used for training versus the full 90-head training set. The medians and 25th percentiles did not significantly change after 20 heads. Figure 7 shows that PIPRR predictions tolerate a moderate amount of noise in the features without significant degradation in homogeneity of the shim predictions. Box plots of test B_1^+ CoVs are shown with increasing feature SNR, which is reported as a function of the equivalent B_1^+ map SNR. Linearly increasing feature SNR results in approximately exponentially decreasing CoVs. A feature SNR of 40 or above produces approximately the same level of homogeneity in the predictions as noise-free features, at 3.4% and 3.1%, respectively. Supporting Figure S6 shows that the SAR penalties were approximately flat across SNR levels. Figure 8 shows the test set B_1^+ CoVs versus the feature classes included in the KRR model, beginning with the feature class of highest importance, which is the product of the mask centroids and the B_1^+ map DC Fourier coefficients, and adding in the other feature classes in order of importance, which was determined as described in the Methods. The CoVs level off after five feature classes are included, and the most important classes involve products of B_1^+ map DC coefficients and tissue mask features. Supporting Figure S7 shows that the SAR penalties were

approximately flat as additional feature classes were included. Figure 9 shows how the homogeneity of 24-coil PIPRR shims depends on the number of coils whose B_1^+ map DC coefficients are included as features in the prediction. Average B_1^+ CoVs increased approximately linearly with a reduced number of coils, but remained less than 0.06 even with no coils. Supporting Figure S8 shows that the SAR penalties were approximately constant as the number of coils was decreased.

Discussion

Summary and implications of results

Computational experiments demonstrated that RF shim PIPRR can predict SAR-efficient tailored RF shims with homogeneous B_1^+ patterns in 100 simulated heads with population-representative dimensions. The predicted shims were consistently more homogeneous than CP mode and NN-predicted shims. They had 50% higher B_1^+ CoVs on average than DD shims (0.031 versus 0.021) but still very low CoVs overall, and 12% lower average SAR penalties. Compared to DD shims, which require full B_1^+ maps of all coils in a slice, RF shim prediction with PIPRR required only a tissue mask and the DC coefficients of the coils' B_1^+ maps in a slice. A tissue mask can be derived from a fast low-flip angle gradient-recalled echo scan, which is commonly performed as a prescan step for B_0 shimming or receive sensitivity mapping. The B_1^+ map DC coefficients could be measured with a single excitation for each coil and slice, which would be between 10–100× faster than full B_1^+ mapping since no in-slice phase encoding would be required. It was further shown that inhomogeneity degraded slowly as the number of B_1^+ map coefficients was reduced, indicating that coils could be skipped altogether. Additionally, the B_1^+ inhomogeneity of PIPRR-predicted shims rose slowly as the SNR of the features was decreased, indicating that it is not highly sensitive to noise. The computational burden of making a set of RF shim predictions with PIPRR is also much lower than for direct design of RF shims (approximately 5 ms to both compute the features and predict the shims, versus approximately 30 seconds to directly design the shims). Overall, the PIPRR method could alleviate the prescan and computational burdens of slice-specific RF shimming for high-field MRI. While its impact on the RF shimming workflow would be greatest for a large number of transmit channels, we expect that it would apply equally well to the 8-channel transmit systems that are widely installed on ultra-high field scanners today.

Importance of iterative training

PIPRR is fairly unique among machine learning techniques for its iterative re-design of its own training data. The need for this was demonstrated by the very poor performance of shims predicted by conventional KRR when it was trained on DD shims, as was shown in Figures 4 and 5. Even the best KRR slice had a null in the field pattern, and the shims were overall less homogeneous than CP mode. The training error of KRR (Equation 4) was over 200× higher than for PIPRR, averaged over folds. This resulted from the fact that the MLS

shim design problem is non-convex with many local minima and infinite solutions (since any phase-shifted solution is also a solution), which makes it very difficult for any machine learning method to learn the relationships between the features and the shims, without a very high order model that could require a prohibitive amount of data to train. PIPRR overcomes this problem by regularizing the training shims so that they can be more easily fit by KRR. Early in this work, we investigated whether MLS shim solutions in a training set could be made more similar by removing the average phase shifts between them prior to learning the KRR weights, but this did not significantly improve the KRR training error or predictions (results not shown). An alternative approach may be to regularize differences in target phase patterns or RF shims between slices and subjects, but this would make the MLS problem more complicated and may overly constrain the designs without significantly improving KRR training error. It is possible to use other machine learning methods in the same iterative framework introduced by PIPRR; KRR was chosen for this work because it is a simple, fast learning method that predicted shims of similar quality to DD shims. With non-linear feature transforms (which were shown to be the most important feature classes), it was able to accurately relate features to shims and model the variations in shims between slices and subjects. More sophisticated learning methods, such as random forests [26] or neural networks [27], could also be used within this framework, but would in most cases require more training data and would require significantly more computation than KRR, which may be prohibitive for iterative training.

Extensions and future work

This study used simulations to validate and characterize the PIPRR algorithm in 100 human heads, using a 24-element transmit coil array at 7T. The next step in development is to implement and validate it in vivo. Training based on numerical simulations of a coil array should be compared to training based on B_1^+ map scans of a population of subjects with the same transmit coil. Numerical simulation-based training would be more convenient since the learner could be more easily adapted to changes in the coil or its electromagnetic environment. The method should also be extended to predict shims for slices with arbitrary orientations (coronal, sagittal, and oblique). An in vivo validation will further inherently consider several realistic factors that were beyond the scope of this initial work, such as variations in coil matching and coupling between subjects, variable tissue distributions within the head, variable head centering and angulation inside the coil array, and differences between approaches for measuring B_1^+ map coefficients and defining the tissue masks.

Addressing these real-world issues may require different or additional feature sets; for example, predicting shims for arbitrary orientations may require the addition of slice angles as features.

Another important consideration for human implementation is SAR. As shown in Figure 5b, PIPRR predictions had somewhat lower SAR penalty values than DD shims. In this work, the training shims were regularized for low SAR but it was not explicitly constrained, since the shims were not designed for a specific pulse sequence. While it is possible to implement strict constraints in the training shim design for a specific flip angle and TR, it is not yet clear how to constrain the predictions to also meet these constraints, without dilating RF

pulses or increasing TR. One approach might be to predict shims as a weighted average of SAR-constrained shims, where the weights collectively have norm one. This could be achieved using, e.g., a softmax function at the output of a neural network. Among existing methods for local SAR-regularized pulse design [??], the SAR regularization approach used here is most related to the method of Sbrizzi et al [?]. That method is based on iteratively-reweighted least-squares, where a local SAR-regularized pulse design is repeated, and between repetitions the regularization parameters applied to the local SAR penalties are updated based on whether they are much larger or smaller than a reference SAR penalty value. This has the effect of reducing and homogenizing local SAR values while discouraging large outliers. The approach taken in this work can be viewed as a first iteration of the Sbrizzi method, before reweighting is applied. It would be straightforward to incorporate Sbrizzi's reweighting scheme in the DD and PIPRR algorithms, which should improve the tradeoff between maximum local SAR and B_1^+ CoV.

PIPRR may be amenable to use outside the brain, in the extremities and abdomen. These applications may be challenged by the relatively larger variations in body sizes and shapes compared to the head, more variable body positioning with respect to the coil array, and variable positioning of the coil array itself, if it is not rigidly attached to the bore or patient bed. These factors may necessitate more sophisticated models that include additional features (such as non-DC B_1^+ Fourier coefficients) and higher-order feature maps (such as squared features, in addition to the feature cross-terms used here), or more sophisticated learning approaches (such as neural networks).

Finally, PIPRR could be extended to predict not just RF shims, but full RF waveforms. For example, spokes [28, 29] and k_T -points [30] pulses are effectively trains of RF shims, so it should be possible to predict each shim in the train. It may also be possible to predict parameterized gradient waveforms for such pulses, such as the gradient moments between the subpulses. Direct design of these pulses requires considerably more computation than RF shimming, especially for large-tip-angles, so the ability to predict them rapidly would have a greater impact on the parallel transmission workflow. More finely sampled waveforms such as spirals could also be predicted, though it is not yet clear how the target pattern could be flexibly resized or positioned in the imaged volume.

Conclusions

This simulation study showed that the RF shim PIPRR method predicts SAR-efficient tailored slice-specific RF shims with low B_1^+ inhomogeneity. Because it does not require full B_1^+ mapping, it would save considerable prescan time compared to directly designed RF shims, and computation of predicted shims requires only feature calculation and a single matrix multiplication. The PIPRR method could therefore improve the parallel transmission workflow by alleviating the prescan and computational burdens of slice-specific RF shimming for high-field MRI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grants R01 EB016695, R01 DA019912, and R25 CA136440. It was conducted in part using the resources of the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University, Nashville, TN.

References

1. Van de Moortele PF, Akgun C, Adriany G, Moeller S, Ritter J, Collins CM, Smith MB, Vaughan JT, Ugurbil K. B_1 destructive interferences and spatial phase patterns at 7 T with a head transceiver array coil. *Magn Reson Med*. 2005; 54:1503–1518. [PubMed: 16270333]
2. Yang QX, Wang J, Zhang X, Collins CM, Smith MB, Liu H, Zhu XH, Vaughan JT, Ugurbil K, Chen W. Analysis of wave behavior in lossy dielectric samples at high field. *Magn Reson Med*. 2002; 47:982–989. [PubMed: 11979578]
3. Grissom WA. Improving high-field MRI using parallel excitation. *Imaging Med*. 2010; 2:675–693.
4. Vaughan JT, Garwood M, Collins CM, Liu W, DelaBarre L, Adriany G, Andersen P, Merkle H, Goebel R, Smith MB, Ugurbil K. 7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images. *Magn Reson Med*. 2001; 46:24–30. [PubMed: 11443707]
5. Ibrahim TS, Lee R, Baertlein BA, Abduljalil AM, Zhu H, Robitaille PML. Effect of RF coil excitation on field inhomogeneity at ultra high fields: A field optimized TEM resonator. *Magn Reson Imaging*. 2001; 19:1339–47. [PubMed: 11804762]
6. Mao W, Smith MB, Collins CM. Exploring the limits of RF shimming for high-field MRI of the human head. *Magn Reson Med*. 2006; 56:918–922. [PubMed: 16958070]
7. Setsompop K, Wald LL, Alagappan V, Gagoski BA, Adalsteinsson E. Magnitude least squares optimization for parallel radio frequency excitation design demonstrated at 7 Tesla with eight channels. *Magn Reson Med*. 2008; 59:908–915. [PubMed: 18383281]
8. Hoyos-Idrobo A, Weiss P, Massire A, Amadon A, Boulant N. Variant strategies to solve the magnitude least squares optimization problem in parallel transmission pulse design and under strict SAR and power constraints. *IEEE Trans Med Imaging*. 2014; 33:739–48. [PubMed: 24595346]
9. Nehrke K, Börner P. DREAM - A novel approach for robust, ultrafast, multislice B_1 mapping. *Magn Reson Med*. 2012; 68:1517–26. [PubMed: 22252850]
10. Gras V, Vignaud A, Amadon A, Le Bihan D, Boulant N. Universal pulses: A new concept for calibration-free parallel transmission. *Magn Reson Med*. 2017; 77:635–643. [PubMed: 26888654]
11. Gras V, Boland M, Vignaud A, Ferrand G, Amadon A, Mauconduit F, Le Bihan D, Stocker T, Boulant N. Homogeneous non-selective and slice-selective parallel-transmit excitations at 7 Tesla with universal pulses: A validation study on two commercial RF coils. *PLoS One*. 2017; 12:e0183562. [PubMed: 28827835]
12. Mirfin C, Glover P, Bowtell R. Optimisation of parallel transmission radiofrequency pulses using neural networks. *Proc 2nd UK7T Ultra High Field Network Symposium*; 2017;
13. Ianni JD, Cao Z, Grissom WA. RF shim design by supervised learning. *Proc Intl Soc Magn Reson Med Ultra High Field Workshop*; 2016; 12
14. Ianni JD, Cao Z, Grissom WA. Predicting multi-coil RF shims via machine learning. *Proc i2i Workshop*, New York University; 2016. 6
15. Ianni JD, Cao Z, Grissom WA. SAR-efficient RF shim prediction via machine learning. *Proc Intl Soc Magn Reson Med*. 2017; 25:3992.
16. Kassakian PW. PhD dissertation. University of California; Berkeley, CA: 2006. Convex Approximation and Optimization with Applications in Magnitude Filter Design and Radiation Pattern Synthesis.
17. Katscher U, Vernickel P, Graesslin I, Börner P. RF shimming using a multi-element transmit system in phantom and in vivo studies. *Proc Intl Soc Mag Reson Med*. 2007; 15:1693.

18. Kerr AB, Zhu Y, Pauly JM. Phase Constraint Relaxation in Parallel Excitation Pulse Design 1. *Proc Intl Soc Magn Reson Med*. 2007; 15:1694.
19. Wu X, Schmitter S, Auerbach EJ, Moeller S, Urbil K, Van de Moortele PF. Simultaneous multislice multiband parallel radiofrequency excitation with independent slice-specific transmit B1 homogenization. *Magn Reson Med*. 2013; 70:630–638. [PubMed: 23801410]
20. Wu X, Schmitter S, Auerbach EJ, Urbil K, Van de Moortele PF. A generalized slab-wise framework for parallel transmit multiband RF pulse design. *Magn Reson Med*. 2016; 75:1444–1456. [PubMed: 25994797]
21. Lewis AS, Malick J. Alternating projections on manifolds. *Math Oper Res*. 2008; 33:216–234.
22. Christ A, Kainz W, Hahn EG, Honegger K, Zefferer M, Neufeld E, Rascher W, Janka R, Bautz W, Chen J, Kiefer B, Schmitt P, Hollenbach HP, Shen J, Oberle M, Szczerba D, Kam A, Guag JW, Kuster N. The Virtual Family—development of surface-based anatomical models of two adults and two children for dosimetric simulations. *Phys Med Biol*. 2010; 55:N23–N38. [PubMed: 20019402]
23. Young JW. Technical report. US Department of Transportation Federal Aviation Administration, Civil Aeromedical Institute; 1993. Head and face anthropometry of adult U.S. civilians.
24. Eichfelder G, Gebhardt M. Local specific absorption rate control for parallel transmission by virtual observation points. *Magn Reson Med*. 2011; 66:1468–1476. [PubMed: 21604294]
25. Gonzalez RC, Woods RE. *Digital Image Processing*. 3. Pearson; 2007.
26. Breiman L. Random forests. *Machine Learning*. 2001; 45:5–32.
27. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. MIT Press; 2016.
28. Zhang Z, Yip CY, Grissom W, Noll DC, Boada FE, Stenger VA. Reduction of transmitter B1 inhomogeneity with transmit SENSE slice-select pulses. *Magn Reson Med*. 2007; 57:842–7. [PubMed: 17457863]
29. Grissom WA, Khalighi MM, Sacolick LI, Rutt BK, Vogel MW. Small-tip-angle spokes pulse design using interleaved greedy and local optimization methods. *Magn Reson Med*. 2012; 68:1553–62. [PubMed: 22392822]
30. Cloos MA, Boulant N, Luong M, Ferrand G, Giacomini E, Le Bihan D, Amadon A. k T-points: Short three-dimensional tailored RF pulses for flip-angle homogenization over an extended volume. *Magn Reson Med*. 2012; 67:72–80. [PubMed: 21590724]

Figure 1.

The 24 channel loop coil, simulated in XFDTD to obtain B_1^+ maps and SAR matrices. The loops were arranged in 3 rows of 8 elements each. The total height of the array was 20.5 cm, and its diameter was 30 cm. One element is highlighted in orange.

Figure 2.

a) Tissue masks for the center transverse slice in five subjects, to demonstrate the variation in head sizes across all subjects. Maximum and minimum head widths and lengths are shown, as well as the median size (middle). b) Central sagittal tissue masks for the subjects with maximum, median, and minimum head height. The numbers next to the names indicate the amplification factor applied to the original Duke or Ella model in the corresponding dimension.

Figure 3.

B_1^+ patterns produced for all the axial slices of a representative Duke head model by circularly polarized (CP) mode, DD shims, NN shims, and PIPRR shims. The displayed slices are the middle slices in the axial, coronal, and sagittal planes. B_1^+ CoVs over all slices are reported for each method, as well as minimum and maximum B_1^+ . Black arrows indicate regions where the Nearest-Neighbor shims had significant B_1^+ discontinuities between adjacent axial slices. Supporting Information Figure S1 shows additional slices.

Figure 4.

Shimmed B_1^+ patterns for the best-, median-, and worst-case (in terms of shimmed B_1^+ coefficient of variation) slices across all test slices, for circularly polarized (CP) mode, DD, Nearest-Neighbor (NN), kernelized ridge regression (KRR) applied to the DD shims, and PIPRR (PIPRR Test). PIPRR training shim patterns are also shown.

Figure 5.

a) B_1^+ pattern inhomogeneity for all test slices for circularly polarized (CP) mode, DD, NN, kernelized ridge regression (KRR) applied to the DD shims, and PIPRR (PIPRR Test). PIPRR training shims are also shown. b) The corresponding SAR penalty terms for all test slices. The values are normalized to the mean SAR penalty of the DD shims. Blue box edges delineate the 25th and 75th percentiles, and medians are indicated by the red bars. Red crosses indicate outliers (values that exceeded the 75th percentile by greater than $1.5 \times$ the difference between the 75th and 25th percentiles). The black whiskers indicate the extent of data not considered outliers.

Figure 6.

Shimmed B_1^+ CoVs of one fold's test set slices, when varying the number of heads included in PIPRR training shim design and learning. The homogeneity of the predicted shim profiles is comparable to those predicted with the full 90-head training set when at least 60 heads are included in the weight learning.

Figure 7.

Shimmed B_1^+ CoVs of all test set slices, with noise of varied amplitude added to the features used for PIPRR prediction. Noise level is reported in terms of equivalent B_1^+ map SNR. The no-noise case is indicated by an SNR of ∞ .

Figure 8.

Feature importance. The box plot shows B_1^+ CoVs of all test set slices, as feature groups are accrued into the final KRR weight learning and testing, in order of importance. The number of features in each class is reported in parentheses next to the class label. Importance was measured as the norm of the weights on each feature class over a range of regularization parameters (β). Cross-terms of mask centroids, B_1^+ map DC Fourier coefficients, Fourier shape descriptors (FSDs), and slice position were the most important features.

Figure 9.

PIPRR test performance with incomplete sets of B_1^+ map DC coefficients. a) Shimmed B_1^+ CoVs of all test set slices, versus the number of coils whose DC Fourier coefficients were included in the PIPRR predictions for all 24 coils. b) Shimmed B_1^+ patterns for the best-, median-, and worst-case predicted test slices when 6 coils' coefficients were included, and when zero coils' coefficients were included (i.e., only the tissue mask size, shape, and position features were used for prediction).