



Harmonizing Diffusion MRI Data Across Magnetic Field Strengths

Suheyyla Cetin Karayumak^(✉), Marek Kubicki, and Yogesh Rathi

Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
{skarayumak, kubicki, yogesh}@bwh.harvard.edu

Abstract. Diffusion MRI (dMRI) data is increasingly being acquired on multiple scanners as part of large multi-center neuroimaging studies. However, diffusion imaging is particularly sensitive to scanner-specific differences in coil sensitivity, reconstruction algorithms, acquisition parameters as well as the scanner magnetic field strength, which precludes joint analysis of such multi-site data. Earlier works on dMRI data harmonization were limited to data acquired on different scanners but with the same magnetic field strength (3T). In this work, we explore the possibility of harmonizing dMRI data acquired on scanners with different magnetic field strengths, i.e., 3T and 7T. We propose a linear and several machine learning based non-linear mapping algorithms that use rotation invariant spherical harmonic (RISH) features to map the dMRI data (the raw signal) between scanners without changing the fiber orientations. We extensively validate our algorithms on *in-vivo* data from the Human Connectome Project (HCP) where we used data from 40 subjects with scans done on both 7T and 3T scanners (10 training + 30 test). Using several quantitative metrics such as the root mean squared error (RMSE) in the harmonized dMRI signal and diffusion measures as well as a fiber bundle overlap measure, our preliminary results on 30 test subjects shows that the convolutional neural network (CNN) based algorithm can reliably harmonize the raw dMRI signal across magnetic field strengths. The algorithms proposed are general and can be used for dMRI data harmonization in multi-site studies.

Keywords: Harmonization · Diffusion MRI · Machine learning (CNN)

1 Introduction

In recent years, several large-scale multi-site neuroimaging studies have been initiated to collect MRI data pertaining to neurodevelopment as well as disease [13, 14] to increase statistical power. However, directly pooling dMRI data

The authors would like to acknowledge the following grants which supported this work: R01MH102377 (PI: Dr. Marek Kubicki), R01MH097979 (PI: Dr. Yogesh Rathi).

acquired from multiple scanners is fraught with problems due to significant differences in dMRI measures of the same subjects scanned on different scanners [15]. On the other hand, better scanner technologies as well as higher field strength scanners are becoming more popular as they provide better contrast and resolution in diffusion-weighted (DW) imaging [3, 11]. For instance, data from a 7T scanner reveals details of tissue properties not visible at 3T [16]. However, data from scanners with different field strengths need to be harmonized to be used jointly.

Recently, several methods have been proposed to harmonize multi-site data, or boost the resolution and quality of dMRI data. Mirzaalian et al. [8] provide a framework for multi-site (3T) harmonization of a single shell (single b-value) dMRI data with similar acquisition parameters (b-values, number of gradients, spatial resolution) and magnetic field strength using rotation invariant spherical harmonics (RISH) features. In [9], the authors use a correction factor for each diffusion tensor derived measure (fractional anisotropy (FA), mean diffusivity (MD)) within a region, while the work in [5] uses a location specific statistical adjustment factor to account for scanner differences. Both these methods perform data harmonization on the final model derived measures (e.g. DTI measures) and not the dMRI signal itself. Consequently, data harmonization has to be done several times independently for each measure, unlike the model-independent method proposed in this work. We also note that, a few works [1, 12] have proposed an image quality transfer method which utilizes nonlinear regression to estimate high resolution DTIs or higher order model parameters. These model specific methods however have not been used in the context of data harmonization, but are potential candidates. Consequently, we compare our methods with the work of [1]. Furthermore, the harmonization between multiple field strengths remains unaddressed.

In this work, we harmonize multi-shell (multiple b-values) dMRI data by predicting 7T-like diffusion MRI signal from 3T data by mapping their corresponding RISH features. In particular, we propose to learn an efficient mapping of multi-shell dMRI signal with different spatial resolution and magnetic field strength (3T and 7T) using deep Convolutional Neural Networks (CNN). We investigate and propose two methods: voxel-wise linear mapping, and patch-based non-linear mapping using deep Convolutional Neural Networks (CNN) which are explained in the following sections.

2 Method

2.1 RISH Features

We represent the dMRI signal \mathbf{S} in a basis of spherical harmonics (SH): $\mathbf{S} \approx \sum_l \sum_m C_{lm} Y_{lm}$, and construct the rotation invariant spherical harmonic (RISH) features which can be appropriately scaled to modify the dMRI signal without changing the principal diffusion directions of the fibers. Thus, our goal is to estimate a voxel-wise linear or a patch-based non-linear mapping of the RISH

features between 3T and 7T data from the same set of subjects, which can then be used on test subjects to validate the quality of the mapping.

The following processing was common for both linear and non-linear methods in Sect. 2.2. Due to differences in spatial resolution between 3T and 7T dMRI data, we first upsample each DW volume using a 7th-order B-spline which was shown to perform better than other interpolation schemes [4]. Next, we use a recently proposed unringing method [6] to remove Gibbs ringing artifact from each DW volume. Five RISH feature maps $C_l^{b,s}(\mathbf{x}; i)$ for each b-value shell with SH orders of $l = \{0, 2, 4, 6, 8\}$ are computed at each voxel location $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ for each scanner s as follows:

$$C_l^{b,s}(\mathbf{x}; t) = \sum_{m=1}^{2l+1} C_{lm}(\mathbf{x})^2, \quad (1)$$

where t is the subject number and $b = \{1000, 2000\}$ is the b-value.

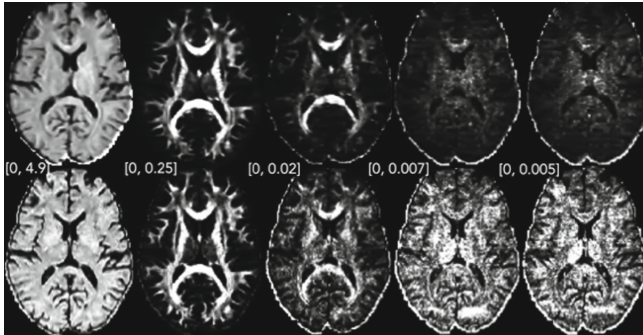


Fig. 1. RISH Features of $b = 1000$ shell for SH orders of $l = \{0, 2, 4, 6, 8\}$ are depicted in each sub-figure from left to right for 3T (top row) and 7T scans (bottom row) for HCP Subject ID: 102311 ([\cdot , \cdot]: the range of intensities chosen for visualization).

Figure 1 shows the RISH features of the same HCP subject scanned on 3T (top) and 7T (bottom) scanner for $b = 1000$. Each RISH feature captures a different aspect (frequency content) of the diffusion signal. Note the significantly increased energy (contrast) in higher order RISH features in 7T data, that is not quite visible in the 3T data.

2.2 Learning the Mapping of RISH Features from 3T to 7T

Voxel-Wise Linear Mapping: Using 3T RISH features as input, our goal is to learn the voxel-wise linear mapping of 3T to 7T. To achieve this, first, the RISH features in the training set are used to create multi-modal RISH feature templates (antsMultiVariateTemplateConstruction [2]). Once the template space

is constructed separately for each shell, we define the expected value of the voxel-wise RISH features as the sample mean $\mathbb{E}_l^{b,s}(\mathbf{x}') \approx \sum_{t=1}^{N_s} \mathbf{C}_l^{b,s}(\mathbf{x}'; t) / N_s$ over the number of training subjects N_s , where s is 3T or 7T scanner and \mathbf{x}' is the voxel location in the template space. Next, we compute the voxel-wise linear (scaling only) maps between RISH features of 3T and 7T data in the template space

using: $\mathcal{S}_l(\mathbf{x}') = \sqrt{\frac{\mathbb{E}_l^{b,7T}(\mathbf{x}')}{\mathbb{E}_l^{b,3T}(\mathbf{x}') + \epsilon}}$. We apply this linear map learned from the training

data set to new subjects from the test data, by non-rigid transformation of scale maps to the subject space. The 7T-like dMRI signal is estimated by scaling the SH coefficients of the signal at each voxel in the subject space as follows: $\hat{\mathbf{C}}_{lm}(\mathbf{x}) = \hat{\mathcal{S}}_l(\mathbf{x}) C_{lm}(\mathbf{x})$, where $\hat{\mathcal{S}}_l(\mathbf{x})$ is the scale map in the subject space and $\hat{\mathbf{C}}_{lm}(\mathbf{x})$ is the scaled SH coefficients. The final diffusion signal is then computed using:

$$\hat{\mathbf{S}}(\mathbf{x}) = \sum_l \sum_m \hat{\mathbf{C}}_{lm}(\mathbf{x}) Y_{lm}. \quad (2)$$

Patch-Based Non-linear Mapping Using Deep CNN: Using 3T RISH features as input, our goal is to learn a nonlinear mapping of 3T to 7T as a patch-wise regression problem. Such mapping can be learned using the paired 3T and 7T RISH features of training data. We first align 3T and 7T data as follows: First, we register b0 maps of 3T and 7T data through rigid registration [2]. The estimated transformation is then applied to each DW volume. Next, the gradient vectors are rotated using the rotation matrix estimated through rigid registration. After 3T and 7T DW data are aligned, we compute RISH features as in Eq. 1. To learn the mapping from 3T to 7T, we construct our deep CNN with five convolutional layers. Specifically, we used an 9×9 RISH feature patch to learn the mapping.

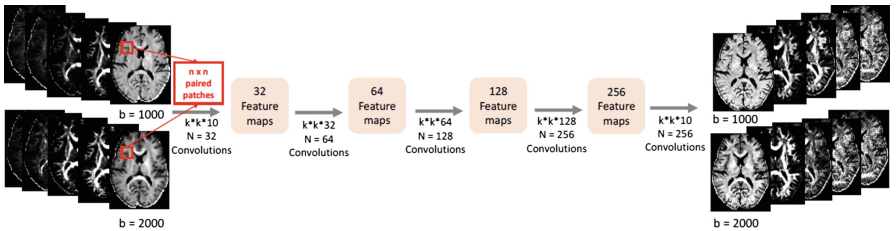


Fig. 2. Our proposed deep CNN architecture for learning the mapping from 3T to 7T RISH features for multi-shell data. We used $k = 3$ as the filter size with five layers.

Figure 2 summarizes our deep CNN architecture. In the first layer, the aim is to learn a feature representation of the input 3T RISH feature patches. It includes convolution filters with size of 32 followed by ReLU activation function. In each layer, RISH features are convolved with a 3×3 kernel with 32, 64, 128,

Table 1. Average RMSE (percentage) in FA, MD, GFA and DWI signal in 30 test subjects using different 3T to 7T mapping techniques for each shell separately.

RMSE: %	Original 3T-7T	Linear	RF-DTI	RF-RISH	CNN-RISH
$b = 1000$					
FA	36.49 ± 2.22	6.94 ± 0.019	6.23 ± 0.023	6.01 ± 0.011	5.41 ± 0.010
MD	40.32 ± 2.67	4.02 ± 0.012	3.44 ± 0.021	3.01 ± 0.008	2.33 ± 0.013
GFA	38.53 ± 2.13	7.11 ± 0.021	-	6.55 ± 0.015	5.95 ± 0.011
DWI	39.33 ± 2.45	7.32 ± 0.023	-	6.61 ± 0.012	5.94 ± 0.010
$b = 2000$					
FA	34.32 ± 2.10	5.97 ± 0.013	5.86 ± 0.019	5.51 ± 0.012	4.95 ± 0.008
MD	35.44 ± 2.14	3.23 ± 0.012	2.97 ± 0.017	2.76 ± 0.014	2.12 ± 0.011
GFA	36.66 ± 2.77	6.53 ± 0.014	-	6.03 ± 0.011	5.23 ± 0.010
DWI	39.43 ± 2.21	7.08 ± 0.021	-	6.51 ± 0.015	5.42 ± 0.012

Table 2. Overlap (percentage) in various fiber bundles: CST, CB, AF and IOFF traced in the original 3T and 7T, and the harmonized 7T-like data in 30 test subjects.

Fiber bundle: %	7T vs 3T	7T-like vs 7T	7T-like vs 3T
CST	93.59 ± 3.36	92.98 ± 3.28	93.90 ± 5.35
CB	95.35 ± 3.03	94.22 ± 5.50	96.60 ± 2.26
AF	96.53 ± 1.48	96.80 ± 1.52	97.61 ± 1.32
IOFF	94.96 ± 2.42	96.35 ± 2.21	95.60 ± 2.81

256 and 256 convolutional filters. After each convolution step, ReLU operation is applied. In training, we used ADAM optimizer with a learning rate 10^{-4} and epoch size is selected as 100.

3 Results

We used 10 HCP subjects [13] as training subjects with dMRI scans obtained from both 7T and 3T scanners. Another independent set of unseen 30 HCP subjects (with data from both 3T and 7T) were used to evaluate the performance of all the methods. 7T data had the following acquisition parameters: 1.05 mm isotropic spatial resolution, two-shells ($b = 1000, 2000$) with 65 gradient directions on each shell; while 3T data had: 1.25 mm isotropic spatial resolution, three-shells ($b = 1000, 2000, 3000$) with 90 gradient directions on each shell. In this work, we learnt the mapping only for $b = 1000$ and $b = 2000$ shells from 3T to 7T. We compared our methods with another non-linear learning method: Regression Forest (RF) method which was presented by [1] to improve DTI data quality (RF-DTI). Note that, this method was not used in the context of data harmonization, yet we found it relevant to compare our work with it. In this

paper, we also introduce RF-RISH (Regression Forest (RF) with RISH features) to provide a fair comparison between our RISH feature based and RF-DTI based method. Using our methods, subject-specific mapping between 3T and 7T was obtained and the final signal was estimated using Eq. 2.

We computed whole brain FA and MD to compare the learning performance between the RISH features based methods and RF-DTI [1]. Root Mean Squared Error (RMSE) on 30 test subjects was computed for DTI specific measures of FA and MD as well non-model specific measure of generalized FA (GFA) and the dMRI signal itself. RMSE was computed between our prediction and the ground truth data that was acquired on 7T from the same set of subjects. Average accuracy and precision values for estimation of FA, MD, GFA and DWI signal are given in Table 1. In Fig. 3(a) top row, we show the estimated FA results using our methods (Linear-RISH, RF-RISH and CNN-RISH) and RF-DTI for $b = 1000$. In Fig. 3(a) bottom row, we show the error maps (RMSE) in FA between the predicted data and the actual scanner acquired 7T data. Figure 3(b) shows error maps in the raw dMRI signal, with most of the error using CNN-RISH confined to the CSF regions of the brain. Even though FA and MD are directly derived from DTI model, RISH features based non-linear methods performed better when compared to RF-DTI. As seen in Table 1 and Fig. 3, our deep CNN-RISH method gives the best performance, with lowest error in several metrics (FA, GFA and dMRI signal error). Thus, our method is tissue model-independent and directly reconstructs the dMRI signal, which can then be used in further analysis.

In order to ensure that our deep CNN method does not change the fiber orientation, we performed whole brain tractography using a multi-tensor unscented Kalman filter (UKF) method [7]. Next, we use the White Matter Query Language (WMQL) [17] to extract specific anatomical white matter bundles from the whole brain tracts. Figure 4 depicts WMQL results for corticospinal tract (CST) and cingulum bundle (CB). After extracting the tracts from the original 3T and 7T, and the harmonized 7T-like data, we used the Bhattacharyya overlap distance (B) to quantify the agreement between the tracts [10]: $B = \frac{1}{3} \left(\int \sqrt{P_h(x)P(x)}dx + \int \sqrt{P_h(y)P(y)}dy + \int \sqrt{P_h(z)P(z)}dz \right)$, where $P(\cdot)$ represents the ground truth probability distribution of the fiber bundle, $P_h(\cdot)$ is the probability distribution of the tracts from the harmonized data and $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ are the fiber coordinates. B is 1 for a perfect match between two fiber bundles and 0 for no overlap at all. In Table 2, we provide the Bhattacharyya overlap measure for: (i) the original 7T vs the original 3T; (ii) the estimated 7T-like vs the original 7T; (iii) the estimated 7T-like vs the original 3T data. Due to the space limitations, we only show the results for CST, CB, arcuate fasciculus (AF) and the inferior occipito-frontal fascicle (IOFF) tracts. We observed very high overlap of 93–97% for all fiber bundles indicating that fiber orientation is preserved by the harmonization algorithm.

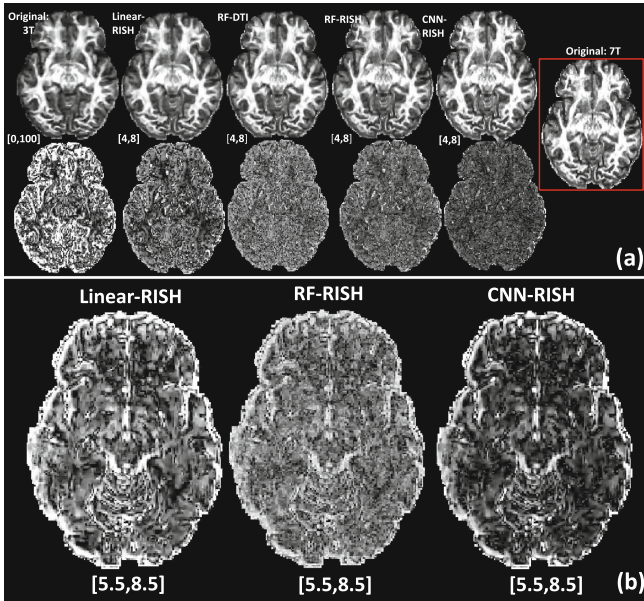


Fig. 3. (a) FA comparison between our methods (Linear-RISH, RF-RISH and CNN-RISH) and RF-DTI for $b = 1000$ are shown in the top row. FA map of the original 7T data is depicted in the rightmost figure. RMSE maps in FA for different methods as well as the original 3T data itself (leftmost) are depicted in the bottom row. All results are computed and shown for the same subject; (b) RMSE maps between DW volumes estimated using our methods (Linear-RISH, RF-RISH and CNN-RISH) and the original 7T data for a single subject. $[\cdot, \cdot]$: the range of intensities chosen for visualization).

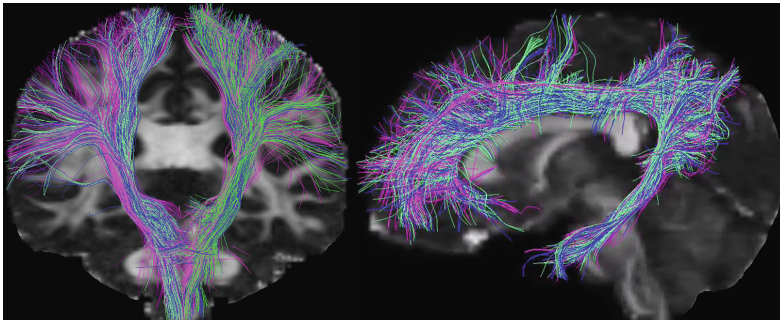


Fig. 4. Significant ($>93\%$) overlap is seen in CST and CB extracted from the original 3T (blue) and 7T (magenta), and harmonized 7T-like (green) dMRI data.

4 Conclusion

In this paper, we proposed a linear and a nonlinear machine learning method to harmonize the raw dMRI data acquired on scanners with very different magnetic field strengths (3T and 7T). We validated our algorithm on 30 test subjects, and demonstrated the efficacy of using this technique to harmonize dMRI data from vastly different scanners in a model-free manner. Even though FA and MD are directly related DTI, both qualitative and quantitative results show that our methods perform better or close (for linear regression) when compared to RF-DTI. The tractography results also prove that our deep CNN method matches the dMRI signal between scanners while preserving the fiber orientations. The proposed method can also be useful to improve the quality and resolution of dMRI data. As a first step, we have demonstrated and validated the utility of this work in harmonizing data from healthy subjects while more validation needs to be done on subjects with gross tissue pathology.

References

1. Alexander, D.C., et al.: Image quality transfer and applications in diffusion MRI. *NeuroImage* **152**(Suppl. C), 283–298 (2017)
2. Avants, B.B., et al.: The optimal template effect in hippocampus studies of diseased populations. *NeuroImage* **49**(3), 2457–2466 (2010)
3. Chilla, G.S., Tan, C.H., Xu, C., Poh, C.L.: Diffusion weighted magnetic resonance imaging and its recent trend—a survey. *Quant. Imaging Med. Surg.* **5**(3), 407 (2015)
4. Dyrby, T.B.: Interpolation of diffusion weighted imaging datasets. *NeuroImage* **103**(Suppl. C), 202–213 (2014)
5. Fortin, J.P., et al.: Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* **161**, 149–170 (2017)
6. Kellner, E., Dhital, B., Kiselev, V., Reiser, M.: Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn. Reson. Med.* **76**(5), 1574–1581 (2015)
7. Malcolm, J.G., Shenton, M.E., Rath, Y.: Filtered multitensor tractography. *IEEE Trans. Med. Imaging* **29**(9), 1664–1675 (2010)
8. Mirzaalian, H., et al.: Multi-site harmonization of diffusion MRI data in a registration framework. *Brain Imaging Behav.* **12**, 1–12 (2016)
9. Pohl, K.M., et al.: Harmonizing DTI measurements across scanners to examine the development of white matter microstructure in 803 adolescents of the NCANDA study. *Neuroimage* **130**, 194–213 (2016)
10. Rath, Y., Gagoski, B., Setsompop, K., Michailovich, O., Grant, P.E., Westin, C.-F.: Diffusion propagator estimation from sparse measurements in a tractography framework. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *MICCAI 2013*. LNCS, vol. 8151, pp. 510–517. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40760-4_64
11. Sotiropoulos, S.N., et al.: Fusion in diffusion MRI for improved fibre orientation estimation: an application to the 3T and 7T data of the human connectome project. *NeuroImage* **134**, 396–409 (2016)
12. Tanno, R., et al.: Bayesian image quality transfer with CNNs: exploring uncertainty in dMRI super-resolution. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) *MICCAI 2017*. LNCS, vol. 10433, pp. 611–619. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66182-7_70

13. Van Essen, D.C., et al.: The WU-Minn human connectome project: an overview. *Neuroimage* **80**, 62–79 (2013)
14. Volkow, N.D., et al.: The conception of the ABCD study: from substance use to a broad NIH collaboration. *Dev. Cogn. Neurosci.* **32**, 4–7 (2017). ISSN 1878-9293
15. Vollmar, C., et al.: Identical, but not the same: Intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0 T scanners. *NeuroImage* **51**, 1384–1394 (2010)
16. Vu, A.T., et al.: High resolution whole brain diffusion imaging at 7T for the human connectome project. *Neuroimage* **122**, 318–331 (2015)
17. Wassermann, D., et al.: The white matter query language: a novel approach for describing human white matter anatomy. *Brain Struct. Funct.* **221**(9), 4705–4721 (2016)