Non-Local Spatial Regularization of MRI T_2 Relaxation Images for Myelin Water Quantification

Youngjin Yoo[1,3] and Roger Tam[2,3]

[1] Department of Electrical and Computer Engineering
[2] Department of Radiology
[3] MS/MRI Research Group,
University of British Columbia, Vancouver, BC, Canada

Abstract. Myelin is an essential component of nerve fibers and monitoring its health is important for studying diseases that attack myelin, such as multiple sclerosis (MS). The amount of water trapped within myelin, which is a surrogate for myelin content and integrity, can be measured *in vivo* using MRI relaxation techniques that acquire a series of images at multiple echo times to produce a T_2 decay curve at each voxel. These curves are then analyzed, most commonly using non-negative least squares (NNLS) fitting, to produce T_2 distributions from which water measurements are made. NNLS is unstable with respect to the noise and variations found in typical T_2 relaxation images, making some form of regularization inevitable. The current methods of NNLS regularization for measuring myelin water have two key limitations: 1) they use strictly local neighborhood information to regularize each voxel, which limits their effectiveness for very noisy images, and 2) the neighbors of each voxel contribute to its regularization equally, which can over-smooth fine details. To overcome these limitations, we propose a new regularization algorithm in which local and non-local information is gathered and used adaptively for each voxel. Our results demonstrate that the proposed method provides more globally consistent myelin water measurements yet preserves fine structures. Our experiment with real patient data also shows that the algorithm improves the ability to distinguish two sample groups, one of MS patients and the other of healthy subjects.

Keywords: T_2 relaxation, quantitative MRI, brain, white matter, myelin, regularization.

1 Introduction

Myelin water fraction (MWF) is an imaging surrogate biomarker for myelin, and is currently the only myelin-specific quantitative MRI measure. Central nervous system (CNS) tissue is inhomogeneous and has two main water environments, one within the myelin, the other being intracellular and extracellular water. MWF is defined as the ratio of myelin water over the total amount of water.

K. Mori et al. (Eds.): MICCAI 2013, Part I, LNCS 8149, pp. 614–621, 2013.

Most MWF studies are done only in white matter (WM), which contains most of the myelin. Normal appearing white matter (NAWM) in MS patients has lower MWF than WM from control subjects, suggesting that myelin loss is a key feature of NAWM pathology in MS. Although multi-echo T_2 relaxation with the non-negative least squares (NNLS) algorithm used to fit a T_2 decay curve with a T_2 distribution [1] is a strong candidate for monitoring demyelination and remyelination in MS patients, one of its current main drawbacks is limited reproducibility due to the instability of the NNLS algorithm. This causes the technique to work well only on data of high SNR; unfortunately, this is often not a practical assumption as T_2 relaxation data tend to be noisy. To improve the robustness of MWF computation, applying various denoising methods to the multi-echo images have been studied [2,3,4,5]. Regularization approaches for NNLS have been also studied, ranging from the conventional Tikhonov regularization [6,7], which is non-spatial, to smoothing methods that work by averaging the neighboring spectra to constrain the NNLS of a given voxel [8]. While the spatial regularization methods have been shown to improve MWF reproducibility, the current approaches only use a small local neighborhood for each voxel, and do not take advantage of more distant neighbors that may have similar T_2 distributions. In addition, these methods assume that neighboring voxels always have similar T_2 distributions, which is not the case in areas with fine details, such as the WM structures near the cortex of the brain, and such details can be lost during regularization.

We present a spatial regularization method that reduces global variability by using information from non-local voxels but also improves fine details by using weighted averaging that avoids combining T_2 distributions that are very different. The algorithm starts with an initial T_2 distribution estimate using the conventional Tikhonov regularized NNLS algorithm. These distributions are then used in a non-local means algorithm [9], using the symmetric Kullback-Leibler (SKL) divergence [10] to quantify the similarity between distributions, to compute the priors for a spatially regularized NNLS, which produces the final T_2 distribution for each voxel. We show that by incorporating non-local means for T_2 distributions into a spatially regularized NNLS framework, an accurate and robust MWF map can be produced. We demonstrate on real MRIs that the method improves the consistency of MWF measurements and the ability to distinguish between MS patients and healthy subjects.

2 Preliminaries

Whittall and MacKay proposed a multi-exponential analysis for MRI [1,7] based on the NNLS algorithm of Lawson and Hanson [11], which inverts a relaxation decay curve to a relaxation time distribution. A set of exponential basis functions characterizes the measured signal y_i

$$y_i = \sum_{j=1}^{M} s_j e^{-t_i/\mathrm{T}_{2j}} = \sum_{j=1}^{M} a_{ij} s_j, \qquad i = 1, 2, \ldots, N, \tag{1}$$

where t_i is the measurement time, M is the number of logarithmically spaced T_2 decay times, N represents the total number of data points (number of acquired T_2 echoes, 32 in our data), s_j is the relative amplitude for each partitioned T_2 time, T_{2j}, and $\mathbf{A} = [a_{i,j}]_{N \times M}$. The set of s_j represents the T_2 distribution for which we need to solve. The NNLS algorithm [11] is used to minimize

$$\chi^2_{\min} = \min_{s \geq 0} \left[\sum_{i=1}^{N} \left| \sum_{j=1}^{M} a_{ij} s_j - y_i \right|^2 \right]. \tag{2}$$

Extra constraints can be incorporated into Eq. 2 to provide more robust fits in the presence of noise [7], using the additional term on the right.

$$\chi^2_r = \min_{s \geq 0} \left[\sum_{i=1}^{N} \left| \sum_{j=1}^{M} a_{ij} s_j - y_i \right|^2 + \mu \sum_{k=1}^{K} \left| \sum_{j=1}^{M} s_j - f_k \right|^2 \right] \tag{3}$$

where $\mathbf{f}$ is the corresponding vector of right-hand side values. With a given target value for the misfit χ^2_r, the parameter μ is automatically determined by the generalized cross-validation approach. From Eq. 3, the T_2 distribution can be estimated as the set of s_j for the decay times T_{2j}. If $\mathbf{f}$ is set to 0, Eq. 3 is equivalent to the conventional Tikhonov regularization, which we abbreviate as rNNLS.

In the closest related work [8] to ours, the T_2 distribution prior $\mathbf{f}$ in Eq. 3 is estimated by averaging the spectra of nine neighboring voxels of each voxel, which achieves spatial regularization but can smooth over fine details.

3 Method

We compute an initial T_2 distribution for each voxel using rNNLS. The regularization parameter μ is selected such that χ^2_r in Eq. 3 is equal to $1.02\chi^2_{\min}$ where $\chi^2_{\min}$ is the unregularized minimum misfit found in Eq. 2. The constant multiplier is set to 1.02 to be consistent with previously reported works using Tikhonov regularization [8,12]. Given initial T_2 distributions $\mathbf{s} = \{\mathbf{s}(i) \mid i \in I\}$ for an image I, a T_2 distribution prior $\mathbf{f}(i)$, for a voxel position i, is computed using the non-local means algorithm [9] as a weighted average of all the T_2 distributions in I,

$$f_k(i) = \sum_{j \in I} w(i,j) s_k(j), \qquad \forall k \in \{1, 2, \ldots, K\}, \tag{4}$$

where the weight $w(i,j)$ depends on the similarity between the distributions at voxel i and j, and satisfies the conditions $0 \leq w \leq 1$ and $\sum_j w(i,j) = 1$. K is set to be the number of logarithmically spaced T_2 decay times. The similarity between the two distributions at i and j are measured from the T_2 distribution

vectors $v(D_i)$ and $v(D_j)$, where D_c denotes a square neighborhood of fixed size D and centered at voxel position c. By treating each T_2 distribution in a voxel as a probability distribution, the similarity between $v(D_i)$ and $v(D_j)$ can be measured by the SKL divergence [10] as follows

$$E(v(D_i), v(D_j)) = \frac{1}{V} \sum_{\substack{\rho \in v(D_i) \\ \varrho \in v(D_j)}} \mathrm{SKL}_{\mathrm{div}}(\mathbf{s}(\rho) \parallel \mathbf{s}(\varrho)) \tag{5}$$

where V is the number of voxels in a square neighborhood, $\mathbf{s}(\rho)$ and $\mathbf{s}(\varrho)$ are normalized versions of the distributions, and the SKL divergence is defined as

$$\mathrm{SKL}_{\mathrm{div}}(\mathbf{s}(\rho) \parallel \mathbf{s}(\varrho)) = \sum_{k=1}^{K} \left(s_k(\rho) \log \left[\frac{s_k(\rho)}{s_k(\varrho)} \right] + s_k(\varrho) \log \left[\frac{s_k(\varrho)}{s_k(\rho)} \right] \right). \tag{6}$$

The weight w used in Eq. 4 is defined as

$$w(i,j) = \frac{1}{\sum_{j \in I} e^{-\left(\frac{E(v(D_i), v(D_j))}{h}\right)^2}} e^{-\left(\frac{E(v(D_i), v(D_j))}{h}\right)^2} \tag{7}$$

where the parameter h controls the degree of filtering.

After computation of the spatial T_2 distribution prior $\mathbf{f}$, the final regularized T_2 distribution is estimated by the NNLS algorithm in Eq. 3. The MWF for each voxel is defined as the percentage of signal having a T_2 between 15 and 50 ms. To vary the degree of regularization, the target value for the misfit χ_r^2 is adjusted relative to χ_{min}^2 in Eq. 2 by a multiplier η, which in turns determines μ. Since the parameters largely depend on noise level, we randomly selected one subject from our data sets for tuning, which should not bias the overall results. We empirically select the parameters η and h to produce the most homogeneous MWF maps in the NAWM of the subject, as measured by the coefficient of variation (CoV). The CoV is generally reduced with higher η and h. We first chose η, starting with $\eta = 1.02$ and increased by 0.01 until the decrease in CoV stabilized. Next we chose h similarly. We use $\eta = 1.07$ and $h = 1.5$ for all of the experiments. For practical purposes, we use a fixed search window of 21×21 and set the neighborhood size for computing the SKL divergence to 7×7, as this has been shown to be effective for non-local denoising of other types of images [9] and structural MRI images [13]. Due to the strong anisotropy of our test data we apply the regularization in 2D on a slice-by-slice basis.

4 Results

In this section, we compare three methods of NNLS regularization for MWF computation: Tikhonov regularization (rNNLS), the spatial regularization presented in [8] (srNNLS), and our proposed non-local spatial regularization (nlsrNNLS). The methods are evaluated under three criteria: the visual quality of the MWF

maps, CoV within the NAWM of healthy subjects, and the ability to distinguish between healthy subjects and MS patients as measured by Student's t-tests and overlapping coefficients (OVLs). All experiments are performed on the NAWM regions of each scan. The T_2 relaxation sequence consists of a 90° excitation pulse followed by 32 slab-selective refocusing pulses (32 echoes, TR = 1200 ms, TE = 10 ms – 320 ms, echo spacing 10 ms), with voxel size = $0.937 \times 0.937 \times 5$mm and image dimensions $256 \times 256 \times 7$ voxels. All images are acquired on a Philips Achieva 3T scanner. To measure CoV, the MRIs of twenty healthy subjects are used. WM segmentation masks are created from registered T_1-weighted images using the FAST software [14]. For the t-tests and OVLs, the scans of 18 healthy subjects and 17 MS patients are used. Each subject was scanned at month 0 (baseline) and month 6. Using a C++ GPU implementation, rNNLS, srNNLS and nlsrNNLS take about 10, 16, and 20 mins respectively to process the data for one patient.

4.1 Visual Assessment

Figure 2 shows a comparison of the visual quality in the MWF maps of a healthy subject and an MS patient. Both srNNLS and nlsrNNLS appear to reduce spatial variability and noise, but nlsrNNLS produces greater spatial coherence overall and is clearly better than both other methods at reconstructing the finer details in the thin WM structures near the cortex. The MS lesions in the MWF map of the MS patient produced by nlsrNNLS are more clearly visible and have more distinct boundaries.

4.2 Coefficient of Variation

In order to compare the variability within the MWF maps produced by each method, we measure the CoV within the WM of each scan of 20 healthy subjects (Figure 1). Both spatial regularization methods significantly improve CoV over rNNLS, with nlsrNNLS being slightly better than srNNLS. The mean CoV values and standard errors are shown in Table 1. There is a statistically significant group difference ($p = 5.12 \times 10^{-5}$) between srNNLS (mean CoV = 1.10) and nlsrNNLS (mean CoV = 0.97).

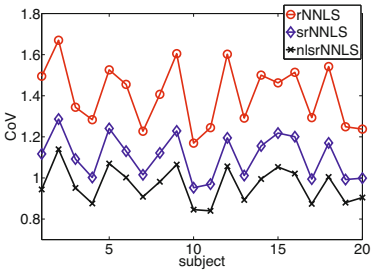

Fig. 1. Comparison of CoV measured on MWF maps of 20 healthy subjects

Table 1. Mean CoV measured on the WM of 20 healthy subjects. The p-value is computed by a two-sample t-test between srNNLS and nlsrNNLS.

	CoV
rNNLS [7]	1.4060±0.1498
srNNLS [8]	1.1047±0.1052
nlsrNNLS	**0.9653±0.0861**
p-value	5.1240×10^{-5}

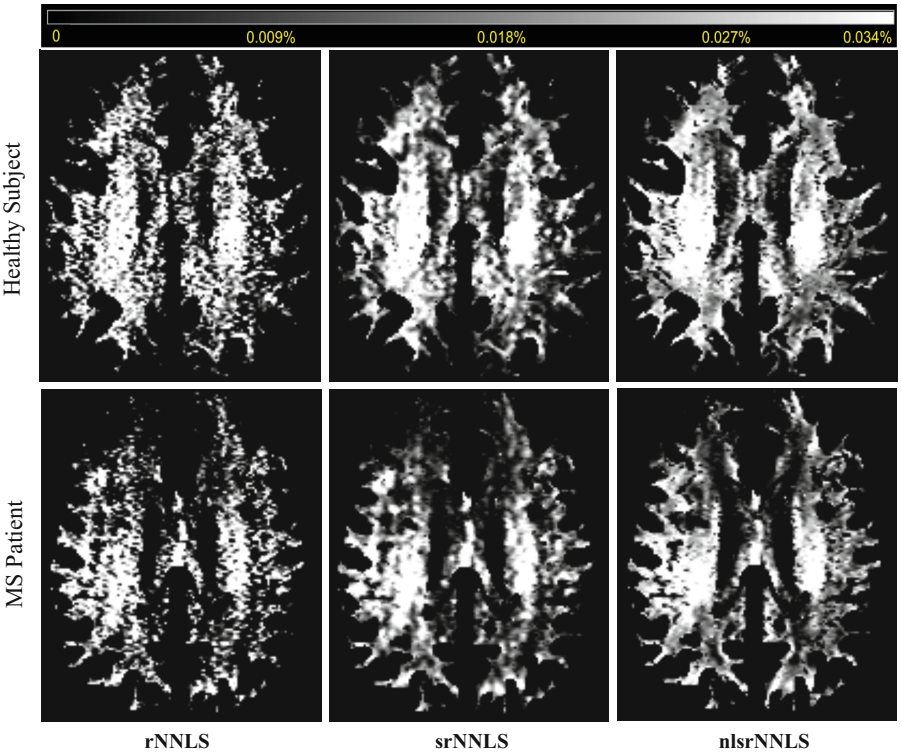

Fig. 2. Comparison of MWF maps produced by rNNLS [7] (left), srNNLS [8] (middle) and nlsrNNLS (right) algorithms in WM for healthy subject (top) and MS patient (bottom). The nlsrNNLS shows an improved visibility of the MWF maps with reduced noise and sharp details preserved.

4.3 Student's *t*-test and Overlapping Coefficient

The data for 18 healthy subjects and 17 MS patients are used to compare the ability of the three regularization methods to distinguish between these two groups. The three mean MWF values computed from each group are shown in Figure 3. The mean MWF values from the MS patients are significantly lower ($p < 0.05$ for all three regularization methods) than those from the healthy subjects at both month 0 and month 6. The p-values computed using t-tests between the MS patients and healthy subjects are shown in Table 2. The nlsrNNLS method yields the lowest p-values in this data set, which suggests it can produce MWF maps that are more sensitive to the differences between MS and healthy brains. To confirm this result, we construct normal distributions for each group and compute the OVL as a second measure of how well the patients and normals can be distinguished. The results from this test for all three regularization methods are summarized in Table 2. Again, the nlsrNNLS yields the best results by producing the lowest overlap values.

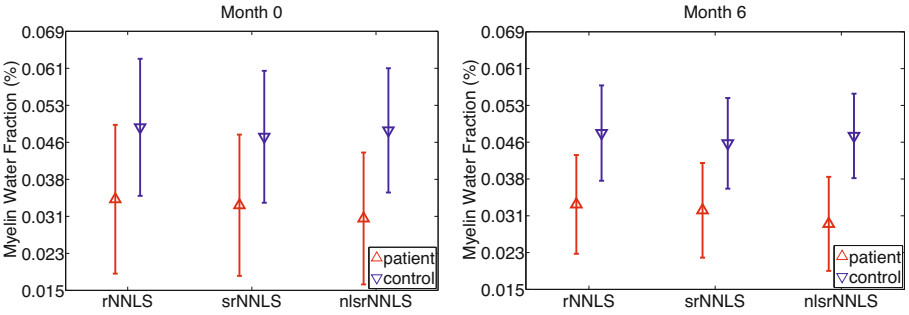

Fig. 3. Mean MWF values produced by the three regularization algorithms from 18 healthy and 17 MS patients at baseline (left) and month 6 (right). Error bars are standard errors. The nlsrNNLS shows the most clear difference between the two groups.

Table 2. Overlapping coefficients and t-test p-values for distinguishing between 18 healthy subjects and 17 MS patients computed using the MWF maps produced by the three regularization algorithms. All p-values are significant (< 0.05), but nlsrNNLS produces MWF values that are most sensitive for distinguishing the two groups.

	Month 0		Month 6	
	OVL	p-value	OVL	p-value
rNNLS [7]	0.6137	7.2×10^{-3}	0.4628	2.0260×10^{-4}
srNNLS [8]	0.6157	7.4×10^{-3}	0.4701	2.4650×10^{-4}
nlsrNNLS	**0.4915**	$\mathbf{4.3543\times10^{-4}}$	**0.3241**	$\mathbf{3.7601\times10^{-6}}$

5 Conclusion

We have presented a non-local spatial regularization for the NNLS algorithm to reliably measure myelin water in WM *in vivo*. The proposed spatial T_2 distribution prior is adaptively estimated based on the SKL divergence, and exploits information from the global WM region. Three different regularization methods have been compared using three main criteria (visual quality, CoV and ability to distinguish between MS patients and normals), and it has been observed that the proposed nlsrNNLS algorithm compares favorably with the other methods. Future work would include further optimization of algorithmic parameters, validation beyond a single scanner, longitudinal/ROIs studies (including lesions), scan-to-scan variability, and investigation of how well nlsrNNLS would work in combination with denoising strategies applied to the multi-echo images [4,5].

Acknowledgements. The authors gratefully acknowledge Drs. Irene Vavasour and Alex MacKay for the use of their data and Drs. Anthony Traboulsee and David Li for helpful discussions. This work was supported by funding from the

University of British Columbia and the Natural Sciences and Engineering Research Council of Canada.

References

1. MacKay, A., Whittall, K., Adler, J., Li, D.K.B., Paty, D., Graeb, D.: In vivo visualization of myelin water in brain by magnetic resonance. Magnetic Resonance in Medicine 31, 673–677 (1994)
2. Jones, C., Whittall, K., MacKay, A.: Robust myelin water quantification: averaging vs. spatial filtering. Magnetic Resonance in Medicine 50, 206–209 (2003)
3. Hwang, D., Chung, H., Nam, Y.: Robust mapping of the myelin water fraction in the presence of noise: Synergic combination of anisotropic diffusion filter and spatially regularized nonnegative least squares algorithm. Journal of Magnetic Resonance Imaging 195, 189–195 (2011)
4. Jang, U., Hwang, D.: High-quality multiple T_2* contrast MR images from low-quality multi-echo images using temporal-domain denoising methods. Medical Physics 39, 468–474 (2012)
5. Kwon, O., Woo, E., Du, Y., Hwang, D.: A tissue-relaxation-dependent neighboring method for robust mapping of the myelin water fraction. NeuroImage, 1–10 (2013)
6. Tikhonov, A.: Solution of incorrectly formulated problems and the regularization method. Soviet Math. Doklady 4, 1035–1038 (1963)
7. Whittall, K., MacKay, A.: Quantitative interpretation of NMR relaxation data. Journal of Magnetic Resonance 84, 134–152 (1989)
8. Hwang, D., Du, Y.: Improved myelin water quantification using spatially regularized non-negative least squares algorithm. Journal of Magnetic Resonance Imaging 30, 203–208 (2009)
9. Buades, A., Coll, B., Morel, J.: A Non-Local Algorithm for Image Denoising. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 60–65. IEEE, Washington, DC (2005)
10. Johnson, D., Sinanovic, S.: Symmetrizing the Kullback-Leibler distance. IEEE Transactions Information Theory (2001)
11. Lawson, C., Hanson, R.: Solving least squares problems. Prentice-Hall, Englewood Cliffs (1974)
12. Levesque, I., Chia, C., Pike, G.: Reproducibility of in vivo magnetic resonance imaging-based measurement of myelin water. Journal of Magnetic Resonance Imaging 32, 60–68 (2010)
13. Manjón, J., Carbonell-Caballero, J., Lull, J., García-Martí, G., Martí-Bonmatí, L., Robles, M.: MRI denoising using non-local means. Medical Image Analysis 12, 514–523 (2008)
14. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Transactions on Medical Imaging 20, 45–57 (2001)