Joint Total Variation ESTATICS for Robust Multi-parameter Mapping

Yaël Balbastre[1](✉), Mikael Brudfors[1], Michela Azzarito[2], Christian Lambert[1], Martina F. Callaghan[1], and John Ashburner[1]

[1] Wellcome Centre for Human Neuroimaging, Queen Square Institute of Neurology, University College London, London, UK
y.balbastre@ucl.ac.uk

[2] Spinal Cord Injury Center Balgrist, University Hospital Zurich, University of Zurich, Zürich, Switzerland

Abstract. Quantitative magnetic resonance imaging (qMRI) derives tissue-specific parameters – such as the apparent transverse relaxation rate $R_2^\star$, the longitudinal relaxation rate R_1 and the magnetisation transfer saturation – that can be compared across sites and scanners and carry important information about the underlying microstructure. The multi-parameter mapping (MPM) protocol takes advantage of multi-echo acquisitions with variable flip angles to extract these parameters in a clinically acceptable scan time. In this context, ESTATICS performs a joint loglinear fit of multiple echo series to extract $R_2^\star$ and multiple extrapolated intercepts, thereby improving robustness to motion and decreasing the variance of the estimators. In this paper, we extend this model in two ways: (1) by introducing a joint total variation (JTV) prior on the intercepts and decay, and (2) by deriving a nonlinear maximum *a posteriori* estimate. We evaluated the proposed algorithm by predicting left-out echoes in a rich single-subject dataset. In this validation, we outperformed other state-of-the-art methods and additionally showed that the proposed approach greatly reduces the variance of the estimated maps, without introducing bias.

1 Introduction

The magnetic resonance imaging (MRI) signal is governed by a number of tissue-specific parameters. While many common MR sequences only aim to maximise the contrast between tissues of interest, the field of quantitative MRI (qMRI) is concerned with the extraction of the original parameters [30]. This interest stems from the fundamental relationship that exists between the magnetic parameters and the tissue microstructure: the longitudinal relaxation rate $R_1 = 1/T_1$ is sensitive to myelin content [10,27,28]; the apparent transverse relaxation rate $R_2^\star = 1/T_2^\star$ can be used to probe iron content [12,21,22]; the magnetization-transfer saturation (MT_{sat}) indicates the proportion of protons bound to macromolecules (in contrast to free water) and offers another metric to investigate

© Springer Nature Switzerland AG 2020
A. L. Martel et al. (Eds.): MICCAI 2020, LNCS 12262, pp. 53–63, 2020.
https://doi.org/10.1007/978-3-030-59713-9_6

myelin loss [14,31]. Furthermore, qMRI allows many of the scanner- and centre-specific effects to be factored out, making measures more comparable across sites [3,9,29,33]. In this context, the multi-parameter mapping (MPM) protocol was developed at 3 T to allow the quantification of R_1, $R_2^\star$, MT_{sat} and the proton density (PD) at high resolutions (0.8 or 1 mm) and in a clinically acceptable scan time of 25 min [14,33]. However, to reach these values, compromises must be made so that the signal-to-noise ratio (SNR) suffers, making the parameter maps noisy; Papp et al. [23] found a scan-rescan root mean squared error of about 7.5% for R_1 at 1 mm, in the absence of inter-scan movement. Smoothing can be used to improve SNR, but at the cost of lower spatial specificity.

Denoising methods aim to separate signal from noise. They take advantage of the fact that signal and noise have intrinsically different spatial profiles: the noise is spatially independent and often has a characteristic distribution while the signal is highly structured. Denoising methods originate from partial differential equations, adaptive filtering, variational optimisation or Markov random fields, and many connections exist between them. Two main families emerge:

1. Optimisation of an energy: $\hat{Y} = \operatorname{argmin}_Y \mathcal{E}_1\left(X - \mathcal{A}(Y)\right) + \mathcal{E}_2\left(\mathcal{G}(Y)\right)$, where X is the observed data, Y is the unknown noise-free data, $\mathcal{A}$ is an arbitrary *forward* transformation (*e.g.*, spatial transformation, downsampling, smoothing) mapping from the reconstructed to the observed data and $\mathcal{G}$ is a linear transformation (*e.g.*, spatial gradients, Fourier transform, wavelet transform) that extracts features of interest from the reconstruction.
2. Application of an adaptive nonlocal filter: $\hat{Y}_i = \sum_{j\in\mathcal{N}_i} w\left(\mathcal{P}_i(X), \mathcal{P}_j(X)\right) X_j$, where the reconstruction of a given voxel i is a weighted average all observed voxels j in a given (possibly infinite) neighbourhood $\mathcal{N}_i$, with weights reflecting similarity between patches centred about these voxels.

For the first family of methods, it was found that the denoising effect is stronger when $\mathcal{E}_2$ is an absolute norm (or sum of), rather than a squared norm, because the solution is implicitly sparse in the feature domain [2]. This family of methods include total variation (TV) regularisation [25] and wavelet soft-thresholding [11]. The second family also leverages sparsity in the form of redundancy in the spatial domain; that is, the dictionary of patches necessary to reconstruct the noise-free images is smaller than the actual number of patches in the image. Several such methods have been developed specifically for MRI, with the aim of finding an optimal, voxel-wise weighting based on the noise distribution [6,7,19,20].

Optimisation methods can naturally be interpreted as a maximum *a posteriori* (MAP) solution in a generative model, which eases its interpretation and extension. This feature is especially important for MPMs, where we possess a well-defined (nonlinear) forward function and wish to regularise a small number of maps. In this paper, we use the ESTATICS forward model [32], which assumes a shared $R_2^\star$ decay across contrasts, with a joint total variation (JTV) prior. JTV [26] is an extension of TV to multi-channel images, where the absolute norm is defined across channels, introducing an implicit correlation between them. TV and JTV have been used before in MR reconstruction (*e.g.*, in compressed-sensing [15], quantitative susceptibility mapping [17], super-resolution[4]). JTV

is perfectly suited for modelling the multiple contrasts in MPMs and increases the power of the implicit edge-detection problem. However, a challenge stems from the nonlinear forward model that makes the optimisation problem nonconvex.

Our implementation uses a quadratic upper bound of the JTV functional and the surrogate problem is solved using second-order optimisation. Positive-definiteness of the Hessian is enforced by the use of Fisher's scoring, and the quadratic problem is efficiently solved using a mixture of multi-grid relaxation and conjugate gradient. We used a unique dataset – five repeats of the MPM protocol acquired, within a single session, on a healthy subject – to validate the proposed method. Our method was compared to two variants of ESTATICS: loglinear [32] and Tikhonov-regularised. We also compared it with the adaptive optimized nonlocal means (AONLM) method [20], which is recommended for accelerated MR images (as is the case in our validation data). In that case, individual echoes were denoised using AONLM, and maps were reconstructed with the loglinear variant of ESTATICS. In our validation, JTV performed consistently better than all other methods.

2 Methods

Spoiled Gradient Echo. The MPM protocol uses a multi-echo spoiled gradient-echo (SGE) sequence with variable flip angles to generate weighted images. The signal follows the equation:

$$S(\alpha, T_R, T_E) = S_0(\alpha, T_R) \exp(-T_E R_2^\star), \tag{1}$$

where α is the nominal flip angle, T_R is the repetition time and T_E is the echo time. PD and T1 weighting are obtained by using two different flip angles, while MT weighting is obtained by playing a specific off-resonance pulse beforehand. If all three intercepts S_0 are known, rational approximations can be used to compute R_1 and $\mathrm{MT}_{\mathrm{sat}}$ maps [13,14].

ESTATICS. ESTATICS aims to recover the decay rate $R_2^\star$ and the different intercepts from (1). We therefore write each weighted signal (indexed by c) as:

$$S(c, T_E) = \exp(\theta_c - T_E R_2^\star), \quad \text{with} \quad \theta_c = \ln S_{0c}. \tag{2}$$

At the SNR levels obtained in practice (>3), the noise of the log-transformed data is approximately Gaussian (although with a variance that scales with signal amplitude). Therefore, in each voxel, a least-squares fit can be used to estimate $R_2^\star$ and the log-intercepts S_c from the log-transformed acquired images.

Regularised ESTATICS. Regularisation cannot be easily introduced in logarithmic space because, there, the noise variance depends on the signal amplitude, which is unknown. Instead, we derive a full generative model. Let us assume that

all weighted volumes are aligned and acquired on the same grid. Let us define the image acquired at a given echo time t with contrast c as $\boldsymbol{s}_{c,t} \in \mathbb{R}^I$ (where I is the number of voxels). Let $\boldsymbol{\theta}_c \in \mathbb{R}^I$ be the log-intercept with contrast c and let $\boldsymbol{r} \in \mathbb{R}^I$ be the $R_2^\star$ map. Assuming stationary Gaussian noise, we get the conditional probability:

$$p\left(\boldsymbol{s}_{c,t} \mid \boldsymbol{\theta}_c, \boldsymbol{r}\right) = \mathcal{N}\left(\boldsymbol{s}_{c,t} \mid \tilde{\boldsymbol{s}}_{c,t}, \sigma_c^2 \boldsymbol{I}\right), \quad \tilde{\boldsymbol{s}}_{c,t} = \exp(\boldsymbol{\theta}_c - t\boldsymbol{r}). \tag{3}$$

The regularisation takes the form of a joint prior probability distribution over $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \ \cdots, \boldsymbol{\theta}_C, \ \boldsymbol{r}]$. For JTV, we get:

$$p\left(\boldsymbol{\Theta}\right) \propto \prod_i \exp\left(-\sqrt{\sum_{c=1}^{C+1} \lambda_c \boldsymbol{\theta}_c^{\mathrm{T}} \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{G}_i \boldsymbol{\theta}_c}\right), \tag{4}$$

where $\boldsymbol{G}_i$ extracts all forward and backward finite-differences at the i-th voxel and λ_c is a contrast-specific regularisation factor. The MAP solution can be found by maximising the joint loglikelihood with respect to the parameter maps.

Quadratic Bound. The exponent in the prior term can be written as the minimum of a quadratic function [2,8]:

$$\sqrt{\sum_c \lambda_c \boldsymbol{\theta}_c^{\mathrm{T}} \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{G}_i \boldsymbol{\theta}_c} = \min_{w_i > 0} \left\{ \frac{w_i}{2} + \frac{1}{2w_i} \sum_c \lambda_c \boldsymbol{\theta}_c^{\mathrm{T}} \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{G}_i \boldsymbol{\theta}_c \right\}. \tag{5}$$

When the weight map $\boldsymbol{w}$ is fixed, the bound can be seen as a Tikhonov prior with nonstationary regularisation, which is a quadratic prior that factorises across channels. Therefore, the between-channel correlations induces by the JTV prior are entirely captured by the weights. Conversely, when the parameter maps are fixed, the weights can be updated in closed-form:

$$w_i = \sqrt{\sum_c \lambda_c \boldsymbol{\theta}_c^{\mathrm{T}} \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{G}_i \boldsymbol{\theta}_c}. \tag{6}$$

The quadratic term in (5) can be written as $\lambda_c \boldsymbol{\theta}_c^{\mathrm{T}} \boldsymbol{L} \boldsymbol{\theta}_c$, with $\boldsymbol{L} = \sum_i \frac{1}{w_i} \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{G}_i$.

In the following sections, we will write the full (bounded) model negative loglikelihood as $\mathcal{L}$ and keep only terms that depend on $\boldsymbol{\Theta}$, so that:

$$\mathcal{L} = \sum_{c,t} \mathcal{L}_{c,t}^{\mathrm{d}} + \mathcal{L}^{\mathrm{p}}, \ \mathcal{L}_{c,t}^{\mathrm{d}} \stackrel{c}{=} \frac{1}{2\sigma_c^2} \|\boldsymbol{s}_{c,t} - \tilde{\boldsymbol{s}}_{c,t}\|^2, \ \mathcal{L}^{\mathrm{p}} \stackrel{c}{=} \frac{1}{2} \sum_c \boldsymbol{\theta}_c^{\mathrm{T}} \boldsymbol{L}_c \boldsymbol{\theta}_c. \tag{7}$$

Fisher's Scoring. The data term (3) does not always have a positive semi-definite Hessian (it is not convex). There is, however, a unique optimum. Here, to ensure that the conditioning matrix that is used in the Newton-Raphson iteration has the correct curvature, we take the expectation of the true Hessian, which is equivalent to setting the residuals to zero – a method known as Fisher's scoring. The Hessian of $\mathcal{L}_{c,t}^{\mathrm{d}}$ with respect to the c-th intercept and $R_2^\star$ map then becomes:

$$\boldsymbol{H}_{c,t}^{\mathrm{d}} = \frac{1}{\sigma^2} \mathrm{diag}\left(\tilde{\boldsymbol{s}}_{c,t}\right) \otimes \begin{bmatrix} 1 & -t \\ -t & t^2 \end{bmatrix}. \tag{8}$$

Misaligned Volumes. Motion can occur between the acquisitions of the different weighted volumes. Here, volumes are systematically co-registered using a skull-stripped and bias-corrected version of the first echo of each volume. However, rather than reslicing the volumes onto the same space, which modifies the original intensities, misalignment is handled within the model. To this end, equation (3) is modified to include the projection of each parameter map onto native space, such that $\tilde{\boldsymbol{s}}_{c,t} = \exp(\boldsymbol{\Psi}_c \boldsymbol{\theta}_c - t\boldsymbol{\Psi}_c \boldsymbol{r})$, where $\boldsymbol{\Psi}_c$ encodes trilinear interpolation and sampling with respect to the pre-estimated rigid transformation. The Hessian of the data term becomes $\boldsymbol{\Psi}_c^{\mathrm{T}} \boldsymbol{H}_{c,t}^{\mathrm{d}} \boldsymbol{\Psi}_c$, which is nonsparse. However, an approximate Hessian can be derived [1], so that:

$$\boldsymbol{H}_{c,t}^{\mathrm{d}} \approx \frac{1}{\sigma^2} \mathrm{diag}\left(\boldsymbol{\Psi}_c^{\mathrm{T}} \tilde{\boldsymbol{s}}_{c,t}\right) \otimes \begin{bmatrix} 1 & -t \\ -t & t^2 \end{bmatrix}. \tag{9}$$

Since all elements of $\tilde{\boldsymbol{s}}_{c,t}$ are strictly positive, this Hessian is ensured to be more positive-definite than the true Hessian in the Löwner ordering sense.

Newton-Raphson. The Hessian of the joint negative log-likelihood becomes:

$$\boldsymbol{H} = \boldsymbol{H}^{\mathrm{d}} + \boldsymbol{L} \otimes \mathrm{diag}\left(\boldsymbol{\lambda}\right). \tag{10}$$

Each Newton-Raphson iteration involves solving for $\boldsymbol{H}^{-1}\boldsymbol{g}$, where $\boldsymbol{g}$ is the gradient. Since the Hessian is positive-definite, the method of conjugate gradients (CG) can be used to solve the linear system. CG, however, converges quite slowly. Instead, we first approximate the regularisation Hessian $\boldsymbol{L}$ as $\tilde{\boldsymbol{L}} = \frac{1}{\min(\boldsymbol{w})} \sum_i \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{G}_i$, which is more positive-definite than $\boldsymbol{L}$. Solving this substitute system therefore ensures that the objective function improves. Since $\boldsymbol{H}^{\mathrm{d}}$ is an easily invertible block-diagonal matrix, the system can be solved efficiently using a multi-grid approach [24]. This result is then used as a warm start for CG. Note that preconditioners have been shown to improve CG convergence rates [5,34], at the cost of slowing down each iteration. Here, we have made the choice of performing numerous cheap CG iterations rather than using an expensive preconditioner.

3 Validation

Dataset. A single participant was scanned five times in a single session with the 0.8 mm MPM protocol, whose parameters are provided in Table 1. Furthermore, in order to correct for flip angles nonhomogeneity, a map of the B_1^+ field was reconstructed from stimulated and spin echo 3D EPI images [18].

Evaluated Methods. Three ESTATICS methods were evaluated: a simple loglinear fit (LOG) [32], a nonlinear fit with Tikhonov regularisation (TKH) and a nonlinear fit with joint total variation regularisation (JTV). Additionally, all echoes were denoised using the adaptive nonlocal means method (AONLM)

Table 1. Sequence parameters of the MPM protocol. The MTw sequence has an off-resonance prepulse (PP): 220°, 4 ms duration, 2 kHz off-resonance.

	FA	TR	TE	Matrix	FOV	PP
T1w	21°	25 ms	[1..8] × 2.3 ms	320 × 280 × 224	256 × 224 × 179.2 mm^3	
PDw	6°	25 ms	[1..8] × 2.3 ms	320 × 280 × 224	256 × 224 × 179.2 mm^3	
MTw	6°	25 ms	[1..6]× 2.3 ms	320 × 280 × 224	256 × 224 × 179.2 mm^3	✓

[20] before performing a loglinear fit. The loglinear and nonlinear ESTATICS fit were all implemented in the same framework, allowing for misalignment between volumes. Regularised ESTATICS uses estimates of the noise variance within each volume, obtained by fitting a two-class Rice mixture to the first echo of each series. Regularised ESTATICS possesses two regularisation factors, one for each intercept and one for the $R_2^\star$ decay, while AONLM has one regularisation factor. These hyper-parameters were optimised by cross-validation (CV) on the first repeat of the MPM protocol.

Leave-One-Echo-Out. Validating denoising methods is challenging in the absence of a ground truth. Classically, one would compute similarity metrics, such as the root mean squared error, the peak signal-to-noise ratio, or the structural similarity index between the denoised images and noise-free references. However, in MR, such references are not artefact free: they are still relatively noisy and, as they require longer sequences, more prone to motion artefacts. A better solution is to use cross-validation, as the forward model can be exploited to predict echoes that were left out when inferring the unknown parameters. We fitted each method to each MPM repeat, while leaving one of the acquired echoes out. The fitted model was then used to predict the missing echo. The quality of these predictions was scored by computing the Rice loglikelihood of the true echo conditioned on the predicted echo within the grey matter (GM), white matter (WM) and cerebro-spinal fluid (CSF). An aggregate score was also computed in the parenchyma (GM+WM). As different echoes or contrasts are not similarly difficult to predict, Z-scores were computed by normalising across repeats, contrasts and left-out echoes. This CV was applied to the first repeat to determine optimal regularisation parameters. We found $\beta = 0.4$ without Rice-specific noise estimation to work better for AONLM, while for JTV we found $\lambda_1 = 5 \times 10^3$ for the intercepts and $\lambda_2 = 10$ for the decay (in s^{-1}) to be optimal.

Quantitative Maps. Rational approximations of the signal equations [13,14] were used to compute R_1 and $\mathrm{MT}_{\mathrm{sat}}$ maps from the fitted intercepts. The distribution of these quantitative parameters was computed within the GM and WM. Furthermore, standard deviation (S.D.) maps across runs were computed for each method.

4 Results

Leave-One-Echo-Out. The distribution of Rice loglikelihoods and Z-scores for each methods are depicted in Fig. 1 in the form of Tukey's boxplots. In the parenchyma, JTV obtained the best score (mean log-likelihood: -9.15×10^6, mean Z-score: 1.19) followed by TKH (-9.26×10^6 and -0.05), AONLM (-9.34×10^6 and -0.41) and LOG (-9.35×10^6 and -0.72). As some echoes are harder to predict than others (typically, early echoes because their absence impacts the estimator of the intercept the most) the log-pdf has quite a high variance. However, Z-scores show that, for each echo, JTV does consistently better than all other methods. As can be seen in Fig. 1, JTV is particularly good at preserving vessels.

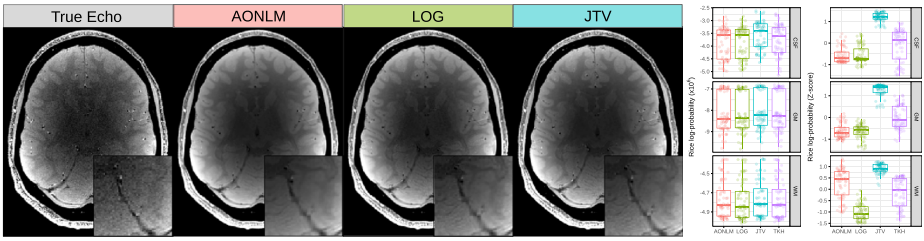

Fig. 1. Leave-one-echo out prediction. Left: the true PDw echo at $T_E = 9.7$ ms from the 5th repeat and three predicted images. Right: boxplots of the Rice log-pdf and corresponding Z-score computed for each method within GM, WM and CSF masks.

Quantitative Maps. R_1, $\mathrm{MT_{sat}}$ and $R_2^\star$ maps reconstructed with each method are shown in Fig. 2, along with mean intensity histograms within GM and WM. Note that these maps are displayed for qualitative purposes; low standard deviations are biased toward over-regularised methods and do not necessarily indicate a better predictive performance. It is evident from the histograms that all denoising methods sharpen the peaks without introducing apparent bias. It can be seen that JTV has lower variance than AONLM in the centre of the brain and higher in the periphery. This is because in our probabilistic setting, there is a natural balance between the prior and the quality of the data. In the centre of the brain, the SNR is lower than in the periphery, which gives more weight to the prior and induces a smoother estimate. The mean standard deviation of AONLM, LOG, JTV and TKH is respectively 9.5, 11.5, 11.5, 9.9 $\times 10^{-3}$ in the GM and 8.6, 12, 9.6, 10 $\times 10^{-3}$ in the WM for R_1, 15, 2, 17, 20 in the GM and 11, 20, 10, 13 in the WM for $R_2^\star$, and 4.6, 5.8, 5.1, 4.5 $\times 10^{-2}$ in the GM and 4.9, 8.2, 4.3, 4.7 $\times 10^{-2}$ in the WM for $\mathrm{MT_{sat}}$. Once again, variance is reduced by all denoising methods compared to the nonregularised loglinear fit. Again, a lower variance does not necessarily indicate a better (predictive) fit, which can only be assessed by the CV approach proposed above.

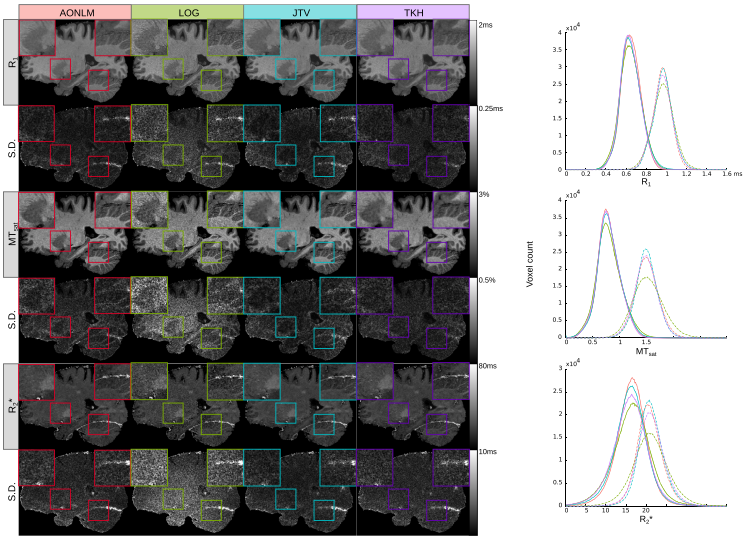

Fig. 2. Quantitative maps. Left: example R_1, $\mathrm{MT_{sat}}$ and $R_2^\star$ maps obtained with each method, and standard deviation (S.D.) maps computed across runs. Right: mean intensity histograms computed within the GM (plain) and WM (dotted) masks.

5 Discussion and Conclusion

In this paper, we introduce a robust, regularisation-based reconstruction method for quantitative MR mapping. The joint total variation prior takes advantage of the multiple MPM contrasts to increase its edge-detection power. Our approach was validated using an unbiased CV scheme, where it compared favourably over other methods, including a state-of-the-art MR denoising technique. It was shown to reduce the variance of the estimated parametric maps over non-regularised approaches, which should translate into increased power in subsequent cross-sectional or longitudinal voxel-wise studies. The use of a well-defined forward model opens the door to multiple extensions: the projection operator could be modified to include other components of the imaging process such as non-homogeneous receive fields or gridding, which would allow for joint reconstruction and super-resolution; parameters that are currently fixed *a priori*, such as the rigid matrices, could be given prior distribution and be optimised in an interleaved fashion; non-linear deformations could be included to account for changes in the neck position between scans; finally, the forward model could be unfolded further so that parameter maps are directly fitted, rather than weighted intercepts. An integrated approach like this one could furthermore include and optimise for other components of the imaging process, such as non-homogeneous transmit fields. In terms of optimisation, our approach should benefit from advances in conjugate gradient preconditioning or other solvers for large linear systems. Alternatively, JTV could be replaced with a patch-based prior.

Nonlocal filters are extremely efficient at denoising tasks and could be cast in a generative probabilistic framework, where images are built using a dictionary of patches [16]. Variational Bayes can then be used to alternatively estimate the dictionary (shared across a neighbourhood, a whole image, or even across subjects) and the reconstruction weights.

Acknowledgements. YB, MFC and JA were funded by the MRC and Spinal Research Charity through the ERA-NET Neuron joint call (MR/R000050/1). MB and JA were funded by the EU Human Brain Project's Grant Agreement No 785907 (SGA2). MB was funded by the EPSRC-funded UCL Centre for Doctoral Training in Medical Imaging (EP/L016478/1) and the Department of Health NIHR-funded Biomedical Research Centre at University College London Hospitals. CL is supported by an MRC Clinician Scientist award (MR/R006504/1). The Wellcome Centre for Human Neuroimaging is supported by core funding from the Wellcome [203147/Z/16/Z].

References

1. Ashburner, J., Brudfors, M., Bronik, K., Balbastre, Y.: An algorithm for learning shape and appearance models without annotations. NeuroImage (2018)
2. Bach, F.: Optimization with sparsity-inducing penalties. FNT Mach. Learn. **4**(1), 1–106 (2011)
3. Bauer, C.M., Jara, H., Killiany, R.: Whole brain quantitative T2 MRI across multiple scanners with dual echo FSE: applications to AD, MCI, and normal aging. NeuroImage **52**(2), 508–514 (2010)
4. Brudfors, M., Balbastre, Y., Nachev, P., Ashburner, J.: MRI super-resolution using multi-channel total variation. In: 22nd Conference on Medical Image Understanding and Analysis, Southampton, UK (2018)
5. Chen, C., He, L., Li, H., Huang, J.: Fast iteratively reweighted least squares algorithms for analysis-based sparse reconstruction. Med. Image Anal. **49**, 141–152 (2018)
6. Coupé, P., Manjón, J., Robles, M., Collins, D.: Adaptive multiresolution non-local means filter for three-dimensional magnetic resonance image denoising. IET Image Process. **6**(5), 558–568 (2012)
7. Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C.: An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Trans. Med. Imaging **27**(4), 425–441 (2008)
8. Daubechies, I., DeVore, R., Fornasier, M., Güntürk, C.S.: Iteratively reweighted least squares minimization for sparse recovery. Commun. Pure Appl. Math. **63**(1), 1–38 (2010)
9. Deoni, S.C.L., Williams, S.C.R., Jezzard, P., Suckling, J., Murphy, D.G.M., Jones, D.K.: Standardized structural magnetic resonance imaging in multicentre studies using quantitative T1 and T2 imaging at 1.5 T. NeuroImage **40**(2), 662–671 (2008)
10. Dick, F., Tierney, A.T., Lutti, A., Josephs, O., Sereno, M.I., Weiskopf, N.: In vivo functional and myeloarchitectonic mapping of human primary auditory areas. J. Neurosci. **32**(46), 16095–16105 (2012)
11. Donoho, D.: De-noising by soft-thresholding. IEEE Trans. Inf. Theory **41**(3), 613–627 (1995)

12. Hasan, K.M., Walimuni, I.S., Kramer, L.A., Narayana, P.A.: Human brain iron mapping using atlas-based T2 relaxometry. Magn. Reson. Med. **67**(3), 731–739 (2012)
13. Helms, G., Dathe, H., Dechent, P.: Quantitative FLASH MRI at 3T using a rational approximation of the Ernst equation. Magn. Reson. Med. **59**(3), 667–672 (2008)
14. Helms, G., Dathe, H., Kallenberg, K., Dechent, P.: High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and T1 relaxation obtained from 3D FLASH MRI. Magn. Reson. Med. **60**(6), 1396–1407 (2008)
15. Huang, J., Chen, C., Axel, L.: Fast multi-contrast MRI reconstruction. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012. LNCS, vol. 7510, pp. 281–288. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33415-3_35
16. Lebrun, M., Buades, A., Morel, J.M.: A nonlocal Bayesian image denoising algorithm. SIAM J. Imaging Sci. **6**(3), 1665–1688 (2013)
17. Liu, T., et al.: Morphology enabled dipole inversion (MEDI) from a single-angle acquisition: comparison with COSMOS in human brain imaging. Magn. Reson. Med. **66**(3), 777–783 (2011)
18. Lutti, A., Hutton, C., Finsterbusch, J., Helms, G., Weiskopf, N.: Optimization and validation of methods for mapping of the radiofrequency transmit field at 3T. Magn. Reson. Med. **64**(1), 229–238 (2010)
19. Manjón, J.V., Coupé, P., Buades, A., Louis Collins, D., Robles, M.: New methods for MRI denoising based on sparseness and self-similarity. Med. Image Anal. **16**(1), 18–27 (2012)
20. Manjón, J.V., Coupé, P., Martí-Bonmatí, L., Collins, D.L., Robles, M.: Adaptive non-local means denoising of MR images with spatially varying noise levels. J. Magn. Reson. Imaging **31**(1), 192–203 (2010)
21. Ogg, R.J., Langston, J.W., Haacke, E.M., Steen, R.G., Taylor, J.S.: The correlation between phase shifts in gradient-echo MR images and regional brain iron concentration. Magn. Reson. Imaging **17**(8), 1141–1148 (1999)
22. Ordidge, R.J., Gorell, J.M., Deniau, J.C., Knight, R.A., Helpern, J.A.: Assessment of relative brain iron concentrations using T2-weighted and T2*-weighted MRI at 3 Tesla. Magn. Reson. Med. **32**(3), 335–341 (1994)
23. Papp, D., Callaghan, M.F., Meyer, H., Buckley, C., Weiskopf, N.: Correction of inter-scan motion artifacts in quantitative R1 mapping by accounting for receive coil sensitivity effects. Magn. Reson. Med. **76**(5), 1478–1485 (2016)
24. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3rd edn. Cambridge University Press, Cambridge, New York (2007)
25. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**(1), 259–268 (1992)
26. Sapiro, G., Ringach, D.L.: Anisotropic diffusion of multivalued images with applications to color filtering. IEEE Trans. Image Process. **5**(11), 1582–1586 (1996)
27. Sereno, M.I., Lutti, A., Weiskopf, N., Dick, F.: Mapping the human cortical surface by combining quantitative T1 with retinotopy. Cereb. Cortex **23**(9), 2261–2268 (2013)
28. Sigalovsky, I.S., Fischl, B., Melcher, J.R.: Mapping an intrinsic MR property of gray matter in auditory cortex of living humans: a possible marker for primary cortex and hemispheric differences. NeuroImage **32**(4), 1524–1537 (2006)
29. Tofts, P.S., et al.: Sources of variation in multi-centre brain MTR histogram studies: Body-coil transmission eliminates inter-centre differences. Magn. Reson. Mater. Phys. **19**(4), 209–222 (2006). https://doi.org/10.1007/s10334-006-0049-8

30. Tofts, P.S.: Quantitative MRI of the Brain, 1st edn. Wiley, Hoboken (2003)
31. Tofts, P.S., Steens, S.C.A., van Buchem, M.A.: MT: magnetization transfer. In: Quantitative MRI of the Brain, pp. 257–298. Wiley, Hoboken (2003)
32. Weiskopf, N., Callaghan, M.F., Josephs, O., Lutti, A., Mohammadi, S.: Estimating the apparent transverse relaxation time (R2*) from images with different contrasts (ESTATICS) reduces motion artifacts. Front. Neurosci. **8**, 278 (2014)
33. Weiskopf, N., et al.: Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation. Front. Neurosci. **7**, 95 (2013)
34. Xu, Z., Li, Y., Axel, L., Huang, J.: Efficient preconditioning in joint total variation regularized parallel MRI reconstruction. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9350, pp. 563–570. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24571-3_67