SIMULTANEOUS BAYESIAN CORRECTION OF SLAB BOUNDARY ARTIFACTS AND BIAS FIELD FOR HIGH RESOLUTION *EX VIVO* MRI

J.E. Iglesias[†][*], *P.M. Paz-Alonso*[†], *G. Lerma-Usabiaga* [†], *R. Insausti*[‡], *K. Miller*[§], *C. Caballero-Gaudes*[†]

[†] Basque Center on Cognition, Brain and Language (BCBL), Donostia - San Sebastián, Spain
[‡] Human Neuroanatomy Laboratory, University of Castilla-La Mancha (UCLM), Albacete, Spain
[§] Oxford University Centre for Functional MRI of the Brain (FMRIB), Oxford, UK

ABSTRACT

Multi-slab MRI overcomes some of the hardware limitations of today's clinical scanners (e.g., memory size), enabling the acquisition of ultra-high resolution *ex vivo* MRI of the whole human brain with high SNR efficiency. However, multi-slab MRI suffers from slab boundary artifacts (SBA) that can greatly bias subsequent analyses. Since SBA heavily interplays with the bias field (BF) present in MRI, we propose a Bayesian method that corrects for SBA and BF simultaneously. The method, which combines a probabilistic brain atlas with an Expectation Maximization inference algorithm, is shown to outperform state-of-the-art SBA and BF correction techniques – even when used in combination.

Index Terms— *Ex vivo* MRI, bias field, slab boundary artifact, Venetian blind, Bayesian image enhancement

1. INTRODUCTION

Ex vivo MR has recently gained popularity in the building of human brain atlases. Since motion artifacts are eliminated, long acquisitions (tens of hours) can be used to obtain ultra-high resolution (UHR) images with high SNR [1, 2, 3]. These studies typically utilize high-field animal scanners and dedicated coils to obtain resolutions as low as 100 μm isotropic. However, the use of such equipment has two main disadvantages. First, the application is limited to the very few sites that possess such specialized resources. And second, such scanners and coils are too small to accommodate whole human brains, which is desirable in order to study human-specific conditions, e.g., psychiatric disorders.

Achieving UHR in *ex vivo* whole brain MR requires 3D acquisition sequences, which are more SNR efficient than their 2D counterparts. Since UHR 3D MRI simultaneously requires a large number of slices and small voxel dimensions, the resolution and volume size that can be achieved with clinical scanners are hampered by their limited hardware specifications. For instance, their limited memory size does not allow the reconstruction of large image volumes.

This limitation can be circumvented with multi-slab 3D MRI, i.e., stacking slabs in the slice-encoding direction. However, RF pulses do not have a rectangular profile across the slab thickness, but exhibit a transition bandwidth between the passband and the stopband, with ripples and side lobes in both bands. Moreover, the flip angle applied to the slices in the transition band is lower than in the passband [4]. These imperfections give rise to the slab boundary artifact (SBA, a.k.a. "Venetian blind"), which consists of a reduction in image intensities at the boundaries between adjacent slabs (see "NO-CORR" in Figs. 1-2). SBA is also common in MRA and diffusion MRI, and biases subsequent analyses on the images.

SBA can be partially mitigated during the MR acquisition with oversampling in the slice direction (e.g., [5]), and can also be partially compensated in the image reconstruction [6, 7]. However, correction of SBA during image postprocessing is still desirable in order to decrease the amount of required overlap slab (thus increasing the efficiency of the MR acquisition), and also to remove effects that could not be corrected in the acquisition or reconstruction.

To our best knowledge, the only post-processing method for SBA correction is *Kholmovski's* [8], which computes an intensity scaling factor for each slice that maximizes the correlation between its histogram and that of the previous slice. However, the performance is considerably affected by intensity inhomogeneities, i.e., the bias field (BF) effect, which appears as spatially smooth multiplicative noise. While the BF can subsequently be corrected with specific algorithms (e.g., N4 [9]), higher performance can potentially be achieved if both artifacts are *simultaneously* considered in the correction.

Here we propose a Bayesian method for simultaneous correction of SBA and BF. We build on ideas from the Bayesian segmentation literature, in which MR images are assumed to be generated by a model consisting of a prior (a probabilistic atlas) and a likelihood (a Gaussian mixture combined with a model of image artifacts). Within this framework, we obtain restored images by "inverting" the model with Bayes' rule. The proposed method is shown to outperform a combination of *Kholmovski's* and N4, and enables UHR scanning of whole brains with standard clinical scanners and coils.

2. METHODS

2.1. Generative model

Let $\boldsymbol{i} = (i_1, \ldots, i_V)^T$ be the log-transformed intensities of the test MRI scan to be corrected (V is its number of voxels). We assume that this scan has been generated by a model very similar to those used in the Bayesian segmentation literature, e.g. [10]. The proposed model requires that a probabilistic atlas has been coarsely (linearly) registered to the native space of the test scan. This atlas defines a prior probability $p(\boldsymbol{s})$ over the segmentation of the scan $\boldsymbol{s} = (s_1, \ldots, s_V)^T$ into four tissue types: white matter, gray matter, cerebrospinal fluid and background (and thus $s_v \in \{1, 2, 3, 4\}$). The prior is assumed to factorize over voxels, such that $p(\boldsymbol{s}) = \prod_{v=1}^{V} p_v(s_v)$, where $p_v(s)$ is a categorical distribution over tissue types at voxel v that is specified by the atlas.

Given the segmentation, the log-transformed intensities are assumed to be conditionally independent samples of four Gaussian mixture models as indexed by $\boldsymbol{s}$, and further corrupted by SBA and BF. These artifacts are modeled as a linear combination of B basis functions $\{\boldsymbol{\Psi}_b\}$ that is added to the signal (the effect is thus multiplicative in the natural domain). The likelihood function connecting $\boldsymbol{i}$ to $\boldsymbol{s}$ is given by:

$$p(\boldsymbol{i}|\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{\theta}) = \prod_{v=1}^{V} p_v(i_v|s_v, \boldsymbol{c}, \boldsymbol{\theta}), \quad \text{where}$$

$$p_v(i|s, \boldsymbol{c}, \boldsymbol{\theta}) = \sum_{c=1}^{C_s} w_{sc} \mathcal{N}\left(i - \sum_{b=1}^{B} c_b \Psi_{bv}; \mu_{sc}, \sigma_{sc}^2\right),$$

where C_s is the number of mixture components of tissue type s; $\mathcal{N}$ is the Gaussian distribution; w_{sc}, μ_{sc} and σ_{sc}^2 are the mixture weight, mean and variance of mixture component c of tissue type s; $\boldsymbol{c} = (c_1, \ldots, c_B)^T$ are the SBA/BF coefficients; and Ψ_{bv} is the value of basis function b at voxel v. We also use $\boldsymbol{\theta}$ to group the Gaussian parameters, i.e., $\boldsymbol{\theta} = \{\{w_{sc}\}, \{\mu_{sc}\}, \{\sigma_{sc}^2\}\}$.

2.2. Bayesian inference

Given this generative model, bias field correction can be cast a Bayesian inference problem: $\hat{\boldsymbol{c}} = \arg\max_{\boldsymbol{c}} p(\boldsymbol{c}|\boldsymbol{i})$. However, the computation of $p(\boldsymbol{c}|\boldsymbol{i})$ involves an intractable integral over the Gaussian parameters $\boldsymbol{\theta}$: $p(\boldsymbol{c}|\boldsymbol{i}) = \sum_s \int_{\boldsymbol{\theta}} p(\boldsymbol{c}, \boldsymbol{s}, \boldsymbol{\theta}|\boldsymbol{i}) d\boldsymbol{\theta}$. Instead, we propose to optimize the SBA/BF coefficients $\boldsymbol{c}$ and Gaussian parameters $\boldsymbol{\theta}$ simultaneously:

$$\{\hat{\boldsymbol{c}}, \hat{\boldsymbol{\theta}}\} = \arg\max_{\{\boldsymbol{c}, \boldsymbol{\theta}\}} p(\boldsymbol{c}, \boldsymbol{\theta}|\boldsymbol{i}) = \arg\max_{\{\boldsymbol{c}, \boldsymbol{\theta}\}} p(\boldsymbol{i}|\boldsymbol{c}, \boldsymbol{\theta}) p(\boldsymbol{c}) p(\boldsymbol{\theta}),$$

where we have assumed independence between $\boldsymbol{c}$ and $\boldsymbol{\theta}$. Further assuming a flat prior for the Gaussian parameters $\boldsymbol{\theta}$, and taking logarithm, the problem becomes:

$$\arg\max_{\{\boldsymbol{c}, \boldsymbol{\theta}\}} \sum_{v=1}^{V} \log\left(\sum_{s=1}^{4} p_v(i_v|s, \boldsymbol{c}, \boldsymbol{\theta}) p_v(s)\right) + \log p(\boldsymbol{c}). \quad (1)$$

The expression in Eq. 1 can be maximized with a generalized expectation maximization (GEM) algorithm, which iterates between an expectation (E) and a maximization (M) step. In the E step, a lower bound to Eq. 1 that touches it at the current estimate of the parameters ($\{\boldsymbol{c}, \boldsymbol{\theta}\}$) is computed, which involves calculating soft assignments to each Gaussian component of each tissue type at each voxel:

$$q_v^{sc} = \frac{w_{sc} \mathcal{N}\left(i_v - \sum_{b=1}^{B} c_b \Psi_{bv}; \mu_{sc}, \sigma_{sc}^2\right) p_v(s)}{\sum_{s'} \sum_{c'} w_{s'c'} \mathcal{N}\left(i_v - \sum_{b=1}^{B} c_b \Psi_{bv}; \mu_{s'c'}, \sigma_{s'c'}^2\right) p_v(s')}.$$

The M step updates the parameters such that the value of the bound is increased; this is guaranteed to increase the value of the log-likelihood (Eq. 1) as well. In our case, if we choose a quadratic penalty for the SBA/BF coefficients (i.e., $p(\boldsymbol{c}) \propto \exp[-\lambda \boldsymbol{c}^T \boldsymbol{c}]$), the updates can be shown to be [10]:

$$\mu_{sc} \leftarrow \frac{\sum_{v=1}^{V} q_v^{sc}(i_v - \sum_{b=1}^{B} c_b \Psi_{bv})}{\sum_{v=1}^{V} q_v^{sc}}$$

$$\sigma_{sc}^2 \leftarrow \frac{\sum_{v=1}^{V} q_v^{sc}(i_v - \sum_{b=1}^{B} c_b \Psi_{bv} - \mu_{sc})^2}{\sum_{v=1}^{V} q_v^{sc}}$$

$$w_{sc} \leftarrow \frac{\sum_{v=1}^{V} q_v^{sc}}{\sum_{c=1}^{C_s} \sum_{v=1}^{V} q_v^{sc}}$$

$$\boldsymbol{c} \leftarrow (\boldsymbol{A}^T \boldsymbol{D} \boldsymbol{A} + 2\lambda \boldsymbol{I})^{-1} \boldsymbol{A}^T \boldsymbol{D} \boldsymbol{r},$$

where $\boldsymbol{I}$ is the identity matrix and:

$$\boldsymbol{A} = [\boldsymbol{\Psi}_1, \cdots, \boldsymbol{\Psi}_B], \quad \boldsymbol{D} = \text{diag}(d_v), \quad d_v = \sum_{s=1}^{4} \sum_{c=1}^{C_s} d_v^{sc},$$

$$d_v^{sc} = \frac{q_v^{sc}}{\sigma_{sc}^2}, \quad \boldsymbol{r} = (r_1, \ldots, r_V)^T, \quad r_v = i_v - \frac{\sum_s \sum_c d_v^{sc} \mu_{sc}}{\sum_s \sum_c d_v^{sc}}.$$

2.3. Choice of basis functions

The model and inference method described in Sections 2.1 and 2.2 above are independent of the choice of basis functions $\{\boldsymbol{\Psi}_b\}$. Here we describe a number of choices for $\{\boldsymbol{\Psi}_b\}$ that encode different SBA/BF correction methods, which will be compared in the experiments in Section 3 below:

- **GLOBAL:** $\{\boldsymbol{\Psi}_b\}$ is a set of 3D, fourth order polynomials with global support (i.e, the whole image domain).

- **SLAB:** an independent set of polynomials with slab-wide support is defined for each slab in the stack.

- **SLAB+SL:** SLAB plus a slice-wise gain, achieved by adding one basis function per slice, such that $\Psi_{bv} = 1$ if voxel v is in the slice, and $\Psi_{bv} = 0$ otherwise [10].

- **INDEP-SL:** an independent set of 2D, fourth order polynomials is defined for each slice in the volume. Smoothness within slabs is not guaranteed by the model, but preserved in practice by the smoothness of the data and of the statistical atlas.

3. EXPERIMENTS AND RESULTS

3.1. MRI data and manual delineations

We used MRI data from two selected postmortem cases from the body donor program of the UCLM medical school. The brain extraction and fixation (intravascular) were performed by R.I. The MRI acquisition was carried out on a 3T Siemens Magnetom TIM Trio scanner using a 12 channel receiver coil. The samples were scanned in vacuum bags filled with Fluorinert, in order to minimize the impact of air bubbles and susceptibility artifacts. We used a 3D multi-slab bSSFP sequence with TE/TR = 5.3/10.6 ms and flip angle 35°. The entire volume was covered with four axial slabs with 112 slices each, and 57% slice oversampling to minimize slab aliasing, at $0.25 \times 0.25 \times 0.25$ mm voxel resolution (matrix size $720 \times 720 \times 448$). Images were acquired with RF increments of 0°, 90°, 180° and 270° (90 minutes per RF phase) and averaged to reduce banding artifacts. This protocol was repeated 10 times to increase SNR (total time: 60 hours). For evaluation purposes, manual annotations of the cerebral white matter were made by J.E.I. on the two cases. Rather than labeling the whole volumes, he traced the white matter on 10 equidistant coronal slices on each of the two cases.

3.2. Experimental setup

Seven competing methods where evaluated in this study: no correction (**"NO-CORR"**); N4 bias field correction [9] (**"N4"**), which has been shown to outperform the popular N3 algorithm; *Kholmovski's* algorithm (**"KHOL"**); a combination of the both (**"KHOL+N4"**), in which N4 is applied to the output from *Kholmovski's* method (this approach produced better results than running the algorithms in the opposite order); and four versions of our proposed method, using the four different sets of basis functions described in Section 2.3 (**"GLOBAL"**, **"SLAB"**, **"SLAB+SL"**, **"INDEP-SL"**).

The performance of the methods was assessed with two different metrics. We evaluated the BF correction with the widely used coefficient of variation ($CV = \sigma/\mu$) of the white matter intensities. Because slices at the boundaries of the slabs do not have a large impact on the CV (since they are overweighted by the rest of slices), we used a different metric to quantitatively evaluate the SBA correction.

More specifically, we took the slices at the boundaries of the slabs (4 slabs yield 3 boundaries, and thus 6 bordering slices in total), and computed the distribution (mean and variance) of the white matter intensities in them. These 6 slices are affected *the most* by SBA. Then, we considered the 10 most-central slices within each slab (40 slices in total), which are affected *the least* by SBA. We also computed the distribution of white matter intensities in this region. Finally, we used the Hellinger distance H between the two distributions as a proxy for the performance of the methods; the better the

SBA correction, the smaller we would expect H to be. Assuming Gaussianity, the (squared) Hellinger distance is:

$$H^2 = 1 - \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{4(\sigma_1^2 + \sigma_2^2)}\right),$$

where μ_1, μ_2 and σ_1^2, σ_2^2 are the means and variances of the two distributions to compare. Note that both CV and H were computed in the natural domain, even if the correction was carried out using log-transformed image intensities.

3.3. Implementation details

Throughout the experiments, we used the linear LPBA40 atlas [11] to define the anatomical priors. We used Elastix [12] to register the atlas to the test scans using an affine transform optimized at 1 mm resolution with mutual information. For the proposed method, each slice was downsampled by a factor of 4 in order to speed up the algorithm; the resulting estimated correction field was upsampled back to the native resolution. The number of mixture components was set to $C_s = 2$ for all four tissue types. The regularization parameter λ was set to 1. Finally, we shall note that our implementation optimizes the SBA/BF coefficients c one slab at the time for **SLAB** and **SLAB+SL**; and one slice at the time for **INDEP-SL**.

3.4. Results

Table 1 shows the mean CV and H for the different approaches. N4 provides a considerable decrease in the CV compared with the original image, though H increases since it is not the goal of the algorithm to correct for SBA. *Kholmovski's* method, on the other hand, successfully reduces H, but only produces a limited BF correction. When N4 is added to the pipeline, both CV and H decrease compared with the use of N4 or *Kholmovski's* algorithm alone.

The version of our method with global support takes advantage of the statistical atlas to reduce the CV further than N4, while its effect on H is similar. When basis functions with limited support are used, our methods outperform the *Kholmovski*/N4 combination. The polynomials with slab-wide support greatly decrease H, while producing a smaller decrease in CV with respect to the global version, as well. As the complexity of the basis functions increases (slabs plus slices; and finally 2D polynomials with support limited to single slices), CV and (especially) H consistently decrease.

Even though the differences in CV and H might appear modest at first, the qualitative results in Figs. 1-2 clearly illustrate the superiority of the proposed methods with respect to the baseline approaches. In Fig. 1, the generative model better estimates the BF, providing much more uniform images (e.g., note the bright temporal regions and dark subthalamic regions given that N4 cannot correct). In Fig. 2, the model with independent 2D polynomials for each slice eliminates the SBA much more effectively than the other competing algorithms.

Metric	NO-CORR	N4	KHOL	KHOL+N4	GLOBAL	SLAB	SLAB+SL	INDEP-SL
CV	0.209	0.125	0.138	0.108	0.091	0.085	0.083	0.082
H	0.141	0.182	0.112	0.099	0.181	0.098	0.077	0.063

Table 1. Means across the test cases for CV of the white matter and for the H distance between the intensity distributions at the boundaries and in the centers of the slabs. In both cases, lower is better.
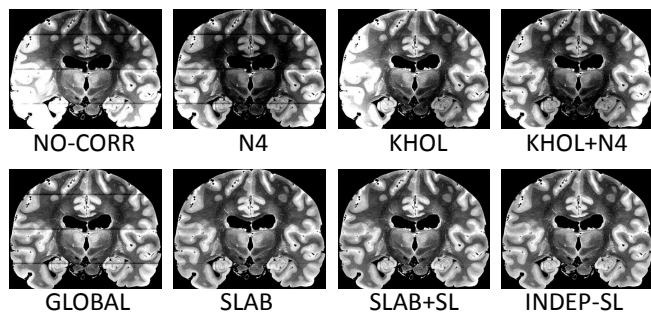
Fig. 1. Coronal slice of sample scan, corrected with the different competing methods.
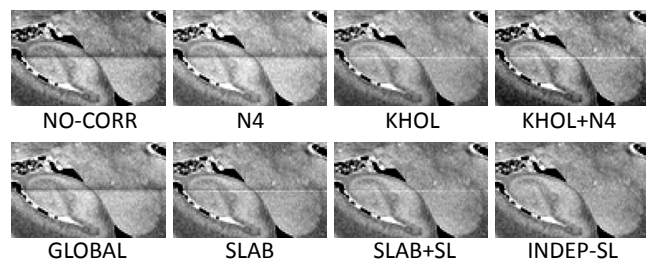
Fig. 2. Close-up of the hippocampal head and amygdala in a sagittal slice of a sample scan (input and outputs).

4. CONCLUSION

A Bayesian method for simultaneous correction of SBA and BF artifacts has been presented in this study. The method capitalizes on the relation between these two artifacts, as well as on prior knowledge encoded in a probabilistic atlas, to outperform standard methods that correct the artifacts independently (even when used in conjunction). The algorithm was used to produce high-quality, UHR *ex vivo* images using a clinical scanner, and successfully corrected for the BF and SBA that was present despite the large slice oversampling factor (57%). The method is computationally efficient; our naive Matlab implementation runs in less than 5 minutes despite the large size of the scans and the need for linear registration.

The main limitation of this study is its small sample size – a problem that is inherent to *ex vivo* human brain studies, due to the scarcity of samples and the long scanning times. Testing the method on a larger dataset, combining it with reconstruction-based SBA correction, and application to other modalities (e.g., diffusion MRI) remain as future work.

5. REFERENCES

[1] PA Yushkevich, BB Avants, et al., "A high-resolution computational atlas of the human hippocampus from postmortem magnetic resonance imaging at 9.4 T," *Neuroimage*, vol. 44, pp. 385–398, 2009.

[2] KL Miller, CJ Stagg, et al., "Diffusion imaging of whole, post-mortem human brains on a clinical MRI scanner," *Neuroimage*, vol. 57, pp. 167–181, 2011.

[3] JE Iglesias, J Augustinack, et al., "A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI," *Neuroimage*, vol. 115, pp. 117, 2015.

[4] MA Bernstein, Franklin King K, and Zhou Xiaohong J, *Handbood of MRI pulse sequences*, Elsevier, 2004.

[5] DL Parker, C Yuan, and DD Blatter, "MR angiography by multiple thin slab 3D acquisition," *Magn Reson Med*, vol. 17, pp. 434–451, 1991.

[6] AT Van, M Aksoy, et al., "Slab profile encoding (PEN) for minimizing slab boundary artifact in three-dimensional diffusion-weighted multislab acquisition," *Magn Reson Med*, vol. 73, pp. 605–613, 2015.

[7] W Wu, PJ Koopmans, et al., "Reducing slab boundary artefacts in 3D multi-slab diffusion MRI using nonlinear inversion for slab profile encoding (NPEN)," Magn Reson Med (in press).

[8] EG Kholmovski, AL Alexander, and DL Parker, "Correction of slab boundary artifact using histogram matching," *J Magn Reson Imaging*, vol. 15, pp. 610–7, 2002.

[9] NJ Tustison, BB Avants, et al., "N4ITK: improved N3 bias correction," *IEEE T Med Imaging*, vol. 29, pp. 1310–1320, 2010.

[10] K Van Leemput, F Maes, et al., "Automated model-based bias field correction of MR images of the brain," *IEEE T Med Imaging*, vol. 18, pp. 885–896, 1999.

[11] DW Shattuck, M Mirza, et al., "Construction of a 3D probabilistic atlas of human cortical structures," *Neuroimage*, vol. 39, pp. 1064–1080, 2008.

[12] S Klein, M Staring, et al., "Elastix: a toolbox for intensity-based medical image registration," *IEEE T Med Imaging*, vol. 29, pp. 196–205, 2010.