

Joint Bayesian Compressed Sensing with Prior Estimate

Berkin Bilgic¹, and Elfar Adalsteinsson^{1,2}

¹EECS, Massachusetts Institute of Technology, Cambridge, MA, United States, ²Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA, United States

INTRODUCTION: In diagnostic MRI, it is routine to acquire multiple images of the same region of interest (ROI) with different contrast preparations. In such multi-contrast acquisitions, joint Bayesian image reconstruction [1] exploits the mutual information across the shared ROI for improved image quality in accelerated acquisitions with undersampling of each contrast. As acquisition times vary among different contrasts, the overall scan time for joint multi-contrast imaging can be minimized for a fixed amount of undersampling by modulating the degree of undersampling among the different contrast preparations. Here, we extend the joint Bayesian framework to asymmetric undersampling schemes where one contrast image is fully sampled while other contrasts are undersampled.

THEORY: Given L undersampled images $\{\mathbf{x}_i\}_{i=1,L} \in \mathbb{C}^N$ acquired with different contrasts and a fully-sampled image \mathbf{x}_{prior} , a sparse representation is obtained by taking the spatial gradients in k -space: $\mathbf{F}_\Omega \delta_i = (1 - e^{-2\pi j k}) \mathbf{y}_i \equiv \mathbf{z}_i$, where $\mathbf{F}_\Omega \in \mathbb{C}^{M \times N}$ is the undersampled Fourier operator, $\{\delta_i\}_{i=1,L}$ are the image gradients, k is the k -space index and $\{\mathbf{y}_i\}_{i=1,L}$ are the k -space data. The gradient of the prior image is directly computed as $\delta_{prior} = \mathbf{F}^{-1}\{(1 - e^{-2\pi j k/n}) \mathbf{y}_{prior}\}$. We utilize both vertical and horizontal gradients, and omit the distinction for simplicity. The data are modeled to be corrupted by complex Gaussian noise with variance σ^2 , yielding the data likelihood $p(\mathbf{z}_i | \delta_i, \sigma^2) = \mathcal{N}(\mathbf{F}_\Omega \delta_i, \sigma^2 \mathbf{I})$. Joint Bayesian CS [1,2] places a Gaussian prior across each pixel of the L images to couple them, $p(\delta_i | \gamma_i) = \mathcal{N}(0, \gamma_i \mathbf{I})$, where $\delta_i \in \mathbb{C}^L$ is the vector formed by taking the i^{th} pixel in each image and γ_i is a hyperparameter controlling the variance. By multiplicative combination of all pixels, full prior distribution is obtained, $p(\delta | \gamma) = \prod_{i=1,N} p(\delta_i | \gamma_i)$. Combining the likelihood and the prior with the Bayes' rule, posterior for the i^{th} image becomes $p(\delta_i | \mathbf{z}_i, \gamma) = \mathcal{N}(\mu_i, \Sigma)$, with (i) $\Sigma = \Gamma - \Gamma \mathbf{F}_\Omega^H \mathbf{A}^{-1} \mathbf{F}_\Omega \Gamma$ and (ii) $\mu_i = \Gamma \mathbf{F}_\Omega^H \mathbf{A}^{-1} \mathbf{z}_i$ where $\mathbf{A} \equiv \sigma^2 \mathbf{I} + \mathbf{F}_\Omega \Gamma \mathbf{F}_\Omega^H$ and $\Gamma \equiv \text{diag}(\gamma)$. The posterior distribution is fully characterized if the hyperparameters γ are estimated, which can be done with an EM-type algorithm by iteratively applying Eqs. (i) & (ii) followed by the update $\gamma_i^{new} = \|\mu_i\|^2 / (L - L \Sigma_{ii} / \gamma_i)$. By using the prior image to initialize the EM iterations, $\gamma_i^{initial} = |\delta_{prior,ii}|^2$, the known sparsity support of δ_{prior} facilitates the recovery of the undersampled images. After estimating the vertical and horizontal gradients, we find images $\{\mathbf{x}_i\}_{i=1,L}$ consistent with these and the k -space data $\{\mathbf{y}_i\}_{i=1,L}$ by solving a least squares problem.

METHODS: Bayesian CS with prior was applied to two datasets, which were also reconstructed with the CS algorithm by Lustig et al. [3] using total variation penalty with an optimal regularization parameter that yielded the smallest normalized root-mean-square error (NRMSE). The first set consists of T2-weighted images obtained with two different TE's using a TSE sequence (212x212 pixels, 1x1x3 mm³, TR=6000, TE₁=27, TE₂=94 ms). An early echo slice was retrospectively undersampled with a random 2D pattern using acceleration $R = 4$ while the late echo image was kept fully sampled to serve as prior. The second dataset is derived from the SRI24 atlas [4] that features proton density (PD), T2 and T1 weighted scans at 200x200 size. Single slices from the T2 and T1 weighted images were undersampled along phase encoding with acceleration $R = 4$, while the PD image was kept fully sampled to supply prior information. An approximate solution to the large-scale matrix inversion \mathbf{A}^{-1} in Eq. (i) was computed iteratively by Lanczos algorithm with partial reorthogonalization [5] for the Bayesian CS algorithm.

RESULTS: Fig. 1 depicts the TSE dataset reconstruction results, for which Lustig et al.'s algorithm yielded 9.3% NRMSE, while Bayesian CS with prior information had 5.8% error. Results for the SRI24 dataset are given in Fig. 2. Here, Lustig et al.'s method yielded 9.5% NRMSE, and the error was 4.3% for Bayesian CS that jointly reconstructed T2 and T1 images with the help of fully-sampled PD image. Joint Bayesian CS [1] *without* using a prior had 4.9% error (not shown). All error plots are scaled 10x.

DISCUSSION: The presented method makes use of the known sparsity support of a fully-sampled image only to initialize Bayesian CS iterations, and hence avoids imposing this support on the reconstructed images. Acquiring a fully-sampled prior is desirable in cases where one imaging sequence is significantly faster than the other contrast weightings, e.g. an MP-RAGE acquisition along with other contrasts.

REFERENCES: [1] Bilgic et al., MRM 2011; [2] Wipf and Rao, IEEE T Signal Proces 2007; [3] Lustig et al., MRM 2007; [4] Rohlfing et al., Hum Brain Mapp 2010; [5] Larsen, PROPACK software package

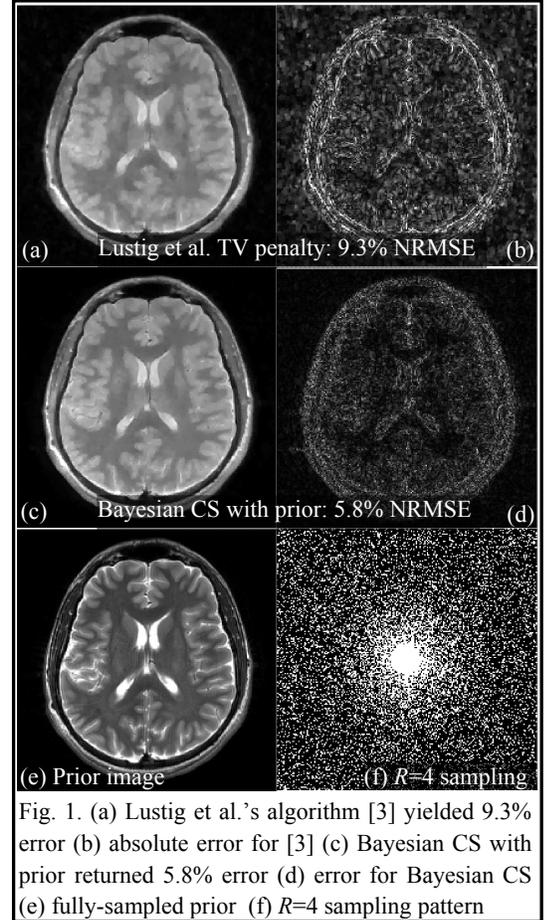


Fig. 1. (a) Lustig et al.'s algorithm [3] yielded 9.3% error (b) absolute error for [3] (c) Bayesian CS with prior returned 5.8% error (d) error for Bayesian CS (e) fully-sampled prior (f) $R=4$ sampling pattern

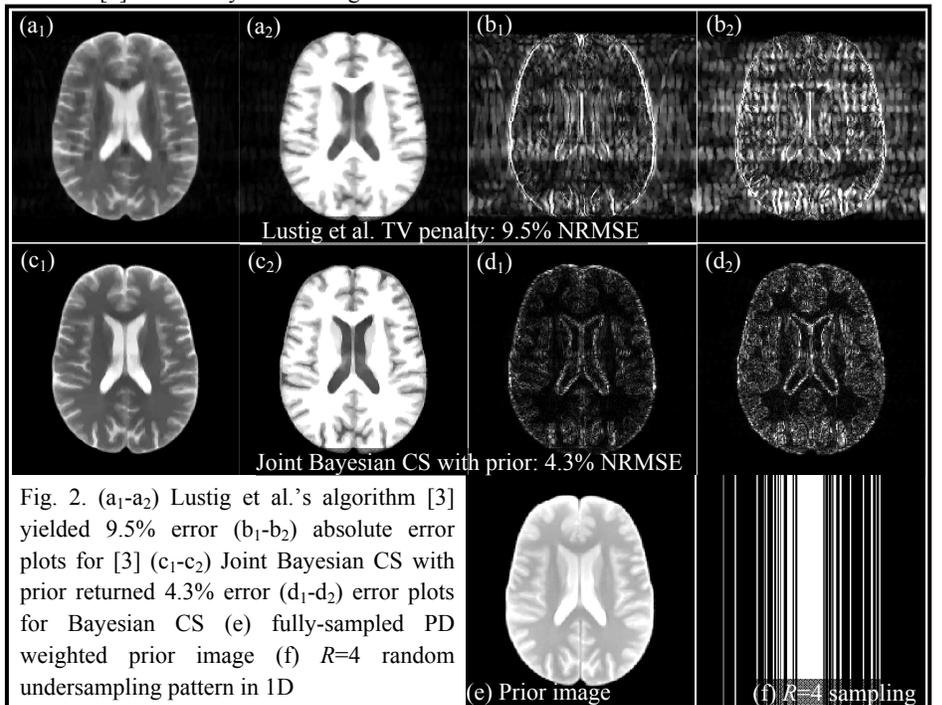


Fig. 2. (a₁-a₂) Lustig et al.'s algorithm [3] yielded 9.5% error (b₁-b₂) absolute error plots for [3] (c₁-c₂) Joint Bayesian CS with prior returned 4.3% error (d₁-d₂) error plots for Bayesian CS (e) fully-sampled PD weighted prior image (f) $R=4$ random undersampling pattern in 1D