Published in final edited form as:

Magn Reson Imaging. 2018 October ; 52: 118–130. doi:10.1016/j.mri.2018.06.008.

Content-Aware Compressive Magnetic Resonance Image Reconstruction

Daniel S. Weller^{a,1,*}, Michael Salerno^{a,b,1}, and Craig H. Meyer^{a,1}

^aUniversity of Virginia, Charlottesville, VA, USA 22904

^bUniversity of Virginia Health System, Charlottesville, VA, USA 22908

Abstract

This paper describes an adaptive approach to regularizing model-based reconstructions in magnetic resonance imaging to account for local structure or image content. In conjunction with common models like wavelet and total variation sparsity, this content-aware regularization avoids oversmoothing or compromising image features while suppressing noise and incoherent aliasing from accelerated imaging. To evaluate this regularization approach, the experiments reconstruct images from single- and multi-channel, Cartesian and non-Cartesian, brain and cardiac data. These reconstructions combine common analysis-form regularizers and autocalibrating parallel imaging (when applicable). In most cases, the results show widespread improvement in structural similarity and peak-signal-to-error ratio relative to the fully sampled images. These results suggest that this content-aware regularization can preserve local image structures such as edges while providing denoising power superior to sparsity-promoting or sparsity-reweighted regularization.

Keywords

magnetic resonance imaging; image reconstruction; compressed sensing

1. Introduction

Magnetic resonance imaging (MRI) poses significant challenges balancing acquisition speed, desired spatial and temporal resolution, and contrast and signal-to-noise ratio. Regularized inverse problem approaches like compressed sensing [1, 2, 3] and low rank modeling [4, 5, 6, 7], enable higher quality reconstructions from substantially reduced or incomplete sets of measurements. These regularizers treat all parts of an image the same, regardless of which contain salient content. So a reconstruction can over-smooth details in one part of an image while reducing noise or aliasing in other parts. To improve upon the quality achieved by regularized reconstructions, we derive a “content-aware” extension of this framework, employing an automatic and adaptive measure of local structural complexity to enhance the perceived quality of such images. A similar measure first appears in [8], where it automatically selects denoising parameters to maximize local gradient coherence, a

¹Declarations of interest: none.

*Corresponding author dweller@virginia.edu (Daniel S. Weller), ms5pc@hscmail.mcc.virginia.edu (Michael Salerno), cmeyer@virginia.edu (Craig H. Meyer).

no-reference measure of image quality. In [9], we use that measure to regularize neuron enhancement in microscope images, limiting oversmoothing of the neuron while reducing background clutter. Given undersampled k-space data, the algorithm in this work must reconstruct images using content-aware regularization. This contrasts with [9], which enhances pre-existing images. In addition, we modify this measure to emphasize the structure of the image content, rather than its quality. This idea is related to image saliency [10, 11, 12], a common component of segmentation algorithms. This technique improves the quality of regularized MRI reconstructions from noisy, incomplete k-space data encountered in many applications.

1.1. Our Contribution

Our content-aware reconstruction strategy for MR images revisits the compressed sensing style of inverse-problem-based reconstruction. We take our measure of image content/structure complexity and weight the penalty function for each of the transform coefficients according to the structure present in the local region upon which that coefficient depends. However, since we do not know the specific image structures a priori, we update and adapt the weights according to the content present in the output images from intermediate iterations of the reconstruction algorithm. Our experiments analyze the empirical convergence properties of our implementation of this algorithm and compare this method against existing compressed sensing reconstructions from the literature.

1.2. Related Work

Other work also modifies the inverse problem formulation of image reconstruction to incorporate prior knowledge of important image features. The prior-knowledge embedding in [13] reweights the sparse transform coefficients according to the expected sparsity based on prior reconstructions of longitudinal data. The approach in [14] weights the reconstruction and sparsifying transform to emphasize preselected image regions of particular importance, improving perceived image quality as a result. However, we seek to produce an improved reconstruction, even when such prior information is unavailable. Our method also avoids restricting improvements to preselected regions of interest. Instead, we consider improving the complete image.

Other reweighting schemes are present in the literature that can reflect image content. For instance, weighting sparsifying transform coefficients by the inverse of the coefficients' magnitudes improves support recovery accuracy in compressed sensing reconstruction by approximating the nonconvex 0-norm [15]. This technique makes it easier to resolve the signal support iteratively. This approach has been applied to multi-channel MRI reconstruction [16] and cardiac MRI [17] with analysis and synthesis transforms. When the sparse transform coefficients are highly correlated with the presence of content in the image, as in the case of total variation regularization, these techniques incidentally measure content. But, as shown in Figure 1(d), this is not the case for other common transforms like the wavelet, where much of the interior brain structure appears lost in the weighting. Additionally, even for total variation, the sparsity reweighting of image structures varies more with the structure intensity than our method, which could lead to differences in results. Thus, while sparsity (and image models in general) depend on the image contents, adapting

the regularizer to sparsity is not the same as adapting the regularizer to image structure or content complexity. Another approach reweights sparsity using g-factor map information about spatially-varying noise levels in the reconstruction [18]. While g-factors can be influenced by the magnetic field loading by the body, these maps rarely reflect actual image content.

Data-driven approaches like blind compressive sensing [19, 20], dictionary learning [21, 22], and transform learning [23] train data-driven models or representations of reconstructed images. Nonlinear extensions like manifold learning [24, 25] and deep learning [26, 27, 28, 29] can describe a wide range of latent features common across a large set of training images. Another deep learning approach [30] automatically tunes regularization parameters on a pixel-wise basis for CT reconstruction based on training data. Our approach does not need to be trained; instead we rely on a simple measure of content applied directly during the reconstruction. The patchwise or convolutional elements of many of these reconstructions promote universal application of the learned models. For training such models, a perceptual metric [31] is content-aware in how it computes approximation errors. However, the perceptual measure in [31] assumes the training data are clear of noise and artifacts, which is not ideal for an iterative reconstruction algorithm. Task-driven approaches in [32] also treat different image regions differently, but not necessarily according to perceived content. With our approach, we limit regularization in structurally complex regions, by adapting during the reconstruction without requiring a clean reference or ground truth. This paper does not specifically target data-driven modeling, but content complexity-based regularization likely can be applied to such models.

1.3. Outline

The next section describes the content-aware reconstruction algorithm and derives the optimization method used in the experiments that follow. Section 3 describes the datasets used, the experimental design, and the methods used to evaluate the results. Section 4 presents the reconstructed images, computation times, and quality evaluations using the proposed and competing methods; these results are discussed in Section 5. The paper concludes by summarizing the contributions of this work.

2. Theory

Given an unknown image x , represented as a column vector with N elements, we form a series of measurements at M different k-space locations at L different receiver channels. We represent the (ML) -element vector $\bar{y}$ by stacking length- M sub-vectors $y_1, \dots, y_L$ corresponding to measurements from each receiver channel. Each of these sub-vectors results from the discrete space Fourier transform (DSFT) F_Ω , sampled at k-space frequencies $\Omega = \{\omega_1, \dots, \omega_M\}$. The elements of the DSFT matrix are $[F_\Omega]_{m,n} = \frac{1}{\sqrt{N}} e^{-j\omega_m \cdot (n - n_0)}$ where n_0 is the voxel location of the origin (center) of the image, and “ $\cdot$ ” is the dot product of the location and frequency vectors. The Fourier transform y reflects the voxel-wise product of the unknown image x with a coil sensitivity profile forming the diagonal of the matrix S_ℓ for $\ell = 1, \dots, L$. We stack these coil sensitivity matrices to form the matrix

$\bar{\mathbf{S}} = [\mathbf{S}_1; \dots; \mathbf{S}_L]$. Denote $\bar{\mathbf{F}}_{\Omega} = \mathbf{I}_{L \times L} \otimes \mathbf{F}_{\Omega}$, the Kronecker product of an $L \times L$ identity matrix with the DSFT. Based on an additive circularly-symmetric complex Gaussian thermal noise model for noise η , uncorrelated across frequency and correlated across receiver channels with covariance matrix Σ , the measurements $\bar{\mathbf{y}} = \bar{\mathbf{F}}_{\Omega} \bar{\mathbf{S}} \mathbf{x} + \eta$. This generative model yields the iterative SENSE reconstruction $\hat{\mathbf{x}}$ described in [33]:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\| \bar{\mathbf{F}}_{\Omega} \bar{\mathbf{S}} \mathbf{x} - \bar{\mathbf{y}} \right\|^2. \quad (1)$$

When the k-space samples are nonuniformly spaced, a vector of density-correction factors $\mathbf{w}$ compensates for the variable weighting of k-space [34] for a single coil channel. Weighting the data fit term by $\mathbf{W} = \mathbf{I}_{L \times L} \otimes \text{diag}(\mathbf{w})$,

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\| \bar{\mathbf{F}}_{\Omega} \bar{\mathbf{S}} \mathbf{x} - \bar{\mathbf{y}} \right\|_{\mathbf{W}}^2, \quad (2)$$

Where $\|\mathbf{y}\|_{\mathbf{W}}^2 = \mathbf{y}' \mathbf{W} \mathbf{y}$ and $\mathbf{W}$ is symmetric positive semidefinite.

In general, the inverse problem is ill-posed if there are insufficient measurements or receiver channels to yield a unique solution. Regularized reconstructions add transform sparsity [35] or low rank penalties (that can also facilitate calibrationless parallel imaging [36, 37, 38, 39]). The total variation matrix $\text{TV}(\mathbf{x})$ and/or wavelet transform Ψ frequently regularize MRI reconstruction:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \left\| \bar{\mathbf{F}}_{\Omega} \bar{\mathbf{S}} \mathbf{x} - \bar{\mathbf{y}} \right\|_{\mathbf{W}}^2 + \beta \sum_p |\Psi \mathbf{x}|_p + \lambda \sum_p \left\| [\text{TV}(\mathbf{x})]_{p,:} \right\|_2. \quad (3)$$

Each column of $\text{TV}(\mathbf{x})$ represents the finite-differences transform of the image $\mathbf{x}$ in a certain direction. Thus, each row approximates a gradient vector for a voxel in $\mathbf{x}$. Here, β and λ are nonnegative tuning parameters, and p indexes the coefficients of the wavelet and/or total variation matrices. In compressed sensing, the sparsity terms can also be expressed as constraints on (1) or (2), which would be equivalent to selecting λ , β as Lagrange multipliers for those constraints. This paper focuses on the unconstrained form in (3).

These regularizers derive *local* representations for the image: total variation reflects the local gradient; wavelets are multiresolution filters with finite support. Even learned dictionaries and convolutional network kernels represent local common structure. Thus, this work introduces a coefficient-wise weighting of these sparse representations to increase the degree of regularization in regions exhibiting less structure. In the context of denoising or enhancement [9], this approach is motivated by the desire to suppress noise outside a structure, while retaining that structure's texture or character. A similar approach is considered here. Let q_n measure the content contained in the local region around voxel n in

an image x (e.g., the partial solution during an iterative reconstruction). This content measure is based on the directionality of the local gradient field. A high degree of directionality indicates an edge or corner. After approximating the image gradients via first-order differences, producing a group of row vectors for each image voxel, we stack these gradient vectors for the voxels in a neighborhood to form a gradient matrix. In two-dimensions, for a 3×3 neighborhood, our gradient matrix would measure 9×2 , and this matrix would have two nontrivial singular values σ_1 σ_2 . High directionality is indicated by $\sigma_1 \gg \sigma_2$, leading to the normalized coherence $\frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2}$. To distinguish strong directionality indicative of content boundaries from weaker edges and noise, we compute a threshold σ_{thresh} on the difference $\sigma_1 - \sigma_2$ and scale the coherence by $(\sigma_1 - \sigma_2)/\sigma_{thresh}$ when that difference is below the threshold (values above the threshold are scaled by one) to form our content measure q_n . This reduces the normalized coherence of weak structure according to how weak it is, while preserving the normalized coherence of strong structure. To choose σ_{thresh} , we use one-dimensional two-class entropy maximization [40, 41, 42] over the set of differences $\sigma_1 - \sigma_2$. This computation works well if the values are relatively bimodal (e.g., strong and weak structure) and avoids relying on second order statistics (c.f., Otsu's method [43]). The resulting map of q_n 's reflects the local structural coherence around each image voxel n . Figure 1(a) shows this image-space map for a real brain image.

For three dimensions, or a time series, we consider q_n to represent structure when gradient coherence identifies an edge (two dominant gradient directions), or to a lesser extent, the face (one dominant gradient direction) of a surface. The gradient matrix now contains 3-element row vectors for each point in the neighborhood. A $3 \times 3 \times 3$ neighborhood would produce a 27×3 local gradient matrix with nontrivial singular values σ_1 σ_2 σ_3 . In 3D, our normalized coherence generalizes to $\frac{\sigma_1 + \sigma_2 - 2\sigma_3}{\sigma_1 + \sigma_2 + \sigma_3}$, and the difference threshold is computed on the numerator of this expression. We produce the q_n map through the same process used for 2D images.

To weight the coefficients of a given transform, we need to migrate these image-space content weights to the transform space. The core idea is that a transform coefficient is local, so the relevant content Q_p for the p th transform coefficient averages the content measures q_n over all the voxels n in the local image region $\mathcal{N}_p$ for that transform coefficient:

$$Q_p = \frac{1}{|\mathcal{N}_p|} \sum_{n \in \mathcal{N}_p} q_n. \quad (4)$$

Note that this differs significantly from using the sparsifying transform directly on the image-space content map. For instance, applying a finite differences transform would zero-out significant content in regions larger than a voxel.

Our reconstruction then weights the absolute value/norm of the p th coefficient/row by the scale factor

$$\kappa_p = \frac{1 - Q_p}{\frac{1}{P} \sum_{p'=1}^P (1 - Q_{p'})}. \quad (5)$$

The k 's for each transform are normalized over all P coefficients/rows for that transform to average a constant (one). This is the same as the unweighted case, simplifying comparison of β, λ tuning parameter values between the unweighted and content-weighted reconstructions. Figure 1(b)–(e) compares sparsity-reweighting [15] against the κ maps for both total variation and wavelet transforms of a real brain image.

The content-aware reconstruction for (3) is

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \left\| \bar{\mathbf{F}}_{\Omega} \bar{\mathbf{S}} \mathbf{x} - \bar{\mathbf{y}} \right\|_{\mathbf{W}}^2 + \beta \sum_{p=1}^P \kappa_{p,1} |\Psi \mathbf{x}|_p + \lambda \sum_{p=1}^P \kappa_{p,2} \left\| [\text{TV}(\mathbf{x})]_{p,:} \right\|_2. \quad (6)$$

We add the subscripts 1 and 2 to the k 's in (6) to emphasize that the weights for wavelet and TV transforms are computed separately.

2.1. Implementation

For fixed k 's, the optimization problem in (6) closely resembles the unweighted and sparsity-reweighted versions, so we use the same nonlinear conjugate gradient (NLCG) optimization [44] with backtracking linesearch for these methods. While popular methods like alternating direction method of multipliers (ADMM) [45, 46, 47] may accelerate convergence versus NLCG methods in other compressive parallel imaging reconstructions [48, 49, 50], ADMM does not guarantee monotonicity, unlike linesearch-based NLCG. Having non-monotonic behavior can introduce instability into the reweighting procedure, leading to divergence. To initialize both reconstructions, we use density-compensated gridding:

$$\mathbf{x}^{(0)} = (\bar{\mathbf{S}}' \bar{\mathbf{S}})^{-1} \bar{\mathbf{S}}' \bar{\mathbf{F}}'_{\Omega} \mathbf{W} \bar{\mathbf{y}}, \quad (7)$$

with the nonuniform Fast Fourier transform (NUFFT) [51] taking the place of the analytical DSFT. Note $\bar{\mathbf{S}} = \mathbf{I}_{N \times N}$ for single-channel reconstruction ($L = 1$).

The nonlinear conjugate gradient method operates on “corner-rounded” approximations of the absolute value and 2-norm joint sparsity terms in (3) and (6) (see [1]). For $\varepsilon > 0$ chosen to be small compared to the squared magnitudes of $\Psi \mathbf{x}$ and $\text{TV}(\mathbf{x})$, we use

$$|\Psi \mathbf{x}|_p \approx \sqrt{[\Psi \mathbf{x}]_p^2 + \varepsilon} - \sqrt{\varepsilon}, \quad (8)$$

and

$$\left\| [\text{TV}(\mathbf{x})]_{p,:} \right\|_2 \approx \sqrt{\left\| [\text{TV}(\mathbf{x})]_{p,:} \right\|_2^2 + \varepsilon} - \sqrt{\varepsilon}. \quad (9)$$

Unlike the sparsity-promoting functions they replace, these functions are differentiable everywhere. The gradient $\nabla(\mathbf{x})$ for the objective function $f(\mathbf{x})$ then uses these approximations.

A few details regarding our NLCG-based implementation follow. First, $\nabla(\mathbf{x})$ contains the Gram matrix $\bar{\mathbf{F}}'_{\Omega} \mathbf{W} \bar{\mathbf{F}}_{\Omega}$ from the data fidelity term. This matrix is Toeplitz, and matrix-vector multiplications can be computed rapidly via circulant embedding [52, 53]. Our selected NLCG method computes the conjugate descent direction $\Delta_x^{(i)}$ for the i th iteration using the standard nonnegative form of the Polak-Ribiere formula. The literature [54, 55] establishes global convergence of this optimization method. A simple backtracking linesearch, enforcing the Wolfe conditions, computes the step size $\alpha^{(i)} > 0$ in each iteration. Then, the image update becomes $\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \alpha^{(i)} \Delta_x^{(i)}$ for $i = 0, 1, 2, \dots$. Convergence is determined by measuring both the image difference $\|\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}\|_2$ and the objective function difference $|f(\mathbf{x}^{(i)}) - f(\mathbf{x}^{(i+1)})|$. If both are less than a threshold, the algorithm is stagnating, and if the gradient at that point is small enough, the algorithm has reached a local minimum. Since the reconstruction is convex (for fixed k 's), this local minimum is also a global minimizer of the reconstruction problem with weights held fixed.

Since the ideal reconstruction is not available to us, we adapt the content-aware κ weights during successive steps of the NLCG implementation. Initially, we use constant k 's equal to one. Alternatively, we could estimate q_n values for the initial image $\mathbf{x}^{(0)}$. But, that image tends to have significant incoherent artifacts, so the k 's from such an image do not necessarily reflect true content. In dynamic imaging, the initial images do not share data over time, so they may have particularly significant aliasing. Then, as the reconstruction converges sufficiently to produce an improved image $\mathbf{x}$, that image is used to recompute the κ weights. That set of weights is held fixed, and the reconstruction continues until the image nearly converges again. This process repeats sufficiently many times, until updating the weights would produce little change in $f(\mathbf{x})$. At this point, the weights are held fixed one more time until convergence.

3. Material and Methods

This section describes the data and methods used to perform experiments evaluating the content-weighted reconstruction against the equivalent unweighted reconstruction. The data are described next, followed by the evaluation procedure, including the error metrics used.

3.1. Data Sources

The reconstruction comparison experiments use four collections of real data including a fully sampled reference image as ground truth. Depicted in Figure 2, these include a pair of

T1-weighted brain images described in [56] and [50], a clinical cardiac cine video [57], and a continuous cardiac imaging video [58].

Previous works [56, 50] describe the T1-weighted brain datasets used here. The first dataset is a T1-weighted 3D magnetization prepared rapid gradient echo (MP RAGE) acquisition of a $256 \times 256 \times 176$ mm volume (readout in the head-foot direction) with 1.0 mm isotropic (in 3D) resolution from a Siemens 3T Tim Trio scanner with a 32-channel vendor-supplied receive array head coil. This 8-minute 3D k-space acquisition obtained 256×176 phase encodes with $TR=2.53$ s, $TI=1.1$ s, $TE=3.45$ ms, $BW=190$ Hz/pixel, and $FA=7$ degrees. After taking an inverse Fourier transform along the readout (head-foot) direction, we retain a 64 mm slab in the cerebrum, use virtual coil compression [59, 60] to reduce the number of channels from 32 to 13 (95% signal energy preserved) for this region, and reconstruct the remaining 2D k-space in a slice-by-slice fashion. As these Cartesian k-space data were acquired with the readout direction orthogonal to the slice plane, we undersample by retaining a central 32×32 -readout k-space region (also used for ESPIRiT calibration [61]) and selecting readouts from the remaining k-space using first-order polynomial variable density random sampling. When viewed in 2D k-space for slice-by-slice reconstruction, these selected readouts appear as random points, with the same sampling mask used for each slice of the 64 mm slab. To vary the signal-to-noise ratio (SNR) level of this retrospectively undersampled data, complex Gaussian noise is added, where the SNR level is computed relative to the image energy in the 13-channel virtual coil images of the fully sampled volume of the 64 mm slab. This noise is correlated across virtual coil channels according to the 13-channel noise covariance matrix, formed by transforming the previously measured 32-channel noise covariance.

The second dataset is a single 3 mm-thick slice from a 10-slice 3D stack-of-spirals spoiled gradient echo acquisition with a GE 3T Discovery scanner and vendor-supplied 8-channel receive array head coil. The slice field of view (FOV) measures 300×300 mm with 256×256 pixels at full resolution (1.17×1.17 mm). Each spiral uses 120 interleaves with 832 samples/leaf and parameters $TR=10.2$ ms, $TE=6.64$ ms, $BW=125$ kHz, and $FA=16$ degrees. Since the FOV is much greater than the size of the brain, we crop the FOV (post-reconstruction) when computing error metrics. While a separate calibration scan may be used to obtain coil sensitivity maps, that information is not provided with the existing dataset. Instead, we crop a 28×28 central region in k-space for calibration. This region is not otherwise used in the reconstruction. We then use a random subset of the 120 original interleaves to undersample the data without changing the image contrast. These data have complex Gaussian noise added to produce the desired SNR, and are correlated across receive channels according to the measured noise covariance matrix provided. The SNR for this dataset is measured with respect to the combined energy of the true multi-channel image.

The next dataset is a previously imaged clinical cardiac cine scan obtained retrospectively from anonymized DICOM data. The scan was performed on a 30 year old male subject using a Siemens 3T Skyra scanner and a vendor-supplied 18-channel body coil. This set of DICOM images contains three reconstructed coil-combined cine videos after ECG-based gating over 25 cardiac phases with a measured RR-interval of 1.02 s averaged over 12 heart beats. To produce a set of three short-axis 356×380 mm slices (6 mm thick) with 1.48 mm

isotropic resolution, the steady state free precession acquisition used 2D Cartesian partial Fourier sampling (85% phase encodes) with TR=35.88 ms, TE=1.22 ms, BW=930 Hz/pixel, and FA=39 degrees. The resulting cine frames measure 240×256 pixels each. The reconstructed DICOM images were retrospectively undersampled via keeping eight lines at the center of k-space for every cardiac phase frame, and randomly selecting phase encodes from the entire 25-frame sequence according to a first-order variable-density weighting. The resulting 2D+“time” sampling mask effectively shares high-frequency data across cardiac phases, consistent with accelerated dynamic imaging practice. The same sampling mask is used for each slice, which are otherwise obtained and reconstructed separately. Complex Gaussian noise is added to the single-channel simulated k-space to achieve the desired SNR level, which is measured with respect to the energy of the true image sequence for that slice.

The final dataset is a continuous acquisition of a 2D cardiac image obtained using a spiral-out k-space trajectory with a balanced steady state free precession pulse sequence [58]. The original dataset contains 35 time frames of continuously acquired (ungated) cardiac data, at a frame rate of 8.31 Hz. Each frame was generated from 32 spiral interleaves (720 samples each) with TR=3.76 ms, TE=0.99 ms, and FA=50 degrees, obtained on a Siemens 1.5T Avanto scanner with four channels saved from an 18-channel vendor-supplied body coil. The desired FOV was 340×340 mm, with a slice thickness of 8 mm, and a matrix size of 128×128 pixels/frame. The effective in-plane spatial resolution computed from the theoretical point spread function of the non-Cartesian sequence was reported as 3.20×3.20 mm [58]. We process only the first cardiac cycle of this continuous acquisition, reconstructing only the first seven frames. We employ a golden-angle approach [62, 63] for retrospective resampling of these images, rotating the coordinates of the first spiral interleaf from the real acquisition according to the golden angle to produce the desired reduced number of interleaves in each time frame. Since each leaf is designed the same as the original acquisition leaf, just rotated, the image contrast remains the same as the original image, making retrospective resampling easier. These data are resampled from three of the four array channels (the fourth is discarded due to low signal from being too distant from the heart). Then complex noise are added to the resampled data to produce measurements at the desired SNR, correlated across channels according to the estimated noise covariance matrix. The SNR level is computed relative to the energy of the full set of multi-channel video images. Although a separate calibration scan for measuring coil sensitivities would have been desirable, such auxiliary data were not available. Instead, a 20×20 central k-space region is obtained from the fully sampled k-space interleaves for the first cardiac phase frame. These are used only to estimate the coil sensitivities.

3.2. Evaluation

Evaluation of the proposed content-aware reconstruction is tricky given that squared error measures are notoriously insensitive to local structural degradation [64]. We still report peak-signal-to-error ratio (PSER) values (in decibels) to ensure our method does not amplify noise:

$$PSER = 20 \log_{10} \frac{\max_n \|\mathbf{x}_{true}\|_n}{\|\hat{\mathbf{x}} - \mathbf{x}_{true}\|_2 / \sqrt{N}}. \quad (10)$$

However, we rely more on the mean structural similarity image index (MSSIM) [65]. This index reflects the average structural consistency over an image, accounting for local luminosity and noise characteristics estimated directly from the reconstructed and reference images. In each experiment, the fully sampled reconstruction, before adding noise, is used as a “ground truth.” Due to FOV limitations of the original spiral acquisition used to acquire the real cardiac video, we compute PSER and MSSIM over the reduced field of view corresponding to the white box in Figure 2(d), to avoid penalizing either method for suppressing spiral aliasing artifacts on the outer boundaries of the image frames. We do the same for the brain and other heart video, to focus on the organ of interest rather than the whole FOV. These MSSIM values are on a scale of zero to one, with larger values reflecting improved structural consistency. While the MSSIM and gradient-coherence measures of structural consistency are derived with similar goals in mind, these measures are completely different in their implementations, and MSSIM describes consistency relative to a reference image, while gradient-consistency is an absolute measure, obtained without a reference. Using these error metrics, we reconstruct each of the four datasets for different acceleration factors and SNR levels. For both brain datasets, a pair of common regularizers are used: 1-norm sparsity with an orthogonal Daubechies-4 wavelet with four levels, and a spatial total variation (TV) penalty. The combination of these regularizers tends to suppress artifacts common with either regularizer individually: blockiness in wavelet regularized reconstructions, and oversmoothing in TV-penalized reconstructions. For the clinical cardiac cine video, we employ total variation, across both space and time, to account for the variation is used for the continuously acquired cardiac video. For NLCG, corner-rounding $\epsilon = 10^{-6}$ is used for each penalty function. A uniform sensitivity map $\mathbf{S} = 1$ is used for single-channel reconstructions; otherwise, we estimate a single set of coil sensitivities $\mathbf{S}$ via ESPIRiT with a 7×7 k-space window and standard null space threshold of 0.02. For each experiment, we perform each reconstruction (conventional, reweighted, content-aware) up to 500 iterations, or until the changes in objective function value and reconstructed image vector $\hat{\mathbf{x}}$ both converge within 10^{-6} . For the continuous cardiac acquisition (the last dataset), the image quality actually degrades for all methods if we run it for too many iterations. So for just the fourth dataset, we run 50 iterations instead of up to 500 iterations. Calculating the content measure q_n uses a 3×3 -pixel local neighborhood ($3 \times 3 \times 3$ for time series) to estimate the local gradient coherence. These values are updated every 25 iterations, which is approximately how many iterations it takes for the reconstructions tested to reduce the objective function by several orders of magnitude for fixed weights.

3.3. Convergence Analysis

To see how the content-aware reconstruction converges, we use a simulated image from the Brainweb database (<http://brainweb.bic.mni.mcgill.ca/brainweb/>) [66], in order to have an exact ground truth. This axial Brainweb image is obtained from a simulation of a normal

anatomical model in Talairach-space with a 1 mm isotropic grid and 1 mm slice thickness of a spoiled fast low angle shot (FLASH) with TR=18 ms, TE=10 ms, FA=30 degrees, and no noise or intensity nonuniformity added. The resulting 256×256 pixel image covers a 256×256 mm FOV. For this field of view, we simulate an eight-channel parallel receive head array coil using the Image Reconstruction Toolbox [67], which simulates receive coil channels according to [68]. We then subsample k-space in a random 2D variable-density Cartesian sampling pattern (indicative of readouts from an underlying 3D imaging sequence), retaining a 26×26 fully sampled region at the center of k-space to reduce overall aliasing energy and permit direct estimation of the coil sensitivities via ESPIRiT. Circularly symmetric complex Gaussian noise are added to the multi-channel subsampled k-space, using a synthetic noise covariance matrix based on electric coupling of adjacent ring coils [69] (an approximation given ours is a circular array), and scaled to produce noise at the desired SNR (with respect to the fully-sampled coil images).

We analyze the reconstruction of this Brainweb data in two ways. The first visualizes the convergence in a Cauchy sense, to determine whether the norm between two successive iterates of x continue to get closer together over time. The next measures the objective function $f(x^{(i)})$ difference before and after updating the k 's. Ideally, we want the effect on the objective function to get smaller as the reconstruction progresses. In practice, we stop updating k 's halfway through these 500-iteration reconstructions, and let NLCG converge.

3.4. Parameter Selection

Hand-tuning regularization parameters β and λ can bias or lead to suboptimal results. Instead, we tune the reconstructions for the best possible result in terms of squared error. Given the available data, we do this in different ways for our datasets. For the 3D Cartesian brain data, our slab has 64 slices, so we optimize our tuning parameters (β and λ) for each algorithm via up to fifty iterations of the Nelder-Mead simplex method, using just the middle slice. The initial simplex is initialized to cover the range of 10^{-4} to 0.1, logarithmically, for each parameter to be tuned. We then reuse those optimized parameter values for the other 63 slices in the stack, without additional tuning. For the Cartesian clinical cine video, we have videos for three independently acquired slices, so we tune the parameter (no wavelet, so just λ) for each method with just the middle slice based on squared error and apply those parameters to reconstructions of the other slices without additional tuning. In both of these cases, we consider this approach to be practical, as it is similar to tuning parameters on previously obtained images and applying those parameters to new datasets without knowing the ground truth for the new data (cross validation).

For the spiral brain reconstruction, we have just one slice, so we use automatic parameter selection (non-Cartesian Monte Carlo WSURE [70]) instead. Since this automatic method does not require ground truth, it is practical for real-world use. For robustness, we use a coarse-to-fine grid search for both β and λ , starting with five values logarithmically spaced between 10^{-3} and 0.1, and refining the search space in each successive step by zooming in (still five values) around the WSURE-optimal parameter pair (β , λ) in the current search space. The implementation we use zooms in twice, achieving log-scale precision down to

$10^{0.125}$ in each dimension. For this dataset, we reuse the parameter pair tuned for the conventional reconstruction in both reweighted reconstructions without additional tuning.

The continuous cardiac acquisition consists of spiral data of a single slice, and WSURE estimation does not appear stable for even the conventional algorithm in this case. So, to determine the parameter λ (temporal TV, no wavelet) that achieves the best squared error for the fourth dataset, we return to the Nelder-Mead simplex method, minimizing squared error using the fully-sampled ground truth, where squared error is constrained to the region around the heart shown in Figure 2(d). While this parameter tuning approach is not ideal or practical, and it may overfit the reconstructions to this image, we notice the optimized parameter value does not vary significantly for this dataset as we accelerate or reduce SNR. So limited hand tuning could replace knowing the best parameter.

The tuned parameter values used in all the comparison experiments are provided in Section 2 of the supplementary material.

3.5. Quality Comparisons

Using our evaluation metrics (PSER and MSSIM), we compare the reconstruction quality on the four datasets described previously. These reconstructions are performed for a range of undersampling factors and noise levels, to understand the trend in reconstruction quality as a function of each. By comparing against the conventional reconstruction, we isolate the effect of the proposed content-aware weighting on the reconstruction quality. Our comparison also differentiates our approach versus an existing reweighted sparse reconstruction [15]. These reconstructions of retrospectively undersampled data predict real-world performance of our content-aware method.

4. Results

This section describes the results of our convergence analysis and quality comparisons. These results are discussed in the next section.

4.1. Convergence Analysis

Figure 3 plots on a log-scale the norm of the difference between images at successive iterations, $\|\mathbf{x}^{(j)} - \mathbf{x}^{(j-1)}\|_2$. Each plot represents a different reconstruction of the Brainweb data, undersampled by the specified factor of R, and with noise added at the specified SNR. In each case, we ran the reconstructions for up to five hundred iterations, stopping early in the case of stagnation (so $\mathbf{x}^{(j)} = \mathbf{x}^{(j-1)}$). In the early iterations, these lines are nearly the same slope, with the conventional reconstruction converging faster only in the later iterations, if at all. Both lines in the R=8, SNR = 20 dB case appear to track each other for the duration of the reconstruction. We ran the conventional reconstruction for as many iterations as the content-aware reconstruction in this test, but as depicted by the blue lines in the figure, these reconstructions tend to converge earlier than the content-aware ones. In each case, we also timed the reconstructions. Reconstruction iterations are mostly the same, differing only as to how many backtracking steps are needed for the linesearch, and including the time to update $\mathbf{k}$'s every 25 iterations for the content-aware reconstruction. Even with these differences, means and standard deviations for iteration times are not much different between methods:

0.13 ± 0.020 s and 0.13 ± 0.024 s for the proposed and conventional reconstructions, respectively (averages are pooled over the four reconstructions shown in the plot). These computation times were generated using MATLAB (The Mathworks, Natick, MA) with a dual-core Intel Core i7–6600U processor and 32 GB of RAM.

Since the objective function value changes with κ , it does not make sense to plot $f(\mathbf{x})$ in a similar way. Instead, we directly measure the changes in $f(\mathbf{x})$ due to κ during each update (every 25 iterations). For the first 250 iterations of each of the four cases shown in Figure 3, we plot the absolute difference in $f(\mathbf{x})$ values before and after updating the $\mathbf{k}$'s. Figure 4 graphs these differences for each case as a function of iteration. After significant changes from the initial updates, the change in $f(\mathbf{x})$ appears to decay in the later iterations. The largest initial changes are observed for the lower SNR (13 dB) test cases (red asterisk and purple square), and these differences appear among the largest (but still insignificant) values after 100 iterations.

4.2. Reconstruction Quality Comparisons

Next, we performed quality comparisons for each real dataset. First, we compare content-aware versus sparsity-reweighted and conventional reconstructions on retrospectively undersampled data from a pair of real T1-weighted brain images. For the 64-slice 3D brain volume, reconstructed slice-by-slice, Figure 5 depicts trends in image quality metric (PSER and MSSIM) values for different undersampling/acceleration factors (R) and signal-to-noise ratio (SNR) levels. As the SNR level decreases, the MSSIM values for the sparsity-reweighted and the conventional reconstructions drop off much faster than for the content-aware reconstruction (thick dashed lines). This reduction in MSSIM also appears more significant at higher acceleration factors. We also observe the content-aware method maintains a slight lead in PSER relative to the other methods. For the case of 6-times undersampled data with 6 dB SNR, a rather challenging reconstruction, we plot slice 36/64 in Figure 6. The content-aware reconstruction appears smoother in the central white matter region, and is not quite as noisy outside the brain. Reconstructions of all 64 slices are depicted in Section 3 of the supplementary material.

For the spiral brain dataset, Figure 7 depicts trends in image quality metric (PSER and MSSIM) values for different undersampling/acceleration factors (R) and signal-to-noise ratio (SNR) levels. While the trend is a bit less clear than in the 3D Cartesian case, we still observe a consistent advantage in PSER and MSSIM values over a wide range of accelerations and SNR levels. Here, the difference appears more pronounced in PSER than in MSSIM, even though the WSURE parameter selection method attempts to select the best squared error. Figure 8 illustrates small differences at a lower acceleration, but the difference images clearly highlight reduced error towards the skull and meninges areas in the content-aware reconstruction versus the other methods.

Next, we compare reconstructions of cardiac data. First up is a retrospective reconstruction of clinical cardiac cine videos of three short-axis slices. The plots in Figure 9 suggest that the content-aware method consistently improves PSER and MSSIM values (for all three slices combined) over a wide range of accelerations (R) and SNR levels. The most significant gains also appear at the lowest SNR levels. The images in Figure 10 show

examples for 3-times undersampled data with 16 dB SNR from the third slice. Even at this level, the three reconstructions appear blurred, but the content-aware reconstruction has larger MSSIM, suggesting the structure is more similar to the ground truth.

The continuous cardiac acquired data (at a frame rate of 8.3 Hz) were undersampled and reconstructed using SENSE parallel imaging with temporal total variation regularization. The content-aware and conventional variants produce results with similar PSER values, but the MSSIM values for the content-aware version are mostly a bit higher, across different SNR levels and undersampling factors (R). These trends are depicted in Figure 11. We also include reconstructed images in Figure 12, which unsurprisingly shows blurring especially around the myocardium at end-systole. However, the structural similarity remains quantitatively higher for the content-aware reconstruction than the conventional and sparsity-reweighted methods.

5. Discussion

The experiments described in this paper provide numerous insights into the capabilities and functionality of the content-aware regularization method. First, we discuss the convergence of the content-aware reconstruction using NLCG optimization. Then we discuss the comparisons on real (retrospective) data. Then we make some overall interpretations regarding the suitability of the method for accelerated MRI reconstruction.

The convergence analysis performed for the reconstruction of a synthetic Brainweb image from undersampled data with different degrees of acceleration (R) and noise levels (SNR) suggests a couple insights into the behavior of the algorithm with adaptive κ weights. First, the similar trends in convergence in a Cauchy sense for the image estimates $x^{(i)}$ for conventional and content-aware reconstruction suggest the NLCG method is not adversely affected by adapting the weights. Typically, a conjugate gradient algorithm will enjoy acceleration due to conjugate directions navigating the graph of the objective function more efficiently than traditional steepest descent. If we change the objective function through reweighting k 's, the conjugacy of these directions would be affected. However, the trends are very similar until late in the reconstruction (after 300400 iterations), suggesting that updating the weights has a minor effect on convergence, especially early on. This observation could mean that the conjugacy of successive descent directions is mostly driven by the data fit term in (6) for appropriate values of λ and β . However, we would expect a more significant effect if λ or β increase substantially. Of course, at that level, image quality may be suboptimal for either reconstruction independent of convergence rate. Another important observation is that while the change in objective function due to updating k 's decreases in later iterations, it does not get arbitrarily small. However, this issue does not appear to affect overall convergence in $x^{(i)}$, suggesting the reconstruction actually becomes less sensitive to k 's in later iterations. But, it emphasizes the importance of only updating κ for a limited number of iterations and allowing the overall reconstruction to stabilize.

The comparisons also reveal important insights about the content-aware method's behavior. The presence of structure should guide the degree of regularization in an image region. As we observe in Figure 1, the structural content varies significantly in real images, and

depending on the transform, the content weighting can either be similar or very different from a sparsity-based weighting. Our idea is that regularizing less heavily on image structure would allow increased regularization in noisy, less structured areas, suppressing overall noise while retaining structure. The results indicating improved MSSIM in most cases suggest our approach appears to mitigate the common tradeoff between noise reduction and oversmoothing. This is particularly important when reducing noise close to focal elements of images, and in smooth regions of those images. The illustration of improved SSIM values in the smoother central region of the brain in Figure 6 is evidence of this increased smoothing while preserving structure. Observing a noticeable difference in SSIM values confirms the value of the quality improvement.

The PSER and MSSIM values for the spiral brain reconstruction appear less stable than for the other reconstructed datasets. This appears to be a limitation of automatic parameter selection, and some amount of additional hand tuning may be necessary to ensure consistently high quality for either of these reweighting methods. These results suggest that new automatic parameter selection methods are needed for such algorithms, especially since squared error is not a great indicator of image structural consistency.

While the cine and continuous cardiac acquisitions both show measurable improvements in MSSIM values, the structural differences in the reconstructed images are not as clearly visible as for the brain images. Indeed, the total variation regularizations for these cardiac reconstructions tend to produce blurrier images than the wavelet-and-total variation combinations used in the brain datasets, even at low accelerations. So, while reweighting is helpful, new regularization models may be even more useful in cardiac imaging.

Given the differences between PSER and MSSIM values for these reconstructions, further investigation with larger data sets is warranted to measure the statistical significance of these quality differences. The increase in PSER and MSSIM values observed over multiple slices from the same dataset, even without tuning parameters for each slice, suggests larger-scale analyses are feasible.

6. Conclusions

This paper introduced a novel content-aware reweighting scheme for sparsity-regularized inverse problems used in MRI reconstruction. We tested this method on retrospectively undersampled brain and cardiac MRI data and demonstrated improved structural quality (reported by MSSIM) for these datasets across a range of acceleration factors and noise levels. We discussed the implications of this work, including the similar convergence with content-aware regularization, and conclude here that the content-aware method is suitable for a range of reconstructions and can complement work in improving image modeling.

In future work, we aim to extend this work to different types of regularization penalties, such as data-driven transforms, and explore real-world applications of our approach, especially in cardiac and body imaging, where acquisition speed is important, and the field of view contains a wide range of content. We also plan to explore other uses of our measure of

content awareness such as focused training of deep neural networks and as an alternative to existing local saliency measures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We acknowledge Yang Yang and other members of Dr. Salerno's research group, who aided in the retrospective collection of datasets used in this paper.

Funding: This work was supported by the National Institutes of Health [grant number R21EB022309 and R01HL131919] and the Thomas F. and Kate Miller Jeffress Memorial Trust, Bank of America, Trustee.

References

- [1]. Lustig M, Donoho D, Pauly JM, Sparse MRI: The application of compressed sensing for rapid MR imaging, *Mag. Res. Med* 58 (2007) 1182–95.
- [2]. Lustig M, Donoho DL, Santos JM, Pauly JM, Compressed sensing MRI, *IEEE Sig. Proc. Mag* 25 (2008) 72–82.
- [3]. Otazo R, Kim D, Axel L, Sodickson DK, Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI, *Mag. Res. Med* 64 (2010) 767–76.
- [4]. Akcakaya M, Basha TA, Goddu B, Goepfert LA, Kissinger KV, Tarokh V, Manning WJ, Nezafat R, Low-dimensional-structure selflearning and thresholding: Regularization beyond compressed sensing for MRI Reconstruction, *Mag. Res. Med* 66 (2011) 756–767.
- [5]. Lingala SG, Hu Y, DiBella E, Jacob M, Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR, *IEEE Trans. Med. Imag* 30 (2011) 1042–54.
- [6]. Zhao B, Haldar JP, Christodoulou AG, Liang Z-P, Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints, *IEEE Trans. Med. Imag* 31 (2012) 1809–20.
- [7]. Otazo R, Cands E, Sodickson DK, Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components, *Magnetic Resonance in Medicine* 73 (2015) 11251136.
- [8]. Zhu X, Milanfar P, Automatic parameter selection for denoising algorithms using a no-reference measure of image content, *IEEE Trans. Im. Proc* 19 (2010) 3116–32.
- [9]. Liang H, Acton ST, Weller DS, Content-aware neuron image enhancement, in: *Proc. IEEE Intl. Conf. on Image Processing*, 2017, pp. 3510–3514. doi:10.1109/ICIP.2017.8296935.
- [10]. Itti L, Koch C, Computational modelling of visual attention, *Nature Rev. Neurosci* 2 (2001) 194–203. [PubMed: 11256080]
- [11]. Kadir T, Brady M, Saliency, scale and image description, *Intl. J. Comp. Vision* 45 (2001) 83–105.
- [12]. Hou X, Zhang L, Saliency detection: A spectral residual approach, in: *Proc. IEEE Conf. on Comp. Vision and Pattern Recognition*, 2007, pp. 1–8. doi: 10.1109/CVPR.2007.383267.
- [13]. Weizman L, Eldar YC, Bashat DB, Compressed sensing for longitudinal MRI: An adaptive-weighted approach, *Medical Physics* 42 (2015) 5195–5208. [PubMed: 26328970]
- [14]. Oh H, Lee S, Visually weighted reconstruction of compressive sensing MRI, *Magnetic Resonance Imaging* 32 (2014) 270–280. [PubMed: 24332887]
- [15]. Candes EJ, Wakin MB, Boyd S, Enhancing sparsity by reweighted l1 minimization, *J. Fourier Anal. and Appl* 14 (2008) 877–905.
- [16]. Chang CH, Ji J, Improved compressed sensing mri with multi-channel data using reweighted l1 minimization, in: *Proc. Int'l. Conf. IEEE Engr. in Med. and Biol. Soc.*, 2010, pp. 875–878. doi: 10.1109/IEMBS.2010.5627890.

- [17]. Chen S-W, Chao S-C, A reweighted ℓ_1 -minimization based compressed sensing for the spectral estimation of heart rate variability using the unevenly sampled data, PLoS ONE 9 (2014) e99098. [PubMed: 24922059]
- [18]. Huang F, Chen Y, Yin W, Lin W, Ye X, Guo W, Reykowski A, A rapid and robust numerical algorithm for sensitivity encoding with sparsity constraints: Self-feeding sparse SENSE, Magnetic Resonance in Medicine 64 (2010) 1078–1088. [PubMed: 20564598]
- [19]. Gleichman S, Eldar YC, Blind compressed sensing, IEEE Trans. Info. Theory 57 (2011) 6958–75.
- [20]. Lingala SG, Jacob M, Blind compressive sensing dynamic MRI, IEEE Trans. Med. Imag 32 (2013) 1132–1145.
- [21]. Ravishankar S, Bresler Y, MR image reconstruction from highly undersampled k-space data by dictionary learning, IEEE Trans. Med. Imag 30 (2011) 1028–1041.
- [22]. Caballero J, Price AN, Rueckert D, Hajnal JV, Dictionary learning and time sparsity for dynamic MR data reconstruction, IEEE Trans. Med. Imag 33 (2014) 979–994.
- [23]. Ravishankar S, Bresler Y, Learning sparsifying transforms, IEEE Trans. Sig. Proc 61 (2013) 1072–86.
- [24]. Usman M, Atkinson D, Kolbitsch C, Schaeffter T, Prieto C, Manifold learning based ECG-free free-breathing cardiac CINE MRI, Journal of Magnetic Resonance Imaging 41 (2015) 1521–1527. [PubMed: 25124545]
- [25]. Poddar S, Jacob M, Dynamic MRI using smoothness regularization on manifolds (SToRM), IEEE Transactions on Medical Imaging 35 (2016) 1106–1115. [PubMed: 26685228]
- [26]. Yang Y, Sun J, Li H, Xu Z, Deep ADMM-net for compressive sensing MRI, in: Adv. in Neural Info. Proc. Sys 29, 2016, pp. 10–18. URL: <http://papers.nips.cc/paper/6406-deep-admm-net-for-compressive-sensing-mri.pdf>.
- [27]. Schlemper J, Caballero J, Hajnal JV, Price A, Rueckert D, A deep cascade of convolutional neural networks for dynamic MR image reconstruction, IEEE Transactions on Medical Imaging 37 (2018) 491–503. [PubMed: 29035212]
- [28]. Yang G, Yu S, Dong H, Slabaugh G, Dragotti PL, Ye X, Liu F, Arridge S, Keegan J, Guo Y, Firmin D, DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction, IEEE Transactions on Medical Imaging 37 (2018) 1310–1321. [PubMed: 29870361]
- [29]. Dar SUH, Cukur T, A transfer-learning approach for accelerated MRI using deep neural networks, 2017 URL: <http://arxiv.org/abs/1710.02615>, arxiv 1710.02615.
- [30]. Shen C, Gonzalez Y, Chen L, Jiang SB, Jia X, Intelligent parameter tuning in optimization-based iterative CT reconstruction via deep reinforcement learning, 2017 URL: <http://arxiv.org/abs/1711.00414>, arxiv 1711.00414.
- [31]. Rehman A, Rostami M, Wang Z, Brunet D, Vrscaj ER, SSIM-inspired image restoration using sparse representation, EURASIP J. Adv. Sig. Proc 2012 (2012) 1–12.
- [32]. Mairal J, Bach F, Ponce J, Task-driven dictionary learning, IEEE Trans. Patt. Anal. Mach. Int 34 (2012) 791–804.
- [33]. Pruessmann KP, Weiger M, Bornert P, Boesiger P, Advances in sensitivity encoding with arbitrary k-space trajectories, Mag. Res. Med 46 (2001) 638–51.
- [34]. Hoge RD, Kwan RKS, Pike GB, Density compensation functions for spiral MRI, Mag. Res. Med 38 (1997) 117–28.
- [35]. Liang D, Liu B, Wang J, Ying L, Accelerating SENSE using compressed sensing, Mag. Res. Med 62 (2009) 1574–84.
- [36]. Majumdar A, Ward RK, Calibration-less multi-coil MR image reconstruction, Mag. Res. Im 30 (2012) 1032–45.
- [37]. Trzasko J, Manduca A, A unified tensor regression framework for calibrationless dynamic, multi-channel MRI reconstruction, in: Proc. Intl. Soc. Mag. Res. Med., 2013, p. 603 URL: <http://submissions.miramcd.com/isrmr2013/proceedings/files/0603.PDF>.
- [38]. Shin PJ, Larson PEZ, Ohliger MA, Elad M, Pauly JM, Vigneron DB, Lustig M, Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion, Magnetic Resonance in Medicine 72 (2014) 959–970. [PubMed: 24248734]

- [39]. Haldar JP, Zhuo J, P-LORAKS: Low-rank modeling of local k-space neighborhoods with parallel imaging data, *Mag. Res. Med* 75 (2016) 1499–1514.
- [40]. Pun T, A new method for grey-level picture thresholding using the entropy of the histogram, *Signal Processing* 2 (1980) 223–237.
- [41]. Kapur JN, Sahoo PK, Wong AKC, A new method for gray-level picture thresholding using the entropy of the histogram, *Computer Vision, Graphics, and Image Processing* 29 (1985) 273–285.
- [42]. Abutaleb AS, Automatic thresholding of gray-level pictures using two dimensional entropy, *Computer Vision, Graphics, and Image Processing* 47 (1989) 22–32.
- [43]. Otsu N, A threshold selection method from grey-level histograms, *IEEE Trans. Sys. Man. Cyb* 9 (1979) 62–66.
- [44]. Fletcher R, Reeves CM, Function minimization by conjugate gradients, *Comput. J* 7 (1964) 149–54.
- [45]. Glowinski R, Marrocco A, Sur l'approximation, par elements finis d'ordre un, et la resolution, par penalisation-dualite d'une classe de problemes de dirichlet non lineaires, *Modelisation Mathematique et Analyse Numerique* 9 (1975) 41–76.
- [46]. Gabay D, Mercier B, A dual algorithm for the solution of nonlinear variational problems via finite-element approximations, *Comput. Math. Appl* 2 (1976) 17–40.
- [47]. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. & Trends in Machine Learning* 3 (2011) 1–122.
- [48]. Ramani S, Fessler JA, Parallel MR image reconstruction using augmented Lagrangian methods, *IEEE Trans. Med. Imag* 30 (2011) 694–706.
- [49]. Ramani S, Fessler JA, Accelerated nonCartesian SENSE reconstruction using a majorize-minimize algorithm combining variable-splitting, in: *Proc. IEEE Intl. Symp. Biomed. Imag*, 2013, pp. 704–7. doi:10.1109/ISBI.2013.6556572.
- [50]. Weller DS, Ramani S, Fessler JA, Augmented Lagrangian with variable splitting for faster non-Cartesian L1-SPIRiT MR image reconstruction, *IEEE Trans. Med. Imag* 33 (2014) 351–361.
- [51]. Fessler JA, Sutton BP, Nonuniform fast Fourier transforms using min-max interpolation, *IEEE Trans. Sig. Proc* 51 (2003) 560–74.
- [52]. Dembo A, Mallows CL, Shepp LA, Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with application to covariance estimation, *IEEE Trans. Info. Theory* 35 (1989) 1206–12.
- [53]. Baron CA, Dwork N, Pauly JM, Nishimura DG, Rapid compressed sensing reconstruction of 3D non-cartesian MRI, *Magnetic Resonance in Medicine* 79 (2018) 2685–2692. [PubMed: 28940748]
- [54]. Gilbert JC, Nocedal J, Global convergence properties of conjugate gradient methods for optimization, *SIAM J. Optim* 2 (1992) 21–42.
- [55]. Dai Y, Han J, Liu G, Sun D, Yin H, Yuan Y-X, Convergence properties of nonlinear conjugate gradient methods, *SIAM J. Optim* 10 (1999) 345–58.
- [56]. Weller DS, Polimeni JR, Grady L, Wald LL, Adalsteinsson E, Goyal VK, Denoising sparse images from GRAPPA using the nullspace method (DESIGN), *Mag. Res. Med* 68 (2012) 1176–89.
- [57]. Jeelani H, Martin J, Vasquez F, Salerno M, Weller DS, Image quality affects deep learning reconstruction of MRI, in: *Proc. IEEE Intl. Symp. Biomed. Imag*, 2018, pp. 357–360. doi: 10.1109/ISBI.2018.8363592.
- [58]. Feng X, Salerno M, Kramer CM, Meyer CH, Non-cartesian balanced steady-state free precession pulse sequences for real-time cardiac MRI, *Magnetic Resonance in Medicine* 75 (2016) 1546–1555. [PubMed: 25960254]
- [59]. Buehrer M, Pruessmann KP, Boesiger P, Kozerke S, Array compression for MRI with large coil arrays, *Mag. Res. Med* 57 (2007) 1131–9.
- [60]. Huang F, Vijayakumar S, Li Y, Hertel S, Duensing GR, A software channel compression technique for faster reconstruction with many channels, *Mag. Res. Im* 26 (2008) 133–41.

- [61]. Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, Vasanawala SS, Lustig M, ESPIRiT- an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA, *Mag. Res. Med* 71 (2014) 990–1001.
- [62]. Winkelmann S, Schaeffter T, Koehler T, Eggers H, Doessel O, An optimal radial profile order based on the golden ratio for time-resolved MRI, *IEEE Trans. Med. Imag* 26 (2007) 68–76.
- [63]. Tolouee A, Alirezaie J, Babyn P, Compressed sensing reconstruction of cardiac cine MRI using golden angle spiral trajectories, *J. Mag. Res* 260 (2015) 10–19.
- [64]. Wang Z, Bovik AC, Mean squared error: Love it or leave it? A new look at signal fidelity measures, *IEEE Sig. Proc. Mag* 26 (2009) 98–117.
- [65]. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Im. Proc* 13 (2004) 600–612.
- [66]. Kwan RK-S, Evans AC, Pike GB, MRI simulation-based evaluation of image-processing and classification methods, *IEEE Trans. Med. Imag* 18 (1999) 1085–97.
- [67]. Fessler JA, Michigan image reconstruction toolbox, 2014 URL: <https://web.eecs.umich.edu/~fessler/code/>.
- [68]. Grivich MI, Jackson DP, The magnetic field of current-carrying polygons: An application of vector field rotations, *Amer. J. Phys* 68 (2000) 469–74.
- [69]. Roemer PB, Edelstein WA, Hayes CE, Souza SP, Mueller OM, The NMR phased array, *Mag. Res. Med* 16 (1990) 192–225.
- [70]. Ramani S, Weller DS, Nielsen J-F, Fessler JA, Non-Cartesian MRI reconstruction with automatic regularization via Monte-Carlo SURE, *IEEE Trans. Med. Imag* 32 (2013) 1411–1422

Figure 1:

Different reweighting strategies are illustrated for the spiral T1-weighted brain [50] used in this paper. Cropped and enlarged to show detail, the first map (a) shows q_n values for each pixel in the true image, on a scale from 0 (black) to 1 (white), where larger values indicate dominant structure. The other maps in the top row are the sparsity-reweighted (b) and content-weighted (c) κ weight maps for the total variation penalty. The bottom row (d), (e) are the same maps for the discrete wavelet transform penalty. These κ weight maps also are on a scale from 0 (black) to 1 (white) and weight the ℓ_1 penalties for these two transforms.

Figure 2:

This paper reconstructs four different datasets. The first dataset contains 64 slices (two shown in (a)) of a T1-weighted 3D brain acquisition from [56]. The second dataset is a spiral T1-weighted brain (b) from [50]. The third dataset is a clinical cardiac cine scan used in [57], and the fourth is a continuously acquired cardiac scan described in [58]. End-diastolic and end-systolic frames are shown for each cardiac dataset. The parallel imaging datasets are displayed using sum-of-squares combination of coils. The white boxes indicate the regions of interest for these reconstructions.

Figure 3:

The differences between images for current and previous iterations $\|x(i) - x^{(i-1)}\|$ are plotted versus iteration i for the proposed content-aware algorithm (thick, red lines) and the conventional reconstruction (thin, blue lines). These plots reflect reconstructions of the Brainweb image from Rx -undersampled data with the listed signal-to-noise ratio (SNR).

Figure 4:

The objective function differences $|f(x^{(i)}) - f(x^{(i-1)})|$ after each update of the content weights ($\kappa, \nu, \kappa, 2$) are plotted versus iteration i . This plot depicts the convergence of these weights for the first 250 iterations of a 500-iteration Brainweb reconstruction with the specified data undersampling (R) and signal-to-noise ratio (SNR). After 250 iterations, the overall reconstruction is allowed to converge without further updates (see Figure 3).

Figure 5:

The peak-signal-to-error ratio (PSER) and mean structural similarity image index (MSSIM) values are higher (better) for the Cartesian brain volume reconstructed using the proposed content-aware approach (thick, dashed lines) than those images reconstructed using the conventional (thin, solid lines) or sparsity-reweighted (thin, dashed lines) approaches. This improvement is most significant at low signal-to-noise ratio (SNR) levels and at higher acceleration factors. (t-b: $R = 3, 4, 5, 6$).

Figure 6:

The ground truth slice 36/64 is depicted for the 3D Cartesian brain volume in (a). The content-aware reconstructed image (d) is compared visually versus conventional (b) and sparsity-reweighted (c) reconstructions for 6-times undersampled data with 6 dB signal-to-noise ratio (SNR). The difference images for the conventional, sparsity-reweighted, and content-aware reconstructions are shown in (e)-(g), respectively. Less noise is present along the skull boundary and inside the brain. Peak-signal-to-error (PSER) and mean structural similarity image index (MSSIM) values are below each image. Note that these reconstructions use parameters tuned for another slice in the same volume.

Figure 7:

The peak-signal-to-error ratio (PSER) and mean structural similarity image index (MSSIM) values for the white box containing the brain in the reconstructed images of retrospectively undersampled spiral brain data are compared for our content-aware approach (thick, dashed lines) and the conventional (thin, solid lines) and sparsity-reweighted (thin, dashed lines) approaches. The PSER and MSSIM values appear mostly higher for the content-aware method, especially at low signal-to-noise ratio (SNR) levels and high acceleration factors. We show results for various acceleration/undersampling factors (t-b: R=3, 4, 5, 6).

Figure 8:

The (c) content-aware reconstructed spiral brain image (ground truth in Figure 2(b)) for $R=3$, $SNR=10$ dB tends to have better quality than the (a) conventional and (b) sparsity-reweighted reconstructions. As a result, the difference image in (f) has fewer degradations than the maps in (d) and (e), especially in the skull region. PSER and MSSIM values are reported for the brain region (see the white box in Figure 2(b)).

Figure 9:

The peak-signal-to-error ratio (PSER) and mean structural similarity image index (MSSIM) values are higher (better) for the Cartesian clinical cardiac cine videos of three slices reconstructed using our content-aware approach (thick, dashed lines) than those reconstructed using the conventional (thin, solid lines) or sparsity-reweighted (thin, dashed lines) approaches. This trend holds over various input acceleration/undersampling factors and signal-to-noise ratio (SNR) levels (t-b: R = 3, 4, 5, 6). These values correspond to the region around the heart (white box area in Figure 2(c))

Figure 10:

End-systole and end-diastole images are shown side-by-side for the third slice of the ground truth clinical cine scan (a). The conventional (b), sparsity-reweighted (c), and content-aware (d) reconstructions are compared visually for three-times retrospectively undersampled k-space with 16 dB signal-to-noise ratio (SNR). Note that these reconstructions were tuned for the second slice, and are un-modified for this cine reconstruction. Note the degradation in the papillary muscles present in all reconstructions, indicating a limitation of the total variation sparsity used. Peak-signal-to-error (PSER) and mean structural similarity image index (MSSIM) values for the entire time series for the third slice are below each image pair.

Figure 11:

The peak-signal-to-error ratio (PSER) and mean structural similarity image index (MSSIM) values are shown for the retrospectively undersampled spiral continuous cardiac acquisition reconstructed using the content-aware (thick, dashed lines) and the conventional (thin, solid lines) and sparsity-reweighted (thin, dashed lines) approaches. While PSER values are little changed, the structural quality reflected by MSSIM increases in the region around the heart (white box area in Figure 2(d)) for most cases. These quality measures are plotted over various input undersampling factors (t-b: R = 3, 4, 5, 6) and signal-to-noise ratio (SNR) levels.

Figure 12:

The ground truth end-diastole and end-systole phases (a) are shown for the seven-frame continuous cardiac spiral acquisition from Figure 2(d). The conventional (b), sparsity-reweighted (c), and content-aware (d) reconstructions from 16 dB signal-to-noise ratio (SNR) data undersampled by a factor of three are compared visually at these two phases. The end-systolic phase occurs at the time of maximum volume change between frames for this continuous acquisition, so the result using temporal total variation is blurred relative to the ground truth. Peak-signal-to-error (PSER) and mean structural similarity image index (MSSIM) values are below each image.