Magnetic Resonance Imaging 55 (2019) 36–45

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Original contribution

Motion-compensated reconstruction of magnetic resonance images from undersampled data

Daniel S. Weller[a,*,1], Luonan Wang[a,1], John P. Mugler III[b,1], Craig H. Meyer[b,1]

[a] University of Virginia, Charlottesville, VA 22904, USA
[b] University of Virginia School of Medicine, Charlottesville, VA 22908, USA

ARTICLE INFO

Keywords:
Motion correction
Image reconstruction
Body imaging
Model-based reconstruction

ABSTRACT

Magnetic resonance imaging of patients who find difficulty lying still or holding their breath can be challenging. Unresolved intra-frame motion yields blurring artifacts and limits spatial resolution. To correct for intra-frame non-rigid motion, such as in pediatric body imaging, this paper describes a multi-scale technique for joint estimation of the motion occurring during the acquisition and of the desired uncorrupted image. This technique regularizes the motion coefficients to enforce invertibility and minimize numerical instability. This multi-scale approach takes advantage of variable-density sampling patterns used in accelerated imaging to resolve large motion from a coarse scale. The resulting method improves image quality for a set of two-dimensional reconstructions from data simulated with independently generated deformations, with statistically significant increases in both peak signal to error ratio and structural similarity index. These improvements are consistent across varying undersampling factors and severities of motion and take advantage of the variable density sampling pattern.

1. Introduction

Motion affects a wide range of magnetic resonance imaging (MRI) acquisitions. For instance, breathing or respiratory motion can move internal organs like the heart, liver, or kidneys during a scan. Cardiac motion can distort images of the heart as well. Gating or triggering (with a respiratory belt, electrocardiogram, or navigator) can suppress these motion effects, but at the cost of less efficient acquisitions. Inconsistent breathing, arrhythmia, or inaccurate physiological monitoring also can diminish the effectiveness of such methods. Additionally, such methods do not address other sources of motion, such as patient movement (bulk motion). Certain patient populations, including young children, cancer patients, and those with neurological or neurodegenerative disorders have difficulty lying still for extended periods and would be expected to move during longer scans. To deal with all these types of motion, for fast, high resolution MRI, motion-corrected reconstructions are leading the way.

In this work, we are interested in suppressing bulk and respiratory motion that can affect high-resolution abdominal images, such as those used to monitor and plan treatment for liver cancer (e.g., hepatoblastoma) in young children. These images can suffer from intra-frame motion blur due to both bulk motion and respiratory motion. While sedation can address the former, doing so precludes breath-holds, poses risks for young children, and requires an anesthesiologist to be present. The need for an anesthesiologist makes scheduling such exams challenging. Even without sedation, repeated breath-holds may be difficult for these children, limiting the achievable resolution and image quality of abdominal images. Alternative imaging techniques such as X-ray computed tomography pose radiation-related risks to these subjects, motivating un-sedated imaging techniques for pediatric MRI. This work, focusing on intra-image (or intra-frame) motion, complements other recent developments addressing inter-frame motion in body imaging, including in dynamic imaging [1–3], multishot imaging [4], and in 4D flow imaging [5–7].

To handle both bulk and respiratory motion, this work focuses on non-rigid correction of intra-frame motion. This contrasts with rigid or translational motion correction that is common for intra-frame motion using k-space navigators (e.g., autofocus [8]). Parallel imaging and sharpness criteria like gradient entropy extend the autofocus approach to non-rigid motion [9,10]. However, these navigators have limited compatibility with spiral-out pulse sequences that can be used for fast imaging. Another approach, MOCCA [11], senses motion without

* Corresponding author.
E-mail addresses: dweller@virginia.edu (D.S. Weller), lw3mz@virginia.edu (L. Wang), jpm7r@virginia.edu (J.P. Mugler), cmeyer@virginia.edu (C.H. Meyer).

[1] Declarations of Interest: none for Daniel S. Weller and Luonan Wang. Both John P. Mugler III and Craig H. Meyer have received funding from Siemens Healthineers.

https://doi.org/10.1016/j.mri.2018.09.008
Received 22 May 2018; Received in revised form 16 August 2018; Accepted 8 September 2018

navigator sequences, but is primarily used for gating. Inter-frame non-rigid motion correction methods avoiding navigators include non-rigid registration [2–4], k-t FOCUSS [12], MASTeR [13], and BLOSM [14], rank constraints [15,16], and the distance-sensitive regularization in PRICE [17] and other methods [18]. But, using these methods to correct intra-frame motion unnecessarily reduces the amount of data per reconstructed image by a factor equal to the number of groups or "frames", and these methods do not specify how to combine the reconstructed time series into a final image. Alternatively, low-resolution image frames can be used for rigid or non-rigid registration, to resolve motion before reconstructing high-resolution video that takes advantage of spatiotemporal redundancy [19].

Incorporating motion estimation into image reconstruction is a highly flexible motion compensation solution. K-space pre-processing is not flexible enough to correct non-rigid motion [20], and non-rigid image registration post-reconstruction [21,22] cannot address intra-frame motion. Methods such as GRICS [23–25] correct respiratory motion via a parametric model of deformable motion in the reconstruction. GRICS requires specifying respiratory phase or other pseudo-navigational information, but avoids traditional navigation. Instead of constraining our deformable motion to a fixed parametric model, we use multi-scale motion estimation, which was previously applied to image registration [26]. This approach is geared to address arbitrary motion without specific tracking. This model-based reconstruction technique relies on variable sampling densities present in non-Cartesian acquisitions (e.g., spirals) used for highly accelerated imaging. Other methods take advantage of radial sampling to aid motion estimation in time-series imaging [27–31]. Many other works also address motion correction and are described in reviews on medical image registration [32] and motion correction methods [33].

A conference abstract [34] described a single-scale version of our motion-compensated reconstruction. Fig. 1 depicts the reconstruction of an image corrupted by such motion by grouping the undersampled k-space data acquired sequentially in time, and jointly estimating the non-rigid motion present for each group along with the reconstructed image. That framework is extended here to use multi-scale motion estimation and parallel imaging for undersampled non-Cartesian k-space data. This paper also includes a comparison against an existing intra-frame motion correction technique (non-rigid autofocus with locally linear translations [9]) and an illustration of the multi-scale motion estimation technique. This paper also illustrates the advantage of non-constant sampling density for this method. These experiments operate with synthetic motion, so a ground truth image with identical contrast and resolution is available for comparison. An additional experiment with real motion indicates that the proposed method can suppress some motion even when correcting motion for each spiral readout individually.

The next section describes the multi-scale non-rigid intra-frame motion compensation method in detail. Section 3 describes the data and evaluation criteria used to validate and compare this method. Experimental results follow in Section 4 and are discussed in Section 5.

2. Theory

When acquiring a single image (or a frame of a video or time series), if the time-scale of intra-frame motion is comparable to the acquisition time for that image, and if the spatial-scale of that motion is comparable to the resolution of the image, that motion can introduce blur or other artifacts. These artifacts are caused by the image content (e.g., organs or tissue) being in different positions for each k-space readout. To reconstruct images without these artifacts, this method adjusts the forward model to include these intra-frame displacements, modeled as non-rigid motion. Ideally, we would estimate such a deformation for every single k-space readout (like autofocus can via navigation). In practice, individual readouts do not contain sufficient data to calculate these deformations. Instead, we group readouts consecutive in time into a single subset and assign all the readouts in a given subset the same non-rigid motion. The group size should be selected to reflect the time scale of discernable motion; using too many groups would unnecessarily trade off data availability for minimal motion reduction.

Defining an unknown "source" image as I to be the target of the reconstruction, we let the transformation $T(w_g)$ map the source image to the deformed state corresponding to the gth group of k-space m_g. This transformation $T(w_g)$ corresponds to a b-spline-based local interpolation of an image onto a new grid that corresponds to adding the displacement vectors w_g to the original (Cartesian) grid locations. Each displacement vector contains coordinate offsets in each direction, nominally on a pixel-by-pixel basis. However, to constrain the motion to be smooth, and limit the number of unknowns, we use b-splines on a lower-resolution grid to describe the coordinate offsets with a reduced set of coefficients. So,

$$[w_g]_x(x,y) = \sum_{i,j} c_{g,x}[i,j]\beta\left(\frac{x}{m_x} - i\right)\beta\left(\frac{y}{m_y} - j\right), \quad (1)$$

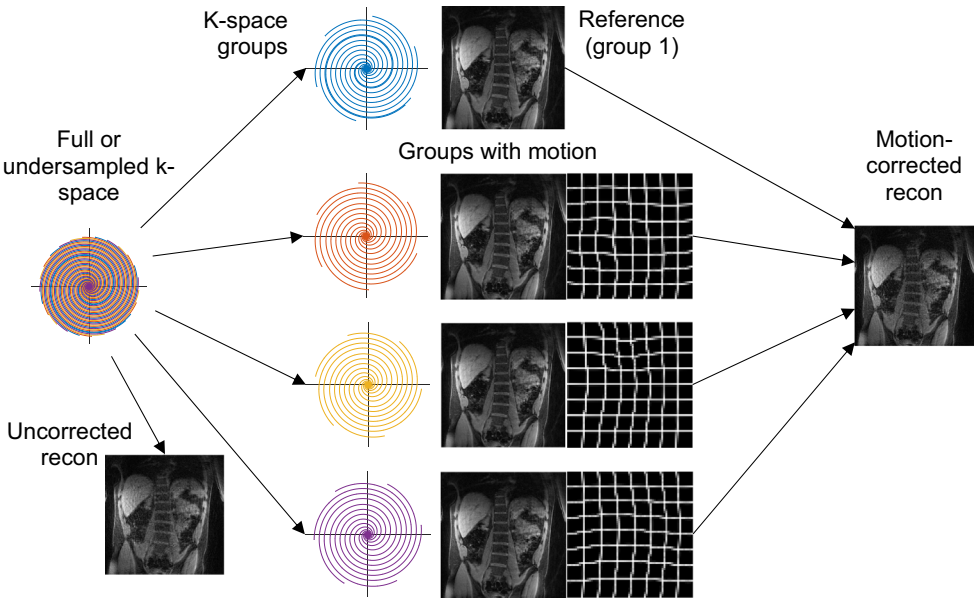

Fig. 1. In this work, full or undersampled k-space corresponding to a single image (or frame) is simulated with non-rigid motion in each of four subgroups (different color spiral readouts). Without motion correction, an uncorrected reconstruction produces the blurred image depicted in the bottom left. The motion-corrected reconstruction accounts for different non-rigid motion in each group and produces an image consistent with the reference group (nominally group 1) from all acquired k-space. The deformation maps and deformed images are shown for groups 2–4 to depict the extent of non-rigid motion in this example; this prior knowledge is not used in the reconstruction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$[w_g]_y(x,y) = \sum_{i,j} c_{g,y}[i,j]\beta\left(\frac{x}{m_x} - i\right)\beta\left(\frac{y}{m_y} - j\right) \quad (2)$$

represent the x- and y-displacements, respectively, of the deformation field of the gth group at coordinates (x,y), in terms of the low-resolution cubic b-spline function $\beta()$ with stretching factors m_x and m_y, and with coefficient matrices $c_{g,x}$ and $c_{g,y}$. The amount of resolution reduction depends on these stretching factors (see the discussion of the multi-scale approach in Section 2.2). Corresponding expressions for 3-D motion have three-dimensional cubic b-splines and a third set of coefficients $c_{g,z}$. For simplicity, we continue the development considering 2-D motion on a 2-D image. We assume the first group of k-space is consistent with the source image we are reconstructing and hold $w_1 = 0$. Then, a single-image reconstruction from G groups of k-space would attempt to solve the following minimization problem:

$$\underset{I,w_1=0,c_{2,x},c_{2,y}\ldots,c_{G,x},c_{G,y}}{\arg\min} \sum_{g=1}^{G} \|m_g - F_g T(w_g(c_{g,x}, c_{g,y}))I\|^2 \quad (3)$$

The matrix F_g is the Cartesian or non-Cartesian Fourier transform, subsampled to produce the k-space values for just the readouts of the gth group. In practice, this reconstruction is ill-posed, since we do not constrain the deformation coefficients, and the entire set of k-space may be insufficient to specify a unique image on its own. We use two approaches to address this issue: regularization and multi-scale (or multi-resolution) motion correction. The former enforces smoothness of the motion or deformation field while taking advantage of parallel imaging redundancy (when multi-channel data are available). The latter takes advantage of variable sampling density with non-Cartesian spiral or radial trajectories.

2.1. Regularization

This paper describes two forms of regularization: a quasi-quadratic constraint on the motion and a null space penalty for SPIRiT [35] parallel imaging (if applicable). The first constrains the b-spline-coefficient-based finite differences approximation to the Jacobian matrix (first-order partial derivatives) of the deformation vector field. This penalty, described in [26], is piecewise-smooth, and limits motion changes that would violate invertibility or produce extreme changes in the volumes of local structures (as reflected by the determinant of the Jacobian matrix). We use first-order backward difference-based approximations

$$p_{g,xx}[i,j] \approx \frac{\partial c_{g,x}[i,j]}{\partial x},\ p_{g,xy}[i,j] \approx \frac{\partial c_{g,y}[i,j]}{\partial x},\ p_{g,yx}[i,j] \approx \frac{\partial c_{g,x}[i,j]}{\partial y},\ p_{g,yy}[i,j] \approx \frac{\partial c_{g,y}[i,j]}{\partial y} \quad (4)$$

for the x- and y-derivatives of the x- and y-displacement coefficients. The penalty function is

$$\phi(p) = \begin{cases} \frac{1}{2}(p-\tau_1)^2, & p \le \tau_1, \\ 0, & \tau_1 < p \le \tau_2, \\ \frac{1}{2}(p-\tau_2)^2, & p > \tau_2. \end{cases} \quad (5)$$

This piecewise-smooth penalty function ensures that the change in motion from one location to the next is neither so small (negative) nor so large (positive) that one pixel would not move too close to a neighbor. Appropriate values for the τ's use small $\tau_1 < 0$ and much larger $\tau_2 \gg -\tau_1$ for p_{xx} and p_{yy}, and relatively small $\tau_1 = -\tau_2$ for p_{xy} and p_{yx}. This choice permits small decreases in volume and larger increases in volume. Decreases in volume should be constrained more aggressively, as too significant decreases lead to ill-conditioning or even non-invertibility of the deformation. Generalization to three-dimensional motion is straightforward using a 3×3 matrix of partial derivatives for each direction and each displacement [26]. The motion-regularized form of the motion-compensated reconstruction is

$$\underset{I,w_1=0,c_{2,x},c_{2,y}\ldots,c_{G,x},c_{G,y}}{\arg\min} \sum_{g=1}^{G} \|m_g - F_g T(w_g(c_{g,x}, c_{g,y}))I\|^2 + \lambda_1 \sum_{g,i,j} \phi(p_{g,xx}[i,j]) + \phi(p_{g,xy}[i,j]) + \phi(p_{g,yx}[i,j]) + \phi(p_{g,yy}[i,j]) \quad (6)$$

The deformations w_g and partial derivatives $p_{g,xx}$, $p_{g,xy}$, etc. are linear functions of $c_{g,x}$ or $c_{g,y}$. The regularization tuning parameter λ_1 trades data-consistency for this motion constraint. We hold this parameter fixed to $\lambda_1 = 20$. The flat part of this regularizer ensures that setting λ_1 too large does not change the motion estimates too much.

However, motion-constraining regularization is insufficient in the case where undersampling may be present, such as in accelerated imaging. In the parallel imaging case, the image I is actually a set of images, one for each channel in the receive array coil. Then, the linear relationship between adjacent k-space frequencies across coils in SPIRiT parallel imaging yields a null space constraint in either the image domain or frequency domain [35]. Adding to the optimization problem a quadratic penalty on the difference between the set of coil images and the per-pixel SPIRiT combination of that set minimizes violation of this null space constraint. The parallel imaging-regularized reconstruction is

$$\underset{\boldsymbol{I},w_1=0,c_{2,x},c_{2,y}\ldots,c_{G,x},c_{G,y}}{\arg\min} \sum_{g=1}^{G} \|\boldsymbol{m}_g - F_g T(w_g(c_{g,x}, c_{g,y}))\boldsymbol{I}\|_F^2 + \lambda_2 \|SPIRiT(\boldsymbol{I}) - \boldsymbol{I}\|_F^2 + \lambda_1 \sum_{g,i,j} \phi(p_{g,xx}[i,j]) + \phi(p_{g,xy}[i,j]) + \phi(p_{g,yx}[i,j]) + \phi(p_{g,yy}[i,j]). \quad (7)$$

The k-space readout groups $\boldsymbol{m}_g$ and image set $\boldsymbol{I}$ in Eq. (7) are matrices, with each column representing a different coil channel. Thus, the Frobenius norm $\|\cdot\|_F$ replaces the vector 2-norm. This approach is also compatible with other forms of regularization on the image set, such as transform-domain sparsity, but such regularizers are well-studied and beyond the scope of this paper. Although intra-frame motion affects the coil sensitivities, and hence the SPIRiT operator calibration, the very low resolution of the SPIRiT kernel and the calibration data limits the severity of motion artifacts on the kernel. Fig. 2 portrays calibrated kernels from real breath-held and free-breathing data. This comparison suggests that real motion produces small differences in these calibrations. As in conventional SPIRiT, we set $\lambda_2 = 1$.

2.2. Multi-scale processing

In the literature, including [26], multi-scale or multi-resolution motion estimation methods are more robust and efficient than single-scale, high-resolution motion estimation. This is due to a number of factors, including the nonconvexity of motion estimation, the diminishing effect of motion on low-resolution images, and the reduced problem size (fewer motion coefficients for each group) at coarse scales. Multi-scale processing can also motivate employing variable-density sampling strategies that undersample higher frequencies more than lower frequencies. Such strategies are easily realized with spiral or radial k-space trajectories, due to the repeated sampling at/near DC. This greater sampling density reduces aliasing energy in low-resolution images, facilitating accurate motion estimation at the coarse scale. Thus, to extend the framework to multi-scale processing, we estimate a set of coefficients $\{c_{g,x}^{(s)}[i,j], c_{g,y}^{(s)}[i,j] : \forall\, i,j\}$ for the scale s and vary the scale in coarse-to-fine order, increasing the number of coefficients in each successive level. Starting at the lowest resolution, or coarsest scale, means we begin with having the fewest motion coefficients to measure. At this scale, we use the largest stretching factors m_x and m_y. Then, after the estimation of the coarse-scale motion (and the associated image) has stabilized, we increase the resolution by a factor of two (dividing m_x and m_y by two, and doubling the number of motion

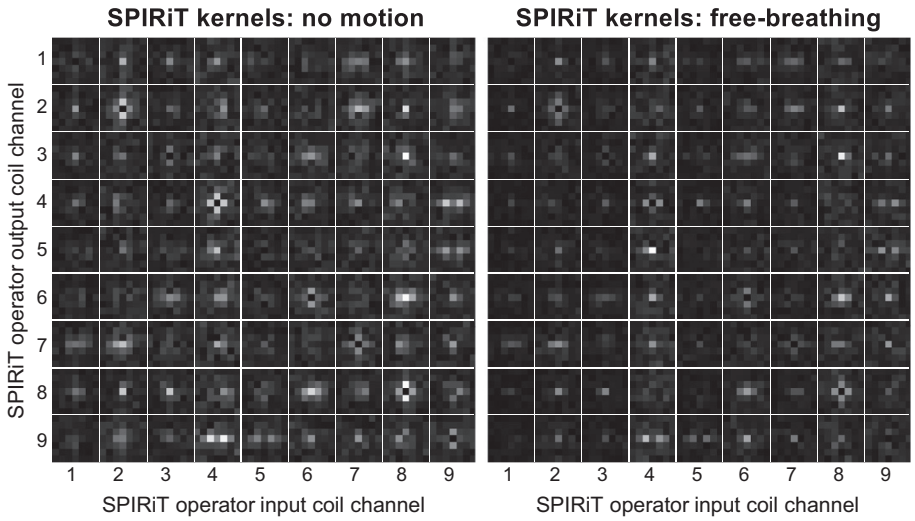

Fig. 2. SPIRiT kernels (7 × 7) are calibrated for the nine-channel virtual coils obtained for both breath-held (left) and free-breathing (right) real abdominal imaging data with real motion used in this work. The SPIRiT operator describes kernels for every input/output pair of coil channels. Each row represents an output channel, and each column represents an input channel. These kernels are very similar between the two cases, suggesting that SPIRiT kernel calibration with very low resolution kernels is reasonably robust to motion artifacts.

coefficients in each group) and restart the estimation process. We initialize the estimates with the image from the previous stage (image resolution does not change) and sets of higher-resolution deformation coefficients for each group interpolated using cubic b-splines from the previous lower-resolution result. Increasing the resolution of just the deformation coefficient sets means that the initial stages are better at capturing bulk motion of the subject and the organs, while the later stages mainly refine motion estimates on the boundaries of different local motion areas. For the parallel imaging case, we use the same SPIRiT calibration at each level, since the image remains the same size. To summarize the algorithm:

1. Initialize all groups' motion to zero, and reproduce a blurred initial image from the first group of k-space data and (if available) parallel imaging. A single SPIRiT iteration (multi-channel), or a non-iterative density-compensated gridded reconstruction (single-channel), produces an initial image with some aliasing artifacts. Start motion estimation at the coarsest scale.
2. Estimate deformation coefficients $\{c_{g,x}^{(s)}[i,j], c_{g,y}^{(s)}[i,j] : \forall\, i,j\}$ at the current scale s for each group via nonlinear conjugate gradients [36] with a second-order Newton step size.
3. Update the image using k-space for all the groups and (if available) SPIRiT regularization via the LSQR method [37].
4. Repeat steps 2–3 at the current scale until either the desired number of alternating iterations has been reached, or stopping criteria are achieved (the change in objective function value and the change in motion coefficients and image values are both below a fixed threshold).
5. Compute motion coefficients for a finer scale (twice the resolution) via cubic b-spline interpolation and repeat steps 2–4 for the finer scale. Repeat this process until the desired scale of motion is reached.

This algorithm is similar to the regularized multi-scale method in [26], but it is modified significantly for MRI reconstruction, with multiple groups of k-space instead of separate images.

3. Material and methods

The objectives of these experiments are (1) to ascertain the effect of multi-scale processing on the quality of motion estimates, (2) to compare the motion-compensated reconstructions of abdominal MR images against existing motion-correcting reconstructions, and (3) to demonstrate the proposed method on real motion. To accomplish these objectives, the experiments use breath-held and free-breathing pilot data sets acquired from consented human subjects under a protocol approved by the Institutional Review Board of the University of Virginia (protocol #17557). All reconstructions are evaluated via visual comparison. The motion-corrected reconstructions from the breath-held data with synthetic motion added retrospectively are also evaluated according to the peak signal to error ratio (PSER) and the structural similarity (SSIM) index perceptual quality metric [38].

3.1. Data sets

The data were acquired on a Siemens Avanto 1.5 T scanner using a T1-weighted gradient echo (VIBE) pulse sequence with a 3D stack of spirals k-space trajectory. The original acquisition prescribed 20 kz partitions. After taking an inverse Fourier transform in the z-direction, the extreme two slices on both ends are removed to mitigate wrap-around aliasing unrelated to motion, resulting in 16 coronal slices, each 5 mm thick. For each kz-partition, the acquired data contains N = 74 spiral interleaves from 12 coil channels (of a Siemens spine and body array coil) with a nominal field of view (FOV) of 400 × 400 mm and a base resolution of 1.56 × 1.56 mm. These dual density interleaves were acquired with a 1/FOV spacing at the center of k-space and N/FOV spacing for the outer part per interleaf. For these data sets, the data were acquired in 30 s (a single breath hold for the breath-held scans). Two additional interleaves were acquired for each kz partition to support automatic off-resonance correction [39]. The same scan prescription was used for both the breath-held and free-breathing (real motion) acquisitions.

From a breath-held ground truth data set, we introduced synthetic motion and resampled k-space along the same readout pattern. Since the emphasis of this research is on correcting intra-frame motion, we introduced different synthetic motion for each 1/4 of k-space, yielding a blurred reconstruction without motion correction. To form this synthetic motion, and to avoid the “inverse crime” of using our motion model to synthesize the motion, we created high-resolution, smooth distortions on a voxel-wise basis around the locations of internal organs like the liver and kidneys that are known to move in free-breathing and pediatric patient imaging. The free-breathing data have their own real motion and do not have any additional motion added.

3.2. Evaluation criteria

For the data incorporating known synthetic motion, direct comparison of the reconstructed image to a motion-free ground truth is possible. Given the ground truth, we compare image quality using the peak signal to error ratio (PSER) defined as $PSER = 20\log_{10}\frac{\sqrt{N}\cdot\|x_{true}\|_\infty}{\|x - x_{true}\|_2}$

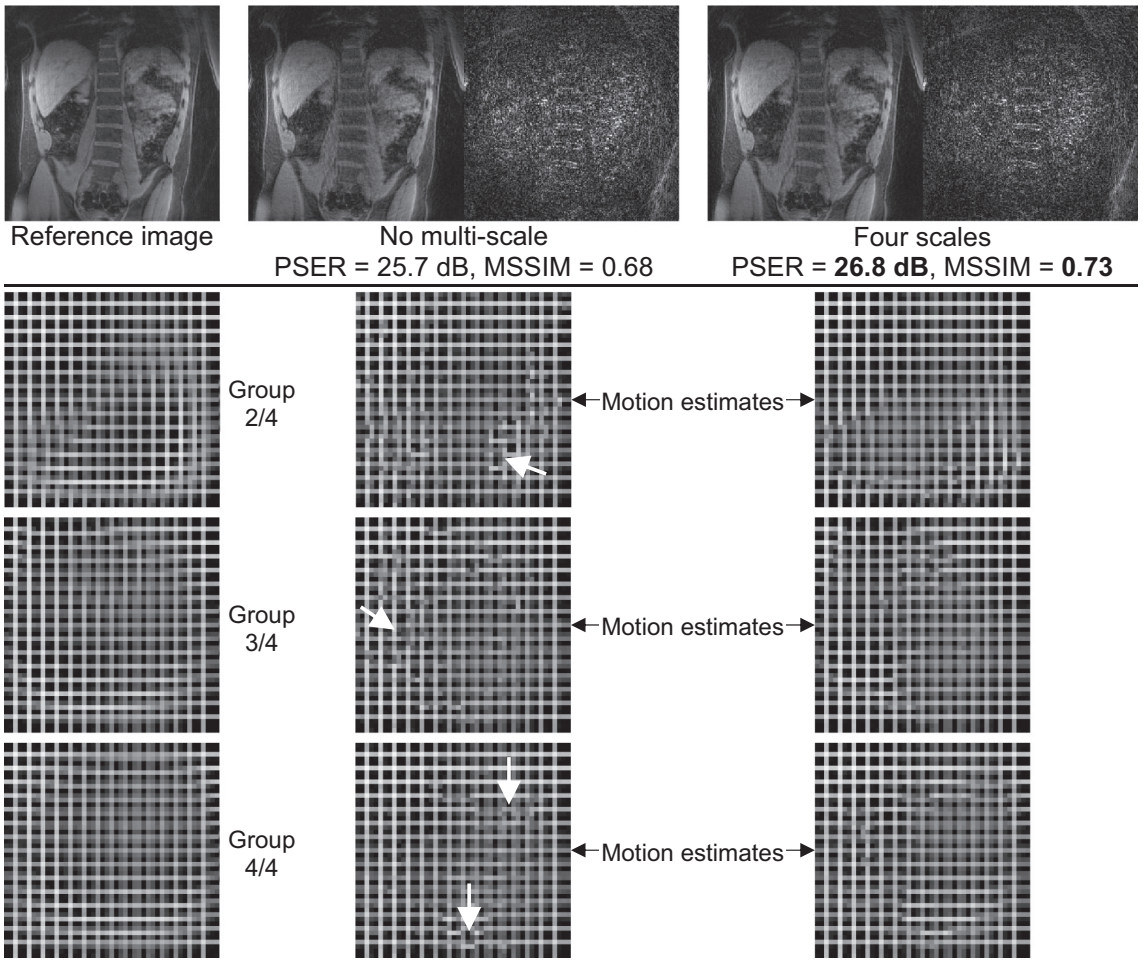

Fig. 3. A 256 × 256 pixel reference image (top-left) is perturbed by four groups of simulated motion during acquisition (group 1 is the reference). This experiment compares motion estimation with and without multi-scale processing. Both methods find mappings from the original image to the transformed data for each group. In the top part, the final reconstruction and difference image using no multi-scale processing shows more artifacts than the reconstruction with four levels of multi-scale processing. The motion estimates for groups 2–4 are shown below. The white arrows (middle column) indicate spurious errors in the motion estimates generated without multi-scale processing.

and the structural similarity index (SSIM). This second measure is sensitive to image sharpness, which is the criterion measured by gradient entropy [40] for competing methods like autofocus navigation [9]. Reconstructions with either synthetic or real motion are also compared visually.

4. Results

The first experiment concerns the validation of the multi-scale reconstruction method on synthetic (simulated) non-rigid motion. Starting from 2×-undersampled dual-density k-space data corrupted by synthetic motion (divided into four groups), we used both our single-scale method [34] and the proposed multi-scale motion estimation to measure the non-rigid motion consistent with the sampled k-space. We then assessed the difference between the reconstructed images and the reference (uncorrupted) image. The original image as well as the computed results for both single- and multi-scale motion estimation are shown in Fig. 3. The multi-scale motion estimation captures bulk motion at the coarsest scale, and it provides a mostly accurate estimate of the motion at the fine scale. The single-scale (no multi-scale) motion estimation produces small motion errors throughout the imaging plane. The four-scale motion estimates are much smoother and more consistent with the true deformations. The largest errors in the reconstructed image estimate are present around the lumbar vertebrae. The area around the liver and outer wall of the body have relatively small errors. Details like the venous structure of the liver have little error as well.

The second experiment compares our motion correction to an existing non-rigid motion correction method that relies on navigator data, again using simulated motion with a ground truth. The existing autofocus method [9] uses parallel imaging to approximate non-rigid motion with locally linear translations. This approximation is consistent with a low-resolution deformation field. To reproduce images from incomplete measurements, this autofocus method is combined with SPIRiT reconstruction, the same parallel imaging reconstruction used to deal with incomplete data in our method. As our method does not require navigators to measure in-plane motion, demonstrating similar or improved performance on undersampled k-space data helps justify the use of the proposed reconstruction for non-rigid motion correction. To begin, we fully sample k-space and interleave the 74 readouts/kz-partition into four groups, three of which are corrupted by motion. Employing the same SPIRiT reconstruction as before yields a highly blurred image with significant motion-amplified aliasing artifacts. Then, we reconstruct the image using our method, alternating SPIRiT and multi-scale motion estimation. We also synthesize kx and ky butterfly navigators in k-space and employ non-rigid group-wise autofocus reconstruction, with and without SPIRiT parallel imaging, for comparison. Figs. 4 and 5 depict all these reconstructions for two of the 16 slices. Each figure also includes the original (ground truth) image, along with difference images (relative to the ground truth) for the motion-

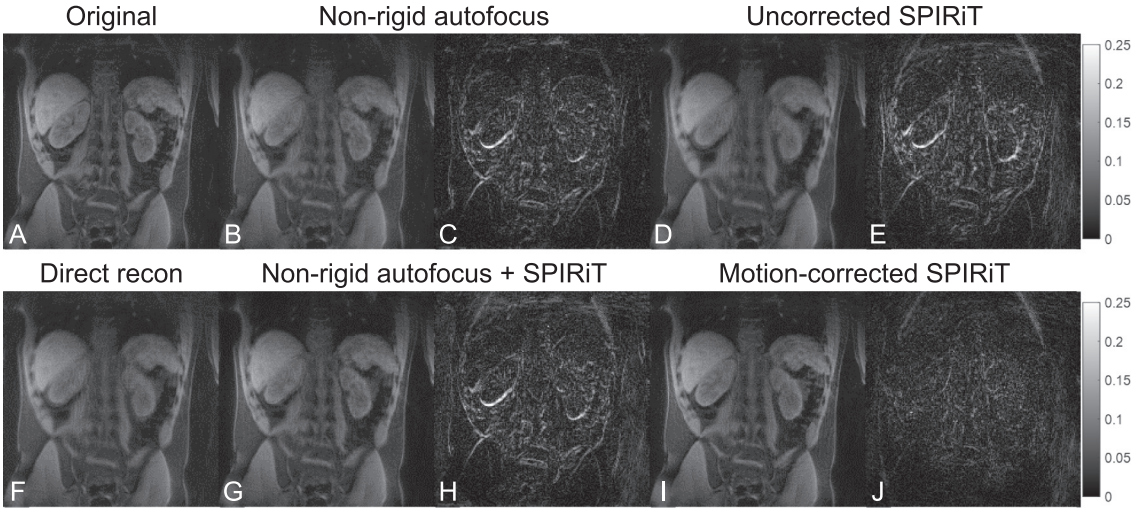

Fig. 4. From fully-sampled k-space data, the conventional SPIRiT reconstruction (D) produces a distorted image even with relatively small motion. Our motion-corrected SPIRiT reconstruction (I) addresses the issue of non-rigid motion, producing an image with substantially reduced aliasing energy (difference image: J). our method outperforms (non-rigid) autofocus with locally linear translations, with (G) or without (B) SPIRiT. Neither autofocus reconstruction reduces aliasing as effectively (difference images: C, H). Our method also does not rely on navigators to track the motion.

correcting reconstructions. The image reconstructed with our method appears a bit blurred versus the original image, but it is a significant improvement over the blurring incurred by the other reconstruction methods. The outlines and most fine-scale details of our reconstruction match those of the uncorrupted reconstruction. The image quality is also improved over the other reconstructions, with significantly reduced coherent structure in the difference image.

From the fully sampled case, we clearly observe improved motion correction with the proposed method. In Fig. 6, we depict these reconstructions for one of the slices, with undersampling factors (R) of $R = 2$ and $R = 3$. The proposed reconstruction has noticeably reduced errors versus the other methods in both cases. The next part of this experiment measures both mean PSER and MSSIM values across all sixteen slices. We first measure the statistical significance over all comparisons using single-factor ANOVA. Since these p-values are significant, we perform post-hoc matched-pairs *t*-tests with Bonferroni correction on the PSER and SSIM values depicted in Fig. 7. For $R = 1$, the fully sampled case, we observe a significant improvement in image quality using our method according to both metrics, and the difference remains significant for $R = 2$ and $R = 3$. The p-values for both are listed in Table 1. Thus, we observe that our method outperforms the others as undersampling increases, although we become less competitive with SPIRiT at high undersampling. We surmise undersampling-related errors dominate the reconstruction error at high acceleration.

Next, Fig. 8 shows mean SSIM and PSER values across the 16 slices for different intensities of synthetic motion. We measure the statistical significance of these means being different for each motion intensity using the same tests as before, and we find statistically significant improvement for our joint reconstruction in most cases (p-values listed in Table 2). At very high levels of motion, the proposed method becomes less competitive, suggesting that our method has an upper limit in how much motion can be effectively corrected.

Fig. 9 portrays the difference between motion/4 reconstructions for both constant- and dual-density spiral trajectories. Our method performs significantly better with dual-density data than constant-density data, as the low-resolution information used to initialize the coarsest scale of motion estimation does not suffer from the aliasing energy present with constant-density data.

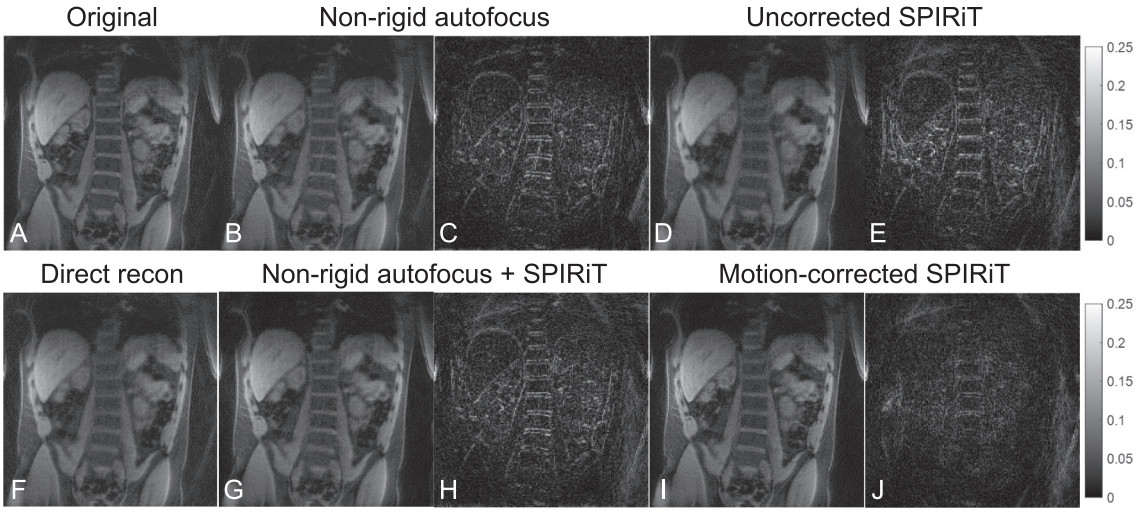

Fig. 5. This figure portrays the same reconstructions and difference images as Fig. 4, for another of the 16 slices in the data set with synthetic motion. The specific motions for the groups in this slice are different from the motions for the other slice. The improved image quality over uncorrected and autofocus-guided SPIRiT reconstruction supports the generality of our method.

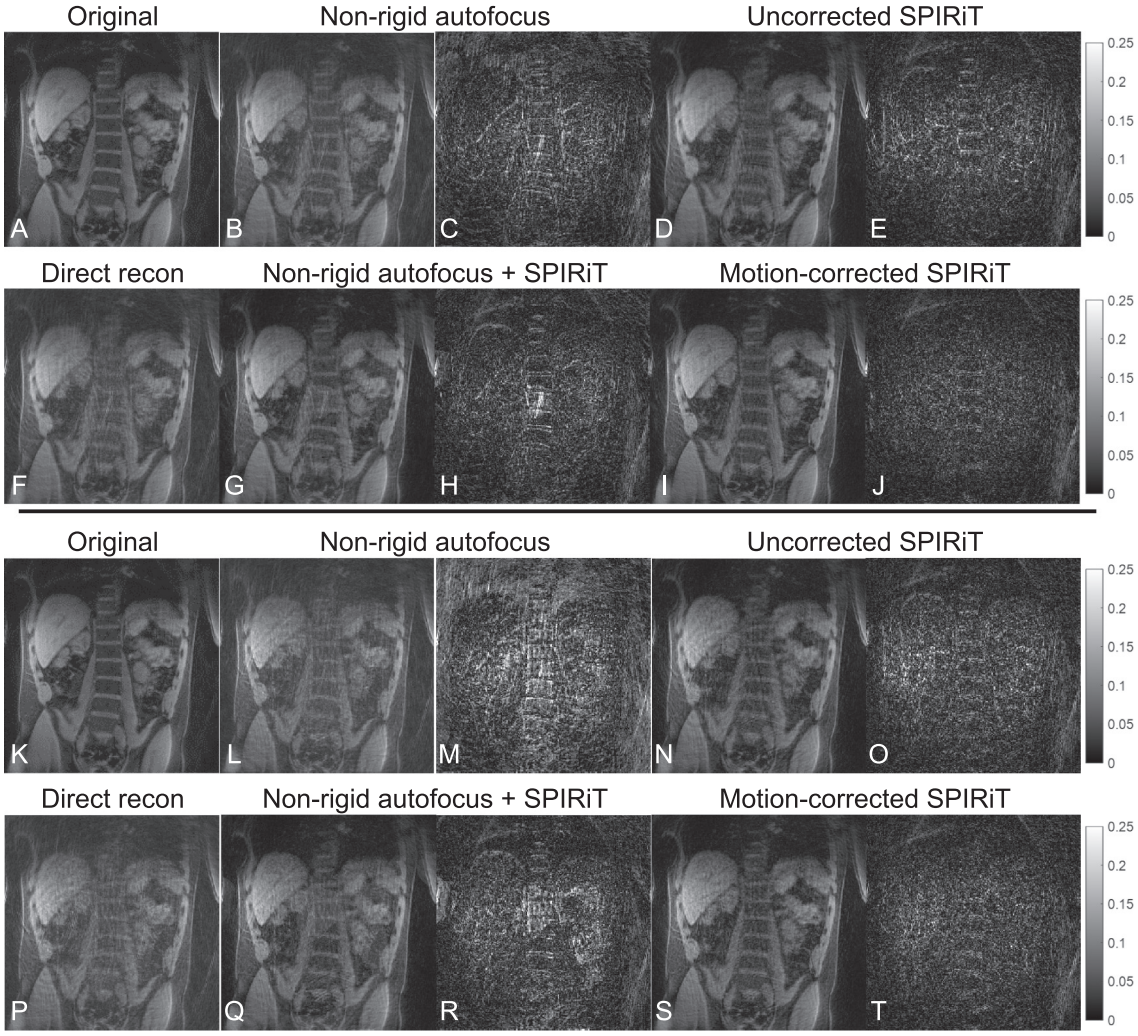

Fig. 6. This figure depicts the reconstructions for the same slice in Fig. 5, with increased undersampling (top two rows: 2× undersampling; bottom two rows: 3× undersampling). As the undersampling factor increases, the joint reconstruction (I, J, S, T) shows less degradation than the other methods.

Fig. 10 depicts the result of applying the proposed motion-compensated reconstruction to free-breathing data (with real motion) acquired using the same paradigm as the breath-held data used previously. The subject was encouraged to breathe normally during the 30-s scan. As a result, the direct reconstruction of these data is highly blurred by motion, compared to the ideal breath-held image in the figure. The proposed reconstruction sharpens some of the low-resolution features of the reconstruction, but significant artifacts remain.

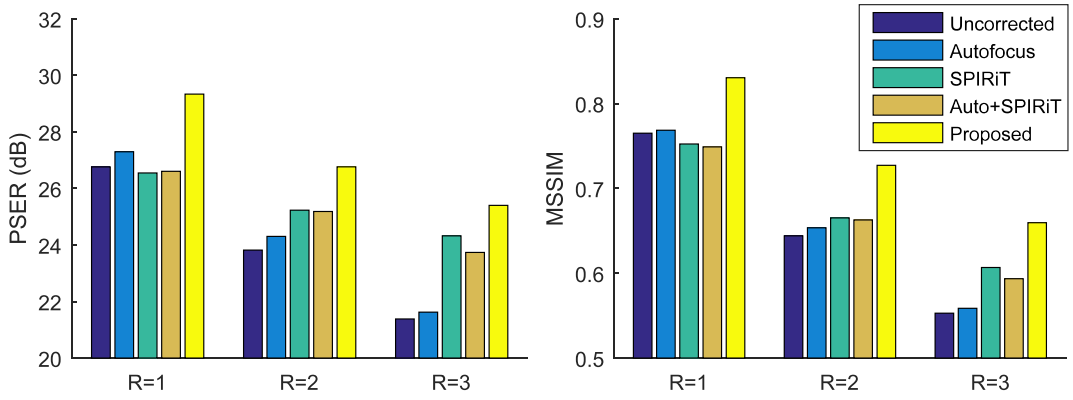

Fig. 7. Varying the degree of undersampling (R), these charts plot the peak-signal-to-error ratio (PSER) and mean structural similarity (MSSIM) index, averaged over the 16 slices with independently generated synthetic motion. These plots indicate a significant increase in PSER and MSSIM for the proposed joint reconstruction method over autofocus, SPIRiT, and autofocus + SPIRiT, as well as over an uncorrected direct reconstruction. Table 1 lists the p-value statistics for both repeated measures ANOVA and *t*-tests for the proposed method versus the others, correcting for multiple comparisons.

Table 1
This table describes the statistical significance of the joint reconstruction (JR) improvement over uncorrected direct reconstruction (UC), autofocus (AF), SPIRiT (SP), and autofocus + SPIRiT (AS), for varying degrees of undersampling, whose average peak-signal-to-error ratio (PSER) and mean structural similarity (MSSIM) index values are depicted in Fig. 6. Significant values with $p < 0.05$ are in bold.

PSERs MSSIMs	p (ANOVA)	p (JR > UC)	p (JR > AF)	p (JR > SP)	p (JR > AS)
R = 1	**1.2e−5**	**2.3e−4**	**0.0055**	**5.6e−5**	**8.3e−5**
	3.4e−6	**5.3e−4**	**0.0012**	**2.4e−5**	**9.9e−6**
R = 2	**2.3e−9**	**1.2e−8**	**4.0e−7**	**0.0025**	**0.0017**
	1.5e−6	**2.4e−6**	**3.1e−5**	**6.0e−4**	**3.3e−4**
R = 3	**1.5e−19**	**9.9e−9**	**9.9e−9**	**0.031**	**1.6e−4**
	5.1e−11	**1.0e−8**	**1.2e−8**	**0.0024**	**7.9e−5**

5. Discussion

The first experiment depicts the advantage of multi-scale motion estimation, a core component of our method. Multi-scale estimation is beneficial mainly due to the diminished effect of motion on lower resolution images and the reduced problem size. As we are measuring fewer motions spaced farther apart from each other, the constraints on consecutive deformation coefficients to preserve invertibility are less severe, and there are fewer coefficients to estimate. This strategy mitigates the problem of finding an appropriate region of attraction for the full-scale nonconvex motion estimation problem by capturing the large motions first. The resulting motion estimates are much smoother and more consistent with their neighbors than the single-scale estimation, likely due to the interpolating behavior of the multi-scale method when transitioning between scales. The resulting image benefits from this smoothness of motion estimation, causing the reconstruction error to be more noise-dominated than in the single-scale case.

The second experiment in this paper demonstrates the superiority of joint motion estimation and image reconstruction versus existing methods such as autofocus-based navigation. Our reconstruction has some obvious advantages: the variety of motion estimates across the image is not constrained to the number, size, and location of coil channels available and the heterogeneity of their sensitivity profiles. Working with real body array coils where only several coils may effectively illuminate a particular slice means that many of the candidate motion estimates from autofocus-based navigation are modeling local noise variations instead of local linear translations. The autofocus method likely would perform better in arrays with more variability and greater receiver sensitivity in the image regions of interest. By testing multiple slices separately with different independently generated motions for each one, we show the generality of our reconstruction method and establish the statistical significance of improvement by our method in terms of both squared error and perceptual quality measures.

Starting with fully sampled data separates the effects of undersampling and motion on the reconstruction, isolating image quality differences due to motion from those due to sampling. Introducing undersampling into this experiment afterwards suggests the image quality improvement remains for moderate levels of undersampling. However, at high enough accelerations, all the SPIRiT-based motion-compensated reconstructions should have similar performance, as undersampling-related errors begin to dominate. By varying the motion severity, we also demonstrate that our method can address both smaller and larger motions. The largest (/2) motion listed in the table corresponds to an average Euclidean displacement of > 5 mm, which is over three pixels at base resolution, and a maximum displacement of 2.9 cm, suggesting our method can handle particularly large motions that we may observe in uncooperative patients or un-sedated young children. As the motion severity increases, our iterative approach tends to degrade slower than existing methods, likely due to the multi-scale process used to handle large motions. The reconstruction error for our method is mostly noise, likely due to using parallel imaging without image-based regularization. Using analytical or data-driven image models would mitigate the noise amplification and permit accurate motion estimation at even greater accelerations with improved initialization at the coarsest scale.

However, the illustration in Fig. 10 using real breathing motion suggests more work is needed to obtain images comparable to breath-held data with this particular acquisition. Since the acquisition and breathing motion are in reality three-dimensional, our 2-D correction is inherently limited in its ability to suppress motion artifacts. Furthermore, the readouts for all 20 kz partitions for a single spiral interleaf rotation angle takes 400 ms, mandating motion estimation be carried out for each interleaf individually. This reconstruction uses as many groups as there are leaves (74) and achieves some motion suppression without substantial sampling-related aliasing. Grouping interleaves differently, and taking advantage of redundancy in the breathing

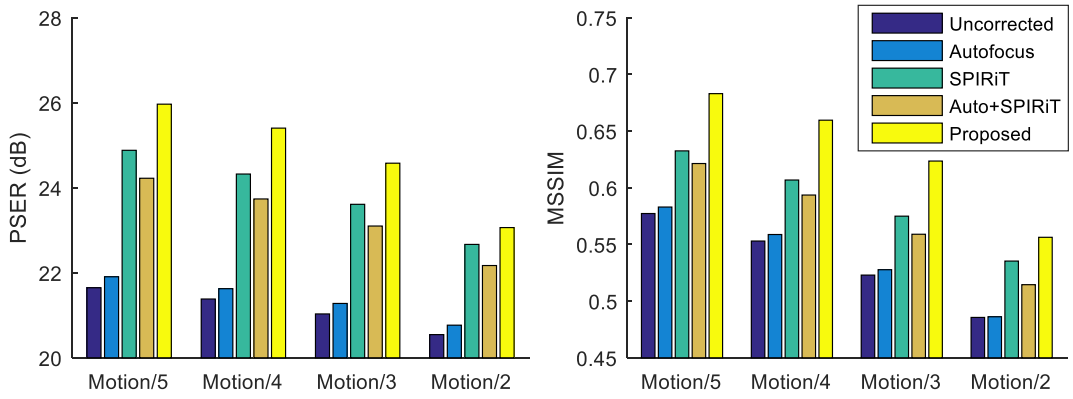

Fig. 8. Varying the severity of motion, these charts plot the peak-signal-to-error ratio (PSER) and mean structural similarity (MSSIM) index, averaged over the 16 slices with independently generated synthetic motion. These plots indicate a significant increase in PSER and MSSIM for the proposed joint reconstruction method over autofocus, SPIRiT, and autofocus + SPIRiT, as well as over an uncorrected direct reconstruction. Table 1 lists the p-value statistics for both repeated measures ANOVA and *t*-tests for the proposed method versus the others, correcting for multiple comparisons.

Table 2
This table describes the statistical significance of the joint reconstruction (JR) improvement over uncorrected direct reconstruction (UC), autofocus (AF), SPIRiT (SP), and autofocus + SPIRiT (AS), for varying motion severities, whose average peak-signal-to-error ratio (PSER) and mean structural similarity (MSSIM) index values are depicted in Fig. 8. Significant values with $p < 0.05$ are in bold.

PSERs MSSIMs	p (ANOVA)	p (JR > UC)	p (JR > AF)	p (JR > SP)	p (JR > AS)
Motion/5	**6.6e−24**	**9.9e−9**	**9.9e−9**	**0.012**	**9.9e−6**
	5.4e−12	**1.0e−8**	**1.0e−8**	**0.0019**	**8.7e−5**
Motion/4	**1.5e−19**	**9.9e−9**	**9.9e−9**	**0.031**	**1.6e−4**
	5.1e−11	**1.0e−8**	**1.2e−8**	**0.0024**	**7.9e−5**
Motion/3	**7.1e−14**	**9.9e−9**	**1.0e−8**	0.15	**0.0058**
	4.7e−9	**3.7e−8**	**1.1e−7**	**0.014**	**4.2e−4**
Motion/2	**7.3e−8**	**3.7e−6**	**2.6e−5**	0.90	0.29
	4.7e−5	**3.4e−4**	**3.9e−4**	0.69	0.081

motion would enable more effective grouping of spiral readouts, reducing and simplifying the motion estimation problem. Effective strategies for such 3-D sampling will be the subject of further research.

A disadvantage of our method is that it relies on low aliasing in the low-resolution acquisition for adequate motion estimation. The significant quality degradation in Fig. 9 with a constant density spiral readout is likely due to aliasing corrupting even the coarse scales of motion estimation. However, threefold undersampling is sufficient to make up for the oversampling necessary to achieve full-FOV density in the center of k-space. We intend to exploit these model-based techniques in future work.

Another disadvantage is the question of grouping k-space data. It is not known a priori how rapidly motion changes in a real acquisition. This means that other methods must be used to detect significant motion changes, to determine a suitable grouping of k-space. MOCCA and similar methods use parallel imaging to identify periods of significant motion, as motion alters the relative intensities across coil channels. Such measures should suffice in this setting as well. In addition to establishing proper timing, the number of groups is effectively a trade-off between the time resolution of motion tracking and the amount of information available to resolve each non-rigid motion. We can control this trade-off through the multi-scale approach by limiting the spatial resolution of the coarsest scale. As reducing the number of readouts per group tends to affect the aliasing energy less near the center of k-space for spiral trajectories, such a technique would permit extending our work to larger numbers of groups if necessary (as in the real breathing motion experiment). Also, MOCCA or similar tracking can point out similar motions disjoint in time, permitting merging groups reflecting repetitive or periodic motion (e.g., normal breathing) to limit the number of motion estimates as well.

For dealing with 3-D imaging, self-navigation requires 3-D k-space trajectories (see [41] for an early overview) capable of capturing movement in all three dimensions over time. Such an acquisition would extend our reconstruction to 3-D motion correction through the three-dimensional spline representation noted in Section 2. However, such motion estimates likely would have reduced spatial resolution versus the 2-D estimates used here to keep the problem size manageable. Another approach would introduce infrequent kz-navigators into the stack-of-spirals pulse sequence and use these navigators to pre-correct 3-D k-space for displacements in the z-direction before taking an inverse Fourier transform in the z-direction and reconstructing each 2-D slice individually. This approach would provide only limited 3-D motion correction, but it would avoid fully 3-D iterative reconstructions, which are computationally intensive in general.

6. Conclusion

To conclude, we introduced the problem of correcting intra-frame non-rigid motion during an acquisition. We derived a multi-scale, motion-regularized iterative reconstruction that jointly estimates non-rigid motion among groups of k-space interleaves as well as the desired “source” image. We then evaluated this method against SPIRiT reconstruction with and without non-rigid autofocus motion correction and demonstrated that our method yields significant improvements in both PSER and SSIM quality metrics. We discussed these results and identified limitations of our method and paths for future development for practical 2-D and 3-D image reconstruction in the presence of intra-frame motion.

Acknowledgements

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

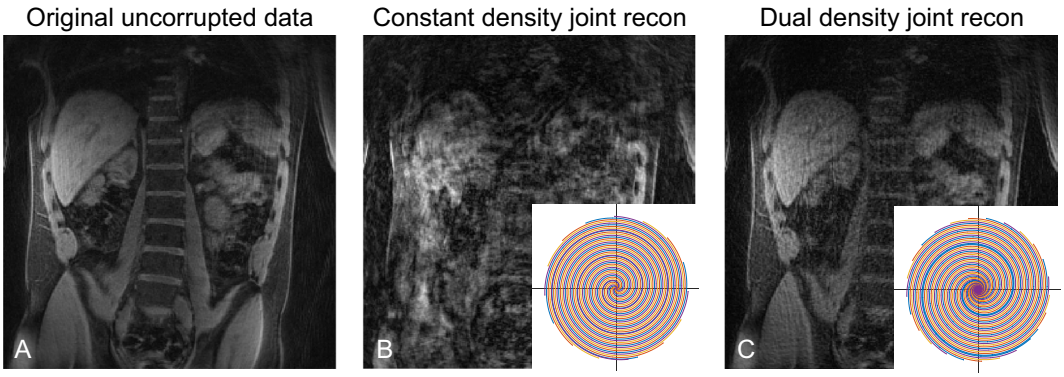

Fig. 9. Our multi-scale joint reconstruction relies on initialization with relatively complete low-resolution information, as aliasing artifacts at low-resolution can mislead motion estimation at low resolution. The difference in reconstruction quality between motion-compensated joint reconstructions from 3×-undersampled constant (middle) and dual (right) density data clearly illustrates this advantage.

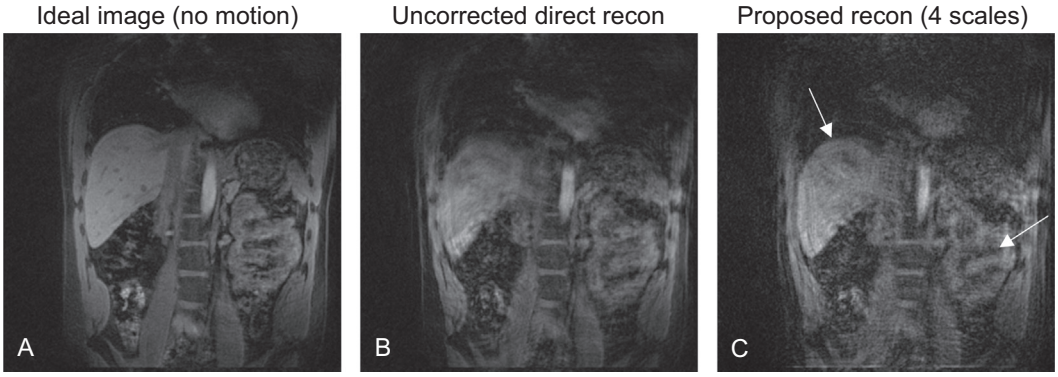

Fig. 10. This experiment compares both direct reconstruction (B) and the proposed joint reconstruction (C) of free-breathing data with real motion (not simulated) against an idealized breath-held image (A). While both images are significantly degraded versus the ideal image, the joint reconstruction provides noticeable improvements in the delineation of various image features (white arrows).

References

[1] Cheng JY, Zhang T, Ruangwattanapaisarn N, Alley MT, Uecker M, Pauly JM, et al. Free-breathing pediatric MRI with nonrigid motion correction and acceleration. J Magn Reson Imaging 2015;42(2):407–20.

[2] Adluru G, DiBella EVR, Schabel MC. Model-based registration for dynamic cardiac perfusion MRI. J Magn Reson Imaging Nov. 2006;24(5):1062–70.

[3] Lingala SG, DiBella E, Jacob M. Deformation corrected compressed sensing (DC-CS): a novel framework for accelerated dynamic MRI. IEEE Trans Med Imaging Jan. 2015;34(1):72–85.

[4] Batchelor PG, Atkinson D, Irarrazaval P, Hill DLG, Hajnal J, Larkman D. Matrix description of general motion correction applied to multishot images. Magn Reson Med Nov. 2005;54(5):1273–80.

[5] Cheng JY, Zhang T, Alley MT, Uecker M, Lustig M, Pauly JM, et al. Comprehensive multi-dimensional MRI for the simultaneous assessment of cardiopulmonary anatomy and physiology. Sci Rep July 13, 2017;7(1):5330.

[6] Feiyu C, Tao Z, Cheng JY, Xinwei S, Pauly JM, Vasanawala SS. Autocalibrating motion-corrected wave-encoding for highly accelerated free-breathing abdominal MRI. Magn Reson Med Nov. 2017;78(5):1757–66.

[7] Zucker EJ, Cheng JY, Anshul H, Michael C, Vasanawala SS. Free-breathing pediatric chest MRI: performance of self-navigated golden-angle ordered conical ultrashort echo time acquisition. J Magn Reson Imaging Jan. 2017;47(1):200–9.

[8] Atkinson D, Hill DG, Stoyle PR, Summers P, Keevil S. An autofocus algorithm for the automatic correction of motion artifacts in MR images. Information Processing in Medical Imaging. 1230. 1997. p. 341–54.

[9] Cheng JY, Alley MT, Cunningham CH, Vasanawala SS, Pauly JM, Lustig M. Nonrigid motion correction in 3D using autofocusing with localized linear translations. Magn Reson Med 2012;68(6):1785–97.

[10] Ingle RR, Wu HH, Addy NO, Cheng JY, Yang PC, Hu BS, et al. Nonrigid autofocus motion correction for coronary MR angiography with a 3D cones trajectory. Magn Reson Med 2014;72(2):347–61.

[11] Hu P, Hong S, Moghari MH, Goddu B, Goepfert L, Kissinger KV, et al. Motion correction using coil arrays (MOCCA) for free-breathing cardiac cine MRI. Magn Reson Med 2011;66(2):467–75.

[12] Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUSS: a general compressed sensing framework for high resolution dynamic MRI. Magn Reson Med 2009;61(1):103–16.

[13] Asif MS, Hamilton L, Brummer M, Romberg J. Motion-adaptive spatio-temporal regularization for accelerated dynamic MRI. Magn Reson Med 2013;70(3):800–12.

[14] Chen X, Salerno M, Yang Y, Epstein FH. Motion-compensated compressed sensing for dynamic contrast-enhanced MRI using regional spatiotemporal sparsity and region tracking: block low-rank sparsity with motion-guidance (BLOSM). Magn Reson Med 2014;72(4):1028–38.

[15] Shaw JL, Christodoulou AG, Sharif B, Li D. Ungated, free-breathing native T1 mapping in multiple cardiac phases in under one minute: a proof of concept. Proc Int Soc Magn Reson Med 2016:3149.

[16] Christodoulou AG, Shaw JL, Bi X, Sharif B, Li D. Non-ECG first-pass myocardial perfusion T1 mapping with low-rank tensor cardiovascular MR multitasking. Proc Int Soc Magn Reson Med 2017:450.

[17] Mohsin YQ, Lingala SG, Dibella E, Jacob M. Accelerated dynamic MRI using patch regularization for implicit motion compensation. Magn Reson Med 2017;77(3):1238–48.

[18] Yoon H, Kim KS, Ye JC. Motion compensated compressed sensing dynamic MRI with low-rank patch-based residual reconstruction. Proc. IEEE International Symposium on Biomedical Imaging (ISBI). 2013. p. 314–7.

[19] Kellman P, Chefd'hotel C, Lorenz CH, Mancini C, Arai AE, McVeigh ER. High spatial and temporal resolution cardiac cine MRI from retrospective reconstruction of data acquired in real time using motion correction and resorting. Magn Reson Med 2009;62(6):1557–64.

[20] Jeelani H, Yang Y, Salerno M, Weller DS. Evaluation of k-space and image-space motion correction schemes for CMR perfusion. 20th Annual SCMR Scientific Sessions. 2017.

[21] Jin N, da Silveira JS, Jolly M, Firmin DN, Mathew G, Lamba N, et al. Free-breathing myocardial T2* mapping using GRE-EPI and automatic non-rigid motion correction. J Cardiovasc Magn Reson Dec. 2015;17(1):113.

[22] Loktyushin A, Nickisch H, Pohmann R, Schölkopf B. Blind multirigid retrospective motion correction of MR images. Magn Reson Med 2015;73(4):1457–68.

[23] Odille F, Vuissoz P, Marie P, Felblinger J. Generalized reconstruction by inversion of coupled systems (GRICS) applied to free-breathing MRI. Magn Reson Med 2008;60(1):146–57.

[24] Fayad H, Odille F, Schmidt H, Würslin C, Küstner T, Felblinger J, et al. The use of a generalized reconstruction by inversion of coupled systems (GRICS) approach for generic respiratory motion correction in PET/MR imaging. Phys Med Biol March 2015;60(6):2529–46.

[25] Odille F, Menini A, Escanyé JM, Vuissoz PA, Marie PY, Beaumont M, et al. Joint reconstruction of multiple images and motion in MRI: application to free-breathing myocardial T_2 quantification. IEEE Trans Med Imaging 2016;35(1):197–207.

[26] Chun SY, Fessler JA. A simple regularizer for B-spline nonrigid image registration that encourages local invertibility. IEEE J Sel Top Sign Proces 2009;3(1):159–69.

[27] Adluru G, McGann C, Speier P, Kholmovski EG, Shaaban A, DiBella EVR. Acquisition and reconstruction of undersampled radial data for myocardial perfusion magnetic resonance imaging. J Magn Reson Imaging 2009;29(2):466–73.

[28] Jung H, Park J, Yoo J, Ye JC. Radial k-t FOCUSS for high-resolution cardiac cine MRI. Magn Reson Med 2010;63(1):68–78.

[29] Feng L, Grimm R, Block KT, Chandarana H, Kim S, Xu J, et al. Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. Magn Reson Med 2014;72(3):707–17.

[30] Feng L, Axel L, Chandarana H, Block KT, Sodickson DK, Otazo R. XD-GRASP: golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing. Magn Reson Med 2016;75(2):775–88.

[31] Yang H, Sharif B, Pang J, Kali A, Bi X, Cokic I, et al. Free-breathing, motion-corrected, highly efficient whole heart T2 mapping at 3T with hybrid radial-cartesian trajectory. Magn Reson Med 2016;75(1):126–36.

[32] Viergever MA, Maintz JBA, Klein S, Murphy K, Staring M, Pluim JPW. A survey of medical image registration – under review. Med Image Anal 2016;33:140–4.

[33] Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. J Magn Reson Imaging 2015;42(4):887–901.

[34] Wang L, Weller DS. Joint motion estimation and image reconstruction using alternating minimization. Proc Int Soc Magn Reson Med 2016:1800.

[35] Lustig M, Pauly JM. SPIRiT: iterative self-consistent parallel imaging reconstruction from arbitrary k-space. Magn Reson Med 2010;64(2):457–71.

[36] Fletcher R, Reeves CM. Function minimization by conjugate gradients. Comput J January 1964;7(2):149–54.

[37] Paige CC, Saunders MA. LSQR: an algorithm for sparse linear equations and sparse least squares. ACM Trans Math Softw March 1982;8(1):43–71.

[38] Wang Zhou, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13(4):600–12.

[39] Chen W, Meyer CH. Semiautomatic off-resonance correction in spiral imaging. Magn Reson Med 2008;59(5):1212–9.

[40] McGee KP, Armando M, Felmlee JP, Riederer SJ, Ehman RL. Image metric-based correction (autocorrection) of motion effects: analysis of image metrics. J Magn Reson Imaging February 2000;11(2):174–81.

[41] Irarrazabal P, Nishimura DG. Fast three dimensional magnetic resonance imaging. Magn Reson Med 1995;33(5):656–62.