

# Blind Multirigid Retrospective Motion Correction of MR Images

Alexander Loktyushin,<sup>1\*</sup> Hannes Nickisch,<sup>2</sup> Rolf Pohmann,<sup>3</sup> and Bernhard Schölkopf<sup>4</sup>

**Purpose:** Physiological nonrigid motion is inevitable when imaging, e.g., abdominal viscera, and can lead to serious deterioration of the image quality. Prospective techniques for motion correction can handle only special types of nonrigid motion, as they only allow global correction. Retrospective methods developed so far need guidance from navigator sequences or external sensors. We propose a fully retrospective nonrigid motion correction scheme that only needs raw data as an input.

**Methods:** Our method is based on a forward model that describes the effects of nonrigid motion by partitioning the image into patches with locally rigid motion. Using this forward model, we construct an objective function that we can optimize with respect to both unknown motion parameters per patch and the underlying sharp image.

**Results:** We evaluate our method on both synthetic and real data in 2D and 3D. In vivo data was acquired using standard imaging sequences. The correction algorithm significantly improves the image quality. Our compute unified device architecture (CUDA)-enabled graphic processing unit implementation ensures feasible computation times.

**Conclusion:** The presented technique is the first computationally feasible retrospective method that uses the raw data of standard imaging sequences, and allows to correct for nonrigid motion without guidance from external motion sensors.

**Magn Reson Med** 73:1457–1468, 2015. © 2014 Wiley Periodicals, Inc.

**Key words:** retrospective motion correction; autofocusing; gradient-based optimization; nonrigid motion; graphic processing unit

healthy cooperative subjects, while patients with movement disorders, elderly people and children almost inevitably move unless sedated. Physically constraining the subject usually provides only limited reduction of motion, and has no effect on motion due to physiological function of the organism, which is unavoidable even if the subject is sedated. Especially respiration and heartbeat are of primary concern for thoracic and abdominal MR imaging. Motion artifacts due to respiration can be reduced with breath-held scans (1), but this imposes many limitations on the acquisition, as scan times have to stay within tolerable breath-holding periods. Further examples of physiological motion include bowel movement, cardiovascular pulsation, and even within the brain, nonrigid pulsations in the thalamus are observed (2).

Usually, physiological motion leads to complex artifacts, which cannot be properly addressed by motion correction methods that assume the scanned object to be a rigid body. Prospective motion correction methods suffer from the global correction problem (see (3)), and are restricted to global affine motions, which include nonrigid shearing and scaling (4). Most of the research on nonrigid motion correction has focused on developing retrospective methods, which correct for motion once the data is acquired.

The early research of (5) shows that it is possible to correct for nonrigid motion in kinetic joint studies. The motion estimation does not require the use of additional measurements, and is driven by the projections from a segmented radial acquisition. In their seminal paper, Batchelor et al. (6) derive a general description of effects of arbitrary (also nonrigid) motion during MR imaging applicable to standard sequences. Motion degradation is modeled as a linear process described by matrix-vector multiplications. The authors show that it is possible to find an estimated inverse of the linear system, and thus to retrospectively recover the underlying sharp image. Their approach is of theoretical interest, because they found it computationally prohibitive to estimate the unknown motion parameters. To solve this problem, external motion reference signals obtained from navigators were proposed (7). This idea turned out to be fruitful, and was later used for nonrigid motion correction in coronary MRI (8), cardiac CINE (9–11), and liver MRI (12).

A further important step was made by Odille et al. (13), who treat the image reconstruction and nonrigid motion parameter estimation as two coupled inverse problems. Still, models for patient motion driven by external sensors such as electrocardiography (ECG) or navigators need to be used to make the motion parameter estimation feasible. In follow-up research (14), it is shown how to reduce the number of unknown motion

## INTRODUCTION

Motion correction is among the most challenging unsolved problems in MRI. Motion artifacts can have a particularly devastating effect on high-resolution images, which require a patient to restrict motion to below the millimeter range over several minutes. Being stationary in the scanner for a long time is challenging even for

<sup>1</sup>Max Planck Institute for Intelligent Systems, Empirical Inference Department, Tübingen, Germany.

<sup>2</sup>Philips Research Europe, Hamburg, Germany.

<sup>3</sup>Max Planck Institute for Biological Cybernetics, High-Field Magnetic Resonance Department, Tübingen, Germany.

<sup>4</sup>Max Planck Institute for Intelligent Systems, Empirical Inference department, Tübingen, Germany.

\*Correspondence to: Alexander Loktyushin, M.Sc., Max Planck Institute for Intelligent Systems, Spemannstr. 41, 72076 Tübingen, Germany. E-mail: aloktyus@tuebingen.mpg.de

Received 6 February 2014; revised 31 March 2014; accepted 4 April 2014  
DOI 10.1002/mrm.25266

Published online 23 April 2014 in Wiley Online Library (wileyonlinelibrary.com).

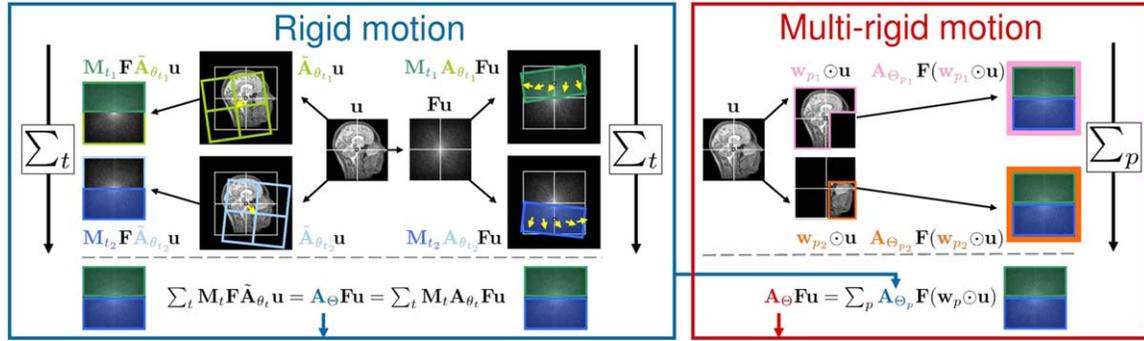


FIG. 1. Illustration of the forward model. Left panel, rigid: An unknown underlying image  $\mathbf{u}$  is affected by a linear transformation, which is given by the operator  $\tilde{\mathbf{A}}_{\theta_t}$  in the spatial domain (left part of the panel), or by  $\mathbf{A}_{\theta_t}$  in the Fourier domain (right part) at every time point  $t$ . The observed image (bottom row) is obtained by summation (over the entire acquisition time  $t = 1..T$ ) of k-space segments extracted by the trajectory matrix  $\mathbf{M}_t$ . Right panel, multirigid: The image is split into  $P$  patches (brain/jaw) by means of spatial windows  $\mathbf{w}_p$ . For each spatial window, a rigid motion (left panel) is applied with a distinct set of motion parameters  $\Theta_p$ . The observed image (bottom row) is obtained by summing over spatial patches.

parameters with the use of adaptive meshes, which involves grouping voxels with similar displacements and thus encoding local smoothness information. But still information from external motion sensors is needed for reconstruction.

A recent work (15) proposes to approximate nonrigid motion as simple linear translations in small spatial regions to reduce the problem complexity. This allows for sufficiently fast implementation of a nonrigid motion correction algorithm, which uses butterfly navigators to estimate the unknown motion parameters. More recently (16), nonrigid motion correction was combined with compressed sensing to address the problem of respiratory motion during free-breathing cardiac acquisitions. Again, a navigator signal is required to estimate the motion.

In this article, we propose a retrospective nonrigid motion correction method that does not require motion information from external sensors, thus being a purely image-data-driven method. We assume locally rigid motion and provide a generic quality functional together with an efficient gradient-based optimization algorithm to recover the underlying sharp image. Our work aims at nonrigid physiological motion compensation in routine clinical imaging (e.g., in abdomen scans) in the medium term. In this study, we make a step forward, and in our in vivo experiments, we correct for multirigid finger motion in wrist imaging.

A Matlab implementation of our new method (GradMC2) together with a few test examples is available at: <http://mloss.org/software/view/430/>.

## METHODS

We base the present method on our previous rigid motion model (17), which we restate in the beginning of this section. We then explain how to combine multiple locally rigid motions in image patches into a joint global nonrigid motion model, and introduce and discuss the concept of image patch separability. Next, we formulate the trade-off between data fidelity and restored image quality measures, and introduce our nested-loop optimization algorithm to recover the sharp image and the underlying

motion parameters. At the end of the section, we detail the technical aspects of our real data experiments.

### Rigid Motion in MRI

In the following, we restate the rigid motion forward model used by (17) to describe the image degradation due to rigid motion in MR scans. Let  $\mathbf{F} \in \mathbb{C}^{N \times N}$  be the orthonormal Fourier transformation matrix, let  $\mathbf{u} \in \mathbb{C}^N$  denote a column vector containing the unknown sharp volume with  $N = n_x \cdot n_y \cdot n_z$  voxels;  $\mathbf{m}_t \in [0, 1]^N$  with  $\sum_{t=1}^T \mathbf{m}_t = \mathbf{1}$  are binary masking coefficients mapping a particular segment of k-space to a given acquisition trajectory, and  $\tilde{\mathbf{A}}_{\theta_t} \in \mathbb{R}^{N \times N}$  is a rigid motion transformation matrix parameterized by a vector pair of translations and rotations  $\theta_t \in \mathbb{R}^3 \times [0, 2\pi]^3$  at acquisition time  $t$ . The motion transformation matrix  $\tilde{\mathbf{A}}_{\theta_t}$  and the Fourier transformation matrix  $\mathbf{F}$  do not commute in general. But since we know that rotation and translation have equivalents in k-space (18) there exists an operator  $\mathbf{A}_{\theta_t}$  such that  $\mathbf{F}\tilde{\mathbf{A}}_{\theta_t} = \mathbf{A}_{\theta_t}\mathbf{F}$ , which we use in our implementation. Assuming additive measurement noise  $\boldsymbol{\varepsilon}$ , the acquisition in k-space can be written as

$$\mathbf{y} = \sum_{t=1}^T \text{diag}(\mathbf{m}_t)\mathbf{A}_{\theta_t}\mathbf{F}\mathbf{u} + \boldsymbol{\varepsilon} = \mathbf{A}_{\Theta}\mathbf{F}\mathbf{u} + \boldsymbol{\varepsilon} \in \mathbb{C}^N \quad [1]$$

where  $\mathbf{m}_t$  selects the part of k-space that is recorded by the scanner at time  $t$  and  $\text{diag}(\cdot)$  builds a diagonal matrix from a vector. The summation is carried out over the total number of acquisitions  $T = n_y \cdot n_z$ . See the left panel of Figure 1 for a graphical illustration.

Let the matrix  $\Theta = [\theta_1, \dots, \theta_T]$  contain the parameters for the entire motion trajectory. As the masking vector  $\mathbf{m}_t$  is binary, the motion matrix  $\mathbf{A}_{\Theta}$  has a convenient block structure

$$\mathbf{A}_{\Theta} := \sum_{t=1}^T \text{diag}(\mathbf{m}_t)\mathbf{A}_{\theta_t} = \begin{bmatrix} [\mathbf{A}_{\theta_1}]_{\mathbf{m}_1} \\ [\mathbf{A}_{\theta_2}]_{\mathbf{m}_2} \\ \vdots \\ [\mathbf{A}_{\theta_T}]_{\mathbf{m}_T} \end{bmatrix} \in \mathbb{C}^{N \times N} \quad [2]$$

allowing for efficient matrix-vector multiplications with  $\mathbf{A}_\Theta$  at a computational effort of  $\mathcal{O}(N)$ . For the computational details on the efficient implementation of the forward model, see ((17), section Methods).

In the following, we demonstrate how a set of forward motion matrices representing locally rigid motion in image patches can be combined to obtain a multirigid forward model.

### Nonrigid Motion in MRI

Nonrigid body motion leads to image distortions that cannot be described by a single *global* pair of translation and rotation vectors. We consider (approximately rigid but) *local* motion inside an image patch, which we will call *multirigid* in the following. We use the term “multirigid” instead of the more common “nonrigid” to emphasize that we view a general nonrigid motion as the extreme case, where each point of the body of interest possesses a unique motion vector. Note that any linear operation on the volume can be modeled by multirigid transformations if only the patches are made small enough, i.e., the number of patches  $P$  is in the order of the number of voxels  $N$ . We illustrate our model in Figure 1.

Formally, we partition the image volume into  $P$  spatial patches using windows  $\mathbf{w}_p \in [0, 1]^N$ . This description is similar to the Efficient Filter Flow framework (19) that aims to solve a nonstationary deconvolution problem. We do neither impose a particular shape nor require connectedness of the patches. The patches could—at least theoretically—be estimated from the measured k-space data  $\mathbf{y}$ , however, in this article, we assume that they are given (see Discussion section). Doing the normalization  $\sum_{p=1}^P \mathbf{w}_p = \mathbf{1}$ , we can interpret the windows  $\mathbf{w}_p$  as the probabilities of voxels to belong to the spatial patch  $p$ . Implementation wise, we define the spatial windows using binary masks, which we convolve with a Gaussian kernel to enforce a smooth transition between the motion parameters of neighboring patches. Note that the width of the Gaussian kernel governs the influence region and the steepness of the motion parameter transition. The k-space equivalent of the pointwise multiplication with a spatial window  $\mathbf{w}_p$  is a convolution matrix  $\mathbf{W}_p$  such that  $\mathbf{W}_p \mathbf{F} \mathbf{u} = (\mathbf{F} \mathbf{w}_p) \star (\mathbf{F} \mathbf{u}) = \mathbf{F}(\mathbf{w}_p \odot \mathbf{u})$ , where  $\star$  is the convolution operation and  $\odot$  denotes pointwise multiplication. Each of the  $P$  patches has an associated motion trajectory parameterized by  $\Theta_p$ . We can write the multirigid forward model as

$$\mathbf{y} = \sum_{p=1}^P \mathbf{A}_{\Theta_p} \mathbf{F}(\mathbf{w}_p \odot \mathbf{u}) + \boldsymbol{\varepsilon} = \mathbf{A}_\Theta \mathbf{F} \mathbf{u} + \boldsymbol{\varepsilon} \in \mathbb{C}^N, \quad [3]$$

where we use  $P$  different rigid motion matrices  $\mathbf{A}_{\Theta_p}$  to construct the multirigid transformation  $\mathbf{A}_\Theta = \sum_{p=1}^P \mathbf{A}_{\Theta_p} \mathbf{W}_p$  with an overall number of  $6 \cdot T \cdot P$  free parameters. Note that for  $P=1$ , we have  $\mathbf{w}_p = \mathbf{1} \Rightarrow \mathbf{W}_p = \mathbf{I}$  that is we recover the rigid motion model introduced in the previous section as a special case.

### Fourier Domain versus Spatial Domain

Instead of doing the computations in the Fourier domain using  $P$  rigid motion transformations as depicted in the

right panel of Figure 1, we can swap the summation over time  $t$  and patches  $p$  and equivalently evaluate the forward model in the spatial domain

$$\mathbf{y} = \sum_{t=1}^T \mathbf{F} \tilde{\mathbf{A}}_{\Theta_t} \mathbf{u} + \boldsymbol{\varepsilon}, \quad \tilde{\mathbf{A}}_{\Theta_t} = \sum_{p=1}^P \tilde{\mathbf{A}}_{\Theta_t, p} \text{diag}(\mathbf{w}_p) \quad [4]$$

leaving the result unaltered but choosing a different computational route. Here, the main difference is the computational complexity of (a) the spatial approach and (b) the Fourier approach:

- Method (a) is dominated by  $T$  fast Fourier transforms  $\mathcal{O}(\frac{2}{3} N \log N)$  following the application of the oversampled multirigid transformation  $\mathcal{O}(2^D N)$  yielding  $\mathcal{O}(T \cdot N \cdot (\frac{2}{3} \log N + 2^D))$ .
- Method (b) is governed by the complexity of  $P$  rigid models where double oversampling is used for the Fourier transform and an additional double oversampling for the rotation interpolation leaving an overall  $\mathcal{O}(P \cdot N \cdot (4^D + 2^D \log N))$ .

Setting  $n := n_x = n_y = n_z = 200$ ,  $N = n^3$  as data size,  $D=3$  as data dimension,  $T=n$  as the number of motion parameters (piecewise constant motion trajectory), and equating the complexities of (a) and (b), we find  $P \approx 19$ . For  $n = 100$ , we obtain  $P \approx 10$  which illustrates how complex the nonrigid motion can become until spatial computations of method (a) are more efficient. As modern processors and graphic processing units (GPUs) allow parallel computations, the resulting trade-off depends on the degree of parallelization. Both the  $P$  interpolated rotations in (b) and the summation over the  $T$  independent Fourier transformed motion transformations in (a) can be done in parallel. In our case of a moderate number of patches  $P \leq 6$  and sufficiently large images, the Fourier approach is computationally beneficial.

### Motion Separability

It is well known that the artifacts due to motion manifest as ghosts that propagate in phase-encode direction, and blur in the direction(s) of motion (see (20,21)). In case of pure translations, the motion degradation is described by a convolution with some point-spread-function, which has a large support (attributed to global ghosts) in phase-encode direction, and much smaller support (resulting in local blur) in frequency encode direction.

This has strong implications when it comes to non-rigid motion correction. Consider a 2D multirigid problem, where the image is half-split into two rectangular patches. If the image is split in frequency encode direction, we call the problem *separable*. This means that it is possible to segment the image into two patches in spatial domain by means of spatial windows, and then treat the two resulting images as two separate rigid motion correction problems, which can be addressed by existing retrospective rigid motion correction methods. In particular, the problem can be solved using our inverse-based approach (17) in the order of seconds time. It is important to note that as blurring also propagates in frequency encode direction, the problem is—strictly speaking—not completely separable. However, as the support of point-

spread-function in frequency encode direction is small the use of spatial windows with smooth cut-off/overlap allows to treat the problem as approximately separable subject to small localized errors near the window boundary. In many cases, where the boundary between rigid parts has no diagnostic value these small errors might be tolerated. So far we were assuming that the image is affected by pure translational motion. It is important to note that a strong rotational motion (in a range of more than a couple of degrees) makes the problem nonseparable irrespective of whether the patches are arranged in phase or frequency encode directions.

In the nonseparable case, the patches are arranged in phase-encode direction, which leads to a considerable “artifact cross-talk” between the adjacent patches resulting a genuine multirigid problem. We see the separable case as a “trivial” one, where a solution might be obtained using rigid motion correction methods. In this article, our goal is to solve the nonseparable problem, and in our experiments we arrange the patches (and assume multiple motion trajectories) in phase-encode direction. We explicitly test the property of separability in a dedicated experiment (see Results section).

### Objective Function

To recover both unknown motion parameters  $\Theta$  and the underlying sharp image  $\mathbf{u}$  from the motion-corrupted k-space measurement  $\mathbf{y}$ , we need to formalize the trade-off between *data fidelity*, i.e., how well the data  $\mathbf{y}$  is reproduced by the forward operator  $\mathbf{A}_\Theta$  and *regularization* of both  $\mathbf{u}$  and  $\Theta$ . Regularization here has two meanings: The regularizer for  $\mathbf{u}$  measures the image quality, and the regularizer for  $\Theta$  softly constrains the vast space of motion parameters, so that only a tiny relevant subspace needs to be considered. Our objective function is

$$\psi(\mathbf{u}, \Theta) = \phi(\mathbf{u}) + \lambda \cdot \|\mathbf{r}_\Theta\|^2 + \mu \cdot \sum_{p=1}^P \text{tr}(\Theta_p \mathbf{D}_t^\top \mathbf{D}_t \Theta_p^\top), \quad [5]$$

$$\text{with } \mathbf{r}_\Theta = \mathbf{A}_\Theta \mathbf{F} \mathbf{u} - \mathbf{y}$$

where  $\lambda \in \mathbb{R}_+$  and  $\mu \in \mathbb{R}_+$  are scalar weights for the data-fidelity term and the motion trajectory regularizer, respectively. We measure data fidelity by the squared norm of the k-space residual  $\mathbf{r}_\Theta$ , which corresponds to the squared error. Regularization of the motion parameters is achieved by requiring small temporal derivatives independently for each of the  $P$  motion trajectories where  $\mathbf{D}_t \in \{0, \pm 1\}^{T \times T}$  is a discrete temporal derivative matrix and  $\text{tr}(\cdot)$  is the trace operator.

Using the full objective (5), we can correct for both strong rotations and multirigid motion. In this work, we use the sum of the gradient’s absolute values  $\phi(\mathbf{u}) = \|\mathbf{G}\mathbf{u}\|_1$ , with  $\mathbf{G} = [\mathbf{D}_x, \mathbf{D}_y, \mathbf{D}_z] \in \{0, \pm 1\}^{3N \times N}$ , and  $\|\mathbf{v}\|_1 = \sum_{i=1}^N |v_i|$  as an image regularizer. We use this  $l_1$ -norm of the gradient image metric for two reasons:

1. Optimization of (5) w.r.t.  $\mathbf{u}$  is a convex penalized least squares problem with a single global optimum.
2. The metric is widely used and compares well to other nonconvex metrics in empirical studies (22).

For a discussion of the connection between the previously used inverse-based approach and Eq. [5], see the Appendix.

### Optimization Scheme

Using the objective function (5), we want to recover both the unknown sharp image  $\mathbf{u}$  and the motion parameters  $\Theta$  simultaneously. Let us define the mapping  $\psi_*$  obtained by evaluating the objective  $\psi$  at the optimal image  $\mathbf{u}_*$  by

$$\psi_*(\Theta) = \psi(\mathbf{u}_*, \Theta), \quad \text{with } \mathbf{u}_* = \arg \min_{\mathbf{u}} \psi(\mathbf{u}, \Theta).$$

Empirically, small changes in  $\Theta$  can lead to big changes in the optimal value  $\mathbf{u}_*$ . Also, optimization w.r.t.  $\mathbf{u}$  is convex and simple, whereas  $\psi$  is a nonconvex function in  $\Theta$ . This suggests the special structure of an optimization scheme, which we detail in the Algorithm 1. The purpose of the *annealing loop* is to progressively relax the importance of the regularization terms of both the image  $\mathbf{u}$  and motion parameters  $\Theta$  by increasing the weight of the data-fidelity term. In the *motion loop*  $\psi_*$  is optimized w.r.t. motion parameters  $\Theta$ ; here we use the fact that the forward model allows to compute the motion parameter gradient in closed form. The most inner *image loop* involves the computation of the current best estimate of the image  $\mathbf{u}_*$  obtained by optimizing the convex objective function (5). Additionally, for strong motions, a multiscale approach is used, which is implemented by the outermost *multiscale loop*. The idea is to first solve a simpler low-resolution problem, and then use the estimated motion parameters as initializations for higher resolution scales. This has an effect of avoiding bad local minima, which the optimizer is prone to stuck in given the nonlinear objective function. To do the optimization, we use the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm [BFGS] (LBFGS) package (23).

---

#### Algorithm 1 Nested-loop optimization algorithm.

---

**Input:**  $\mathbf{y} \in \mathbb{C}^N, N = n_x \cdot n_y \cdot n_z$  // Corrupted volume  
 $\mathbf{s} = [s_1, s_2, \dots, s_K]$  // **Multiscale schedule**  
 $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_L]$  // **Annealing schedule**  
 $N_\Theta$  // Number of motion estimation steps  
 $N_u$  // Number of image estimation steps

**Output:**  $\mathbf{u}_*$  // Restored image

Start from zero motion  $\Theta \leftarrow \mathbf{0}$ .

**For**  $\mathbf{s} = [s_1, s_2, \dots, s_K]$  **do** // **Multiscale loop**  
  Use only k-space center of the raw data.  
  **For**  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_L]$  **do** // **Annealing loop**  
    **For**  $\ell = 1, \dots, N_\Theta$  **do** // **Motion estimation loop**  
      Start from zero image  $\mathbf{u} \leftarrow \mathbf{0}$ .  
      // **Image estimation loop**  
      Find  $\mathbf{u}_* \leftarrow \arg \min_{\mathbf{u}} \psi(\mathbf{u}, \Theta)$ .  
      Compute the motion gradient  
       $\nabla \psi_* \leftarrow \partial \psi(\mathbf{u}_*, \Theta) / \partial \Theta$ .  
      Make a conjugate gradient step along  $\nabla \psi_*$ .  
    **End**  
  **End**  
  Initialize central frequency part of  $\Theta$  on next scale.  
**End**

---

## Implementation

### Reweighted Data-Fidelity Term

The literature on blind deconvolution (24,25) suggests that it is beneficial to compute the data-fidelity term in the gradient domain for blur kernel estimation (corresponds to our motion estimation loop in Alg. 1). We also adopt this approach and use a data fidelity of the form  $\|\mathbf{G}\mathbf{F}^H\mathbf{r}_\Theta\|^2$  in the motion estimation loop of Alg. 1. Here,  $\mathbf{r}_\Theta = \mathbf{A}_\Theta\mathbf{F}\mathbf{u} - \mathbf{y}$  is the k-space residual and  $\mathbf{G} = [\mathbf{D}_x, \mathbf{D}_y, \mathbf{D}_z] \in \{0, \pm 1\}^{3N \times N}$  computes the gradient using finite differences. As the gradient computation operation is diagonal in k-space and the Fourier matrix  $\mathbf{F}$  is orthonormal (leaving the norm invariant), the data-fidelity term becomes the weighted sum  $\|\mathbf{d}_x \odot \mathbf{r}_\Theta\|^2 + \|\mathbf{d}_y \odot \mathbf{r}_\Theta\|^2 + \|\mathbf{d}_z \odot \mathbf{r}_\Theta\|^2$  instead of the plain norm  $\|\mathbf{r}_\Theta\|^2$  in Eq. [5]. The effect of the reweighting is similar to a preconditioner in a linear system: High frequencies are magnified and low frequencies receive less emphasis. This comes in line with an intuition that motion can be best estimated from the localized high-frequency structures such as edges and cusps. Without the reweighting—due to the second-order structure of natural (and also medical) images—the norm  $\|\mathbf{r}_\Theta\|^2$  would be dominated by the low frequencies. The reweighting is not a prerequisite for our method to work, but the reconstruction results for 3D volumes are better and the optimization is faster.

### Forward Model Operator

In case of pure translations, the matrix  $\mathbf{A}_{\Theta_p}$  is diagonal with the elements given by the Fourier translation identity  $\mathcal{F}[g(x+a)] = \int_{-\infty}^{\infty} g(x)e^{-2\pi ik(x+a)}dx = \mathcal{F}[g(x)] \cdot e^{-2\pi ika}$ . For rotational motion, the matrix  $\mathbf{A}_{\Theta_p}$  has non-zero off-diagonal elements, which correspond to interpolation coefficients used for regridding. The number of non-zero elements in each row depends on the number of neighbors used for interpolation. The interpolation quality can be increased by using more neighbors, and also by using an oversampled grid, which can be obtained by zero-padding in the spatial domain. In our implementation, we are using 2-fold oversampling and  $4^D$  neighboring points, where  $D$  is the number of dimensions.

### GPU-Based Parallel Computation

It is important to note that the matrix  $\mathbf{A}_{\Theta_p}$  never needs to be constructed explicitly, an operation that can easily become memory prohibitive when dealing with large data volumes. What needs to be implemented are the matrix-vector operations: multiplications with  $\mathbf{A}_{\Theta_p}$ , its transpose and derivative with respect to motion parameters. Computationally these operations amount to point-wise multiplications with translational phase ramps, and interpolations to estimate the values on rotation-deformed grids. As these operations involve local multiplications and additions in frequency domain they can be efficiently handled by the computational architectures capable of massively parallel computations. We use the GPUs for this purpose, and implement our method in compute unified device architecture (CUDA) C.

Not only the operations with  $\mathbf{A}_{\Theta_p}$  matrix can benefit from a parallel computation. In fact the objective func-

tion (5) involves further ingredients that are subject to parallelization such as computation of spatial gradients, fast Fourier transforms, and image metric evaluation. To have the maximum computational gain, we implement and run the entire pipeline of Alg. 1 on GPU. In particular, this allows to minimize the memory transactions from the graphic card to the main memory, which can be a bottleneck when transferring large data volumes.

The most computationally active part of the algorithm is the inner image estimation block, which is iterated  $LN_\Theta$  times, and itself involves repetitive application of the motion operator. As the motion parameters do not change within this block, this allows to precompute both translational phase coefficients and interpolation weights before starting the block and store them in GPU memory. This makes the image estimation loop computationally cheap, and the entire algorithm operable in practical time.

### Computation Times

Figure 2 compares computation times for different setups. The simplest class of problems involves rigid bodies affected by weak (translation  $< 3$  mm, rotation  $< 4^\circ$ ), but commonly encountered motion. Such problems can be addressed with our fast inverse-based approach, which only requires a single loop for estimation of motion parameters. Going beyond this regime (multirigid mode and/or strong motion) requires additional computational work. The Fourier transforms still constitute the bottleneck, however, we now perform optimization in several nested loops, which increases the computation time by a factor depending on the number of outer loop iterations. Thus, there is a trade-off between the computation time and an achievable level of complexity of the motion correction problem to be solved. Still, using GPUs the computation times even for complicated classes of problems are tolerable (significantly below 1 h), and are going to improve due to the rapid advances in modern GPU technology.

### Data Acquisition

Imaging experiments were performed on a 3 T scanner (Magnetom Trio, Siemens Healthcare, Erlangen, Germany), using single-channel volume coils designed for wrist examinations. In a non-human set of experiments, we imaged kiwi fruits, which feature high-contrast fine-detailed regular internal structure, and are thus very helpful in assessing the ability of the correction algorithm to improve the visual quality. To induce the motion, we used a special MR-compatible actuator. 2D kiwi data had a matrix size of  $384 \times 384$  over a field-of-view (FOV) of  $150 \times 150$  mm<sup>2</sup> with a slice thickness of 3 mm. Fast length adjustment of short reads (FLASH) images with an echo time of 5.9 ms, a repetition time of 250 ms, and a flip angle of  $50^\circ$  were acquired within 1.5 min.

In vivo images were acquired with approval of the local ethics board. The hand of a human subject was imaged, who was told to move an index finger, while trying to keep the rest of the hand stationary. A standard FLASH sequence was used in both two and three

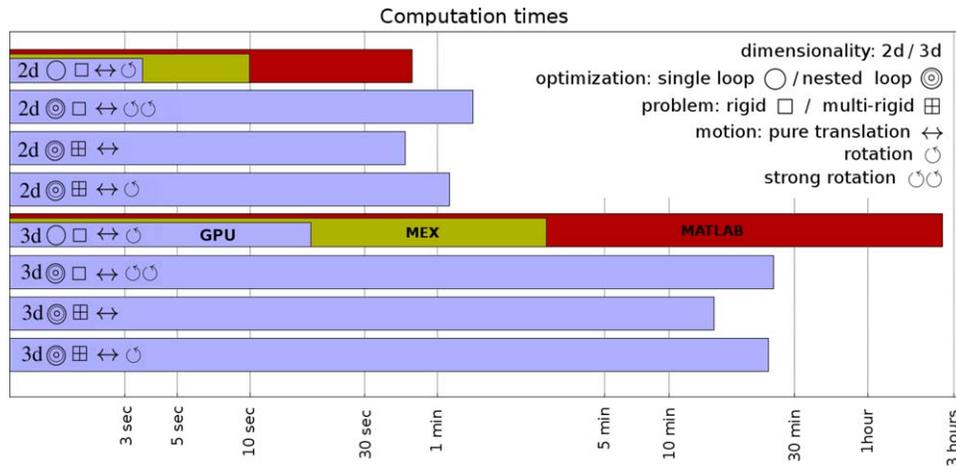


FIG. 2. Computation times (in log scale) for different motion correction problem classes. The problems differ in the dimensionality (2D,  $192 \times 192$  vs. 3D,  $192 \times 192 \times 16$ ) of the raw data, the kind of motion (global rigid motion  $\square$  versus multirigid motion  $\boxplus$ ), and the type of motion (translation  $\leftrightarrow$ , rotation  $\circ$ , strong rotation  $\odot$ ). The different problem classes require different flavors of the algorithm (single-loop  $\circ$  direct inverse objective vs. nested-loop  $\odot$  full objective optimization). For some (feasible) problems, we also show CPU computation times for C(MEX) and MATLAB implementations.

dimensions. 2D images with a matrix size of  $384 \times 384$  over a FOV of  $180 \times 180 \text{ mm}^2$ , a echo time of 5.9 ms, a repetition time of 250 ms, and a flip angle of  $50^\circ$  were acquired within 1 min. 2D images with a multifinger motion were acquried within 2 min, a matrix size of  $320 \times 320$  over a FOV of  $256 \times 256 \text{ mm}^2$ , a echo time of 7.1 ms, a repetition time of 500 ms, a flip angle of  $60^\circ$ , and a slice thickness of 3 mm. 3D wrist data had a matrix size of  $256 \times 256 \times 24$  over a FOV of  $180 \times 180 \times 36 \text{ mm}^3$ , a echo time of 3.1 ms, a repetition time of 20 ms, and a flip angle of  $20^\circ$  and were acquired within 2 min.

**RESULTS**

Simulated Data

We first used Eq. [1] to simulate motion-corrupted datasets, where we know both ground-truth image and motion trajectory. In the first experiment (see Figure 3), we examined the convergence properties of the nested-loop Algorithm 1. We evaluated the correction progress by tracking the error in the estimated image and trajectory on each outer loop iteration. As a ground-truth we used a 2D image (size of  $192 \times 192$ ) of the monkey brain,

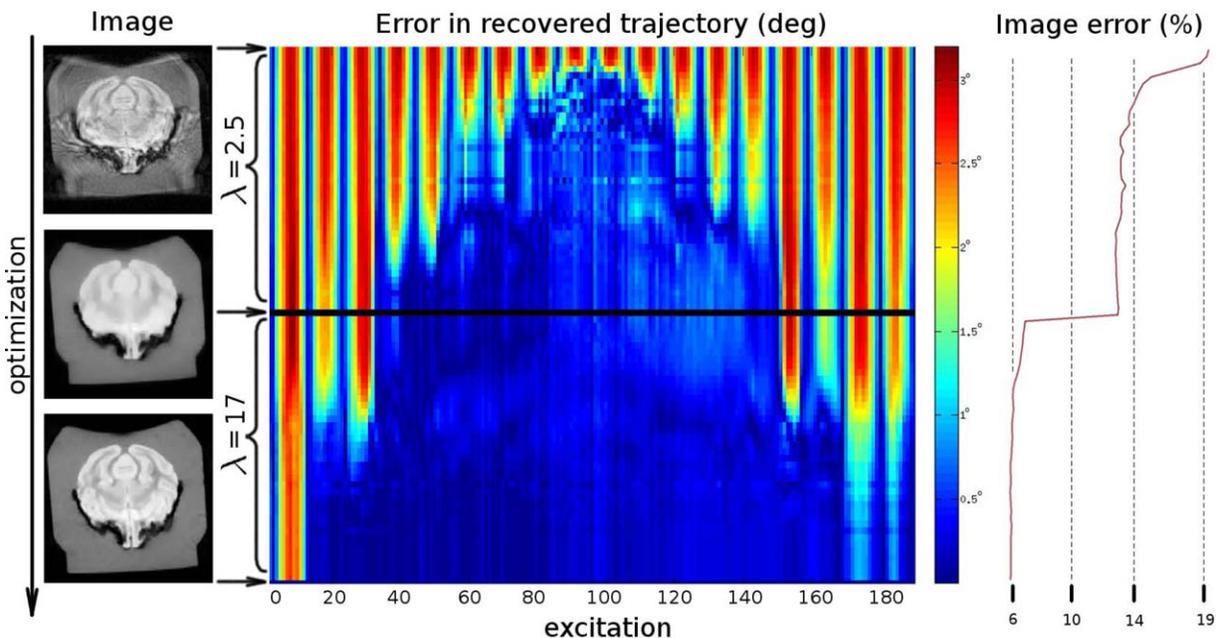
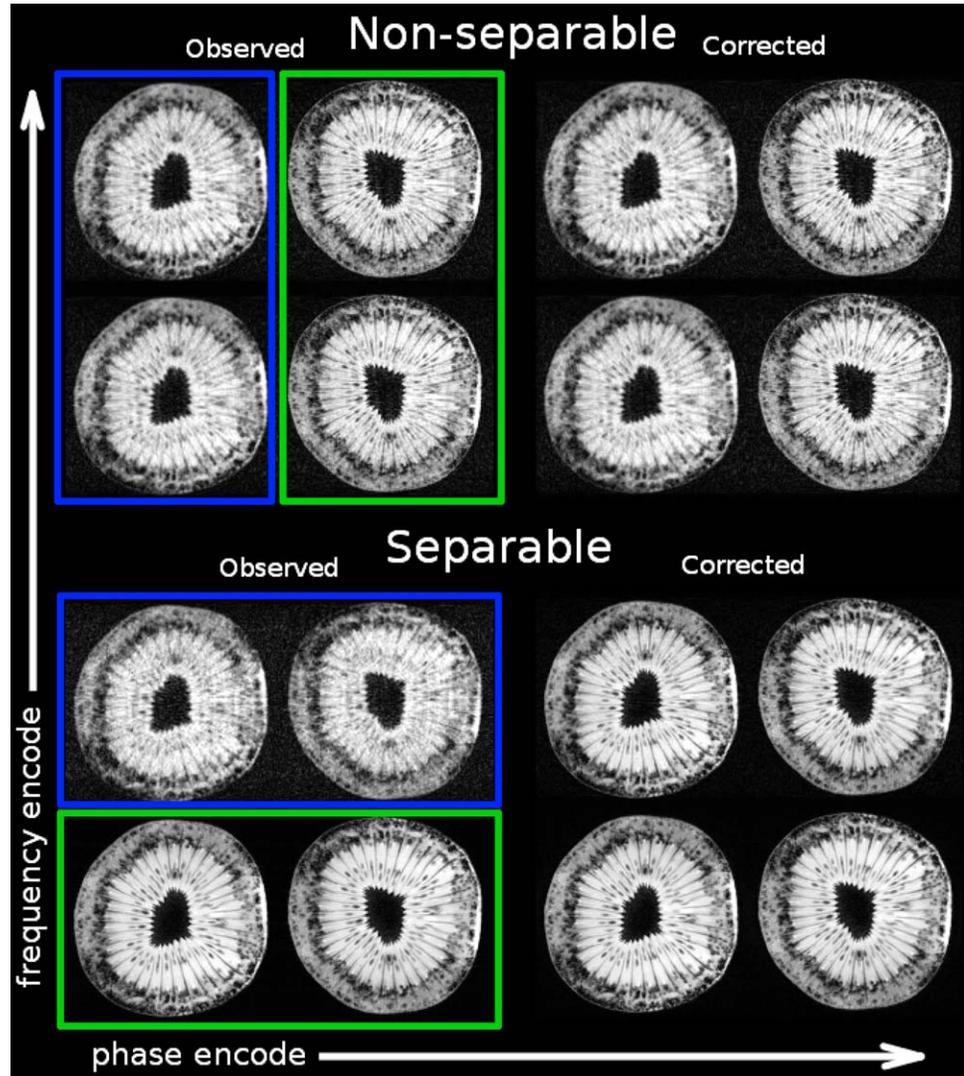


FIG. 3. Analysis of the motion correction process w.r.t. the error decrease in the reconstructed image and motion trajectory. The optimization time line is going from top to bottom. The regularization parameter annealing scheme consists of two epochs delimited with a black line. Left: the images before, in the middle (after the first annealing epoch), and after correction. Center: the change in the absolute value of the difference between estimated motion parameters and ground-truth; the color coding (red to blue) indicates the error decrease. Right: Normalized RMSE in reconstructed image computed w.r.t. ground-truth.

FIG. 4. Rigid motion correction in case of nonseparable (top) and separable (bottom) problems. In both cases, the same random and bi-directional translational motion was applied to patch enclosed in a blue frame. No motion was applied to green-frame patch. The left part of the figure shows the simulated motion-affected images, and the right part the results of motion correction.



and set the motion trajectory to a global (rigid, single patch) rotation of pure sine form with an amplitude of 2 degrees. In the center of Figure 3, we show for each iteration of the algorithm the error (absolute value of the difference) in the estimated rotational motion parameters computed w.r.t. ground-truth. The optimization time arrow is pointing from the top to the bottom of Figure 3, and the improvement (color coded from red to blue) toward the end of optimization can be seen. Periodically occurring low-error motion estimates (best seen in the very first iterations) are due to the fact that motion parameters were initialized to zeros, thus for zero-crossings of the sine trajectory they will be initially valid. We used two annealing steps  $\lambda = [2.5, 17]$ , which split the optimization process into two epochs (each epoch takes 50 iterations). In the first epoch (top half of Fig. 3), the algorithm gradually improved the motion parameter estimates corresponding to low-frequency excitations. This is because the heavily weighted regularization term eliminates high-frequency information in the image. During the second epoch, less weight was put

into regularization term, and the algorithm was able to recover high-frequency motion parameters.

The effects of the two epochs are easily seen in the recovered images. In the left column of the Figure 3, from top to bottom we show the motion-corrupted image, the image estimated after first epoch (notice oversmoothing due to strong regularization), and the image recovered after the second epoch. Additionally, we computed the normalized root mean-squared error between estimated images and ground-truth, and plot it in the right part of Figure 3.

Our next experiment (see Figure 4) probes the separability property that we introduced in the Methods section. We first study a nonseparable case, and split the image into two patches (blue, green) in phase-encode direction. We used our multirigid forward model to apply the translational motion (uniformly random bi-directional motion in a range of 4 pixels) to the blue patch. As ghosts propagate globally (from the nearby blue patch) one can see the decrease of image quality in the green patch, although it is not directly affected by

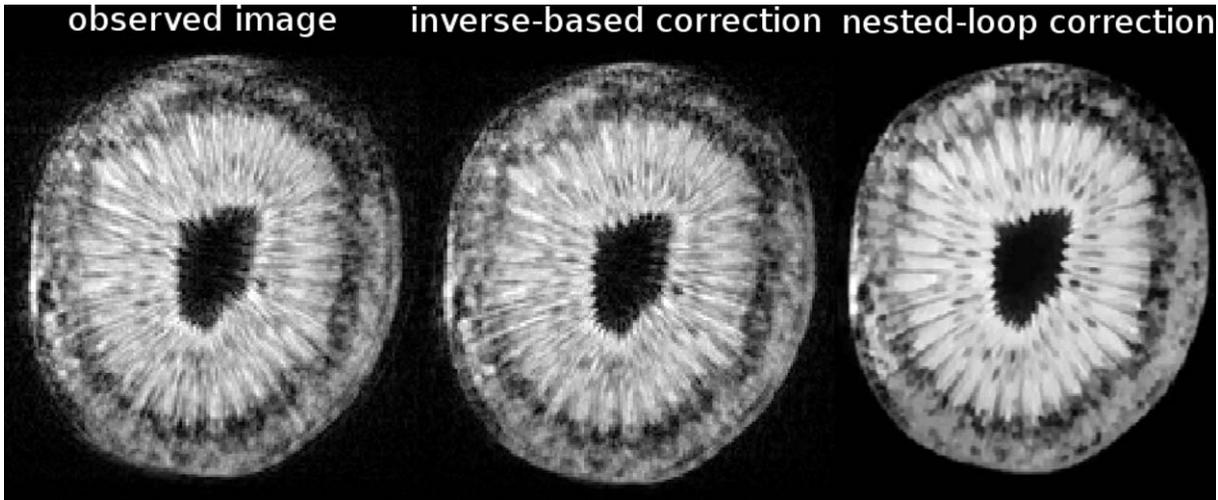


FIG. 5. Real data. Left to right: Image of a kiwi corrupted by rotational motion; restoration using the inverse-based objective; correction using the nested-loop algorithm.

motion. We then use spatial windows to extract the contents from each patch and correct them separately using our rigid inverse-based approach (17). No improvement

in image quality can be observed. In the separable case, the patches were arranged in frequency encode direction. The same motion trajectory as used in the nonseparable

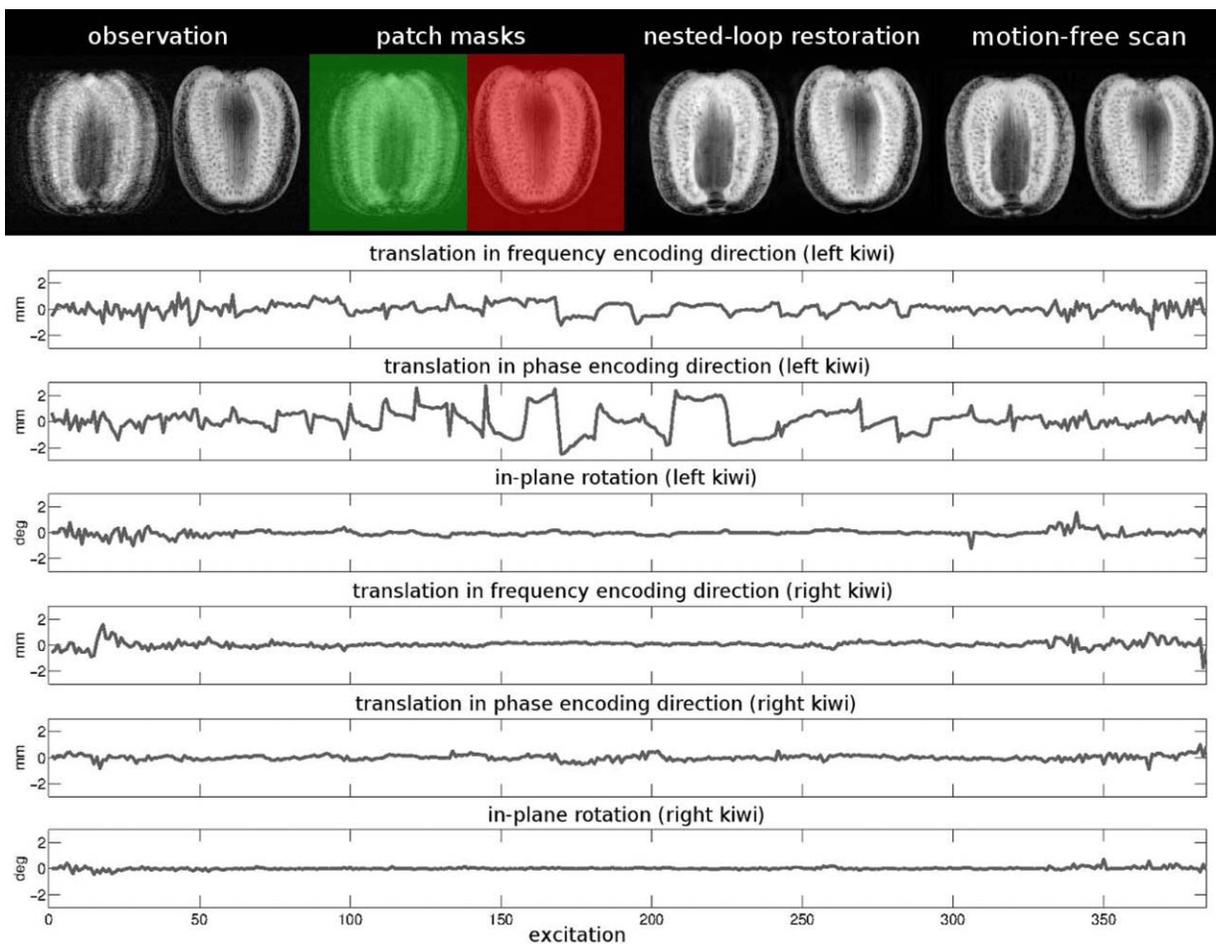


FIG. 6. Real data. Multirigid motion correction was performed on the image split into two patches (indicated by red/green overlays). Top: from left to right, we show the observed image (of kiwi fruits), the spatial masks used, motion correction result, and ground-truth image from the motion-free scan. Bottom: the trajectories for both patches recovered by the algorithm.

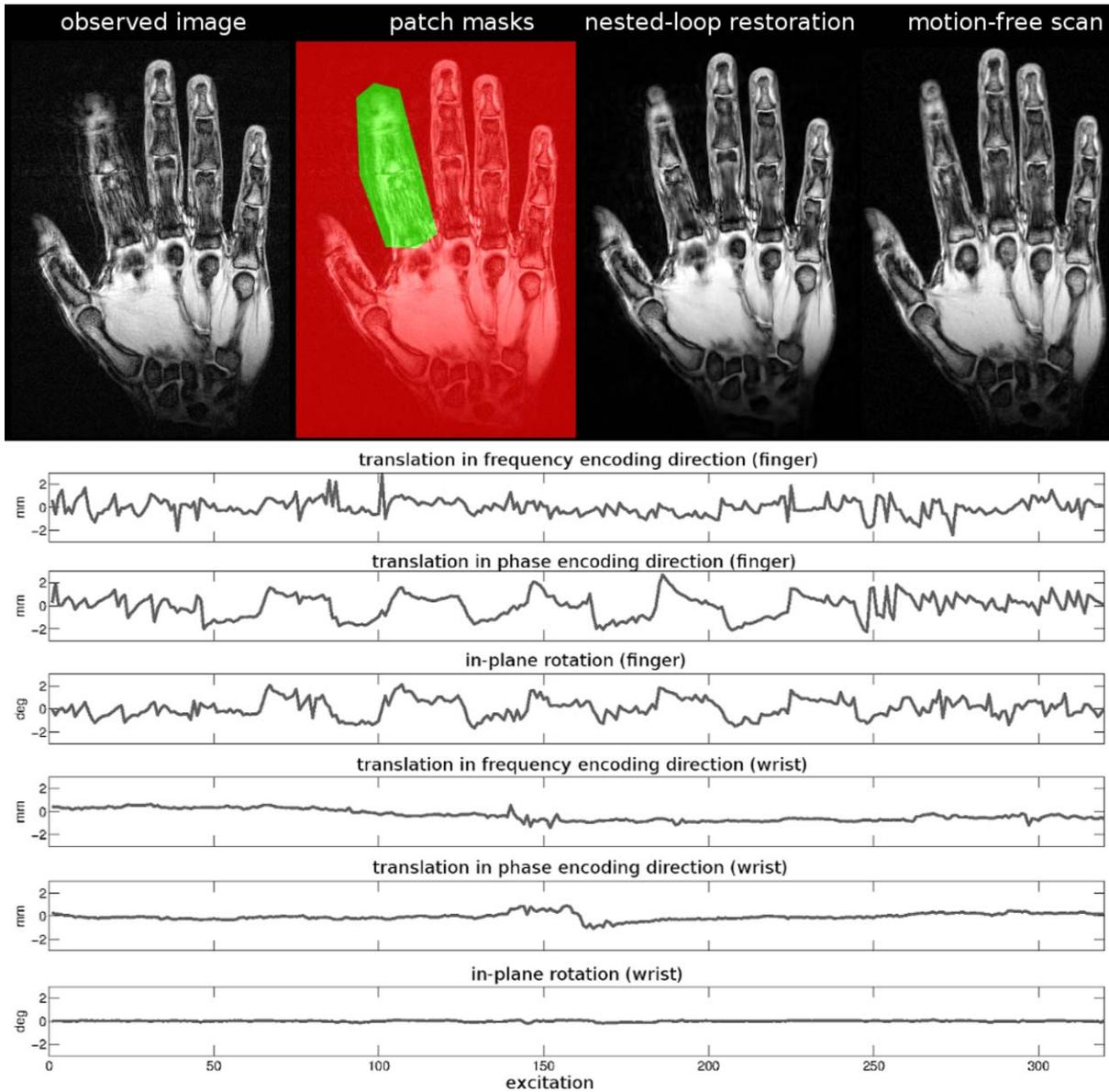


FIG. 7. Real data. Multirigid motion correction was performed on in vivo data, where the image of the human wrist was acquired. During the experiment, the subject was moving the index finger, and was trying to keep the hand stationary. Top: from left to right, we show the observed image, the spatial masks used, motion correction result, and ground-truth image from the motion-free scan. Bottom: the trajectories for both patches recovered by the algorithm.

experiment was applied to the blue patch. The green patch was not affected by motion. In contrast to the non-separable case, a great improvement of image quality is now evident.

#### Real Data

Similarly to the simulated data experiments, we first consider the global single-patch problem involving strong rotational motion. Figure 5 demonstrates the ability of our method to correct real data affected by strong rigid motion, and its superiority over our previous inverse-based method. We used  $N_{\Theta} = N_{\mathbf{u}} = 50$  for

motion/image estimation loops, and the annealing schedule of  $\lambda = [e^1, e^2, e^3]$ .

For our next experiments, we used a wrist coil to acquire images of kiwi fruits and a hand of a human subject. Note that the spatial windows (indicated by color-coded overlays) were arranged in phase-encode direction leading to hard, nonseparable problems. For the first real data experiment (see Figure 6), we acquired the image of two kiwi fruits. During the acquisition, we were using MR-neutral actuator to displace the left kiwi fruit. Doing the motion correction, we split the image into two patches, and used our multirigid nested-loop algorithm. To evaluate the quality of reconstructed images, we additionally acquired a motion-free ground-truth image. As

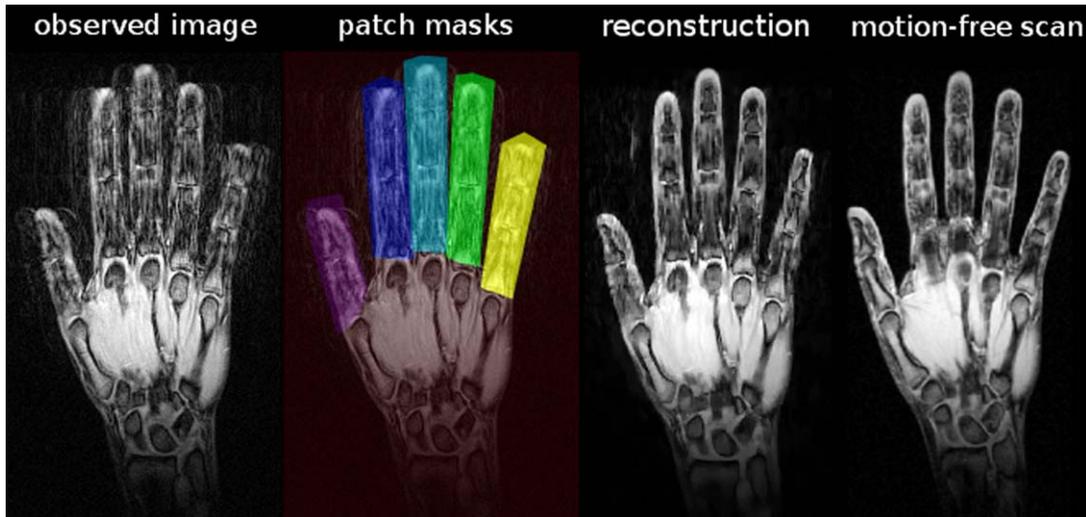


FIG. 8. Real data. Multirigid motion correction was performed on the image split into  $P=6$  patches. The human subject was moving all six fingers during the experiment.

can be seen in Figure 6, the algorithm was able to remove strong artifacts, and reconstructed results look very similar to ground-truth.

In the next experiment, we acquired in vivo data of the human hand (see Figure 7). The subject was asked to move the index finger and keep the wrist stationary. Also in this case with a more biologically plausible motion our algorithm was able to significantly improve the image quality. Additionally, in the bottom part of the figures, we plot the motion trajectories recovered by the algorithm. The motion parameters estimated in the left kiwi and index finger patches indicate a strong motion. Please note, however, that the motion trajectories recovered in nonmotion patches also contain significant variation. This is due to the fact that blind autofocusing motion correction approaches suffer from an ambiguity in the recovered motion parameters, which does not influence the image quality of the recovered image (17).

Additionally, if strong regularization of the image is used, this has an effect of penalizing high-frequency information, and means that the recovered image is tolerant to variation in high-frequency excitation motion parameters.

Next, we tested our algorithm on a more complicated problem, which involved multiple patches (Fig. 8). The subject was moving all fingers simultaneously trying to displace them in a disorganized manner, so that each finger has a unique motion trajectory. Also in this case major improvement of the image quality is possible. Doing the reconstructions we used the following algorithm parameters:  $N_{\Theta} = N_u = 50$ ,  $\lambda = [e^1, e^2]$ .

Finally, we tested our method on a 3D dataset, which was also corrupted by multirigid motion. In Figure 9, we show four slices from the motion-corrupted, corrected, and motion-free volumes. A significant improvement of the image quality can be observed in the reconstructed

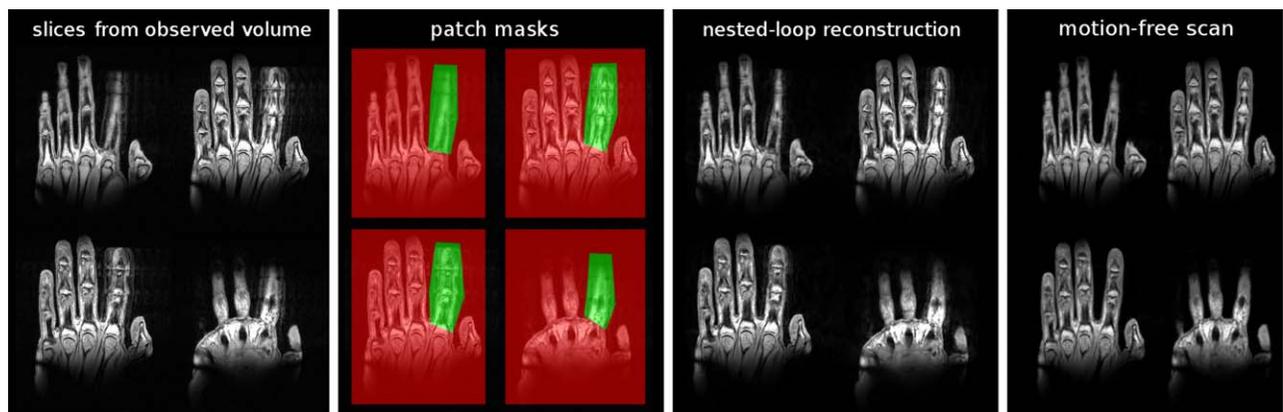


FIG. 9. Real data. Multirigid motion correction was performed on the 3D image of the human hand. During the experiment, the subject was moving the index finger. Left: four slices from the observed 3D dataset (containing 32 slices in total). Middle: spatial windows. Right: multirigid motion correction result, and the slices from the reference volume not affect by motion.

result. We used the following algorithm parameters:  $N_{\Theta} = 100$ ,  $N_{\mathbf{u}} = 50$ ,  $\lambda = [e^1, e^2]$ , and three multiscale iterations.

## DISCUSSION

In this proof-of-concept study, we have experimentally demonstrated the performance of multirigid motion correction using a fully blind retrospective method. Our technique is an extension of our previous work; the main innovation being the use of a forward model based objective function.

We assume that patch masks to be given as an input (although in theory they could be obtained by discrete optimization) either by the user or an automatic process like a coarse segmentation of the 3D image. Of course, this means that before using our algorithm the patches need to be provided (either by their coordinates or by a graphical interface). This is a difficult problem in itself if the multirigid motion is complex, or if the observation is covered by strong artifacts. In a practical setting, e.g., knee or hand imaging, the patch splitting can be obtained by segmentation or atlas registration so that the degrees of freedom of the imaged body parts are reflected in the parts of the segmentation. Another way to avoid the need of user-created masks is a regular patch splitting of the image. With increasing patch density, any nonrigid motion can be described, an extreme case being the vector field over all voxels. At this extreme the problem is, of course, computationally very hard to solve, furthermore, it is severely underconstrained given the mere number of parameters to optimize. Further research will go into the direction of coupling the motion parameters over patches for reducing the effective number of parameters to optimize. In particular, upgrading our algorithm to use an adaptive patch splitting (14) looks like a reasonable next step.

Compared to our previous inverse-based approach (17), the method presented in this article is computationally more demanding. Modern graphic cards render retrospective motion correction problems that were considered almost intractable years ago solvable in a reasonable time now. Although, in our current implementation, we are already experiencing massive time savings due to parallelization there is still room for improvement by code optimization. For example, when performing line searches in the objective landscape, one could avoid to compute the gradient direction on each step.

Our new algorithm has more parameters to tune compared to our previous inverse-based procedure, which is robust enough to be used out of the box. The tuning is a price to pay for being able to solve the more complex multirigid problem. In our current research, we faithfully report all the parameters that we were using in our experiments. Future research might allow to simplify the choice of parameters.

In this study, the validity of our approach was demonstrated on somewhat artificial instances of multirigid motion. Our ultimate goal is to address the problem of motion correction in abdomen data sets, i.e., images of liver or kidneys, which are affected by breathing. This is challenging for many reasons, first, motion due to breath-

ing is quite strong, and second, the task of splitting the abdomen data into patches is nontrivial. Usually there are many different tissues in the FOV, and they are likely to possess different modes of motion—i.e., spine, and walls of the abdomen are likely not to be moving, while the internal organs are affected by translations, and possibly small rotations.

While in the current study, we assumed the data were acquired with a single coil, this may be a difficult setting compared to multirigid motion correction from multicoil data. This is due to the fact that spatial sensitivity profiles of the coils naturally disentangle the effects (artifacts) of nonrigid motion over the spatial domain that otherwise propagate globally in phase-encode direction. Recent nonrigid motion correction methods (see Introduction) operate on data from multiple coils and should strongly benefit from this property. The bigger challenge is that in practice the input to the reconstruction algorithm would likely be a magnitude-only DICOM image. Thus, not only the information from each coil is not accessible but also the spatial phase in the image is gone. To our knowledge, there is no method for doing retrospective motion correction of both rigid and nonrigid motion based on pure magnitude data.

In summary, we have developed a retrospective method to correct a multirigid motion that does not need external information on motion as an input. At the core of our new method is an analytic nonlinear optimization of objective function of both the image and motion subject to fidelity and regularization constraints. We have demonstrated the ability of the method to correct for motion on both simulated and real data sets.

## APPENDIX

### Connection to Inverse-Based Approach

Assuming  $P = 1$ ,  $\mu = 0$ ,  $\mathbf{A}_{\Theta}$  to be invertible and its inverse being a valid motion matrix  $\mathbf{A}_{\Theta'} := \mathbf{A}_{\Theta}^{-1}$ , and substituting  $\mathbf{u} = \mathbf{F}^H \mathbf{A}_{\Theta} \mathbf{z}$  we can rewrite the objective  $\psi(\mathbf{z}, \Theta) = \phi(\mathbf{F}^H \mathbf{A}_{\Theta} \mathbf{z}) + \lambda \cdot \|\mathbf{z} - \mathbf{y}\|^2$ , where  $\mathbf{z}$  is an auxiliary variable. For large<sup>1</sup>  $\lambda$  we obtain  $\mathbf{y} = \mathbf{z}$  and hence  $\psi(\Theta) = \phi(\mathbf{F}^H \mathbf{A}_{\Theta} \mathbf{y})$ , which is the direct inverse objective function we used in our previous work (17, Eq. 9). Although it allows for fast motion correction, the inverse-based objective has clearly a number of limitations, i.e., it does not allow for valid reconstruction in case rotational motion is stronger than few degrees.

## REFERENCES

1. Plathow C, Ley S, Zaporozhan J, Schöbinger M, Gruenig E, Puderbach M, Eichinger M, Meinzer HP, Zuna I, Kauczor HU. Assessment of reproducibility and stability of different breath-hold manoeuvres by dynamic MRI: comparison between healthy adults and patients with pulmonary hypertension. *Eur Radiol* 2006;16:173–179.
2. Soellinger M, Rutz AK, Kozerke S, Boesiger P. 3d cine displacement-encoded MRI of pulsatile brain motion. *Magn Reson Med* 2009;61:153–162.

<sup>1</sup>From the theory of regularization paths, we know that we do not need to drive  $\lambda$  to infinity but there exists a threshold beyond which the optimizer will always obey  $\mathbf{z} = \mathbf{y}$ .

3. Maclaren J, Herbst M, Speck O, Zaitsev M. Prospective motion correction in brain imaging: a review. *Magn Reson Med* 2012;69:621–636.
4. Manke D, Rösch P, Nehrke K, Börner P, Dössel O. Model evaluation and calibration for prospective respiratory motion correction in coronary MR angiography based on 3d image registration. *IEEE Trans Med Imaging* 2002;21:1132–1141.
5. Schäffter T, Rasche V, Carlsen IC. Motion compensated projection reconstruction. *Magn Reson Med* 1999;41:954–963.
6. Batchelor P, Atkinson D, Irarrazaval P, Hill D, Hajnal J, Larkman D. Matrix description of general motion correction applied to multishot images. *Magn Reson Med* 2005;54:1273–1280.
7. Odille F, Cindea N, Mandry D, Pasquier C, Vuissoz PA, Felblinger J. Generalized MRI reconstruction including elastic physiological motion and coil sensitivity encoding. *Magn Reson Med* 2008;59:1401–1411.
8. Schmidt J, Buehrer M, Boesiger P, Kozerke S. Nonrigid retrospective respiratory motion correction in whole-heart coronary MRA. *Magn Reson Med* 2011;66:1541–1549.
9. Odille F, Uribe S, Batchelor PG, Prieto C, Schäffter T, Atkinson D. Model-based reconstruction for cardiac cine MRI without ECG or breath holding. *Magn Reson Med* 2010;63:1247–1257.
10. Vuissoz PA, Odille F, Fernandez B, Lohezic M, Benhadid A, Mandry D, Felblinger J. Free-breathing imaging of the heart using 2d cine-GRICS (generalized reconstruction by inversion of coupled systems) with assessment of ventricular volumes and function. *J Magn Reson Imaging* 2012;35:340–351.
11. Hansen M, Sørensen T, Arai A, Kellman P. Retrospective reconstruction of high temporal resolution cine images from real-time MRI using iterative motion correction. *Magn Reson Med* 2012;68:741–750.
12. White MJ, Hawkes DJ, Melbourne A, Collins DJ, Coolens C, Hawkins M, Leach MO, Atkinson D. Motion artifact correction in free-breathing abdominal MRI using overlapping partial samples to recover image deformations. *Magn Reson Med* 2009;62:440–449.
13. 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;60:146–157.
14. Menini A, Vuissoz PA, Felblinger J, Odille F. Joint reconstruction of image and motion in MRI: implicit regularization using an adaptive 3D mesh. *Med Image Comput Assist Interv* 2012;15:264–271.
15. Cheng JY, Alley MT, Cunningham CH, Vasawala SS, Pauly JM, Lustig M. Nonrigid motion correction in 3d using autofocusing with localized linear translations. *Magn Reson Med* 2012;68:1785–1997.
16. Usman M, Atkinson D, Odille F, Kolbitsch C, Vaillant G, Schaeffter T, Batchelor PG, Prieto C. Motion corrected compressed sensing for free-breathing dynamic cardiac MRI. *Magn Reson Med* 2012;70:504–516.
17. Loktyushin A, Nickisch H, Pohmann R, Schölkopf B. Blind retrospective motion correction of MR images. *Magn Reson Med* 2013;70:1608–1618.
18. Bracewell RN. *The Fourier transform and its applications*. New York: McGraw Hill; 1999.
19. Hirsch M, Sra S, Schölkopf B, Harmeling S. Efficient filter flow for space-variant multiframe blind deconvolution. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, California, USA, 2010. pp 607–614.
20. Wood M, Henkelman R. MR image artifacts from periodic motion. *Med Phys* 1985;12:143–151.
21. Morelli J, Runge V, Ai F, Attenberger U, Vu L, Schmeets S, Nitz W, Kirsch J. An image-based approach to understanding the physics of MR artifacts. *Radiographics* 2011;31:849–866.
22. McGee K, Manduca A, Felmlee J, Riederer S, Ehman R. Image metric-based correction (autocorrection) of motion effects: analysis of image metrics. *J Magn Reson Imaging* 2000;11:174–181.
23. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. *SIAM J Sci Stat Comput* 1995;16:1190–1208.
24. Fergus R, Singh B, Hertzmann A, Roweis ST, Freeman WT. Removing camera shake from a single photograph. *ACM Trans Graphics* 2006;25:787–794.
25. Cho S, Lee S. Fast motion deblurring. *ACM Trans Graphics* 2009;28:145:1–145:8.