Self-supervised motion-corrected image reconstruction network for 4D magnetic resonance imaging of the body trunk

Kustner, Thomas; Pan, Jiazhen; Gilliam, Christopher; Qi, Haikun; Cruz, Gastao; Hammernik, Kerstin; Blu, Thierry; Rueckert, Daniel; Botnar, René; Prieto, Claudia; Gatidis, Sergios

DOI:
10.1561/116.00000039

APSIPA Transactions on Signal and Information Processing, 2022, 11, e12

Original Paper

Self-Supervised Motion-Corrected Image Reconstruction Network for 4D Magnetic Resonance Imaging of the Body Trunk

Thomas Küstner[1], Jiazhen Pan[2], Christopher Gilliam[3], Haikun Qi[4], Gastao Cruz[5], Kerstin Hammernik[2,6], Thierry Blu[7], Daniel Rueckert[2,6], René Botnar[5,8], Claudia Prieto[5,8] and Sergios Gatidis[1*]

[1] *Medical Image And Data Analysis (MIDAS.lab), Department of Diagnostic and Interventional Radiology, University Hospital Tübingen, Tübingen, Germany*

[2] *Lab for Artificial Intelligence in Medicine, Technical University of Munich, Munich, Germany*

[3] *RMIT University, Melbourne, Australia*

[4] *School of Biomedical Engineering, ShanghaiTech University, Shanghai, China*

[5] *School of Biomedical Engineering and Imaging Sciences, King's College London, St. Thomas' Hospital, London, UK*

[6] *Biomedical Image Analysis Group, Department of Computing, Imperial College London, London, UK*

[7] *Chinese University Hong Kong, Hong Kong*

[8] *Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Santiago, Chile*

*Corresponding author: Thomas Küstner, thomas.kuestner@med.uni-tuebingen.de. The authors would like to thank Brigitte Gückel for study coordination as well as Carsten Groeper and Gerd Zeger for data acquisition.

Received 29 July 2021; Revised 27 December 2021

ISSN 2048-7703; DOI 10.1561/116.00000039

ABSTRACT

Respiratory motion can cause artifacts in magnetic resonance imaging of the body trunk if patients cannot hold their breath or triggered acquisitions are not practical. Retrospective correction strategies usually cope with motion by fast imaging sequences under free-movement conditions followed by motion binning based on motion traces. These acquisitions yield sub-Nyquist sampled and motion-resolved k-space data. Motion states are linked to each other by non-rigid deformation fields. Usually, motion registration is formulated in image space which can however be impaired by aliasing artifacts or by estimation from low-resolution images. Subsequently, any motion-corrected reconstruction can be biased by errors in the deformation fields. In this work, we propose a deep-learning based motion-corrected 4D (3D spatial + time) image reconstruction which combines a non-rigid registration network and a 4D reconstruction network. Non-rigid motion is estimated in k-space and incorporated into the reconstruction network. The proposed method is evaluated on in-vivo 4D motion-resolved magnetic resonance images of patients with suspected liver or lung metastases and healthy subjects. The proposed approach provides 4D motion-corrected images and deformation fields. It enables a $\sim 14\times$ accelerated acquisition with a 25-fold faster reconstruction than comparable approaches under consistent preservation of image quality for changing patients and motion patterns.

Keywords: Motion-compensated image reconstruction, Magnetic Resonance Imaging, Image registration, Deep learning reconstruction

1 Introduction

In clinical diagnostics, magnetic resonance imaging (MRI) is a valuable and versatile tool to assess anatomy and functional processes within the human body in a non-invasive manner. However, MRI is prone to several artifacts which can deteriorate image quality significantly. Due to its long acquisition time, motion is one of the major extrinsic factors influencing image quality. Patient and physiological motion induces aliasing along the phase-encoding direction and/or blurring of the image content, where the appearance depends on the imaging trajectory.

Motion visualization, estimation and correction are thus important tasks when processing MRI data. Fast and accurate motion estimation and tracking is required to enable prospective or retrospective motion correction techniques which can be applied to, e.g., image guided interventions [50], cardiac assessment [48] or magnetic resonance (MR)-based motion correction of positron emission tomography (PET) data [38, 53]. Several prospective and retrospective motion correction methods have been developed which include fast imaging sequences [17, 45], tracking of motion by sensors (MR navigators [22, 30, 66], cameras [46], respiratory belts or electrocardiogram [75]), application of motion-robust acquisition schemes [3], prospectively corrected acquisitions [67], and motion-resolved imaging [10, 19, 31, 40, 72].

Global translations or rotations of stiff structures describe rigid motion and arises from movement of body parts like head motion. Rigid motion can be modelled in k-space as linear phase drifts and incorporated into acquisition as prospective correction scheme or into reconstruction as retrospective correction scheme. Non-rigid motion, i.e., local deformations of tissues, mainly occurs in the thorax and abdominal region caused by physiological motion. However, local deformations in image space are related to non-trivial changes in the acquired k-space, which imposes further challenges in motion correction. Correction of non-rigid motion usually involves two steps: image reconstruction and image registration from motion-resolved data.

Motion-resolved data acquisition for these applications are usually accelerated by Parallel Imaging or Compressed Sensing techniques yielding sub-Nyquist sampled (in the following denoted as subsampled) k-space data. In order to reconstruct aliasing-free images these methods rely on reconstruction schemes that, for example incorporate sparsity or low-rank constraints to solve the ill-posed problem [45, 56]. Fixed sparsity assumptions in Compressed Sensing are often too restrictive and incapable of fully modelling spatio-temporal dynamics. Careful fine-tuning between regularization and data consistency is required and especially in highly subsampled cases residual aliasing may remain in the image (under-regularization) or staircasing and blurring artifacts can occur (over-regularization) which affect the image registration.

After reconstruction, non-rigid motion fields can be estimated in image space from reconstructed images by solving a registration problem. A particular interest and challenge lies in the derivation of reliable motion fields which capture the spatio-temporal non-rigid deformations, such as respiratory or cardiac movement. The non-rigid motion estimation problem can be formulated in image space using diffusion-based [69], parametric spline-based [61] or optical flow-based methods [23].

Instead of performing these two steps sequentially, motion-compensated image reconstruction schemes [1, 2, 11, 21, 54, 55, 71] integrate both motion field estimation and motion correction into the reconstruction process. These methods require reliable motion-resolved images from which the motion fields can be estimated. Motion field estimation can be controlled or supported by external motion surrogate signals [11, 54], initial motion field estimates [1, 2], from motion-aliased images [21] or low-frequency image contents [71].

Moreover, spatio-temporal redundancies can be exploited to achieve an aliasing-free image [5, 8, 27, 32, 43, 44, 52]. While these methods have been proven to be more robust against registration errors, they can require a significantly increased computational demand and/or limit imaging acceleration.

In case of highly subsampled data, aliasing artifacts in the reconstructed images can impair the registration process as reconstruction errors can propagate into the image registration and/or low-resolution images do not provide sufficient information for accurate registration. Moreover, higher subsampling leads to a challenging ill-posed reconstruction problem for which inherent spatio-temporal redundancies need to be better exploited. Recently deep-learning based reconstruction methods have been promoted to target this area [7, 9, 12, 13, 18, 20, 24, 26, 29, 33, 42, 49, 59, 62, 64, 68, 73, 74, 77]. Network inputs thereby differ from single-coil 2D image [12, 64, 73] and/or k-space [12, 68, 77] to multi-coil 2D image [7, 41, 49, 63], 2D k-space [6, 9, 20, 24, 63, 73] or low-resolution 3D k-space [47]. The works studied static imaging [7, 12, 13, 18, 20, 26, 41, 49, 68, 73, 74, 77] and dynamic imaging [33, 59, 62, 64], i.e., exploiting spatio-temporal dynamics. Recently, works investigated the possibility to combine reconstruction networks with image-based registrations [25, 58, 65].

The novel contributions of this work are the integration of an unrolled reconstruction and registration network into a motion-corrected reconstruction network. We extend our previously proposed reconstruction network [33] for a self-supervised motion-corrected 4D (3D spatial + time) reconstruction. We thereby exploit spatio-temporal redundancies as implicit motion handling via separable spatial-temporal convolution filters and as explicit motion handling via motion fields derived from the motion registration network [34–36]. The non-rigid motion information is directly extracted from the subsampled k-space data, based on the idea of optical flow registration [16, 37]. The proposed method which operates on the subsampled k-space data is compared against an image-based registration paired with a motion-corrected iterative SENSE reconstruction [2]. We investigate the proposed approach in 36 patients with suspected liver or lung metastases and 20 healthy subjects for retrospectively subsampled data of 3D motion-resolved MR imaging.

2 In-Vivo 4D MR Acquisition

To investigate the proposed motion-corrected reconstruction network, motion-resolved k-space data were obtained on a cohort of 36 patients (60 ± 9 years, 20 female) with suspected liver or lung metastases and 20 healthy subjects (31 ± 4 years, 9 female) [38]. The study was approved by the local ethics committee and all subjects gave written consent.

A 3D T1 weighted spoiled gradient echo sequence was acquired on a 3T PET/MR (Biograph mMR; Siemens Healthcare) in coronal orientation with a variable-density Poisson Disc subsampling [39] for an acquisition time of 90 seconds (prospectively subsampled) and 300 seconds (reference). The remaining imaging parameters were TE = 1.23 ms, TR = 2.60 ms, bandwidth = 890 Hz/px and a flip angle of 7°. A matrix size of $N_x \times N_y \times N_z = 256 \times 256 \times 144$ (RO $\times$ PE $\times$ 3D $\Leftrightarrow$ HF $\times$ LR $\times$ AP) was acquired covering a field-of-view of $500 \times 500 \times 360\,\mathrm{mm}^3$. A 2D MR self-navigation signal ($256 \times 8 \times 1$, RO$\times$PE$\times$3D) was acquired each 200 ms serving as gating signal. MR data was retrospectively gated into $N_t = 8$ respiratory gates, ranging from end-expiratory to end-inspiratory position, with a Gaussian view-sharing amongst neighbouring gates. An average acceleration factor per motion gate of $\sim 14\times$ (prospectively subsampled) and $\sim 2\times$ (reference) was obtained [38]. The coil sensitivity map was estimated from the time-averaged fully sampled calibration center region by ESPIRIT [70] with virtual coil compression to a common size of $N_{ch} = 8$.

3 Non-Rigid Registration in k-Space

For the motion-compensated reconstruction, a reliable estimation of motion fields is required. Although most (intensity-based) registration algorithms are to a certain extent robust to blurring and/or noise amplification, they are very sensitive to coherent aliasing artefacts. A motion field estimation from the subsampled k-space data can avoid this potential impairment from aliasing and blurring. The non-rigid registration follows the concept of the LAP algorithm [15]. The key idea of LAP is that any non-rigid deformation can be regarded as local translational displacements. A local translation on the other hand can be regarded as an all-pass operation in Fourier space, i.e., non-rigid motion can be modelled as local all-pass filter operations.

Under the assumption of local brightness consistency and a motion flow continuum, the optical flow equation of a displacement can be stated in discrete (discrete domain is used throughout the manuscript) Fourier space

$$\rho_f(\underline{x}) = \rho_m(\underline{x} - \underline{u}_{\underline{x},m}) \Longleftrightarrow \nu_f(\underline{k}) \simeq \nu_m(\underline{k}) e^{-j\underline{u}^{\mathrm{T}}_{\underline{x},m}\underline{k}} \tag{1}$$

for deforming a moving image ρ_m to a fixed image ρ_f via a deformation field $u_{x,m} = [u_x, u_y, u_z]^{\mathrm{T}}$ at image position $\underline{x} = [x, y, z]^{\mathrm{T}}$ and of motion state m (i.e., warping moving state m to fixed state f), with $\nu_f(\underline{k})$ and $\nu_m(\underline{k})$ being the k-spaces of the fixed and moving image at k-space sampling location $\underline{k} = [k_x, k_y, k_z]^{\mathrm{T}}$ (underscore notation denotes a vector, $[\cdot]^{\mathrm{T}}$ is the transposed vector and $j = \sqrt{-1}$ represents the imaginary unit). The linear phase ramp can be regarded as an all-pass filter $H(\underline{k}) = e^{-j\underline{u}^{\mathrm{T}}\underline{k}} = F(\underline{k})/F(-\underline{k})$ that can be split into a forward $F(\underline{k})$ and backward filter $F(-\underline{k})$ which is all-pass by design, i.e., $|H(\underline{k})| = 1$ [14]. Global non-rigid deformation is thus modelled as local translational displacements and hence the problem of non-rigid image registration is transformed to estimating the appropriate local all-pass filter for different $\underline{x}$ in a cubic window W. The phase-modulated (for various $\underline{x}$ positions) at motion state i tapering function $\mathrm{T}(\underline{k})$ for the local window W

$$W(\underline{x}) \cdot \rho_i(\underline{x}) \Longleftrightarrow \mathrm{T}(\underline{k}) * \nu_i(\underline{k}) \tag{2}$$

allows us to conclude the non-rigid registration formulation in k-space (Fourier domain)

$$\begin{aligned} &\min_{\{c_n\}} \sum_{\underline{k} \in R^3} (\mathrm{T}(\underline{k}) * (F(\underline{k})\nu_f(\underline{k})), \mathrm{T}(\underline{k}) * (F(-\underline{k})\nu_m(\underline{k})))^2 \\ &\text{s.t. } F(\underline{k}) = F_0(\underline{k}) + \sum_{n=1}^{N_F} c_n F_n(k) \quad \forall \underline{k} \in \mathbb{R}^3 \end{aligned} \tag{3}$$

At each k-space position $\underline{k}$, the N_F optimal filters $c_n F_n$ are estimated [4] by minimizing the dissimilarity between ν_m and ν_f. The deformation field $\underline{u}$ of motion state m in the image domain can be directly derived from the all-pass filter

$$\underline{u}_m = j \left. \frac{\partial \ln H(\underline{k})}{\partial \underline{k}} \right|_{\underline{k}=\underline{0}} = 2 \left[\frac{\sum_{\underline{x}} x f(\underline{x})}{\sum_{\underline{x}} f(\underline{x})}, \frac{\sum_{\underline{x}} y f(\underline{x})}{\sum_{\underline{x}} f(\underline{x})}, \frac{\sum_{\underline{x}} z f(\underline{x})}{\sum_{\underline{x}} f(\underline{x})} \right]^{\mathrm{T}} \tag{4}$$

The optimal all-pass filters $F(\underline{k})$ that solve Equation (3) are learnt by the convolutional filters in the registration network. It is expected that each convolutional layer learns an optimal filter $c_n F_n$ that achieves diffeomorphic and smooth flows. Real and imaginary part of the moving and fixed k-space are passed through a succession of $N_F = 6$ 3×3 convolutional filters with dyadic increase in kernel size (starting kernel size 64) and leaky ReLU activation function. In the last layer a fully connected regression is performed on the average pooled feature map to estimate the in-plane deformations u_{11}, u_{21} at the given central location of the input patch after the first run. To obtain a 3D deformation field $\underline{u}$, the registration is also performed on an orthogonal direction, yielding u_{12}, u_{22} that are merged with the previous run to obtain $u_x = u_{11}$, $u_y = 0.5(u_{21} + u_{12})$, $u_z = u_{22}$. The whole non-rigid deformation field $\underline{u}_{\underline{x},m}$ warping motion state m to the fixed state f is obtained by estimating the deformations u_x, u_y, u_z at all voxel locations $\underline{x}$. This principle follows the idea of approximating a global non-rigid flow by local translational deformations.

4 Motion-Corrected Image Reconstruction

The proposed motion-compensated network architecture consists of two sub-networks as depicted in Figure 1: A 3D non-rigid registration network [36] which

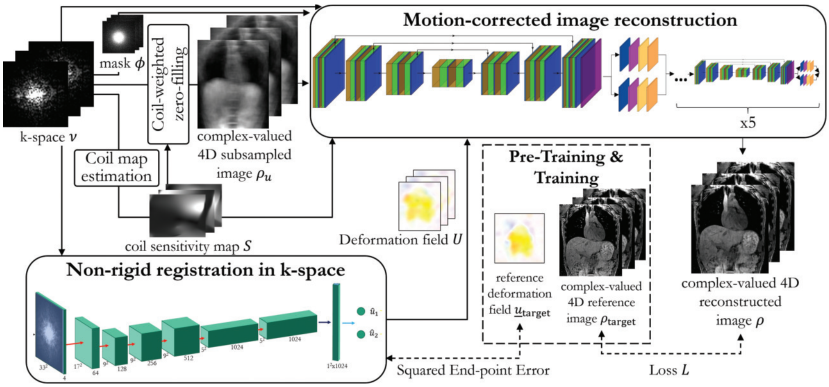

Figure 1: Proposed motion-compensated 4D reconstruction network which consists of two sub-networks: A non-rigid registration network directly operating on motion-resolved k-spaces and a reconstruction network employing 3D spatial and 1D temporal convolutions to exploit spatio-temporal redundancies. The reconstruction network consists of six cascaded UNet regularizers with intermittent data consistency blocks from a coil-weighted zero-filled image ρ_u. The estimated deformation fields U, the auto-calibrated coil sensitivity maps S, the k-space ν and the sampling mask ϕ are incorporated in the data consistency blocks to reconstruct a motion-corrected image ρ for each motion state. The registration network is pre-trained in a supervised manner deploying a squared end-point error loss to reference deformation fields derived via Local All-Pass. The joint motion-compensated reconstruction network is then trained in a self-supervised manner to minimize the combined complex-valued photometric and mean-squared error loss L.

provides the motion fields and a $(3+1)$D reconstruction network [33] which includes the estimated motion in the data consistency block to reconstruct an aliasing-free and motion-corrected image. Subsampled and motion-resolved k-spaces $\nu \in \mathbb{C}^{N_x N_y N_z N_t N_{Ch}}$ serve as input from which 3D non-rigid deformation fields $\underline{u} \in \mathbb{R}^{N_x N_y N_z N_t N_t \cdot 3}$ between all motion state pairs are estimated. N_x, N_y and N_z reflect the 3D spatial dimensions, N_t the temporal direction, and N_{Ch} the channels of the multi-coil MR receiver array. The coil sensitivity map $S \in \mathbb{C}^{N \times N}$ with $N = N_x N_y N_z N_{Ch}$ is derived from the k-space ν. The SENSE combined subsampled 4D image $\rho_u \in \mathbb{C}^{N_x N_y N_z N_t}$ is reconstructed to an aliasing-free and motion-corrected image $\rho \in \mathbb{C}^{N_x N_y N_z N_t}$ for each motion state.

A physics-based unrolled reconstruction is used for motion-corrected reconstruction [33], consisting of cascaded $(3+1)$D UNets and intermittent data consistency blocks. The network operates on multi-coil complex-valued 4D (3D + time) data. It introduces a series of 3D spatial and 1D temporal complex-valued convolutional filters paired with motion field warping in the data consistency blocks. The input to the network is the complex-valued subsampled and motion-resolved image ρ which was reconstructed with a coil-weighted zero-filling, as well as the acquired k-space ν, the sampling mask

ϕ, the coil sensitivity map S and the 3D motion fields $\underline{u}$ from the registration network between all pairs of the motion-resolved data restacked into the sparse matrix $U \in \mathbb{R}^{N \times N}$ with $N = N_x N_y N_z N_t \cdot 3$. The output of the motion-corrected reconstruction is a motion-resolved image in which each motion state utilized information from the remaining motion states. The unconstrained motion-compensated MR reconstruction problem is thus given by

$$\min_{\rho} \mathcal{R}(\rho; \Theta) + \lambda \|E\rho - \nu\|_2^2 \tag{5}$$

where $E = \phi FSU$ is the encoding operator, F denotes the discrete Fourier transform, $\|\cdot\|_2$ is the ℓ_2 norm and $\lambda > 0$ is the data consistency weighting parameter. The regularizer $\mathcal{R}(\rho; \Theta)$ is expressed by the residual denoising $(3+1)$D UNet mappings $f_{\text{UNet}}(\rho; \Theta)$

$$\mathcal{R}(\rho; \Theta) = \|\rho - f_{\text{UNet}}(\rho; \Theta)\|_2^2 \tag{6}$$

with learnable parameters Θ. Combining Equations (5) and (6) allows us to formulate the alternating reconstruction algorithm

$$\rho^{(i+1)} = \arg\min_{\rho} \left\|\rho - z^{(i)}\right\|_2^2 + \lambda \|E\rho - \nu\|_2^2 \tag{7}$$

$$z^{(i)} = f_{\text{UNet}}\left(\rho^{(i)}; \Theta\right) \tag{8}$$

for step-wise unrolled updates at stage (i) of the image ρ. Subproblem (7) can be solved using conjugate gradient descent

$$\rho^{(i+1)} = \left(\lambda E^{\text{H}} E + I\right)^{-1} \left(\lambda E^{\text{H}} \nu + z^{(i)}\right) \tag{9}$$

which is incorporated as the data consistency block in the motion-corrected reconstruction network. The encoding operator E thereby explicitly steers sharing of spatio-temporal information amongst all motion states via the motion fields in U obtained from the k-space registration network. Inverse motion mappings U^{H} as required in Equation (9) consider the backward motion field obtained from the registration network.

The $(3+1)$D UNet mappings perform an implicit spatio-temporal information sharing via the $(3+1)$D convolutional filters. The motion fields in the data consistency blocks enable an explicit spatio-temporal sharing.

The $(3+1)$D UNet $f_{\text{UNet}}(\rho; \Theta)$ has four encoding and decoding stages which consist each of two pairs of complex-valued spatial convolutional layers of size $3 \times 3 \times 3 \times 1$ $(k_x \times k_y \times k_z \times t)$ followed by a temporal convolution of size $1 \times 1 \times 1 \times 2$ and complex ReLU activation. A complex convolution is performed. A dyadic increase in channel size is selected between scales, starting from 8 for the first scale. Residual paths between encoder/decoder improve convergence. In the encoder branch the last convolutional layer per stage uses a stride of 2 for down-sampling between stages while transposed convolutions are performed in the decoder side for up-sampling.

The registration network resulted in $\sim$ 25 million trainable parameters, and for six cascaded $(3+1)$D UNets, the motion-corrected reconstruction network results in $\sim$ 5.8 million trainable parameters.

The registration network was pre-trained in a supervised manner on pairs of moving and fixed k-space inputs with the corresponding target motion fields $\underline{u}_{\text{target}}$ derived from the iterative SENSE reconstructed images of the reference acquisition via image-based LAP [16, 37]. Flows were augmented by smoothing, translating, rotating and shearing. In total 15,000 training samples were generated which resulted after tapering in $\sim$ 150 million training samples. The squared end-point error (sEPE)

$$\text{sEPE} = L_{\text{Reg}} = \sum_{i \in \{x,y,z\}} \left(u_{\text{target},i} - u_i\right)^2 \tag{10}$$

was deployed as the training loss. Training was performed with an Adam optimizer [28] (learning rate $2.5 \cdot 10^{-4}$ with learning rate scheduler, batch size 64) over 150 k iterations on a Nvidia V-100 GPU (32 GB VRAM).

Afterwards, the proposed motion-corrected network was trained in a self-supervised manner (for registration network) on retrospectively subsampled reference data. The registration network was initialized with the pre-trained weights. Training data for the motion-compensated reconstruction network was generated by an iterative SENSE reconstruction [57] of the reference data which was retrospectively binned and subsampled by variable-density Poisson Disc for accelerations in the range of $2\times$ to $20\times$. The respective "fully-sampled" ($2\times$ accelerated) reference data served as target image ρ_{target}. A complex-valued mean-squared error and photometric loss

$$\begin{aligned} L = &\left\| [\text{Re}(\rho), \text{Im}(\rho)]^{\text{T}} - \left[\text{Re}\left(\rho_{\text{target}}\right), \text{Im}\left(\rho_{\text{target}}\right)\right]^{\text{T}} \right\|_2^2 \\ &+ 0.5 \cdot \sum_{t=1}^{N_t} \sum_{s \neq t} \left\| \rho_t - T_I(\rho_s, \underline{u}_s) \right\|_1 \end{aligned} \tag{11}$$

with warping T_I (bilinear interpolation) of the 3D image ρ_s at motion state s via the deformation field $\underline{u}_s$ into the 3D image ρ_t at motion state t is used as training loss to yield close agreement to the target image ρ_{target}. The loss is optimized by Adam [28] (learning rate 10^{-4}, batch size 16) and fixed data consistency parameter $\lambda = 10^{-3}$ on three Nvidia V-100 GPU (3×32 GB VRAM) for 40 epochs. The code is made publicly available under: https://github.com/midas-tum.

Overall, 50 subjects (33 patients and 17 healthy subjects) were used for training and 6 subjects (3 patients and 3 healthy subjects) were used in testing.

For training, the reference data were used while for testing the prospectively subsampled data were taken. A 5-fold cross-validation was performed.

5 Evaluation and Experiments

The proposed motion-corrected reconstruction framework was evaluated on prospectively subsampled data in 6 subjects. Motion fields were estimated from subsampled k-space data ($\sim 14\times$ accelerated) with the registration network and reconstructed with the motion-corrected $(3+1)$D reconstruction network. For comparison, two image-based 3D registrations using the image-based LAP (denoted as *imageLAP*) [16, 37, 38] and *NiftyReg* [51] were combined with a motion-corrected iterative SENSE reconstruction [2]. Registrations of these methods were performed on initial iterative SENSE reconstructed images. Comparative methods were run on an Intel Xeon E5-2697 CPU.

The end-point error $\text{EPE} = \left\|\underline{u} - \underline{u}_{\text{target}}\right\|_2$ and end-angulation error $\text{EAE} = \arg(\underline{u}, \underline{u}_{\text{target}})$ between the estimated motion field $\underline{u}$ of the prospectively subsampled acquisition was compared with the target motion field $\underline{u}_{\text{target}}$ obtained from an image-based LAP registration of the reference acquisition ($2\times$ accelerated). Structural similarity index (SSIM) [76] and normalized root MSE $(\text{NRMSE}) = 1/N\sqrt{\text{MSE}}$ (N being the number of voxels) were calculated between the motion-corrected image ρ and the target ρ_{target} of the reference scan at end-expiratory position. All quantitative results are reported as mean $\pm$ one standard deviation over all voxel positions, test subjects and cross-validations.

6 Results

The motion-corrected reconstruction in a healthy subject of the proposed framework is shown in Figure 2 in comparison to the image-based *imageLAP* and *NiftyReg*. The obtained motion fields are overlaid on the reconstructed and motion-corrected images as a vector field (forward deformation) pointing from end-expiratory ($t = 1$) to end-inspiratory state ($t = 8$). Additionally, the deformation fields are illustrated in coronal and sagittal orientation. We observe a significant portion of motion in superior-inferior and anterior-posterior direction. The registration network does not estimate any motion in the image background whereas the image-based methods try to match background voxels. The proposed approach thus helps to minimize background bleeding into image content and reduces noise amplification in the reconstruction. Reconstruction from subsampled images was possible in all cases with a markedly improved visual image quality and sharpness of the proposed approach. For the *imageLAP* and *NiftyReg*, residual blurring of the diaphragm at the lung-liver interface was

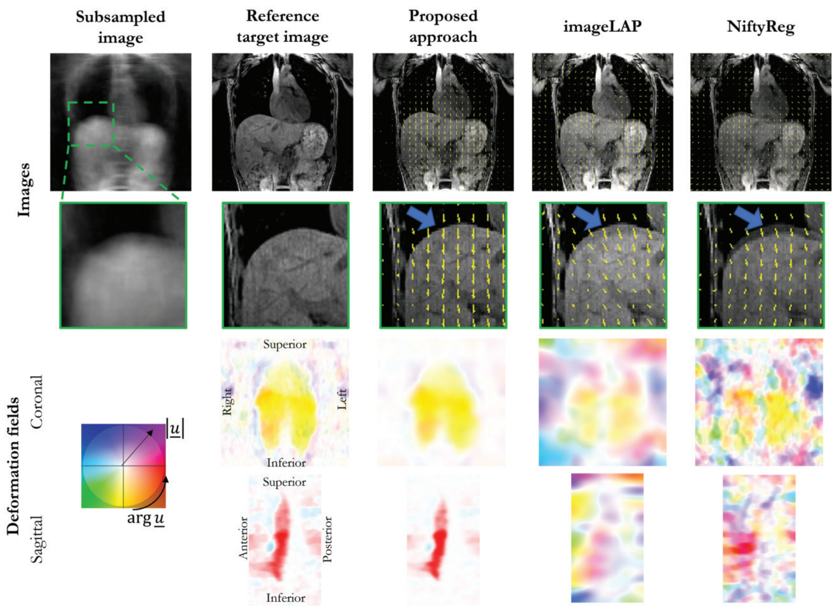

Figure 2: End-expiratory state of motion-corrected and reconstructed images in a healthy test subject for the proposed approach, image-based LAP registration with motion-corrected iterative SENSE reconstruction and NiftyReg registration with motion-corrected iterative SENSE reconstruction. Subsampled images were prospectively acquired with a subsampling of 16× reflecting an acquisition time of 90 seconds for $N_t = 8$ motion states. The reference target image represents the 2× accelerated acquisition of 300 seconds. Zoomed images of the liver dome are depicted. Deformation fields in coronal orientation are overlaid on motion-corrected images. Color-coded forward deformation fields are shown in coronal and sagittal orientation pointing from end-expiratory to end-inspiratory state. Blue arrows indicate residual motion blurring at the lung-liver interface.

observed (pointed out by arrows). Visually improved reconstruction quality was obtained with the proposed approach.

Images of a patient with pancreas carcinoma and liver metastasis in liver segment V are shown in Figure 3 for the proposed approach in comparison to *imageLAP* and *NiftyReg*. The proposed approach provides clear delineation of the liver lesion comparable to the reference scan as pointed out by the arrows. Image-based approaches suffer from residual blurring. Good quality images were obtained in an accelerated acquisition of ~ 90 seconds, reducing scan time and rendering clinically feasible for respiratory motion-compensated imaging of the body trunk.

Motion-resolved images over all $N_t = 8$ motion states of a patient with neuroendocrine tumor are depicted in Figure 4. The forward and backward deformation fields are overlaid on the respective obtained motion-corrected images of the proposed approach. Good and consistent image quality amongst

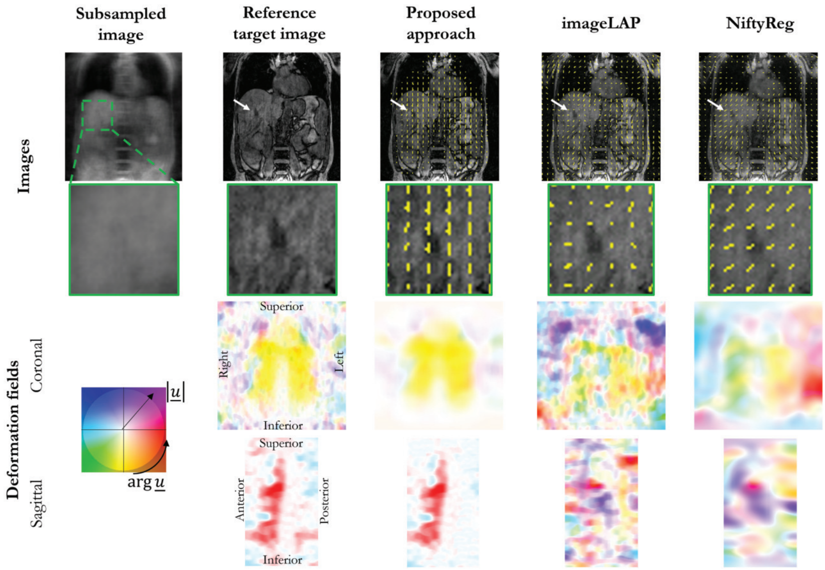

Figure 3: End-expiratory state of motion-corrected and reconstructed images in a patient with pancreas carcinoma and liver metastasis (pointed out by arrows). The proposed approach, image-based LAP registration with motion-corrected iterative SENSE reconstruction and NiftyReg registration with motion-corrected iterative SENSE reconstruction are shown. Zoomed images of the liver metastasis in liver segment V are depicted. Forward deformation fields in coronal orientation are overlaid on motion-corrected images. Color-coded forward deformation fields are shown in coronal and sagittal orientation pointing from end-expiratory to end-inspiratory state.

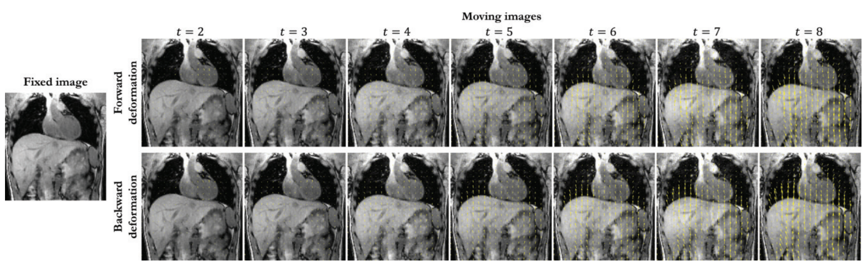

Figure 4: Reconstructed motion-corrected images of proposed approach in a patient with neuroendocrine tumor in the liver over complete respiratory cycle. Forward (end-expiratory state $t = 1$ to end-inspiratory state $t = 8$) and backward deformation fields in relation to fixed image ($t = 1$) are overlaid on motion-corrected moving images.

all motion states can be observed in our proposed approach, highlighting the benefit of explicit (via deformation fields in data consistency layer) and implicit (3D + 1D convolutional filters) temporal sharing. Diffeomorphic flows were obtained that enable a respiratory cyclic consistent motion sharing, i.e., during

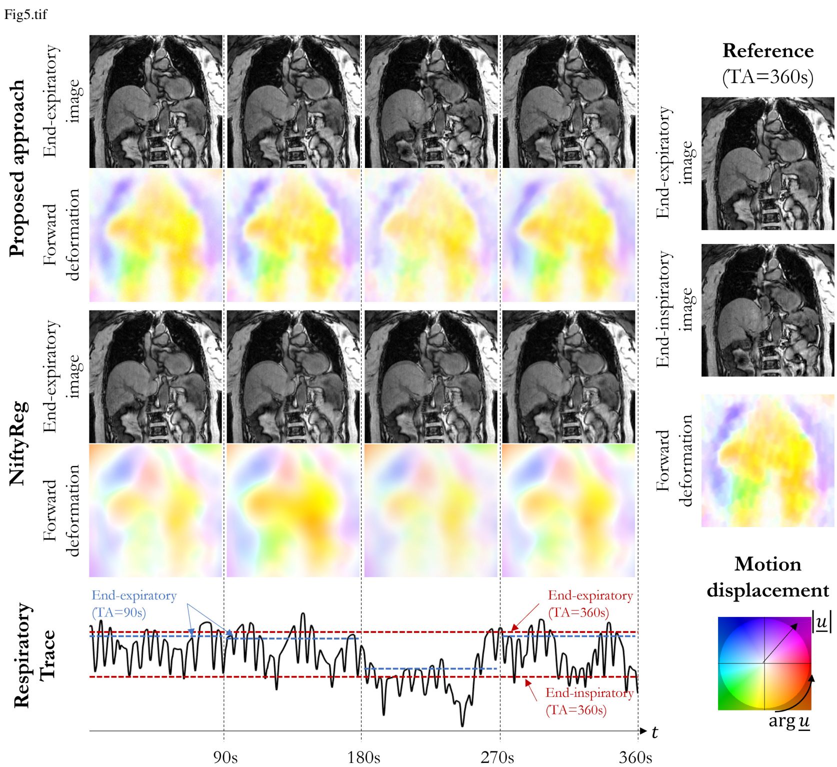

Figure 5: Respiratory non-rigid motion estimation in a patient with caecal carcinoma showing strong irregular breathing patterns. Motion displacement is estimated by the proposed LAPNet in k-space in comparison to image-based non-rigid registration by NiftyReg (cubic B-Splines). Motion-corrected images are reconstructed from an extended scan of 360 seconds as ~ 90 seconds long scans, corresponding to a $\sim 20\times$ subsampling. End-expiratory images and forward deformations (end-expiratory to end-inspiratory) are shown for these cases. In the respiratory trace, horizontal dashed lines mark the end-expiratory and end-inspiratory bins of the respective 90 s scans and of the reference 360 s scan. Reference images depict the end-expiratory and end-inspiratory image reconstructed by iterative SENSE and reference flow depict the imageLAP registration.

joint training only backward deformations can be estimated and forward fields can be inverted therefrom which saves computational costs.

The impact of changing motion patterns for a patient with caecal carcinoma can be seen in Figure 5 for an extended acquisition time of TA $= 360$ seconds. The extracted self-navigation signal reveals a strong irregular breathing pattern. Motion-corrected images were reconstructed from shorter ~ 90 seconds portions of the acquisition, reflecting in an $\sim 20\times$ acceleration per motion bin for this subject. The proposed approach was able to track the motion consistently and provided motion-corrected images with high quality. The image-based solutions

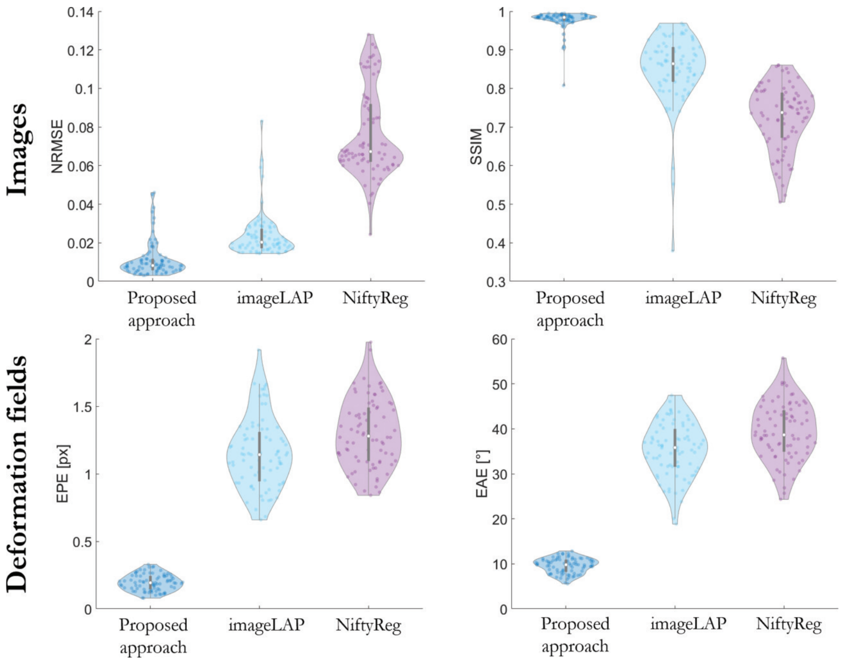

Figure 6: Quantitative analysis over test subjects for reconstructed motion-corrected images and deformations fields of the proposed approach, image-based LAP registration with motion-corrected iterative SENSE reconstruction and NiftyReg registration with motion-corrected iterative SENSE reconstruction. Reconstructed motion-corrected images of subsample data (acquisition time 90 seconds) were compared against reference target images (acquisition time 300 seconds). Deformation fields were compared against reference deformation fields derived via Local All-Pass from the reference target images. Violin plots depict mean (central point), positive and negative standard deviation around mean (vertical central bar) and distribution of values.

were impaired by residual blurring along the diaphragm. Reconstruction to end-expiratory bin was possible with realistic captured deformations which can be appreciated by the 90 seconds scan (180–270 seconds) for which end-expiratory binned image resembles closely the end-inspiratory binned reference image of the complete scan (TA = 360 seconds).

Quantitative analysis of the motion-corrected reconstruction is summarized in Figure 6 and Table 1. Violin plots in Figure 6 indicate an improved quantitative performance of the proposed approach over the image-based solutions with reduced error in the motion-corrected images. Deformation fields which were extracted from the subsampled acquired data show a high similarity to reference deformation for the proposed approach. The proposed approach outperforms both image-based approaches. Any errors in the registration originating from residual aliasing or blurring can propagate into the motion-

Table 1: Quantitative analysis of end-point error (EPE) and end-angulation error (EAE) between estimated deformation field and target deformation field obtained from image-based LAP in reference scan, in prospectively subsampled acquisition ($\sim 14\times$). Structural similarity index (SSIM) and normalized root mean-squared error (NRMSE) are calculated between the motion-compensated reconstructed image and the end-expiratory target image of the reference scan. Inference times for 4D test cases are reported. Metrics are reported as mean $\pm$ one standard deviation for all voxels and test subjects. Best performance is indicated in bold.

	Proposed approach	*imageLAP*	*NiftyReg*
EPE	$\mathbf{0.17 \pm 0.26}$	0.97 ± 1.70	1.34 ± 1.31
EAE	$\mathbf{7.9^\circ \pm 9.9^\circ}$	$35.5^\circ \pm 22.6^\circ$	$40.7^\circ \pm 25.3^\circ$
SSIM	$\mathbf{0.96 \pm 0.04}$	0.88 ± 0.07	0.81 ± 0.03
NRMSE	$\mathbf{0.005 \pm 0.001}$	0.017 ± 0.008	0.023 ± 0.01
Inference registration	$\mathbf{30 \pm 2}$ **s**	$231 \pm 8\,\mathrm{s}$	$301 \pm 10\,\mathrm{s}$
Inference reconstruction	$\mathbf{5 \pm 1}$ **s**	$610 \pm 6\,\mathrm{s}$	$608 \pm 5\,\mathrm{s}$
Inference total	$\mathbf{35 \pm 2}$ **s**	$841 \pm 8\,\mathrm{s}$	$909 \pm 9\,\mathrm{s}$

corrected reconstruction yielding a reduced image quality metric. Consistent and reproducible results were obtained with k-space based registration over the complete cohort. Pre-training of the registration network required $\sim$ 12 hours. Training duration of the proposed method was around $\sim$ 296 hours. Overall, motion-corrected reconstruction of the proposed method took on average $\sim$ 35 seconds (registration $\sim$ 30 seconds, reconstruction $\sim$ 5 seconds), for *imageLAP* $\sim$ 841 seconds (registration $\sim$ 231 seconds, reconstruction $\sim$ 610 seconds) and for *NiftyReg* $\sim$ 909 seconds (registration $\sim$ 301 seconds, reconstruction $\sim$ 608 seconds), yielding a 25-times faster reconstruction with the proposed approach.

7 Discussion

In this work, we proposed the combination of a deep-learning non-rigid k-space registration network with a deep-learning reconstruction network for motion-corrected MR image reconstruction. We investigated the possibility of directly estimating the non-rigid deformation in k-space without the need of a prior image reconstruction. The obtained deformation fields were subsequently incorporated into the motion-corrected reconstruction to enhance spatio-temporal information sharing. An unrolled physics-based reconstruction network was used with a cascade of $(3 + 1)$D convolutional layers and intermittent data consistency blocks.

Deformation fields and motion patterns can be different in the reference scan and in the subsampled acquisition. Test subjects were selected which showed good agreement in motion patterns between reference and subsampled acquisition to perform a quantitative analysis with minimal bias. The reference target motion field is provided by an image-based LAP registration of the target reference scan. Generation of accurate and reliable ground truth motion fields, as well as their evaluation, still remain an open challenge [60]. Therefore, the registration network was pre-trained in a supervised manner and subsequently fine-tuned jointly together with the reconstruction network in a self-supervised way.

As the registration method for *imageLAP* and for retrieving the target deformation field were the same, we can expect close agreement in quantitative metrics. Hence, any deviation can be mainly attributed to residual aliasing and blurring in the initially reconstructed subsampled image. Deviations can thus be a first indicator on how residual aliasing and blurring may influence registration performance when registration is performed in image space. A comparison between *imageLAP* and the registration network (proposed approach) shows the impact of performing a registration on accelerated data in image space and in k-space. Further detailed analysis of the k-space based LAP registration on in-vivo data is reported in [36]. The comparison of the image-based LAP to the k-space based LAP registration on simulated motion flows can be found in [35].

The non-rigid registration network provided accurate deformation fields which were in close agreement to the target deformation and resembled the underlying motion. K-space registration showed high agreement with reference motion. For highly accelerated acquisitions, image-based registration can fail whereas k-space registration still provides satisfactory performance. The quantitative analysis of the motion fields yielded good agreement in non-rigid k-space registration with minimal errors which subsequently contributed to reduced errors in the reconstruction.

Continuous and smooth deformations for consecutive motion states were obtained with k-space registration. Matching forward and backward deformation fields ensured a respiratory cyclic consistent warping. The registration in k-space was less affected by background noise than the image-based version and deformation fields were concentrated to the actual image content. Static regions (e.g., spine) were not deformed and the largest flow occurred in the liver, lung and spleen along superior-inferior direction. The network-based registration was less computationally demanding than the image-based versions. Pre-training the registration network enabled good initialization for the joint self-supervised training.

The reconstruction network utilized an efficient spatio-temporal redundancy sharing with the proposed $(3+1)$D convolutional filters. In contrast to a full 4D convolution operation, less trainable parameters were required. Moreover, the estimated deformation fields were incorporated to guide and share samples in the data consistency as well. A conjugate gradient data consistency was formulated to solve the multi-coil complex-valued processing. The amount of unrolled reconstruction stages and the regularization parameter were chosen empirically to provide a trade-off between performance, trainable parameters and training duration. The proposed approach yields motion-corrected images for each respiratory phase and the deformation fields which enable further analysis of the underlying motion.

We acknowledge several limitations of this study. We performed a comparison against established image-based registration and motion-corrected reconstruction techniques. In the future, other deep learning techniques should be compared [58] as well as the impact of the separable convolution filters [33] shall be investigated. A supervised pre-training is performed for the non-rigid registration network which may be impaired by the image-based registration ground-truth. In addition, a full 3D registration may also be beneficial to capture the local 3D deformation. However, the obtained results in this study did not indicate any performance loss of the pseudo 3D training scheme. The proposed approach was only tested for respiratory motion in T1-weighted imaging of the body trunk. Future studies will investigate its generalizability to different imaging applications and sequences. A Cartesian subsampling was performed which results in incoherent aliasing artifacts along phase-encoding directions. For radial or spiral subsampling, different aliasing artifacts will manifest in the image and may require a retraining.

8 Conclusion

A deep-learning based motion-corrected reconstruction network was proposed which combines a non-rigid k-space registration network with a (3+1)D reconstruction network. Non-rigid registration in k-space is feasible and provides reliable deformation fields, especially for highly accelerated imaging for which image-based registration is impaired. Incorporating the deformation fields into the reconstruction network allows for efficient utilization of spatio-temporal information. The proposed approach was in close agreement with the ground-truth and provided 4D motion-corrected images and deformation fields within ~ 35 seconds.

Financial Support

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2180 – Project number 390900677 and EXC-Number 2064/1 – Project number 390727645. The work was supported by EPSRC grants EP/P032311/1, EP/P007619/1, and EP/P001009/1.

Ethical Standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Biographies

Thomas Küstner (Member, IEEE; Junior Fellow, ISMRM) received his PhD from the University of Stuttgart, Germany, in 2017. From 2018 to 2020 he was with the School of Biomedical Engineering and Imaging Sciences at Kings College London, United Kingdom. Since 2020 he leads the group of medical imaging and data analysis (MIDAS.lab) in the University Hospital of Tübingen, Germany. He is a junior fellow of the society for magnetic resonance in medicine. His research focuses on multi-parametric and multi-modality deep learning methods in acquisition, reconstruction and analysis for patient-centered workflows and with particular focus on MR-based motion imaging, correction and reconstruction.

Jiazhen Pan is currently pursuing his PhD at Lab for Artificial Intelligence in Medicine, Technical University of Munich. Before that, he received his MSc from the University of Stuttgart and from 2019 to 2020, he worked as visiting researcher at the School of Biomedical Engineering and Imaging Sciences at Kings College London. His research interest lies in optical flow estimation, MR Imaging registration and reconstruction.

Christopher Gilliam is a Research Fellow in the School of Engineering at RMIT University, Australia. Previously he was a Postdoctoral Fellow in the Image and Video Processing Group, Department of Electronic Engineering at The Chinese University of Hong Kong. He graduated from Imperial College London, UK, with a First Class Honours MEng in Electrical and Electronic Engineering in 2008 and with a PhD in Electrical and Electronic Engineering in 2013. His research interests include: sampling theory and its application in image processing, in particular Image-Based Rendering (IBR); Sparse sampling involving signals with finite rate of innovation (FRI); Motion estimation and registration for image sequences, in particular biological and medical imaging; and Sensor management. For further details please visit https://sites.google.com/site/cwsgilliam

Haikun Qi received her Ph.D. training in cardiovascular MRI at Tsinghua University from 2013 to 2018. Then she continued her post-doc training at School of Biomedical Engineering and Imaging Sciences at King's College

London. In early 2021, she joined the School of Biomedical Engineering at ShanghaiTech University as an assistant professor. She is currently working on developing next-generation cardiovascular MRI techniques, including 3D high-resolution single-sequence multi-modality cardiac MRI and exploration of deep learning to enhance the performance of data acquisition, motion correction, image reconstruction and disease diagnosis.

Gastao Cruz received his Ph.D. in 2016 from King's College London, where he remained as a Research Associate to date. He is a junior fellow of the society for magnetic resonance in medicine and his research interest lies in accelerated acquisition and reconstruction methods for MRI, motion correction, and quantitative multi-parametric mapping for tissue characterization.

Kerstin Hammernik is a postdoctoral researcher in the Lab for Artificial Intelligence in Healthcare and Medicine, Technical University of Munich, Germany and at the Department of Computing, Imperial College London, United Kingdom. In 2019, she received her Ph.D. degree in Computer Science from Graz University of Technology, Austria. During her Ph.D., she spent four months as a research intern at the Center for Advanced Imaging Innovation and Research, New York University School of Medicine, USA. Her research interests are inverse problems and physics-based deep learning in medical imaging, with a special focus on fast MRI acquisition and reconstruction for cardiac and musculoskeletal applications.

Thierry Blu (Fellow, IEEE) received the "Diplome d'ingénieur" degrees from the École Polytechnique, Palaiseau, France, in 1986, and Télécom Paris (ENST), France, in 1988, and the Ph.D. degree in electrical engineering from ENST, Paris, France, for a study on iterated rational filterbanks, applied to wideband audio coding, in 1996. He was with the Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne, Switzerland, from 1998 and 2007. He is currently a Professor with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong. His research interests include wavelets, approximation and sampling theory, sparse representations, image denoising, biomedical imaging, optics, and wave propagation. Dr. Blu was a member of the IEEE Signal Processing Theory and Methods Technical Committee (2008–2013) and an Associate Editor for the IEEE Transactions on Image Processing (2002–2006), the IEEE Transactions on Signal Processing (2006–2010), and Elsevier's Signal Processing (2008–2011). He is currently on the board of EURASIP Journal on Image and Video Processing, SIAM Journal on Imaging Sciences, and a member of the IEEE Bio Imaging and Signal Processing Technical Committee.

Daniel Rueckert (Fellow, IEEE) gained a M.Sc. from Technical University Berlin in 1993, a Ph.D. from Imperial College in 1997, followed by a post-doc at King's College London. In 1999 he joined Imperial College and became a full Professor in 2005. Since 2020 he is Alexander von Humboldt Professor for AI in Medicine and Healthcare at the Technical University of Munich. He has published more than 500 journal and conference articles with over 50,000 citations and has graduated over 50 Ph.D. students as well as supervised over 40 post-doc. He has been an associate editor of IEEE Transactions on Medical Imaging and is a member of the editorial board of Medical Image Analysis. He has served as a member of organising and programme committees at numerous conferences, including General Co-chair of MMBIA 2006 and FIMH 2013 as well as Programme Co-Chair of MICCAI 2009, ISBI 2012 and WBIR 2012. In 2014, he was elected as a Fellow of the MICCAI society and in 2015 he was elected as a Fellow of the Royal Academy of Engineering and as a fellow of the IEEE. More recently he has been elected as a Fellow of the Academy of Medical Sciences (2019).

René Botnar (Fellow ISMRM and SCMR) received his B.Sc. (1990) and M.Sc. in Physics (1992) from the Karlsruhe Institute of Technology, Germany, and his Ph.D. from the Swiss Federal Institute of Technology, Zurich, Switzerland, in 1997. From 1996 to 1997 he was a Postdoctoral Research Associate in the Department of Radiology at the University of Zurich. In 1997, he joined the Cardiac Magnetic Resonance Center at the Beth Israel Deaconess Medical Center and Harvard Medical School as Clinical Scientist for Philips Medical Systems. In 2003, he became the Scientific Director of the Cardiac MR Center and was appointed as Assistant Professor of Medicine at Harvard Medical School, Boston, USA. In 2005, he became a Full Professor of Biomedical Imaging at the Technische Universität München, Germany and at the end of 2007, he joined the School of Biomedical Engineering and Imaging Sciences at King's College London, UK, as Chair of Cardiovascular Imaging, and since 2014 as Head of the Biomedical Engineering Department. Dr. Botnar was a board member of the Society for Cardiovascular Magnetic Resonance from 2008 to 2011 is currently Associate Editor of the Journal of Cardiac Magnetic Resonance (JCMR) and Senior Editor of the Journal of Molecular Imaging and Biology. He has authored more than 295 peer-reviewed original papers, 45 review articles and 25 book chapters in the field of CMR. He also holds 12 patents and is an editor of a CMR textbook on Cardiovascular Magnetic Resonance Imaging.

Claudia Prieto received a M.Sc. and a Ph.D. from the Pontificia Universidad Católica de Chile in 2005 and 2007, respectively; followed by a post-doc at King's College London. In 2011 she joined King's College London as a Lecturer and became a full Professor in 2020. Her research focuses on acquisition, reconstruction and motion correction methods for multiparametric and quantitative magnetic resonance imaging. She has published more than

100 journal articles in these topics, and has supervised over 30 Ph.D. students and post-docs. She is Deputy Editor of Magnetic Resonance in Medicine and SCMR Board of Trustee. She has served as a member of organising and programme committees at several conferences, including SCMR 2020 and 2022, and ESMRMB 2019 and 2020.

Sergios Gatidis received his M.D. in 2011 at the University of Tübingen and his M.Sc. in Mathematics in 2014 at the University of Hagen. 2017 he was appointed Assistant Professor and 2020 Associate Professor of Radiology at the Department of Radiology of the University Hospital Tübingen. His research focuses on automated analysis of multiparametric medical image data.

References

[1] D. Atkinson, D. L. Hill, P. N. Stoyle, *et al.*, "Automatic Correction of Motion Artifacts in Magnetic Resonance Images Using An Entropy Focus Criterion," *IEEE Transactions on Medical Imaging*, 16(6), 1997, 903–10.

[2] P. Batchelor, D. Atkinson, P. Irarrazaval, *et al.*, "Matrix Description of General Motion Correction Applied to Multishot Images," *Magnetic Resonance in Medicine*, 54(5), 2005, 1273–80.

[3] K. T. Block, M. Uecker, and J. Frahm, "Undersampled Radial MRI with Multiple Coils. Iterative Image Reconstruction Using A Total Variation Constraint," *Magnetic Resonance in Medicine*, 57(6), 2007, 1086–98.

[4] T. Blu, P. Moulin, and C. Gilliam, "Approximationorder of the Lap Optical Flow Algorithm," in *IEEE International Conference on Image Processing (ICIP)*, 2015, 48–52.

[5] J. Caballero, A. N. Price, D. Rueckert, *et al.*, "Dictionary Learning and Time Sparsity for Dynamic MR Data Reconstruction," *IEEE Transactions on Medical Imaging*, 33(4), 2014, 979–94.

[6] F. Chen, J. Y. Cheng, V. Taviani, *et al.*, "Data-Driven Self-Calibration and Reconstruction for Non-Cartesian Wave-Encoded Single-Shot Fast Spin Echo Using Deep Learning," *Journal of Magnetic Resonance Imaging*, 51(3), 2020, 841–53.

[7] F. Chen, V. Taviani, I. Malkiel, *et al.*, "Variable-Density Single-Shot Fast Spin-Echo MRI with Deep Learning Reconstruction by Using Variational Networks," *Radiology*, 289(2), 2018, 366–73.

[8] X. Chen, M. Salerno, Y. Yang, *et al.*, "Motion-Compensated Compressed Sensing for Dynamic Contrast-Enhanced MRI Using Regional Spatiotemporal Sparsity and Region Tracking: Block Low-Rank Sparsity with Motion-Guidance (BLOSM)," *Magnetic Resonance in Medicine*, 72(4), 2014, 1028–38.

[9] J. Y. Cheng, M. Mardani, M. T. Alley, *et al.*, "DeepSPIRiT: Generalized Parallel Imaging using Deep Convolutional Neural Networks," in *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*, 2018, 570.

[10] J. Y. Cheng, T. Zhang, N. Ruangwattanapaisarn, *et al.*, "Free-Breathing Pediatric MRI with Nonrigid Motion Correction and Acceleration," *Journal of Magnetic Resonance Imaging*, 42(2), 2015, 407–20.

[11] L. Cordero-Grande, R. P. A. Teixeira, E. J. Hughes, *et al.*, "Sensitivity Encoding for aligned Multishot Magnetic Resonance Reconstruction," *IEEE Transactions on Computational Imaging*, 2(3), 2016, 266–80.

[12] T. Eo, Y. Jun, T. Kim, *et al.*, "KIKI-net: Cross-Domain Convolutional Neural Networks for Reconstructing Undersampled Magnetic Resonance Images," *Magnetic Resonance in Medicine*, 80(5), 2018, 2188–201.

[13] N. Fuin, A. Bustin, T. Küstner, *et al.*, "A Variational Neural Network for Accelerating Free-breathing Whole-Heart Coronary MR Angiography," in *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*, 2019, 478.

[14] C. Gilliam and T. Blu, "Local All-Pass Filters for Optical Flow Estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, 1533–7.

[15] C. Gilliam and T. Blu, "Local All-Pass Geometric Deformations," *IEEE Transactions on Image Processing*, 27(2), 2018, 1010–25.

[16] C. Gilliam, T. Küstner, and T. Blu, "3D Motion Flow Estimation Using Local All-Pass Filters," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2016, 282–5.

[17] M. A. Griswold, P. M. Jakob, R. M. Heidemann, *et al.*, "Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)," *Magnetic Resonance in Medicine*, 47(6), 2002, 1202–10.

[18] K. Hammernik, T. Klatzer, E. Kobler, *et al.*, "Learning a Variational Network for Reconstruction of Accelerated MRI Data," *Magnetic Resonance in Medicine*, 79(6), 2018, 3055–71.

[19] F. Han, Z. Zhou, E. Han, *et al.*, "Self-Gated 4D Multiphase, Steady-State Imaging with Contrast Enhancement (MUSIC) Using Rotating Cartesian K-space (ROCK): Validation in Children with Congenital Heart Disease," *Magnetic Resonance in Medicine*, 78(2), 2017, 472–83.

[20] Y. Han, L. Sunwoo, and J. C. Ye, "k-space Deep Learning for Accelerated MRI," *IEEE Transactions on Medical Imaging*, 39(2), 2019, 377–86.

[21] M. W. Haskell, S. F. Cauley, B. Bilgic, *et al.*, "Network Accelerated Motion Estimation and Reduction (NAMER): Convolutional Neural Network Guided Retrospective Motion Correction Using A Separable Motion Model," *Magnetic Resonance in Medicine*, 82(4), 2019, 1452–61.

[22] M. Henningsson, P. Koken, C. Stehning, *et al.*, "Whole-Heart Coronary MR Angiography with 2D Self-Navigated Image Reconstruction," *Magnetic Resonance in Medicine*, 67(2), 2012, 437–45.
[23] B. K. Horn and B. G. Schunck, "Determining Optical Flow," *Artificial Intelligence*, 17(1–3), 1981, 185–203.
[24] S. A. H. Hosseini, C. Zhang, S. Weingärtner, *et al.*, "Accelerated Coronary MRI with sRAKI: A Database-Free Self-Consistent Neural Network k-space Reconstruction for Arbitrary Undersampling," *PLoS ONE*, 15(2), 2020.
[25] Q. Huang, D. Yang, H. Qu, *et al.*, "Dynamic MRI Reconstruction with Motion-Guided Network," in *Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning, Proceedings of Machine Learning Research*, 2019, 275–84.
[26] C. M. Hyun, H. P. Kim, S. M. Lee, *et al.*, "Deep Learning for Undersampled MRI Reconstruction," *Physics in Medicine and Biology*, 63(13), 2018, 135007.
[27] H. Jung, J. C. Ye, and E. Y. Kim, "Improved k-t BLAST and k-t SENSE Using FOCUSS," *Physics in Medicine and Biology*, 52(11), 2007, 3201–26.
[28] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv preprint arXiv:1412.6980*, 2014.
[29] F. Knoll, K. Hammernik, C. Zhang, *et al.*, "Deep Learning Methods for Parallel Magnetic Resonance Image Reconstruction," *IEEE Signal Process Mag*, 37(1), 2020, 128–40.
[30] T. Kober, J. P. Marques, R. Gruetter, *et al.*, "Head Motion Detection Using FID Navigators," *Magnetic Resonance in Medicine*, 66(1), 2011, 135–43.
[31] T. Küstner, A. Bustin, O. Jaubert, *et al.*, "Fully Self-Gated Free-Running 3D Cartesian Cardiac CINE with Isotropic Whole-Heart Coverage in less than 2 min," *NMR in Biomedicine*, 2020, e4409.
[32] T. Küstner, A. Bustin, O. Jaubert, *et al.*, "Isotropic 3D Cartesian Single Breath-Hold CINE MRI with Multi-Bin Patch-Based Low-Rank Reconstruction," *Magnetic Resonance in Medicine*, 2020.
[33] T. Küstner, N. Fuin, K. Hammernik, *et al.*, "CINENet: Deep Learning-Based 3D Cardiac CINE MRI Reconstruction with Multi-Coil Complex-Valued 4D Spatio-Temporal Convolutions," *Scientific Reports*, 10(1), 2020, 13710.
[34] T. Küstner, J. Pan, C. Gilliam, *et al.*, "Deep-Learning based Motion-Corrected Image Reconstruction in 4D Magnetic Resonance Imaging of the Body Trunk," in *Proceedings of the IEEE Asia-Pacific Signal and Information Processing Association (APSIPA)*, 2020.

[35] T. Küstner, J. Pan, H. Qi, *et al.*, "LAPNet: Deep-Learning based Non-Rigid Motion Estimation in k-space from Highly Undersampled Respiratory and Cardiac Resolved Acquisitions," in *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*, 2021.

[36] T. Küstner, J. Pan, H. Qi, *et al.*, "LAPNet: Non-Rigid Registration derived in k-space for Magnetic Resonance Imaging," in *IEEE Transactions on Medical Imaging*, 2021, 1–1.

[37] T. Küstner, M. Schwartz, S. Gatidis, *et al.*, "Local All-Pass Filter (LAP): Efficient Optical Flow-Based Image Registration," in *Proceedings of the ISMRM Workshop on Motion Correction in MRI and MRS*, 2017.

[38] T. Küstner, M. Schwartz, P. Martirosian, *et al.*, "MR-Based Respiratory and Cardiac Motion Correction for PET Imaging," *Medical Image Analysis*, 42, 2017, 129–44.

[39] T. Küstner, C. Würslin, S. Gatidis, *et al.*, "MR Image Reconstruction Using a Combination of Compressed Sensing and Partial Fourier Acquisition: ESPReSSo," *IEEE Transactions on Medical Imaging*, 35(11), 2016, 2447–58.

[40] T. Küstner, C. Würslin, M. Schwartz, *et al.*, "Self-Navigated 4D Cartesian Imaging of Periodic Motion in the Body Trunk Using Partial k-space Compressed Sensing," *Magnetic Resonance in Medicine*, 78(2), 2017, 632–44.

[41] D. Lee, J. Yoo, S. Tak, *et al.*, "Deep Residual Learning for Accelerated MRI Using Magnitude and Phase Networks," *IEEE Transactions on Biomedical Engineering*, 65(9), 2018, 1985–95.

[42] D. J. Lin, P. M. Johnson, F. Knoll, *et al.*, "Artificial Intelligence for MR Image Reconstruction: An Overview for Clinicians," *Journal of Magnetic Resonance Imaging*, 2020.

[43] S. G. Lingala, Y. Hu, E. DiBella, *et al.*, "Accelerated Dynamic MRI Exploiting Sparsity and Low-Rank Structure: kt SLR," *IEEE Transactions on Medical Imaging*, 30(5), 2011, 1042–54.

[44] F. Liu, D. Li, X. Jin, *et al.*, "Dynamic Cardiac MRI Reconstruction Using Motion Aligned Locally Low Rank Tensor (MALLRT)," *Magnetic Resonance Imaging,* [Epub ahead of print], 2019.

[45] M. Lustig, D. Donoho, J. M. Pauly, *et al.*, "Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging," *Magnetic Resonance in Medicine*, 58(6), 2007, 1182–95.

[46] J. Maclaren, M. Herbst, O. Speck, *et al.*, "Prospective Motion Correction in Brain Imaging: A Review," *Magnetic Resonance in Medicine*, 69(3), 2013, 621–36.

[47] M. O. Malavé, C. A. Baron, S. P. Koundinyan, *et al.*, "Reconstruction of Undersampled 3D Non-Cartesian Image-Based Navigators For Coronary MRA Using An Unrolled Deep Learning Model," *Magnetic Resonance in Medicine*, 82(2), 2020, 800–12.

[48] D. Manke, K. Nehrke, P. Börnert, *et al.*, "Respiratory Motion in Coronary Magnetic Resonance Angiography: A Comparison of Different Motion Models," *Journal of Magnetic Resonance Imaging*, 15(6), 2002, 661–71.

[49] M. Mardani, E. Gong, J. Y. Cheng, *et al.*, "Deep Generative Adversarial Neural Networks for Compressive Sensing MRI," *IEEE Transactions on Medical Imaging*, 38(1), 2018, 167–79.

[50] J. R. McClelland, J. M. Blackall, S. Tarte, *et al.*, "A Continuous 4D Motion Model from Multiple Respiratory Cycles for Use in Lung Radiotherapy," *Medical Physics*, 33(9), 2006, 3348–58.

[51] M. Modat, G. R. Ridgway, Z. A. Taylor, *et al.*, "Fast Free-Form Deformation Using Graphics Processing Units," *Computer Methods and Programs in Biomedicine*, 98(3), 2010, 278–84.

[52] Y. Q. Mohsin, S. G. Lingala, E. DiBella, *et al.*, "Accelerated Dynamic MRI Using Patch Regularization for Implicit Motion Compensation," *Magnetic Resonance in Medicine*, 77(3), 2017, 1238–48.

[53] C. Munoz, R. Neji, G. Cruz, *et al.*, "Motion-Corrected Simultaneous Cardiac Positron Emission Tomography and Coronary MR Angiography with High Acquisition Efficiency," *Magnetic Resonance in Medicine*, 79(1), 2018, 339–50.

[54] F. Odille, P. A. Vuissoz, P. Y. Marie, *et al.*, "Generalized Reconstruction by Inversion of Coupled Systems (GRICS) Applied to Free-Breathing MRI," *Magnetic Resonance in Medicine*, 60(1), 2008, 146–57.

[55] F. Ong and M. Lustig, "Joint Non-Rigid Motion Estimation and Image Reconstruction via Sparse Blind Deconvolution," in *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*, 2017, 3937.

[56] R. Otazo, E. Candes, and D. K. Sodickson, "Low-Rank Plus Sparse Matrix Decomposition for Accelerated Dynamic MRI with Separation of Background and Dynamic Components," *Magnetic Resonance in Medicine*, 73(3), 2015, 1125–36.

[57] K. P. Pruessmann, M. Weiger, P. Boernert, *et al.*, "Advances in Sensitivity Encoding with Arbitrary k-space Trajectories," *Magnetic Resonance in Medicine*, 46(4), 2001, 638–51.

[58] H. Qi, R. Hajhosseiny, G. Cruz, *et al.*, "End-to-end Deep Learning Nonrigid Motion-Corrected Reconstruction For Highly Accelerated Free-Breathing Coronary MRA," *Magnetic Resonance in Medicine*, 86(4), 2021, 1983–96.

[59] C. Qin, J. Schlemper, J. Caballero, *et al.*, "Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction," *IEEE Transactions on Medical Imaging*, 38(1), 2019, 280–90.

[60] T. Rohlfing, C. Maurer, W. O'Dell, *et al.*, "Modeling Liver Motion and Deformation During the Respiratory Cycle Using Intensity-Based Free-Form Registration of Gated MR Images," *Medical Imaging*, 4319, 2001.

[61] D. Rueckert, L. I. Sonoda, C. Hayes, *et al.*, "Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images," *IEEE Transactions on Medical Imaging*, 18(8), 1999, 712–21.

[62] C. M. Sandino, P. Lai, S. S. Vasanawala, *et al.*, "Accelerating Cardiac Cine MRI Using a Deep Learning-Based ESPIRiT Reconstruction," *Magnetic Resonance in Medicine*, 85(1), 2021, 152–67.

[63] C. M. Sandino, P. Lai, S. S. Vasanawala, *et al.*, "DL-ESPIRiT: Improving Robustness to SENSE Model Errors in Deep Learning-Based Reconstruction," in *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*, 2019, 659.

[64] J. Schlemper, J. Caballero, J. V. Hajnal, *et al.*, "A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction," *IEEE Transactions on Medical Imaging*, 37(2), 2018, 491–503.

[65] G. Seegoolam, J. Schlemper, C. Qin, *et al.*, "Exploiting Motion for Deep Learning Reconstruction of Extremely-Undersampled Dynamic MRI," in *MICCAI*, 2019.

[66] S. Skare, A. Hartwig, M. Martensson, *et al.*, "Properties of a 2D Fat Navigator for Prospective Image Domain Correction of Nodding Motion in Brain MRI," *Magnetic Resonance in Medicine*, 73(3), 2015, 1110–9.

[67] O. Speck, J. Hennig, and M. Zaitsev, "Prospective Real-Time Slice-by-Slice Motion Correction for fMRI in Freely Moving Subjects," *Magnetic Resonance Materials in Physics, Biology and Medicine*, 19(2), 2006, 55.

[68] J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for Compressive Sensing MRI," in *Proceedings of the Advances in Neural Information Processing Systems*, 2016, 10–8.

[69] J.-P. Thirion, "Image Matching as a Diffusion Process: An Analogy with Maxwell's Demons," *Medical Image Analysis*, 2(3), 1998, 243–60.

[70] M. Uecker, P. Lai, M. J. Murphy, *et al.*, "ESPIRiT – An Eigenvalue Approach to Autocalibrating Parallel MRI: Where SENSE Meets GRAPPA," *Magnetic Resonance in Medicine*, 71(3), 2014, 990–1001.

[71] M. Usman, S. Latif, M. Asim, *et al.*, "Retrospective Motion Correction in Multishot MRI using Generative Adversarial Network," *Scientific Reports*, 10(1), 2020, 4786.

[72] M. Usman, B. Ruijsink, M. S. Nazir, *et al.*, "Free Breathing Whole-Heart 3D CINE MRI with Self-Gated Cartesian Trajectory," *Magnetic Resonance Imaging*, 38, 2017, 129–37.

[73] S. Wang, Z. Su, L. Ying, *et al.*, "Accelerating Magnetic Resonance Imaging via Deep Learning," in *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*, 2016, 514–7.

[74] G. Yang, S. Yu, H. Dong, *et al.*, "DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction," *IEEE Transactions on Medical Imaging*, 37(6), 2018, 1310–21.
[75] M. Zaitsev, J. Maclaren, and M. Herbst, "Motion Artifacts in MRI: A Complex Problem with Many Partial Solutions," *Journal of Magnetic Resonance Imaging*, 42(4), 2015, 887–901.
[76] W. Zhou, A. C. Bovik, H. R. Sheikh, *et al.*, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, 13(4), 2004, 600–12.
[77] B. Zhu, J. Z. Liu, S. F. Cauley, *et al.*, "Image Reconstruction by Domain-Transform Manifold Learning," *Nature*, 555(7697), 2018, 487–92.