Robust motion correction for cardiac T1 and ECV mapping using a T1 relaxation model approach

Sofie Tilborghs[a,d,*], Tom Dresselaers[b,d], Piet Claus[c,d], Guido Claessen[c], Jan Bogaert[b,d], Frederik Maes[a,d], Paul Suetens[a,d]

[a]*Department of Electrical Engineering, ESAT/PSI, KU Leuven, Leuven, Belgium*
[b]*Department of Imaging and Pathology, Radiology, KU Leuven, Leuven, Belgium*
[c]*Department of Cardiovascular Sciences, KU Leuven, Leuven, Belgium*
[d]*Medical Imaging Research Center, UZ Leuven, Herestraat 49 - 7003, 3000 Leuven, Belgium*

Abstract

T1 and ECV mapping are quantitative methods for myocardial tissue characterization using cardiac MRI, and are highly relevant for the diagnosis of diffuse myocardial diseases. Since the maps are calculated pixel-by-pixel from a set of MRI images with different T1-weighting, it is critical to assure exact spatial correspondence between these images. However, in practice, different sources of motion e.g. cardiac motion, respiratory motion or patient motion, hamper accurate T1 and ECV calculation such that retrospective motion correction is required. We propose a new robust non-rigid registration framework combining a data-driven initialization with a model-based registration approach, which uses a model for T1 relaxation to avoid direct registration of images with highly varying contrast. The registration between native T1 and enhanced T1 to obtain a motion free ECV map is also calculated using information from T1 model-fitting. The method was validated on three datasets recorded with two substantially different acquisition protocols (MOLLI (dataset 1 (n=15) and dataset 2 (n=29)) and STONE (dataset 3 (n = 210))), one in breath-hold condition and one free-breathing. The average Dice coefficient increased from $72.6\pm12.1\%$ to $82.3\pm7.4\%$ (P<0.05) and mean boundary error decreased from $2.91\pm1.51mm$ to $1.62\pm0.80mm$ (P<0.05) for motion

---

*Corresponding author
Email address:* `sofie.tilborghs@kuleuven.be` (Sofie Tilborghs)

correction in a single T1-weighted image sequence (3 datasets) while average Dice coefficient increased from 63.4±22.5% to 79.2±8.5% (P<0.05) and mean boundary error decreased from 3.26±2.64mm to 1.77±0.86mm (P<0.05) between native and enhanced sequences (dataset 1 and 2). Overall, the native T1 SD error decreased from 67.32±32.57ms to 58.11±21.59ms (P<0.05), enhanced SD error from 30.15±25ms to 22.74±8.94ms (P<0.05) and ECV SD error from 10.08±9.59% to 5.42±3.21% (P<0.05) (dataset 1 and 2).

Keywords: cardiac T1 mapping, cardiac ECV mapping, motion correction, image registration

---

1. Introduction

Accurate measurement of myocardial T1 and extra cellular volume (ECV) using cardiac MRI is highly relevant for the diagnosis of diffuse myocardial diseases such as diffuse fibrosis, amyloidosis and Anderson Fabry disease (Schelbert and Messroghli (2016); Haaf et al. (2017)). Compared to the conventional late gadolinium enhanced (LGE) images where diagnosis is based on the subjective assessment of relative contrast differences, T1 and ECV mapping allow quantitative characterization of the myocardium. In T1 mapping, a map of the longitudinal relaxation time (T1) is obtained by fitting an exponential curve to each pixel in a sequence of T1-weighted images, acquired over multiple heart cycles, using a dedicated imaging protocol e.g. the Modified Look-Locker Inversion recovery protocol (MOLLI) (Messroghli et al. (2004)) or the free-breathing multislice slice-interleaved T1 sequence (STONE) (Weingärtner et al. (2015)). In MOLLI, 2D T1-weighted images are typically acquired using two or three inversion pulses with multiple read-outs after each inversion pulse, while in STONE, a single image is acquired after each inversion, in theory eliminating disturbance of the relaxation curve by multiple read-outs. Additionally, an ECV map (Kellman et al. (2013)) can be constructed by combining the T1 map before (native) and after (enhanced) gadolinium contrast injection.

Motion artifacts in the pixel-wise T1 map, and consequently in the pixel-

2

wise ECV map, appear when the sequence of T1-weighted images is misaligned due to involuntary patient motion, cardiac motion and/or respiratory motion. Despite ECG-gating and acquisition in mid- or end-diastole, artifacts can result from heart-rate related triggering errors or drifts, which appear as stochas-
25 tic non-rigid motion over the images. Respiration related motion is generally more continuous in time and is often limited by imposing breath hold during acquisition, although also a trend towards free-breathing acquisition, using a respiratory navigator, is notable. However, even in breath-held acquisitions, respiratory-related motion artifacts cannot be eliminated completely due to the
30 subject's inability or noncompliance to repeated breath-hold instructions in a usually long cardiac MRI exam, diaphragmatic drift or septal shifts resulting from pressure differences between both ventricles during a breath-hold. Retrospective non-rigid motion correction is thus imperative for correct quantification of T1 in the whole myocardium.

35 The use of off-the-shelf state-of-the-art generic data-driven image registration approaches (Viergever et al. (2016)) for this application is complicated by the intrinsic complexity of the image data, including contrast inversion, partial volume effects and signal nulling for images acquired near the zero crossing of the T1 relaxation curve (Xue et al. (2012)), such that dedicated approaches are
40 needed. Furthermore, obtaining a motion-free set of T1-weighted images requires aligning more than two images implying that also a registration strategy (i.e. pairwise to reference or groupwise) should be specified.

In Roujol et al. (2015), a local non-rigid registration framework based on optical flow was developed that simultaneously estimated the motion field and
45 the intensity variation to cope with the large variations in contrast. All images are transformed to a reference image, which is chosen to be the image with the largest or second largest inversion time (TI). In Huizinga et al. (2016), a groupwise method for quantitative MRI, including T1 mapping, was proposed whereby all images are registered simultaneously to a mean space by
50 minimization of a cost function based on principal component analysis (PCA), assuming a non-specific low dimensional signal model. This method was vali-

3

dated on human cardiac data in Tao et al. (2018). Guyader et al. (2016) uses a theoretically-derived total-correlation-based metric instead of the empirically-determined PCA-based cost in the same framework. Recently, Zhang et al. (2018) proposed to perform motion correction in T1 mapping by maximizing the similarity of normalized gradient fields between the reference image (the image with longest TI) and any other image, combined with an elastic regularizer. The same algorithm is used to perform registration between the native and enhanced T1 map.

Model-based approaches on the other hand exploit the underlying T1 relaxation model, such that direct registration between images with largely different or inverted contrast can be avoided. In Xue et al. (2012), motion free-synthetic images resembling the original contrasts are constructed based on a crude initial T1 estimate to guide the registration. This initial T1 map is estimated from two images with respectively the shortest and longest TI, which have been registered to avoid motion artifacts. Alignment between the synthetic and original images is achieved with a fast variational non-rigid image registration framework. In Van De Giessen et al. (2013), the error on the exponential curve fitting, which is assumed to increase in case of misalignment, is directly used as registration criterion to spatially align the images obtained from a Look-Locker sequence. In this acquisition sequence, no ECG-gating is performed and more T1-weighted images (typically 30) are used. In practice, this is implemented as an iterative approach, which alternately fits the exponential curve pixel-by-pixel through the datapoints and minimizes the error between the model intensities and the real intensities. Both steps use the sum-of-squared differences as optimization metric. A limitation of these model-based algorithms is however that a good initialization for T1 is required, which in practice involves a sufficiently accurate initialization of the registration. Furthermore, relying on the model is complicated by the need for signal polarity restoration when magnitude-reconstructed images are used.

A third category of registration methods, besides data-driven and model-based approaches, is a segmentation-based registration approach. In such ap-

4

proach, correspondence of segmentations of myocardium or other structures among the different images is used as similarity criterion for motion correction. An example for native T1 mapping was recently presented in El-Rewaidy et al. (2018) where a segmentation of the myocardium in every T1-weighted image is obtained using an active shape model (ASM). Registration is subsequently achieved by estimating rigid- and non-rigid transformation parameters to align all contours to the reference contour, which is chosen to be the contour of the image with maximal contrast (shortest TI). A registration based on a global segmentation of the myocardium alone has as main disadvantage that correspondence of images is only enforced at endo- and epicardium, which is an important limitation in case of focal diseases (e.g. focal fibrosis or myocardial infarction).

Related to the problem of motion correction for T1 mapping is motion correction in myocardial perfusion imaging (Gupta et al. (2012)), for which a challenge has been organized at STACOM 2014 (Pontre et al. (2016)). In perfusion imaging, also a set of T1-weighted images with different contrast is available, although there are typically more images ($\sim$40-60), thus requiring a longer breathhold (45-50s) (Gupta et al. (2012)). Furthermore, whereas for T1 mapping the intensity of the left ventricle (LV) and right ventricle (RV) blood pool can be assumed to be identical in one image, this is not the case in perfusion imaging. For this problem, also model-based approaches have been presented e.g. by Likhite et al. (2015). In their method, they construct model images based on a diffusion model, which requires the arterial input function from the RV blood pool. The model is thus not only dependent on the intensity values of a single pixel as is the case for T1 mapping and requires a prior segmentation.

In this paper, we present a model-based registration method for T1 mapping, which iteratively minimizes the errors on the T1 curve fit by registering each image to its corresponding model derived from estimated parameters during T1 curve fitting. Compared to existing model-based algorithms, we propose a robust data-driven initial registration to avoid large bias in the initial T1 estimate in cases with large motion. Robustness is assured by automatically

determining an optimal registration sequence for each image and by exploiting specific knowledge about the acquisition to identify the low-signal images in the sequence that are difficult to register, requiring specific attention. Additionally, we choose to use a two parameter exponential model instead of a three parameter model during registration to further increase robustness. Finally, we also integrate co-registration of the native and enhanced images in the framework to obtain motion free ECV maps. We validated our algorithm on different in-vivo datasets using both segmentation-based error metrics and metrics directly representing the reliability of T1 and ECV values. An initial version of this work was presented in Tilborghs et al. (2017a) and Tilborghs et al. (2017b). New contributions in this paper are:

- Two parameter instead of three parameter model for model-based registration to increase robustness and convergence

- New method for co-registration, which uses the parameters of the two parameter exponential model and thereby optimally integrates the information contained in both scans

- Validation on a publicly available dataset acquired using a free-breathing multi-slice protocol (El-Rewaidy et al. (2018))

2. Methods

2.1. Registration framework

The proposed algorithm combines a data-driven initialization with a model-based refinement. The data-driven initialization is required to obtain a robust first estimate for the pixelwise T1 before exploiting the T1-relaxation model. This combined method is applied to both native and enhanced scans separately. A motion free ECV map is additionally obtained by co-registration. An overview of the method is given in Fig. 1.
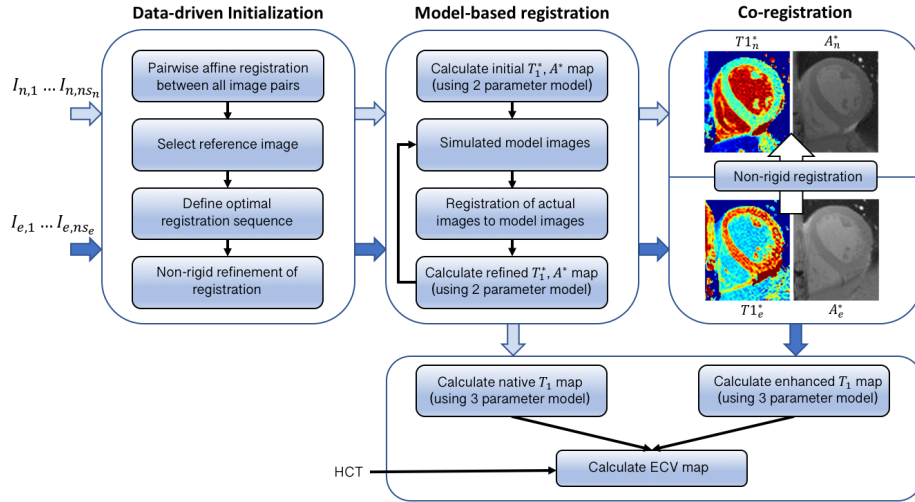
Figure 1: Overview of registration framework. During the data-driven initialization and model-based registration, native (n, light connection arrows) and enhanced (e, dark connection arrows) images are handled separately. ns_n and ns_e are the number of native and enhanced images respectively. The co-registration warps the enhanced images in the native image space. After registration, motion corrected native T1 map, enhanced T1 map and ECV map are calculated.

2.1.1. Motion correction for T1 mapping

Data-driven initialization. The data-driven initialization considers all pairwise registrations between all images I_1 to I_{ns} in the scan simultaneously and applies a global optimization approach to find an optimal reference image and optimal registration sequences to minimize registration failures. The optimization is based on a symmetric measure $C(a, b)$ for the registration affinity between any two images I_a and I_b, which takes into account 1) the consistency $C_c(a, b)$ between their forward $T^{a \rightarrow b}$ and backward $T^{b \rightarrow a}$ affine registration transformation and 2) the similarity C_s of each image with all other images in the scan after affine registration, which acts as a penalty term to discourage registration of dissimilar images:

$$C(a, b) = (C_c(a, b) + C_c(b, a)) + w \cdot (C_s(a) + C_s(b)), \qquad (1)$$

7

where

$$C_c(a,b) = \left\| (T^{b \to a} \cdot T^{a \to b} - \mathbf{I}) \cdot c_a \right\| \tag{2}$$

with c_a a set of samples in image I_a on which the registrations are evaluated and

$$C_s(a) = \frac{1}{\sum_i MI(T(a), i) + MI(a, T(i))} \tag{3}$$

the inverse of the total mutual information (MI) of image I_a with any other image I_i. The two terms are balanced by the weight w to have a more or less equal contribution to the cost $C(a,b)$. The Floyd-Warshall algorithm (Floyd (1962)) is subsequently applied to the cost matrix C to determine the sequence of pairwise image registrations for which the total accumulated cost $C_{opt}(a,b)$ to transform any image I_a to any image I_b is minimal. For example, the cost $C(a,b)$ of transforming $I_a \to I_b$ might be larger than the sum of the cost to transform $I_a \to I_c$ and $I_c \to I_b$, resulting in $C_{opt}(a,b) = C(a,c) + C(c,b)$ in this case. Finally, the cost for choosing every image I_a as reference image is calculated as $C_{tot}(a) = \sum_i C_{opt}(i,a)$. The image I_r for which $C_{tot}(r)$ is minimal, is selected as reference image.

The required pairwise registrations are subsequently refined by non-rigid registration. Each image is then warped to the reference image by concatenating the transformations according to its optimal registration sequence.

Because images with low signal were found to be difficult to register with any other image, they are treated differently to increase registration robustness. These images are automatically detected in the scan based on their average signal intensity over the region of interest (ROI) (Sec. 2.3) and are initially only aligned affinely with the reference image, based on the affine registrations of their adjacent images in acquisition order, assuming the affine motion component to be continuous.

Model-based registration. The intensities $s(t)$ in corresponding pixels of the images of a T1-weighted scan, acquired at timepoints t after the inversion pulse,

8

follow an exponential model of T1 relaxation. For the MOLLI acquisition protocol (Messroghli et al. (2004)), a three parameter model is commonly used:

$$s(t) = A - B \cdot e^{-\frac{t}{T_{app}}}, \tag{4}$$

where A and B are dimensionless parameters, T_{app} is the apparent T1 and the true T1 is $T1 = T_{app} \cdot (\frac{B}{A} - 1)$. From the estimated parameters A, B and T_{app} in Eq. 4, ideal model images I_t^m can be derived for every timepoint t in the sequence with similar contrast as the original images I_t, but with reduced motion artifacts (Fig. 2b). Hence, it is expected that in case of misalignment, the difference between I_t^m and I_t will increase (Mewton et al. (2011)). Following this hypothesis, refinement of the registration can be achieved by transforming I_t to minimize a similarity cost between I_t^m and I_t and iterating the estimation of model parameters and the transformation step (Tilborghs et al. (2017b)). However, since in practice relatively few images are obtained in the initial, most steeply ascending part of the exponential T1 relaxation curve, the intensities in the images with low inversion time t will dominate the estimation of A, B and T_{app}. In case of misalignment of one or few of such images, the appearance of corresponding model images will be dominated by the actual image itself, such that no or only slow convergence of motion correction can be achieved.

To improve convergence, we propose to use a more robust model (Fig. 2) to calculate our model images I_t^m, which contains only two parameters (Weingärtner et al. (2015)):

$$s(t) = A^* \cdot (1 - 2 \cdot e^{-\frac{t}{T1^*}}), \tag{5}$$

where A^* is a dimensionless parameter and $T1^*$ is an estimation for T1. The optimal parameters are estimated using the Nelder-Mead simplex direct search algorithm (Lagarias et al. (1998)). Because magnitude-reconstructed images are used in our experiments, the exponential curve cannot be fitted directly through the image intensities as the polarity of the signals is unknown. Hence, we use a multifitting approach (Nekolla et al. (1992)) where for each pixel, five curve
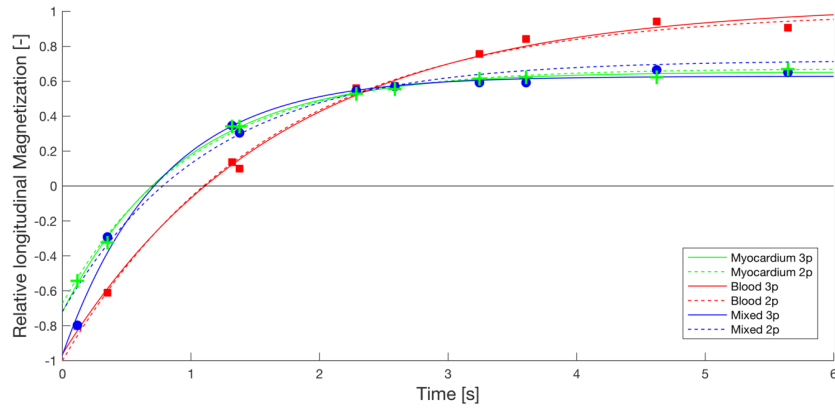
(a)
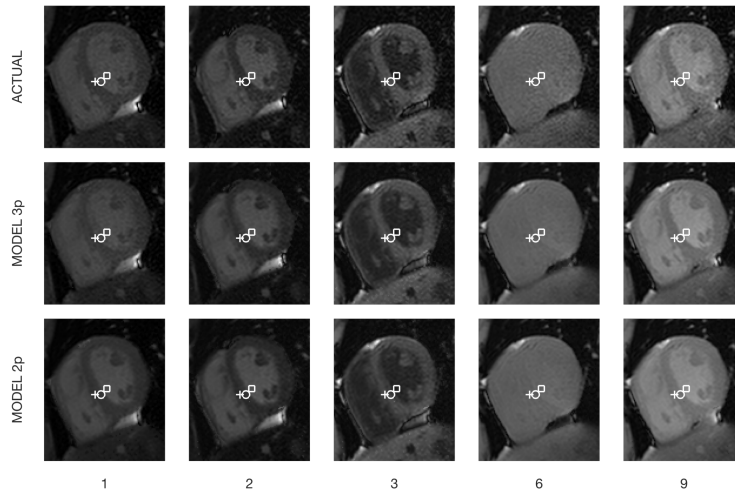
(b)

Figure 2: (a) Curve fitting using the three (Eq. 4) and two parameter (Eq. 5) model for three different pixels: myocardial pixel (green), LV blood pixel (red) and mixed pixel where the first datapoint belongs to the LV blood pool and all other are part of the myocardium due to misalignment of the images (blue). The two parameter curve shows an improved correction of the first datapoint of the mixed pixel. (b) Five out of ten T1-weighted images acquired using the MOLLI protocol: actual images, three and two parameter model images with the myocardium (cross), blood (square) and mixed (circle) pixels indicated.

205 fittings are performed (none, one, two, three or four first points inverted). The curve fitting with the lowest squared error is considered to be the correct one.

10

We iteratively repeat the calculation of model images using the two parameter model and the non-rigid registration of the actual images to the estimated model images. In every step, a new $T1^*$ map and new model images with reduced motion artifacts are obtained.

2.1.2. Motion correction for ECV mapping

To obtain a motion free ECV map, alignment of the native and enhanced T1 map is required. We propose to perform a joint registration of $T1^*$ maps and images of the A^* parameter of Eq. 5 by minimizing the sum of mutual information between both image pairs:

$$C_{co}(T_{co}, T1_n^*, A_n^*, T1_e^*, A_e^*) = MI(T1_n^*, T_{co}(T1_e^*)) + MI(A_n^*, T_{co}(A_e^*)) \quad (6)$$

where T_{co} is the transformation applied to the enhanced images and the subscripts n and e refer to the native and enhanced scan respectively. By exploiting the complementary information in both images, a more robust motion correction can be achieved. Furthermore, since both images are readily available from the model-based registration for motion free T1 maps, no additional T1 calculation step has to be performed.

2.2. Construction of T1 and ECV maps

After motion correction, motion free T1 maps can be constructed using the most appropriate relaxation model for the adopted acquisition protocol, e.g. the three parameter model (Eq. 4) for MOLLI acquisitions or the two parameter model (Eq. 5) for STONE acquisitions. The chosen model for the final T1 map is independent of the model used for registration.

Motion free ECV maps can additionally be obtained by combining native and enhanced T1 maps:

$$ECV = (1 - HCT) \cdot \left(\frac{\frac{1}{T1_e} - \frac{1}{T1_n}}{\frac{1}{T1_{blood,e}} - \frac{1}{T1_{blood,n}}} \right) \quad (7)$$

where HCT is the hematocrit of the patient, i.e. the percentage of red blood cells in the blood. The subscripts n and e refer to native and enhanced respectively

and *blood* to reference blood T1 values obtained from a manually selected region in the left ventricular blood pool.

2.3. Implementation

²³⁵   The proposed framework is implemented in Matlab and has been executed on an iMac with 3.5 GHz Intel Core i5 processor and 32 GB RAM. The pairwise registrations are calculated using Elastix (Klein et al. (2010)).

Prior to registration, a rectangular ROI around the heart is manually selected on the first image of each sequence in order to focus the registration on the heart ²⁴⁰ only and to decrease computational complexity. The co-registration additionally uses a registration mask representing the heart. This mask is automatically segmented from the native T1 map obtained after initial data-driven registration using low level segmentation techniques.

The non-rigid transformations are parameterized using a B-spline (Rueck- ²⁴⁵ ert et al. (1999)) regularized with a rigidity penalty (Staring et al. (2007)), which constrains deformation identically over the whole image domain. The mutual information similarity measure is used for data-driven initialization and co-registration. For the set of samples required to evaluate the pairwise affine registrations (Eq. 2) in the data-driven initialization only one sample, the center ²⁵⁰ point in the ROI, is used, as using more points was not found to influence over- all performance. The model-based motion correction minimizes the normalized cross-correlation between every corresponding actual image and model image since their intensities are expected to be similar. We perform five steps of the model-based motion correction.

²⁵⁵   Computation time to obtain two registered T1 maps (native and enhanced, typically about 10 images per scan) and a registered ECV map with this im- plementation, which is not optimized for speed, is about fifteen minutes. The model-based registration, including a new curve fitting in every iteration, takes about 80s per iteration and thereby consumes major part of the time.

2.4. Evaluation

2.4.1. Phantom experiment: two versus three parameter model

To demonstrate the improved performance of our two parameter model registration versus a more straightforward three parameter model registration, we compared the two methods on a physical phantom. The phantom was constructed using three cylinders representing respectively the native T1 of blood, myocardium and fat (Fig. 3a). Two native MOLLI-scans, shifted 5mm in the vertical image direction, consisting of 10 T1-weighted images (repetition time $= 2.06ms$, echo time $= 0.93ms$, FOV $= 300 \times 300mm^2$, acquisition matrix $=$ 152x150, reconstructed pixel size $= 1.17$x1.17mm, slice thickness $= 8mm$, flip angle $= 35°$ and sensitivity encoding factor $= 2$) and assuming a heart rate of 66 BPM were acquired with a 1.5T MR scanner (Ingenia, Philips Healthcare, Best, the Netherlands) following the 5s(3s)3s MOLLI scheme. The two scans are used to create 10 new sequences where for each sequence, one image is displaced with respect to the others. We applied the model-based step of our registration method (thus without data-driven initialization) to each of the new sequences and calculated the remaining average displacement error of the shifted image after registration over a square region of interest.

2.4.2. In vivo validation

For the validation of our registration framework we used three different datasets: (1) a dataset of healthy, young athletes (n $=$ 15, males (M) $=$ 7, age $= 16$+/-1 years), (2) a dataset containing all T1-weighted image sequences from our CMR database with histologically confirmed cardiac amyloidosis in a period of two years (n $=$ 29, M $=$ 24, age $= 72$+/-10 years) and (3) the publicly available dataset of free-breathing T1 mapping data of El-Rewaidy et al. (2018) (n $=$ 210, M $=$ 134, age $= 57$+/-14 years). The use of datasets 1 and 2 was approved by the ethical committee of our hospital (S60774). All datasets were processed identically.

The 2D+time scans of dataset 1 and 2 were recorded using a 1.5 T MR scanner (Achieva (athletes) and Achieva or Ingenia (amyloidosis patients), Philips
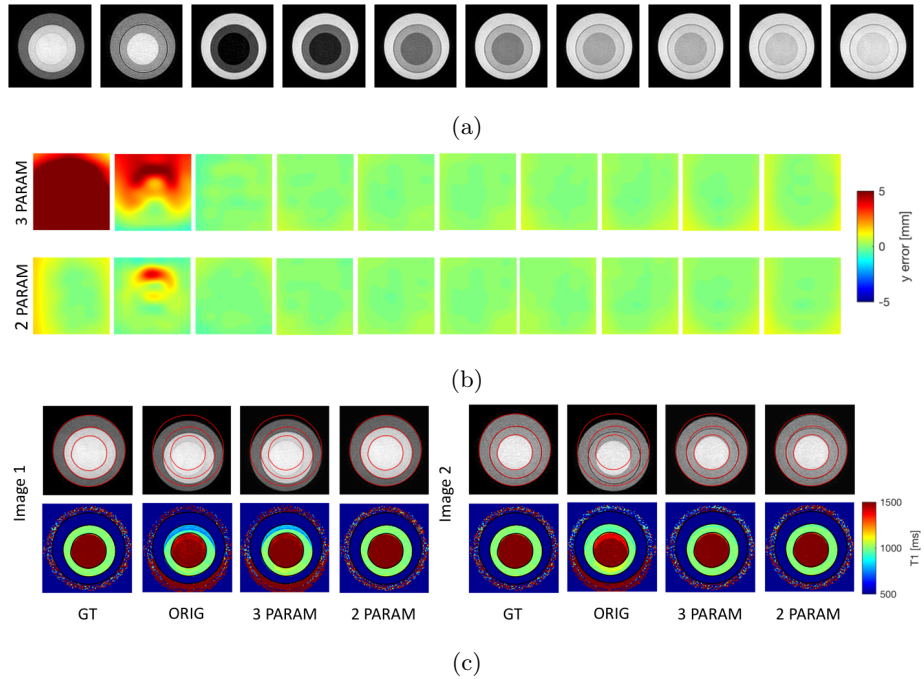
(a)

(b)

(c)

Figure 3: (a) T1-weighted phantom images in order of increasing TI. (b) Residual displacement, for an initial translation of 5mm in the vertical direction of each image separately, after registration using a three (top) or two parameter (bottom) model. (c) Result when shifting the first image (shortest TI) only (left panel) or the second image (2nd shortest TI) only (right panel). Both panels show from left to right: ground truth image; original image (after initial displacement); recovered image with three parameter model; recovered image with two parameter model. The bottom row shows the corresponding (three parameter fitted) T1 maps. The larger residual displacements of the three parameter model shown in Fig. 3b result in a disturbed T1 map, especially for image 1.

Healthcare, Best, the Netherlands) and a 6-channel (Achieva) or 32-channel (Ingenia) cardiac phased array receiver coil. For the enhanced scans, 0.15 mmol/kg of gadobutrol (Gadovist, Bayer Schering) was administered to the subjects. The 5s(3s)3s and the 4s(1s)3s(1s)2s MOLLI schemes (Messroghli et al. (2004); Kellman and Hansen (2014)), imposing breath-hold, were used for the native and enhanced scans respectively. Following these schemes, 7 to 16 images per native scan and 9 to 16 images per enhanced scan were obtained (repetition time = $2.24\pm0.10ms$, echo time = $1.03\pm0.05ms$, flip angle = $35°$, sensitivity encoding

factor = 2). The typical image geometry was FOV = $300 \times 300mm^2$, acquisition matrix = 152×150, reconstructed pixel size = $1.17 \times 1.17mm^2$ and slice thickness = $10mm$, although in about 12% of the image sequences a different FOV, matrix and pixel size was used e.g. FOV = $348 \times 348mm^2$, acquisition matrix = 176×174 , reconstructed pixel size = $1.09 \times 1.09mm^2$. The mid-cavity short-axis images of dataset 1 and short-axis (SAX) images at three levels (base, mid and apical) as well as horizontal long axis (HLA) views of dataset 2 were used for validation.

The third dataset contains native SAX images at five levels from apical to basal for 210 patients with known or suspected cardiovascular disease, acquired using the STONE sequence (Weingärtner et al. (2015)). The scans were recorded under free-breathing condition using a 1.5 T MR scanner (Achieva, Philips Healthcare, Best, the Netherlands), a 32-channel cardiac phased array receiver coil and a respiratory navigator using prospective tracking with navigator tracking slice factor of 0.6. Each scan contains 11 T1-weighted images with repetition time = $2.7ms$, echo time = $1.37ms$, flip angle = $70°$, sensitivity encoding factor = 1.5, FOV = $360 \times 351mm^2$, acquisition matrix = 172×166, reconstructed pixel size = $1.13 \times 1.13mm^2$ and slice thickness = $8mm$.

The endo- and epicardium were manually delineated on every (non-corrected) T1-weighted image by two different, independent observers to assess the inter-observer variability (dataset 1), or one observer (dataset 2 + 3). After registration, the segmentations were propagated to the reference image space by applying the calculated transformations to binary images representing the myocardium.

For perfect image alignment and assuming no segmentation errors, the overlap of the myocardial segmentations should be optimal. This is assessed by the Dice similarity coefficient DSC (Dice (1945)) averaged over all possible image pairs (k, l) in a native or enhanced scan:

$$DSC_M = \frac{\sum_{k=1}^{ns} \sum_{l=1, l \neq k}^{ns} DSC(k, l)}{ns \cdot (ns - 1)}, \tag{8}$$

with

$$DSC(k,l) = \frac{2 \cdot (S_k \bigcap S_l)}{S_k + S_l}, \tag{9}$$

where S_k is the myocardium segmentation of image k and ns is the number of images in the scan. To evaluate the result of co-registration, $DSC(k,l)$ averaged over all native-enhanced image pairs is used:

$$DSC_{M,co} = \frac{\sum_{k=1}^{ns_n} \sum_{l=1}^{ns_e} DSC(k,l)}{ns_n \cdot ns_e}. \tag{10}$$

Furthermore, the global Dice similarity coefficient (Huizinga et al. (2016)), which assesses the overlap of all images simultaneously and is thereby sensitive to the misregistration of a single image, is evaluated:

$$DSC_G = \frac{ns \cdot (S_1 \bigcap S_2 ... \bigcap S_{ns})}{S_1 + S_2 + ... + S_{ns}}. \tag{11}$$

Additionally, the mean distance between endo- and epicardial contours of different images is calculated (Mean Boundary Error, MBE) and averaged over all possible image pairs identically as for DSC_M and $DSC_{M,co}$.

To assess the influence of the motion correction on actual T1 and ECV values, accuracy of native and enhanced T1 values as well as of ECV values is determined for the different cardiac segments proposed by the American Heart Association (AHA) (Cerqueira et al. (2002)) before and after registration. The different segments are obtained by manual delineation of the reference image. Segmental T1 and ECV accuracy are calculated similarly as in El-Rewaidy et al. (2018) to be the mean of the difference between the pixelwise T1/ECV values in the segment of interest and a ground truth T1/ECV value for this segment. To calculate the ground truth T1/ECV values, a ground truth image alignment is first obtained by warping all T1-weighted images to the reference image (see Sec. 2.1.1) using a thin-plate spline transformation model to exactly align the manually delineated contours. The pixelwise T1/ECV values of these warped images are calculated inside the manual segmentation of the reference image and a single ground truth value per segment is obtained by averaging. Additionally, standard deviation (SD) error, calculated according to Kellman et al. (2013), is used. This metric represents the SD of T1/ECV estimate in T1/ECV units

16

for every pixel separately to construct so-called SD maps (see Fig. 4a) and is obtained by transforming the SD of the error on the curve fit into the SD of the estimated parameters (i.e. T1, ECV). The transformation of SD was derived analytically in Kellman et al. (2013). The SD error was averaged per segment and reported similarly as for accuracy. We further calculated the septal T1 and ECV values (segments 8 and 9) in dataset 1, using the contours of the reference image, to define their range in this healthy group.

The statistical significance of all obtained results is assessed using the two-sided Wilcoxon signed rank test with a significance level of 5%.

2.4.3. Robustness

Robustness of the different steps of our approach is shown through four separate experiments: 1) To demonstrate the need for data-driven initialization, DSC_M, DSC_G and MBE are additionally calculated for a registration approach where the data-driven initialization is omitted. 2) We illustrate the improved performance of our data-driven initialization compared to a simpler initialization, which performs a pairwise registration of each image to the image with maximal TI, with an example. 3) The robustness of a two parameter model compared to a three parameter model is evaluated on the native scans of dataset 1 (n = 15). Similar to the phantom experiment, data-driven initialization was omitted. 4) The robustness of our co-registration approach is shown for two cases for which native and enhanced scans are difficult to align using one of the two most straightforward registration approaches: registration of native and enhanced T1-weighted image with maximal TI and registration of native and enhanced T1 map directly.

3. Results

3.1. Phantom experiment

Fig. 3a and 3b show respectively the different T1-weighted images in the phantom scan (in order of increasing TI) and the remaining displacement error

17

in vertical direction after model-based registration with a two or three parameter model when the corresponding image is initially translated. The translation of images 3 to 10 can be equally well corrected by both models, but while the two parameter model is able to obtain an average accuracy smaller than $1mm$ for all images, the three parameter model completely fails to restore alignment if the first image (lowest TI) is displaced (Fig. 3c).

3.2. In vivo validation

The DSC_M, DSC_G and MBE for the three datasets and for different observers or slices are given in Tab. 1-4. All results are given as mean value $\pm$ standard deviation (std). The three datasets are considerably different in prior motion, which is reflected in their DSC_M, DSC_G and MBE values before motion correction.

Dataset 1 contains notable motion, in particular also septal shift (e.g. Fig. 4a). For native scan, enhanced scan and co-registration, significant improvement is obtained after registration. Furthermore, a significant difference can be observed between the values for the different observers: the inter-observer DSC, evaluated before motion correction, is equal to $79.0 \pm 5.1\%$ and $76.9 \pm 6.7\%$ for native and enhanced images respectively. Additionally, the endocardial MBE equals $1.15 \pm 0.36mm$ and $1.30 \pm 0.39mm$ and the epicardial MBE $1.64 \pm 0.36mm$ and $1.71 \pm 0.42mm$ respectively. These values for the inter-observer variability are typically lower (DSC) or higher (MBE) than the average results obtained after registration.

Dataset 2 is generally characterized by limited motion and a thicker myocardium resulting in larger initial DSC_M and DSC_G. Statistically significant increase of DSC_M was only obtained for the native mid-cavity SAX image set, for the enhanced HLA image set and for the co-registration of apical and mid-cavity SAX. DSC_G only improved for co-registration of apical and mid-cavity SAX image sets. A statistically significant decrease in MBE was additionally obtained for the enhanced mid-cavity SAX image set. It can also be observed that registration accuracy, as assessed by DSC_M and DSC_G (MBE) based on

Table 1: DSC_M, DSC_G and MBE for dataset 1 before motion correction ($ORIG$), after model-based motion correction without data-driven initialization (REG_{noDI}) and after the proposed motion correction (REG). The results were calculated using the manual contours of two independent observers (Obs 1 and Obs 2). Statistically significant improvement ($P < 0.05$) of REG vs $ORIG$ is indicated with gray background. Statistically significant improvement ($P < 0.05$) of REG vs REG_{noDI} is indicated with ∗. For every experiment, the data of 15 subjects was used.

Dataset 1	Native		Enhanced		Co-reg	
	Obs 1	Obs 2	Obs 1	Obs 2	Obs 1	Obs 2
$DSC_M[\%]$						
$ORIG$	79.1±9.2	84.8±7.7	76.4±17.3	81.1±1.3	39.4±20.7	52.5±21.5
REG_{noDI}	85.8±4.0	87.5±2.8	84.7±8.6	87.1±7.1	80.1±5.2	83.2±3.8
REG	85.7±3.7	87.0±3.4	86.0±5.8	88.0±4.9	80.5±4.0	83.6±2.8
$DSC_G[\%]$						
$ORIG$	53.1±17.9	63.7±16.3	49.1±26.8	58.8±24.4	11.3±9.8	24.4±15.2
REG_{noDI}	62.1±9.5	67.1±7.2	57.2±18.1	64.5±18.5	42.1±15.9	50.4±15.3
REG	61.1±8.3	64.9±8.7	60.2±13.7	66.9±12.2	41.3±13.2	51.24±9.9
$MBE[mm]$						
$ORIG$	1.17±0.48	1.21±0.53	1.40±1.29	1.58±1.41	3.76±1.81	3.74±1.96
REG_{noDI}	0.68±0.18	0.86±0.17	0.75±0.57	0.96±0.58*	0.97±0.29	1.20±0.27
REG	0.68±0.16	0.88±0.22	0.67±0.31	0.88±0.39	0.94±0.22	1.15±0.17

manually defined contours, is lower (higher) for apical slices, which is likely due to the more complicated and variable appearance (e.g. papillary muscles) of the apex complicating both the registration as well as the manual segmentation.

Dataset 3, acquired under free-breathing condition, contains the largest motion of the three datasets. Statistically significant improvement after motion correction is observed for all slices for DSC_M, DSC_G and MBE. Similarly to dataset 2, DSC_M (DSC_G) and MBE are respectively lowest and highest for apical slices. Furthermore, MBE after motion correction is higher for dataset 3 compared to dataset 1 and 2 since free-breathing data is more challenging to register e.g. because of a higher prevalence of out-of-plane motion. The results also show that DSC_G is a more conservative measure compared to DSC_M.

Table 2: DSC_M, DSC_G and MBE for native and enhanced scans of dataset 2 before motion correction ($ORIG$), after model-based motion correction without data-driven initialization (REG_{noDI}) and after the proposed motion correction (REG). Statistically significant improvement ($P < 0.05$) of REG vs $ORIG$ is indicated with gray background. No significant difference between REG_{noDI} and REG was observed. The number of scans used for every experiment (n) is indicated.

Dataset 2		SAX apex	SAX mid	SAX base	HLA
Native	n	16	30	17	28
$DSC_M[\%]$	$ORIG$	82.9±7.4	87.2±7.1	88.6±5.4	89.4±4.9
	REG_{noDI}	83.4±6.8	89.2±4.5	89.9±2.8	90.7±2.9
	REG	83.2±7.0	89.7±2.9	89.8±2.7	90.0±3.0
$DSC_G[\%]$	$ORIG$	56.2±16.9	67.1±16.4	67.9±16.0	72.1±13.0
	REG_{noDI}	53.8±18.5	70.0±12.1	72.2±7.6	75.1±8.2
	REG	53.2±19.3	71.1±9.0	71.4±7.5	72.5±8.3
$MBE[mm]$	$ORIG$	1.22±0.38	1.27±0.78	1.09±0.44	1.16±0.49
	REG_{noDI}	1.14±0.32	1.04±0.46	0.97±0.23	1.01±0.25
	REG	1.15±0.32	0.98±0.24	0.98±0.23	1.08±0.2
Enhanced	n	12	22	11	23
$DSC_M[\%]$	$ORIG$	82.9±7.9	84.9±10.0	87.4±9.0	87.5±6.1
	REG_{noDI}	83.9±5.5	88.0±5.6	88.3±8.1	89.6±3.7
	REG	83.7±5.5	88.7±4.0	88.7±6.7	89.5±3.8
$DSC_G[\%]$	$ORIG$	55.3±17.5	61.7±21.1	67.6±17.7	66.3±16.0
	REG_{noDI}	53.9±14.6	65.0±15.3	68.0±18.0	70.4±9.9
	REG	52.6±15.1	66.6±12.8	69.1±14.9	69.4±11.2
$MBE[mm]$	$ORIG$	1.34±0.67	1.50±0.99	1.26±0.79	1.37±0.66
	REG_{noDI}	1.23±2.35	1.18±0.54	1.13±0.53	1.11±0.34
	REG	1.24±0.36	1.12±0.42	1.10±0.44	1.13±0.35

The final purpose of motion correction for T1 and ECV mapping is to obtain more accurate T1 and ECV values. The visual improvement of T1 and ECV maps after motion correction is shown for the three datasets in Fig. 4 and 5. Fig. 4a shows native T1, enhanced T1 and ECV maps, together with their corresponding SD errormaps (Kellman et al. (2013)) before and after motion correction for one subject of dataset 1, which clearly shows septal shift motion.

Table 3: DSC_M, DSC_G and MBE for co-registration of dataset 2 before motion correction ($ORIG$), after model-based motion correction without data-driven initialization (REG_{noDI}) and after the proposed motion correction (REG). Statistically significant improvement ($P < 0.05$) of REG vs $ORIG$ is indicated with gray background. No significant difference between REG_{noDI} and REG was observed. The number of scans used for every experiment (n) is indicated.

Dataset 2		SAX apex	SAX mid	SAX base	HLA
Co-reg	n	9	19	10	21
$DSC_M[\%]$	$ORIG$	54.3±19.6	70.0±17.5	72.0±20.2	75.3±15.4
	REG_{noDI}	71.0±10.3	82.1±6.1	81.6±9.3	80.9±7.9
	REG	70.7±10.2	82.8±4.6	81.5±9.4	80.2±8.0
$DSC_G[\%]$	$ORIG$	25.5±19.6	40.2±20.8	45.8±25.7	49.9±18.5
	REG_{noDI}	32.2±17.2	51.6±14.5	53.7±2.08	56.8±10.7
	REG	30.7±18.6	52.7±13.2	54.3±18.2	54.5±10.1
$MBE[mm]$	$ORIG$	4.68±5.23	1.71±0.65	2.58±1.44	2.64±1.80
	REG_{noDI}	2.25±1.90	1.71±0.65	1.78±0.84	2.10±1.07
	REG	2.24±0.90	1.64±0.49	1.78±0.87	2.14±1.02

For dataset 2, one SAX mid-cavity case and one HLA case of two different patients are shown in Fig. 4b. The increased native T1, increased ECV and higher heterogeneity in the myocardium are all in agreement with the diagnosis of cardiac amyloidosis in these patients. Since dataset 3 is acquired under free-breathing condition, the increase in image quality after motion correction is even more pronounced (Fig 5).

Mean septal (native) T1 and ECV in dataset 1 significantly decreased respectively from $1046\pm71ms$ and $37\pm9\%$ before registration to $1010\pm24ms$ and $26\pm2\%$ after registration. The accuracy and SD error before and after motion correction are given in Fig. 6 and 7 according to the AHA convention (Cerqueira et al. (2002)). Due to frequent occurrence of septal motion in dataset 1, the improvement in enhanced T1 accuracy is significant for segments 8 and 9. Improvement in ECV accuracy is significant for all segments in dataset 1, with an average accuracy below 1.9% for each segment. This can be explained by

Table 4: DSC_M, DSC_G and MBE for dataset 3 before motion correction ($ORIG$), after model-based motion correction without data-driven initialization (REG_{noDI}) and after the proposed motion correction (REG). Statistically significant improvement ($P < 0.05$) of REG vs $ORIG$ is indicated with gray background, and statistically significant improvement ($P < 0.05$) of REG vs REG_{noDI} with $*$. All five views of 196 patients were used for validation.

Dataset 3	1 (apical)	2	3	4	5 (basal)
$DSC_M[\%]$					
$ORIG$	64.9±11.5	68.8±11.1	71.3±10.2	72.2±10.4	73.8±9.9
REG_{noDI}	75.6±9.5	79.5±7.8	82.0±6.6∗	82.1±7.4∗	83.1±6.9∗
REG	75.8±9.2	80.2±6.7	82.8±5.4	83.0±5.9	83.9±5.5
$DSC_G[\%]$					
$ORIG$	17.9±16.1	23.9±17.6	27.6±18.4	29.7±18.2	32.3±18.9
REG_{noDI}	39.6±17.6	47.4±15.6	53.2±14.4*	52.7±15.9*	54.8±15.6
REG	39.3±17.3	48.2±15.0	54.5±13.3	53.9±15.2	55.7±14.7
$MBE[mm]$					
$ORIG$	3.28±1.49	3.19±1.39	3.19±1.36	3.34±1.47	3.15±1.35
REG_{noDI}	2.04±1.09	1.83±0.91*	1.72±0.81*	1.83±0.92*	1.74±0.86*
REG	2.01±1.73	1.73±0.76	1.61±0.64	1.71±0.74	1.63±0.73

the larger motion between native and enhanced scans compared to the motion between the images in one scan, illustrating that registration is crucial for ECV mapping. Accuracy is also improved in dataset 2, except for the apical segments. Furthermore, the SD error decreases significantly for all segments, except for native T1 of segment 11 in dataset 1 and ECV of segment 2 in dataset 2. Due to the increased motion in dataset 3, significant improvement after motion correction is obtained for every segment for both accuracy and SD error.

3.3. Robustness

Additionally, the motion correction was performed without data-driven initialization (REG_{noDI} in Tab. 1-4 and Fig. 5 and 7). This did not show any statistically significant deterioration compared to the proposed method for the datasets acquired under breath-hold condition (dataset 1 and 2), except for the
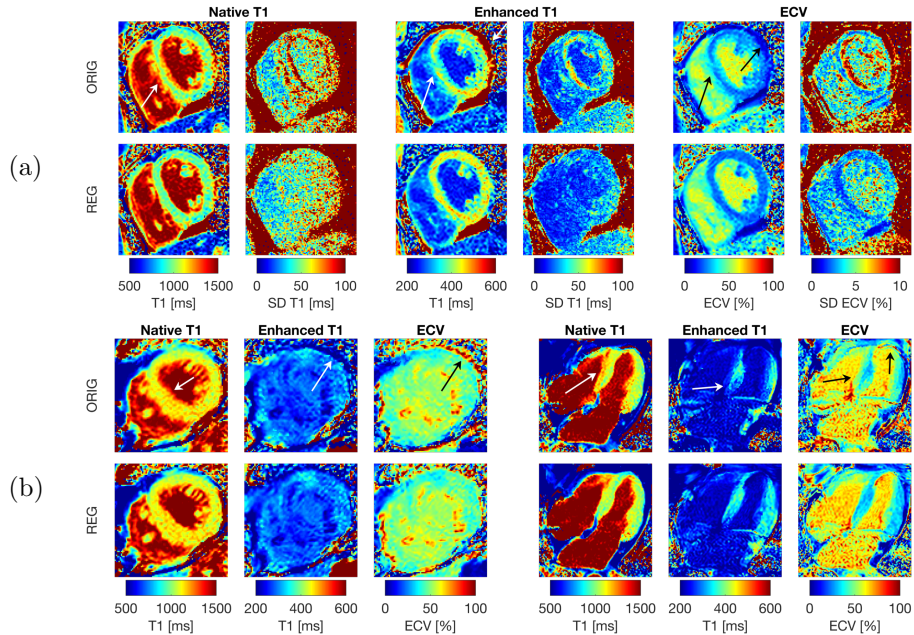
Figure 4: (a) From left to right: native T1, enhanced T1 and ECV maps and corresponding SD errormaps (Kellman et al. (2013)) of dataset 1 before ($ORIG$) and after (REG) registration. In this subject, septal shift motion can be observed before registration. (b) From left to right: native T1, enhanced T1 and ECV maps before ($ORIG$) and after (REG) registration of the mid-cavity short-axis slice of one patient (left) and the HLA slice of another patient (right) of dataset 2. In amyloidosis patients, the native T1 and ECV in the myocardium are elevated, as apparent in the case on the left.

MBE of enhanced scans of observer 2 in dataset 1. Because of the increased initial motion in dataset 3, initialization is more important in this dataset. For ₄₅₅ this dataset, omitting the data-driven initialization resulted in a significant deterioration of accuracy in 15 out of 28 segments and of SD error in 11 out of 28 segments (Fig. 7).

Fig. 8 shows two examples of T1-weighted image sets before and after motion correction and illustrates the wide variety in content and global image ₄₆₀ appearance. Especially the low-signal images (fourth image in both sets) are difficult to register with general purpose registration methods, but are successfully handled by our method. This is further illustrated in Fig. 9, which shows
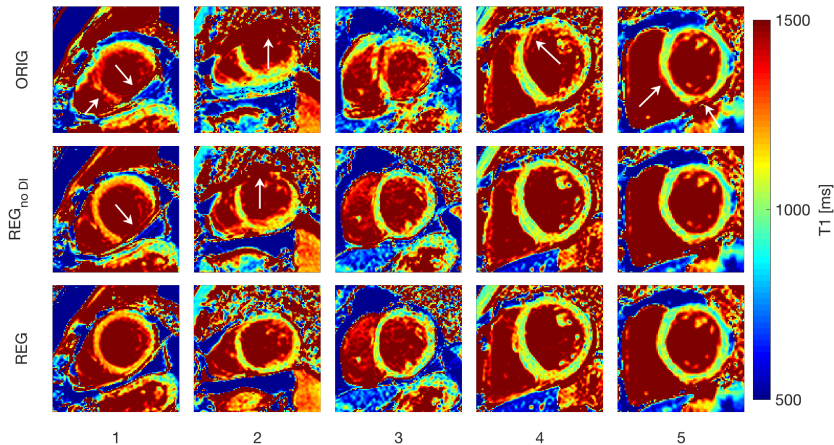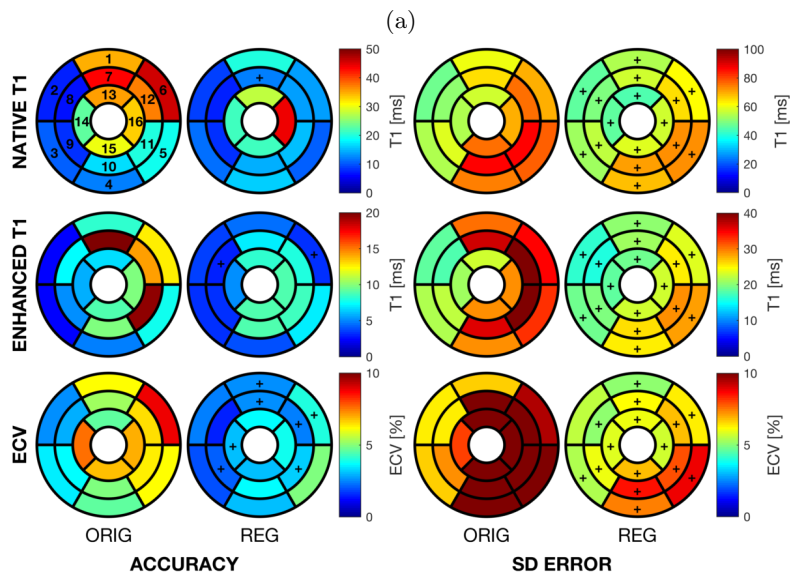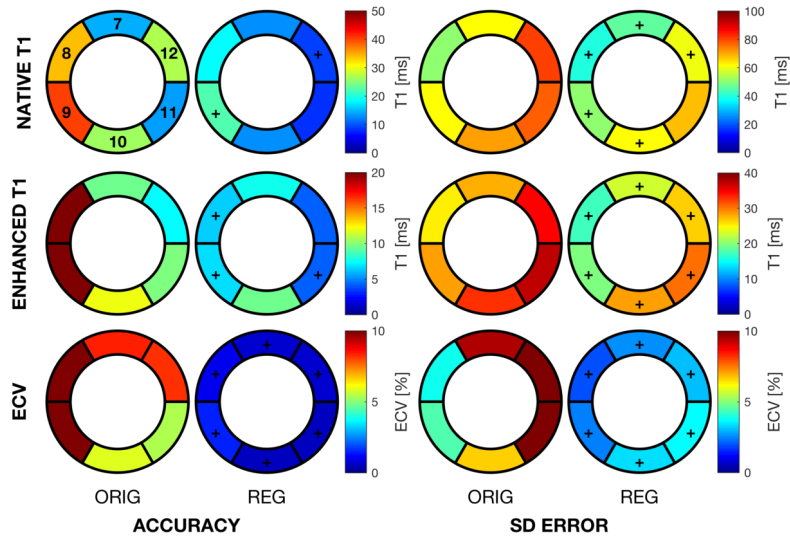
23

Figure 5: Native T1 maps before motion correction ($ORIG$), after model-based motion correction without data-driven initialization (REG_{noDI}) and with the proposed motion correction method combining data-driven initialization and model-based registration (REG) of dataset 3 for five different patients and five different slices from apical to basal (left to right). Because dataset 3 contains free-breathing data, motion correction is crucial for construction of T1 maps.

the result of our method and of a similar approach with an alternative data-driven initialization. This alternative initialization registers every image to the image with maximal TI. In the latter case, the second image (low signal image) is erroneously deformed.

The robustness of our two parameter model-based approach compared to a more straightforward three parameter model-based approach is evaluated on the native scans of dataset 1, using respective model-based approaches without data-driven initialization. The improvement with the two parameter model was significant for DSC_M ($85.8 \pm 4.0\%$ vs $83.8 \pm 5.9\%$), DSC_G ($62.1 \pm 9.5\%$ vs $56.7 \pm 15.6\%$) and MBE ($0.68 \pm 0.18mm$ vs $0.78 \pm 0.29mm$).

Fig. 10 illustrates the robustness of our model-based similarity function for co-registration (Eq. 6) compared to more standard co-registration approaches, including registration of native and enhanced images with longest TI or registration of native and enhanced T1 map directly. In Fig. 10a, the epicardium is difficult to perceive in the long TI images due to fluid around the heart whereas

24

Figure 6: Accuracy (left) and SD error (right) of native T1 (top), enhanced T1 (middle) and ECV (bottom) before (ORIG) and after (REG) registration for (a) dataset 1 and (b) dataset 2. Segments which show a statistically significant improvement ($P < 0.05$) after registration are indicated with +.
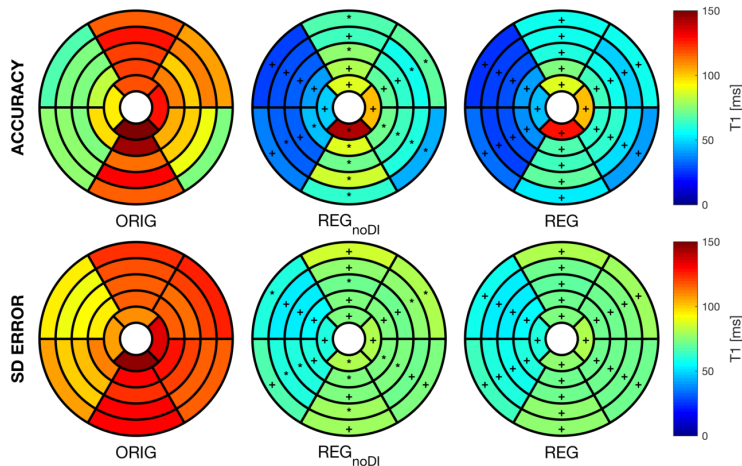
Figure 7: Accuracy (top) and SD error (bottom) of native T1 before registration (ORIG), after model-based motion correction without data-driven initialization (REG_{noDI}) and after the proposed registration (REG) for dataset 3. Segments which show a statistically significant improvement after registration ($P < 0.05$) are indicated with +. Segments in REG_{noDI}, which show both statistically significant improvement after registration ($P < 0.05$) and significant deterioration compared to the proposed method ($P < 0.05$) are indicated with *. T1 and SD error are calculated with the two parameter model.

in Fig. 10b the registration of the two T1 maps directly will fail due to the increased enhanced myocardial T1 shortening in amyloidosis patients. Joint registration of A^* and $T1^*$ map avoids these issues in both cases.

4. Discussion

In this paper, we presented a method for motion correction for T1 and ECV mapping, which combines a model-based approach with a data-driven initialization. In model-based registration approaches, the underlying T1 relaxation model is exploited such that direct registration between images with largely different or inverted contrast can be avoided, which is the main challenge when performing registration on a set of T1-weighted images. A difficulty in this model-based approach is that the model is estimated for every pixel separately from a set of images, which is generally suboptimally aligned before motion correction, such that motion artifacts will also be present in the model images in
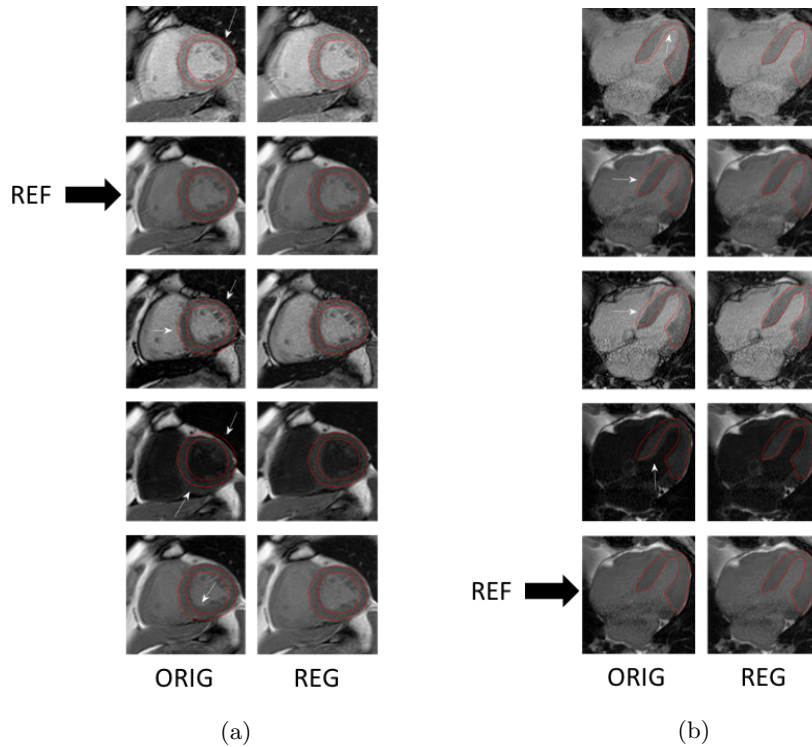
Figure 8: Five images of a native MOLLI scan from two different patients from dataset 2: (a) SA and (b) HLA images (in order of acquisition) before (ORIG) and after (REG) the proposed motion correction. The second (SA) and last (HLA) image shown were automatically selected as reference image. The reference contour is displayed on all images.

case of large motion (e.g. Fig. 2b, second image of three parameter model). To circumvent motion artifacts in the model, Xue et al. (2012) proposes to do an initial registration between the image with the shortest TI and the image with the largest TI and uses these images to calculate the initial model images. How-
495 ever, decreasing the number of sample points for curve fitting will negatively affect the accuracy of the estimated parameters, resulting in model images with a contrast that deviates from the actual image. Consequently, Xue et al. (2012) adds an additional minimization process to create the final synthetic images used for registration, in which also similarity of contrast to the actual images
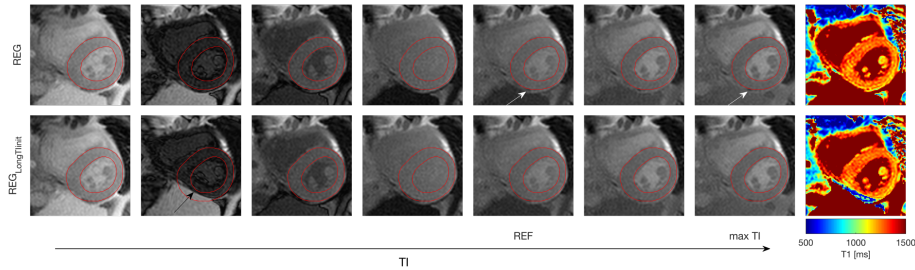500 is enforced. The construction of reliable model images based on curve fitting

27

Figure 9: T1-weighted images, in order of increasing TI, and T1 map after the proposed motion correction (top, REG) and after registration with an alternative data-driven initialization (bottom, $REG_{LongTIinit}$). This alternative initialization registers every image to the image with maximal TI. For $REG_{LongTIinit}$, misregistration is observed in the second image (black arrow), which results in a disturbed T1 map. The chosen reference image according to our method (REF) shows an increased contrast at the epicardium (white arrows) compared to the image with maximal TI. Only images of the first inversion are shown.

alone, is thus essentially a compromise between on the one hand a higher similarity in contrast with the actual images by using more images and on the other hand a motion-free model by using only the images that are initially aligned by e.g. a standard data-driven registration method.

⁵⁰⁵ In our method, the model-based registration is preceded by a global data-driven initialization approach such that the initial motion is already partially corrected and convergence of the model-based registration is accelerated. The data-driven approach tries to estimate the suitability of each image pair for pairwise registration to avoid registration failures. However, this is insufficient ⁵¹⁰ for accurate motion correction on itself since the appearance of images is too variable. For example, fat is perceived as high intensity structure in nearly all images, except for images with very low TI. Images that are prone to registration failure are discarded for data-driven registration to increase robustness. Different to our approach, Van De Giessen et al. (2013) performs model-based ⁵¹⁵ registration of a Look-Locker image sequence without initialization.

Furthermore, as shown in Fig. 2 and Fig. 3, using a two instead of three parameter model for construction of model images decreases the motion artifacts
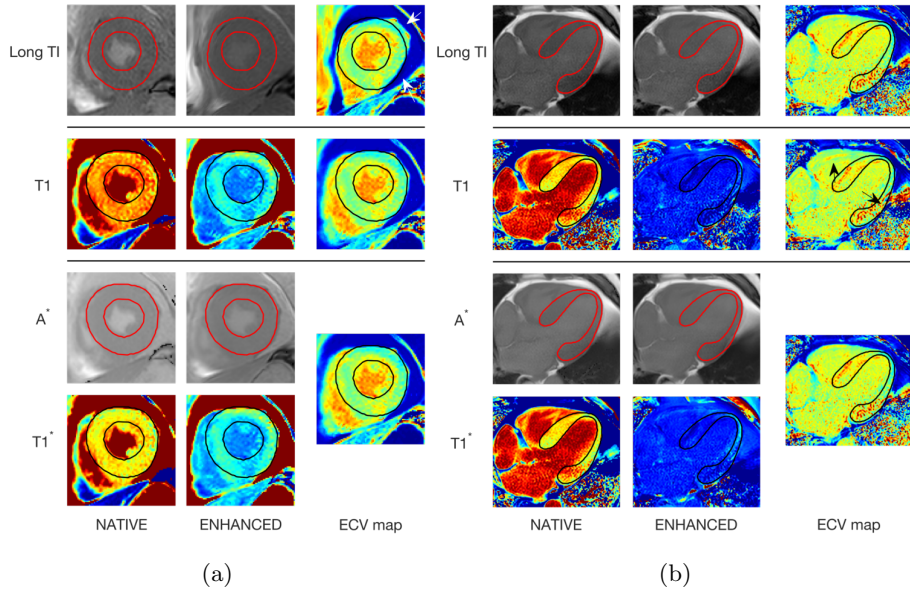
28

Figure 10: Three different approaches for co-registration: registration of native and enhanced image with maximal TI (top), registration of native and enhanced T1 map (middle) and joint registration of $T1^*$ and A^* images (Eq. 6). The images that are registered in every approach are shown for a patient with fluid around the heart (a) and an amyloidosis patient (b). The resulting ECV map after registration is also shown.

in the model images and thus increases the robustness of motion correction, especially in images with low TI. The three parameter model however theoretically better describes the evolution in successive T1-weighted MOLLI images than the two parameter model, such that intuitively, motion correction accuracy will be further increased when a model-based registration step with the three parameter model is performed after our proposed method. Since our initial experiments did not show clinically relevant improvement with this approach, it was not analyzed further. For the data acquired using the STONE sequence, Weingärtner et al. (2015) showed that the two parameter model is more precise but less accurate than the three parameter model.

A drawback of our current method is the relatively long processing time. The cost function in data-driven initialization is constructed using the consistency

of pairwise forward and backward affine registration transformations, implying that ns^2 affine registrations should be performed of which only few ($\sim ns$) are finally used for transformation. However, this has only a limited effect on the total computation time since the main part of it is attributed to the repeated pixelwise curve fittings in each iteration step. This dominance of the curve fitting process on computation time was also reported by Xue et al. (2012) for their method. In this respect, we expect the computation time for both methods (in similar implementations) to relate to each other proportionally to the number of iterations performed, which was five for our method and two for Xue et al. (2012). To circumvent the long curve fitting times, the possibility to use convolutional neural networks (CNNs) for T1 map regression, as was already done for diffusion brain imaging (Golkov et al. (2016)), could be investigated in the future. Another point for further improvement is the fact that we rely on a fixed number of iterations in our current approach, while depending on the prior motion and available contrasts in the image set, convergence speed might vary. It would therefore be useful to include a stopping condition, e.g. based on the SD error inside the heart (see Fig. 4), that automatically evaluates the quality of image alignment and the improvement of image alignment in subsequent iteration steps. Since discernible motion is present in only about $40 - 56\%$ of MOLLI breath-held datasets according to Xue et al. (2012), Xue et al. (2013) and Roujol et al. (2015)), such a quality evaluation metric might also be useful to judge whether motion correction is a priori necessary.

For the registration of native and enhanced T1 maps to obtain a motion-free ECV map, two different approaches are found in literature: (1) registration of native and enhanced T1 map directly (Zhang et al. (2018)), (2) registration of native and enhanced images with the longest TI (Kellman et al. (2012)). We proposed to perform co-registration by optimizing a similarity function, which is the sum of the mutual information between images of the A^* and $T1^*$ parameters of Eq. 5 to optimally exploit the information of a complete set of T1-weighted images, since both images contain complementary information. This

especially improves the registration result in cases where different structures appear equally intense in one of the images. For example, the epicardium can be difficult to perceive in images with long TI when the heart is surrounded by fluid (Fig. 10a) or the enhanced T1 map can have low contrast between myocardium and blood pool in case of amyloidosis (Fig. 10b). The reason to choose the $T1^*$ and A^* images over the $T1$, A and B images, which could be obtained from Eq. 4 is that both image sets contain similar information, that the former is readily available in our framework and that including both A and B images will lead to increasing computation time while the additional benefit was found to be very limited in an initial experiment.

As no direct registration ground truth is available, we validated our method using manual segmentations, as Xue et al. (2012); Roujol et al. (2015); Huizinga et al. (2016); Guyader et al. (2016); Van De Giessen et al. (2013); El-Rewaidy et al. (2018). These are inherently biased by the segmentation strategy of the observer as can be appreciated from the large inter-observer variability on dataset 1 (Tab. 1). The obtained DSC_M (DSC_G) of observer 1 is consistently smaller than that of observer 2. Furthermore, since it is difficult to delineate the myocardium in low contrast images or low signal images, also intra-observer variability is likely in these segmentations, which can introduce additional bias on DSC_M, DSC_G and MBE. Nevertheless, we included the segmentations of all images as alignment of these images is also crucial for accurate T1 estimation since their pixel intensities are located in the initial, most steeply ascending part of the exponential T1 relaxation curve (see Sec. 2.1.1). In our results, we found significant improvement of DSC_M, DSC_G and MBE after motion correction. However, the significance of the difference between the values before and after motion correction is not only dependent on the registration algorithm itself, but also on the amount of motion initially present in the images. For data with little or no motion, the purpose cannot be to significantly increase DSC, but rather not to deteriorate it. In Tab. 1-4, both DSC_M and DSC_G are shown where DSC_M represents the average overlap between any two images and DSC_G gives

31

the percentage of myocardial area in which there is overlap for all images used to construct a T1/ECV map. In this sense, DSC_G might seem a more valuable measure to judge the quality of T1 and ECV but it might be too rigorous since misalignment of every image has not the same effect on calculated T1 and ECV. To obtain a more valuable measure of motion correction, which directly represents the improvement of T1 or ECV accuracy, the deviation in T1 due to misalignment of every image should be determined and should be taken into account in the validation measure. Due to the large variation in image contrast, TI and T1 among images and patients, this is however a difficult task. As an alternative, segmental T1 and ECV accuracy and SD error were calculated. To calculate the accuracy, a ground-truth reference T1 and ECV value is needed, which we obtained by assuming perfect image alignment if the manual contours exactly overlap, while as discussed before, these contours suffer from inter- and intra-observer variability. Additionally, since correspondence of the contours is only enforced at endo- and epicardium, no constraints are available from the data itself on how the transformation should act inside the myocardium. Hence, we forced this transformation to be smooth by using a thin-plate spline model. Furthermore, the SD error for ECV map is a scaled sum of the native and enhanced T1 error and is therefore limited in capturing registration errors between native and enhanced T1, for example, around the septum.

We compared our results with the results of El-Rewaidy et al. (2018) by using their dataset and manual contours (dataset 3) and found similarly improvement of DSC (5% increase in their DSC and 11% in our DSC_M on dataset 3) and MBE (a decrease of $1mm$ and $1.5mm$ respectively), although the actual values differ. This is probably due to a slightly different definition of both metrics. A thorough comparison of performance of the data-driven method of Roujol et al. (2015), ARCTIC, is difficult, but we will highlight two differences. First, whereas our method requires only an indication of a ROI around the heart, which should be easy to automate in the future, ARCTIC requires a delineation of LV in the reference image for both steps of their algorithm (global

32

⁶²⁵ motion estimation and local non-rigid motion estimation). Second, we defined an automatic procedure to determine the reference image in de data-driven initialization. Roujol et al. (2015) in contrast heuristically defined the image with the second longest TI (native images) or longest TI (enhanced images) to be the reference. Although these are valid choices, they are chosen for the specific ⁶³⁰ acquisition sequences and thus not transferable to other acquisitions since the performance of registration algorithms is dependent on the choice of reference image as demonstrated in Huizinga et al. (2016). Furthermore, ARCTIC treats the registration of every T1-weighted image as a separate problem, instead of our groupwise approach. This second comment also applies to the method of ⁶³⁵ Zhang et al. (2018), who chooses the image with the longest TI as reference. In contrast, Huizinga et al. (2016) did propose a groupwise approach that is based on the assumption of a non-specified low signal model in the data. Because this method directly minimizes a similarity metric over all images, it is more computationally efficient than our two-step iterative approach of model-fitting ⁶⁴⁰ and registration. However, we showed that in case of large motion, initialization is crucial before application of a model-based method. Contrary to Xue et al. (2013) who uses phase information for motion correction and Menini et al. (2015) who performs motion correction jointly with MR image reconstruction, our algorithm is completely retrospective. An advantage of our framework is ⁶⁴⁵ also that if new acquisition schemes are developed that require a different model (e.g. Shao et al. (2015)), it is in principle possible to simply adapt the fitting function.

⁶⁵⁰ Every attempt to increase the accuracy of the T1 and ECV technique will impact clinical practice since a smaller range of normal and pathological T1 and ECV values will allow to better detect subtle changes not related to inaccuracies of the technique (e.g. influence of breathing). If our motion correction algorithm is able to decrease these ranges, this will allow to improve early detection of certain disease, follow up of disease progression and monitoring of the impact of ⁶⁵⁵ therapies. In an experiment on a small group of healthy subjects (dataset 1), we

demonstrated that the range of septal T1 (std from 71 to $24ms$) and ECV (std from 9 to $2ms$) decreased after motion correction. More research is necessary to thoroughly evaluate the impact of motion correction on clinical outcome. Co-registration of native and enhanced scans is generally more critical to obtain

₆₆₀ accurate ECV values compared to the registration for T1 mapping with breath-held acquisitions, since patient motion is more likely to occur during the time between the two scans ($\sim 15min.$) and since, for the breath-held acquisitions, both scans are acquired in a different breath-hold. As a result, clinical use of ECV maps lags the use of T1 maps and only few software tools for research

₆₆₅ purpose exist (e.g. Altabella et al. (2017)). Instead, one ECV value for the septal region is usually calculated from one septal native and enhanced T1 value. Motion correction allows to construct more reliable ECV maps and thereby promotes the adoption of ECV maps in the clinic.

5. Conclusion

₆₇₀ In this paper, we presented a robust model-based method for motion correction to obtain artefact-free T1 and ECV maps. Validation on 3 datasets showed a significant increase in DSC_M, decrease in MBE and improved accuracy for T1 and ECV mapping in cases with notable motion. Motion correction in T1 and ECV mapping is highly clinically relevant in cases with substantial motion

₆₇₅ such as in respiratory-gated free-breathing data but also for ECV mapping in breath-held acquisitions.

6. Acknowledgement

₆₈₀ **References**

Altabella, L., Borrazzo, C., Carnì, M., Galea, N., Francone, M., Fiorelli, A., Di Castro, E., Catalano, C., Carbone, I., 2017. A feasible and automatic free tool

for T1 and ECV mapping. Physica Medica 33, 47–55. URL: `http://dx.doi.org/10.1016/j.ejmp.2016.12.002`, doi:`10.1016/j.ejmp.2016.12.002`.

Cerqueira, M.D., Weissman, N.J., Dilsizian, V., Jacobs, A.K., Kaul, S., Laskey, W.K., Pennell, D.J., Rumberger, J.A., Ryan, T., Verani, M.S., 2002. Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart: A Statement for Healthcare Professionals From the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. Journal of the American Society of Echocardiography 15, 463–467.

Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecology 26, 297–302.

El-Rewaidy, H., Nezafat, M., Jang, J., Nakamori, S., Fahmy, A.S., Nezafat, R., 2018. Nonrigid active shape model-based registration framework for motion correction of cardiac T 1 mapping. Magnetic Resonance in Medicine 80, 780–791. doi:`10.1002/mrm.27068`.

Floyd, R.W., 1962. Algorithm 97: Shortest path. Communications of the ACM 5, 345.

Golkov, V., Dosovitskiy, A., Sperl, J.I., Menzel, M.I., Czisch, M., Sämann, P., Brox, T., Cremers, D., 2016. q-Space Deep Learning: Twelve-Fold Shorter and Model-Free Diffusion MRI Scans. IEEE Transactions on Medical Imaging 35, 1344–1351. doi:`10.1109/TMI.2016.2551324`, arXiv:`1602.07360`.

Gupta, V., Kirili, H.A., Hendriks, E.A., van der Geest, R.J., van de Giessen, M., Niessen, W., Reiber, J.H., Lelieveldt, B.P., 2012. Cardiac MR perfusion image processing techniques: A survey. Medical Image Analysis 16, 767–785. doi:`10.1016/j.media.2011.12.005`.

Guyader, J.M., Huizinga, W., Fortunati, V., Poot, D.H., Kranenburg, M.V., Veenland, J.F., Paulides, M.M., Niessen, W.J., Klein, S., 2016. Total Correlation-Based Groupwise Image Registration for Quantitative MRI. IEEE

Computer Society Conference on Computer Vision and Pattern Recognition Workshops , 626–633doi:10.1109/CVPRW.2016.84.

Haaf, P., Garg, P., Messroghli, D.R., Broadbent, D.A., Greenwood, J.P., Plein, S., 2017. Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: a comprehensive review. Journal of Cardiovascular Magnetic Resonance 18, 89. doi:10.1186/s12968-016-0308-4.

Huizinga, W., Poot, D., Guyader, J.M., Klaassen, R., Coolen, B., van Kranenburg, M., van Geuns, R., Uitterdijk, A., Polfliet, M., Vandemeulebroucke, J., Leemans, A., Niessen, W., Klein, S., 2016. PCA-based groupwise image registration for quantitative MRI. Medical Image Analysis 29, 65–78. doi:10.1016/j.media.2015.12.004.

Kellman, P., Arai, A.E., Xue, H., 2013. T1 and extracellular volume mapping in the heart: estimation of error maps and the influence of noise on precision. Journal of Cardiovascular Magnetic Resonance 15, 56. doi:10.1186/1532-429X-15-56.

Kellman, P., Hansen, M.S., 2014. T1-mapping in the heart: accuracy and precision. Journal of Cardiovascular Magnetic Resonance 16, 2. doi:10.1186/1532-429X-16-2, arXiv:NIHMS150003.

Kellman, P., Wilson, J.R., Xue, H., Ugander, M., Arai, A.E., 2012. Extracellular volume fraction mapping in the myocardium, part 1: evaluation of an automated method. Journal of Cardiovascular Magnetic Resonance 14, 63. doi:10.1186/1532-429X-14-63.

Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J., 2010. elastix: A Toolbox for Intensity-Based Medical Image Registration. IEEE Transactions on Medical Imaging 29, 196–205. doi:10.1109/TMI.2009.2035616.

Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E., 1998. Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions. SIAM Journal on Optimization 9, 112–147. doi:10.1137/S1052623496303470.

Likhite, D., Adluru, G., DiBella, E., 2015. Deformable and Rigid Model-Based Image Registration for Quantitative Cardiac Perfusion, in: Lecture Notes in Computer Science: Statistical Atlases and Computational Models of the Heart: Imaging and Modelling Challenges: 5th International Workshop, STACOM 2014 Held in Conjunction with MICCAI 2014 Boston, MA, USA, September 18., pp. 41–50. doi:10.1007/978-3-319-14678-2.

Menini, A., Slavin, G.S., Stainsby, J.A., Ferry, P., Felblinger, J., Odille, F., 2015. Motion correction of multi-contrast images applied to T 1 and T 2 quantification in cardiac MRI. Magn Reson Mater Phy 28, 1–12. doi:10.1007/s10334-014-0440-9.

Messroghli, D.R., Radjenovic, A., Kozerke, S., Higgins, D.M., Sivananthan, M.U., Ridgway, J.P., 2004. Modified look-locker inversion recovery (MOLLI) for high-resolution T 1 mapping of the heart. Magnetic Resonance in Medicine 52, 141–146. doi:10.1002/mrm.20110.

Mewton, N., Liu, C.Y., Croisille, P., Bluemke, D., Lima, J.A.C., 2011. Assessment of myocardial fibrosis with cardiovascular magnetic resonance. Journal of the American College of Cardiology 57, 891–903. doi:10.1016/j.jacc.2010.11.013.

Nekolla, S., Gneiting, T., Syha, J., Deichmann, R., Haase, A., 1992. T1 Maps by K-Space Reduced Snapshot-Flash MRI. Journal of Computer Assisted Tomography 16, 327–332. doi:10.1097/00004728-199203000-00031.

Pontre, B., Cowan, B.R., DiBella, E., Kulaseharan, S., Likhite, D., Noorman, N., Tautz, L., Tustison, N., Wollny, G., Young, A.A., Suinesiaputra, A., 2016. An Open Benchmark Challenge for Motion Correction of Myocardial Perfusion MRI. IEEE Journal of Biomedical and Health Informatics 2194, 1–1. URL: http://ieeexplore.ieee.org/document/7542127/, doi:10.1109/JBHI.2016.2597145.

Roujol, S., Foppa, M., Weingärtner, S., Manning, W.J., Nezafat, R., 2015. Adaptive Registration of Varying Contrast-Weighted Images for Improved

Tissue Characterization (ARCTIC): Application to T 1 Mapping. Magnetic Resonance in Medicine 73, 1469–1482. doi:`10.1002/mrm.25270`.

Rueckert, D., Sonoda, L.I., Hayes, C., G Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. IEEE Transactions on Medical Imaging 18, 712–721.

Schelbert, E.B., Messroghli, D.R., 2016. State of the Art: Clinical Applications of Cardiac T1 Mapping. Radiology 278, 658–676.

Shao, J., Nguyen, K.l., Natsuaki, Y., Spottiswoode, B., Hu, P., 2015. Instantaneous Signal Loss Simulation (InSiL): An Improved Algorithm for Myocardial T 1 Mapping Using the MOLLI Sequence. Journal of Magnetic Resonance Imaging 41, 721–729. doi:`10.1002/jmri.24599`.

Staring, M., Klein, S., Pluim, J.P.W., 2007. A rigidity penalty term for nonrigid registration. Medical Physics Med. Phys 34, 4098–4108.

Tao, Q., van der Tol, P., Berendsen, F.F., Paiman, E.H., Lamb, H.J., van der Geest, R.J., 2018. Robust motion correction for myocardial T1 and extracellular volume mapping by principle component analysis-based groupwise image registration. Journal of Magnetic Resonance Imaging 47, 1397–1405. doi:`10.1002/jmri.25863`.

Tilborghs, S., Dresselaers, T., Bogaert, J., Claus, P., Maes, F., Suetens, P., 2017a. Evaluation of a Novel Model-Based Registration Algorithm for T1 and ECV Mapping on Clinical Data, in: ISMRM Workshop on Magnetic Resonance Imaging of Cardiac Function, 17-20 August 2017.

Tilborghs, S., Dresselaers, T., Claus, P., Claessen, G., Bogaert, J., Maes, F., Suetens, P., 2017b. Robust Model-Based Registration of Cardiac MR Images for T1 and ECV Mapping, in: Pop, M., Wright, G.A. (Eds.), Functional Imaging and Modelling of the Heart. FIMH 2107. Lecture Notes in Computer Science, Springer International Publishing, Cham. pp. 42–50.

URL: http://link.springer.com/10.1007/978-3-319-59448-4, doi:10.1007/978-3-319-59448-4.

Van De Giessen, M., Tao, Q., Van Der Geest, R.J., Lelieveldt, B.P.F., 2013. Model-Based Alignment of Look-Locker MRI Sequences for Calibrated Myocardial Scar Tissue Quantification, in: IEEE 10th International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1038–1041. doi:10.0/Linux-x86_64.

Viergever, M.A., Maintz, J.A., Klein, S., Murphy, K., Staring, M., Pluim, J.P., 2016. A survey of medical image registration - under review. Medical Image Analysis 33, 140–144. doi:16/S1361-8415(01)80026-8.

Weingärtner, S., Roujol, S., Akçakaya, M., Basha, T.A., Nezafat, R., 2015. Free-breathing multislice native myocardial T1 mapping using the slice-interleaved T1 (STONE) sequence. Magnetic Resonance in Medicine 74, 115–124. doi:10.1002/mrm.25387.

Xue, H., Greiser, A., Zuehlsdorff, S., Jolly, M.P., Guehring, J., Arai, A.E., Kellman, P., 2013. Phase-Sensitive Inversion Recovery for Myocardial T 1 Mapping with Motion Correction and Parametric Fitting. Magnetic Resonance in Medicine 69, 1408–1420. doi:10.1002/mrm.24385.

Xue, H., Shah, S., Greiser, A., Guetter, C., Littmann, A., Jolly, M.P., Arai, A.E., Zuehlsdorff, S., Guehring, J., Kellman, P., 2012. Motion Correction for Myocardial T1 Mapping Using Image Registration with Synthetic Image Estimation. Magnetic Resonance in Medicine 57, 1644–1655. doi:10.1002/mrm.23153.

Zhang, S., Le, T.T., Kabus, S., Su, B., Hausenloy, D.J., Cook, S.A., Chin, C.W., Tan, R.S., 2018. Cardiac magnetic resonance T1 and extracellular volume mapping with motion correction and co-registration based on fast elastic image registration. Magnetic Resonance Materials in Physics, Biology and Medicine 31, 115–129. doi:10.1007/s10334-017-0668-2.