

PCA-based groupwise image registration for quantitative MRI

Huizinga, W; Poot, DHJ; Guyader, JM; Klaassen, R; Coolen, BF; van Kranenburg, M; van Geuns, RJM; Uitterdijk, A; Niessen, WJ; More Authors

DOI

[10.1016/j.media.2015.12.004](https://doi.org/10.1016/j.media.2015.12.004)

Publication date

2016

Document Version

Final published version

Published in

Medical Image Analysis

Citation (APA)

Huizinga, W., Poot, DHJ., Guyader, JM., Klaassen, R., Coolen, BF., van Kranenburg, M., van Geuns, RJM., Uitterdijk, A., Niessen, WJ., & More Authors (2016). PCA-based groupwise image registration for quantitative MRI. *Medical Image Analysis*, 29, 65-78. <https://doi.org/10.1016/j.media.2015.12.004>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

**Green Open Access added to [TU Delft Institutional Repository](#)
as part of the Taverne amendment.**

More information about this copyright law amendment
can be found at <https://www.openaccess.nl>.

Otherwise as indicated in the copyright section:
the publisher is the copyright holder of this work and the
author uses the Dutch legislation to make this work public.



PCA-based groupwise image registration for quantitative MRI[☆]



W. Huizinga^{a,*}, D.H.J. Poot^{a,b}, J.-M. Guyader^a, R. Klaassen^c, B.F. Coolen^d, M. van Kranenburg^{e,f}, R.J.M. van Geuns^{e,f}, A. Uitterdijk^f, M. Polfliet^{g,h}, J. Vandemeulebroucke^{g,h}, A. Leemansⁱ, W.J. Niessen^{a,b}, S. Klein^a

^a Biomedical Imaging Group Rotterdam, Departments of Radiology & Medical Informatics, Erasmus MC, Rotterdam, The Netherlands

^b Quantitative Imaging Group, Department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

^c Department of Medical Oncology, Academic Medical Center, Amsterdam, The Netherlands

^d Department of Radiology, Academic Medical Center, Amsterdam, The Netherlands

^e Department of Radiology, Erasmus MC, Rotterdam, The Netherlands

^f Department of Cardiology, Erasmus MC, Rotterdam, The Netherlands

^g Vrije Universiteit Brussel, Department of Electronics and Informatics (ETRO), Brussels, Belgium

^h iMinds, Department of Medical IT, Ghent, Belgium

ⁱ Image Sciences Institute, University Medical Center Utrecht, The Netherlands

ARTICLE INFO

Article history:

Received 9 February 2015

Revised 10 December 2015

Accepted 10 December 2015

Available online 19 December 2015

Keywords:

Groupwise image registration

Quantitative MRI

Motion compensation

Principal component analysis

ABSTRACT

Quantitative magnetic resonance imaging (qMRI) is a technique for estimating quantitative tissue properties, such as the T_1 and T_2 relaxation times, apparent diffusion coefficient (ADC), and various perfusion measures. This estimation is achieved by acquiring multiple images with different acquisition parameters (or at multiple time points after injection of a contrast agent) and by fitting a qMRI signal model to the image intensities. Image registration is often necessary to compensate for misalignments due to subject motion and/or geometric distortions caused by the acquisition. However, large differences in image appearance make accurate image registration challenging. In this work, we propose a groupwise image registration method for compensating misalignment in qMRI. The groupwise formulation of the method eliminates the requirement of choosing a reference image, thus avoiding a registration bias. The method minimizes a cost function that is based on principal component analysis (PCA), exploiting the fact that intensity changes in qMRI can be described by a low-dimensional signal model, but not requiring knowledge on the specific acquisition model. The method was evaluated on 4D CT data of the lungs, and both real and synthetic images of five different qMRI applications: T_1 mapping in a porcine heart, combined T_1 and T_2 mapping in carotid arteries, ADC mapping in the abdomen, diffusion tensor mapping in the brain, and dynamic contrast-enhanced mapping in the abdomen. Each application is based on a different acquisition model. The method is compared to a mutual information-based pairwise registration method and four other state-of-the-art groupwise registration methods. Registration accuracy is evaluated in terms of the precision of the estimated qMRI parameters, overlap of segmented structures, distance between corresponding landmarks, and smoothness of the deformation. In all qMRI applications the proposed method performed better than or equally well as competing methods, while avoiding the need to choose a reference image. It is also shown that the results of the conventional pairwise approach do depend on the choice of this reference image. We therefore conclude that our groupwise registration method with a similarity measure based on PCA is the preferred technique for compensating misalignments in qMRI.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Quantitative magnetic resonance imaging (qMRI) is a technique that enables the estimation of tissue properties from a series of

images acquired with different imaging parameters or acquired at multiple time points after injection of a contrast agent. Because these tissue properties can be indicators of the biological state of the tissue and their change during disease, their precise and accurate estimation is important. Examples of such tissue properties are the relaxation parameters T_1 and T_2 , the mean diffusivity (MD), the apparent diffusion coefficient (ADC) and K^{trans} , a measure of capillary permeability. They are estimated by fitting a low-dimensional signal model (qMRI model) to the acquired MR images, see e.g. (Tofts, 2003). Typically five to over a hundred images

[☆] This paper was recommended for publication by Dr. Nicholas Ayache.

* Corresponding author. Tel.: +31 107044884.

E-mail address: wykehuizinga@gmail.com, w.huizinga@erasmusmc.nl (W. Huizinga).

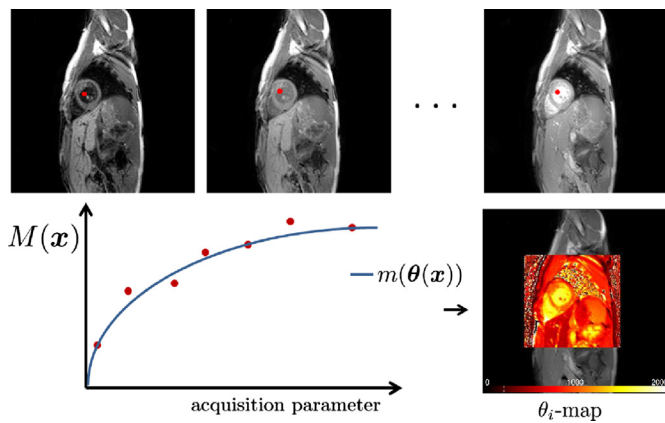


Fig. 1. A scheme showing the concept of qMRI. On the top there are three contrast-varying images of a heart. The red dots indicate one specific pixel location \mathbf{x} in each of the images and their intensities are plotted in the graph below the images. A model $m(\theta)$ is fitted to the pixel intensities, where θ contains the tissue properties of interest. The model is fitted for each pixel in the image so that a map of the element(s) in θ can be created. Such a map, showing T_1 , is shown on the right of the graph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are acquired, depending on the tissue properties of interest. Fig. 1 illustrates the concept of qMRI.

The fitting procedure assumes an anatomical correspondence between the images in the acquired series. However, due to patient motion and/or geometric distortions caused by the acquisition this correspondence can be lost, which may lead to erroneous parameter estimation, especially at tissue boundaries. Corrections during acquisition, such as gating or breath-holding, do not always give the desired effect and can significantly increase the acquisition time. Another solution is to align the images prior to fitting the qMRI model. This alignment can be achieved with image registration techniques. However, image registration for qMRI imposes two main challenges: firstly, the contrast is different for each of the images in the acquired series, complicating registration based on the image intensities. Secondly, often more than two images need to be registered in the context of qMRI. In this second case one can choose a pairwise registration approach in which all images are registered to a chosen reference image. To deal with contrast changes that occur in a series of qMRI images the pairwise approach is commonly used with a metric based on mutual information (MI), because this metric is robust against intensity changes in the images (Bron et al., 2013; Mangin et al., 2002; Guyader et al., 2014). However, a major disadvantage of this approach is that the choice of reference image will influence the result of the registration, which we will demonstrate in this paper. To circumvent the need to choose a reference image one can use a so-called groupwise registration approach. In such an approach all images are simultaneously registered to a mean space. Moreover, with this approach the information of all images is taken into account during the registration. This improves consistency over a pairwise registration approach, as shown in Metz et al. (2011) for groupwise registration of dynamic CT images.

We distinguish two categories of registration methods for qMRI data, discussed in the following two paragraphs: (a) model-based methods that use the qMRI model to register the images (Xue et al., 2012; Buonaccorsi et al., 2007; Andersson and Skare, 2002; Hallack et al., 2014; Bhushan et al., 2011), and (b) data-driven methods, which do not rely on the qMRI model (Hamy et al., 2014; Melbourne et al., 2007).

The model-based registration method proposed in Xue et al. (2012) for the registration of T_1 data, uses the fitted parameters of the qMRI model to generate reference images for all images in

the dataset. The registration is done in a pairwise fashion using a cross-correlation similarity metric. A similar approach was used by Buonaccorsi et al. (2007) for dynamic contrast-enhanced (DCE) images. Andersson and Skare (2002) and Hallack et al. (2014) directly minimize the residual error of the qMRI model fit. For the registration of DCE images (Bhushan et al., 2011) propose to maximize the joint posterior probability between the intensities estimated by the model and the true data. All these approaches eliminate the requirement to choose a reference image and are robust to the appearance differences in the image. However, these model-based methods assume that the images adhere exactly to the qMRI model, which is not always true in all structures that are present in the images due to noise and acquisition artifacts or when the model is too simple to represent the image intensities. For example, the DT model may not fully describe the signal in all voxels of the brain due to the presence of multiple fiber orientations, as shown by Behrens et al. (2007) and Jeurissen et al. (2013).

Data-driven methods for qMRI registration have been proposed by Hamy et al. (2014) and Melbourne et al. (2007). Hamy et al. (2014) proposed robust data decomposition registration for registering DCE data. With the assumption that the low rank components are free from local contrast changes or artefacts, they use a sparse and low-rank decomposition and register the low-rank components to the mean of all low-rank components in the current resolution. Melbourne et al. (2007) proposed a progressive principal component registration for DCE data. This method registers the series of images to an artificial series, which is generated from the first principal components of the original images. The registration is repeated and at each iteration a new series is generated and more principal components are added in the reconstruction.

Groupwise registration methods, not specifically developed for the registration of qMRI data, have been proposed by Liu et al. (2014), Miller et al. (2000), Metz et al. (2011), Wachinger and Navab (2013), and Bhatia et al. (2007). Liu et al. (2014) use a sparse/low-rank decomposition, similar to Hamy et al. (2014), to register brain images between subjects, with the assumption that the low-rank components are free from lesions and pathologies. Miller et al. (2000) proposed a method based on voxel-wise entropy. While this method is robust against intensity variation, qMRI acquisitions often have few measurements, making it difficult to estimate a well-defined probability density function, which is needed for entropy calculation. Metz et al. (2011) proposed a groupwise dissimilarity metric based on voxel-wise variance, implicitly assuming small intensity differences between the images. Due to the large intensity variations in qMRI it is not expected that this method is suitable for qMRI applications, however we investigate if this is indeed the case. Wachinger and Navab (2013) proposed a sum of accumulated pairwise estimates (APE), such as the sum of normalized correlation coefficients of all possible image pairs in the series, as a similarity metric. Even though this method has not yet been validated for qMRI data, it is designed to align images with different contrasts. Bhatia et al. (2007) proposed a groupwise extension of the pairwise mutual information registration approach, in which the sum of the mutual information between the voxel-wise mean of all images and each image is maximized. A downside of such a metric for qMRI is that contrast at edges may be reduced after computing the mean image, due to the large intensity differences among the images.

In this paper we propose a generic data-driven groupwise registration approach which by design is suitable for a wide range of qMRI applications without explicitly requiring the applicable qMRI model. In our approach we exploit the fact that in qMRI the intensity changes according to a low-dimensional acquisition model. When the images are not aligned, the complexity of the data is increased, i.e. the data can no longer be described by the acquisition model. We propose two dissimilarity metrics based on this

principle. These metrics use PCA to quantify the amount of misalignment in the qMRI series. Because of the groupwise formulation of the registration, the need to choose a reference image is eliminated. We evaluated the generic applicability of the method on five different challenging applications of qMRI: T_1 mapping in a porcine heart, combined T_1 and T_2 mapping in carotid arteries, ADC mapping in the abdomen, diffusion tensor mapping in the brain and dynamic contrast-enhanced mapping in the abdomen. As the groupwise approaches of Wachinger and Navab (2013), Bhatia et al. (2007), and Metz et al. (2011) have potential for qMRI data we use them as reference methods. The model-based approach by e.g. (Hallack et al., 2014) is implemented for T_1 mapping and will be used as a reference method for this application. Finally we also compare our method to the commonly used MI-based pairwise registration approach.

Preliminary results of this registration approach were presented at the MICCAI workshop on Computational Diffusion MRI and Brain Connectivity, Japan 2013 (Huizinga et al., 2013), and at the Workshop on Biomedical Image Registration, London 2014 (Huizinga et al., 2014). In this work, we considerably expand previous studies, and include the following new elements:

- a novel version of the dissimilarity metric, free of any user-defined parameters;
- an extended range of qMRI applications on which the methods are evaluated;
- quantitative evaluation of registration accuracy based on manually annotated landmarks, overlap of manual segmentations, and smoothness of the deformation field; and
- comparison to a larger variety of reference registration methods.

2. Methods

2.1. Quantitative magnetic resonance imaging

In a qMRI examination multiple images are acquired in the same subject. Let $M_g(\mathbf{x})$ for $g \in \{1 \dots G\}$ be a series of G images and \mathbf{x} a 2D or 3D spatial coordinate. The intensity at \mathbf{x} for each image M_g is predicted by a low-dimensional qMRI model m_g :

$$M_g(\mathbf{x}) = m_g(\boldsymbol{\theta}(\mathbf{x})) + \epsilon(\mathbf{x}), \quad (1)$$

where $\boldsymbol{\theta}$ is a Γ -dimensional vector with tissue properties and ϵ is the noise at coordinate \mathbf{x} . In our applications, the number of tissue properties Γ ranges from three to seven and $\Gamma < G$, which is why we call the qMRI model low-dimensional. Each qMRI application follows a different model m_g . An example of such a qMRI model m_g is the modified Look-Locker inversion recovery model proposed by Messroghli et al. (2004):

$$m_g(\boldsymbol{\theta}) = |A(1 - Be^{-Tg/T_1^*})|, \quad (2)$$

where $\boldsymbol{\theta} = (A, B, T_1^*)$ and Tg the inversion time for image M_g . The parameter of interest, T_1 , is calculated using $T_1 = T_1^*(B - 1)$. For this specific model $\Gamma = 3$, as $\boldsymbol{\theta}$ contains three parameters. The functions m_g that are used for the experiments in this paper are presented in Section 3.

2.2. Registration frameworks: pairwise and groupwise

Prior to estimating $\boldsymbol{\theta}$ the images M_g need to be registered. Due to motion and/or geometric distortions caused by the acquisition correspondence between the images is lost, i.e. a spatial coordinate \mathbf{x} does not correspond to the same anatomical location in each of the images. In the pairwise registration approach, one reference image $M_R(\mathbf{x})$ is chosen and all other images $M_g(\mathbf{x})$ for $g \neq R$ are registered to $M_R(\mathbf{x})$. In our registration method, the transformation

is modeled by a set of transform parameters $\boldsymbol{\mu}$. For each image $M_g(\mathbf{x})$ there is a transformation $T_g(\mathbf{x}; \boldsymbol{\mu}_g)$. The pairwise registration is formulated as the minimization of a dissimilarity metric \mathcal{D} with respect to $\boldsymbol{\mu}_g$:

$$\hat{\boldsymbol{\mu}}_g = \arg \min_{\boldsymbol{\mu}_g} \mathcal{D}(\boldsymbol{\mu}_g), \quad (3)$$

which is repeated for all $g \neq R$. \mathcal{D} measures the dissimilarity of $M_R(\mathbf{x})$ and $M_g(T_g(\mathbf{x}; \boldsymbol{\mu}_g))$.

In the groupwise registration framework the images $M_g(\mathbf{x})$ for all g are registered simultaneously to a mean space. We formulate groupwise registration as the minimization of a dissimilarity metric \mathcal{D} with respect to $\boldsymbol{\mu}$:

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} \mathcal{D}(\boldsymbol{\mu}), \quad (4)$$

where $\boldsymbol{\mu}$ is a vector containing all $\boldsymbol{\mu}_g$. Here, \mathcal{D} measures the dissimilarity of all transformed images $M_g(T_g(\mathbf{x}; \boldsymbol{\mu}_g))$ with respect to each other. The parameters $\boldsymbol{\mu}_g$ are simultaneously optimized for all g .

2.3. Proposed dissimilarity metrics

We present two novel groupwise dissimilarity metrics. Let the images M_g be represented as columns of an $N \times G$ matrix \mathbf{M} , where N is the number of voxels in one image M_g . A row of \mathbf{M} can be considered as a datapoint in a G -dimensional space. Note that when M_g is noise-free for all g , these datapoints lie in a – possibly non-linear – Γ -dimensional subspace, where Γ is the number of free qMRI model parameters. Fig. 2a shows an intensity scatter plot of three images generated using a one-dimensional ($\Gamma = 1$) non-linear model: $m_g(\boldsymbol{\theta}) = 1 - e^{-a\boldsymbol{\theta}}$, where $\boldsymbol{\theta} = a$. We see that the points lie on a curved line, i.e. a non-linear Γ -dimensional subspace. Fig. 2b illustrates the effect of a small misalignment: the intensity scatter plot becomes more dispersed.

The correlation matrix of the data points in \mathbf{M} is defined as:

$$\mathbf{K} = \frac{1}{N-1} \boldsymbol{\Sigma}^{-1} (\mathbf{M} - \overline{\mathbf{M}})^T (\mathbf{M} - \overline{\mathbf{M}}) \boldsymbol{\Sigma}^{-1}, \quad (5)$$

where $\boldsymbol{\Sigma}$ is a diagonal matrix with the standard deviations of each column of \mathbf{M} as diagonal elements and $\overline{\mathbf{M}}$ is a matrix with in each element of that column the column-wise average of \mathbf{M} . The dimension of the subspace can be estimated by an eigenvalue decomposition of \mathbf{K} i.e., by a PCA. The key idea behind the proposed dissimilarity metric is that, when motion is present in the images, the data no longer adheres to the presumed qMRI model and the eigenvalue spectrum of \mathbf{K} changes. We choose to perform PCA on the correlation matrix \mathbf{K} , instead of the covariance matrix, to be insensitive to arbitrary intensity scaling between images. Fig. 3a shows the eigenvalue spectrum of \mathbf{K} for an aligned set of synthetic, noise-free images, created with the qMRI model of Eq. (2), and the deformed set of images (Fig. 3b). We can observe that once the images are deformed, the higher eigenvalues increase. Our method aims to transform the images M_g such that the eigenvalue spectrum of \mathbf{K} approaches the spectrum of an aligned set of images.

Let λ_j be the j th eigenvalue of \mathbf{K} , with $\lambda_j > \lambda_{j+1}$, i.e. the eigenvalues with a lower index have a higher value. The first dissimilarity metric we propose is defined as the difference between the sum of all eigenvalues (which is equal to the trace of \mathbf{K} , which is equal to the constant G) and the sum of the L highest eigenvalues:

$$\mathcal{D}_{\text{PCA}}(\boldsymbol{\mu}) = \sum_{j=1}^G K_{jj}(\boldsymbol{\mu}) - \sum_{j=1}^L \lambda_j(\boldsymbol{\mu}) = G - \sum_{j=1}^L \lambda_j(\boldsymbol{\mu}), \quad (6)$$

where the dependence on $\boldsymbol{\mu}$ has been made explicit to clarify that \mathbf{K} (and thus λ_j) is computed based on the deformed images

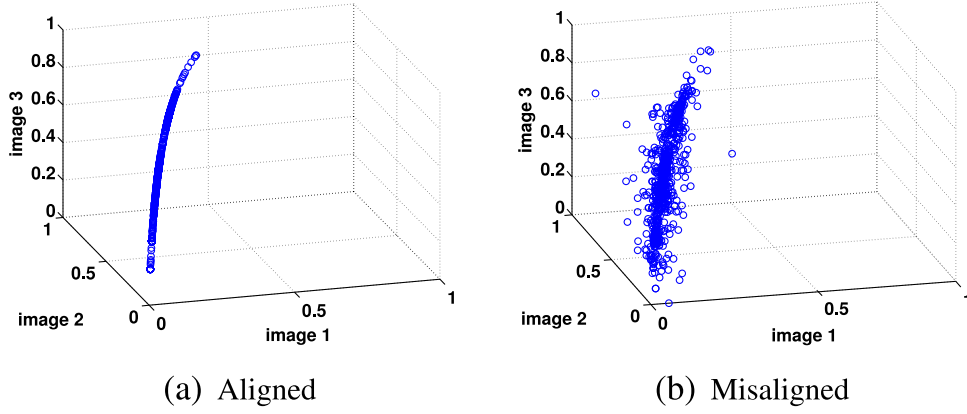


Fig. 2. Scatter plots of three aligned images (a) and three misaligned images (b) that are generated with a one-dimensional non-linear model.

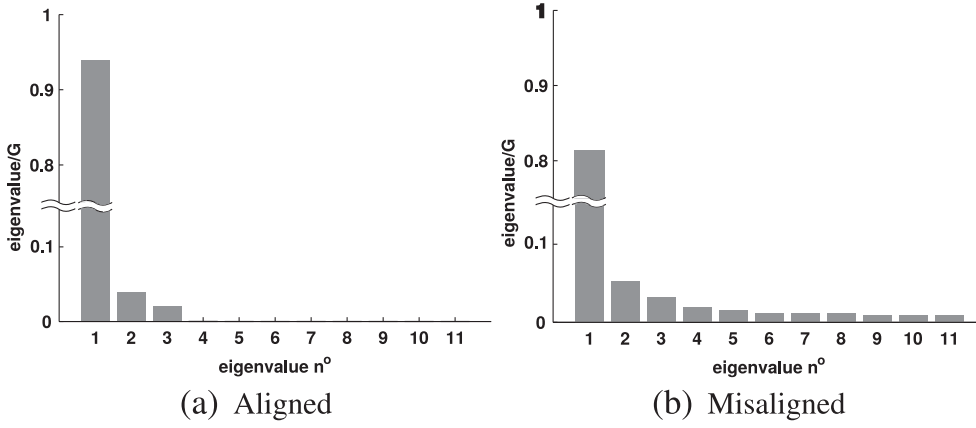


Fig. 3. Eigenvalue spectra of aligned (a) and misaligned (b) set of images. Note that $\Gamma = 3$, so we expect three dominant eigenvalues in the aligned case. Note that in (a) eigenvalues 4–11 are not exactly zero but very small ($< 10^{-3}$) and therefore not visible in the plot.

$M_g(\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_g))$. The constant $1 \leq L \leq G$ is a user-defined parameter. For different qMRI models, a different value of L must be chosen. A good initial guess is $L = \Gamma$, assuming that the non-linear Γ -dimensional subspace can be approximated by a Γ -dimensional hyperplane. Peng et al. (2010) also assume that well-aligned images are linearly correlated but they use a sparse and low-rank decomposition to directly minimize the rank of \mathbf{M} with a trade-off parameter for sparsity. Since the rank of \mathbf{M} is equal to the number of non-zero eigenvalues of \mathbf{K} , the methods are related, but in our method the rank is not minimized but can be controlled by L , which is preferable in qMRI, since the rank depends on the qMRI model. The dimension of the subspace in qMRI may be less than Γ when the parameters in $\boldsymbol{\theta}$ are correlated, or higher than Γ due to the non-linearity of the acquisition models. That is why we propose a second dissimilarity metric that circumvents the need to choose L :

$$\mathcal{D}_{\text{PCA2}}(\boldsymbol{\mu}) = \sum_{j=1}^G j \lambda_j(\boldsymbol{\mu}). \quad (7)$$

In $\mathcal{D}_{\text{PCA2}}$ the eigenvalues with the highest values (and the lowest indices) have the lowest weight. Given that $\lambda_1 + \lambda_2 + \dots + \lambda_j = G$, $\mathcal{D}_{\text{PCA2}}$ promotes that as much variance as possible is explained by a few large eigenvectors.

2.3.1. Metric derivatives

Minimization of the dissimilarity metric with gradient based optimizers requires the derivative of the metric with respect to $\boldsymbol{\mu}$. To differentiate Eqs. (6) and (7) with respect to $\boldsymbol{\mu}$ we use the ap-

proach of van der Aa et al. (2007):

$$\frac{\partial \mathcal{D}_{\text{PCA}}}{\partial \boldsymbol{\mu}} = - \sum_{j=1}^L \frac{\partial \lambda_j}{\partial \boldsymbol{\mu}} = - \sum_{j=1}^L \mathbf{v}_j^T \frac{\partial \mathbf{K}}{\partial \boldsymbol{\mu}} \mathbf{v}_j, \quad (8)$$

and

$$\frac{\partial \mathcal{D}_{\text{PCA2}}}{\partial \boldsymbol{\mu}} = \sum_{j=1}^G j \frac{\partial \lambda_j}{\partial \boldsymbol{\mu}} = \sum_{j=1}^G j \mathbf{v}_j^T \frac{\partial \mathbf{K}}{\partial \boldsymbol{\mu}} \mathbf{v}_j, \quad (9)$$

where \mathbf{v}_j^T is the j th eigenvector of \mathbf{K} . Similarly to van der Aa et al. (2007) we ignore the unlikely repetition of eigenvalues. For repeated eigenvalues, linear combinations of eigenvectors are also an eigenvector, which invalidates the above. Using Eqs. (5) and (8) we obtain the derivative of \mathcal{D}_{PCA} with respect to an element μ_p :

$$\begin{aligned} \frac{\partial \mathcal{D}_{\text{PCA}}}{\partial \mu_p} = & - \frac{2}{N-1} \sum_{i=1}^L \left[\mathbf{v}_i^T \boldsymbol{\Sigma}^{-1} (\mathbf{M} - \bar{\mathbf{M}})^T \left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \bar{\mathbf{M}}}{\partial \mu_p} \right) \boldsymbol{\Sigma}^{-1} \mathbf{v}_i \right. \\ & \left. + \mathbf{v}_i^T \boldsymbol{\Sigma}^{-1} (\mathbf{M} - \bar{\mathbf{M}})^T (\mathbf{M} - \bar{\mathbf{M}}) \frac{\boldsymbol{\Sigma}^{-1}}{\partial \mu_p} \mathbf{v}_i \right]. \end{aligned} \quad (10)$$

The above expression is obtained after simplifications and using the fact that

$$\mathbf{v}^T \mathbf{B}^T \mathbf{E} \mathbf{v} = \mathbf{v}^T \mathbf{E}^T \mathbf{B} \mathbf{v} \quad (11)$$

for two matrices \mathbf{B} and \mathbf{E} and vector \mathbf{v} . The derivative of $\boldsymbol{\Sigma}^{-1}$ with respect to μ_p is equal to

$$\frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \mu_p} = - \frac{\boldsymbol{\Sigma}^{-3}}{N-1} \text{diag} \left[(\mathbf{M} - \bar{\mathbf{M}})^T \left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \bar{\mathbf{M}}}{\partial \mu_p} \right) \right] \quad (12)$$

and $\partial \mathbf{M} / \partial \mu_p$ and $\partial \bar{\mathbf{M}} / \partial \mu_p$ are computed using

$$\frac{\partial M_g(\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_g))}{\partial \mu_p} = \left(\frac{\partial M_g}{\partial \mathbf{x}} \right)^T \bigg|_{\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_g)} \left(\frac{\partial \mathbf{T}_g}{\partial \mu_p} \right) \bigg|_{(\mathbf{x}; \boldsymbol{\mu}_g)}. \quad (13)$$

The same steps are used to calculate $\partial \mathcal{D}_{\text{PCA2}} / \partial \mu_p$. It was verified that the value of $\partial \bar{\mathbf{M}} / \partial \mu_p$, the derivative of the mean intensities, was negligibly small and could therefore be ignored in the actual implementation.

2.4. Transformation models

In our experiments (see Section 4) we use two different transformation models: a non-rigid transformation model in which deformations are modeled by cubic B-splines, proposed by Rueckert et al. (1999), and an affine transformation model. Similar to Wachinger and Navab (2013) we used an exponential mapping of the affine matrix for parameterization.

2.5. Optimization

An adaptive stochastic gradient descent (ASGD) optimization method, proposed by Klein et al. (2009), was used, which randomly samples positions in image space at each iteration in order to reduce computation time. Sampling was done off the voxel grid, which was shown to be necessary to reduce interpolation artifacts, as proposed by Klein et al. (2010). A multi-resolution strategy was used. In such a strategy the image is Gaussian-blurred with a certain standard deviation and at each level the standard deviation is decreased, such that large deformations are corrected first and finer deformations are corrected at higher levels. When a B-spline transformation was used, the control point spacing was also decreased at each resolution level. The number of random samples, the number of resolution levels, and the number of iterations per resolution level are user-defined parameters. Linear interpolation was used to interpolate the images during registration, to limit computation time. Cubic B-spline interpolation was used to produce the final motion-compensated images.

For the groupwise framework, the average deformation of the images was constrained to be zero by the approach of Balci et al. (2007): the average derivative of the dissimilarity metric with respect to its parameters $\boldsymbol{\mu}_g$ is subtracted from each derivative to $\boldsymbol{\mu}_g$, i.e. the derivatives are centered to zero mean.

$$\frac{\partial \mathcal{D}^*}{\partial \boldsymbol{\mu}_g} = \frac{\partial \mathcal{D}}{\partial \boldsymbol{\mu}_g} - \frac{1}{G} \sum_{g'} \frac{\partial \mathcal{D}}{\partial \boldsymbol{\mu}_{g'}}, \quad (14)$$

where $\partial \mathcal{D}^* / \partial \boldsymbol{\mu}_g$ is the zero-centered derivative.

2.6. Reference dissimilarity metrics

We compared the proposed method with three other methods: a pairwise method that uses a MI-based dissimilarity metric, and four groupwise methods. The pairwise MI dissimilarity used was proposed by Thévenaz and Unser (2000). The number of histogram bins used to calculate the probability functions is set to 32 in all experiments.

Wachinger and Navab (2013) proposed accumulated pairwise estimates (APE) as a family of metrics. One of the metrics they propose is the sum of squared normalized correlation coefficients. This can be written as the sum of the squared elements of the correlation matrix \mathbf{K} . We implemented this metric as follows:

$$\mathcal{D}_{\text{APE}}(\boldsymbol{\mu}) = 1 - \frac{1}{G} \sqrt{\sum_i \sum_j K_{ij}(\boldsymbol{\mu})^2}. \quad (15)$$

Metz et al. (2011) proposed the sum of the variances, assuming no intensity changes between images. The metric is defined as:

$$\mathcal{D}_{\text{VAR}}(\boldsymbol{\mu}) = \frac{1}{NG} \sum_{i=1}^N \sum_{g=1}^G [M_g(\mathbf{T}_g(\mathbf{x}_i; \boldsymbol{\mu}_g)) - \tilde{M}(\mathbf{x}_i; \boldsymbol{\mu})]^2, \quad (16)$$

where $\tilde{M}(\mathbf{x}_i; \boldsymbol{\mu}) = \frac{1}{G} \sum_g M_g(\mathbf{T}_g(\mathbf{x}_i; \boldsymbol{\mu}_g))$.

The groupwise mutual information based method proposed by Bhatia et al. (2007) was implemented as:

$$\mathcal{D}_{\text{G-MI}}(\boldsymbol{\mu}) = - \sum_{g=1}^G [H(\tilde{M}(\cdot; \boldsymbol{\mu})) + H(M_g(\mathbf{T}_g(\cdot; \boldsymbol{\mu}_g))) - H(\tilde{M}(\cdot; \boldsymbol{\mu}), M_g(\mathbf{T}_g(\cdot; \boldsymbol{\mu}_g)))]], \quad (17)$$

where $H(\cdot)$ represents the marginal entropy function and $H(\cdot, \cdot)$ the joint entropy function. For this method the number of histogram bins was also set to 32 in all experiments.

Andersson and Skare (2002) and Hallack et al. (2014) proposed to minimize the residual error of the model fit:

$$\mathcal{D}_{\text{T}_1}(\boldsymbol{\mu}) = \frac{1}{N} \sum_{i=1}^N \min_{\boldsymbol{\theta}} \sum_{g=1}^G (M_g(\mathbf{T}_g(\mathbf{x}_i; \boldsymbol{\mu}_g)) - m_g(\boldsymbol{\theta}(\mathbf{x}_i)))^2. \quad (18)$$

We implemented this method for the model given by Eq. (2).

2.7. qMRI fitting method

Having registered the images M_g , the qMRI parameters $\boldsymbol{\theta}$ can be estimated by fitting the model m_g at each voxel. The qMRI model is fitted using a maximum-likelihood (ML) estimator that takes into account the Rician noise of MRI data. The estimation procedure is defined as (Poot and Klein, 2015):

$$\hat{\boldsymbol{\theta}}, \hat{\varepsilon} = \arg \max_{\boldsymbol{\theta}, \varepsilon} (\ln p(\mathbf{M} | m(\boldsymbol{\theta}), \varepsilon)), \quad (19)$$

where p is the probability density function of the Rician distribution, \mathbf{M} are the images, m is the qMRI model with parameters $\boldsymbol{\theta}$, and ε is the noise level, which is estimated per voxel. To get an accurate estimate of the noise level, the bias caused by the reduced degrees of freedom of the residual is corrected according to Poot and Klein (2015):

$$\hat{\varepsilon}^{\text{modif}} = \hat{\varepsilon} \sqrt{G / (G - \Gamma)}. \quad (20)$$

When $G - \Gamma$ is low, there are not enough degrees of freedom left in a voxel to estimate the noise level precisely. In these cases we therefore regularized the noise level with a log prior $S(\varepsilon)$ promoting a spatially smooth noise level field (Poot and Klein, 2015):

$$\hat{\boldsymbol{\theta}}, \hat{\varepsilon} = \arg \max_{\boldsymbol{\theta}, \varepsilon} (\ln p(\mathbf{M} | m(\boldsymbol{\theta}), \varepsilon) - S(\varepsilon)). \quad (21)$$

The qMRI fitting method is publicly available at fitMRI.bigr.nl.

3. Data

We evaluated the registration methods on a purely synthetic dataset, real CT lung datasets, five real qMRI datasets and five synthetic datasets derived from the real data. The subsections below describe all datasets.

3.1. SYNTH-MODEL

A purely synthetic image was generated to demonstrate the advantages of the proposed PCA-based methods when registering qMRI data. The generated images are linear combinations of five basis images, representing five independent ‘tissue’ properties in a qMRI experiment:

$$m_g(\boldsymbol{\theta}) = \mathbf{w}_g^T \boldsymbol{\theta}, \quad (22)$$

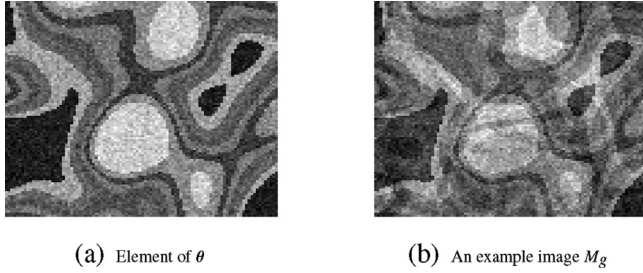


Fig. 4. (a) One of the five ‘tissue’ properties, i.e. an element of θ , (b) one of the G weighted images.

where θ contains five ‘tissue’ properties. Each property is spatially correlated and contains large and small structures. An example ‘tissue’ property map is shown in Fig. 4a. The weights are optimized such that the sum of the pairwise mutual information of the images is minimized. The intrinsic dimension of the synthetic data can however be perfectly resolved by a principal component analysis. We expect that especially the proposed PCA-based methods can successfully register the images. An example of an image M_g is shown in Fig. 4b. The code to generate the SYNTH-MODEL is included as supplementary material.

3.2. Synthetic qMRI

To evaluate how the methods perform with different qMRI models, in a setting with known ground truth, we created synthetic data based on real qMRI data. To save computation time, we extracted a 2D slice from a single subject, for each of the qMRI applications studied in this paper (see Sections 3.4–3.8). These slices were fitted and the obtained θ and the acquisition parameters were used to simulate the contrast with the qMRI model that belongs to the data. Rician noise was added to the synthetic qMRI data such that the signal-to-noise (SNR) ratio was equal to 10, where the SNR is defined as the mean signal intensity over all images divided by the σ -parameter of the Rician distribution. The pixel spacings in the first two dimensions and G for the synthetic datasets are equal to their values in the corresponding real datasets. All synthetic qMRI datasets are written with a prefix ‘s-’.

3.3. CT-LUNG

Ten 4D CT lung datasets from the DIR-LAB database (Castillo et al., 2009) were used to demonstrate the results of the methods on datasets with only slight intensity changes. Note that this can be seen as a special case of quantitative imaging where a 1-parameter model ($\Gamma = 1$) describes the intensity with a constant:

$$m_g(\theta) = c, \quad (23)$$

with $\theta = c$, which is trivially estimated by computing the mean over M_g . The voxel size was approximately $1.1 \times 1.1 \times 2.5$ mm³. The size of the first five images was around $256 \times 256 \times 100$ and of the last five images around $512 \times 512 \times 128$. For all datasets $G = 10$.

3.4. T1MOLLI-HEART

Quantification of T_1 relaxation is important for the characterization of myocardial tissue, which is useful to assess both ischemic and nonischemic heart muscle diseases (Xue et al., 2012). A popular cardiac T_1 mapping method uses the modified Look-Locker inversion recovery (MOLLI) sequence, described in Messroghli et al. (2004). The qMRI model for MOLLI is given by Eq. (2). Nine T1MOLLI-HEART datasets from porcine hearts with a transmural myocardial infarction of the lateral wall were acquired using single-slice acquisition. For each subject $G = 11$ two-dimensional images of size 512×512 were acquired. The pixel size was around 0.7×0.7 mm². Inversion times T_{I_g} ranged from 82 to 3866 ms. Fig. 5a shows an example image.

3.5. T1VFA-CAROTID

Quantifying the relaxation parameters T_1 and T_2 can help to assess the composition of atherosclerotic plaque in the carotid arteries (Coolen et al., 2015). Carotid plaque composition has been shown to be associated with the occurrence of cerebrovascular events (Takaya et al., 2006). Coolen et al. (2015) proposed a quantitative carotid plaque imaging method based on a 3D improved motion-sensitized driven-equilibrium prepared black-blood turbo field echo sequence. The qMRI model of this variable flip-angle approach is given by:

$$m_g(\theta) = \left| A \sin(\alpha_g) \frac{1 - e^{-TR/T_1}}{1 - \cos(\alpha_g)e^{-TR/T_1}} e^{-TE_g/T_2} \right|, \quad (24)$$

with $\theta = (A, T_1, T_2)$. α_g the flip-angle and TE_g the T_2 preparation time per image M_g and TR the repetition time (Coolen et al., 2015). Values for the acquisition parameters were $\alpha_g \in [6, 4, 15, 15, 15]$ degrees, $TE_g \in [11.5, 11.5, 11.5, 26, 45]$ ms and $TR = 10$ ms. Eight T1VFA-CAROTID datasets were acquired. For each subject, $G = 5$ three-dimensional images of size $224 \times 223 \times 36$ with voxel size $0.7 \times 0.7 \times 0.7$ mm³ were acquired. One of these five images was an anatomical reference scan. The other four scans were used for the combined T_1 and T_2 fitting. Although this method gave accurate T_2 mapping results when analyzed in a ROI (Coolen et al., 2015), the acquisitions were not optimal yet for true voxelwise T_2 fitting, possibly due to the rather low maximum TE_g . In our experiments, it proved difficult to perform accurate voxelwise T_2 fits on the current data, regardless of the registration strategy. Therefore, we only report evaluations of T_1 values. More acquisition details can be found in Coolen et al. (2015) and an example image is shown in Fig. 5b. Please note that although the T1MOLLI-HEART and the T1VFA-CAROTID experiments both result in a T_1 estimate, a different model m_g is underlying the acquisition.

3.6. ADC-ABDOMEN

The ADC provides quantitative information related to the diffusion of water molecules in biological tissues. Pathophysiological processes such as cancer are known to have an impact on cell density, which translates into different diffusion properties and therefore the ADC is an interesting biomarker to assess these diseases

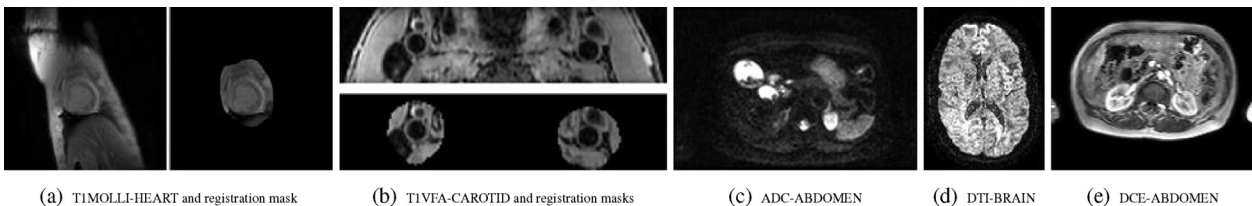


Fig. 5. Examples of real datasets and, if used, corresponding registration masks.

Table 1
Acquisition details of the DTI-BRAIN data.

Dataset	# $b = 0$ s/mm ²	# DWIs	G	b -value [s/mm ²]	Voxel size [mm ³]
1	10	60	70	700	$2.0 \times 2.0 \times 2.0$
2	1	60	61	1200	$1.75 \times 1.75 \times 2.0$
3	1	32	33	800	$1.75 \times 1.75 \times 2.0$
4	1	32	33	800	$1.75 \times 1.75 \times 2.0$
5	1	45	46	1200	$1.72 \times 1.72 \times 2.0$

(Guyader et al., 2014). In this work, we evaluate ADC mapping in the abdomen. The ADC acquisition model is the following:

$$m_g(\theta) = B_0 e^{-b_g \mathbf{u}_g^T \mathbf{D} \mathbf{u}_g}, \quad (25)$$

with $\theta = (B_0, D_{11}, D_{22}, D_{33})$ (Guyader et al., 2014). The vector in the direction of the applied gradient is given by \mathbf{u}_g , \mathbf{D} is a 3×3 symmetric diffusion tensor and b_g is the so-called b -value. The ADC is given by

$$\text{ADC} = \text{trace}(\mathbf{D})/3, \quad (26)$$

hence only the diagonal elements of \mathbf{D} are required. For acquiring the ADC data, an interleaved multi-slice diffusion weighted acquisition sequence was used. Five ADC-ABDOMEN datasets were acquired. For each subject $G = 19$ images of size $256 \times 224 \times 40$ were acquired with a voxel size of $1.48 \times 1.48 \times 5$ mm³ with b -values of 0, 100, 150, 200, 300, 500 and 900 s/mm². Diffusion weighting was applied in three orthogonal directions aligned with the read, phase and slice directions. More details about the acquisition settings of these datasets can be found in Guyader et al. (2014). Fig. 5c shows an example of a diffusion weighted image.

3.7. DTI-BRAIN

Measuring the diffusion tensor of water in tissues can provide parameters that help to characterize tissue composition, the physical properties of tissue constituents, tissue microstructure and its architectural organization. It is often used to map the white matter tissue structure (Basser and Jones, 2002; Jones and Leemans, 2011; Tournier et al., 2011). The qMRI model belonging to DT imaging equals:

$$m_g(\theta) = B_0 e^{-b \mathbf{u}_g^T \mathbf{D} \mathbf{u}_g}, \quad (27)$$

with $\theta = (B_0, D_{11}, D_{12}, D_{13}, D_{22}, D_{23}, D_{33})$ for the DT mapping. The vector in the direction of the applied gradient is given by \mathbf{u}_g , \mathbf{D} is a 3×3 symmetric diffusion tensor and b is the so-called b -value. In this paper we calculate the mean diffusivity (MD), which is given by:

$$\text{MD} = \text{trace}(\mathbf{D})/3. \quad (28)$$

Five DTI-BRAIN datasets were acquired with a diffusion weighted echo planar imaging sequence. Details about the five datasets are provided in Table 1 and details of the acquisition settings can be found in, respectively, Leemans et al. (2006), de Geeter et al. (2012), Wang et al. (2012), van der Aa et al. (2011), and Reijmer et al. (2012). Fig. 5d shows an example of a diffusion weighted image.

3.8. DCE-ABDOMEN

DCE imaging is an established method for assessing microvascular changes associated with disease in tissues. Examples of tissue diseases where DCE imaging is used are cancer, inflammatory conditions, cerebral ischemia, and cardiac ischemia (Parker et al., 2006). In DCE imaging the contrast in the images varies due to the injected contrast agent. The acquisition parameters T_R and α are not varied. In this work, we study DCE imaging

in the abdomen. The qMRI model belonging to the DCE acquisition equals:

$$m_g(\theta) = S_0 \frac{(1 - \cos \alpha) e^{-T_R/T_{10}}}{1 - e^{-T_R/(T_{10} + C(\theta, t_g) T_1)}}, \quad (29)$$

with $C(\theta, t_g)$ the contrast agent concentration in the tissue, described by the extended Tofts model (Tofts et al., 1999):

$$C(\theta, t_g) = C_p(t_g - \Delta) v_p + C_p(t_g - \Delta) * K^{\text{trans}} e^{-\frac{K^{\text{trans}}}{v_e} t_g}, \quad (30)$$

where $\theta = (K^{\text{trans}}, v_e, v_p, \Delta)$, $C_p(t)$ is the blood plasma concentration described by the general arterial input function by Parker et al. (2006), and $*$ denotes the convolution operator. The time delay Δ was estimated separately, prior to fitting θ . K^{trans} is a measure of capillary permeability, v_p is the blood volume fraction, v_e is the volume fraction of extracellular, extravascular space and S_0 is the signal without contrast agent. Known variables were the acquisition time t_g , the flip-angle α , the repetition time T_R , the longitudinal relaxation parameter without contrast T_{10} , and the longitudinal relaxation parameter of the contrast agent T_1 . The values for the known acquisition parameters were $\alpha = 20$ degrees, $T_R = 3.19$ ms, $T_{10} = 725$ ms, and $T_1 = 200$ ms. Five DCE-ABDOMEN datasets were acquired with a spoiled gradient echo sequence. Per subject $G = 160$ images of size $160 \times 160 \times 30$ were acquired with a voxel size of $2.5 \times 2.5 \times 2.5$ mm³. More acquisition details can be found in Klaassen et al. (2014). An example image is shown in Fig. 5e.

4. Experiments

The proposed registration methods were implemented in the publicly available registration package Elastix (Klein et al., 2010). All datasets were registered with six intensity-based dissimilarity metrics and, in addition, both the real and synthetic T1MOLLI-HEART datasets were registered with a model-based metric.

1. \mathcal{D}_{MI} (pairwise)
2. \mathcal{D}_{VAR} (groupwise)
3. \mathcal{D}_{APE} (groupwise)
4. $\mathcal{D}_{\text{G-MI}}$ (groupwise)
5. \mathcal{D}_{T_1} (groupwise) (T1MOLLI model)
6. \mathcal{D}_{PCA} (groupwise)
7. $\mathcal{D}_{\text{PCA2}}$ (groupwise)

To account for deformations caused by heart-pulsations and breathing we used a B-spline transformation model for the CT-LUNG data, T1MOLLI-HEART, T1VFA-CAROTID, ADC-ABDOMEN and DCE-ABDOMEN experiments. For ADC-ABDOMEN and DCE-ABDOMEN datasets, results are also reported for an affine transformation. To account for deformations caused by head motion and eddy current distortions we used an affine transformation model for application DTI-BRAIN. Two types of experiments were done:

1. The synthetic images were deformed by applying a known transformation and registered with all methods;
2. The real CT-LUNG and qMRI data were registered with all methods.

The parameters of the transformation applied to the SYNTH-MODEL data were drawn from a normal distribution, such that the initial deformation was approximately six pixels for the affine transformation and three pixels for the B-spline transformation. The parameters of the transformations applied to the synthetic qMRI datasets were drawn from normal distributions, such that the initial deformation was approximately one to two pixels. For the cases with non-rigid deformations, the control point spacing σ for the initial and estimated B-spline transformation was equal in all dimensions: $\sigma = [\nu, \nu]$, with the following values for ν :

Table 2

Summary of the settings used in the image registration experiments, for all datasets.

qMRI application	G	Γ	L	ν [mm]	Reference image	Registration mask	Evaluation ROI
SYNTH-MODEL	20	5	5	16	First image of the series	–	–
CT-LUNG	10	1	1	6, 13 and 20	First image of the series	Lungs	Red
T1MOLLI-HEART	11	3	3	32, 64 and 128	Slice 1, 4, 7 and 11	See Fig. 5a	Myocardium
T1VFA-CAROTID	5	3	1	8, 16 and 32	Anatomical scan	See Fig. 5b	Carotid artery wall
ADC-ABDOMEN	19	4	4	32, 64 and 128	$b = 0$ s/mm ² image	–	Spleen
DTI-BRAIN	See Table 1	7	7	–	$b = 0$ s/mm ² image	–	Parenchyma
DCE-ABDOMEN	160	4	4	32, 64 and 128	Non-contrast weighted image	–	Pancreas

- SYNTH-MODEL: $\nu = 16$ mm
- s-T1MOLLI: $\nu = 64$ mm
- s-T1VFA: $\nu = 16$ mm
- s-ADC: $\nu = 64$ mm
- s-DCE: $\nu = 64$ mm

The values of ν were chosen such that the simulated deformations were realistic. For the T1MOLLI-HEART, T1VFA-CAROTID experiments and their synthetic datasets we used registration masks, loosely drawn around the heart and the carotid arteries, to reduce the influence of surrounding organs. Examples are shown in Fig. 5a and b. For T1VFA-CAROTID, registration was performed separately for the left and right carotid. Lung masks were used in the CT-LUNG experiment, similar to Metz et al. (2011). For the cases with non-rigid deformations, different values for the control point spacing σ for the B-spline transformation were used. The values for ν are provided in Table 2.

Before further processing of the ADC-ABDOMEN data, within-image motion artifacts due to interleaved acquisition were corrected with the methods of Guyader et al. (2014).

For all registrations we used the following settings:

- two resolutions,
- 1000 iterations per resolution,
- 2048 random coordinate samples per resolution.

In the pairwise registration framework, a reference image must be chosen. The images that are used as reference for each application are shown in Table 2. For the T1MOLLI-HEART data there is no obvious reference image, so in this case we used multiple images of each dataset as a reference image, which enables us to evaluate the effect of the reference image on the results.

The \mathcal{D}_{PCA} method requires the user to set the parameter L . This parameter is shown in Table 2 for all qMRI applications.

4.1. Evaluation measures

As the synthetic data are aligned by design, they can be used as a ground truth to evaluate registration accuracy. A known transformation $\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_{\text{init},g})$ was applied to the data and corrected using the registration methods. The accuracy of the registration method can be evaluated by calculating the resulting residual deformation field $\mathbf{d}(\mathbf{x})$:

$$\mathbf{d}_g(\mathbf{x}) = \mathbf{T}_g(\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_{\text{init},g}); \hat{\boldsymbol{\mu}}_g) - \mathbf{x}, \quad (31)$$

with $\mathbf{T}_g(\mathbf{x}; \hat{\boldsymbol{\mu}}_g)$ the transformation estimated by the registration method. The lower $\mathbf{d}(\mathbf{x})$ for all \mathbf{x} , the more accurate the registration method. However, the constraint, Eq. (14), was not applied to the initial transformation, so we subtract the mean of the deformation field: $\mathbf{d}_g^*(\mathbf{x}) = \mathbf{d}_g(\mathbf{x}) - \frac{1}{G} \sum_{g'=1}^G \mathbf{d}_{g'}(\mathbf{x})$. We report the mean and standard deviation of $\|\mathbf{d}_g^*(\mathbf{x})\|$ over all \mathbf{x} and g . Secondly, we measured how the uncertainty of the fitted parameter of interest of the qMRI model was altered by the registration method (see Section 4.1.4).

For the real data, no ground truth was available. We therefore used, besides a visual inspection, the following evaluation measures, which can be summarized as:

1. how well manually defined landmarks correspond in the images (CT-LUNG, T1VFA-CAROTID and DCE-ABDOMEN);
2. how well anatomical regions of interest (ROI) overlap in all images (T1MOLLI-HEART and ADC-ABDOMEN);
3. the smoothness of the deformation field, evaluated within an ROI;
4. the effect of the registration method on the uncertainty of the fitted parameter of the qMRI model, evaluated within an ROI.

To facilitate a fair comparison of all registration methods for the different evaluation measures, the images resulting from the groupwise registration metrics \mathcal{D}_{VAR} , \mathcal{D}_{APE} , \mathcal{D}_{G-MI} , \mathcal{D}_{T_1} , \mathcal{D}_{PCA} and \mathcal{D}_{PCA2} were transformed to the space of the reference image $M_R(\mathbf{x})$ as used by \mathcal{D}_{MI} . To that purpose, the inverse transformation $\mathbf{T}_R^{-1}(\mathbf{x}; \boldsymbol{\mu}_R)$ was computed. The manually outlined structures that are used as ROIs for all qMRI applications are mentioned in Table 2. For the DTI-BRAIN data, neither landmarks nor structures could be reliably identified on each of the diffusion weighted images, so for this application no overlap or point correspondence was calculated. In the following sections the evaluation methods are explained in more detail.

4.1.1. Point correspondence

The correspondence between landmarks was evaluated with the mean target registration error (mTRE):

$$\text{mTRE}_l = \frac{1}{G-1} \sum_{g=1, g \neq R}^G \|\mathbf{p}_g - \mathbf{T}_g(\mathbf{T}_R^{-1}(\mathbf{p}_R; \boldsymbol{\mu}_R); \boldsymbol{\mu}_g)\|, \quad (32)$$

where G is the number of images in which the landmark l is annotated, \mathbf{p}_R the landmark in the reference image, \mathbf{p}_g the landmark annotated in image g , and $\mathbf{T}_g(\mathbf{T}_R^{-1}(\mathbf{p}_R; \boldsymbol{\mu}_R); \boldsymbol{\mu}_g)$ the transformation of point \mathbf{p}_R from the reference space to the space of image g . The transformations from reference space were used to facilitate the comparison with the non-groupwise \mathcal{D}_{MI} approach. For the CT-LUNG data anatomical landmarks are publicly available. For the DCE-ABDOMEN and T1VFA-CAROTID data the landmarks were manually annotated. Multiple landmarks were annotated and therefore we report the mean over the mTRE_l for all l . For CT-LUNG, 75 landmarks were annotated in the first five timepoints and 300 landmarks were annotated in the first and fifth timepoint. The mTRE is calculated for the 75 landmarks in the first five timepoints and for the 300 landmarks in inhale and exhale phase. The mean of these two mTREs is reported. In the T1VFA-CAROTID data four landmarks per dataset were determined, two per registration mask, and they were defined in all five images. In the DCE-ABDOMEN five landmarks per dataset were determined and they were defined in 50 of the 160 images. Due to sliding motion, low resolution and high contrast differences the five landmarks were annotated in 50 different images, which results in an unequal number of landmarks per image of the DCE-ABDOMEN

data. Note that instead of an mTRE_l , i.e. an mTRE per landmark l , we could have computed an mTRE_g , i.e. an mTRE per image g . For datasets CT-LUNG and T1VFA-CAROTID the mean over mTRE_l for all l is equal to the mean over mTRE_g for all g , but for the DCE-ABDOMEN data this is not the case due to the unequal number of landmarks per image.

4.1.2. Overlap

We extended the Dice coefficient to measure the overlap between more than two segmentations as:

$$\text{Dice}_G = G \frac{|S_1 \cap S_2 \dots \cap S_G|}{|S_1| + |S_2| \dots + |S_G|}, \quad (33)$$

where S_g is the segmentation in the g^{th} image and G the number of segmentations. Note that Dice_G is sensitive to the misregistration of a single image, which is important in qMRI since a single misregistered image can severely reduce the quality of the qMRI model fit.

In the T1MOLLI-HEART experiment we measured the overlap of the manually outlined myocardium, outlined on six to nine images. Not all images were outlined because manual segmentation on some images proved to be difficult due to low contrast between the myocardium and surrounding tissues. For the ADC-ABDOMEN we measured the overlap of the manually outlined spleen in $b = 0 \text{ s/mm}^2$ image and in diffusion weighted images $b = 100 \text{ s/mm}^2$ in phase direction, $b = 100 \text{ s/mm}^2$ in slice direction, $b = 150 \text{ s/mm}^2$ in slice direction, $b = 200 \text{ s/mm}^2$ in read direction, $b = 300 \text{ s/mm}^2$ in read direction, $b = 500 \text{ s/mm}^2$ in phase direction and $b = 900 \text{ s/mm}^2$ in phase direction.

4.1.3. Smoothness of the transformation

Extreme and non-smooth deformations are unexpected for the experiments we conducted. Hence, the smoothness of the deformation field can be used to identify such undesirable transformations. It is obtained by calculating the mean of the standard deviation of the determinant of $\partial \mathbf{T}_g / \partial \mathbf{x}$ over all \mathbf{x} for all images:

$$\text{STD}_{|\partial \mathbf{T}_g / \partial \mathbf{x}|} = \frac{1}{G} \sum_{g=1}^G \text{STD}(|\partial \mathbf{T}_g / \partial \mathbf{x}|), \quad (34)$$

where $\text{STD}(|\partial \mathbf{T}_g / \partial \mathbf{x}|)$ is the standard deviation of $|\partial \mathbf{T}_g / \partial \mathbf{x}|$ over all $\mathbf{x} \in \text{ROI}$.

4.1.4. Uncertainty estimation of the qMRI fit

The uncertainty of the estimated qMRI model parameters θ was quantified by the Cramér–Rao lower bound (CRLB), which provides a lower bound for the variance of the ML estimated parameters (Sijbers et al., 1999; Cavassila et al., 2001; Rao, 1946; Poot and Klein, 2015). The CRLB of a function $f(\theta)$ is given by:

$$\text{CRLB}_f(\theta, \varepsilon) = \left(\frac{\partial f}{\partial \theta} \right)^T \mathbf{I}^{-1}(\theta, \varepsilon) \left(\frac{\partial f}{\partial \theta} \right), \quad (35)$$

where $\mathbf{I}(\theta, \varepsilon)$ is the Fisher information matrix. For Rician distributed data, no closed-form expression for the Fisher information exists. To evaluate the CRLB, we used the method by Poot et al. (2013), which approximates the integrals present in the evaluation of the Fisher information to double precision accuracy. To use the CRLB as indicator of misalignment we adopt the measure proposed by Bron et al. (2013), which is the 90% percentile of the square root of the CRLB ($90\% \sqrt{\text{CRLB}}$) over an ROI. This measure identifies misalignment, because misalignment may result in biologically implausible values of the estimated parameters, especially at tissue boundaries. Additionally, the model will fit less accurately to the data, resulting in a higher estimated noise level and thus higher CRLB (Bron et al., 2013). A 90% percentile was chosen because registration errors are mainly visible at the edges of a region and

Table 3

Results of the SYNTH-MODEL experiment. Mean \pm standard deviation of $\|\mathbf{d}_g^*(\mathbf{x})\|$ [mm].

	Affine	B-spline
(-)	6.2 \pm 3.7	3.0 \pm 1.6
\mathcal{D}_{MI}	7.7 \pm 8.9	9.3 \pm 7.1
\mathcal{D}_{VAR}	39.7 \pm 30.7	34.0 \pm 21.5
\mathcal{D}_{APE}	8.4 \pm 4.9	21.5 \pm 13.3
$\mathcal{D}_{\text{G-MI}}$	20.1 \pm 14.0	16.9 \pm 12.2
\mathcal{D}_{PCA}	0.1 \pm 0.1	0.6 \pm 0.6
$\mathcal{D}_{\text{PCA2}}$	0.1 \pm 0.3	1.5 \pm 1.9

Table 4

Results of the synthetic qMRI experiments. Mean \pm standard deviation of $\|\mathbf{d}_g^*(\mathbf{x})\|$ [mm].

	s-T1MOLLI	s-T1VFA	s-ADC	s-DTI	s-DCE
(-)	1.55 \pm 1.10	1.02 \pm 0.79	0.98 \pm 0.63	1.32 \pm 0.91	1.18 \pm 0.64
\mathcal{D}_{MI}	1.83 \pm 1.94	0.43 \pm 0.38	1.80 \pm 2.03	1.46 \pm 1.21	3.03 \pm 5.34
\mathcal{D}_{VAR}	10.01 \pm 10.02	1.64 \pm 1.44	5.03 \pm 4.85	9.88 \pm 11.54	1.67 \pm 5.93
\mathcal{D}_{APE}	2.71 \pm 2.88	0.14 \pm 0.12	1.14 \pm 0.85	1.77 \pm 2.06	0.96 \pm 0.94
$\mathcal{D}_{\text{G-MI}}$	1.35 \pm 1.54	0.41 \pm 0.34	1.26 \pm 1.01	1.23 \pm 1.45	0.93 \pm 1.02
\mathcal{D}_{T_1}	0.23 \pm 0.21	-	-	-	-
\mathcal{D}_{PCA}	0.30 \pm 0.28	0.15 \pm 0.13	0.76 \pm 0.57	0.16 \pm 0.16	1.16 \pm 3.19
$\mathcal{D}_{\text{PCA2}}$	0.79 \pm 0.93	0.12 \pm 0.11	0.73 \pm 0.57	0.23 \pm 0.28	0.75 \pm 1.26

Table 5

90% $\sqrt{\text{CRLB}}$ of T_1 [ms], ADC [$\mu\text{m}^2/\text{ms}$], MD [$\mu\text{m}^2/\text{ms}$] and K^{trans} [min^{-1}] in ROI of synthetic datasets.

	s-T1MOLLI	s-T1VFA	s-ADC	s-DTI	s-DCE
(-)	148	>1000	0.243	0.063	1.79
\mathcal{D}_{MI}	50	542	0.046	0.023	4.05
\mathcal{D}_{VAR}	121	>1000	0.043	0.021	2.76
\mathcal{D}_{APE}	75	206	0.054	0.018	0.92
$\mathcal{D}_{\text{G-MI}}$	47	390	0.093	0.007	0.39
\mathcal{D}_{T_1}	25	-	-	-	-
\mathcal{D}_{PCA}	27	203	0.039	0.008	1.31
$\mathcal{D}_{\text{PCA2}}$	34	222	0.028	0.008	0.32

many voxels in the ROI will not make a difference; however we do not want to be sensitive to a very low number of extremely high outliers. Please note that when the 90% $\sqrt{\text{CRLB}}$ is high, the data is likely to be misaligned. However, when the 90% $\sqrt{\text{CRLB}}$ is low, it is not certain if the data is well aligned, which is why this measure should preferably be used in combination with other criteria.

5. Results

The results in Tables 3–11 are color-coded. A red-to-white color scale is used to visually depict the (rank of the) results, where red indicates a worse result and white indicates a better result.

5.1. Results on synthetic data

The mean and standard deviation of $\|\mathbf{d}_g^*(\mathbf{x})\|$ of the SYNTH-MODEL images are shown in Table 3. The left column shows the mean and standard deviation of $\|\mathbf{d}_g^*(\mathbf{x})\|$ in the case of $\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_{\text{init},g})$ being an affine transformation, and in the right column the case of $\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_{\text{init},g})$ being a B-spline transformation. Only the proposed methods \mathcal{D}_{PCA} and $\mathcal{D}_{\text{PCA2}}$ were successful in aligning the images.

Table 4 shows the mean \pm standard deviation of $\|\mathbf{d}_g^*(\mathbf{x})\|$ of the synthetic qMRI data. This evaluation measure is calculated within

Table 6

Results of the CT-LUNG experiment. Mean and standard deviation over all subjects for all evaluation measures and control point spacings 6 mm, 13 mm, and 20 mm.

mTRE [mm]	6	13	20
(-)	6.72 ± 2.51	6.72 ± 2.51	6.72 ± 2.51
\mathcal{D}_{MI}	1.78 ± 0.40	1.43 ± 0.23	1.45 ± 0.21
\mathcal{D}_{VAR}	1.59 ± 0.59	1.42 ± 0.39	1.47 ± 0.37
\mathcal{D}_{APE}	1.47 ± 0.60	1.40 ± 0.37	1.45 ± 0.34
\mathcal{D}_{G-MI}	1.37 ± 0.38	1.28 ± 0.21	1.31 ± 0.21
\mathcal{D}_{PCA}	1.47 ± 0.60	1.40 ± 0.37	1.45 ± 0.34
\mathcal{D}_{PCA2}	1.72 ± 0.78	1.56 ± 0.55	1.59 ± 0.49

STD $ \partial T_g/\partial x $ [%]	6	13	20
(-)	0 ± 0	0 ± 0	0 ± 0
\mathcal{D}_{MI}	28 ± 8	15 ± 4	11 ± 3
\mathcal{D}_{VAR}	14 ± 4	9 ± 2	7 ± 2
\mathcal{D}_{APE}	12 ± 3	8 ± 2	7 ± 2
\mathcal{D}_{G-MI}	13 ± 2	9 ± 2	7 ± 2
\mathcal{D}_{PCA}	12 ± 3	8 ± 2	7 ± 2
\mathcal{D}_{PCA2}	10 ± 2	7 ± 2	6 ± 1

the registration mask when that is present or else on the entire image. For all experiments except DCE-ABDOMEN, \mathcal{D}_{VAR} showed the worst registration accuracy compared to all other methods. For s-T1MOLLI, \mathcal{D}_{T_1} performed best, which is to be expected, since, in this simulation example, the intensities perfectly adhere to the T1-

MOLLI model. For the remaining applications either \mathcal{D}_{PCA} or \mathcal{D}_{PCA2} outperformed the other methods.

Table 5 shows the $90\% \sqrt{CRLB}$ of the tissue property of interest of the synthetic datasets, before and after registration, evaluated in a specified ROI. For s-T1MOLLI the model-based metric \mathcal{D}_{T_1} performed best. For all other applications both \mathcal{D}_{PCA} or \mathcal{D}_{PCA2} achieved the lowest $90\% \sqrt{CRLB}$ of the tissue property.

5.2. Results on real data

The results of the CT-LUNG experiment are shown in Table 6. The best control point spacing for this data is 13 mm. Method \mathcal{D}_{G-MI} had the best performance, and \mathcal{D}_{PCA} performed similar to \mathcal{D}_{MI} , \mathcal{D}_{VAR} and \mathcal{D}_{APE} . Additionally, to allow for a direct comparison, we computed the mean of $mTRE_l$ for all l of \mathcal{D}_{VAR} on the subset of CT-LUNG used by Metz et al. (2011). In this subset we observed 1.26(0.33) where 1.26(0.27) was reported by Metz et al. Slight differences may be due to different mask and sampling strategies.

The results of the experiments on the real T1MOLLI-HEART data are shown in Table 7. Even though the lowest $90\% \sqrt{CRLB}$ is obtained for \mathcal{D}_{T_1} , the methods \mathcal{D}_{PCA} and \mathcal{D}_{PCA2} have the highest Dice $_G$ and the lowest STD $|\partial T_g/\partial x|$. \mathcal{D}_{VAR} performed worst on all evaluation measures. Also notice that the result of \mathcal{D}_{MI} depends heavily on the choice of reference image. The best results are obtained for reference image 11, which is the image acquired with the longest inversion time and therefore having the highest SNR.

The results on the T1VFA-CAROTID data are shown in Table 8. Because the registration ran separately for the left and right carotid, the mean over the two carotids is taken per subject. For

Table 7

Results of the T1MOLLI-HEART experiment. Mean and standard deviation over all subjects for all evaluation measures and control point spacings 32 mm, 64 mm, and 128 mm and for reference images 1, 4, 7 and 11.

Dice $_G$ [%]	32				64				128			
	1	4	7	11	1	4	7	11	1	4	7	11
(-)	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8	48 ± 8
\mathcal{D}_{MI}	26 ± 11	22 ± 10	22 ± 15	39 ± 14	37 ± 11	34 ± 16	19 ± 15	52 ± 7	44 ± 14	40 ± 13	24 ± 17	51 ± 7
\mathcal{D}_{VAR}	2 ± 3	4 ± 6	5 ± 6	4 ± 5	16 ± 12	16 ± 12	16 ± 12	16 ± 12	35 ± 17	35 ± 17	35 ± 17	35 ± 17
\mathcal{D}_{APE}	20 ± 10	21 ± 10	22 ± 11	21 ± 10	48 ± 11	48 ± 10	48 ± 11	48 ± 11	48 ± 9	49 ± 9	48 ± 9	48 ± 8
\mathcal{D}_{G-MI}	36 ± 15	36 ± 15	36 ± 15	36 ± 15	41 ± 7	42 ± 8	42 ± 7	41 ± 8	43 ± 10	43 ± 10	43 ± 10	43 ± 10
\mathcal{D}_{T_1}	44 ± 15	45 ± 15	45 ± 15	44 ± 14	50 ± 13	51 ± 13	50 ± 13	50 ± 13	50 ± 10	51 ± 9	51 ± 10	51 ± 10
\mathcal{D}_{PCA}	54 ± 9	54 ± 9	54 ± 9	54 ± 9	53 ± 7	53 ± 7	53 ± 7	53 ± 8	51 ± 9	50 ± 9	51 ± 9	51 ± 9
\mathcal{D}_{PCA2}	52 ± 13	52 ± 13	52 ± 13	52 ± 13	52 ± 11	52 ± 10	52 ± 11	53 ± 10	52 ± 7	52 ± 8	52 ± 8	52 ± 8

STD $ \partial T_g/\partial x $ [%]	32				64				128			
	1	4	7	11	1	4	7	11	1	4	7	11
(-)	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
\mathcal{D}_{MI}	24 ± 8	28 ± 8	29 ± 9	10 ± 4	7 ± 2	7 ± 2	8 ± 3	3 ± 1	1 ± 0	1 ± 0	1 ± 0	1 ± 0
\mathcal{D}_{VAR}	82 ± 31	60 ± 20	57 ± 21	78 ± 27	16 ± 14	9 ± 5	10 ± 6	11 ± 6	1 ± 1	1 ± 0	1 ± 0	1 ± 0
\mathcal{D}_{APE}	17 ± 7	17 ± 7	18 ± 8	13 ± 6	3 ± 2	3 ± 3	4 ± 4	2 ± 2	0 ± 0	0 ± 0	0 ± 0	0 ± 0
\mathcal{D}_{G-MI}	10 ± 3	12 ± 11	15 ± 13	7 ± 4	2 ± 1	2 ± 2	2 ± 2	1 ± 1	0 ± 0	0 ± 0	0 ± 0	0 ± 0
\mathcal{D}_{T_1}	15 ± 9	11 ± 5	14 ± 8	8 ± 3	3 ± 2	3 ± 2	3 ± 2	2 ± 1	0 ± 0	0 ± 0	0 ± 0	0 ± 0
\mathcal{D}_{PCA}	6 ± 4	4 ± 2	4 ± 1	4 ± 1	2 ± 1	1 ± 1	2 ± 1	1 ± 1	0 ± 0	0 ± 0	0 ± 0	0 ± 0
\mathcal{D}_{PCA2}	6 ± 2	5 ± 2	6 ± 2	4 ± 1	1 ± 1	1 ± 1	2 ± 1	1 ± 1	0 ± 0	0 ± 0	0 ± 0	0 ± 0

90% $\sqrt{CRLB} T_1$ [ms]	32				64				128			
	1	4	7	11	1	4	7	11	1	4	7	11
(-)	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19	92 ± 19
\mathcal{D}_{MI}	119 ± 26	157 ± 61	467 ± 610	93 ± 19	97 ± 16	103 ± 18	146 ± 52	83 ± 13	91 ± 14	95 ± 16	121 ± 29	81 ± 12
\mathcal{D}_{VAR}	523 ± 1044	279 ± 220	>1000	164 ± 62	179 ± 202	118 ± 28	173 ± 111	111 ± 25	124 ± 70	98 ± 20	102 ± 25	97 ± 19
\mathcal{D}_{APE}	94 ± 20	93 ± 25	92 ± 22	89 ± 17	85 ± 12	82 ± 14	84 ± 17	82 ± 10	86 ± 12	81 ± 12	84 ± 15	82 ± 10
\mathcal{D}_{G-MI}	90 ± 14	88 ± 15	88 ± 24	85 ± 15	86 ± 14	81 ± 12	85 ± 21	79 ± 11	89 ± 7	84 ± 8	83 ± 11	80 ± 5
\mathcal{D}_{T_1}	83 ± 18	76 ± 17	76 ± 19	80 ± 19	82 ± 19	77 ± 18	76 ± 19	75 ± 13	80 ± 15	77 ± 14	76 ± 15	75 ± 13
\mathcal{D}_{PCA}	83 ± 17	79 ± 17	78 ± 11	85 ± 16	87 ± 16	84 ± 20	83 ± 13	85 ± 13	88 ± 14	89 ± 21	103 ± 50	85 ± 10
\mathcal{D}_{PCA2}	81 ± 15	79 ± 16	79 ± 12	79 ± 12	83 ± 12	79 ± 14	80 ± 10	81 ± 10	84 ± 12	82 ± 15	82 ± 10	81 ± 11

Table 8

Results of the T1VFA-CAROTID experiment. Mean and standard deviation over all subjects for all evaluation measures and control point spacings 8 mm, 16 mm, and 32 mm.

mTRE [mm]	8	16	32
(-)	1.47 ± 0.54	1.47 ± 0.54	1.47 ± 0.54
\mathcal{D}_{MI}	1.26 ± 0.44	1.22 ± 0.43	1.23 ± 0.45
\mathcal{D}_{VAR}	2.65 ± 0.79	1.84 ± 0.36	1.16 ± 0.37
\mathcal{D}_{APE}	1.24 ± 0.55	1.11 ± 0.41	1.10 ± 0.43
\mathcal{D}_{G-MI}	1.16 ± 0.48	1.08 ± 0.42	1.10 ± 0.43
\mathcal{D}_{PCA}	1.25 ± 0.56	1.11 ± 0.42	1.10 ± 0.43
\mathcal{D}_{PCA2}	1.13 ± 0.46	1.08 ± 0.39	1.10 ± 0.43
STD $_{ \partial T_g/\partial x }$ [%]	8	16	32
(-)	0 ± 0	0 ± 0	0 ± 0
\mathcal{D}_{MI}	7 ± 1	2 ± 0	0 ± 0
\mathcal{D}_{VAR}	49 ± 25	6 ± 3	1 ± 0
\mathcal{D}_{APE}	6 ± 1	2 ± 1	0 ± 0
\mathcal{D}_{G-MI}	5 ± 1	1 ± 0	0 ± 0
\mathcal{D}_{PCA}	6 ± 1	2 ± 1	0 ± 0
\mathcal{D}_{PCA2}	5 ± 1	1 ± 0	0 ± 0
90% $\sqrt{CRLB} T_1$ [ms]	8	16	32
(-)	>1000	>1000	>1000
\mathcal{D}_{MI}	530 ± 136	501 ± 83	523 ± 93
\mathcal{D}_{VAR}	>1000	>1000	848 ± 605
\mathcal{D}_{APE}	539 ± 154	498 ± 94	530 ± 95
\mathcal{D}_{G-MI}	514 ± 134	491 ± 83	515 ± 82
\mathcal{D}_{PCA}	540 ± 154	498 ± 93	530 ± 94
\mathcal{D}_{PCA2}	532 ± 154	510 ± 110	523 ± 87

Table 9

Results of the ADC-ABDOMEN experiment. Mean and standard deviation over all subjects for all evaluation measures and control point spacings 32 mm, 64 mm, and 128 mm and an affine transformation.

Dice $_G$ [%]	32	64	128	Affine
(-)	70 ± 4	70 ± 4	70 ± 4	70 ± 4
\mathcal{D}_{MI}	61 ± 18	64 ± 16	73 ± 8	72 ± 7
\mathcal{D}_{VAR}	48 ± 25	50 ± 27	58 ± 17	52 ± 14
\mathcal{D}_{APE}	69 ± 5	72 ± 4	73 ± 5	73 ± 5
\mathcal{D}_{G-MI}	58 ± 10 *	69 ± 5 *	72 ± 5 *	71 ± 5
\mathcal{D}_{PCA}	65 ± 13	71 ± 5	71 ± 4	72 ± 4
\mathcal{D}_{PCA2}	75 ± 7	75 ± 5	73 ± 5	71 ± 5
STD $_{ \partial T_g/\partial x }$ [%]	32	64	128	Affine
(-)	0 ± 0	0 ± 0	0 ± 0	-
\mathcal{D}_{MI}	25 ± 11	8 ± 3	2 ± 1	-
\mathcal{D}_{VAR}	33 ± 13	8 ± 5	2 ± 1	-
\mathcal{D}_{APE}	32 ± 15	7 ± 4	1 ± 1	-
\mathcal{D}_{G-MI}	27 ± 11 *	7 ± 3 *	1 ± 0 *	-
\mathcal{D}_{PCA}	11 ± 3	3 ± 1	1 ± 0	-
\mathcal{D}_{PCA2}	8 ± 2	3 ± 1	0 ± 0	-
90% $\sqrt{CRLB} ADC$ [$\mu m^2/ms$]	32	64	128	Affine
(-)	1.37 ± 0.83	1.37 ± 0.83	1.37 ± 0.83	1.37 ± 0.83
\mathcal{D}_{MI}	0.23 ± 0.05	0.25 ± 0.05	0.29 ± 0.05	0.52 ± 0.35
\mathcal{D}_{VAR}	0.20 ± 0.08	0.23 ± 0.10	0.35 ± 0.16	0.43 ± 0.21
\mathcal{D}_{APE}	0.42 ± 0.06	0.42 ± 0.13	0.40 ± 0.14	0.50 ± 0.29
\mathcal{D}_{G-MI}	0.44 ± 0.10*	0.42 ± 0.14*	0.37 ± 0.05*	0.46 ± 0.12
\mathcal{D}_{PCA}	0.15 ± 0.04	0.23 ± 0.06	0.46 ± 0.27	0.55 ± 0.30
\mathcal{D}_{PCA2}	0.23 ± 0.03	0.27 ± 0.05	0.41 ± 0.18	0.50 ± 0.26

*The registration of one subject failed and was not included in the results.

Table 10

Results of the DTI-BRAIN experiment. Mean and standard deviation of the 90% \sqrt{CRLB} MD over all subjects.

90% \sqrt{CRLB} MD [$\mu m^2/ms$]	Affine
(-)	0.096 ± 0.029
\mathcal{D}_{MI}	0.084 ± 0.028
\mathcal{D}_{VAR}	1.930 ± 4.000
\mathcal{D}_{APE}	0.085 ± 0.029
\mathcal{D}_{G-MI}	0.120 ± 0.039
\mathcal{D}_{PCA}	0.085 ± 0.029
\mathcal{D}_{PCA2}	0.084 ± 0.028

Table 11

Results of the DCE-ABDOMEN experiment. Mean and standard deviation over all subjects for all evaluation measures and control point spacings 32 mm, 64 mm, and 128 mm and an affine transformation.

mTRE [mm]	32	64	128	Affine
(-)	8.49 ± 4.54	8.49 ± 4.54	8.49 ± 4.54	8.49 ± 4.54
\mathcal{D}_{MI}	6.73 ± 2.02	6.46 ± 2.32	6.47 ± 2.37	7.35 ± 3.55
\mathcal{D}_{VAR}	14.92 ± 5.65	6.98 ± 1.45	6.29 ± 1.98	7.01 ± 2.10
\mathcal{D}_{APE}	13.46 ± 5.48	6.86 ± 2.33	6.36 ± 2.37	6.91 ± 3.05
\mathcal{D}_{G-MI}	7.51 ± 1.47	6.41 ± 2.37	6.41 ± 2.38	7.31 ± 3.79
\mathcal{D}_{PCA}	6.21 ± 2.25	6.11 ± 2.32	6.24 ± 2.37	6.54 ± 2.69
\mathcal{D}_{PCA2}	5.89 ± 2.23	5.99 ± 2.17	6.18 ± 2.27	6.62 ± 2.75
STD $_{ \partial T_g/\partial x }$ [%]	32	64	128	Affine
(-)	0 ± 0	0 ± 0	0 ± 0	-
\mathcal{D}_{MI}	20 ± 9	4 ± 2	1 ± 1	-
\mathcal{D}_{VAR}	22 ± 14	4 ± 1	1 ± 0	-
\mathcal{D}_{APE}	21 ± 12	4 ± 1	1 ± 0	-
\mathcal{D}_{G-MI}	17 ± 6	3 ± 1	1 ± 0	-
\mathcal{D}_{PCA}	11 ± 4	4 ± 2	1 ± 0	-
\mathcal{D}_{PCA2}	6 ± 3	2 ± 1	0 ± 0	-
90% $\sqrt{CRLB} k^{trans}$ [min^{-1}]	32	64	128	Affine
(-)	2.84 ± 2.30	2.84 ± 2.30	2.84 ± 2.30	2.84 ± 2.30
\mathcal{D}_{MI}	3.85 ± 2.41	3.64 ± 4.13	2.54 ± 2.58	3.79 ± 4.11
\mathcal{D}_{VAR}	3.08 ± 2.68	1.38 ± 0.74	1.28 ± 0.86	1.54 ± 0.95
\mathcal{D}_{APE}	2.18 ± 1.56	1.49 ± 1.30	1.59 ± 1.68	1.63 ± 1.24
\mathcal{D}_{G-MI}	1.61 ± 1.05	1.42 ± 1.15	1.37 ± 1.23	2.08 ± 2.05
\mathcal{D}_{PCA}	1.69 ± 1.48	1.52 ± 1.18	1.46 ± 1.09	1.52 ± 1.17
\mathcal{D}_{PCA2}	1.17 ± 0.87	1.27 ± 0.92	1.38 ± 1.16	1.81 ± 1.70

all methods the lowest 90% $\sqrt{CRLB} T_1$ is obtained at $\nu = 16$ mm, of which \mathcal{D}_{G-MI} performed best. At $\nu = 16$ mm a similar point correspondence and transformation smoothness are obtained by all methods except for \mathcal{D}_{VAR} .

The results of the real ADC-ABDOMEN data are shown in Table 9. The 90% \sqrt{CRLB} ADC in the ROI is decreased after registration for all methods and all control point spacings. Extreme deformations, especially at higher b-values cause the lower overlap, which can be seen from the relatively high STD $_{|\partial T_g/\partial x|}$ for low control point spacings. \mathcal{D}_{PCA2} performed best in terms of Dice $_G$ overlap and has the lowest STD $_{|\partial T_g/\partial x|}$ and 90% \sqrt{CRLB} ADC for $\nu = 32$ mm.

The results of the DTI-BRAIN data are shown in Table 10. Both \mathcal{D}_{VAR} and \mathcal{D}_{G-MI} increase the 90% \sqrt{CRLB} MD. A similar 90% \sqrt{CRLB} MD is obtained for the remaining methods.

The results of the DCE-ABDOMEN data are shown in Table 11. The best 90% $\sqrt{CRLB} k^{trans}$ was obtained for \mathcal{D}_{PCA2} for $\nu = 32$ mm, whereas \mathcal{D}_{MI} performed worst on that measure with $\nu = 32$ mm. The point correspondence is worst for \mathcal{D}_{VAR} at $\nu = 32$ mm. Overall \mathcal{D}_{PCA2} showed the best point correspondence and the lowest STD $_{|\partial T_g/\partial x|}$.

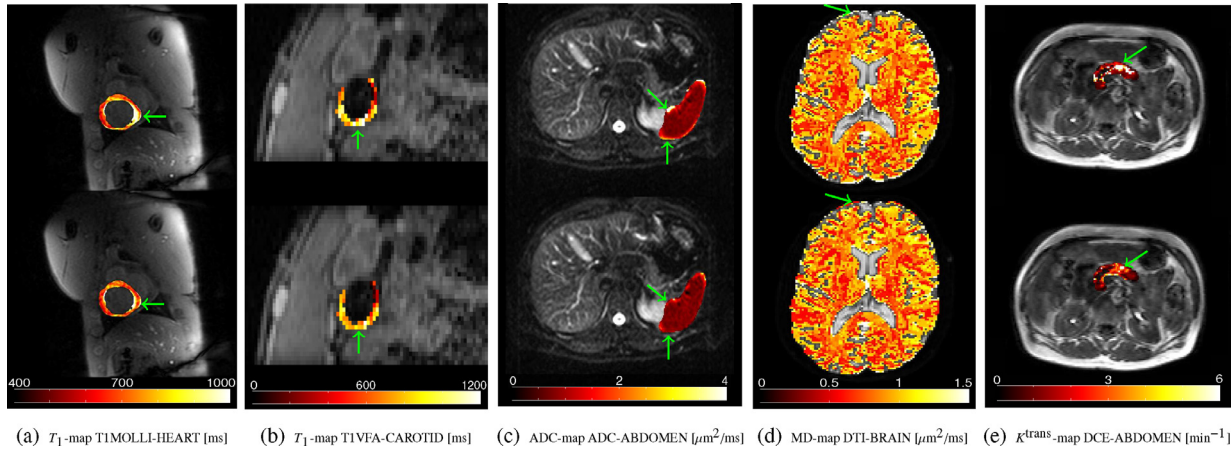


Fig. 6. Tissue maps, generated before (top) and after registration (bottom) with $\mathcal{D}_{\text{PCA}2}$. The green arrows indicate the differences in the tissue maps before and after registration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.3. Visual inspection of qMRI data

Fig. 6 shows tissue maps before registration and after registration with $\mathcal{D}_{\text{PCA}2}$ of specific datasets that are used in the experiments. The difference in the estimated tissue maps is clearly visible, especially at tissue boundaries, showing the importance of registration as a pre-processing step prior to fitting the qMRI data. For each qMRI application we extracted slices from an original real qMRI dataset and a motion corrected dataset (with $\mathcal{D}_{\text{PCA}2}$), which can be viewed online with the Elsevier viewer, which is available at the online version of this manuscript.

6. Discussion

We proposed two general registration methods for qMRI and we compared our methods to five other state-of-the-art registration methods. The following sections discuss in more detail the results of the synthetic and real data, some limitations of the evaluation, and possibilities for future work.

6.1. Synthetic data

A purely synthetic dataset was created to clearly demonstrate the advantages that the proposed PCA-based methods have. In this example with low pairwise mutual information between the images, the experiments confirmed that the methods based on mutual information (\mathcal{D}_{MI} , $\mathcal{D}_{\text{G-MI}}$), or groupwise correlation (\mathcal{D}_{APE}) failed to register the data. The proposed PCA-based methods gave very good registration results in this challenging example. It should be emphasized that, as the images in this synthetic example are linear combinations of a few basis images, it is expected that methods based on PCA give good results. Nevertheless this example provides insight in the functionality of the registration methods.

On synthetic data created using existing qMRI models, the PCA-based methods perform well in all experiments for both evaluation measures, whereas the other methods either fail or have only good results in a selection of the experiments.

For the s-T1VFA-CAROTID and T1VFA-CAROTID we used $L = 1$ instead of $L = \Gamma = 3$ for metric \mathcal{D}_{PCA} . Visual inspection of initial results led to this choice. A possible explanation why $L = 1$ showed better results is that the relatively small registration mask and the small number of acquired images could lead to only one dominant eigenvector. The second proposed metric, $\mathcal{D}_{\text{PCA}2}$, eliminates the necessity to choose L and shows results comparable to \mathcal{D}_{PCA} .

6.2. Real data

In the experiments where a B-spline transformation was used, we investigated the influence of different values for the control point spacing. All experiments show that the proposed PCA-based metrics show less extreme deformations and have better results than the other methods, which is particularly the case for low control point spacings.

The CT-LUNG data was used to evaluate the proposed methods on data without intensity variation. The best result on the CT-LUNG data is obtained with $\mathcal{D}_{\text{G-MI}}$, possibly due to the slight intensity variations that are present in the images. However, \mathcal{D}_{MI} , \mathcal{D}_{VAR} , \mathcal{D}_{APE} and \mathcal{D}_{PCA} showed competitive results.

In the TIMOLLI-HEART experiment we used different images as a reference image. The results show that the pairwise approach using \mathcal{D}_{MI} heavily depends on the choice of reference image, while the groupwise approaches do not use a reference image during registration (and hence have a consistent performance across reference images).

\mathcal{D}_{T_1} performed well on the synthetic data, but on real data the Dice_G was lower and the deformations were less smooth than those of the proposed methods. Furthermore, this method needs customization for different qMRI models.

6.3. Limitations of the evaluation

Evaluation of image registration results is usually a hard task because no ground truth is available. By using both synthetic and real qMRI data we extensively evaluated the registration methods. Registration accuracy was measured in terms of a mean and standard deviation of a residual deformation field (synthetic data), point correspondence or a groupwise Dice overlap (real data), the smoothness of the obtained deformation (real data), and the uncertainty of the qMRI parameter of interest (real and synthetic data).

It was impossible to accurately outline the ROIs on all images of the TIMOLLI-HEART data, because in some images there was not enough contrast between the tissues. This may result in an overly optimistic overlap measure, since those images might also be harder to register. In the DCE-ABDOMEN data, it proved difficult to manually annotate landmarks in the images, due to sliding motion of the various organs combined with the breathing motion. These problems, associated with the lack of a ground truth, were alleviated by reporting multiple evaluation measures, and by visually inspecting the data. In this way, a relatively complete picture of registration performance was given.

Smoothness and periodicity over time, which were modeled explicitly by Metz et al. (2011), were not assumed or enforced in any of the experiments in our work. It could, however, be incorporated into our PCA-based method in the same way as proposed by Metz et al. (2011). Both in DCE-ABDOMEN and ADC-ABDOMEN we observed sliding motion. We did not explicitly account for this in the transformation model we used. However, since our methods can be extended with any other transformation model, one could use e.g. the transformation models proposed by Delmon et al. (2013) and Berendsen et al. (2014), which are designed to account for sliding motion.

No explicit regularization of the deformation field was used in the registration experiments. Different amounts of implicit regularization were compared by using different values for the control point spacings of the B-spline transformation model. Adding a regularization term such as bending energy may even further improve the results.

The CRLB was used as a measure to detect voxels that are fitted with a high uncertainty due to e.g. misalignment. Taking the 90% percentile over a voxel-wise fitted ROI is a good measure for registration accuracy and the values we report should be seen as such. It should be stressed that, since the 90% percentile is a rather conservative statistic, the $90\% \sqrt{\text{CRLB}}$ values reported could be misleading when seen as an indicator of the precision typically attained. For that purpose, the entire distribution of CRLB values in the ROI would have to be used.

6.4. Future work

Since the qMRI models are typically non-linear, a linear dimension-reduction technique such as PCA can only approximate the dimension of the subspace. It would therefore be interesting to investigate replacing PCA by a non-linear dimension reduction technique such as kernel-PCA or Laplacian eigenmaps, similar to e.g. Baumgartner et al. (2014). However, this is computationally demanding and it may not result in better registrations if the minimum of the dissimilarity metric remains at the same location in the transform parameter space.

The proposed registration methods are potentially applicable to any set of images where the voxel intensities are defined by a low-dimensional model. Possible other applications in medical imaging are perfusion computed tomography (CT) and dynamic positron emission tomography (PET). Perfusion CT is used for e.g. calculation of the myocardial blood flow index to detect functionally significant coronary lesions (Rossi et al., 2014). In perfusion CT a contrast agent is injected in the coronary arteries and the heart is imaged at several time points, giving each image a different contrast. Dynamic PET is used for e.g. estimating the energy consumption in tumors. In dynamic PET a radio-active tracer is injected and multiple time frames are imaged to assess both the spatial and temporal pattern of the tracer uptake (Muzi et al., 2012). Both applications acquire multiple time points in which the contrast changes and a low-dimensional model is fitted to the data to obtain quantitative information about tissues. Therefore, the proposed methods could potentially be used to correct the motion in such data.

7. Conclusion

We proposed two novel generic groupwise registration methods for qMRI. We evaluated our methods on five different applications of qMRI and compared it to five other state-of-the-art registration methods. We showed the advantage of a groupwise approach versus a pairwise approach and showed that our proposed methods have better or equal registration accuracy as the other methods. Both methods that were proposed showed good results. When no

or little intensity changes are expected we would recommend using \mathcal{D}_{PCA} and choose $L = 1$ for minimal computational complexity and good registration results. For compensating misalignments in qMRI the preferred dissimilarity metric is $\mathcal{D}_{\text{PCA}2}$, because it is free of any user-defined parameters.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under Grant agreement no. 601055, VPH-DARE@IT.

The research of Dr. A. Leemans is supported by VIDU Grant 639.072.411 from the Netherlands Organisation for Scientific Research (NWO).

The authors would like to thank Dr. H.M.M. van Beusekom for helping with the acquisition of the T1MOLLI-HEART data. The acquisition of the T1MOLLI-HEART data was financially supported by Agentschap NL (Grant no. ISO43050) (SENTER-NOVEM): “A novel approach to myocardial regeneration” to Dr. H.M.M. van Beusekom et al.

The authors would like to thank N.M. deSouza, MD, PhD, L. Bernardin, MD, and N. Douglas, MSc, Institute of Cancer Research, London, UK, for providing the ADC-ABDOMEN data. The ADC-ABDOMEN data were acquired in the context of the Quic-Concept project, which has received support from the Innovative Medicines Initiative Joint Undertaking (<http://www.imi.europa.eu>) under Grant agreement number 115151, resources of which are composed of financial contribution from the European Union’s Seventh Framework Programme (FP7/2007–2013) and EFPIA companies’ in kind contribution.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.media.2015.12.004](https://doi.org/10.1016/j.media.2015.12.004).

References

- Andersson, J., Skare, S., 2002. A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. *NeuroImage* 16 (1), 177–199.
- Balci, S., Golland, P., Shenton, M., Wells, M., 2007. Free-form B-spline deformation model for groupwise registration. In: *Proceedings of Statistical Registration – MICCAI Workshop*, pp. 23–30.
- Basser, P., Jones, D., 2002. Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review. *Nucl. Magn. Reson. Biomed.* 15, 456–467.
- Baumgartner, C., Kolbitsch, C., Balfour, D., Marsden, P., McClelland, J., Rueckert, D., King, A., 2014. High-resolution dynamic MR imaging of the thorax for respiratory motion correction of PET using groupwise manifold alignment. *Med. Image Anal.* 18 (7), 939–952.
- Behrens, T., Berg, H.J., Jbabdi, S., Rushworth, M., Woolrich, M., 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *NeuroImage* 34 (1), 144–155.
- Berendsen, F., Kotte, A., Viergever, M., Pluim, J., 2014. Registration of organs with sliding interfaces and changing topologies. In: *Proceedings of SPIE Medical Imaging 9034: Medical Imaging 2014: Image Processing*, vol. 9034. SPIE, p. 90340E.
- Bhatia, K., Hajnal, J., Hammers, A., Rueckert, D., 2007. Similarity metrics for groupwise non-rigid registration. In: *Lecture Notes in Computer Science – MICCAI 2007*. Springer, pp. 544–552.
- Bhushan, M., Schnabel, J., Risser, L., Heinrich, M., Brady, J., Jenkinson, M., 2011. Motion correction and parameter estimation in dceMRI sequences: Application to colorectal cancer. In: *Lecture Notes in Computer Science MICCAI 2011*, vol. 6891, pp. 476–483.
- Bron, E., van Tiel, J., Smit, H., Poot, D., Niessen, W., Krestin, G., Weinans, H., Oei, E., Kotek, G., Klein, S., 2013. Image registration improves human knee cartilage T1 mapping with delayed gadolinium-enhanced MRI of cartilage (dGEMRIC). *Eur. Radiol.* 23 (1), 246–252.
- Buonaccorsi, G., O’Connor, J., Counce, A., Roberts, C., Cheung, S., Watson, Y., Davies, K., Hope, L., Jackson, A., Jayson, G., Parker, G., 2007. Tracer kinetic model-driven registration for dynamic contrast enhanced MRI time series data. *Magn. Reson. Med.* 5 (58), 1010–1019.
- Castillo, R., Castillo, E., Guerra, R., Johnson, V., McPhail, T., Garg, A., Guerrero, T., 2009. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Phys. Med. Biol.* 54 (7), 1849.

- Cavassila, S., Deval, S., Huegen, C., van Ormondt, D., Graveron-Demilly, D., 2001. Cramér-Rao bounds: an evaluation tool for quantitation. *Nucl. Magn. Reson. Biomed.* 14, 278–283.
- Coolen, B., Poot, D., Liem, M., Smits, L., Gao, S., Kotek, G., Klein, S., Nederveen, A., 2015. Three-dimensional quantitative T1 and T2 mapping of the carotid artery: Sequence design and in vivo feasibility. *Magn. Reson. Med.* <http://www.ncbi.nlm.nih.gov/pubmed/25920036>.
- Delmon, V., Rit, S., Pinho, R., Sarrut, D., 2013. Registration of sliding objects using direction dependent B-splines decomposition. *Phys. Med. Biol.* 58 (5), 1303–1314.
- de Geeter, N., Crevecoeur, G., Dupré, L., Van Hecke, W., Leemans, A., 2012. A DTI-based model for TMS using the independent impedance method with frequency-dependent tissue parameters. *Phys. Med. Biol.* 57, 2169–2188.
- Guyader, J.-M., Bernardin, L., Douglas, N., Poot, D., Niessen, W., Klein, S., 2014. Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion MR images of the abdomen. *J. Magn. Reson. Imaging* 42 (2), 315–330.
- Hallack, A., Chappell, M., Gooding, M., Schnabel, J., 2014. A new similarity metric for groupwise registration of variable flip angle sequences for improved T_{10} estimation in DCE-MRI. In: *Proceedings of WBIR 2014. Lecture Notes in Computer Science*. Springer, pp. 154–163.
- Hamy, V., Dikaio, N., Punwani, S., Melbourne, A., Latifoltojar, A., Makanyanga, J., Chouhan, M., Helbren, E., Menys, A., Taylor, S., Atkinson, D., 2014. Respiratory motion correction in dynamic MRI using robust data decomposition registration—application to DCE-MRI. *Med. Image Anal.* 18 (2), 301–313.
- Huizinga, W., Metz, C., Poot, D., de Groot, M., Niessen, W., Leemans, A., Klein, S., 2013. Groupwise registration in diffusion weighted MRI for correcting subject motion and eddy current distortions using a PCA based dissimilarity metric. In: *Proceedings of Computational Diffusion MRI and Brain Connectivity – MICCAI Workshop*, pp. 163–174.
- Huizinga, W., Poot, D., Guyader, J.-M., Smit, H., van Kranenburg, M., van Geuns, R.-J., Uitterdijk, A., van Beusekom, H., Coolen, B., Leemans, A., Niessen, W., Klein, S., 2014. Non-rigid groupwise image registration for motion compensation in quantitative MRI. In: *Proceedings of Workshop on Biomedical Image Registration*, pp. 184–193.
- Jeurissen, B., Leemans, A., Tournier, J., Jones, D., Sijbers, J., 2013. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Hum. Brain Map.* 34 (11), 2747–2766.
- Jones, D., Leemans, A., 2011. Diffusion tensor imaging. *Methods Mol. Biol.* 711, 127–144.
- Klaassen, R., Gurney-Champion, O., ter Voert, E., Heerschap, A., Bijlsma, M., Besselink, M., van Tienhoven, G., C.Y. Nio, J., Punt, C., Wilminck, J., van Laarhoven, H., Nederveen, A., 2014. Motion correction of high temporal 3T dynamic contrast enhanced MRI of pancreatic cancer – preliminary results. In: *Proceedings of the 22st Annual Meeting International Society for Magnetic Resonance in Medicine*, p. 3667.
- Klein, S., Pluim, J.P., Staring, M., Viergever, M., 2009. Adaptive stochastic gradient descent optimisation for image registration. *Int. J. Comput. Vis.* 81 (3), 227–239.
- Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J., 2010. elastix: a toolbox for intensity based medical image registration. *IEEE Trans. Med. Imaging* 29 (1), 196–205.
- Leemans, A., Sijbers, J., Backer, S.D., Vandervliet, E., Parizel, P., 2006. Multiscale white matter fiber tract coregistration: a new feature-based approach to align diffusion tensor data. *Magn. Reson. Med.* 55, 1414–1423.
- Liu, X., Neithammer, M., Kwitt, R., McCormick, M., Aylward, S., 2014. Low-rank to the rescue – atlas-based analysis in the presence of pathologies. In: *Proceedings of MICCAI 2014. Lecture Notes in Computer Science*, vol. 8673.
- Mangin, J.-F., Poupona, C., Clark, C., Bihana, D.L., Bloch, I., 2002. Distortion correction and robust tensor estimation for MR diffusion imaging. *Med. Image Anal.* 6 (3), 191–198.
- Melbourne, A., Atkinson, D., White, M., Collins, D., Leach, M., Hawkes, D., 2007. Registration of dynamic contrast-enhanced MRI using a progressive principal component registration (PPCR). *Phys. Med. Biol.* 52 (17), 5147–5156.
- Messroghli, D., Radjenovic, A., Kozerke, S., Higgins, D., Sivananthan, M., Ridgway, J., 2004. Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart. *Magn. Reson. Med.* 52, 141–146.
- Metz, C., Klein, S., Schaap, M., van Walsum, T., Niessen, W., 2011. Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. *Med. Image Anal.* 15 (2), 238–249.
- Miller, E., Matsakis, N., Viola, P., 2000. Learning from one example through shared densities on transforms. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 464–471.
- Muzi, M., O'Sullivan, F., Mankoff, D., Doot, R., Pierce, L., Kurland, B., Linden, H., Kinahana, P., 2012. QIN: quantitative assessment of dynamic PET imaging data in cancer imaging. *Magn. Reson. Imaging* 30 (9), 1203–1215.
- Parker, G., Roberts, C., Macdonald, A., Buonaccorsi, G., Cheung, S., Buckley, D., A.Jackson, Y.Watson, Davies, K., Jayson, G., 2006. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magn. Reson. Med.* 56 (5), 993–1000.
- Peng, Y., Ganesh, A., Wright, J., Ma, Y., 2010. RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images. In: *Proceedings of IEEE conference on Computer Vision and Pattern Recognition, CVPR*.
- Poot, D., Klein, S., 2015. Detecting statistically significant differences in quantitative MRI experiments, applied to diffusion tensor imaging. *IEEE Trans. Med. Imaging* 34 (5), 1164–1176.
- Poot, D., Kotek, G., Niessen, W., Klein, S., 2013. Bias correction of maximum likelihood estimation in quantitative MRI. *Proc. SPIE* 8669 86691F-1–86691F-6.
- Rao, C., 1946. Minimum variance and the estimation of several parameters. *Math. Proc. Camb. Philos. Soc.* 43, 280–283.
- Reijmer, Y., Leemans, A., Heringa, S., Wieland, I., Jeurissen, B., Koek, H., Biessels, G., Wang, Z., 2012. Improved sensitivity to cerebral white matter abnormalities in Alzheimer's disease with spherical deconvolution based tractography. *PLoS One* 7, 1371.
- Rossi, A., Dharampal, A., Wragg, A., Davies, L., van Geuns, R., Anagnostopoulos, C., Klotz, E., Kitslaar, P., Broersen, A., Mathur, A., Nieman, K., Hunink, M., de Feyter, P., Petersen, S., Pugliese, F., 2014. Diagnostic performance of hyperaemic myocardial blood flow index obtained by dynamic computed tomography: does it predict functionally significant coronary lesions? *Eur. Heart J. – Cardiovasc. Imaging* 15 (1), 85–94.
- Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D., 1999. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging* 18, 712–721.
- Sijbers, J., den Dekker, A., Raman, E., Dyck, D.V., 1999. Parameter estimation from magnitude MR images. *Int. J. Imaging Syst. Technol.* 10, 109–114.
- Takaya, N., Yuan, C., Chu, B., Saam, T., Underhill, H., Cai, J., Tran, N., Polissar, N., Isaac, C., Ferguson, M., Garden, G., Cramer, S., Maravilla, K., Hashimoto, B., Hatsukami, T., 2006. Association between carotid plaque characteristics and subsequent ischemic cerebrovascular events: a prospective assessment with MRI – initial results. *Stroke* 37 (3), 818–823.
- Thévenaz, P., Unser, M., 2000. Optimization of mutual information for multiresolution image registration. *IEEE Trans. Med. Image Process.* 9 (12), 2083–2099.
- Tofts, P., 2003. Quantitative MRI of the Brain: Measuring Changes Caused by Disease. John Wiley and Sons.
- Tofts, P., Brix, G., Buckley, D., Evelhoch, J., Henderson, E., Knopp, M., Larsson, H., Lee, T.-Y., Mayr, N., Parker, G., Port, R., Taylor, J., Weisskoff, R., 1999. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: Standardized quantities and symbols. *J. Magn. Reson. Imaging* 10, 223–232.
- Tournier, J., Mori, S., Leemans, A., 2011. Diffusion tensor imaging and beyond. *Magn. Reson. Med.* 65 (6), 1532–1556.
- van der Aa, N., Leemans, A., Northington, F., van Straaten, H., van Haastert, I., Groenendaal, F., Benders, M., de Vries, L., 2011. Does diffusion tensor imaging-based tractography at 3 months of age contribute to the prediction of motor outcome after perinatal arterial ischemic stroke. *Stroke* 42, 3410–3414.
- van der Aa, N., ter Morsche, H., Mattheij, R., 2007. Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem. *Electr. J. Linear Algebra* 16, 300–314.
- Wachinger, C., Navab, N., 2013. Simultaneous registration of multiple images: similarity metrics and efficient optimization. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (5), 1221–1233.
- Wang, H., Hsu, J., Leemans, A., 2012. Diffusion tensor imaging of vascular Parkinsonism: structural changes in cerebral white matter and the association with clinical severity. *Arch. Neurol.* 69, 1340–1348.
- Xue, H., Shah, S., Greiser, A., Guetter, C., Littmann, A., Jolly, M., Arai, A., Zuehlsdorff, S., Guehring, J., Kellman, P., 2012. Motion correction for myocardial T1 mapping using image registration with synthetic image estimation. *Magn. Reson. Med.* 67 (6), 1644–1655.