Westfälische
Wilhelms-Universität
Münster

Fachbereich Mathematik und Informatik

Diplomarbeit in Mathematik

Mass-preserving Registration of Medical Images

eingereicht von
Lars Ruthotto

Münster, 05. März 2010

Gutachter
Prof. Dr. Martin Burger
Priv.-Doz. Dr.Carsten Wolters

Abstract

Image Registration is one of today's most challenging tasks in medical imaging. This thesis develops and investigates tailored registration approaches for two important problems in real life medicine demanding mass-preserving transformations. A variational approach to susceptibility correction of echo planar images (EPI) is proposed, validated on phantom data and applied to functional magnetic resonance images (fMRI) and diffusion tensor images (DTI). Secondly, gated positron emission tomography (PET) images are corrected for respiratory motion using a novel non linear mass preserving registration algorithm.

Acknowledgments

I want to thank everyone that made this thesis possible and especially...

- Carsten Wolters without whom I would definitely never have written a thesis about medical imaging. Being my ‘HiWi’-boss he accompanied and supported me all the way, gave me the right literature and the right problems.

- Martin Burger for his experience and assistance whenever I was struggling with discretizations and function spaces.

- Fabian Gigengack who - as my precedessor - introduced me to image registration, returned - as a PhD student - with the PET data of Chapter 3 and quickly created the VAMPIRE together with me.

- Harald Kugel who patiently explained me everything about spins, echoes and much more MRI stuff and measured the interesting MR images in Chapter 2.

- Bernd Fischer’s group from the institute of mathematics in LÃ¼beck. Especially Janine Olesch for the good cooperation that led to the EPI correction approch presented in Chapter 2, the proceedings [22] and hopefully many more nice projects.

- Jan Modersitzki for making his *Flexible Algorithm for Image Registration FAIR* toolbox freely available.
- Jann-Philipp Zocher, Felix Lucka, Johannes Vorwerk and Sven Wagner for their corrections and annotations that made this thesis readable.

- My parents Dagmar and Johannes for their mental and financial support at all times.

Contents

Introduction

Image registration is one of today's most challenging tasks in medical imaging. Its main objective is to make two images that are recorded at different points in time, with different perspectives, by different devices and/or of different subjects more alike, cf. [12]. There is a variety of clinical and scientifical problems that rely on an accurate image registration. One example is the combination of information from multiple images, acquired using different modalities, like PET and CT. This procedure is also referred to as *image fusion*. One uni-modal example is the evaluation of a medical treatment, where images taken before and after surgery have to be compared.

From the mathematical point-of-view image registration is also a challenging and interesting problem. According to Hadamard a problem is *well posed*, if there exists an unique solution that furthermore depends continuously on the data. Otherwise a problem is called *ill-posed*. Note that there are in general many - even trivial - ways to align the two images. Due to this under-determinedness, image registration is a typical example for an ill-posed inverse problem, cf. [11].

The under-determinedness in image registration led to the inclusion of more and more prior knowledge into the registration task.

Similarity vs. Regularity

In image registration one always faces the same trade-off: On the one hand one wants to render both images as similar as possible while on the other hand a meaningful and regular transformation is desired. Therefore, it is at least as interesting to study the transformation as to compare the resulting images. In the current thesis we aim to ensure the regularity of the transformation by the use of regularisation techniques.

Mass-Preserving Registration

In its very sense registration means finding a geometric transformation y that optimally aligns the points in the *template* image $\mathcal{T}$ to the corresponding points in the *reference* image $\mathcal{R}$. The space in which $\mathcal{R}$ is taken is then called the *reference frame*. Once both images are in one common frame, their content can be compared or combined.

In general, the transformation may affect the total amount of intensity - that is the sum over all grey values - of the template image $\mathcal{T}$. For there are registration problems

implicitly or explicitly demanding such a *mass-preservation*, an extension of the existing registration theory is desirable.
Mass preserving registration problems have been studied using techniques of optimal mass transport by Rehman et al. in [25] and by Zhu et al in [34]. In contrast, this thesis chooses addresses this problem with image registration tools.

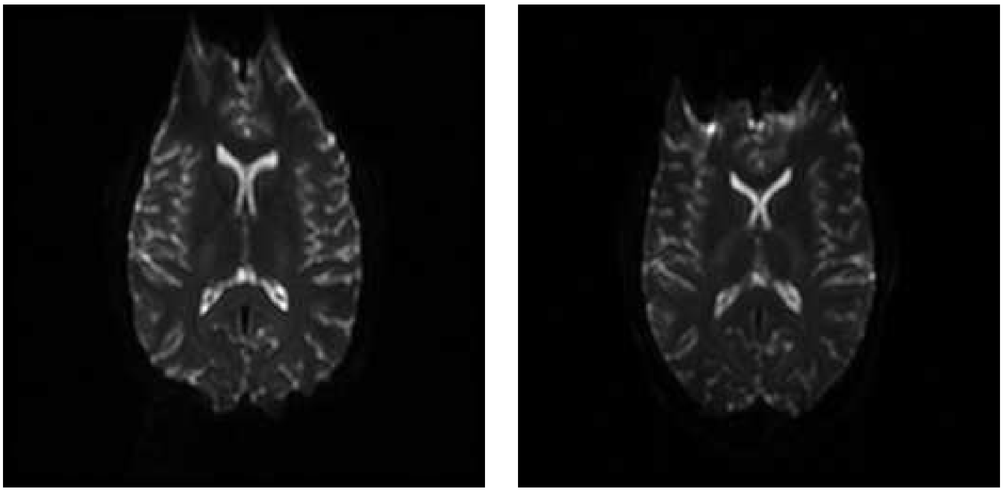
Scope of this thesis

The scope of this thesis is to present two mass-preserving registration problems occurring in medical imaging. For both problems we will propose and analyse a tailored non-linear variational registration approach.

Susceptibility correction of EPI

Chapter 2 is devoted to the correction of *echo planar images* (EPI). This ultra fast *magnetic resonance imaging* sequence is very sensitive to field inhomogeneities caused for instance by susceptibility differences of the object being imaged. These field inhomogeneities affect the spatial reconstruction. The displacement occurs mainly along the *phase encoding direction* and altering the gradients sign inverts the displacement. Because the lost signal can not be regained a mass-preserving transformation along the phase-encoding direction has to be estimated out of two reference scans with altered gradient signs. An example for this is depicted below in Figure 0.1.

We will study this problem from a variational view point. After providing an existence



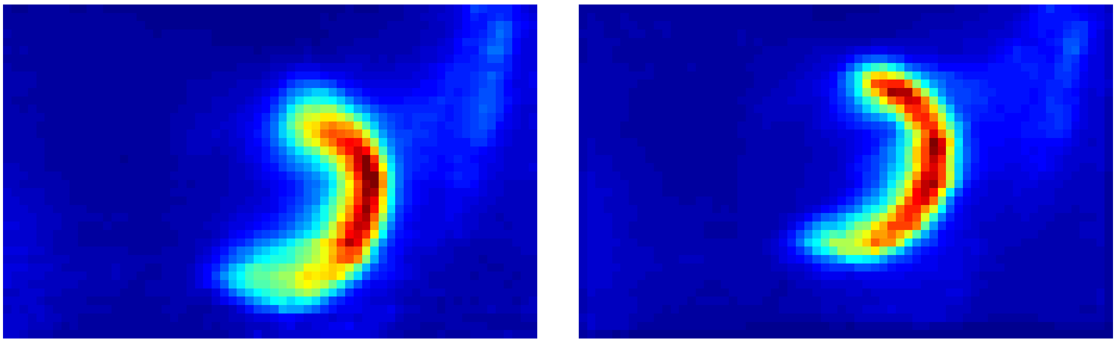
(a) EPI image with positive phase encoding gradient (b) EPI image with negative phase encoding gradient

Figure 0.1: Transversal slice of two echo planar images acquired with altered gradient sign. Note, that the inhomogeneity affects the localization along the phase-encoding direction, which is here along the anterior-posterior axis - i.e. from bottom to top.

theory for the minimization problem in the Sobolev space $W^{1,2}$, we will implement the algorithm as an extension to the freely available *Flexible Algorithms for Image Registration* (*FAIR*) toolbox by Jan Modersitzki in *Matlab* [19]. Finally, we will test the algorithm on three EPI datasets and explain how it might be used for correcting two commonly used imaging sequences, namely *diffusion tensor imaging* (DTI) and *functional magnetic resonance imaging* (fMRI).

Motion Correction in Cardiac Gated PET

In Chapter 3 we will try to correct *respiratory gated positron emission tomography images* for motion. In order to improve the quality and accuracy, images showing different phases of the subject's breathing cycle have to be accurately registered into a common frame. n this common frame the image can then be averaged, resulting in a much sharper image than in Figure 0.2 on the left. For the images of the different phases are acquired within the same time interval, the amount of tracer uptake in each of the images is equal. Hence, it is important to use a mass-preserving transformation model.



(a) PET reconstruction - without motion correction

(b) Gated PET image after motion correction

Figure 0.2: PET images of a human heart. Without gating and motion correction the heart is blurred, due to respiratory and cardiac movement during the acquisition.

The first attempt to register the images with an elastic non-linear registration approach will lead us to the direct inclusion of the mass-preservation into the registration task. We will call the proposed algorithm `VAMPIRE` for VAriational Mass-Preserving Image REgistration. By means of non-linear calculus of variations we will also prove existence in the Sobolev space $W^{1,6}$.

1 Image Registration

Given two images - the *template* image $\mathcal{T}$ and the *reference* image $\mathcal{R}$ - image registration means finding a *reasonable* transformation that - applied to the template image - makes both images more *alike*. This chapter will translate this rather vague formulation of the registration problem into mathematical terms. However, the discussion is restricted to those elements used later on. The interested reader is referred to [18, 19, 23] for excellent and more general presentations of image registration.

Mathematically, images can be represented and handled as functions which will be our starting point in Section 1.1. One way to distinguish registration approaches, is by the used transformation model. The most important class for this thesis, are *non-parametric* approaches where the transformation y is not defined by a set of basis functions and can move each point into an arbitrary direction.

In order to measure how *alike* two images are, we need a distance functional $\mathcal{D}$. Since both registration problems discussed in this thesis are *uni modal* - i.e. their intensity values are correlated - the *sum-of-squared distance* (SSD) and the *normalized-cross-correlation*($\mathcal{NCC}$) are adequate choices and described in Section 1.2.

As mentioned above, we are interested in finding a *reasonable* transformation. What this means depends of course on the specific context and deformation model. Solving the registration problem, we have to trade off the similarity of the images against a meaningful transformation. We measure the smoothness of our transformation y by a functional $\mathcal{S}[y]$ in order to punish unwanted behaviour. This leads to a bias to a desired solution, and can furthermore render the problem less under-determined and thus easier to solve. In our later applications, we will make use of the elastic potential, that is described in Section 1.3.

In the following chapters of this thesis we will include prior knowledge about the shape and the mass-preserving property of the transformation into the registration problem and also prove the existence of optimal transformations. This can be handled by means of variational calculus in an elegant fashion. Therefore, we formulate the registration problem in a variational formulation

$$\min_y \mathcal{J}[y] := \mathcal{D}(\mathcal{T}(y), \mathcal{R}) + \alpha \mathcal{S}[y]. \tag{1.1}$$

In this formulation the regularization parameter $\alpha \in \mathbb{R}^+$ balances between minimizing the data term $\mathcal{D}[y]$ and keeping the deformation meaningful which is judged by $\mathcal{S}[y]$.

It remains to describe how to find an optimal transformation. Section 1.4 describes the *Gauss-Newton* optimization and the *multi-level* strategy implemented in the *FAIR* toolbox [19] to solve registration problems like (1.1). Following a *discretize-then-optimize* strategy, we first discretize the functional before optimizing it. Consequently, each element's discretization is described in the corresponding section.

1.1 Representation of Images

This section introduces images as functions. We will say an image is *continuous* if it is of infinite resolution or more precisely, a function on $\mathbb{R}^3$. However in most applications the resolution is finite and therefore *digital* or *discrete* images/functions are interesting to study, too. In order to support the distinction between continuous and discrete settings we use capitalized calligraphic letters for continuous and capitalized Latin letters for discrete images. Discrete images are typically defined on *cell-centred* grids. We also explain how to evaluate a discrete image on non-grid points by interpolation methods and finally describe *rigid* and *non-parametric* transformations.

Definition 1 (Image)**.** *Let $d \in \mathbb{N}$. A function $\mathcal{I} : \mathbb{R}^3 \rightarrow \mathbb{R}$ is called a three-dimensional grey image, if its support $\Omega \subset \mathbb{R}^3$ is compact and $0 \leq \mathcal{I}(x) < 256$ for all $x \in \mathbb{R}^3$*
The space of all images is denoted by Img$(\mathbb{R}^3)$.

The *image domain* will be denoted by Ω in the remainder of this work. By definition we can treat images as functions. Therefore regularity - e.g. differentiability, measurability - of images is defined in the same way as for functions.

Definition 2 ($\mathcal{L}_2$ norm)**.** *Let $\mathcal{I}$ be an d-dimensional image. Then the $\mathcal{L}_2$-Norm is given by*

$$||\mathcal{I}||_2 := \left(\int_{\mathbb{R}^3} \mathcal{I}(x)^2 dx \right)^{\frac{1}{2}}.$$

Being finite and compactly supported, all reasonable images are elements of $\mathcal{L}_2$ which is a very weak choice for a function space. When calculating derivatives of an image, one would want the image to be continuously differentiable. However, interesting images can, because of their edges, not even be continuous. That is why weaker choices of subspaces of $\mathcal{L}_2$ - for instance $\mathcal{I} \in W^{j,p}$, see [10], or $\mathcal{I} \in$ BV, see [3] - are common in mathematical image processing.
The discussion of continuous images is especially useful in proofs, however, in image processing all relevant images are *digital images* that are of finite resolution. Commonly, the image data is given on a regular spaced *cell-centred grid.*

Definition 3 (Grid)**.** *Let the image domain $\Omega :=]0,1[^3$, and $m_1, m_2, m_3 \in \mathbb{N}$ be the number of discretization points. The points*

$$x_{j_1,j_2,j_3} = (x_{j_1}, x_{j_2}, x_{j_3})^T \in \Omega \cup \partial\Omega, 1 \leq j_l \leq m_l \text{ for } l = 1, 2, 3$$

are called grid points*. The array*

$$X := (x_{j_1,j_2,j_3})_{1 \leq j_l \leq m_l,\, l=1,2,3} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$$

is called grid matrix

Implementation 1 (Grids in *FAIR*)**.** *In the* FAIR *toolbox grids are stored in column vectors in lexicographical ordering. Thus, the vector starts by the x_1 coordinates of all grid points. Then the x_2 coordinates are stored and finally the x_3 coordinates. The coordinate system is supposed to be right-handed, see [19].*

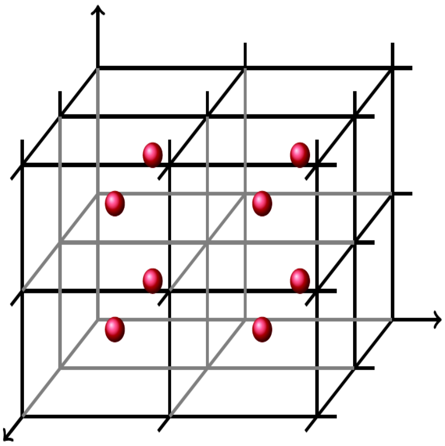


Figure 1.1: cell centred grid by Papenberg [23]

As pointed out above, discrete images are commonly discretized on cell-centred grids.

Example 1 (Cell-Centred Grid)**.** *Let* $\Omega :=]0,1[^3$ *and* $m_1, m_2, m_3 \in \mathbb{N}$ *be some given numbers. For* $1 \leq j_l \leq m_l$ *with* $l = 1, 2, 3$ *the grid points*

$$x_j^{cc} = \left(\frac{2_{j_1} - 1}{2m_1}, \frac{2_{j_2} - 1}{2m_2}, \frac{2_{j_3} - 1}{2m_3} \right)$$

build a cell centred grid or Neumann grid. It is depicted in Figure 1.1.

Definition 4 (Discrete Image)**.** *Let* $n_x, n_y, n_z \in \mathbb{R}^+$ *and* $\Omega = [0, n_x] \times [0, n_y] \times [0, n_z]$ *be the image domain and* $m_x, m_y, m_z \in \mathbb{N}$ *be the number of discretization points along the three dimensions. A* digital *image* I *then corresponds mathematically to a three dimensional real matrix with positive numbers- i.e.* $I \in \mathbb{R}^{m_x \times m_y \times m_z}$*. The elements of* I *are then the grey values at centre of the respective grid cell.*

Remark 1 (Digital images as discrete functions)**.** *When evaluating a continuous image on a cell-centred grid, that is*

$$I(j_1, j_2, j_3) = \mathcal{I}(x_{j_1, j_2, j_3}) \; \forall j_1, j_2, j_3,$$

one can define a projection from the space of all continuous images to the space of all discrete images.

Implementation 2 (Matrix vs. Vector representation)**.** *In* Matlab *images can either be stored in a three dimensional matrix of doubles, or in a row vector. We will make use of both notations. Let* `R`$\in \mathbb{R}^{m_1 \times m_2 \times m_3}$ *be an image array. Then we can get the row vector representation by*

```
R = R(:);
```

To regain the matrix representation we can use

```
R = reshape(R,[m1 m2 m3]);
```

Interpolation

In order to evaluate discrete images at non-grid points we have to use an interpolation scheme. Although there are many possible operators, we restrict the presentation to the linear interpolation operator $\mathcal{I}^{\text{lin}}$ and the 3D cubic spline interpolation $\mathcal{I}^{\text{spline}}$. The interpolation operator defines - based on the discrete image data I - a continuous image on $\Omega \subset \mathbb{R}^3$ which we then can evaluate on non-grid points. For a detailed overview on the different operators and their implementation, see [18, 19, 23].

Definition 5 (Linear 3D Interpolation)**.** *Let I be a digital image, given on a cell centred grid x^{cc}. The linear interpolant $\mathcal{I}^{lin}$ is then given by*

$$\mathcal{I}^{lin}(x;I) = \sum_{k\in\{0,1\}^3} I\left(\frac{\lfloor n_1x_1\rfloor + k_1}{n_1}, \frac{\lfloor n_2x_2\rfloor + k_2}{n_2} + \frac{\lfloor n_3x_3\rfloor + k_3}{n_3}\right)$$
$$\cdot \prod_{j=1}^{3}\left(\frac{(-1)^{k_j}}{n_j}(\lfloor x_jn_j\rfloor + 1 - k_j - x_jn_j)\right).$$

The linear interpolant is not differentiable at grid points. However, in order to achieve fast convergence we want to use a fast optimization scheme with descent directions depending on derivatives of our functional. Those in turn require derivatives of the template image. Especially when starting the optimization with an undeformed grid an interpolation scheme with higher regularity is desired. To this end we introduce the 3D cubic spline interpolation.
The idea is to have a finite set of cubic basis functions, where each function b_j is a translated version of a so-called *mother spline* b. We will first describe the interpolation in one dimension. The interpolation in three dimensions can then be composed, easily.

Definition 6 (Cubic spline basis in 1D)**.** *Let x^{cc} be a cell-centred grid in one dimension. Without loss of generality the cell-centred grid points x_j are equal to j. The mother spline b is then defined as*

$$b(x) = \begin{cases} (x+2)^3 & ,-2 \le x < -1, \\ -x^3 - 2(x+1)^3 + 6(x+1) & ,-1 \le x < 0, \\ x^3 + 2(x-1)^3 - 6(x-1) & ,0 \le x < 1, \\ (2-x)^3 & ,1 \le x < 2, \\ 0 & ,\ \textit{else}\ . \end{cases} \tag{1.2}$$

The cubic spline basis is then given by $\mathcal{B} := \{b^j(x) \mid b^j(x) = b(x-j), j = 1, ..., m_1\}$.

In order to obtain a function out of a digital image I given on the cell-centred grid one expands the interpolant by

$$\mathcal{I}^{\text{spline}}(x) = \sum_{j=1}^{m_1} c_j b^j(x).$$

Where the coefficients c_j are chosen such that the interpolation condition holds - cf. [19] for implementational remarks -

$$I_j = \mathcal{I}^{\text{spline}}(x_j),\ \forall j = 1, ..., m_1.$$

For three dimensional data, the generalisation is now easy to perform.

Definition 7 (Cubic spline interpolation in 3D)**.** *Let x^{cc} be a cell-centred grid and I be a discrete image on x^{cc}. Let the interpolation coefficients $c_j := c_{j_1,j_2,j_3}$ be chosen such that the interpolation condition is satisfied. Then the interpolant of I is a continuously differentiable function*

$$\mathcal{I}^{spline}(x) = \sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \sum_{j_3=1}^{m_3} c_j \prod_{l=1}^{3} b^{j_l}(x_l).$$

The i-th partial derivative is then given by

$$\partial_i \mathcal{I}^{spline}(x) = \sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \sum_{j_3=1}^{m_3} c_j (b^{j_i})'(x_i) \prod_{l \neq i} b^{j_l}(x_l).$$

Where each $(b^{j_i})'(x_i)$ can be computed from equation (1.2).

Even for our high-dimensional data, we will use the cubic B-spline interpolation scheme, since we are interested in a very accurate registration result. We could of course speed up the registration by using linear interpolation, but this might cause difficulties in the optimization algorithm, due to the non-differentiability at grid points. Especially, when starting the optimization with an undeformed grid, choosing spline interpolation should pay off and result in an accurate search direction.

Geometric Transformations

In order to complete the discussion about images, we explain how to apply a geometric transformation y to an image. In continuous setting geometric transformations are changes of the image's coordinate system. Thus, the geometric transformation y is a function from $\mathbb{R}^3$ to $\mathbb{R}^3$.

In discrete setting we can carry out a geometric transformation by interpolating on a transformed grid y^{cc} that is a discrete version of y - which is referred to as the *Eulerian approach* in comparison to the *Lagrangian approach*, described for instance in [9]. When visualizing the transformations, we will show the deformed grid at which we interpolate the template image.

When using elastic regularization in *FAIR* , the optimization variable y^{stg} is discretized on a *staggered grid* and denoted by `yc`.In contrast to the cell-centred grid, the x_1, x_2, x_3 components are obtained as a one dimensional nodal-grid of the respective interval. This has the advantage that the divergence of the deformation can be calculated via short differences. For a more detailed view on the different grid types in *FAIR* , we refer to [23].

Clearly, one can think of many different transformations. The presentation here is limited to two transformations we will use later on - the *rigid* and the *non-parametric* transformation. An overview of the possible transformations can be found for instance in [18, 19, 23].

Rigid Transformation

When registering two images of the head of the same subject, taken at different points in time, it is likely that the subject has moved in between. This movement can typically be described as a rotation and a shift, since the head is supposed to be rigid. In 3D the transformation can then be described by a parameter $\theta \in \mathbb{R}^6$. Therefore the rigid transformation is one example of a *parametric transformations.*
The first three elements of θ describe the rotations along the respective axes. They can be expressed as the product of the orthogonal matrices

$$R_{x_1}(\theta_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_1) & -\sin(\theta_1) \\ 0 & \sin(\theta_1) & \cos(\theta_1) \end{pmatrix}$$

$$R_{x_2}(\theta_2) = \begin{pmatrix} \cos(\theta_2) & 0 & -\sin(\theta_2) \\ 0 & 1 & 0 \\ \sin(\theta_2) & 0 & \cos(\theta_2) \end{pmatrix}$$

$$R_{x_3}(\theta_3) = \begin{pmatrix} \cos(\theta_3) & -\sin(\theta_3) & 0 \\ \sin(\theta_3) & \cos(\theta_3) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In combination with a translation, that is described by $(\theta_4, \theta_5, \theta_6)$ we end up with the following transformation

$$y(x;\theta) = (R_{x_1} \cdot R_{x_2} \cdot R_{x_3}) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} \theta_4 \\ \theta_5 \\ \theta_6 \end{pmatrix}$$

By definition y is then diffeomorphic - i.e. y and its inverse y^{-1} are continuously differentiable - which is a desired feature in image registration.

Non-Parametric Transformation

In contrast to the rigid registration that can be completely described by 6 parameters, we will mainly use *non-parametric* transformations. For each voxel can be moved arbitrarily, they are also called *free form transformations.* Clearly, non-parametric transformations are, in general, not diffeomorphic. Furthermore, the use of non-parametric transformations renders the registration problem under-determined. In order to obtain a transformation that fits to the physical model and to treat the under-determinedness, we introduce similarity measures that penalize certain features of the transformations in Section 1.3. Thereby we can sort out deformations that are not reasonable in the specific context.

1.2 Distance measure

As a second component of our registration problem (1.1) this section defines the distance $\mathcal{D}$ of two images. Generally stated, the distance between the two images can be defined as

$$\mathcal{D}(\mathcal{T},\mathcal{R}) = \int_\Omega w(x)\psi(r(\mathcal{T},\mathcal{R}))dx. \tag{1.3}$$

Here $r : Img^2(\Omega) \to Img(\Omega)$ is the residual or difference function, $w : \mathbb{R} \to \mathbb{R}^+$ is a weighting function and $\psi \in C^2(\mathbb{R},\mathbb{R})$. There are many specifications of r,w and ψ that are sensitive to different image features. As Modersitzki points out in [19, p. 115] there is no universal rule which distance measure to choose. He suggests, however, that the distance measure should be as restrictive as possible in the specific context.
In the applications discussed later the images show the same subject and are of the same modality. Thus, the intensities can be assumed to be correlated and we can compare them by intensity-based measures. We will make use of the sum of squared differences (SSD) or their normalized cross-correlation ($\mathcal{NCC}$) that are described below.

Sum of Squared Distances

The sum of squared differences (SSD) can be derived as a weighted $\mathcal{L}_2$ measure as follows.

Definition 8 (SSD Distance). *Let $\mathcal{T}$,$\mathcal{R}$ be images. Then by choosing*

$$r := \mathcal{T} - \mathcal{R} \text{ and } \psi(x) := \frac{1}{2}x^2$$

we obtain the sum of squared differences (SSD) measure

$$\mathcal{D}^{SSD}(\mathcal{T},\mathcal{R}) = \frac{1}{2}\int_\Omega w(x)(\mathcal{T}(x) - \mathcal{R}(x))^2dx.$$

When registering the template image $\mathcal{T} \in \mathrm{Img}(\Omega)$ to the reference image $\mathcal{R} \in \mathrm{Img}(\Omega)$, we can define a functional on the space of the geometric transformations by

$$\mathcal{D}^{\mathrm{SSD}}(y) = \frac{1}{2}\int_\Omega w(x)(\mathcal{T}(y(x)) - \mathcal{R}(x))^2dx. \tag{1.4}$$

When minimizing $\mathcal{D}^{\mathrm{SSD}}$ we want to apply a fast optimization scheme using derivatives. In the non-parametric case, we need its first variation with respect to y in order to calculate the *direction of steepest descent.*

Theorem 1 (First variation). *Let $\mathcal{T} \in C^1(\Omega,[0,256])$ and $\mathcal{R} \in Img(\Omega)$ be images. Then the first variation of $\mathcal{D}^{SSD}$ (1.4) is given by*

$$\mathcal{D}'(y) = w(x)(\mathcal{T}(y) - \mathcal{R})\nabla\mathcal{T}(y). \tag{1.5}$$

Proof. Let v be a transformation. We calculate the Gateaux derivative of the functional $\mathcal{D}^{\text{SSD}}$ at point y.

$$\begin{aligned} d\mathcal{D}(y,v) &= \lim_{\epsilon \to 0} \frac{1}{\epsilon}(\mathcal{D}(y+\epsilon v) - \mathcal{D}(y)) \\ &= \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_\Omega w(x)(\mathcal{T}(y+\epsilon v) - \mathcal{R})^2 - (\mathcal{T}(y) - \mathcal{R})^2 dx \end{aligned}$$

With a first order Taylor expansion we obtain

$$\begin{aligned} &= \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_\Omega 2w(x)(\mathcal{T}(y) - \mathcal{R})(\epsilon \nabla \mathcal{T}(y) v) + O(\epsilon^2) dx \\ &= \int_\Omega w(x)(\mathcal{T}(y) - \mathcal{R}) \nabla \mathcal{T}(y) v dx. \end{aligned}$$

Then

$$-\mathcal{D}'(y) = -w(x)(\mathcal{T}(y) - \mathcal{R})\nabla \mathcal{T}(y)$$

is also called the *force field* - see [19] - of $\mathcal{D}^{\text{SSD}}$. ☐

Implementation 3 (SSD Distance). *The discretization in* FAIR *is broken down into two steps. First the occurring integrals are discretized by a simple midpoint quadrature. Therefore the volume of a voxel is defined by* `hd = h(1) * h(2) * h(3)`. *The distance between the two digital images in vector shape* `I1` *and* `I2` *is then given by*

```
D = 0.5 * hd * (I1-I2)' * (I1-I2);
```

The first order derivative can be computed as

```
dD = hd * (I1-I2)' * gradI1;
```

Here `gradI1` *is provided by the chosen interpolation scheme.*

Normalized Cross Correlation

In Chapter 3 we will evaluate the registration result by means of the normalized cross-correlation ($\mathcal{NCC}$) of the transformed template and the reference to be conform to previous approaches. However, there is a strong link between the minimization of the SSD and the maximization of the normalized cross-correlation.

Definition 9 (Normalized Cross Correlation). *Let* $\mathcal{T}, \mathcal{R} \in Img(\mathbb{R}^3)$ *be two three dimensional images. The* $r(T,R) \in Img(\mathbb{R}^3)$ *is then given by*

$$r(\mathcal{T},\mathcal{R}) = \frac{\mathcal{T} - \mu(\mathcal{T})}{\sigma(\mathcal{T})} \cdot \frac{\mathcal{R} - \mu(\mathcal{R})}{\sigma(\mathcal{R})}$$

With $\mu(\mathcal{I}) := \int_{\mathbb{R}^3} \mathcal{I}(x)dx$*,* $\sigma^2(\mathcal{I}) := \mu((\mathcal{I} - \mu(\mathcal{I}))^2)$ *and the following choices in* (1.3)

$$\psi(x) = x^2 \text{ and } w(x) = 1$$

we obtain the normalized cross correlation distance measure by

$$\mathcal{NCC} := \int_{\mathbb{R}^3} \left(\frac{\mathcal{T} - \mu(\mathcal{T})}{\sigma(\mathcal{T})} \cdot \frac{\mathcal{R} - \mu(\mathcal{R})}{\sigma(\mathcal{R})} \right)^2 dx$$

Remark 2. *Since*

$$2\mathcal{D}^{SSD}[\mathcal{T}, \mathcal{R}] = ||\mathcal{T}||^2_{\mathcal{L}_2} + |\mathcal{R}||^2_{\mathcal{L}_2} - (\mathcal{T}, \mathcal{R})_{\mathcal{L}_2}$$

and

$$\begin{aligned} \sigma(\mathcal{T})\sigma(\mathcal{R}) \cdot \mathcal{NCC} &= (\mathcal{T} - \mu(\mathcal{T}), \mathcal{R} - \mu(\mathcal{R}))_{\mathcal{L}_2} \\ &= (\mathcal{T}, \mathcal{R})_{\mathcal{L}_2} - \mu(\mathcal{T})\mu(\mathcal{R}) \end{aligned}$$

there is a link between minimizing the SSD and maximizing the $\mathcal{NCC}$*. For a rigid transformation we have* $\det(Dy) = 1$ *and therefore* $||\mathcal{T}||^2_{\mathcal{L}_2}$*,*$||\mathcal{R}||^2_{\mathcal{L}_2}$*,*$\sigma(\mathcal{T}), \sigma(\mathcal{R}), \mu(\mathcal{T})$ *and* $\mu(\mathcal{R})$ *are constants. Then the minimization of* $\mathcal{D}^{SSD}$ *is equivalent to the maximization of* $\mathcal{NCC}$ *[18].*

1.3 Regularization

In order to render the non-parametric image registration problem less ill-posed and favour physically meaningful transformations, we define the similarity functional $\mathcal{S}$. Apart from rendering the problem more convex, we can also incorporate prior knowledge about the transformation that happened between both images. In the following applications we make use of the linear elastic potential. The assumption then is, that the image is painted on an elastic domain. As most parts of the body are deformed elastically the elastic potential is a common choice in medical image registration.
In general, a continuous $\mathcal{L}_2$ based regularizer is defined as

$$S[y - y_{\text{Ref}}] = \frac{1}{2} \int_{\Omega} |\mathcal{B}[y - y_{\text{Ref}}]|^2 dx \tag{1.6}$$

Here $\mathcal{B}$ is a differential operator. In this way we can favour transformations that are similar to y_{Ref}, which is typically chosen as the identity - then we penalize updates - or the resulting transformation of a parametric preregistration.

Definition 10 (Elastic potential)**.** *With the so-called Lamé constants* $\mu, \lambda \in \mathbb{R}^+$*, the elastic differential operator is defined as*

$$\mathcal{B} = \begin{pmatrix} \sqrt{\mu}\nabla & 0 & 0 \\ 0 & \sqrt{\mu}\nabla & 0 \\ 0 & 0 & \sqrt{\mu}\nabla \\ \sqrt{\lambda + \mu}\partial_1 & \sqrt{\lambda + \mu}\partial_2 & \sqrt{\lambda + \mu}\partial_3 \end{pmatrix}$$

Inserting the differential operator into (1.6) leads to the elastic potential

$$\mathcal{S}[u] = \frac{1}{2}\int_{\Omega} ||\mathcal{B}[u]||^2 dx = \frac{1}{2}\int_{\Omega} \mu(\nabla u, \nabla u) + (\lambda + \mu)(\nabla \cdot u)^2 dx$$

While the component wise gradient terms penalize oscillations in the transformation, the divergence is used to indicate volume changes. The derivative of the elastic potential can be computed as

$$d\mathcal{S}[u] = (\mathcal{B}u)^T \mathcal{B}.$$

And the second order derivative is given by

$$d^2\mathcal{S}[u] = \mathcal{B}^T \mathcal{B}.$$

Implementation 4 (Elastic potential)**.** *Similar to the discretization of the SSD distance* FAIR *approximates the integral by a simple midpoint quadrature on a cell-centred grid. The discretization of the differential operator $\mathcal{B}$ is done in two steps. First, the partial differential operators $\partial_k u^j$ are discretized individually as sparse matrices and then concatenated in order to build up $\mathcal{B}$. Especially for high-dimensional data, building the operator is quite slow. That is why the operator is stored as a* `persistent` *variable in the* FAIR *function* `regularizer.`*The matrix is updated only if its size does not correspond to the grid size any more. This happens in a multi-level approach, when changing from a coarser level to a finer one.*

In FAIR *one can reduce the memory consumption by using a matrix free implementation of the finite difference operators.*

1.4 Optimization

So far, we stated image registration as a minimization problem and described all occurring elements. For we follow a *discretize-then-optimize* strategy each element's discretization is already explained leading to a function $f(y) : \mathbb{R}^{3\cdot m_1 \cdot m_2 \cdot m_3} \rightarrow \mathbb{R}$, which is the discrete version of the functional $\mathcal{J}$ in (1.1). We will now go on in developing an optimization scheme for finding a minimizer y^* of the function $f(y)$. Then y^* is the transformation grid that - by interpolation of the template image at y^* - leads to an optimal alignment with respect to the chosen distance and regularization term. A very broad and detailed presentation of numerical optimization can be found in [21]. More information about the implementation in the context of image registration can be found in [19]. In order to find y^* we use a Gauss-Newton method that iteratively minimizes the real-valued function

$$f(y) = \psi(r(y)) + \alpha(\mathcal{B}y)^T \mathcal{B}y.$$

Here, the residual r and ψ are defined as described in Section 1.2 and $\mathcal{B}$ is the differential operator chosen for regularization.

As a starting guess we will either choose y_0 as the identity or equal to the result of a previous parametric registration. In the following, the iterates will be denoted by subindices, for instance $f_k := f(y_k)$.

The Gauss-Newton method belongs to the class of line search methods. That means that in the k-th iteration step one has to choose a search direction p_k first and then to solve the one-dimensional minimization problem

$$\min_{\beta_k} f(y_k + \beta_k p_k).$$

Then the $(k+1)$-th iterate is given as

$$y_{k+1} = y_k + \beta_k p_k.$$

The remainder of this section therefore describes the computation of the search direction p_k, explains the line search solvers and finally, the stopping criteria that are used in the following chapters.

Search Direction

We will determine the search direction by a Gauss-Newton method. In contrast to Newton's method, the goal is to use a numerically cheap approximation of the Hessian $\nabla^2 f$ in order to determine the search direction given by

$$p_k = -(\nabla^2 f)^{-1} \nabla f.$$

In our case, where in the distance functional $\psi(x) := \frac{1}{2} x^T x$, the gradient of our function f is given by

$$\nabla f(y_k) = r(y_k)^T \nabla r(y_k) + \alpha(\mathcal{B}(y_k - y_{\text{Ref}}))^T \mathcal{B}.$$

The second derivative becomes

$$\nabla^2 f(y_k) = \nabla r(y_k)^T \nabla r(y_k) + r(y_k)^T \nabla^2 r(y_k) + \alpha \mathcal{B}^T \mathcal{B}.$$

We will avoid computing the Hessian $\nabla^2 r(y_k)$ and approximate the second derivative by

$$\nabla^2 f(y_k) \approx \nabla r(y_k)^T \nabla r(y_k) + \alpha(\mathcal{B}(y_k - y_{\text{Ref}}))^T \mathcal{B}.$$

The error due to the approximation of the Hessian is relatively small whenever $r(y_k)$ is small - that is close to a solution - and whenever $\nabla^2 r(y_k)$ is small. The latter is the case in regions where r is relatively affine. A special advantage in image registration is, that the gradient of the residual usually contains image derivatives. By approximating the Hessian as shown above, we avoid computing second derivatives which might be affected by the noise present in the data.

For the computation of the search direction p_k we solve the following linear equation system

$$\nabla r(y_k)^T \nabla r(y_k) p_k = -\nabla r(y_k).$$

This will be done iteratively by the preconditioned conjugate gradient method.

Determining the step length

For a given search direction p_k, we now want to choose β_k a minimizer of the one-dimensional problem

$$\min_{\beta_k} \phi(\beta_k) = f(y_k + \beta_k p_k).$$

In solving this sub-problem we face the trade off that we want to find an β_k that leads to a significant reduction of our objective function f, but on the other hand we do not want to spend too much time searching for it. In general, it is too expensive to identify the global minimizer of ϕ and even finding a local minimizer of moderate precision would require too many evaluations of our objective function. Note that especially for three dimensional data, evaluating the objective function is computationally expensive.
Therefore we content ourselves with trying a sequence of candidates for β_k and stopping when a specified condition is satisfied. Here, we will describe three conditions. We will start by the Armijo condition and then present two extensions: the Wolfe and the strong Wolfe condition.

Definition 11 (Armijo condition)**.** *Let f, y_k, p_k be defined as above. Furthermore let $c_1 \in (0,1)$ be a small constant. Then a reasonable step lengths should lead to a sufficient decrease of the objective function, that is*

$$f(y_k + \alpha p_k) \leq f(y_k) + c_1 \alpha \nabla f_k^T p_k. \tag{1.7}$$

As proposed in Nocedal and Wright [21] a typical choice for c_1 is 10^{-4}. The Armijo condition alone is not enough to ensure a reasonable progress, because it is satisfied for all sufficiently small values of β_k. To rule out unacceptable short step length we introduce the curvature condition.

Definition 12 (Curvature condition)**.** *Let c_1 be the constant in Definition 11 and $c_2 \in (c_1, 1)$. The* curvature condition *requires β_k to satisfy*

$$\nabla f(y_k + \beta_k p_k)^T p_k \geq c_2 \nabla f_k p_k.$$

The *Wolfe condition* then is the combination of the Armijo condition and the curvature condition. Satisfying the Wolfe condition, however, does not necessarily mean that β_k is particularly close to a minimizer of ϕ.

Definition 13 (Strong Wolfe condition)**.** *Let c_1 and c_2 be chosen as above. Then the* Strong Wolfe condition *requires β_k to satisfy the Armijo condition and furthermore a modified curvature condition, that is*

$$|\nabla f(y_k + \beta_k p_k)^T p_k| \geq c_2 |\nabla f_k p_k|.$$

The advantage of the strong Wolfe condition is that the derivative $\phi'(\beta_k)$ can not be too positive any more. Therefore we exclude step length that are far away from a stationary point of ϕ.

Stopping Criteria

For $\rho_1, \rho_2, \rho_3 \in \mathbb{R}^+$ a tolerance level, $\epsilon \in \mathbb{R}^+$ the machine precision and $k_{\max}$ the maximum iterations the stopping criteria consists of five conditions. One indicator for a minimum is, that the update of the objective function f becomes very small

$$||f(y_{k+1}) - f_k|| \leq \rho_1(1 + |f(y_0)|). \tag{1.8}$$

Furthermore the $k+1$-th iterate of the optimization variable will be close to k-th iterate. This is controlled by the second condition

$$||y_{k+1} - y_k|| \leq \rho_2(1 + ||y_0||). \tag{1.9}$$

Because of the necessary condition the derivative of the objective function disappears close to a local minimum

$$||\nabla f(y_k)|| \leq \rho_3(1 + |f(y_0)|). \tag{1.10}$$

We will also stop the iteration, when the value of the gradient becomes smaller than the machine precision

$$||\nabla f(y_k)|| \leq \epsilon \tag{1.11}$$

or when the maximum number of iterations is exceeded

$$k \geq k_{\max} \tag{1.12}$$

To sum up, we have reached a local minimum when

$$((1.8) \text{ and } (1.9) \text{ and } (1.10)) \tag{1.13}$$

is true. Including (1.11) and (1.12) we end up with the stopping criterion

$$((1.8) \text{ and } (1.9) \text{ and } (1.10)) \text{ or } (1.11) \text{ or } (1.12) \tag{1.14}$$

Multi-Level Strategy

In order to obtain a fast and robust algorithm we make use of a multi-level strategy. This means that we solve the problem on a coarser grid first. Clearly, calculations on coarser levels are cheaper compared to fine levels. Additionally, details are neglected on coarser levels and the images are smoothed by a Gaussian kernel before interpolation. This improves the stability against local minima. After obtaining a transformation on a coarser level, we prolongate it and use it as a starting point on the next higher level. By these means we will take less risk to end up in a local minimum.

(a) Level 1 (b) Level 2 (c) Level 3 - Full Resolution

Figure 1.2: A typical multi-resolution strategy is depicted. From left to right, the resolution is each time doubled until we reach the full resolution.

Algorithm 1 Multi Level Image Registration with Gauss-Newton

Require: Template $\mathcal{T}$, Reference $\mathcal{R}$, start value y_0

Ensure: Transformation $y*$

```
for $l$ = minLevel to maxLevel do
    $y_0^l = \text{PROLONGATE}(y_{\text{end}}^{l-1})$
    while NOT STOP do
        $f_k = f(y_k, \mathcal{T}^l, \mathcal{R}^l)$
        $\nabla f_k = \nabla f(y_k, \mathcal{T}^l, \mathcal{R}^l)$
        $\nabla^2 f_k \approx (\nabla f_k)^T \nabla f_k + \alpha \mathcal{B}^T B$
        $p^k = -(\nabla^2 f_k)^{-1} \nabla f_k$
        $\beta_k = \text{LINESEARCH}(p^k)$;
        $y_{k+1} = y_k + \beta_k p_k$
        $k \to k+1$
    end while
end for
```

1.5 Summary

In the present chapter we presented image registration in its variational setting, i.e.

$$\min_y \mathcal{J}[y] := \mathcal{D}(\mathcal{T}(y), \mathcal{R}) + \alpha \mathcal{S}[y].$$

After explaining the different elements and their respective discretization we presented a tailored optimization scheme for image registration. Understanding of the presented framework will give us a perfect starting point for solving two real-life problems in medical imaging in the following two chapters.

In both chapters we will develop customized non-parametric registration approaches by incorporating prior knowledge about the specific shape of the transformation into the registration problem. In Chapter 2, that deals with the susceptibility correction of *echo-planar* magnetic resonance images, there a physical distortion model restricts the transformation to be along one axis. In Chapter 3 which is entitled *Motion Correction in gated*

PET we are searching for a non-linear transformation capable of aligning two images showing different phases of the human respiratory cycle.
In addition to the geometric shape of the distortions, we assume that in both cases no signal is lost due to the geometric deformations. This fact is something not yet accounted for by the framework presented in this chapter. In general, the deformed template will not have the same amount of signal or *mass*, i.e.

$$\int_{\mathbb{R}^3} \mathcal{T}(y(x))dx \neq \int_{\mathbb{R}^3} \mathcal{T}(x)dx.$$

As a matter of course, we want to modify the image registration problem to a minimal extent, such that we will still be able to use most of the tools presented in the above chapter.

2 Correction of Echo Planar Images

With our toolbox of mathematical image registration at hand, we will now address the problem of susceptibility distortions in *echo-planar images* (EPI). Since this ultra-fast MRI sequence is widely used in both medical research and clinical applications and the distortions heavily affect the spatial accuracy of those studies, there are already many approaches to this problem. Those approaches can mainly be divided into two classes:

- The *fieldmap* approaches measure the field-inhomogeneity directly by a reference scan and subsequently correct the acquired EPI image [15, 14, 1].
- In contrast to that, the *image-based* approaches make use of the *reversed-gradient* method [5, 26, 30, 2, 1, 20]. Here one compares the obtained echo planar image to an - apart from reversed gradients - identical reference scan. Some image based proposals also aim to register the EPI image to an undistorted structural - e.g. T2 image - of the subject, but taking the physical EPI distortion model - that will be derived later on - into account [28, 33].

Both types of correction approaches are deeply linked, since one can think of the image-based approaches as an inverse formulation of the fieldmap approaches. In [33] Wu et al. compared both kind of approaches and came to the conclusion that fieldmap approaches were superior in infratentorial regions including brainstem and cerebellum, as well as in ventral areas of the temporal lobes. Their image based approach led to superior results in all rostral brain regions.

In the present chapter, we will first explain the physical distortion model for echo-planar images and then derive a new variational image-based approach that is also sketched in [22].

To understand the physical deformation model we will shed some light on the basics of magnetic resonance imaging in Section 2.1. In Subsection 2.1 we will then discuss the effect of field inhomogeneities on spatial encoding in greater detail and formulate a forward model and the experimental setting for our inverse problem. Section 2.2 will enter knowledge about the shape of the deformation into a new functional. This functional will be minimized using the image registration techniques presented in chapter 1. In Section 2.3 we will apply our approach to three problems. After validating our algorithm on a phantom scan, we will give two examples how to apply it to real life problems occurring in diffusion tensor imaging and functional magnetic resonance imaging. We will also develop a complete correction pipeline for those two imaging techniques.

2.1 Basics on Magnetic Resonance Imaging

This section derives and explains the physical deformation model for echo planar imaging and motivates the *reversed gradient* method used later on.
Subsection 2.1 will introduce to MRI by explaining the components of an MRI system and some basic but important terms. Although there are many different imaging sequences we will tailor the presentation to understand echo planar imaging sequences.
Afterwards, Subsection 2.1 will explain how the *spatial encoding* methods used in our EPI sequences work in a perfectly homogeneous magnetic field. To understand how we determine the origin of a signal is of great importance, since the accurate mapping onto the image is desired.
In Subsection 2.1 the assumption of a homogeneous field will be dropped. We will explain the effect of a non-homogeneous field to the spatial encoding in EPI. Afterwards we will describe the mathematical forward model that simulates the distortion of an image measured with a given field-inhomogeneity. Finally, we will present the experimental set-up for our inverse correction algorithm.

Introduction to Magnetic Resonance Imaging

MRI, or *Nuclear Magnetic Resonance Imaging* (NMRI), is a non-invasive imaging technique. As a tomography technique MRI examines the subject slice wise using the magnetic resonance phenomenon of the protons e.g. the hydrogen nuclei which make up 10% of the body-weight [32]. Each proton has a spin which results in a microscopic magnetic moment μ that points in a random direction when the proton is situated in the earth's magnetic field. In this case, the macroscopic, or *bulk magnetization* M - i.e. the sum over all microscopic magnetic moments in a small volume - equals zero.
We will now briefly discuss the components of an MRI system and the principle of signal and echo generation. Since we will only focus on those aspects that are important for our forward distortion model, the interested reader is referred to [29] and [17] for a more detailed presentation.
To simplify our presentation below, we will assume the patient is placed within the scanner with the following alignment: The *left-right* axis equals exactly the x_1 axis. The *posterior-anterior* - i.e. from the back of the head to the brow - axis equals the x_2 axis and the *inferior-superior* - i.e. from foot to head- axis is equal to the x_3 axis. Sometimes we will also use the abbreviation $x = (x_1, x_2, x_3)$. Without loss of generality we assume that the slices are then oriented perpendicular to the x_3 axis (i.e. axial slices).
The voxel size will be denoted in millimetre by $h = (h_1, h_2, h_3)$.

Main Magnet

The main magnet produces a very strong *homogeneous* magnetic field B_0 pointing along the x_3-axis. Because of the non-zero spin the protons then *precess* around the x_3 axis. In

thermal equilibrium, their microscopic magnetic moments can align either parallel to B_0 or anti-parallel. Since the parallel state is the lower energy state slightly more protons align parallel by Boltzmann's law. Their amount is among other things positively dependent on the strength of B_0. The excess of protons in the lower energy state causes the bulk magnetization M of one small volume to point along the x_3-direction. In this case M equals its projection to the x_3-axis, denoted by M_{x_3}.
The *Larmor frequency* - the frequency at which the protons precess - is proportional to the strength of the magnetic field

$$\omega_0 = \gamma B_0. \tag{2.1}$$

Here, γ is the gyromagnetic ratio in [MHz/T] and B_0 is measured in Tesla (T). In our experiments the B_0 strength is 3.0T which is about 50.000 times the strength of the earth magnetic field.
The *resonance condition* states that the frequency at which a precessing magnet resonates to a RF-pulse equals its Larmor frequency.

RF-System

The second important component of an MR system is the *radio frequency* system. Its transmitter coil generates a rotating magnetic field B_1 orthogonal to B_0 with frequency ω_{rf}. The RF pulse enters a forced precession to all magnets that satisfy the resonance condition - i.e. precess at ω_{rf}. All other magnets remain untouched. Therefore, the bulk magnetization M is tipped away from the x_3-axis and gains a transverse component M_{x_1,x_2}.
The RF-system also contains a receiver coil that detects the precessing magnetization and converts it into a complex electric signal $S(t)$. This signal is called *Free Induction Decay* (FID) - that is the *induced* signal caused by *free* precession that shows a characteristic *decay* - and is the mother signal in MRI sequences. Let $\rho(x)$ be the density of the excited protons, whose accurate mapping is desired. For a not further specified field-inhomogeneity $\Delta B(x)$ the *echo signal* is given by

$$S(t) = \int_{\text{object}} \rho(x) e^{-i\gamma \Delta B(x)(t-T_E)} dx. \tag{2.2}$$

By varying the field-inhomogeneity in a controlled fashion - i.e. producing an inhomogeneity $\Delta B(x)$, we are able to spatially encode the signal.
The smaller angle between the x_3-axis and M is called *flip angle* and denoted by α. The flip angle depends on the magnitude of B_1 and the duration of the pulse. We use a 90° pulse that projects M_{x_3} completely into the $x_1 x_2$-plane.
After the pulse, the system returns to its thermal equilibrium state. This rather slow effect, is called *T1-* or *longitudual relaxation* and is accomplished by the much faster *transverse* (or T2) relaxation. Different choices of the repetition time T_R - the time between two RF-pulses - and the echo time T_E - the time between excitation and signal detection - lead to different signal intensities in specific voxels.

Gradient System

Three orthogonal coils and their amplifier produce a linearly varying magnetic field $B(x) := Gx$, with G being a constant field. Its strength is much smaller than the main magnets and in the order of 30-80 mT/m in clinical scanners [17], i.e. a difference of 7.5 to 20 mT on the edge of a ±25 cm field-of view.
To simplify notation, we will define a x_i-gradient to be a gradient that points into the x_i-direction, for $i = 1, 2, 3$.
As shown above, the varying strength of the resulting field $B_0 + Gx$ will change the protons Larmor frequency. Thus a gradient system enables spatial encoding of the received RF resonance signal.

Phase

A *phase* is defined as an angle between the magnetic moments of two precessing magnets. If, for instance, the magnetic moments point into the same direction, their phase is 0°. One also says, the protons are *in-phase* and calls the state when the two protons have a stable phase difference *phase coherence.*
Dephasing means that a *phase shift* - a phase difference between two magnets - is introduced by for instance using a gradient. We already know that, for instance, a positive x_2 - or *phase encoding* -gradient will increase the Larmor frequency along the posterior-anterior axis. Thus, during a short time the anterior and the posterior magnets dephase and obtain a fixed phase difference after switching off the phase encoding gradient.

Echo generation

As mentioned above, the receiver coil measures the FID signal after the RF-pulse. Because of T2 relaxation and dephasing due to field inhomogeneities this signal decays with time. In contrast to the signal loss due to relaxation, the signal loss due to dephasing can be recovered. In order to rephase the spins and thereby generate an echo, there are two approaches. On the one hand *spin echoes* used for the DTI acquisition in Subsection 2.3.3 and on the other hand *gradient echoes* used in the Phantom scan in Subsection 2.3.1 and the fMRI data in Subsection 2.3.2.
In spin echo sequences a 180° refocusing pulse at resonance frequency is used at time T to invert the magnetization vectors of the spins. While the phase differences are inverted as well the spins retain their precession frequency. Consequently, after 2T the spins rephase and the echo signal reaches its maximum.
In order to generate a gradient echo one notes that for instance, by using positive x_1-gradient after the RF-pulse we can dephase the signal in a controlled fashion. The spins on the right get a phase lead compared to the left ones. After excitation a x_1-gradient with negative sign is used to dephase the spins. While recording the signal, we change the sign of the x_1- or *readout* gradient. This makes the spins, that precessed at a lower frequency before catch up the formerly faster ones. Thus the signal *rephases* and the FID

signal increases again.
One important difference between spin- and gradient-echo sequences is that in the former no signal is lost due to field inhomogeneities. The readout gradient in gradient echo sequences can only rephase the spins that were dephased by its negative gradient before. In contrast to that the refocusing pulse inverts any phase difference, i.e. also those due to field imperfections.

Echo-Planar Imaging

In the following, we are interested in *Echo Planar Images* (EPI). Literally, EPI means that by one *echo* one *planar* slice is *imaged.* EPI is the first ultra high-speed imaging technique proposed by Mansfield in 1977 [17]. EPI is typically implemented as *single shot* imaging. However, multi-shot EPI are commonly used as well.
In Section 2.3 we will correct EPI images coming from two different applications. One will be from a *functional MRI* (fMRI) experiment in which the brain of a proband is observed over a time interval. The speed of echo planar imaging is used to generate a time series of high temporal resolution. Functional MRI is of great importance in modern neurocognitive sciences.
In the second case, we will improve the quality of *diffusion weighted images*(DTI). Here, one measures the water diffusion in the head along many directions to accurately estimate a diffusion tensor. To make this feasible, one needs a fast imaging technique. There is a wide range of applications of DTI in research and clinics, see for example [7].

Spatial Encoding

Our goal is to accurately map the density of the excited protons to our continuous image - i.e. $\mathcal{I}(x) = \rho(x)$. However, since ρ is not directly observable, it has to be reconstructed from the received signal $S(t)$. We will now present the spatial encoding methods, we will use later on in our examples. In contrast to the preceding subsection we assume the B_0 field to be perfectly homogeneous.

Slice selection

We start by selectively exciting one specific slice, see Figure 2.1. This can be done by using a x_3-gradient while the RF pulse takes place. Due to the linear field inhomogeneity, the Lamor-frequency increases linearly along x_3 as given by equation (2.1). Because of the resonance condition, the RF-frequency $\omega_{rf}(\tilde{x_3})$ that will selectively excite the slice $(x_1, x_2, \tilde{x_3})$ is given by

$$\omega_{rf}(\tilde{x_3}) = \omega_0 + \gamma G_z \tilde{x_3}. \tag{2.3}$$

In the following, we assume that we are able to excite exactly one slice which is then infinitesimally thin. We model this by the Dirac-delta distribution $\delta(x_3 - \tilde{x_3})$ which

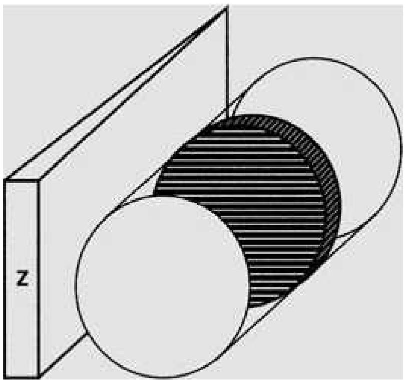


Figure 2.1: Slice selection using the x_3 gradient taken from Weishaupt [29]. A specific frequency selectively excites the shaded slice. The adjacent slices have other resonance frequencies and are therefore not influenced.

changes the representation of our signal in equation (2.2) to

$$S(t, \tilde{x_3}) = \int \mathcal{I}(x)\delta(x_3 - \tilde{x_3})e^{-i\gamma\Delta B(x)(t-T_E)}dx \tag{2.4}$$

Frequency Encoding

To determine from which image column of the excited slice the electric signal originates, we use *frequency encoding.* This time we use a x_1-gradient - i.e. the *frequency encoding direction* is the x_1-axis, i.e. from left to right. Consequently, the magnetic field linearly increases along the x_1-axis which is from right to left , see Figure 2.2. This means that protons on the right precess slower around the x_3-axis than protons on the left hand side. Moreover each column $(\tilde{x_1}, x_2, \tilde{x_3})$ now has a characteristic precession frequency which is given by

$$\omega(\tilde{x_1}, x_2, \tilde{x_3}) = \omega_0 + \gamma G_{x_1}\tilde{x_1}.$$

We can now enter this field inhomogeneity into equation (2.4)

$$S(t, \tilde{x_3}) = \int \mathcal{I}(x)\delta(x_3 - \tilde{x_3})e^{-i\gamma G_{x_1}x(t-T_E)}dx. \tag{2.5}$$

Phase Encoding

Finally, we want to determine the origin of the signal within a given column.The idea of *phase encoding* is to use a x_2 gradient for a short time interval T_{pe} - the *preparatory period* - directly after the RF-pulse. Note, that the system is already excited and the bulk magnetization of the selected slice precesses in the x_1x_2-plane. Tfhe linearly increasing magnetic field strength along the *phase encoding direction* - that is the x_2-axis, i.e. posterior to anterior - makes the anterior protons precess faster than the posterior ones. This leads to an advance - the so called *phase shift* - of the anterior protons compared to the

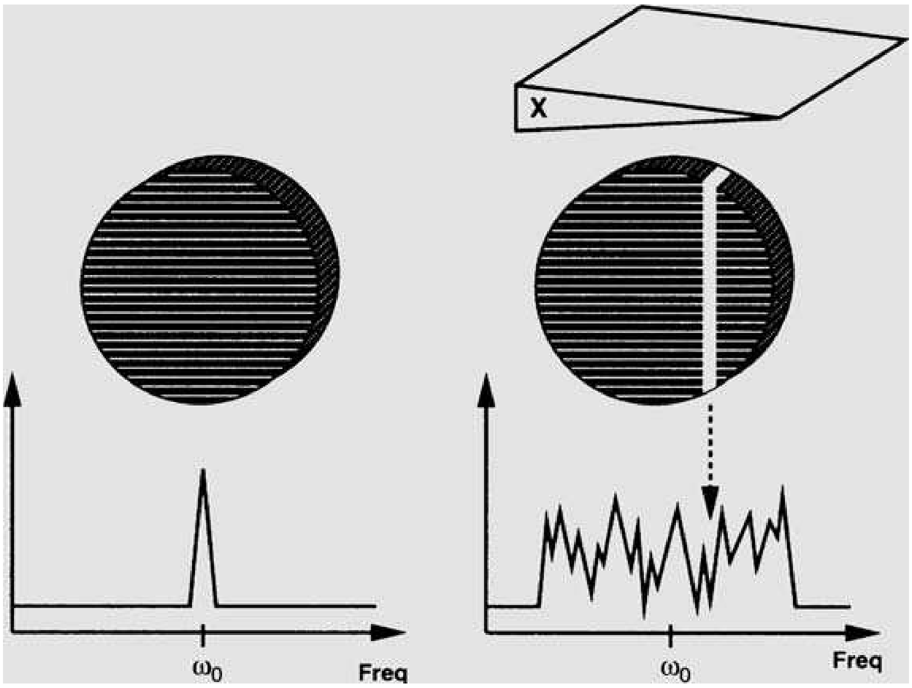


Figure 2.2: Frequency encoding by the x_1-gradient tzaken from Weishaupt [29]: without gradient (left) the only frequency we receive is the Larmor frequency ω_0. Using the gradient (right) we receive a frequency spectrum in which each frequency can be mapped to a specific column.

posterior ones. The extent of the advance depends on both the strength of G_{x_2} and the length of the preparatory period.
Therefore the x_2-component of the protons location can be determined by the phase shift

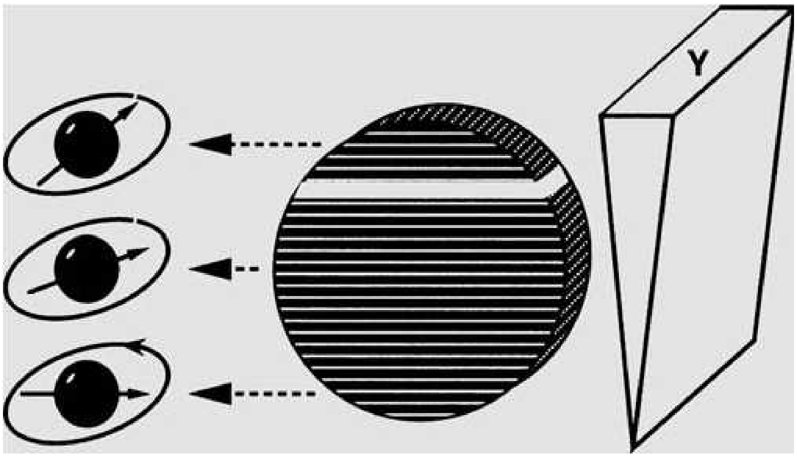


Figure 2.3: Phase encoding using the x_2-gradient by Weishaupt [29]. Each horizontal line gains a specific phase lead, that unambiguously identifies it.

$$\phi(x_1, x_2, x_3) = -\gamma G_{x_2} x_2 T_{Pe}.$$

When including this into the equation of the frequency encoded signal (2.5) to obtain

$$S(t, G_{x_2}, \tilde{x_3}) = \int \mathcal{I}(x)\delta(x_3 - \tilde{x_3})e^{-i\gamma(G_{x_1}x_1(t-T_E)-x_2G_{x_2}T_{pe})}dx. \tag{2.6}$$

To sum it up, phase encoding is done by pre-frequency encoding the signal for a short time interval. Note, that G_{x_2} and T_{pe} have to be chosen such that the dephasing is not more than 180°.

K-Space

We are now able to establish a connection between spatial encoding and the Fourier transform. We define

$$k_{x_1} := G_{x_1}(t - T_E) \text{ and } k_{x_2} := -G_{x_2}T_{pe}.$$

Our spatially encoded signal (2.6) then reads

$$S(k_{x_1}, k_{x_2}, \tilde{x_3}) = \int \mathcal{I}(x)\delta(x_3 - \tilde{x_3})e^{-i\gamma(k_{x_1}x_1 + k_{x_2}x_2)}dx. \tag{2.7}$$

Thus, k_{x_1} and k_{x_2} build up the two dimensional k-space for each given slice $\tilde{x_3}$. Whereas k_{x_1} increases linearly in time, k_{x_2} depends just on the strength and the length of the phase encoding. Thus, when recording the signal for a given time interval and given dephasing one covers one line in the k-space. In equation (2.7) one can see that $S(k_{x_1}, k_{x_2})$ is the Fourier transform of our slice image $\mathcal{I}(x_1, x_2)$. The idea then is to sample the k-space by many excitations with different preparatory periods and calculate the image by the inverse Fourier transform. In EPI however, one aims to sample the k-space with only one echo by additional dephasing and inverting the readout-gradient during the echo.

Effects of Field Inhomogeneities

After explaining spatial encoding in the absence of field imperfection, we can take the final step to our forward distortion model: We examine the effect of a non-linear field-inhomogeneity B_e on the resulting images.
There are two causes for those inhomogeneities. First, the B_1 field produced by the head coil may not be perfectly homogeneous. Secondly, varying magnetic susceptibility - the ability of an object to be magnetized - of different tissue types leads to imperfections that are correlated to the strength of the main magnet [5]. Note, that the latter part of the inhomogeneity varies from subject to subject and furthermore depends on their positions in the scanner.

Slice Selection

As can be seen in the slice selection equation (2.3) the excitation of a specific slice is based on the assumption of linear inhomogeneities. If, for one of the reasons above, the field is perturbed and we excite the system by an RF-pulse of frequency ω_{rf} given by (2.3) instead of exciting the planar slice $(x_1, x_2, \tilde{x_3})$ we would now excite the slice $(x_1, x_2, z'(x))$ where

$$z'(x) = x_3 + \frac{B_e(x)}{G_{x_3}}.$$

With ω_{slice} denoting the bandwidth of the slice gradient in [Hz] and h_3 the voxel size in x_3 direction in mm the distortion reads

$$z'(x) = x_3 + \frac{h_3}{\omega_{\text{slice}}} B_e(x).$$

The RF pulse is then in general not able to selectively excite protons of one planar slice. Since the reconstruction algorithm is unable to account for non-linear inhomogeneities this leads to a shift in x_3-direction.

In our EPI sequences the bandwidth in slice selection is relatively high. This leads only to subvoxel movements.

Frequency Encoding

In the presence of B_e, the mapping of signal frequency to a specific column is not accurate any more. The field inhomogeneity adds a spurious component to the frequency spectrum, given by

$$\omega(x) = \gamma B_0 + \gamma G_{x_1} x_1 + \gamma B_e(x).$$

This changes the signal (2.7) to

$$S(k_{x_1}, k_{x_2}, \tilde{x_3}) = \int \mathcal{I}(x) \delta(x_3 - z'(x)) e^{-i\gamma(k_{x_1} x_1 + k_{x_2} x_2 + B_e(x) t)} dx.$$

We end up with

$$S(k_{x_1}, k_{x_2}, \tilde{x_3}) = \int \mathcal{I}(x) \delta(x_3 - z'(x)) e^{-i\gamma(k_{x_1} x_1 + k_{x_2} x_2')} dx. \tag{2.8}$$

When denoting the receiver bandwidth in [Hz] by ω_{rec}, the distortion along the frequency encoding direction is given as

$$y'(x) = x_1 + \frac{h_1}{\omega_{\text{rec}}} B_e(x).$$

However, distortions in readout-direction are not the major problem in EPI, because the acquisition time between adjacent points in k-space is very short - note, that k_x increases linearly in time [14].

Phase Encoding

The presence of B_e affects the precession frequency of the protons. Clearly, protons with a higher B_e will precess faster relative to other protons at a position where B_e is lower or even negative. As pointed out above, phase encoding assumes that the position of the magnets can be determined by their phase difference. In the absence of B_e this initial phase shift remains after turning the x_2-gradient off. The presence of B_e now adds a spurious component to the controlled dephasing given by

$$\phi(x) = \gamma B_e(x) t.$$

We define the shift along the phase-encoding direction as

$$y'(x) := x_2 + \frac{B_e(x)}{G_{x_2}} t. \tag{2.9}$$

Alternatively, when denoting the excitation bandwidth in [Hz] by ω_{exc}, the shift along the phase encoding direction is given as

$$y'(x) := x_2 + \frac{h_2}{\omega_{\text{exc}}} B_e(x).$$

Here it is important to note that the sign of the shift is altered when the phase encoding gradient is reversed.

Because the time between the acquisition of two adjacent points in the x_2-direction of the k-space is rather high, this effect is more pronounced in EPI. In contrast to the bandwidth in frequency encoding or slice selection direction, the bandwidth is of more than one order smaller. Therefore the field inhomogeneity leads to displacements of several voxels in phase encoding direction.

Forward Model

After describing how a non-linear field-inhomogeneity B_e affects the spatial localization we formulate our forward distortion model.

In order to do so, we will assume that the non-linear field-inhomogeneity $B_e : \mathbb{R}^3 \to \mathbb{R}$ is known and in $C^1(\mathbb{R}^3, \mathbb{R})$. Furthermore, we know from the MRI protocol the values for $\omega_{\text{rec}}, \omega_{\text{exc}}, \omega_{\text{slice}} \in \mathbb{R}$ and for $h \in \mathbb{R}^3$.

Then, we examine the relationship between the exact image - measured in the absence of any non-linear inhomogeneity - and the distorted image. As we found out, the field-inhomogeneity causes distortions along the slice selection and phase-encoding direction. The amount of the shifts in those directions both depend on the strength of B_e. Thus, shifts can only occur along one direction - the *distortion direction* v.

Definition 14 (Coordinate change). *The* distortion direction *in* $\mathcal{I}_1$ *is defined by*

$$v = \left(\frac{h_1}{\omega_{rec}}, \frac{h_2}{\omega_{exc}}, \frac{h_3}{\omega_{slice}}\right)^T.$$

The coordinate change is then given by

$$y : \mathbb{R}^3 \to \mathbb{R}^3,\ y(x) = x + B_e(x)\, v.$$

Remark 3 (Distortion direction in $\mathcal{I}_2$)**.** *Following the reversed gradient method usually assumes to alter the sign of all gradients when acquiring* $\mathcal{I}_2$*. Then the distortions will occur exactly along the same axis just in opposite directions. In practice, however, only the phase encoding gradient is altered. In such cases the distortions in* $\mathcal{I}_2$ *will occur along*

$$v_2 := -Qv$$

where Q *is an orthogonal matrix that inverts the sign of the readout- and slice-selection gradient. This notation can simply be adjusted in cases when also the readout gradient is altered but not the slice encoding gradient.*

Remark 4. *As we have seen, the amount of the geometric deformation can be reduced by increasing the bandwidth. This, however, decreases the signal to noise ratio (SNR) and increases artefacts due to eddy currents in diffusion tensor imaging.*
As pointed out above, the bandwidth in phase encoding direction is of more than one order smaller compared to the other two directions. Therefore v *is almost parallel to the anterior-posterior axis and the phase encoding bandwidth has a big impact on the occurring distortions. If, for instance, the phase encoding gradient is not steep enough,* y *might not be diffeomorphic. Then intensities from different voxels will sum up and both be localized in the same voxel. In those cases the Jacobian determinant will either be zero or not finite.*
That is why the proper choice of the phase encoding bandwidth is a necessary condition for the success of a later correction.

Because of integration by substitution the integral of our signal (2.8) changes to

$$S(k_{x_1}, k_{x_2}, \tilde{x_3}) = \int \mathcal{I}(y(x)) |\det(Dy(x))| \delta(z_1 - \phi(x)) e^{-i\gamma(k_{x_1}x_1' + k_{x_2}x_2')} dx.$$

By the use of the Fourier transformation we can now obtain the very important relationship between our undistorted image i and the deformed and modulated version $\tilde{\mathcal{I}}$ measured in the presence of the field inhomogeneity B_e. This is the final step to our forward distortion model.

Definition 15 (Forward Model)**.** *With the notation of Definition (14) and under the assumption that the coordinate change in Definition 14* y *is diffeomorphic we can define the operator* $\mathbf{S} : Img(\mathbb{R}^3) \to Img(\mathbb{R}^3)$

$$\tilde{\mathcal{I}}(x) = \mathbf{S}(\mathcal{I}; B_e, v)(x) = \mathcal{I}(y(x)) \det(Dy(x))$$

that applies the distortion caused by the field inhomogeneity B_e *on the exact image* $\mathcal{I}$*. Given the distorted image* $\tilde{\mathcal{I}}$*, one can also get the exact one back by*

$$\mathcal{I}(x) = \mathbf{S}^{-1}(\tilde{\mathcal{I}}; B_e, v))(x) = \tilde{\mathcal{I}}(y^{-1}(x)) / \det(Dy(x)).$$

Theorem 2. *Let* $d : \mathbb{R}^3 \to \mathbb{R}$ *be continuously differentiable and* $y(x) := x + dv$ *be diffeomorphic. Then the determinant of* Dy *simplifies to*

$$\det(Dy(x)) = 1 + (\nabla d, v).$$

Proof. In this special case, the transformation becomes

$$y(x) = x + d(x)v = \begin{pmatrix} x_1 + d(x)v_1 \\ x_2 + d(x)v_2 \\ x_3 + d(x)v_3 \end{pmatrix}.$$

The Jacobian then reads

$$Dy = \begin{pmatrix} 1 + \partial_1 dv_1 & \partial_2 dv_1 & \partial_3 dv_1 \\ \partial_1 dv_2 & 1 + \partial_2 dv_2 & \partial_3 dv_2 \\ \partial_1 dv_3 & \partial_2 dv_3 & 1 + \partial_3 dv_3 \end{pmatrix}.$$

With Sarrus' rule the functional determinant can then be calculated as

$$\begin{aligned} \det(Dy) = {} & (1 + \partial_1 v_1)(1 + \partial_2 dv_2)(1 + \partial_3 dv_3) + \partial_2 dv_1 \partial_3 dv_2 \partial_1 dv_3 \\ & + \partial_3 dv_1 \partial_1 dv_2 \partial_2 dv_3 - (1 + \partial_1 v_1) \partial_3 dv_2 \partial_2 dv_3 \\ & - \partial_2 dv_1 \partial_1 dv_2 (1 + \partial_3 dv_3) - \partial_3 dv_1 (1 + \partial_2 dv_2) \partial_1 dv_3 \end{aligned}$$

When expanding, all products of derivatives cancel each other out and we get

$$\begin{aligned} & = 1 + \partial_1 dv_1 + \partial_2 dv_2 + \partial_3 dv_3 \\ & = 1 + \nabla d \, v. \end{aligned}$$

□

Experimental Setting: The Inverse Problem

By definition of our physically motivated forward model, gradients with altered sign lead to deformation in opposite directions and resulting intensity modulation, because the sign of the bandwidth is altered as well. This motivates the acquisition of two images with altered gradients in order to estimate B_e as proposed in [5, 26, 20]. After the acquisitions, we are interested in solving the following inverse problem:

Theorem 3 (Inverse Model). *Given two images $\mathcal{I}_1$ and $\mathcal{I}_2$ that are identical except from the altered gradient. Especially it is assumed that the non-linear field-inhomogeneity B_e is identical in both images and the position and orientation of the subject in the scanner is unchanged.*
We then want to find a suitable inhomogeneity B_e that renders the corrected images as similar as possible, i.e.

$$\mathbf{S}^{-1}(\mathcal{I}_1; B_e, v) = \mathbf{S}^{-1}(\mathcal{I}_2; B_e, -Qv).$$

2.2 Variational Approach for Susceptibility Correction of EPI

Bearing the physical model in mind and having the image registration tools of Chapter 1 at hand, we will now develop a variational approach to the EPI correction problem. To this end we start by translating the inverse model derived in the previous section into an image registration problem. For the field inhomogeneity B_e is scalar and non-linear, we model it by a scalar function $d : \mathbb{R}^3 \to \mathbb{R}$. Combining - i.e. multiplying - d and the distortion direction v and $-Qv$ we obtain a non-parametric transformation $y : \mathbb{R}^3 \to \mathbb{R}^3$, that is restricted along v and $-Qv$, respectively. Clearly, $d(x)$ is the amount the voxel x is shifted along the distortion direction.

In addition to the coordinate change, the presence of B_e and the mass-preservation leads also to a change in intensity. To account for this, we have to multiply the images' grey values by the determinant of the transformation's Jacobian, as described in the section above. This will incorporate the mass preserving property of the EPI distortions directly in our model.

In comparison to the typical image registration setting, it is noticeable, that there is no reference/template relationship. Out of the two deformed images we want to estimate the field inhomogeneity d in order to obtain the undistorted image. Therefore both images are transformed and modulated in opposite directions. On the other hand our problem is not that different from our standard formulation as well. Given two images, we want to estimate a suitable - in our case elastically regularized - transformation y, that makes the transformed images as similar as possible.

This motivates our advance in the remainder of this section. After addressing the problem of rigid body movements between the acquisition of $\mathcal{I}_1$ and $\mathcal{I}_2$ we will derive a new functional that models the EPI correction. Then the variations of our new functional are calculated and the existence of a minimizer in the function space $W^{1,2}(\Omega)$ is shown. Also some notes about the discretization and implementation of our functional are given.

Rigid Body Movement

The physical model assumes that the images are identically except for their altered gradient sign. However, the patient might move between the acquisition of $\mathcal{I}_1$ and $\mathcal{I}_2$. Consequently, we could correct for those movements by a rigid registration estimating the rotation $Q^{\text{rot}} \in O(3)$ and the voxel shift $(\theta_4, \theta_5, \theta_6)'$. Consequently, the distortion direction in the rigidly registered image v has to be rotated by $Q^{\text{rot}} \cdot Q$ when resampling the image volume.

On the one hand, taking the head movement into consideration should increase the accuracy of the correction result.

On the other hand it might be argued that whenever significant head movements occur between both acquisitions, the assumptions of our inverse problem are not satisfied. Note particularly, that the field-inhomogeneity caused by susceptibility differences depends on the subjects position inside the scanner. Thus the two reference scans are taken in different magnetic fields. Additionally, even if there was no head movement at all the huge

susceptibility artefacts might give rise to a spurious head movement. In such cases one might be better off disregarding the motion between both acquisitions.
Since accounting for a possible head movement is rather a notational challenge than a mathematical, we will not neglect it apart from some proofs. Our code is capable of handling head movements by using a rigid registration before performing the susceptibility correction. This also gives us the chance to study the impact of head movements to our correction in future projects.
For notational convenience we will assume that Q already incorporates the possible rotation of the distortion direction.

Deriving the Distance Functional

We incorporate our prior knowledge about the intensity change caused by the transformation in the distance term. More precisely, we change the residual function r in the general distance measure, cf. equation (1.3). For the grey values can be assumed to be correlated, we use the SSD distance to compare the images. We transform the images by interpolating them at

$$\mathcal{I}_1(x + d(x)v) \cdot (1 + \nabla dv) \text{ and } \mathcal{I}_2(x - d(x)Qv) \cdot (1 - \nabla d(Qv)).$$

Note, that usually it is assumed that the subject did not move between both acquisitions - i.e. Q is the identity matrix. In this special case both transformations are performed exactly in opposite directions. Furthermore it should be mentioned that although our two transformations are vector fields, they both just depend on one scalar field - namely d. Later on, this will simplify our optimization problem. The new residual is defined as

$$r(\mathcal{I}_1, \mathcal{I}_2; d) = \mathcal{I}_1(x + d(x)v) \cdot (1 + \nabla dv) - \mathcal{I}_2(x - d(x)\ Qv) \cdot (1 - \nabla d(Qv)).$$

our distance functional then reads

$$\begin{aligned}\mathcal{D}^{\text{EPI}}(d) =& \int_\Omega \psi(r(\mathcal{I}_1, \mathcal{I}_2; d))dx \\ =& \frac{1}{2} \int_\Omega (\mathcal{I}_1(x + d(x)v)(1 + \nabla dv) - \mathcal{I}_2(x - d(x)\, Qv)(1 - \nabla d(Qv)))^2 dx.\end{aligned}$$

Definition 16 (Lagrangian)**.** *Let* $v \in \mathbb{R}^3$. *The associated Lagrangian to our distance functional, is a smooth function*

$$L(p(x), z(x), x) : \mathbb{R}^3 \times \mathbb{R} \times \Omega \to \mathbb{R},$$

where p *will be substituted by* ∇d *and* z *is the placeholder for* d*. In our simplified case -* $v = e_2$ *and* Q *the identity -* L *does not depend on* p_1, p_2

$$\begin{aligned}L(p, z, x) =& \frac{1}{2} \left(\mathcal{I}_1(x + zv)(1 + pv) - \mathcal{I}_2(x - zQv)(1 - p(Qv))\right)^2 \\ :=& \frac{1}{2} r(\mathcal{I}_1, \mathcal{I}_2; p, z, x)^2.\end{aligned}$$

With this notation we observe that

$$\mathcal{D}^{EPI} = \int_\Omega L(p(x), z(x), x)dx.$$

Using the Lagrangian, makes the computations of the derivatives easier and help us concerning the proof of the existence of a minimizer. Thereby, one can also switch to an Euler framework by observing, that a minimizer of our functional fulfils the optimality condition - i.e. a non-linear partial differential equation. This is shown in Evans [10] by the following Lemma.

Lemma 1 (Euler-Lagrange equation). *Let* $L(p(x), z(x), x) : \mathbb{R}^3 \times \mathbb{R} \times \Omega \to \mathbb{R}$ *be a smooth Lagrangian. Then any smooth minimizer* u *of the associated functional*

$$I[w] = \int_\Omega L(Dw, w, x)$$

solves the Euler-Lagrange equation, i.e.

$$-\sum_{i=1}^{3} \left(L_{p_i}(Du, u, x)\right)_{x_i} + L_z(Du, u, x) = 0 \text{ in } \Omega. \tag{2.10}$$

Proof. Let $v \in C_0^\infty(\Omega)$ be a test function. Then we define the real-valued function

$$i(\epsilon) := I[u + \epsilon v].$$

Since u is a minimizer of $I[\cdot]$, we observe that $i(\cdot)$ has a minimum at $\epsilon = 0$. The necessary condition then gives us

$$i'(0) = 0. \tag{2.11}$$

We compute this *first variation* by writing out

$$i(\epsilon) = \int_\Omega L(Du + \epsilon Dv, u + \epsilon v, x).$$

The first derivative of i then reads

$$i'(\epsilon) = \int_\Omega \sum_{i=1}^{3} L_{p_i}(Du + \epsilon Dv, u + \epsilon v, x) v_{x_i} + L_z(Du + \epsilon Dv, u + \epsilon v, x) v dx.$$

By setting $\epsilon = 0$ we obtain with (2.11)

$$0 = i'(0) = \int_\Omega \sum_{i=1}^{3} L_{p_i}(Du, u, x) v_{x_i} + L_z(Du, u, x) v dx.$$

For v has compact support integration by parts gives us

$$0 = \int_\Omega \left[-\sum_{i=1}^{3} (L_{p_i}(Du, u, x))_{x_i} + L_z(Du, u, x) \right] v dx$$

which then leads to the desired non-linear partial differential (2.10), by applying the fundamental theorem of variational calculus. $\square$

Theorem 4 (First variation)**.** *Let* $\mathcal{I}_1, \mathcal{I}_2 \in C^1(\Omega, \mathbb{R})$ *be continously differentiable. To simplify notation we assume* $v = e_2$*,* $Q = id$ *and denote the resampled images by* $\tilde{\mathcal{I}}_1 := \mathcal{I}_1(x + d(x)e_2)$ *and* $\tilde{\mathcal{I}}_2 := \mathcal{I}_2(x - d(x)e_2)$*. Then the first variation of our distance term in divergence form reads*

$$\begin{aligned} d\mathcal{D}^{EPI}(d) = &- \partial_2 \left(r(\mathcal{I}_1, \mathcal{I}_2, d)\ (\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2) \right) \\ &+ r(\mathcal{I}_1, \mathcal{I}_2, d) \left(\partial_2 \tilde{\mathcal{I}}_1 (1 + \partial_2 d) + \partial_2 \tilde{\mathcal{I}}_2 (1 - \partial_2 d) \right). \end{aligned} \tag{2.12}$$

This can further be simplified to

$$d\mathcal{D}^{EPI}(d) = [\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2] \left(-\partial_2 [\tilde{\mathcal{I}}_1 - \tilde{\mathcal{I}}_2] - \partial_2 [(\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2) \partial_2 d] \right). \tag{2.13}$$

Proof. In Definition 16 the Lagrangian of our distance functional is defined as

$$L(p, z, x) = \frac{1}{2} r(\mathcal{I}_1, \mathcal{I}_2; p, z, x)^2$$

Because of the choice of v and Q, L depends only on p_2. As shown in Lemma 1, we can obtain the first variation by

$$d\mathcal{D}^{\mathrm{EPI}} = -(L_{p_2}(p, z, x))_{x_2} + L_z(p, z, x). \tag{2.14}$$

Calculating the derivatives individually, one gets

$$\begin{aligned} L_z =& L_z(p, z, x) \\ =& r(\mathcal{I}_1, \mathcal{I}_2; p_2, z, x)\ \partial_z r(\mathcal{I}_1, \mathcal{I}_2; p_2, z, x) \\ =& r(\mathcal{I}_1, \mathcal{I}_2; p_2, z, x) \left((1 + p_2)\ \partial_2 \tilde{\mathcal{I}}_1 + (1 - p_2)\ \partial_2 \tilde{\mathcal{I}}_2 \right). \end{aligned}$$

Since $L_{p_1} = L_{p_3} = 0$, we just need to calculate

$$L_{p_2} = r(\mathcal{I}_1, \mathcal{I}_2; p_2, z, x) \left(\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2 \right). \tag{2.15}$$

Inserting L_{p_2} and L_z into equation (2.14) yields the first variation in its divergence form (2.12). In order to further simplify the expression we explicitly compute the derivative with respect to x_2.

$$\begin{aligned} (L_{p_2})_{x_2} =& \partial_2 r(\mathcal{I}_1, \mathcal{I}_2; p_2, z, x) \left(\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2 \right) \\ &+ r(\mathcal{I}_1, \mathcal{I}_2; p_2, z, x) \left((1 + p_2)\ \partial_2 \tilde{\mathcal{I}}_1 + (1 - p_2)\ \partial_2 \tilde{\mathcal{I}}_2 \right) \end{aligned}$$

We compute the derivative of the residuum r with respect to x_2

$$\begin{aligned} \partial_2 r =& \partial_2 r(\mathcal{I}_1, \mathcal{I}_2; p_2, z, x) \\ =& \partial_2 \left((1 + p_2)\ \tilde{\mathcal{I}}_1 - (1 - p_2)\ \tilde{\mathcal{I}}_2 \right) \end{aligned}$$

With product rule, we get

$$
\begin{aligned}
&=\partial_2 p_2\,\tilde{\mathcal{I}}_1 + (1+p_2)\partial_2\tilde{\mathcal{I}}_1 + \partial_2 p_2\,\tilde{\mathcal{I}}_2 - (1-p_2)\partial_2\tilde{\mathcal{I}}_2\\
&=(1+p_2)\partial_2\tilde{\mathcal{I}}_1 - (1-p_2)\partial_2\tilde{\mathcal{I}}_2 + \partial_2 p_2\,\tilde{\mathcal{I}}_1 + \partial_2 p_2\tilde{\mathcal{I}}_2\\
&=\partial_2(\tilde{\mathcal{I}}_1 - \tilde{\mathcal{I}}_2) + p_2(\partial_2(\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2)) + \partial_2 p_2(\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2).
\end{aligned}
$$

Using product rule again, yields

$$
=\partial_2\left(\tilde{\mathcal{I}}_1 - \tilde{\mathcal{I}}_2\right) + \partial_2\left((\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2)p_2\right).
$$

This finally gives us the first derivative of our functional

$$
d\mathcal{D}^{\text{EPI}} = [\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2]\left(-\partial_2[\tilde{\mathcal{I}}_1 - \tilde{\mathcal{I}}_2] - \partial_2[\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2)\partial_2 p_2]\right).
$$

□

Remark 5. *The associated* Euler-Lagrange *equation of our distance measure* $\mathcal{D}^{EPI}$ *is given by*

$$
[\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2]\left(-\partial_2[\tilde{\mathcal{I}}_1 - \tilde{\mathcal{I}}_2] - \partial_2[(\tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2)\partial_2 z]\right) = 0. \tag{2.16}
$$

Consequently, a minimizer of $\mathcal{D}^{EPI}$ *solves this non-linear partial differential equation.*

Notation 1 (Directional derivative). *Let* $v \in \mathbb{R}^3$. *To simplify the following notation, we introduce the directional derivative operator*

$$
\nabla_v : C^1(\mathbb{R}^3,\mathbb{R}) \to C(\mathbb{R}^3,\mathbb{R}) \tag{2.17}
$$

$$
f \mapsto v_1\partial_1 f + v_2\partial_2 f + v_3\partial_3 f. \tag{2.18}
$$

Corollary 1. *Under the same assumption on* $\mathcal{I}_1$ *and* $\mathcal{I}_2$ *the theorem can be generalized for arbitrary distortion directions* $v \in \mathbb{R}^3$ *and rotations* $Q \in O(3)$. *With the same techniques as in the proof above, one obtains the first variation in its divergence form as*

$$
\begin{aligned}
d\mathcal{D}^{EPI}(d) = -\sum_{i=1}^{3}\partial_i\left(r(d)\;[v_i\tilde{\mathcal{I}}_1 + (Qv)_i\tilde{\mathcal{I}}_2]\right)\\
+r(d)\left(\nabla_v\tilde{\mathcal{I}}_1\;(1+\nabla_v d) + \nabla_{Qv}\tilde{\mathcal{I}}_2\;(1-\nabla_{Qv} d)\right).
\end{aligned} \tag{2.19}
$$

Implementation 5 (First variation). *The implementation of the first variation is not trivial. Directly implementing the continuous form (2.13) turns out to be not stable enough. This is due to the fact that the partial differential operators in*

$$
d\mathcal{D}^{EPI}(d) = -\sum_{i=1}^{3}\partial_i\left(L_{p_i}(p,z,x)\right) + L_z(p,z,x)
$$

stem from a partial integration argument as we have seen in Lemma 1. Therefore, in a discrete setting, the outer partial derivative operator has to be chosen as the adjoint operator of ∇d. *That is why we implement the first variation in its divergence form (2.19). For computing the gradient of* d *we choose explicit finite differences denoted by*

$$
\nabla d \approx \nabla^+ d.
$$

Thus, when computing the divergence, we have to use implicit finite differences by ∂_i^- *for* $i = 1, 2, 3$.

Theorem 5 (Second variation). *Under the same assumptions as in Theorem 4 the second variation of the functional $\mathcal{D}^{EPI}$ can be approximated by*

$$d^2\mathcal{D}^{EPI}(d) \approx \nabla r(d)^T \nabla r(d),$$

where the derivative of the residual is given by

$$\nabla r(d) = \nabla_v \mathcal{I}_1 \, (1 + \nabla_v d) + \nabla_{Qv} \mathcal{I}_2 \, (1 - \nabla_{Qv} d) + \mathcal{I}_1 \nabla_v + \mathcal{I}_2 \nabla_{Qv}.$$

Simplifying the Elastic Potential

Theorem 6 (Elastic Potential). *Let $d \in C^1(\mathbb{R}^3, \mathbb{R})$ be continuously differentiable and $v \in \mathbb{R}^3$. Let $y(x) := x + d(x)v$ and $y_{Ref}(x) = x$. Let $|v|^2 = 1$. Then with $\mathcal{B}$ being the linear elastic potential operator $\mathcal{B}$ presented in Chapter 1 and*

$$\mathcal{B}_{EPI} = \begin{pmatrix} \sqrt{\mu}\nabla \\ \sqrt{\lambda+\mu}(v_1\partial_1 + v_2\partial_2 + v_3\partial_3) \end{pmatrix}.$$

the following equation holds

$$|\mathcal{B}[y - y_{Ref}]|^2 = |\mathcal{B}_{EPI}[d]|^2.$$

The linear elastic potential for the EPI correction then reads

$$\mathcal{S}^{EPI}(d) = \frac{1}{2}\int_\Omega \mu \, |v|^2 \, |\nabla d|^2 + (\mu + \lambda)(\nabla_v d)^2 dx.$$

Proof. The elastic differential operator for a transformation $y : \mathbb{R}^3 \to \mathbb{R}^3$ is defined in Chapter 1 as

$$\mathcal{B} = \begin{pmatrix} \sqrt{\mu}\nabla & 0 & 0 \\ 0 & \sqrt{\mu}\nabla & 0 \\ 0 & 0 & \sqrt{\mu}\nabla \\ \sqrt{\lambda+\mu}\partial_1 & \sqrt{\lambda+\mu}\partial_2 & \sqrt{\lambda+\mu}\partial_3 \end{pmatrix}.$$

Our transformation has a specific form - namely

$$y(x) = x + d(x)v = \begin{pmatrix} x_1 + d(x)v_1 \\ x_2 + d(x)v_2 \\ x_3 + d(x)v_3 \end{pmatrix}.$$

Inserting this into the differential operator, one gets

$$\mathcal{B}[y - y_{\text{Ref}}] = \begin{pmatrix} v_1\sqrt{\mu}\nabla d \\ v_2\sqrt{\mu}\nabla d \\ v_3\sqrt{\mu}\nabla d \\ \sqrt{\lambda+\mu}v_1\partial_1 d + \sqrt{\lambda+\mu}v_2\partial_2 d + \sqrt{\lambda+\mu}v_3\partial_3 d \end{pmatrix}.$$

This gives us the desired equality

$$|\mathcal{B}[y - y_{\text{Ref}}]|^2 = |\mathcal{B}_{\text{EPI}}[d]|^2.$$

Finally, we compute

$$
\begin{aligned}
|\mathcal{B}_{\text{EPI}}[d]|^2 =& |v_1\sqrt{\mu}\nabla d|^2 + |v_2\sqrt{\mu}\nabla d|^2 + |v_3\sqrt{\mu}\nabla d|^2 + (\mu+\lambda)(\sum_{i=1}^{3} v_i\, \partial_i d)^2 \\
=& \mu(v_1^2+v_2^2+v_3^2)|\nabla d|^2 + (\mu+\lambda)(\nabla_v d)^2.
\end{aligned}
$$

☐

Theorem 7 (First Variation). *Under the same assumptions and using the same notation as in the Theorem before, the first variation of our functional simplifies to*

$$d\mathcal{S}^{EPI}(d) = (\mathcal{B}_{EPI}d)^T\mathcal{B}_{EPI}.$$

Theorem 8 (Second Variation). *Under the same assumptions and using the same notation as in the Theorem before, the second variation of our functional simplifies to*

$$d^2\mathcal{S}^{EPI}(d) = \mathcal{B}_{EPI}^T\mathcal{B}_{EPI}.$$

Definition 17 (EPI correction functional). *Let $\mathcal{I}_1, \mathcal{I}_2 \in Img(\Omega)$ be EPI acquired with reversed gradients. Let $v \in \mathbb{R}^3$ denote the distortion direction and $Q \in O(3)$ the orthogonal matrix such that $\mathcal{I}_2$ is distorted along $-Qv$. Then EPI correction results in the minimization of*

$$\min_d \mathcal{J}^{EPI}(d) := ||\mathcal{I}_1(x+d(x)v)(1+\nabla_v d(x)) - \mathcal{I}_2(x-d(x)Qv)(1-\nabla_{Qv}d(x))||^2 + \alpha\mathcal{S}^{EPI}(d). \tag{2.20}$$

Existence of a Minimizer

Using methods of non-linear functional analysis, we aim to prove the existence of a minimizer for our EPI functional $\mathcal{J}^{\text{EPI}}$ from Definition 17 in the Sobolev space $W^{1,2}(\Omega,\mathbb{R})$ which is also referred to as $H^1(\Omega,\mathbb{R})$, see e.g. Evans [10] Section 5.2. One way to do this is to show that our functional is coercive and weakly lower semicontinuous. In order to deduct these properties from properties of the Lagrangian we will make use of the following theorem.

Theorem 9 (Weak lower semicontinuity). *Assume that the Lagrangian $L : \mathbb{R}^3 \times \mathbb{R} \times \Omega$ is smooth, bounded below and in addition the mapping*

$$p \mapsto L(p, z, x)$$

is convex for each $z \in \mathbb{R}, x \in \Omega$. Then the associated functional $I := \int_\Omega L(p(x), z(x), x)dx$ is weakly lower semicontinuous on $W^{1,2}(\Omega,\mathbb{R})$.

Proof. see Evans [10], p. 446f. ☐

With this result at hand, we can show the existence of a minimizer for our functional, by examining its Lagrangian. We want to apply the following theorem.

Theorem 10 (Existence of minimizer)**.** *Assume that the admissible set $\mathcal{A}$ is non empty.Let the Lagrangian L satisfy the coercivity inequality, i.e. there are constants $\delta > 0, \gamma \geq 0$ such that*

$$L(p, z, x) \geq \delta |p|^2 - \beta \ , \ \forall p \in \mathbb{R}^3, z \in \mathbb{R}, x \in \Omega.$$

Furthermore the mapping

$$p \mapsto L(p, z, x), \ \textit{is convex} \ \ , \forall z \in \mathbb{R}, x \in \Omega$$

Then there exists at least one function $u \in \mathcal{A}$ solving

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Proof. see Evans [10], p. 448f. □

Remark 6. *Because we did not specify any boundary condition, the admissible set $\mathcal{A}$ is the whole function space, i.e.*

$$\mathcal{A} = \{w \in W^{1,2}(\Omega, \mathbb{R})\}.$$

We start by defining the Lagrangian of our joint functional.

Definition 18 (Lagrangian)**.** *Let $\alpha > 0$, $v \in \mathbb{R}^3$ be the distortion direction and $Q \in O(3)$ the rotation of the subject between the scans. Then the Lagrangian of our EPI problem*

$$L(p, z, x) : \mathbb{R}^3 \times \mathbb{R} \times \Omega \to \mathbb{R}$$

is given by

$$\begin{aligned} L(p, z, x) =& \frac{1}{2} \left(\mathcal{I}_1(x + zv)(1 + (p, v)) - \mathcal{I}_2(x - zQv)(1 - (p, Qv)) \right)^2 \\ &+ \frac{\alpha}{2} \left(\mu \, |v|^2 \, |p|^2 + (\mu + \lambda)(p, v)^2 \right). \end{aligned}$$

Lemma 2 (Coercivity)**.** *For $\alpha > 0$ the EPI correction functional $\mathcal{J}^{EPI}$ is coercive on $W^{1,2}(\Omega, \mathbb{R})$.*

Proof. We are interested in an inequality of the form

$$L(p, z, x) \geq \delta |p|^2 - \beta \ , \ \forall p \in \mathbb{R}^3, z \in \mathbb{R}, x \in \Omega$$

with constants $\delta > 0$ and $\beta \geq 0$.

$$\begin{aligned} L(p, z, x) =& \frac{1}{2} \left(\mathcal{I}_1(x + zv)(1 + (p, v)) - \mathcal{I}_2(x - zQv)(1 - (p, Qv)) \right)^2 \\ &+ \frac{\alpha}{2} \left(\mu \, |v|^2 \, |p|^2 + (\mu + \lambda)(p, v)^2 \right) \\ \geq& \frac{\alpha}{2} \left(\mu \, |v|^2 \, |p|^2 + (\mu + \lambda)(p, v)^2 \right) \\ \geq& \frac{\alpha}{2} \mu \, |v|^2 \, |p|^2. \end{aligned}$$

By choosing $\beta = 0$ and $\delta = \frac{\alpha}{2}\mu\, |v|^2$ we have the desired inequality and see that

$$\mathcal{J}^{\mathrm{EPI}}(d) = \int_\Omega L(\nabla d, d, x)dx \geq \delta ||Dd||^2_{L^2}.$$

This completes our proof. □

Lemma 3 (Convexity)**.** *For $\alpha > 0$ the Lagrangian of the EPI correction functional $\mathcal{J}^{EPI}$ is convex in the first variable.*

Proof. We want to show that for fixed $\mathcal{I}_1, \mathcal{I}_2, v, Q, z, x$ the mapping

$$\begin{aligned} p \mapsto & \frac{1}{2}\left(\mathcal{I}_1(x+zv)(1+(p,v)) - \mathcal{I}_2(x-zQv)(1-(p,Qv))\right)^2 \\ & + \frac{\alpha}{2}\left(\mu\, |v|^2\, |p|^2 + (\mu+\lambda)(p,v)^2\right). \end{aligned}$$

is convex. Clearly, the linear elastic potential is convex in the gradient. It remains to show that the mapping is convex in p_i for $i = 1, 2, 3$. We calculate this for the first summand

$$p_i \mapsto \frac{1}{2}\left(\mathcal{I}_1(x+zv)(1+p_iv_i) - \mathcal{I}_2(x-zQv)(1-p_iv_i)\right)^2.$$

In order to simplify notation we define the scalars $i_1 := \mathcal{I}_1(x+zv)$ and $i_2 := \mathcal{I}_2(x-zQv)$. Then we show for given $t \in [0,1]$ and fixed $p, q \in \mathbb{R}$

$$\begin{aligned} & \left(i_1(1+tp+(1-t)q) - i_2(1-tp-(1-t)q)\right)^2 \\ \leq & t\left(i_1(1+p) - i_2(1-p)\right)^2 + (1-t)\left(i_1(1+q) - i_2(1-q)\right)^2. \end{aligned}$$

For the left hand side we compute

$$\begin{aligned} & \left(i_1(1+tp+(1-t)q) - i_2(1-tp-(1-t)q)\right)^2 \\ = & i_1^2\left(1+tp+(1-t)q\right)^2 - 2i_1i_2\left(1+tp+(1-t)q\right)\left(1-tp-(1-t)q\right) \\ & + i_2^2\left(1-tp-(1-t)q\right)^2 \\ = & i_1^2\left(t^2p^2 + 2t(1-t)pq + (1-t)^2q^2 + 1 + 2tp + 2(1-t)q\right) \\ & - 2i_1i_2\left(1 - t^2p^2 + 2t(1-t)pq + (1-t)^2q^2\right) \\ & + i_2^2\left(t^2p^2 + 2t(1-t)pq + (1-t)^2q^2 + 1 - 2tp - 2(1-t)q\right) \\ = & i_1\left((tp+(1-t)q)^2 + 1 + 2tp + 2(1-t)q\right) \\ & - 2i_1i_2\left(1 - (tp+(1-t)q)^2\right) \\ & + i_2^2\left((tp+(1-t)q)^2 + 1 - 2tp - 2(1-t)q\right). \end{aligned}$$

For the right hand side we obtain

$$\begin{aligned} & t\left(i_1(1+p) - i_2(1-p)\right)^2 + (1-t)\left(i_1(1+q) - i_2(1-q)\right)^2 \\ = & i_1^2\left(t(1+p)^2 + (1-t)(1+q)^2\right) - 2i_1i_2\left(t(1-p^2) + (1-t)(1-q^2)\right) \\ & + i_2^2\left(t(1-p)^2 + (1-t)(1-q)^2\right) \\ = & i_1^2\left(tp^2 + (1-t)q^2 + 1 + 2tp + 2(1-t)q\right) \\ & - 2i_1i_2(1 - tp^2 - (1-t)q^2) \\ & + i_2^2\left(tp^2 + (1-t)q^2 + 1 - 2tp - 2(1-t)q\right). \end{aligned}$$

Now we drop the expressions on both sides of the inequality. Our left hand side then becomes

$$\begin{aligned}&i_1^2(tp+(1-t)q)^2+2i_1i_2(tp+(1-t)q)^2+i_2^2(tp+(1-t)q)^2\\=&(i_1+i_2)^2(tp+(1-t)q)^2\\\leq&(i_1+i_2)^2(tp^2+(1-t)q^2).\end{aligned}$$

Therefore the distance term is convex in p_i. As already pointed out, due to the convexity of the elastic potential in p this gives us the convexity of the whole Lagrangian in p. ☐

Theorem 11. *Let $\alpha > 0$. The EPI correction functional $\mathcal{J}^{EPI}$ has at least one minimizer in $W^{1,2}(\Omega, \mathbb{R})$.*

Proof. As shown above, the joint functional is coercive. Furthermore its Lagrangian is convex in p and bounded below. Therefore, by Theorem 10 we have shown the existence of a solution. ☐

Remarks on the Implementation

Including the intensity modulation in our distance measure and the restriction of our transformation along one axis leads will be done by introducing a new objective function with the notation:

```
function [Jc,para,dJ,H] = JTK_symEPIobjFctnV3(I1,I2,omega,m,xc,Q,d,discScheme)
```

Here, `I1` and `I2` are the interpolation coefficients of the respective images. This notation is chosen to be more consistent with the *FAIR* toolbox especially to be able to use the optimization and visualization techniques. Furthermore we make use of the following notation:

```
 dim = numel(m) %dimension of Omega
 m = [m(1:dim)] % number of discretization points
 n = prod(m);
 yRef %reference grid;
 vI1 % distortion direction in I1
vI2 = Q*v; % adjusted distortion direction in I2
```

Instead of interpolating the template image on the deformed grid `yc`, as implemented for the native non-parametric elastic registration algorithm in `NPIRobjFctn.m` we transform both images only along one axis in opposite directions. Our new objective function thus has the deformation `d` as input argument instead of `yc`, where `d` is discretized on a cell-centred grid.

```
[I1c dI1] = inter(I1,xc + kron(vI1,d),omega,m);
[I1c dI2] = inter(I2,xc - kron(vI2,d),omega,m);
size(dI1) = [n 3*n]; sparse matrix
```

Since only directional derivatives of the images occur we compute

```
dI1 = dI1 * vI1;
dI2 = dI2 * vI2;
```

We start by calculating the intensity modulations, that are given by $1 + (\nabla d, v)$ and $1 - (\nabla d, Qv)$. This will be done by using forward finite differences and is implemented matrix-free. Because the second derivative stems from an integration by parts argument, it is evident to choose an adjoint differential operator. The optional argument `discScheme` chooses consistent discretization schemes.

```
dd = computeGradient(d,omega,m);
size(dd) = [n dim];
```

Later, we need the directional derivatives along v and v_2 and the directional derivative of those directional derivatives. Therefore we compute

```
ddvI1 = dd * v;
ddvI2 = dd * v2;
```

Next we modulate the intensities in both images

```
modI2 = I1c .* (1+ddvI1);
modI3 = I2c .* (1-ddvI2);
```

The interpolated and intensity modulated images can then be compared by means of the chosen distance function. We also call the adjusted regularizer to benefit from the restricted transformation as shown in Theorem 6.

```
[Dc,res,~,~,d2psi] = distance(modI1c,modI2c,omega,m,'doDerivative',doDerivative);
[Sc,dS,d2S] = regularizer2(d(:), omega,m,'doDerivative',doDerivative,'v',v);
```

The value of our functional then becomes

```
 Jc = Dc + alpha * Sc;
```

If we need to compute derivatives, we do the following

```
Lz = res.*(dI1.*(1+ddvI1)+dI2.*(1-ddvI2));
Lp = [res.*(v(1)*I1c +v2(1)*I2c) res.*(v(2)*I1c +v2(2)*I2c)
  res.*(v(3)*I1c +v2(3)*I2c)];
Lpx = divergence(Lp);
```

Then the first variation of the distance term becomes

```
 dD = Lz - Lpx;
```

The first variation of our functional reads

```
 dJ = dD' + alpha *  dS;
```

for the first variation of the distance term we need to compute the derivative of the residual with respect to d

```
 dOp1 = compDirectionalDerivative([],omega,m,v); %returns operator
 dOp2 = compDirectionalDerivative([],omega,m,v2); %returns operator
    dr = spdiags(dI1.*(1+modI1)+dI2.*(1-modI2),0,n,n)+...
        spdiags((I1c),0,n,n)*dOp1 + spdiags((I2c),0,n,n)*dOp2;
```

The Hessian of our functional is then estimated by

```
 H = d2psi * (dr'*dr) + alpha *  d2S;
```

2.3 Numerical Experiments

In this section we will apply the derived algorithm to three different EPI measurements. The first dataset contains EPI scans of a water-filled bottle. For this dataset a measured fieldmap is also available. Afterwards we will deal with the correction of EPI that occur in *functional magnetic resonance imaging*(fMRI). Finally, a *diffusion tensor imaging* (DTI) sequence is corrected. For the latter two problems are of great importance in clinical application and research, we will also suggest how to optimally incorporate our correction approach into the image processing pipeline.

Data

The images were acquired on a 3.0 T Philips Intera Achiever (Release 2.5) scanner.
All computations are performed using *Matlab* 2009b on a 64-Bit Linux system with 7 kernels at 2.5 GhZ each in 3D. The amount of memory was 16 GB.

Optimization

We configure our optimizer with the stopping criteria set to $(\rho_1, \rho_2, \rho_3) = (10^{-4}, 10^{-3}, 10^{-3})$ and we use an Armijo condition for the line search, to increase speed and since stronger conditions on the step length, did not improve the results in previous tests. The maximum number of iterations was set to 30.

Determining the regularization parameters

In order to find values for α, μ and λ we use an automatic test protocol. First we set $\mu = 1$ fixed and then try to find optimal values for λ and α. We recall the augmented regularization term from Theorem 6 for the simple case $v = e_2$ and $Q = I$

$$\mathcal{S}^{\mathrm{EPI}}(d) = \frac{1}{2} \int_\Omega \mu \, |\nabla d|^2 + (\mu + \lambda)(\partial_2 d)^2 dx.$$

We note that by choosing λ we have the opportunity to balance the weights of the two summands. While the first summand punishes oscillations in all directions, the last one only penalizes the directional derivative along the distortion direction. In this case, we showed above that this partial derivative is directly linked to the functional determinant, via

$$\det(D(x + e_2 d)) = 1 + \partial_2 d \text{ and } \det(D(x - e_2 d)) = 1 - \partial_2 d.$$

For we are only interested in diffeomorphic and orientation preserving transformations we will require both functional determinants to be positive and greater than 0. This is true if and only if

$$-1 < \partial_2 d(x) < 1 \; \forall x \in \Omega.$$

This is equivalent to

$$(\partial_2 d(x))^2 < 1 \; \forall x \in \Omega.$$

These simple considerations will drive our testing procedure, as we see that - in order to get a folding free grid - it is of great importance to punish the partial derivative along the distortion direction. We will therefore successively increase λ and then aim to find the smallest α that leads to a successful regularization, i.e. a folding free grid. At some point, we will, however, stop to increase λ when the gradients begin to oscillate too much along the other directions. Because of the huge deformations in our images the distortions caused by the field inhomogeneity cannot be expected to be very smooth along the readout and slice selection direction though. In our tests, $\lambda = 30$ was found to lead to good results. Choosing $\lambda << 30$ did not correct the huge deformations , while the resulting transformations achieved by choosing $\lambda >> 30$ oscillated heavily along the undistorted directions.

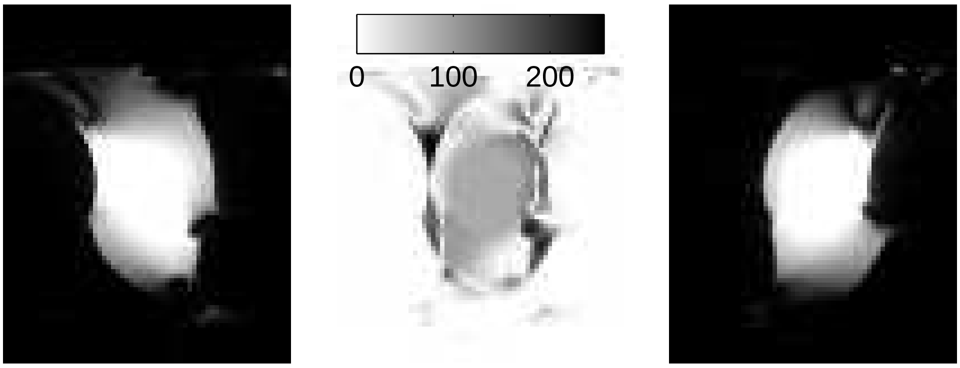
Visualization

For $\lambda = 30$ we vary α between 0 and 500 and visualize the results of this *α-test* in four plots. First of all, we plot the *l-curve* that is often used in inverse problems. The l-curve is a log-log plot with the value for the regularization term on the x-axis and the distance measure on the y-axis. One usually expects the plot to be of “L” form. In theory, one should then pick an alpha close to the bend of the L. In order to choose an α we will, however, follow a different strategy. As pointed out, we are interested in obtaining a diffeomorphic transformation. Thus we will use the first α that leads to a folding-free transformation. In order to examine whether the elastic regularization is able to lead to such a transformation we will plot α against the number of foldings in the second plot. Complementary to the l-curve we will also have a look at the dependence of the distance and regularization terms on alpha.
To demonstrate the obtained transformation, we show the initial images with superimposed *interpolation lines*. Those lines are a feasible visualization, because displacements only occur in one direction. Consequently, instead of showing the whole grid, we omit for clarity the direction along with no displacements occur. The transformation can then be interpreted as straighten the lines and adjusting the intensities for volume changes.

2.3.1 Phantom scan

In our first experiment we evaluate our inverse approach using data from a phantom scan. This dataset might be used in later studies to compare our image-based approach to fieldmap approaches, because we acquired a fieldmap as well.

Data



(a) $\mathcal{I}_1$ - saggital (b) Difference - saggital (c) $\mathcal{I}_2$ - saggital

Figure 2.4: Initial data. The apa (left), app (right) images and their initial difference are shown. Phase encoding was altered along the posterior-anterior axis. This leads to displacements in opposite directions from left to right in the slices shown. Note, that intensities sum up, when voxels are edged together and that intensity decreases when the tissue is stretched vice versa. More slices, transversal views and one further choice for phase encoding is illustrated in the Appendix A.1.

For the validation of our derived algorithm, we use a 1 litre water-filled plastic bottle as a phantom. The susceptibility differences on the water-air interface will lead to significant distortions along the phase-encoding direction.

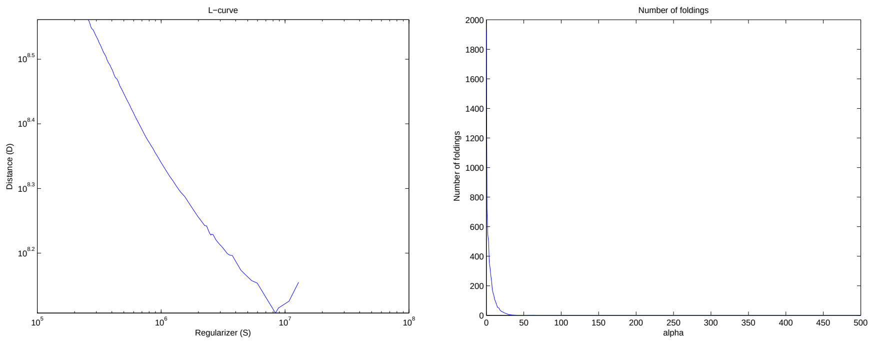
We obtain a phase encoded T2 magnitude image with 100 slices and a in plane resolution of 256×256. Note, that there is only water in the bottle and thus there is only one relaxation time. That is why, the weighting has no impact on the image contrast.

Furthermore, we measure four EPI magnitude images with 80 slices, in plane resolution of 64×64. The field of view is $\Omega = [0, 230] \times [0, 230] \times [0, 288]$ which corresponds to a voxel size of about 3.6 mm^3. The phase encoding direction varies from left/right, right/left, bottom/top and top/bottom. Figure A.1 in the Appendix shows how the choice of the phase encoding direction affects the distortions due to the field inhomogeneity. As our physical model predicts the main distortion occurs along the phase encoding direction. In the other directions the effects are less pronounced which is also indicated by the bandwidth parameters of the scan protocol.

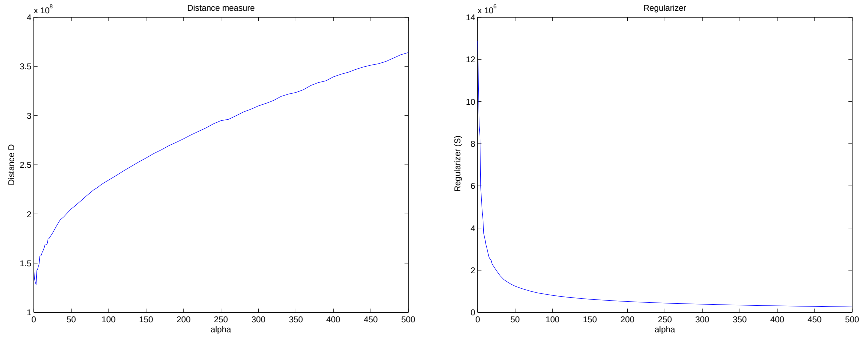
In addition to the huge deformations in both images the Lebesgue $\mathcal{L}_1$ measure varies for instance between the apa and app images, with 507 to 578, respectively. Therefore, being mass-preserving, our algorithm can not reduce the difference of both images to zero.

Finally we measure images that enable us to determine the field inhomogeneity. Once we do it with $\text{TE}_1 = 10$ ms and $\text{TE}_2 = 12$ ms and once with a shorter offset - namely $\text{TE}_1 = 5.7$ ms and $\text{TE}_2 = 9.1$ ms. At each echo time we acquire the magnitude, the real and imaginary part and the phase.

Correction results



(a) l-curve (b) Number of foldings

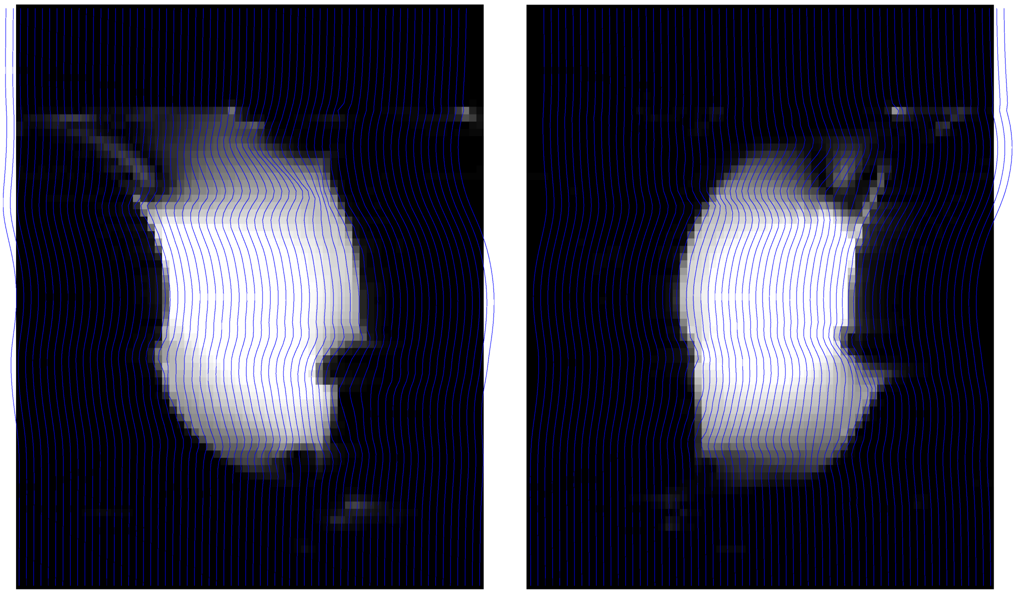


(c) Distance (d) Regularizer

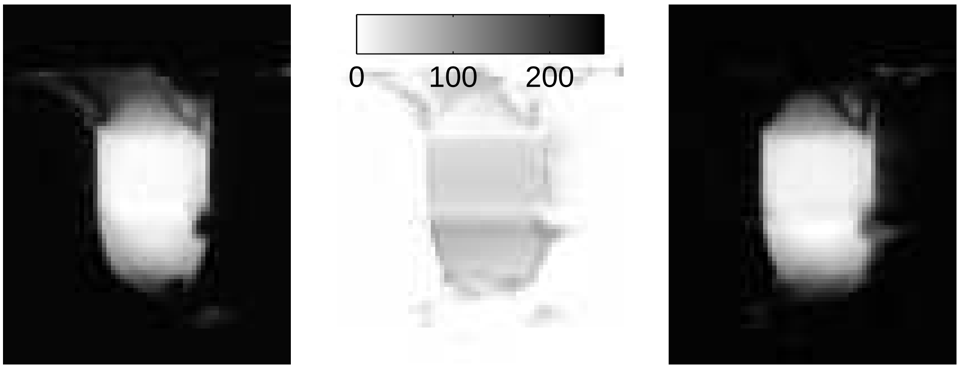
Figure 2.5: Results of α-Test. According to the explanation in 2.3 77 possible α values between 0 and 500 were tested for the correction of the apa and app images of the water bottle. The top-left plot shows the l-curve. As to be expected the more we reduce the distance term, the less elastic the deformation becomes.The top right plot, gives the number of grid foldings in dependence on α. Here we can see that our regularization is capable of supplying a grid free of foldings for sufficiently large α (here,≥ 250). In the bottom row, α is plotted against the distance term and the regularization term, respectively.

In our inverse approach we try to correct `apa` and `app` - i.e. the images with phase encoding along the anterior-posterior axis. By choosing $v = (0,1,0)^T$ and $Q = id$ distortions are only corrected along the phase encoding direction, where the water-fat shift is with 18.636 pix / Hz almost of two orders higher than along the other directions(e.g. readout direction 23.3 pix/Hz). To avoid noise induced deformations at the boundary, we pad the image with zeros in each direction. We make use of the multi-level strategy from Chapter 1. The first multi-level step uses a resolution of $9 \times 9 \times 8$. After 4 steps we end up at the full resolution of $72 \times 72 \times 88$.

The procedure described in Subsection 2.3 above showed us that $\mu = 1$ and $\lambda = 30$ were the best choices. Furthermore $\alpha = 260$ was the first value, that led to a folding-free grid. Using spline interpolation each correction was done in about 5 minutes. For $\alpha = 260$ the functional was reduced by 64.6 % whereas the distance term itself was even reduced by 69.8%. The number of iterations was (6,3,3,3) and resulted in each level in a local minimum. The correction was able to increase the normalized cross-correlation of both images from 90.3 % to 97.5%. The biggest deformation was 17.5 mm and the range of the functional determinant of $\mathcal{I}_1$ modulation was from 0.12 to 1.57. Results are depicted in Figure 2.3.1.



(a) $\mathcal{I}_1$ interpolation grid - saggital (b) $\mathcal{I}_1$ interpolation grid - saggital



(c) $\mathcal{I}_1$ interpolated - saggital (d) final difference - saggital (e) $\mathcal{I}_2$ interpolated - saggital

Figure 2.6: Correction result. The top row depicts the saggital image slices of the initial volume images with a superimposed folding-free interpolation lines. Since deformations only occur along the phase-encoding direction which is from left-right, only the deformed top-down-lines are depicted. Underneath, the resulting undistorted images are shown. Although the difference between the two images could be decreased significantly compared to Figure 2.3.1 the difference image does not disappear. However, the contours in the distance image disappeared more or less. What remains could be a systematic intensity difference due to the initially different $\mathcal{L}_1$ measure. More results are depicted in the Appendix A.1.

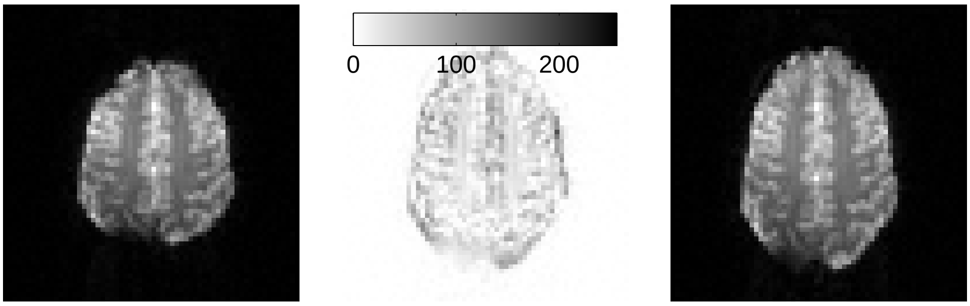
2.3.2 Functional MR Imaging

One major application of EPI is *functional Magnetic Resonance Imaging* (fMRI). Here changes in blood oxygenation and flow are imaged and related to neural activity, because active parts in the brain have a higher oxygen demand. fMRI can then be used to produce activation maps showing which parts of the brain are involved in a particular mental process. One example for a clinical application is the localisation of brain functions in pre-surgical planning. As a relatively young technique fMRI plays a small but growing role in neuroimaging, today [8]. In this subsection we will try to apply our correction algorithm to data that come from a neurocognitive experiment. We will also develop a complete correction pipeline for this data.

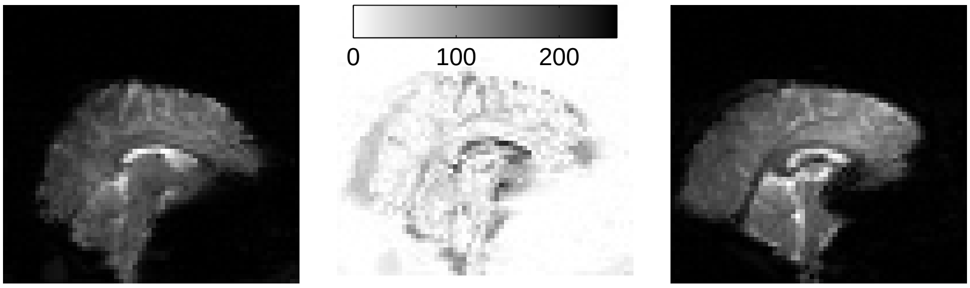
Experiment

Two important functions of cognitive control are conflict detection and inhibition. Using functional MRI, their neurophysiological correlates were studied in a modified Flanker paradigm. As task, subjects had to press a button. According to the direction of a visually presented pair of arrows, they had to use their right or left hand. In a minor case of the trials, subjects had to withhold their responses and change their response hand. For conflict detection and inhibition, the activation of especially the right inferior frontal cortex and the midcingulate cortex are crucially important.

Initial Data



(a) $\mathcal{I}_1$ - transversal (b) Difference - transversal (c) $\mathcal{I}_2$ - transversal



(d) $\mathcal{I}_1$ - saggital (e) Difference - saggital (f) $\mathcal{I}_2$ - saggital

Figure 2.7: Initial data. Transversal and saggital slices of fMRI images before the EPI correction are shown. Left and right rows show apa and app, respectively. In between the difference is depicted. Note the opposite effect of the field-inhomogeneity along the posterior-anterior axis in the two acquisitions. Furthermore the mass-preserving kind of the transformation can be seen, since intensity values sum up, wherever edgings occur and intensity decreases where the tissue is stretched. More slices and views are shown in the Appendix A.2.

The experiment was performed with a healthy male subject on a Philips 3.0 T scanner. The dataset consists of one time series of 170 EPI that show only the subjects brain before the experiment. Afterwards two full head EPI with a resolution of $64 \times 64 \times 80$ - but with the same voxel size of $3.6 \times 3.6 \times 3.6$ millimetres as the 170 EPI before - are recorded with phase encoding in altered directions along the anterior-posterior axis. Then the actual experiment is performed and the subjects brain is imaged meanwhile. This results in 1.135 detail images with positive phase encoding gradients, that have to be corrected. Therefore we aim to estimate the distortions from the two full-head scans.

The initial images is shown in Figure 2.3.2 and in the Appendix in A.2. It can be seen that the field-inhomogeneity leads to altered displacements along the anterior-posterior axis with large magnitude. This clearly illustrates the importance of the correction, since all images have to be registered to a standard space for later comparison. This latter registration will of course be much easier, using images that are undeformed by a mass-preserving transformation. Like in the phantom scan the $\mathcal{L}_1$ measure differs in the two scans with 236.77 and 191.09 for apa and app, respectively. We aim to estimate the EPI distortions using the two full head EPI that are acquired before the actual experiment took place. Thus, we choose $\mathcal{I}_1$ as the app image and $\mathcal{I}_2$ as the apa image.

In order to avoid deformations due to noise at the boundary, we pad the images with zeros.

Our α-test - that is performed using spline interpolation - gives us that $\alpha = 50$, $\mu = 1$ and $\lambda = 30$ lead to a folding-free deformation.

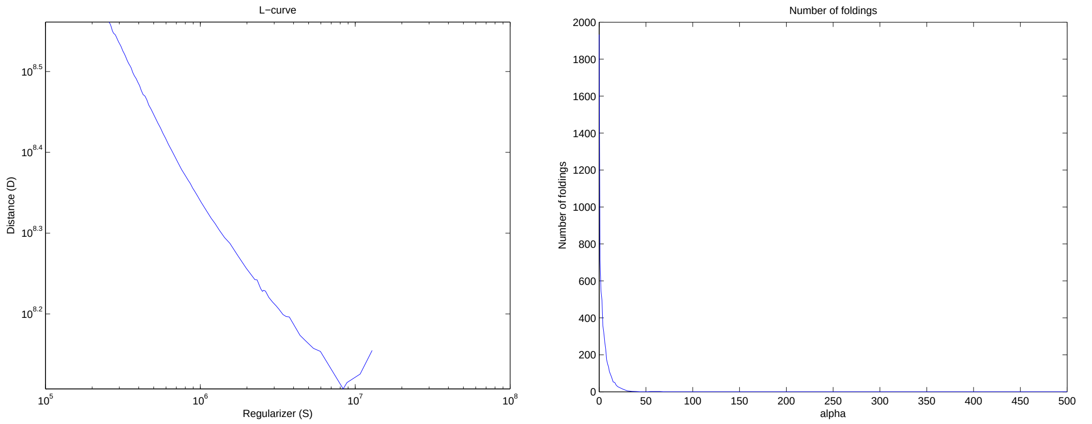
The multilevel strategy for the minimization of our functional is to start at level 4 with a resolution of $9 \times 9 \times 9$ and run the optimization until level 7 where the complete resolution of $72 \times 72 \times 68$ is reached.

Correction Results

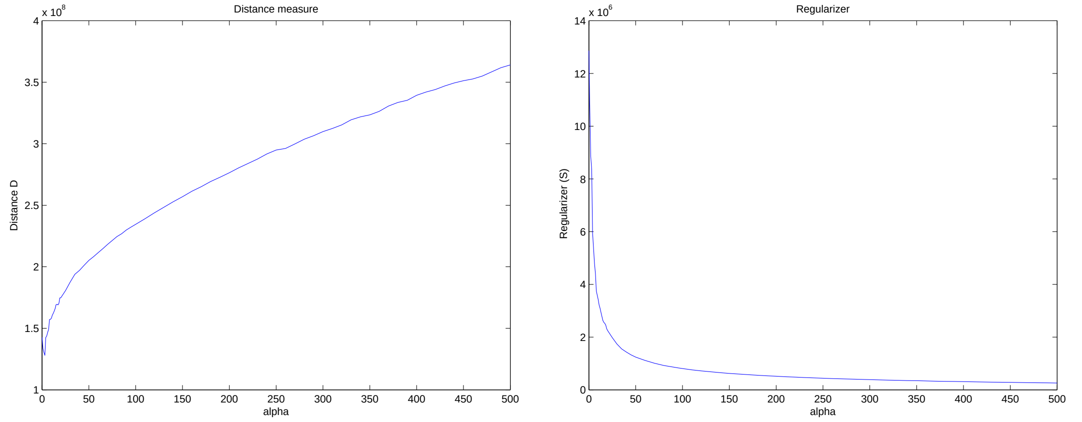
For $\alpha = 50$ the optimization led to a reduction of the functional by 69.6%. The distance term decreased by 67.4 %. This corresponds to an increase of the normalized cross-correlation from 88.6% to 97.8%.

Using the accurate spline interpolation the computational time for the optimization was about 12 minutes and took 5, 7 and 2 iterations on the respective levels.

The largest corrected deformation was 16 mm and the functional determinant of $\mathcal{I}_1$ had a range between 0.114 and 1.887 which means, that no foldings occurred.

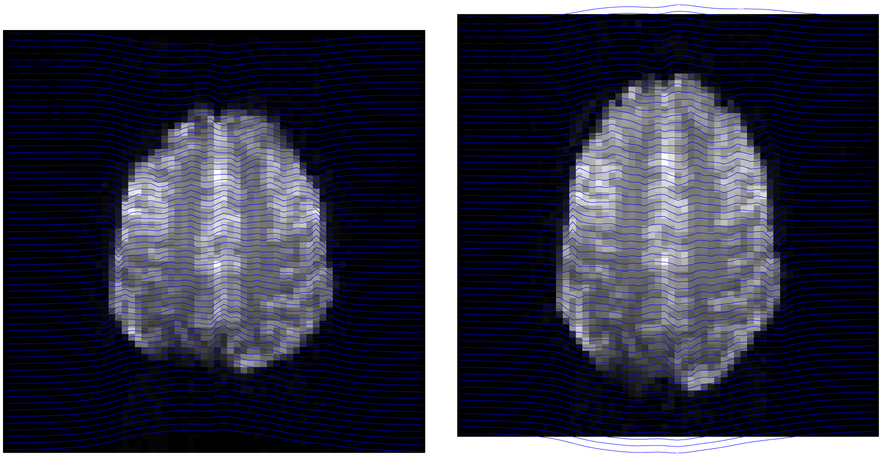


(a) l-curve (b) Number of foldings



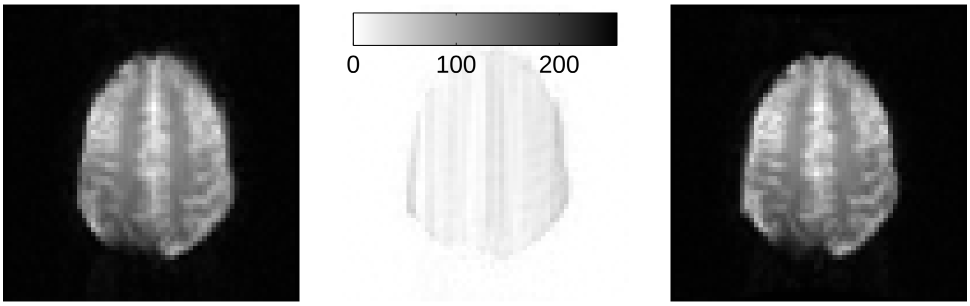
(c) Distance (d) Regularizer

Figure 2.8: Results of α-Test. According to the explanation in 2.3 77 possible α values between 0 and 500 were tested for the correction of the apa and app fMRI images of one healthy subject using linear interpolation. The top-left plot shows the l-curve that is more linearly falling on the log-log scale than "L" shaped and has an increase for weakly regularized problems. This might indicate that the correction is more likely to end up in a local minimum for small choices of α. The top right plot, gives the number of grid foldings in dependence of α. Here we can see that our regularization is capable of supplying a grid that is free of foldings for α sufficiently large (here, ≥ 50). In the bottom row, α is plotted against the distance term and the regularization term, respectively.



(a) Initial $\mathcal{I}_1$ and interpolation lines - transversal

(b) Initial $\mathcal{I}_2$ and interpolation lines - transversal

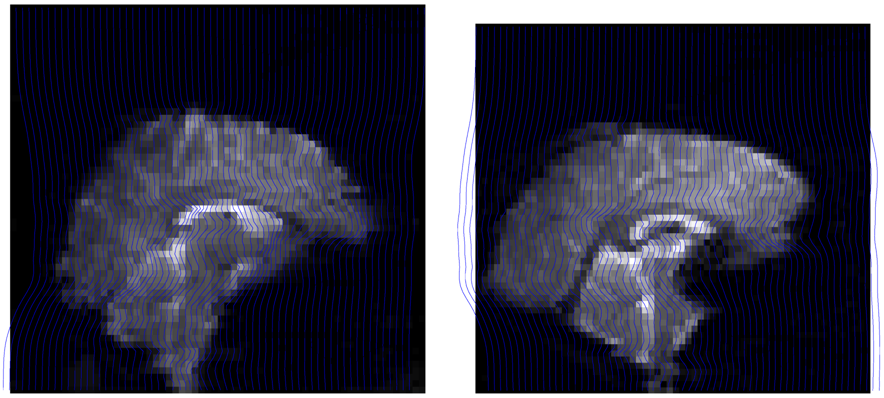


(c) Transformed $\mathcal{I}_1$ - transversal

(d) Difference - transversal

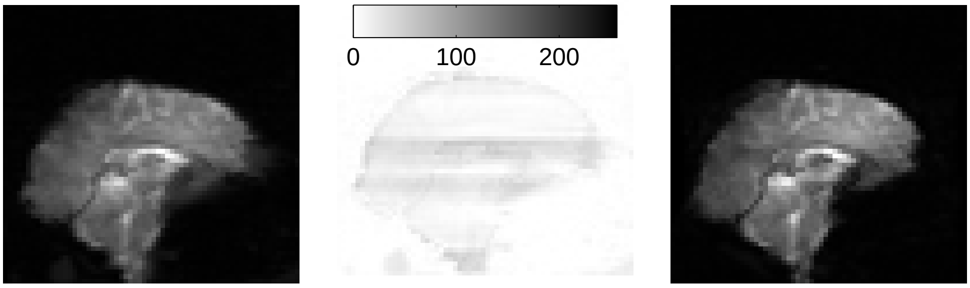
(e) Transformed $\mathcal{I}_2$ - transversal

Figure 2.9: Correction result. Transversal slices after EPI correction are shown. Compared to the initial data in Figure 2.3.2 the severe non-linear deformations of the images could be reduced significantly. However some distance remains even after the correction, which had to be expected because of the images initial $\mathcal{L}_1$ difference. More results are shown in the Appendix A.2.



(a) Initial $\mathcal{I}_1$ and interpolation lines - saggital

(b) Initial $\mathcal{I}_2$ and interpolation lines - saggital



(c) Transformed $\mathcal{I}_1$ - saggital

(d) Difference - saggital

(e) Transformed $\mathcal{I}_2$ - saggital

Figure 2.10: Correction result. Saggital slices after the EPI correction are shown. Compared to the initial data in Figure 2.3.2 the corrected images became much more alike although some difference remains. More results are shown in the Appendix A.2.

Pipeline for fMRI Correction

In order to use our algorithm for the correction of fMRI experiments we implement the following pipeline. From the two full head EPI that are acquired right before the experiment, we estimate the field inhomogeneity. Afterwards we rigidly register all images of the experiment session to the full head image. Once they are all in this reference frame we are able to correct all images with our estimated inhomogeneity. Since the statistical evaluation is not affected by multiplying constants, the Jacobian modulation is optional. For theories sake we, however, decided to apply it.

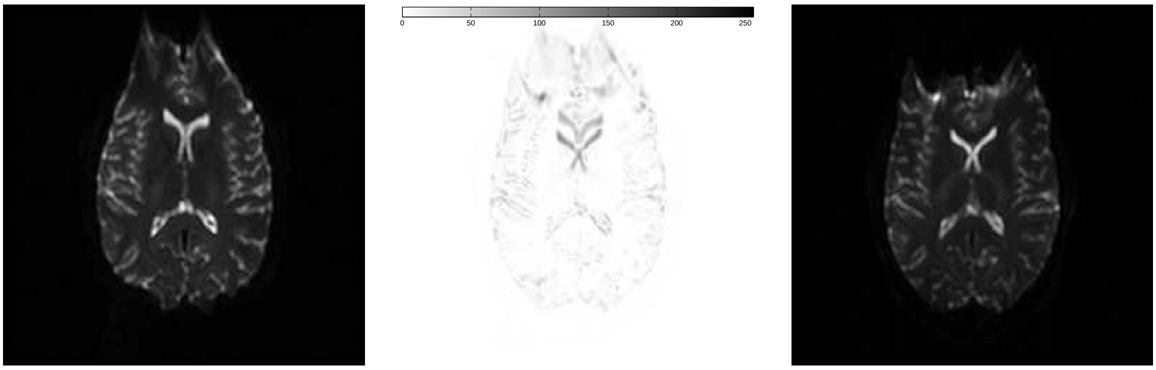
2.3.3 Diffusion Tensor Imaging

Another important MRI sequence that is based on spin-echo echo planar images is *diffusion tensor imaging* (DTI). In contrast to other imaging techniques, with DTI, one aims to acquire a tensor valued image of the human brain that give information about the water diffusion. In order to measure the diffusion one uses n additional diffusion gradients and acquires $n+1$ echo-planar images to measure the diffusivity along the gradient directions.

The first image is measured without any diffusion gradient and therefore called *unweighted* or b=0 weighted image. Those images show an image contrast similar to T2 images. Being echo-planar images, the single DTI images are exposed to the typical distortions. This makes an accurate registration onto a T2 or a T1 image, required for fusion, very hard. Since there are situations where the tensors have to be aligned to a T1 or T2 image, we aim to correct the images for their non-linear deformations first. That is why we have to estimate the distortions out of two reference scans like in the two previous examples. Therefore we will use the information contained in the unweighted DTI images.

Initial Data

In our case we measured the water diffusion coefficients along 20 directions. Additionally we acquired one EPI image without diffusion weighting - commonly denoted by $b = 0$ image. This resulted in 21 three dimensional images each having a resolution of $256 \times 256 \times 36$ and a voxel size of $0.93 \times 0.93 \times 3.60$ mm. In order to improve the signal-to-noise ratio, we acquired one whole dataset for both gradient signs instead of just acquiring a second $b = 0$ image with altered phase-encoding direction. The time to acquire one DTI series was about 7 minutes.

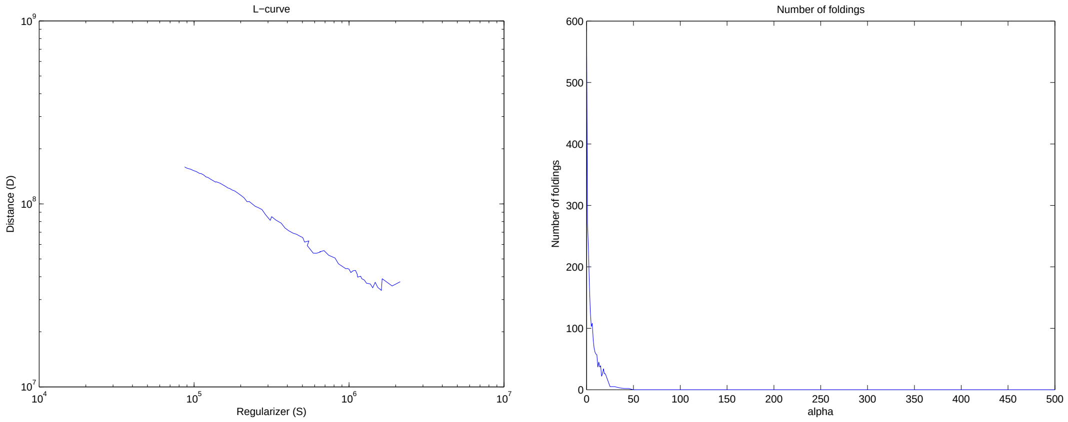


(a) Initial I_1 (b) Initial difference (c) Initial I_2

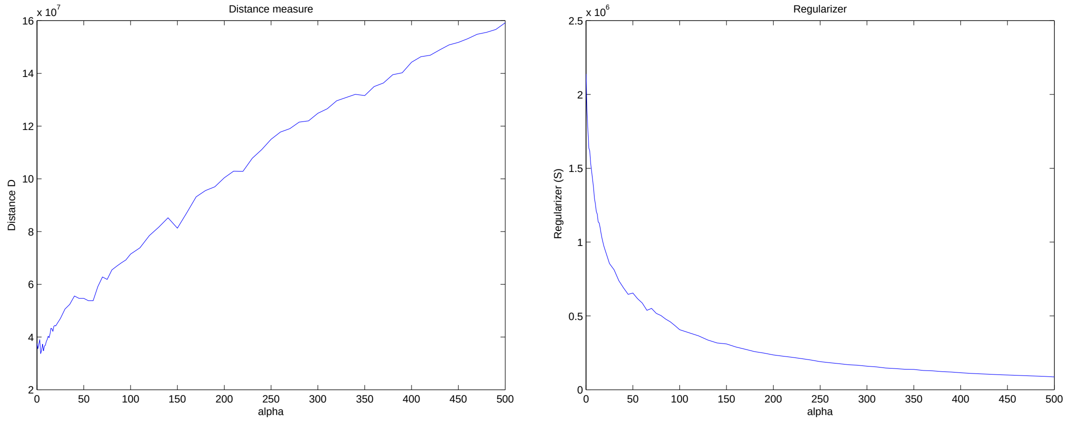
Figure 2.11: Initial Images. The same slice of I_1 and I_2 - that are identically except their altered phase encoding gradients - are shown. One can clearly observe huge displacements along the phase-encoding direction - which is altered along anterior to posterior. Secondly, one notices that the displacement direction is altered in both images. Therefore regions that are edged in I_1 happen to be stretched in I_2. We will exploit this fact in our correction algorithm.

Registration Results

For $\mu = 1$ and $\lambda = 30$ our test script stopped at $\alpha = 16$ with a folding-free grid. For smaller values of α foldings occurred. Using the accurate 3D spline interpolation the optimization of the high dimensional functional took about 80 minutes. On the coarsest level the iteration was stopped after 30 iterations. On the two finer levels, the iteration resulted in a local minimum after 13 and 6 iterations, respectively. The functional could be minimized by 87.4 % which corresponds to a decrease in the SSD distance by 94.8 %. The cross-correlation rose from 63.3 % to 97.9 %. The range of the functional determinant for $\mathcal{I}_1$ is from 0.13 to 1.89 and distortions up to 31.85 mm are corrected.

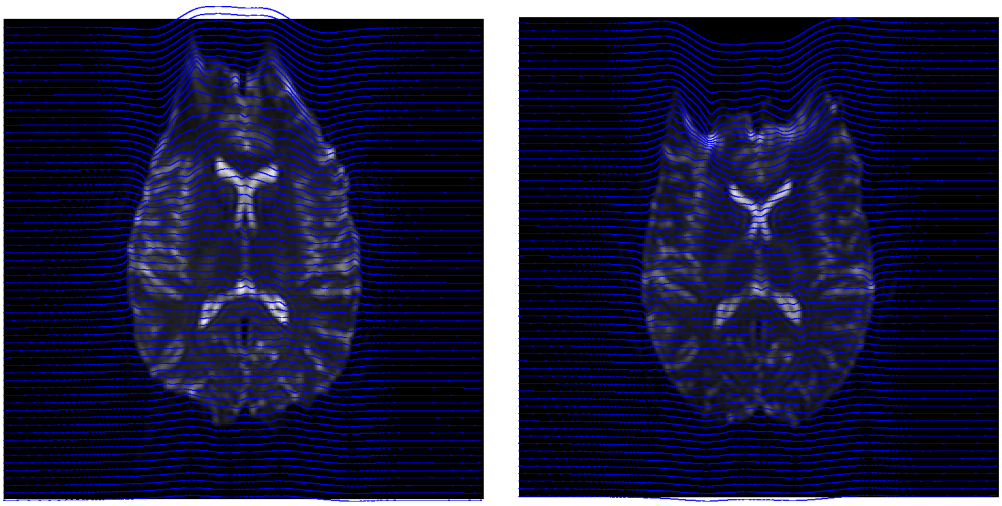


(a) l-curve (b) Number of foldings

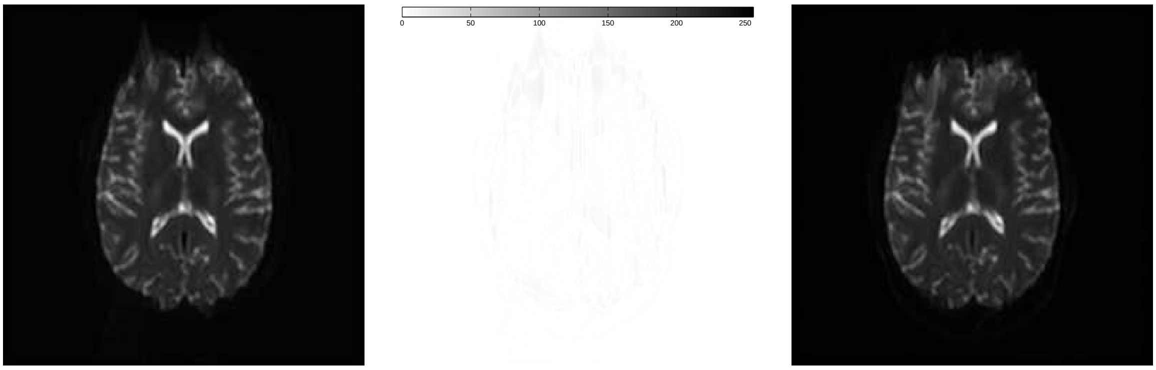


(c) Distance (d) Regularizer

Figure 2.12: Results of α-Test. According to the explanation in 2.3 77 possible α values between 0 and 500 were tested for the correction of the apa and app b=0 weighted DTI images of one healthy subject using linear interpolation. The l-curve is again linearly falling on this log-log scale. Furthermore the correction tends to be more likely end up in a local minimum whenever the problem is weakly regularized indicated by the curve's small increase on the right end. The top right plot, gives the number of grid foldings subject to α. Here we can see that our regularization is capable of supplying a grid free of foldings for α sufficiently large. The bottom row shows the distance term and the regularization term plotted against α, respectively.



(a) Initial I_1 and interpolation lines (b) Initial I_2 and interpolation lines



(c) Final I_1 (d) Final difference (e) Final I_2

Figure 2.13: Correction result for $\alpha = 16$, $\mu = 1$ and $\lambda = 30$. The same transversal slice of I_1 and I_2 as in Figure 2.3.3 is shown. The top row shows the initial images with the interpolation lines superimposed. The bottom row shows the interpolated and intensity modulated images and their difference. Compared to the initial data, the difference is decreased. However, the huge displacements in the frontal regions of several centimetres did not vanish completely. More views are shown in the Appendix, cf. A.3.

Pipeline for DTI Correction

In order to use our algorithm for the correction of DTI datasets we implement the following pipeline. We acquire one dataset for each gradient sign. That is, we obtained 2×21 3D images. In the first post-processing step we estimate the displacement field by using the $b = 0$ weighted images. Afterwards we use the FSL software package [27] to correct for eddy currents. This is done by linearly registering all 3D images to the reference frame that is given by the $b = 0$ image. We are using the affine transformation matrices to combine the affine transformation with our deformation. By these means we interpolate each image only once. Finally, the intensities were modulated and both datasets were averaged. After linearly registering the datasets to the T1 image and interpolating, we estimate the diffusion tensors again by using the FSL software package.

2.4 Discussion

In the present chapter we motivated and modelled a mass-preserving variational approach for the correction of echo-planar images. We showed the existence of a minimizer in the function space $W^{1,2}(\Omega, \mathbb{R})$. Afterwards we implemented the approach and applied it to measured data of three different types.

The algorithm was able to significantly correct the huge distortions in the measured data. This should be an important pre-processing step that should positively affect the accuracy of the respective applications.

2.5 Outlook

Because of the tricky determination of the proper regularization parameters and the fact that the choice is not physically motivated the usefulness of other regularization terms is an interesting questions. For the actual deformations are not very smooth, the regularization by using the elastic potential might negatively affect the correction. It has to be mentioned that not the material properties of the human brain, but the shape of the distortions caused by the field inhomogeneity have to be decisive.

For foldings in the grid can easily be identified by their partial derivatives one might guarantee displacement regularity by using a constrained non-linear optimization technique. This would result in 2 linear constraints for each voxel. Although this might decrease the speed of one single correction the time is easily recaptured because one omits to search for a folding free grid.

Once the deformed coordinate system is estimated, a reconstruction of the raw data in this distorted space might also be an option worth considering.

In future studies a comparison to previous approaches and an application to broader dataset is desired. Especially the evaluation of the effect on the BOLD contrast in fMRI and the reliability on the computed diffusion tensors would be interesting.

3 Motion Correction in Gated PET

In the following chapter we will develop and discuss a novel non-linear mass-preserving registration approach and apply it to the problem of motion correction in *gated Positron Emission Tomography* (PET). After a brief description of the imaging and gating technique, we will try to correct the gates for motion with standard elastic registration methods. Because of a mass-preserving property of gated PET images, however, this will not lead to optimal results and will motivate us to incorporate this knowledge directly into the registration functional.

3.1 Imaging Technique

This section briefly describes the imaging technique used later. For a more detailed view we refer to the book *Emission Tomography* by Wernick and Aarsvold [31].

Emission Tomography

Emission tomography (ET) techniques, such as *single-photon emission computed tomography* (SPECT) or *positron emission tomography* (PET) techniques use radioactive materials to image physiological properties of the body. Literally, using gamma-ray emission - or *tracer principle* - a volumetric image of the body's interior (called *tomography*) is produced. Depending on the choice of the radioactive tracer that is induced into the body, an ET image can for instance represent the spatial distribution of the glucose metabolism, as it will be the case later on.
Emission tomography techniques can be distinguished mainly by the type of radioisotope incorporated in the tracer. SPECT studies use radiopharmaceuticals labelled with a single-photon emitter, whereas PET requires the labelling isotope to be a positron (i.e. anti-electron) emitter.

Positron Emission Tomography

In PET the decay of a positron-emitting isotope leads to the emission of two gamma-rays that travel nearly in opposite directions. The patient is surrounded by a collection of gamma-ray detectors which are connected to circuity that senses the timing of the gamma-ray detections. When two gamma-rays are detected almost simultaneously it is inferred that these two rays must have resulted from the same decay event. Thus, it is assumed that the origin of the decay event lies on the line segment connecting the two participating detectors.

In order to be flexible about the image reconstruction we will use a *list mode* acquisition. In list mode, the time of detection of every gamma-ray event is recorded into a list in addition to its spatial coordinates and other parameters such as photon energy. Furthermore the acquisition is *gated* with respect to the patients breathing cycle - we say it is *respiratory gated.*

Gating

In a gated study the data acquisition is synchronized to the rhythm of the patients cardiac or breathing cycles. For instance, in cardiac studies, one acquires a short sequence of, say 10 frames, that depicts the motion taking place during one cardiac cycle of the patient. This single heartbeat is actually a composite representation formed by averaging over the imaging data collected during a very large number of cardiac cycles. The purpose of this compositing operation is to reduce the effect of noise. Each cardiac cycle is much too short to acquire sufficient gamma-ray counts to produce a useful image, but the composite image has acceptable quality because it is reconstructed from all the counts acquired during the study.

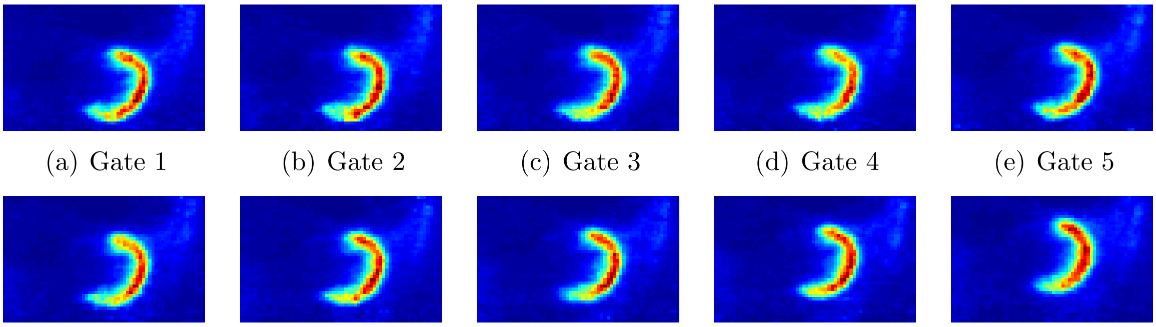
In respiratory gated imaging, a camera is used to detect and divide the breathing cycles into the desired number of gates. A detailed examination and comparison of cardiac and respiratory gating techniques can be found for instance in [4].

Motion Correction

The main objective for gating the images is to avoid image blurring due to the motion that takes place during the long acquisition of PET images. That is why we aim to reconstruct the image in one common reference frame. This step clearly requires an accurate image registration of the gated images in order to detect the motion. In this present chapter we will - under the assumption that the human heart behaves elastically - try to use the non-linear registration techniques presented in Chapter 1. As it turns out, the standard registration approach does not lead to the desired result, because it disregards the assumptions, that no tracer uptake is lost between the different phases of breathing cycle and uses an inappropriate assumption of conserved local intensities. This will motivate us to develop a new mass-preserving non-linear registration algorithm in order to estimate the respiratory motion.

Data

Respiratory gated heart images of a human patient are acquired on a PET/CT scanner. After respiratory gating ten images are reconstructed individually each corresponding to one phase of the breathing cycle. The grey scale images are of a resolution of $64 \times 64 \times 40$. One slice of the data is shown in Figure 3.1.


(a) Gate 1 (b) Gate 2 (c) Gate 3 (d) Gate 4 (e) Gate 5

(f) Gate 6 (g) Gate 7 (h) Gate 8 (i) Gate 9 (j) Gate 10

Figure 3.1: Initial data: Same slice of all ten gates of one subjects heart. The respiratory motion causes an upward movement of the heart from gate one to gate ten.

3.2 Elastic Registration

In order to map all different gates into the reference gate - in our case gate ten - we use non-parametric elastic registration techniques. As described in Chapter 1 this results in the minimization of the following functional:

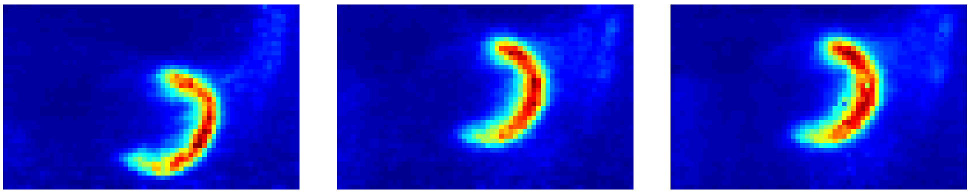
$$\mathcal{J}(y) = ||\mathcal{T}(y) - \mathcal{R}||_2^2 + \alpha \mathcal{S}[y],$$

where $\mathcal{S}$ measures the elastic potential of the non-parametric transformation y and $\alpha \in \mathbb{R}^+$ is a regularization parameter.

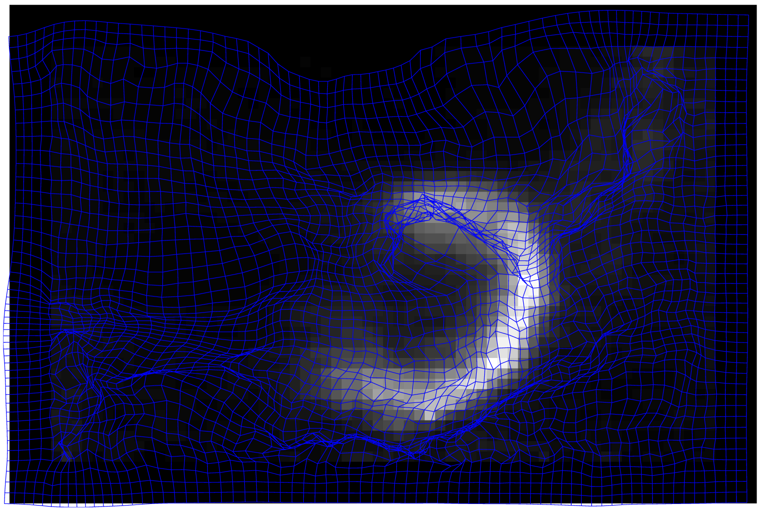
Numerical Set up

Exemplarily, we choose gate one as the template image $\mathcal{T}$ and gate ten to be the reference image $\mathcal{R}$. Those two frames show the most movement and are thus the most challenging to register of the nine possible pairs. We choose $\alpha = 150$ and pad the images with zeros at the boundaries. We perform three multi level steps - $18 \times 18 \times 12$, $36 \times 36 \times 24$ and $72 \times 72 \times 48$ - ending up at full resolution. The maximum number of optimization iterations is chosen to be 30.

Registration Results



(a) Initial Template (b) Reference (c) Final Template



(d) Transformation grid

Figure 3.2: Registration results for standard non-parametric 3D elastic registration of gate four (a) to gate ten (b). At first glance, the transformed template (c) resembles the reference image very well. However, the transformation shows many foldings, because the registration algorithm has to avoid the evaluation of the template image in very bright regions.

After 311 seconds and 10, 8 and 4 iterations the registration stopped in a local minimum. On the one hand the functional was minimized by 97.7% and the distance between the two gates was reduced by 98.3%. Furthermore the normalised cross-correlation increased from 66.4% to 99.4%. On the other hand the transformation grid shows many foldings. The obtained deformation does not resemble the respiratory motion. Clearly, the number of foldings could be decreased - or even avoided - for α sufficiently large, but this would presumably result in a larger remaining distance. The registration result is depicted in Figure 3.2.

Discussion

After registration, the transformed template is very similar to the reference image. Additionally the decrease of the distance and the accompanied increase in the cross-correlation are very promising. However, we were searching for a *reasonable* transformation providing such similarity. From the deformation grid we observe, that the transformation does not resemble the breathing movement very well. This is due to the fact that the transformation model is not appropriate for this problem. As mentioned above - in contrast to respiration - the non-parametric transformation model is in general not mass-preserving, since it only interpolates the template image on a deformed grid without accounting for volume changes. With regard to the very bright regions in the template image, which do not occur in the reference image, the behaviour is quite reasonable. The optimizer avoids as good as possible the evaluation of the template image in such regions. Thus, intensity is lost in those regions. That is one reason why the standard registration algorithm does not lead to a satisfying registration result in this case. Since the gated images are acquired in the same time period, no significant difference in activity is to be expected. Our transformation model does on the other hand not account for this preservation because in general

$$\int_\Omega \mathcal{T}(x)dx \neq \int_\Omega \mathcal{T}(y(x))dx. \tag{3.1}$$

To sum up, the registration result was two-sided. On the one hand we were able to significantly reduce the distance measure and make the two images very alike with respect to their SSD difference and their normalized cross-correlation. However, the transformation is far from being one-to-one as can be seen in Figure 3.2(d). In addition, we did not account for one given assumption, namely the conservation of signal. In the following, this prior knowledge about the images is incorporated directly into the registration problem.

3.3 Elastic Mass-Preserving Registration

As mentioned above, all gates contain almost the same amount of activity. In order to find a reasonable - i.e. one-to-one - and mass-preserving transformation we include this prior knowledge into our functional.

Developing the functional

We will now assume that our transformation $y : \mathbb{R}^3 \to \mathbb{R}^3$ is diffeomorphic. Then we know by the integration by substitution theorem

$$\int_{y(\Omega)} \mathcal{T}(x)dx = \int_\Omega \mathcal{T}(y(x)) \left| det(Dy(x)) \right| dx. \tag{3.2}$$

Where Dy is the Jacobian of our transformation, i.e.

$$Dy = \left(\begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} (y_1, y_2, y_3) \right)^T = \begin{pmatrix} \partial_1 y_1 & \partial_2 y_1 & \partial_3 y_1 \\ \partial_1 y_2 & \partial_2 y_2 & \partial_3 y_2 \\ \partial_1 y_3 & \partial_2 y_3 & \partial_3 y_3 \end{pmatrix}. \tag{3.3}$$

Therefore, when transforming the template image, we will not only apply the geometric transformation y, but also change the intensity by multiplication with the functional determinant $|\det(Dy(x))|$. Thereby, we account for changes in the images intensity due to the volume change of the transformation. Whenever the transformation enlarges the volume of one voxel, the voxels intensity is distributed over a larger domain and thus has to be decreased. By these means we account for mass preservation. Our functional then reads

$$\mathcal{J}^{\mathrm{MP}}(y) = ||\mathcal{T}(y) \cdot |\det(Dy)| - \mathcal{R}||_2^2 + \alpha\mathcal{S}[y]. \tag{3.4}$$

In addition we are interested in diffeomorphic and orientation preserving transformations. Therefore the functional determinant should be positive and we can drop the absolute value

$$\mathcal{J}^{\mathrm{MP}}(y) = ||\mathcal{T}(y) \cdot \det(Dy) - \mathcal{R}||_2^2 + \alpha\mathcal{S}[y]. \tag{3.5}$$

As in the standard elastic registration algorithm above, we make use of a Gauss-Newton optimization scheme. Therefore we need to compute the first variation of our distance functional.

Definition 19 (Lagrangian). *The associated* Lagrangian *to the distance term in (3.5) is a smooth function*

$$L(P, z, x) : \mathbb{R}^{3\times 3} : \mathbb{R}^3 \times \Omega \to \mathbb{R}$$

that is specified for given images $\mathcal{T}, \mathcal{R} \in Img(\Omega)$ *as*

$$\begin{aligned} L(P, z, x) &= \frac{1}{2}\left(\mathcal{T}(z(x))\det(P(x)) - \mathcal{R}(x)\right)^2 \\ &= \frac{1}{2}\left(r(\mathcal{T}, \mathcal{R}; P, z, x)\right)^2. \end{aligned}$$

Since P is the placeholder for the Jacobian of y we introduce the following notation in analogy to (3.3)

$$P := \begin{pmatrix} p_1^1 & p_2^1 & p_3^1 \\ p_1^2 & p_2^2 & p_3^2 \\ p_1^3 & p_2^3 & p_3^3 \end{pmatrix}.$$

Furthermore L_P and L_z denote the derivative of L with respect to P and z respectively, i.e.

$$L_P = \begin{pmatrix} L_{p_1^1} & L_{p_2^1} & L_{p_3^1} \\ L_{p_1^2} & L_{p_2^2} & L_{p_3^2} \\ L_{p_1^3} & L_{p_2^3} & L_{p_3^3} \end{pmatrix} \text{ and } L_z = \begin{pmatrix} L_{z^1} \\ L_{z^2} \\ L_{z^3} \end{pmatrix}.$$

Notation 2 (Matrix divergence). *Let $F : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ be a matrix valued differentiable function on $\mathbb{R}^3$. Then we define its divergence as*

$$\nabla \cdot F(x) = \begin{pmatrix} \nabla \cdot F^1 \\ \nabla \cdot F^2 \\ \nabla \cdot F^3 \end{pmatrix} = \begin{pmatrix} \partial_1 f_1^1 + \partial_2 f_2^1 + \partial_3 f_3^1 \\ \partial_1 f_1^2 + \partial_2 f_2^2 + \partial_3 f_3^2 \\ \partial_1 f_1^3 + \partial_2 f_2^3 + \partial_3 f_3^3 \end{pmatrix}.$$

Lemma 4 (First variation in divergence form). *Let $L : \mathbb{R}^{3\times 3} \times \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ be a smooth Lagrangian function. For any smooth function $w : \Omega \to \mathbb{R}^3$ we define the associated functional as*

$$I[w] := \int_\Omega L(Dw(x), w(x), x)dx.$$

A smooth minimizer $u = (u^1, u^2, u^3)$ of $I[\cdot]$ must then solve the following system of non linear partial differential equations.

$$-\nabla \cdot L_p(Du, u, x) + L_z(Du, u, x) = 0 \text{ in } \Omega.$$

Proof. Let $v = (v^1, v^2, v^3) \in C_0^\infty(\Omega, \mathbb{R}^3)$ be a fixed test function. Then we can define a scalar function

$$i(\epsilon) := I[u + \epsilon v].$$

The necessary condition for u being a minimizer is then

$$i'(0) = 0.$$

Let us now calculate $i'(0)$

$$i'(0) = \int_\Omega \sum_{i=1}^{3} \sum_{k=1}^{3} L_{p_i^k}(Du, u, x) v_{x_i}^k + \sum_{k=1}^{3} L_{z^k}(Du, u, x) v^k dx.$$

Because of the compact support of v, integration by parts in the first summand leads to

$$= \int_\Omega \sum_{i=1}^{3} \sum_{k=1}^{3} - \Big(L_{p_i^k}(Du, u, x) \Big)_{x_i} v^k + \sum_{k=1}^{3} L_{z^k}(Du, u, x) v^k dx.$$

And since this identity is valid for all choices of v^1, v^2, v^3 we can conclude

$$- \sum_{i=1}^{n} \Big(L_{p_i^k}(Du, u, x) \Big)_{x_i} + L_{z^k}(Du, u, x) = 0 \text{ in } \Omega \text{ for } k = 1, 2, 3.$$

This can compactly be written as

$$-\nabla \cdot L_P(Du, u, x) + L_z(Du, u, x) = 0 \text{ in } \Omega.$$

□

Lemma 5 (Derivative of the determinant)**.** *The determinant* $\det : \mathbb{R}^{3\times3} \to \mathbb{R}$ *is a continuously differentiable scalar function. Its derivative at a matrix A is given by*

$$\nabla \det(A) = \begin{pmatrix} \det A_1^1 & -\det A_2^1 & \det A_3^1 \\ -\det A_1^2 & \det A_2^2 & -\det A_3^2 \\ \det A_1^3 & -\det A_2^3 & \det A_3^3 \end{pmatrix}.$$

Where A_i^j stands for the 2×2 matrix one obtains by deleting the i-th column and the j-th row.

Proof. By Laplace's rule we can develop the determinant to its j-th row by

$$\det A = \sum_{i=1}^{3} (-1)^{i+j} a_i^j \det A_i^j.$$

Thereby we can compute the elements of $\nabla \det$ as

$$\frac{\partial \det}{\partial a_i^j} = (-1)^{i+j} \det A_i^j.$$

□

In order to obtain the first variation of our distance term, we will now use the Lagrangian framework.

Theorem 12 (First variation). *Let $\mathcal{T}, \mathcal{R} \in C^1(\Omega, \mathbb{R})$ and $y \in C^2(\Omega, \mathbb{R}^3)$ be continuously differentiable. The first variation of the distance term (3.5) is then given by*

$$d\mathcal{D}(y) = r(y)\det(Dy)\nabla\mathcal{T} - \nabla \cdot (r(y)\mathcal{T}\ \nabla\det(Dy)).$$

where $r(y) := r(\mathcal{T}, \mathcal{R}; y)$ and $\nabla\cdot$ is the matrix divergence defined in Notation 2.

Proof. As we know from Lemma 4 the first variation in its divergence form is given as

$$-\nabla \cdot L_P(P, z, x) + L_z(P, z, x) = 0 \text{ in } \Omega. \tag{3.6}$$

We compute the parts individually. Because of

$$L_{z^k}(P, z, x) = r(\mathcal{T}, \mathcal{R}; P, z, x)\ \det(P)\ \partial_k\mathcal{T}(z)\ ,\ k = 1, 2, 3$$

L_z becomes

$$L_z(P, z, x) = r(\mathcal{T}, \mathcal{R}; P, z, x)\ \det(P)\ \nabla\mathcal{T}(z).$$

For $i = 1, 2, 3$, we have

$$\begin{aligned} L_{p_i^k} &= L_{p_i^k}(P, z, x) \\ &= r(\mathcal{T}, \mathcal{R}; P, z, x)\ \mathcal{T}(z)\ \frac{\partial}{\partial p_i^k}\det(P). \end{aligned}$$

Consequently L_P becomes

$$L_P = r(\mathcal{T}, \mathcal{R}; P, z, x)\ \mathcal{T}(z)\ \nabla\det(P).$$

Inserting L_z and L_P into (3.6) gives the assertion.

□

Implementation 6 (Derivative of the residual). *Following a discretize-then-optimize strategy we need to calculate the gradient of the residual function $r : \mathbb{R}^9 \to \mathbb{R}^3$ defined by*

$$r(y) = \mathcal{T}(y)\det(Dy) - \mathcal{R}.$$

By chain rule we note that

$$\nabla_{y_i} r = \partial_i\mathcal{T}(y)\det(Dy) + \mathcal{T}(y)\ \sum_{j=1}^{3}(\nabla_A\det(Dy))_j^i \cdot \nabla(\partial_j^+ y_i)\ , i = 1, 2, 3.$$

Since ∂_j^+ is a linear differential operator that is implemented as a sparse matrix

$$\nabla(\partial_j^x y_i) = \delta_{i,j}\partial_j^+\ i = 1, 2, 3.$$

In Matlab *this operation can be implemented as a matrix-matrix multiplication via*

```
 dOp =  [d1Op 0 0; 0 d1Op 0; 0 0 d1Op;
 d2Op 0 0; 0 d2Op 0; 0 0 d2Op;
d3Op 0 0; 0 d3Op 0; 0 0 d3Op];
 %diOp is the sparse i-th partial differential operator as a sparse matrix
 gradDetDydiag = spdiags(reshape(gradDetDy,[],9), 0:n:9*(n-1), n, 9*n);
 dr = gradDetDydiag * dOp;
```

Existence of a minimizer

We examine whether our functional actually has one minimum in a suitable function space. We restrict this discussion to the case $\alpha > 0$. In the current formulation we can not expect the existence of a minimizer. Therefore we change the functional to

$$\mathcal{J}^{\mathrm{MP}}(y) = \frac{1}{2}\int_\Omega (\mathcal{T}(y)\det(Dy) - \mathcal{R}(x))^2 dx + \frac{\alpha}{6}\int_\Omega |\mathcal{B}[y - y_{\mathrm{Ref}}]|^6 dx.$$

Note that the exponent of the elastic regularizer is changed which affects the derivatives of our functional and thereby the optimization.
This adjustment, however, allows us to show the existence of a minimizer in the Sobolev space $W^{1,6}(\Omega, \mathbb{R}^3)$. In order to do so, we use the framework of polyconvex Lagrangians to show the lower semi continuity of $\mathcal{J}^{\mathrm{MP}}$. Then the fundamental theorem of Optimization gives us the desired existence.

Definition 20 (Lagrangian)**.** *The Lagrangian of our joint functional is a smooth function*

$$L(P, z, x) : \mathbb{R}^{3\times 3} \times \mathbb{R}^3 \times \Omega \to \mathbb{R},$$

given by

$$L(P, z, x) = \frac{1}{2}\left(\mathcal{T}(z)\det P - \mathcal{R}(x)\right)^2 + \frac{\alpha}{2}\left(\mu \sum_{i,j=1}^{3} |p_i^j|^4 + (\mu + \lambda)\sum_{i=1}^{3} |p_i^i|^4\right).$$

Lemma 6 (Coercivity)**.** *The functional for the mass-preserving elastic registration $\mathcal{J}^{MP}$ is coercive on $W^{1,6}(\Omega, \mathbb{R}^3)$, i.e. there are constants $\delta > 0, \gamma \geq 0$ such that*

$$\mathcal{J}^{MP}(y) \geq \delta ||Dy||_{L^4}^4 - \gamma.$$

Proof. As in Chapter 2, we try to show that there exist constants $\omega > 0, \beta \geq 0$ such that

$$L(P, z, x) \geq \omega |P|^6 - \beta \ , \ \forall P \in \mathbb{R}^{3\times 3}, z \in \mathbb{R}^3, x \in \Omega.$$

By definition

$$\begin{aligned} L(P, z, x) &= \frac{1}{2}\left(\mathcal{T}(z)\det P - \mathcal{R}(x)\right)^2 + \frac{\alpha}{6}\left(\mu \sum_{i,j=1}^{3} |p_i^j|^6 + (\mu + \lambda)\sum_{i=1}^{3} |p_i^i|^6\right) \\ &\geq \frac{\alpha}{6}\left(\mu \sum_{i,j=1}^{3} |p_i^j|^6 + (\mu + \lambda)\sum_{i=1}^{3} |p_i^i|^6\right) \\ &\geq \frac{\alpha}{6}\mu \sum_{i,j=1}^{3} |p_i^j|^6 \\ &= \frac{\alpha}{6}\mu |P|^6. \end{aligned}$$

By defining now $\beta := 0$ and $\omega := \frac{\alpha}{6}\mu$ the proof is complete ☐

Because of the functional determinant in the distance term our Lagrangian is not convex in P. Fortunately, there is the framework of *polyconvexity*, described for instance by Evans in [10], that allows us to use methods of non linear functional analysis in order to show the weak lower semicontinuity of $\mathcal{J}^{\mathrm{MP}}$ in $W^{1,6}$.

Definition 21 (Polyconvexity)**.** *Let L be the Lagrangian of a functional, not necessarily convex in P. Let L be of the shape*

$$L(P, z, x) = F(P, \det P, z, x)$$

where $F : \mathbb{R}^{3\times 3} \times \mathbb{R} \times \mathbb{R}^3 \times \Omega \to \mathbb{R}$ is smooth. If the joint mapping

$$(P, r) \mapsto F(P, r, z, x) \textit{ is convex for each fixed } z \in \mathbb{R}^3, x \in \Omega$$

we say L is polyconvex.

Lemma 7. *The Lagrangian of our joint functional is polyconvex.*

Proof. First of all, L is of the form (21) in the previous definition.

$$F(P, r, z, x) = \frac{1}{2}\left(\mathcal{T}(z)r - \mathcal{R}(x)\right)^2 + \frac{\alpha}{6}\left(\mu \sum_{i,j=1}^{3} |p_i^j|^6 + (\mu + \lambda) \sum_{i=1}^{3} |p_i^i|^6\right).$$

Since F is smooth, only the convexity in P and r remains to be shown. If $\alpha = 0$, the convexity in P is trivial, since it does not occur in the distance term. For $\alpha > 0$, it can be deduced directly from the convexity of the Lagrangian of the elastic potential. The convexity in r can be seen by the following computations restricted to the distance term. Let $i_1 := \mathcal{T}(z), i_2 := \mathcal{R}(x) \in \mathbb{R}$ be fixed and $t \in [0, 1]$ and $r_1, r_2 \in \mathbb{R}$. We then have to show

$$\begin{gathered} F(tr_1 + (1-t)r_2) \leq t\, F(r_1) + (1-t)F(r_2) \\ (i_1(tr_1 + (1-t)r_2) - i_2)^2 \leq t(i_1r_1 - i_2)^2 + (1-t)(i_1r_2 - i_2)^2. \end{gathered}$$

The left hand side is

$$\begin{aligned} (i_1(tr_1 + (1-t)r_2) - i_2)^2 =& t^2(i_1r_1)^2 + 2t(1-t)i_1^2r_1r_2 - 2ti_2i_1r_1 \\ &+ (1-t)^2(i_1r_2)^2 - 2(1-t)i_2i_1r_2 + i_2^2 \end{aligned}$$

whereas the right hand side reads

$$\begin{aligned} t(i_1r_1 - i_2)^2 + (1-t)(i_1r_2 - i_2)^2 =& t(i_1r_1)^2 - 2ti_2i_1r_1 + ti_2^2 + (1-r)^2(i_1r_2)^2 \\ &- 2(1-t)i_2i_1r_2 + (1-t)i_2^2. \end{aligned}$$

After eliminating all terms occurring on both sides one ends up with

$$(ti_1r_1 + (1-t)i_1r_2)^2 \leq t(i_1r_1)^2 + (1-t)(i_1r_2)^2.$$

□

The advantage of the concept of polyconvexity lies in the fact that determinants are weak continuous.

Lemma 8 (Weak continuity of determinants)**.** *Assume* $3 < q < \infty$ *and*

$$u_k \rightharpoonup u \text{ weakly in } W^{1,q}(\Omega, \mathbb{R}^3).$$

Then

$$\det Du_k \rightharpoonup \det Du \text{ weakly in } L^{\frac{q}{3}}(\Omega).$$

Proof. Evans [10], page 454. □

The lemma above motivates the change of the exponent of the elastic differential operator. It would be straight forward to choose $q = 4$ in order to show existence. However, due to the L_2 data term we need weak convergence of the determinant in L_2 and consequently we need $q := 6$.

Lemma 9. *Suppose* $3 < q < \infty$ *and large enough so that all integrals exist. Assume also that L is bounded below and polyconvex. Then* $\mathcal{F}$ *is weakly lower semicontinuous on* $W^{1,q}(\Omega, \mathbb{R}^3)$*.*

Proof. Evans [10], page 456 ff □

Now we can combine our findings and obtain

Theorem 13. *Our functional* $\mathcal{J}^{MP}$ *has at least one minimizer in* $W^{1,6}(\Omega, \mathbb{R}^3)$*.*

Proof. As we have seen above, $\mathcal{J}^{\mathrm{MP}}$ is coercive on $W^{1,6}(\Omega, \mathbb{R}^3)$ and also lower semicontinuous. This proves the existence via the fundamental theorem of optimization. □

Implementation

In order to integrate our new VAriational Mass-Preserving Image REgistration approach into the freely available *FAIR* package by Jan Modersitzki [19] we create a new objective function

```
function [Jc,para,dJ,H] = VAMPIREobjFctn3(T,Rc,omega,m,yRef,yc)
```

Here, `T` are the template images interpolation coefficients, `Rc` is the reference image, `omega` the image domain, `m` the number of discretization points, `yRef` a staggered reference grid and `yc` the current transformation. The output values `Jc`, `dJ` and `H` stand for the value of the objective function and its first and second derivative. Information about the optimization progress is stored in `para`. The grid change operator from staggered to cell-centred grids `P` and the differential operator ∇D are stored as `persistent` variables and only updated on a new level.

```
persistent P dOp
```

The first step in transforming the template image is the interpolation at the deformed grid `yc`.

```
yc = center(yc, m);
% compute interpolated image and derivative
[Tc,dT] = inter(T,omega, yc,'doDerivative',doDerivative);
```

We then compute the intensity modulation `detDy` and its derivative `gradDetDy` in the new function `detFunctions`. The intensity modulation is then applied to the deformed template image.

```
[detDy, ~, ~, gradDetDy] = detFunctions(yc, omega, m, doDerivative);
Tcold   = Tc;
Tc      = Tc .* detDy;
```

Afterwards the SSD distance, the value of the regularization term and if necessary its derivatives are computed.

```
% compute distance measure
[Dc,rc,~,~,d2psi] = distance(Tc,Rc,omega,m,'doDerivative',doDerivative);
% compute regularizer
alpha = regularizer2('get','alpha');
[Sc,dS,d2S] = regularizer2(ys,omega,m,'doDerivative',doDerivative);
% evaluate joint function and return if no derivatives need to be computed
Jc = Dc + alpha *  Sc;
```

If derivatives are required we calculate the first variation in its divergence form. Here `div` implements the matrix divergence defined in Notation 2. Because of the partial integration argument in Lemma 4 it is of great importance to compute `div` using the adjoint finite difference operator of the one chosen in `detFunctions` .

```
Lz = rc' * (spdiags(detDy,0,n,n) * dT);
Lp = repmat(rc .* Tcold, [dim dim]) .* gradDetDy;
divLp = div(Lp, omega, m);
dD = Lz - divLp';
dD = hd * dD; % hd = voxel volume
dJ = dD * P + alpha * dS;
```

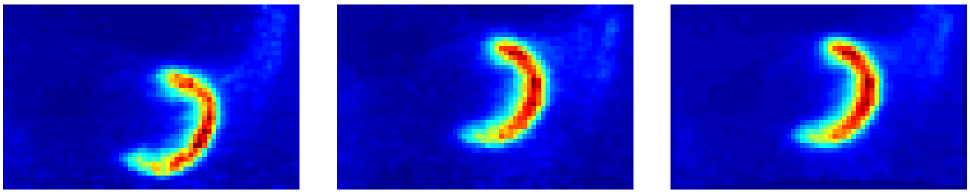
In order to approximate the Hessian for the Gauss-Newton scheme, the residuals derivative is needed. Here we use the differential operator stored in `dOp` as described in Implementation 6.

```
gradDetDydiag = spdiags(reshape(gradDetDy,[],dim^2), 0:n:dim^2*(n-1), n, dim^2*n);
dr = spdiags(detDy,0,n,n) * dT + spdiags(Tcold,0,n,n) * gradDetDydiag * dOp;
dr = dr * P;
H  = d2psi * (dr' * dr) + alpha *  d2S;
```

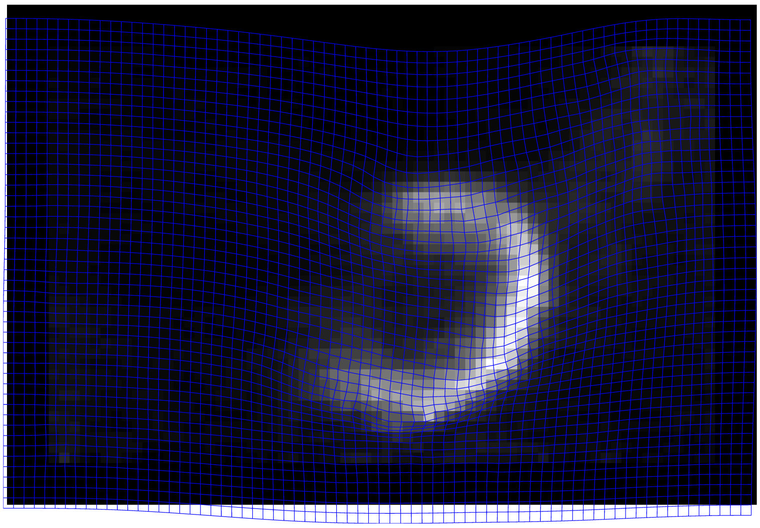
We then minimise the objective function by the methods supplied by *FAIR* using a multilevel strategy.

Registration Results

Using the same dataset, parameters and image pair as in the standard non-linear registration we were able to reduce the new functional by 98.1%. The cross correlation increased to 99.95% and the range of the Jacobian modulation is $[0.33, 2.80]$. Although we can ensure the existence of a solution only in case of a $\mathcal{L}_6$ regularization, we use the same $\mathcal{L}_2$ based elastic potential as in Section 3.2. For our low dimensional data, this led to a smooth and folding-free deformation depicted in Figure(3.3(d)). Most importantly, the transformation does resemble the respiratory motion very accurate.



(a) Initial template (b) Reference (c) Final template



(d) Transformation grid

Figure 3.3: Registration results for mass-preserving registration. The deformed template resembles the reference image very well. Compared to Figure (3.2(d)) the grid has got no foldings. Additionally the deformation seems to reasonably resemble the respiratory motion that occurred between the different gates.

Motivated by this positive result we start registering the other gates onto gate 10 as well. For each gate mass-preserving registration finds smooth and folding-free deformation fields that are able to make the images very similar to the reference gate.

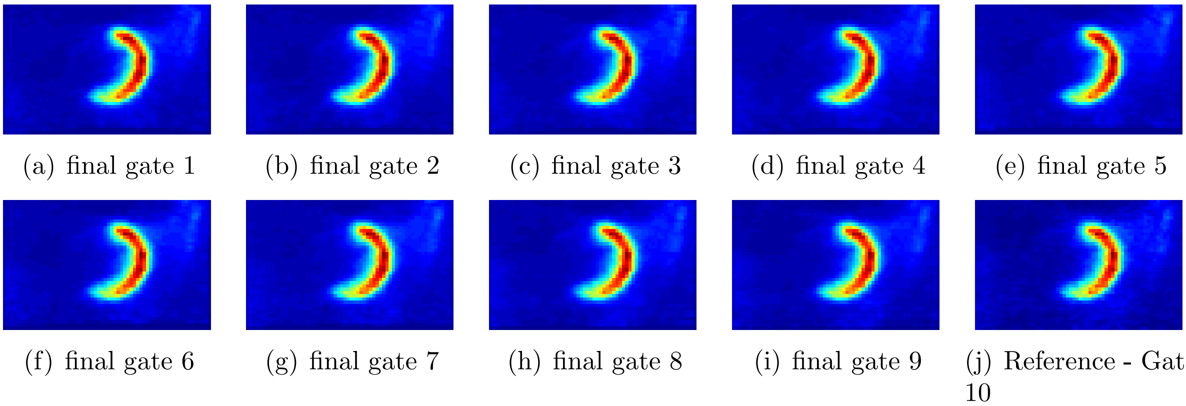

(a) final gate 1 (b) final gate 2 (c) final gate 3 (d) final gate 4 (e) final gate 5
(f) final gate 6 (g) final gate 7 (h) final gate 8 (i) final gate 9 (j) Reference - Gate 10

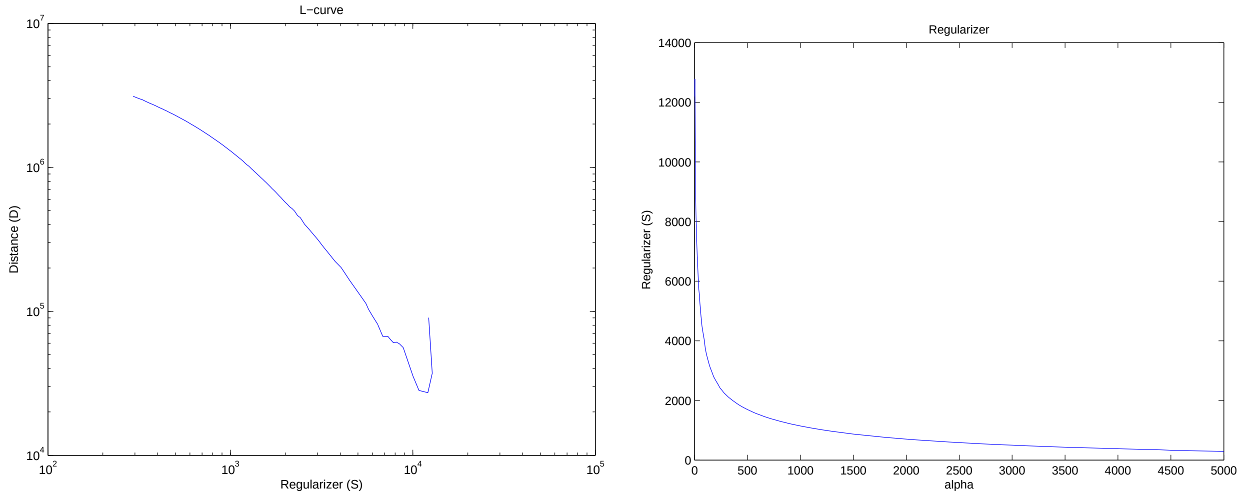
Figure 3.4: Registration result: Same slice of the mass-preserving registered nine gates and the reference gate. The new non-linear registration approach was able to render the images almost identical, which is also indicated by the cross-correlation in Table 3.3. Furthermore the transformation grids were free of foldings and resembled the actual motion.

	Gate 1	Gate 2	Gate 3	Gate 4	Gate 5	Gate 6	Gate 7	Gate 8	Gate 9
cc	0.98976	0.99019	0.99067	0.99116	0.99248	0.99336	0.99415	0.9953	0.99632
ccold	0.50346	0.5216	0.53894	0.56461	0.61275	0.66193	0.73016	0.80629	0.90132
maxvec	7.437	7.3044	6.9075	6.5197	5.9571	5.2871	4.502	3.5268	2.3524

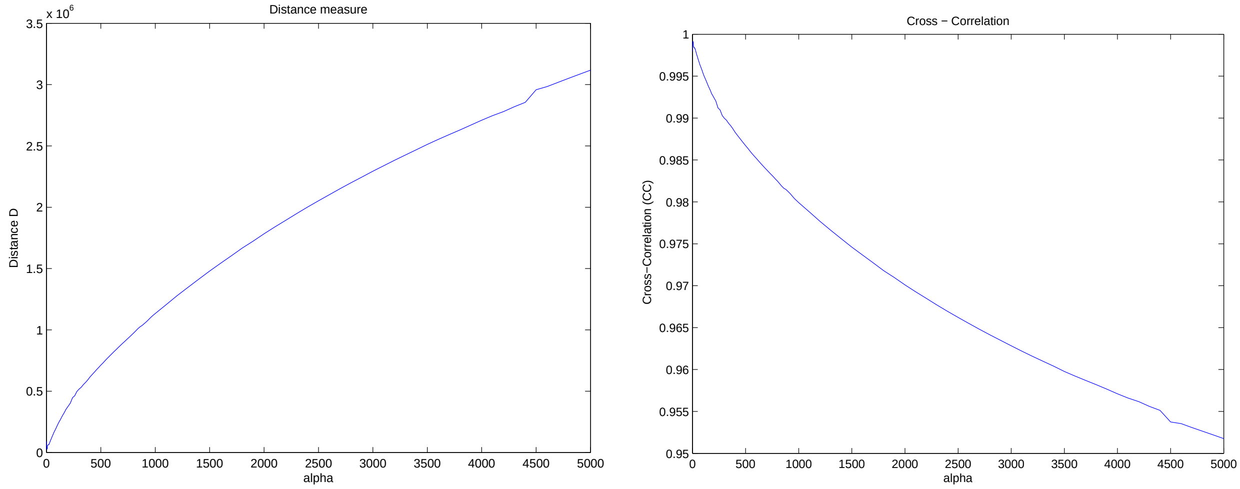
Table 3.1: Registration results. From this table we can infer, that our mass-preserving registration leads in our case to a accurate registration. The normalised cross-correlation could be increased nearly to 100% in each of the nine registration problems. This is especially remarkable for gates that show very different respiratory phases compared to the reference gate, for instance gate 1. Furthermore we like to point out that the deformation grid shows no folding, which can be deducted from the range of the functional determinant.

Sensitivity to alpha

Now the choice of parameters is examined. In our case, the functional 3.4 depends just on the regularization parameter α and on the Lamé constants μ and λ which are set to their default values, i.e. 1 and 0 respectively. In order to examine the sensitivity to alpha, we performed the registration of gate one onto gate ten above for 116 possible values of α between 0 and 5000. In order to save computation time, we will use a linear interpolation scheme.
It turns out that even for $\alpha = 0$ the transformation grid is free of foldings.



(a) L-curve (b) Regularizer

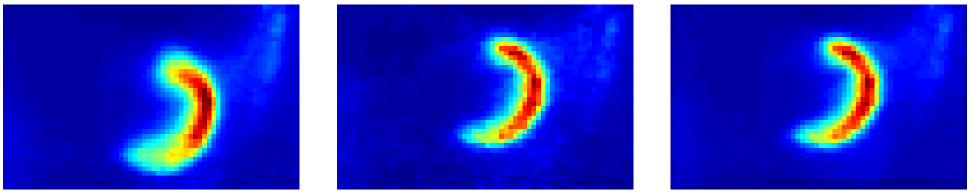


(c) SSD distance (d) Normalised Cross Correlation

Figure 3.5: Results of α-test with 116 choices for α between 0 and 5000. Plot (a) shows the final values of the distance measure and similarity measure for different choices of α on a log-scale. The remaining plots (b) - (d) demonstrate the impact of α on the regularizer $\mathcal{S}$, the distance term $\mathcal{D}$ and the normalised cross-correlation. As expected, the more the distance is reduced, the less elastic the solution becomes indicated by large values of the regularization functional. Furthermore it can be seen that the SSD and NCC behave exactly the same as it should be the case for mass-preserving transformations.

Effect on Image Quality

Finally we return to our initial motivation to correct the respiratory gated PET images for motion in order to enforce the image analogy. Hence, the non-linear registration was a necessary first step in order to estimate the respiratory motion between the gates. The inclusion of the mass-preservation proved to be crucial in order to find reasonable transformations. After applying the transformations to the images of gate one to nine and thereby mapping them to the reference gate ten we can average the samples. As can be seen in the right image in Figure 3.3 the motion corrected image is sharper than the not corrected and contains less noise than the reference gate.



(a) Reconstruction - without motion correction

(b) Reference gate

(c) Reconstruction - with motion correction

Figure 3.6: Effect on image quality. Compared to the reconstruction without motion correction (left), we could increase the sharpness of the image. Furthermore the averaging reduced the background noise compared to the reference gate.

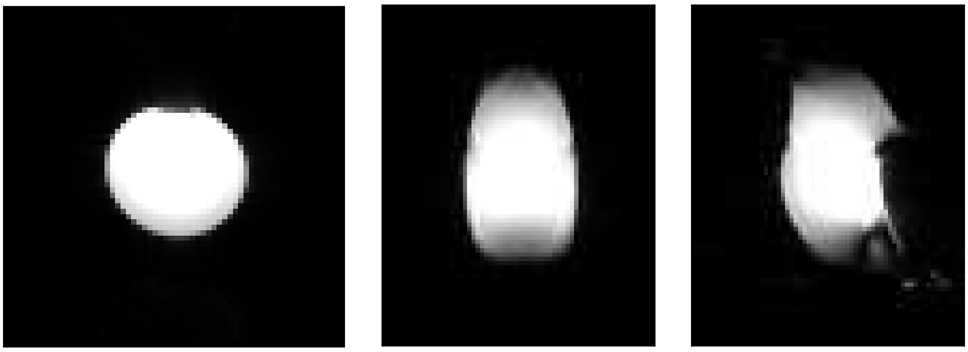
3.4 Discussion and Outlook

In order to approach the problem of motion correction in gated PET, we developed a new non-linear mass-preserving registration algorithm. The simple extension of the non linear registration functional strongly affected the existence theory and the implementation. However, we can prove the existence of a minimizer in the function space $W^{1,6}(\Omega, \mathbb{R}^3)$.
Our example indicates that - at least for gated PET images - it is important to include the mass-preservation constraint into the registration problem in order to obtain reasonable deformation fields. Mass-preserving registration leads to very smooth and folding-free deformations. Our new algorithm was even capable of accurately registering images showing very different phases of the breathing cycle.
In contrast to non mass-preserving registration, even weak regularization resulted in diffeomorphic transformations. Furthermore the transformations resembled the respiratory motion.
The accurate registration also leads to a much sharper averaged image and is therefore well-suited to increase the image quality in gated PET images.
First tests with multiple subject indicate that our algorithm can robustly improve the image quality of both cardiac and respiratory gated PET. In future studies a comparison to optical flow based correction approaches like the one proposed by Dawood et al. in [6] is asired.
In order to avoid resampling, one could also reconstruct the template gates one to nine in the deformed coordinate system as described in [16].
For PET images are more and more used in combination with structural images for instance CT or MRI, a multi-modal enhancement to the presented approach is desirable to enforce the accuracy of image fusion.

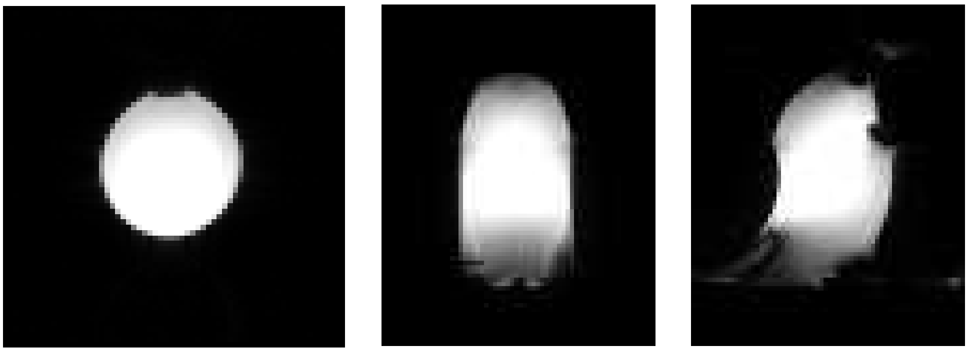
A Appendix

In this chapter more image slices of the experiments with the EPI correction algorithm described in Chapter 2 are shown. It is divided into three sections. In section A.1 the effect of different phase encoding directions on the deformations is shown in Figure A.1. In the preceding Sections and a deeper insight in the correction result of functional MR images and diffusion tensor images is provided, respectively.

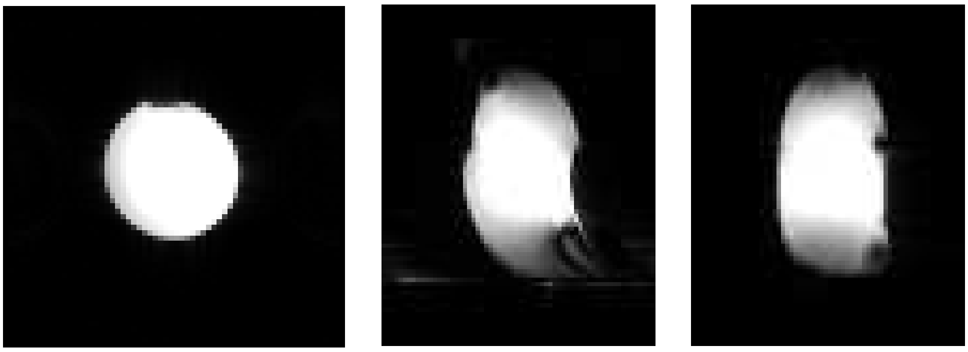
A.1 Phantom scan



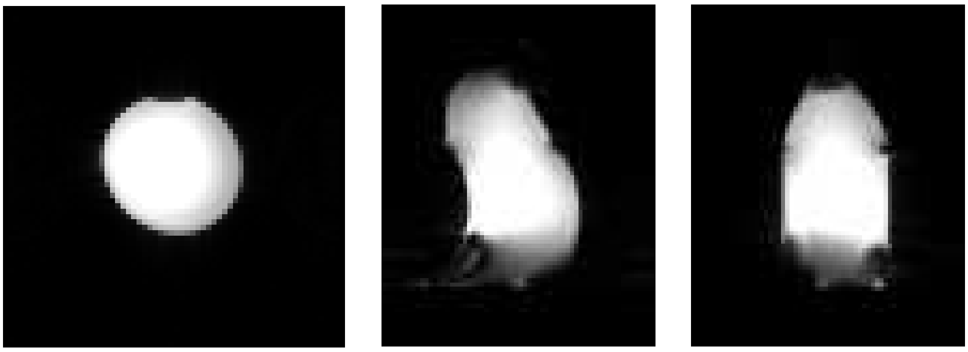
(a) app z (b) app y (c) app x



(d) apa z (e) apa y (f) apa x

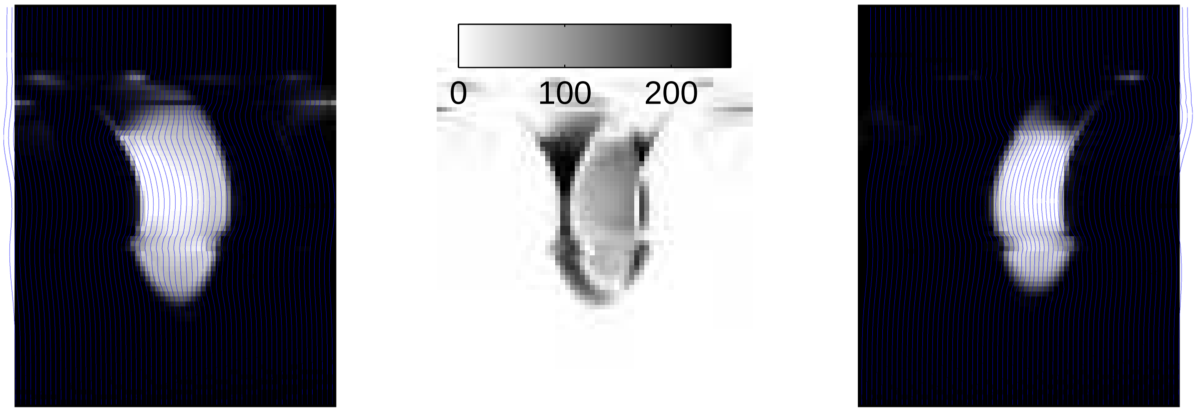


(g) rll z (h) rll y (i) rll x

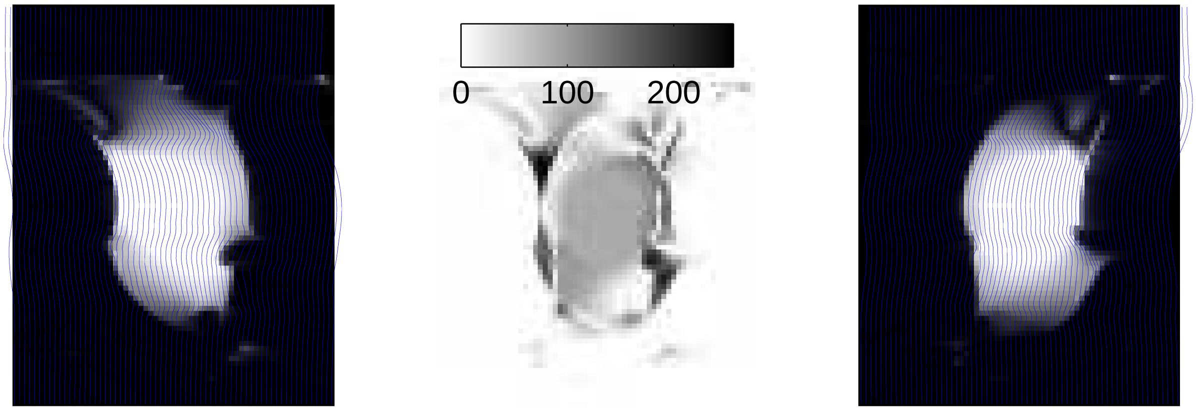


(j) rlr z (k) rlr y (l) rlr x

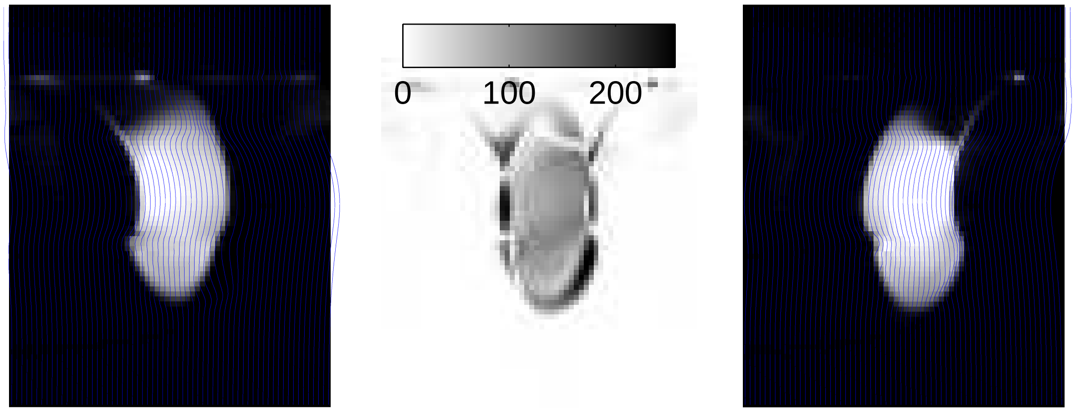
Figure A.1: Acquired echo-planar images of a water-filled bottle. The different phase encoding directions are ordered column wise where different views of the three dimensional volume are given. The first two rows are acquired with phase encoding from anterior/posterior (app) and posterior/anterior (apa), respectively. The distortions in (c) and (f) and in (a) and (d) are reversed. The latter two columns represent the results of left/right (rlr) and right/left (rll) phase encoding. Here the amount left/right shift in (h) and (f) and in (g) and (j) are reversed. Note, that the differences of (b) and (c) and of (i) and (l) might be due to the fact that we show one slice only



(a) $\mathcal{I}_1$ slice 22 initial - saggital (b) Difference slice 22 initial - saggital (c) $\mathcal{I}_2$ slice 22 initial - saggital

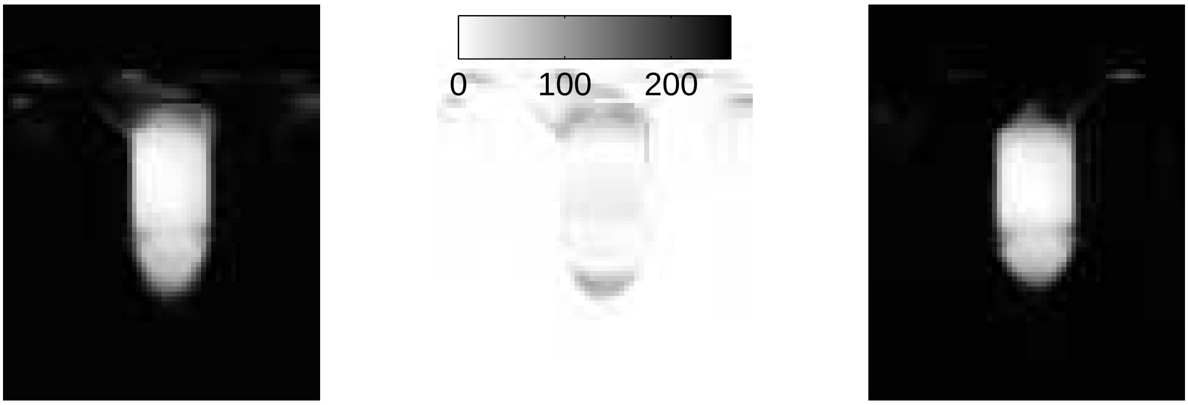


(d) $\mathcal{I}_1$ slice 32 initial - saggital (e) Difference slice 32 initial - saggital (f) $\mathcal{I}_2$ slice 32 initial - saggital

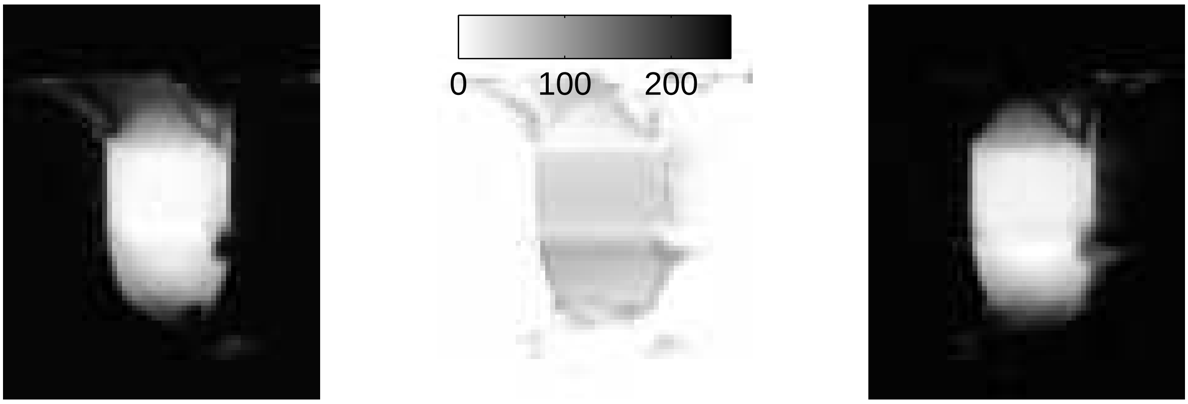


(g) $\mathcal{I}_1$ slice 42 initial - saggital (h) Difference slice 42 initial - saggital (i) $\mathcal{I}_2$ slice 42 initial - saggital

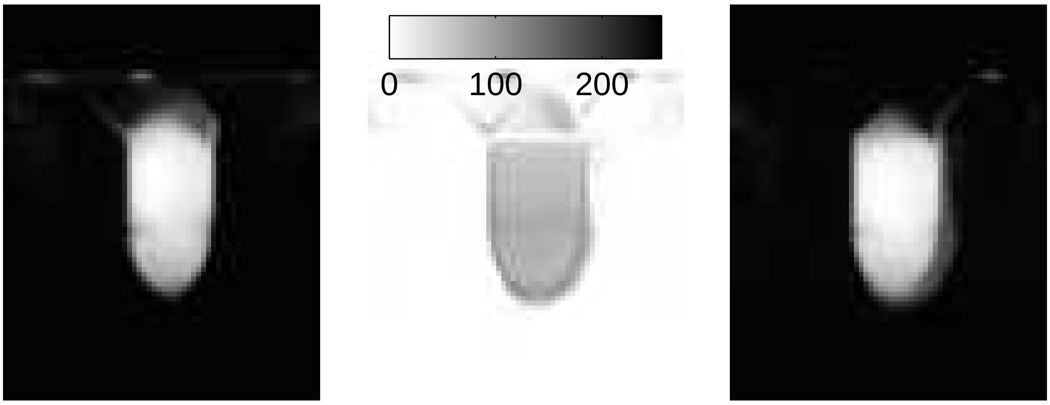
Figure A.2: Initial saggital slices. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 22 final - saggital (b) Difference slice 22 final - saggital (c) $\mathcal{I}_2$ slice 22 final - saggital

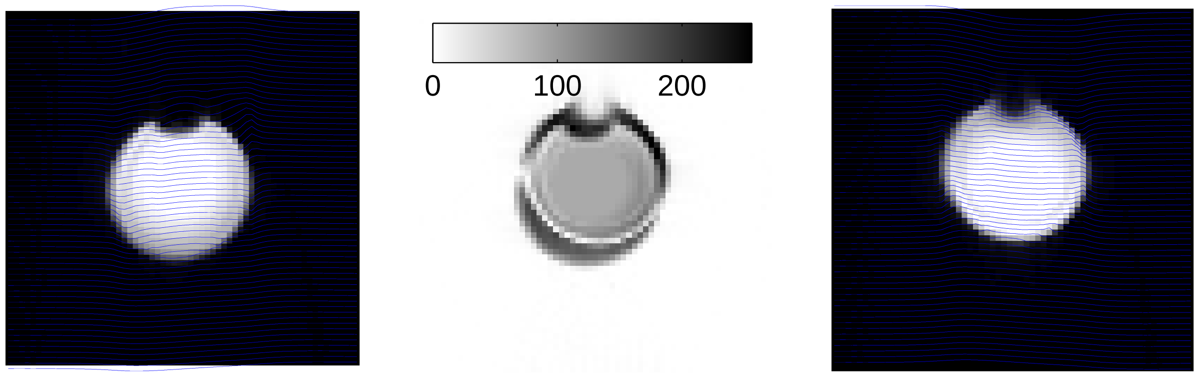


(d) $\mathcal{I}_1$ slice 32 final - saggital (e) Difference slice 32 final - saggital (f) $\mathcal{I}_2$ slice 32 final - saggital



(g) $\mathcal{I}_1$ slice 42 final - saggital (h) Difference slice 42 final - saggital (i) $\mathcal{I}_2$ slice 42 final - saggital

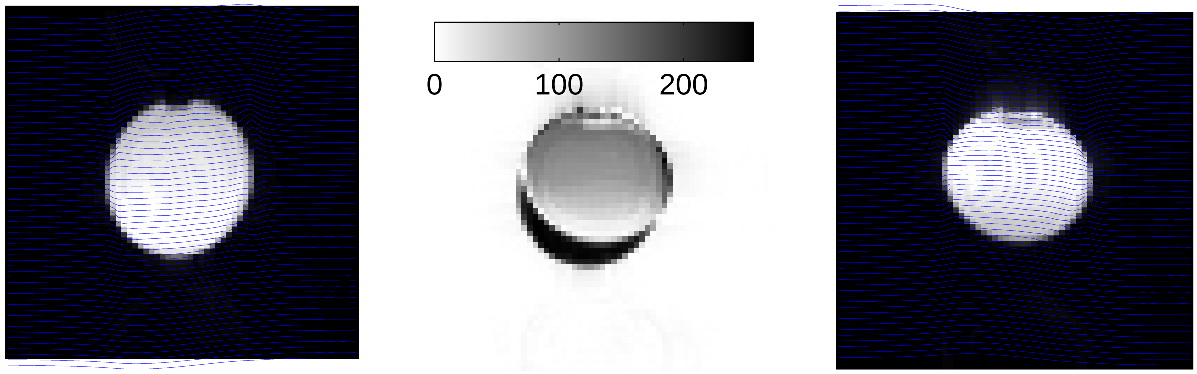
Figure A.3: Final saggital slices for $\alpha = 260$, $\mu = 1$ and $\lambda = 30$. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 32 initial - transversal

(b) Difference slice 32 initial - transversal

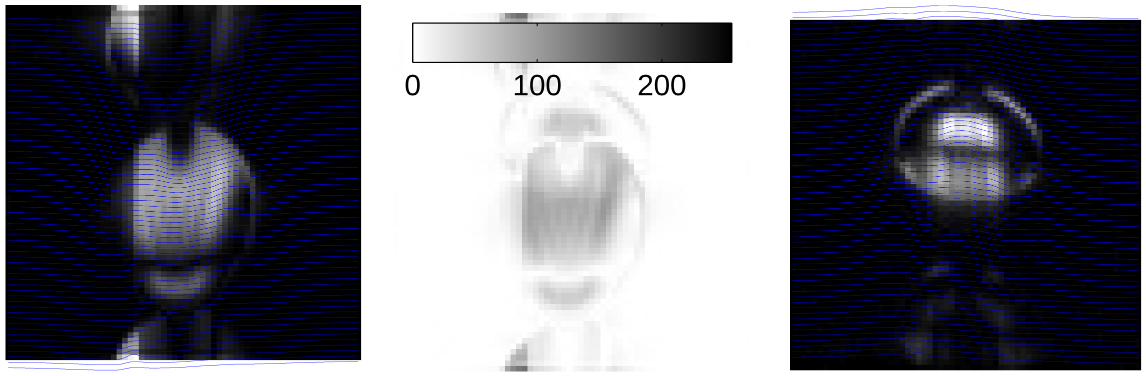
(c) $\mathcal{I}_2$ slice 32 initial - transversal



(d) $\mathcal{I}_1$ slice 50 initial - transversal

(e) Difference slice 50 initial - transversal

(f) $\mathcal{I}_2$ slice 50 initial - transversal

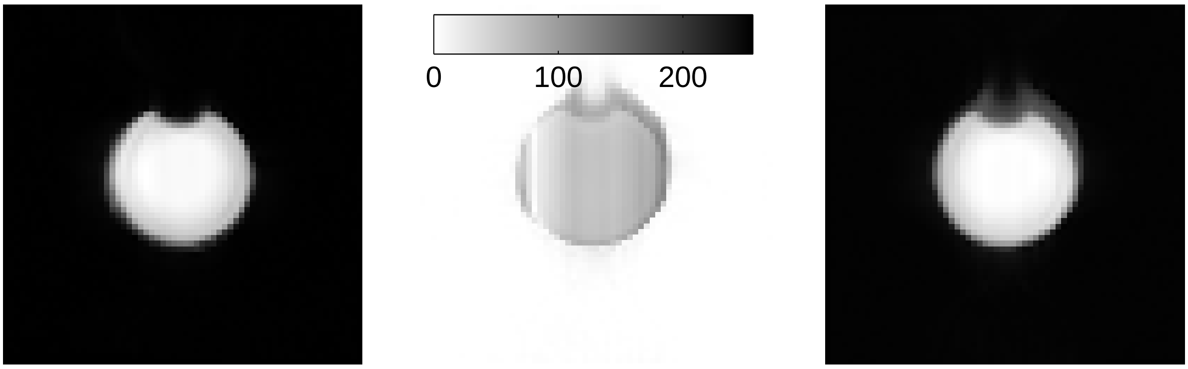


(g) $\mathcal{I}_1$ slice 61 initial - transversal

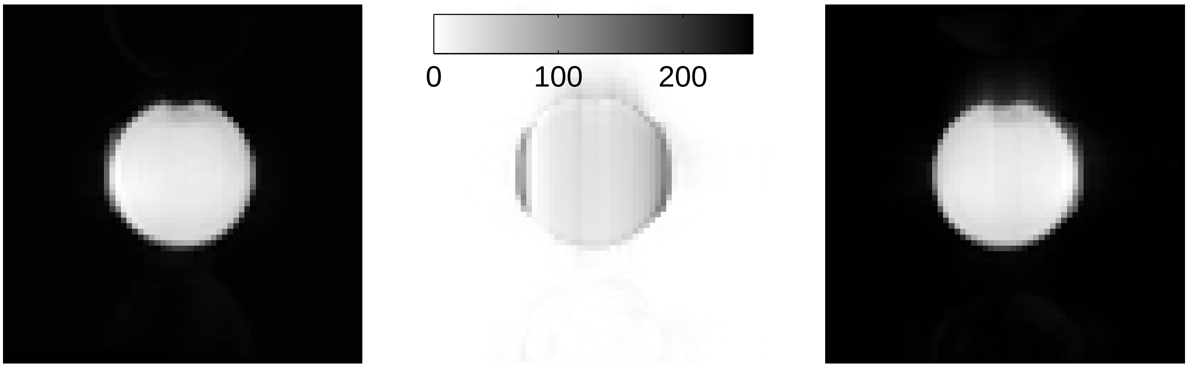
(h) Difference slice 61 initial - transversal

(i) $\mathcal{I}_2$ slice 61 initial - transversal

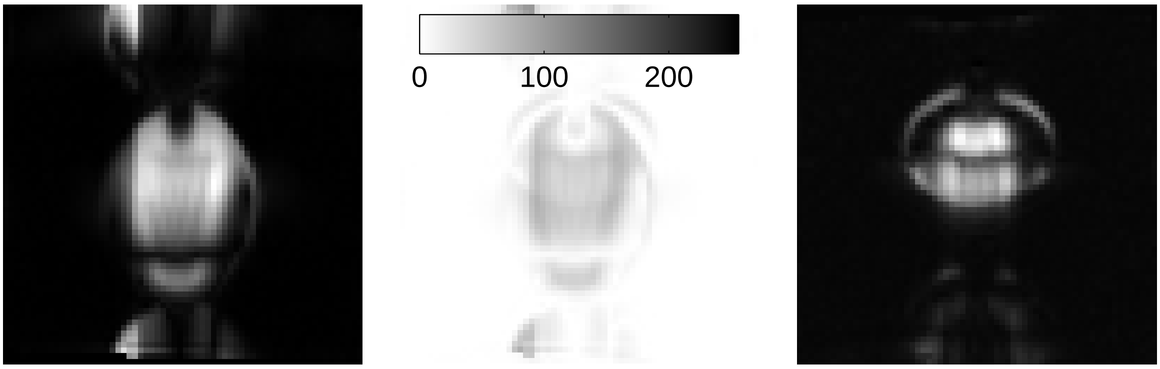
Figure A.4: Initial transversal slices. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 32 final - transversal (b) Difference slice 32 final - transversal (c) $\mathcal{I}_2$ slice 32 final - transversal



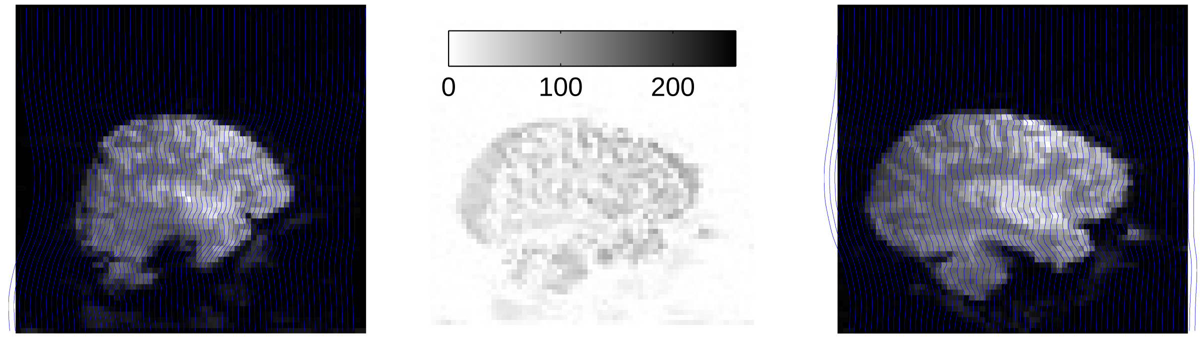
(d) $\mathcal{I}_1$ slice 50 final - transversal (e) Difference slice 50 final - transversal (f) $\mathcal{I}_2$ slice 50 final - transversal



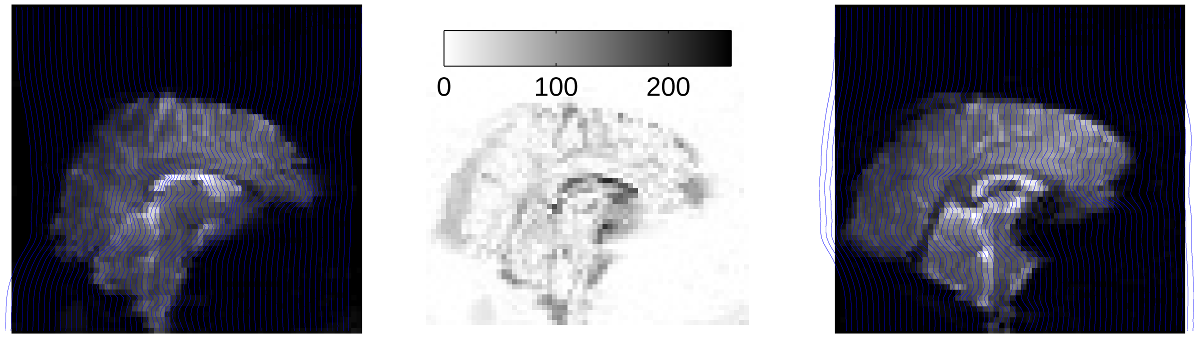
(g) $\mathcal{I}_1$ slice 61 final - transversal (h) Difference slice 61 final - transversal (i) $\mathcal{I}_2$ slice 61 final - transversal

Figure A.5: Final transversal slices for $\alpha = 260$, $\mu = 1$ and $\lambda = 30$. Left Column shows the app images, right the apa images and the middle the difference images.

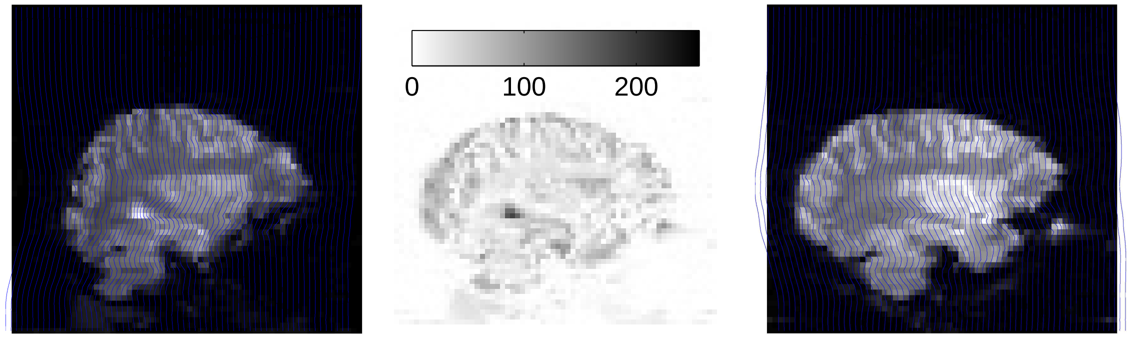
A.2 fMRI



(a) $\mathcal{I}_1$ slice 22 initial - saggital (b) Difference slice 22 initial - saggital (c) $\mathcal{I}_2$ slice 22 initial - saggital

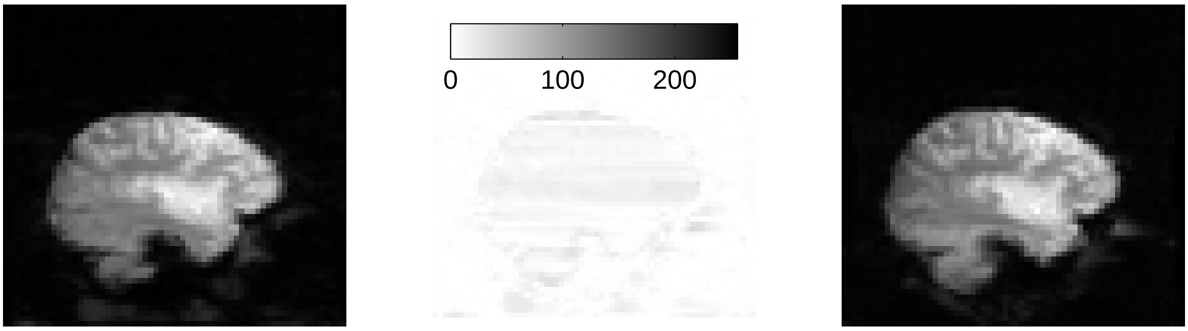


(d) $\mathcal{I}_1$ slice 32 initial - saggital (e) Difference slice 32 initial - saggital (f) $\mathcal{I}_2$ slice 32 initial - saggital

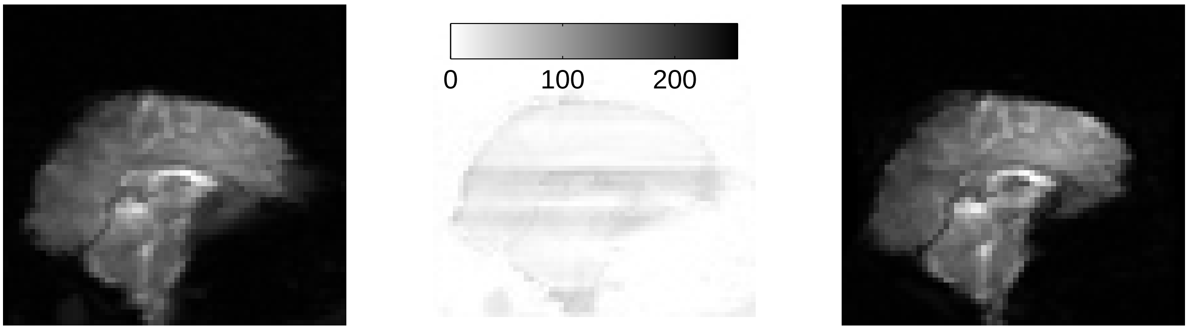


(g) $\mathcal{I}_1$ slice 42 initial - saggital (h) Difference slice 42 initial - saggital (i) $\mathcal{I}_2$ slice 42 initial - saggital

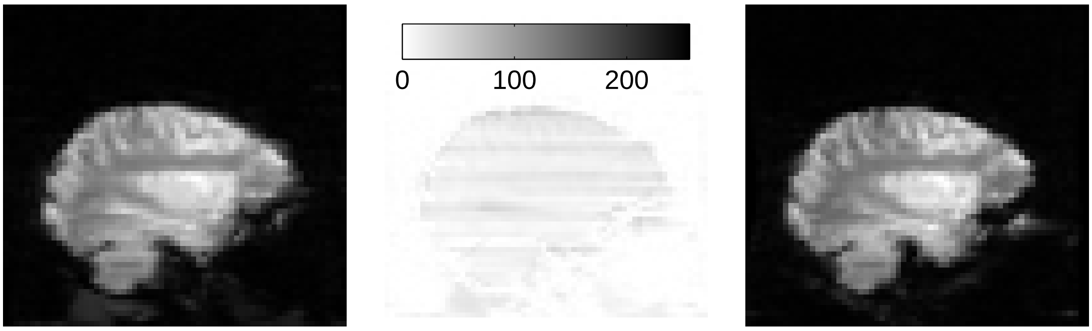
Figure A.6: Initial saggital slices. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 22 final - saggital (b) Difference slice 22 final - saggital (c) $\mathcal{I}_2$ slice 22 final - saggital

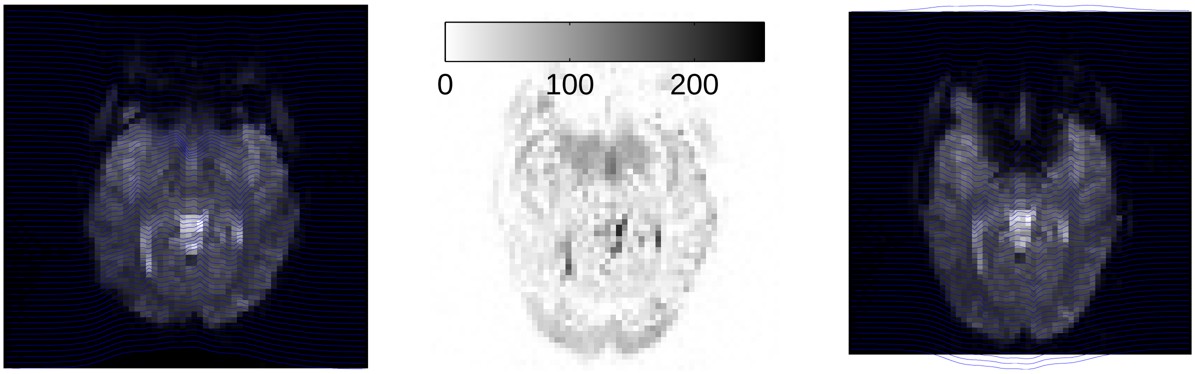


(d) $\mathcal{I}_1$ slice 32 final - saggital (e) Difference slice 32 final - saggital (f) $\mathcal{I}_2$ slice 32 final - saggital



(g) $\mathcal{I}_1$ slice 42 final - saggital (h) Difference slice 42 final - saggital (i) $\mathcal{I}_2$ slice 42 final - saggital

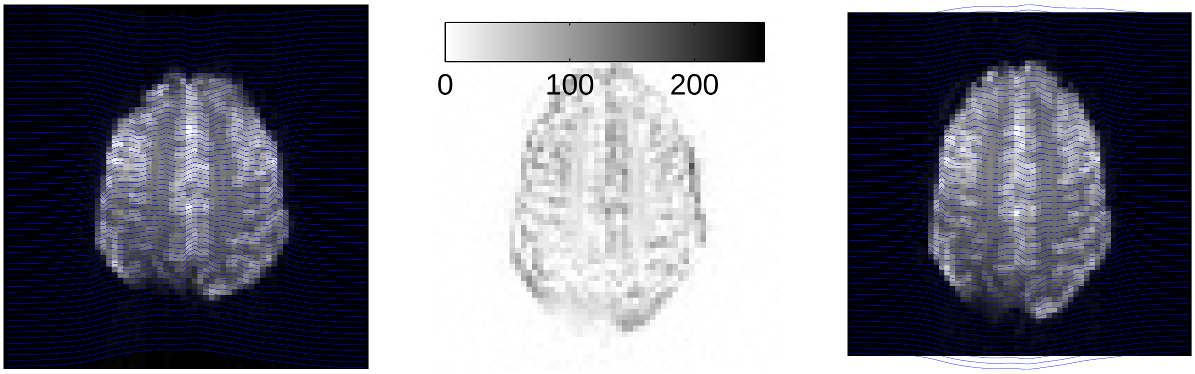
Figure A.7: Final saggital slices for $\alpha = 50$, $\mu = 1$ and $\lambda = 30$. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 22 initial - transversal

(b) Difference slice 22 initial - transversal

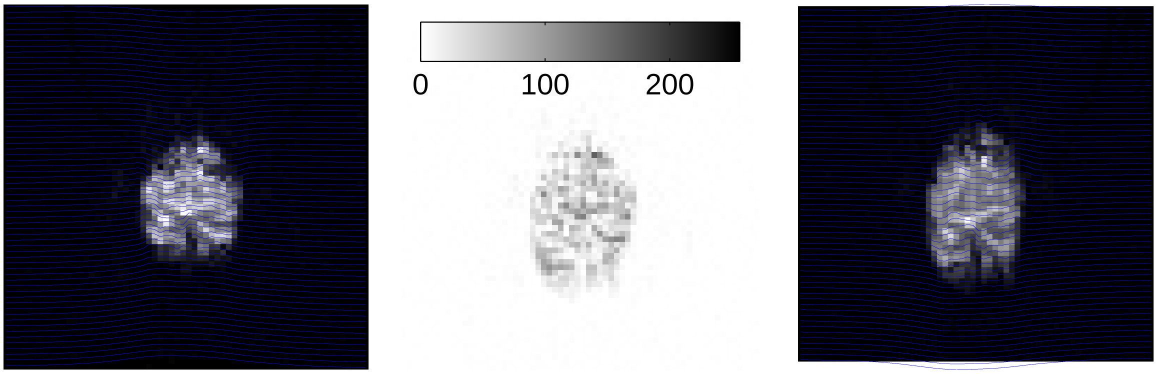
(c) $\mathcal{I}_2$ slice 22 initial - transversal



(d) $\mathcal{I}_1$ slice 35 initial - transversal

(e) Difference slice 35 initial - transversal

(f) $\mathcal{I}_2$ slice 35 initial - transversal

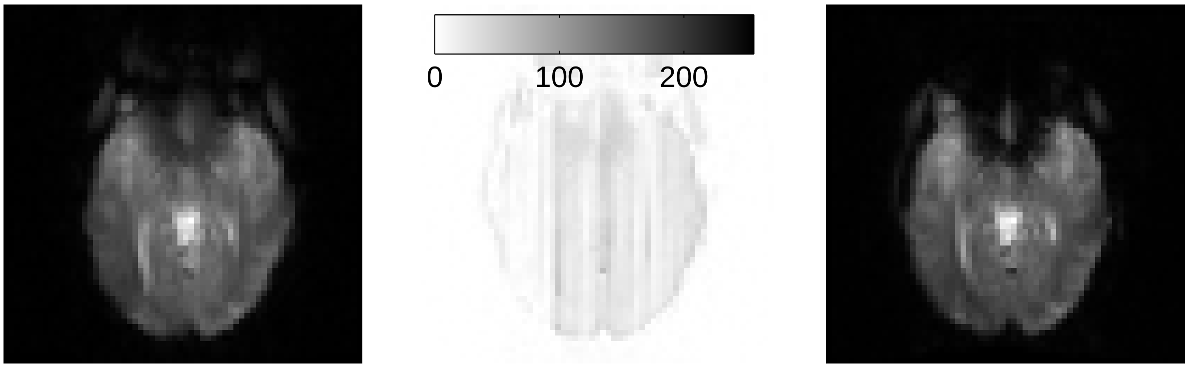


(g) $\mathcal{I}_1$ slice 42 initial - transversal

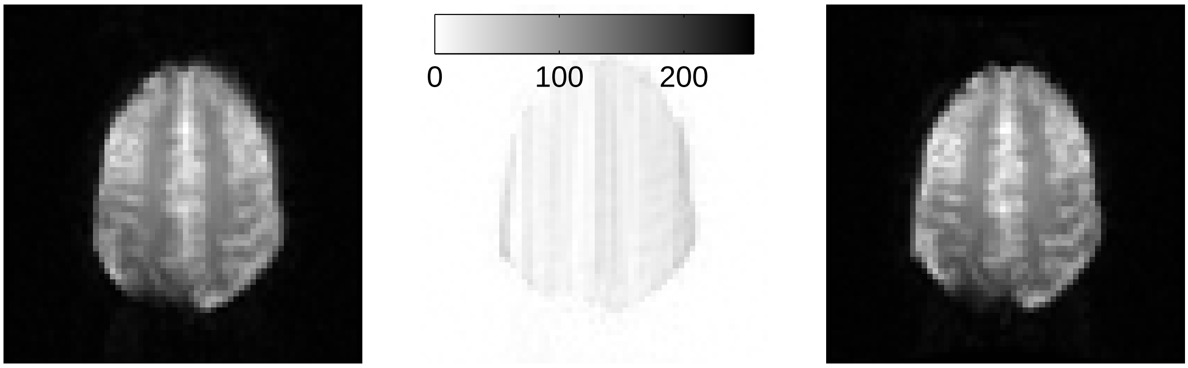
(h) Difference slice 42 initial - transversal

(i) $\mathcal{I}_2$ slice 42 initial - transversal

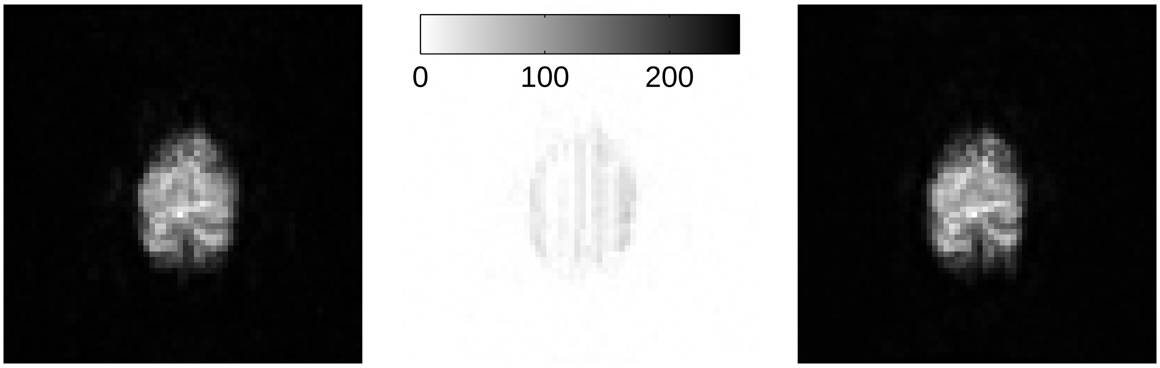
Figure A.8: Initial transversal slices. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 22 final - transversal (b) Difference slice 22 final - transversal (c) $\mathcal{I}_2$ slice 22 final - transversal



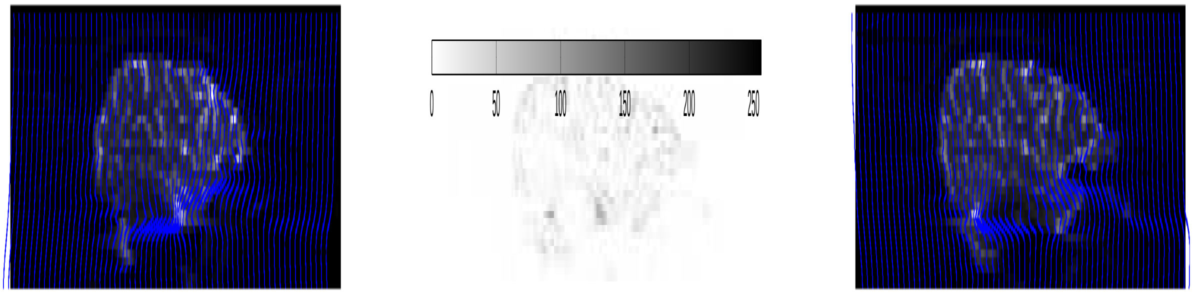
(d) $\mathcal{I}_1$ slice 35 final - transversal (e) Difference slice 35 final - transversal (f) $\mathcal{I}_2$ slice 35 final - transversal



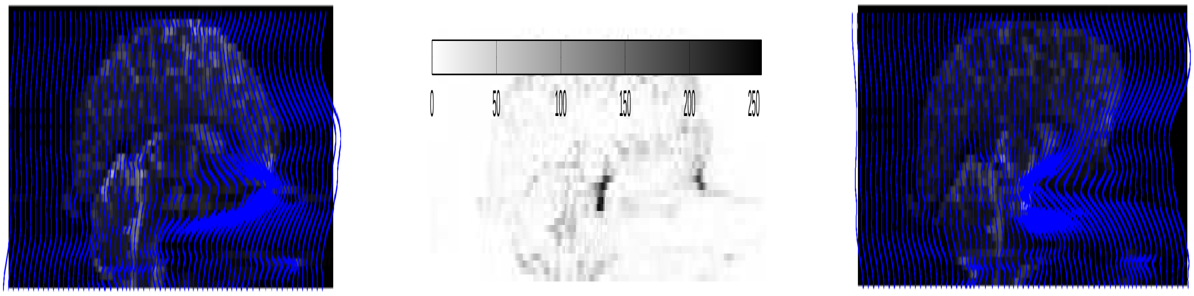
(g) $\mathcal{I}_1$ slice 42 final - transversal (h) Difference slice 42 final - transversal (i) $\mathcal{I}_2$ slice 42 final - transversal

Figure A.9: Final transversal slices $\alpha = 50$, $\mu = 1$ and $\lambda = 30$. Left Column shows the app images, right the apa images and the middle the difference images.

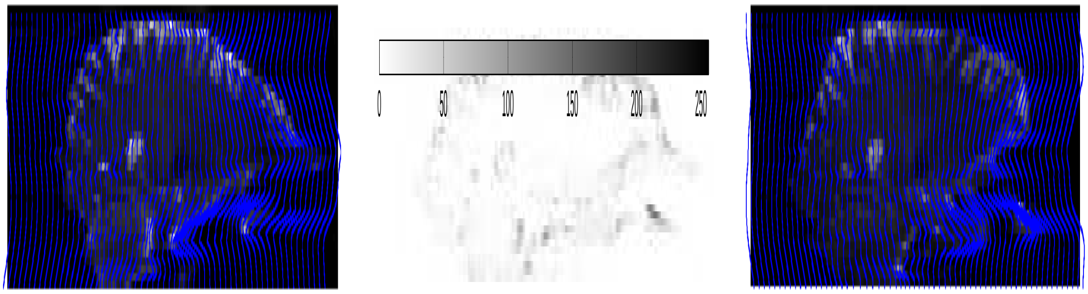
A.3 DTI



(a) $\mathcal{I}_1$ slice 22 initial - saggital (b) Difference slice 22 initial - saggital (c) $\mathcal{I}_2$ slice 22 initial - saggital

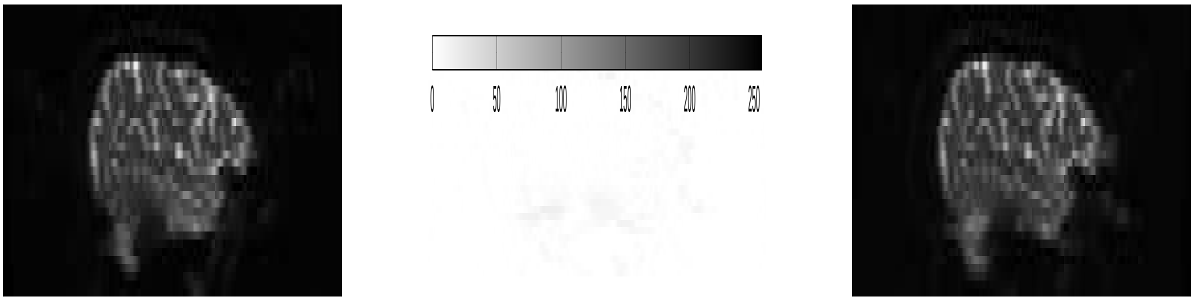


(d) $\mathcal{I}_1$ slice 32 initial - saggital (e) Difference slice 32 initial - saggital (f) $\mathcal{I}_2$ slice 32 initial - saggital



(g) $\mathcal{I}_1$ slice 42 initial - saggital (h) Difference slice 42 initial - saggital (i) $\mathcal{I}_2$ slice 42 initial - saggital

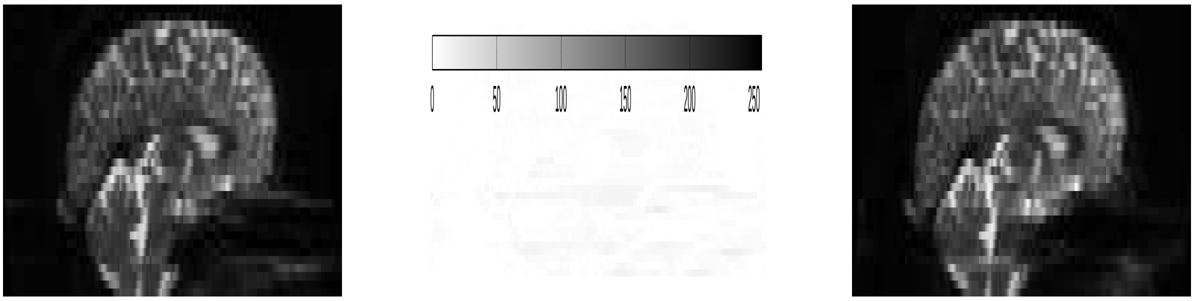
Figure A.10: Initial saggital slices. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 80 final - saggital

(b) Difference slice 80 final - saggital

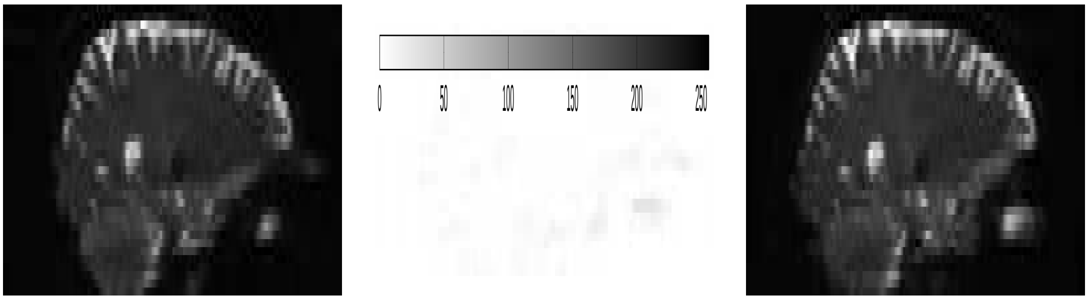
(c) $\mathcal{I}_2$ slice 80 final - saggital



(d) $\mathcal{I}_1$ slice 128 final - saggital

(e) Difference slice 128 final - saggital

(f) $\mathcal{I}_2$ slice 128 final - saggital

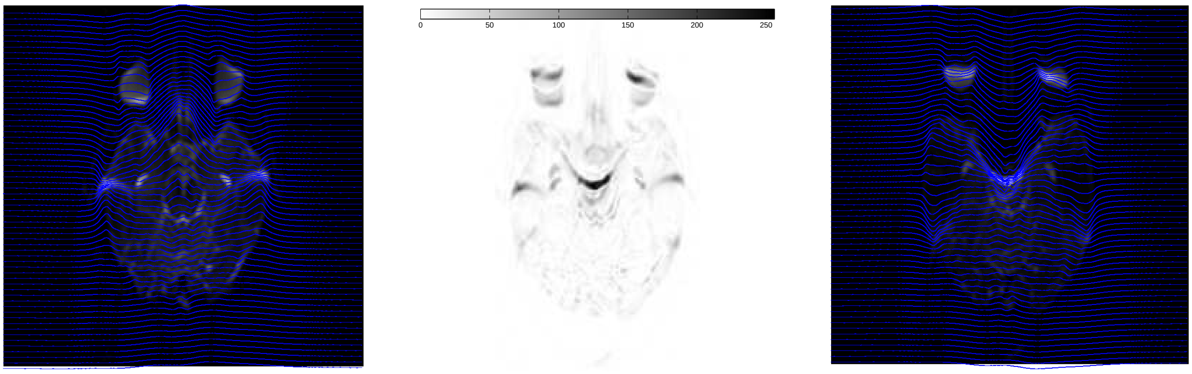


(g) $\mathcal{I}_1$ slice 150 final - saggital

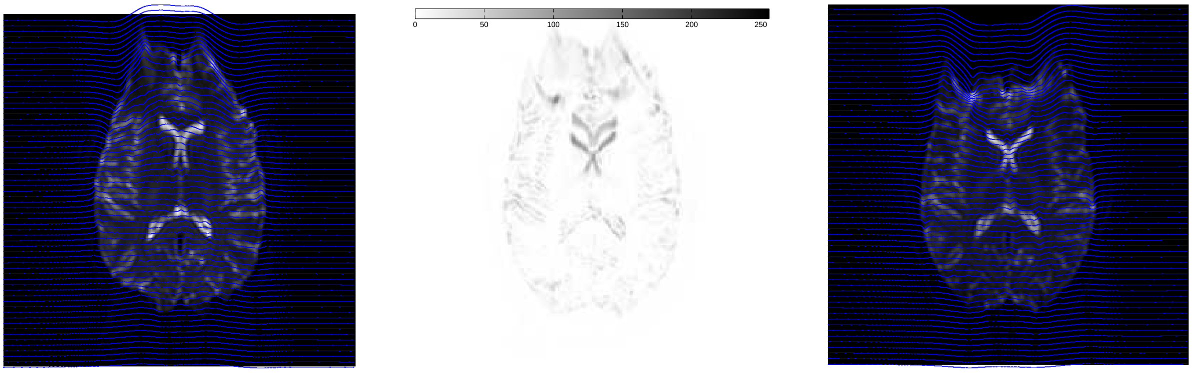
(h) Difference slice 150 final - saggital

(i) $\mathcal{I}_2$ slice 150 final - saggital

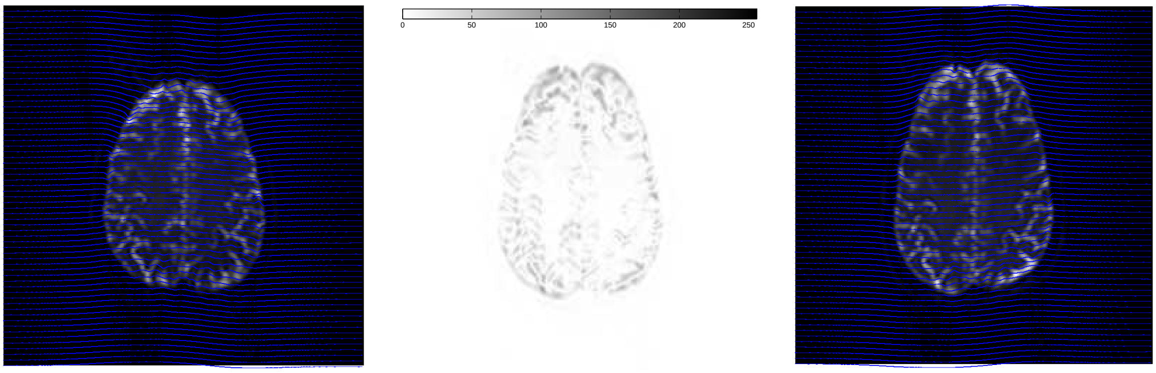
Figure A.11: Final saggital slices for $\alpha = 16$, $\mu = 1$ and $\lambda = 30$. Left Column shows the app imagess, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 11 initial - transversal (b) Difference slice 11 initial - transversal (c) $\mathcal{I}_2$ slice 11 initial - transversal

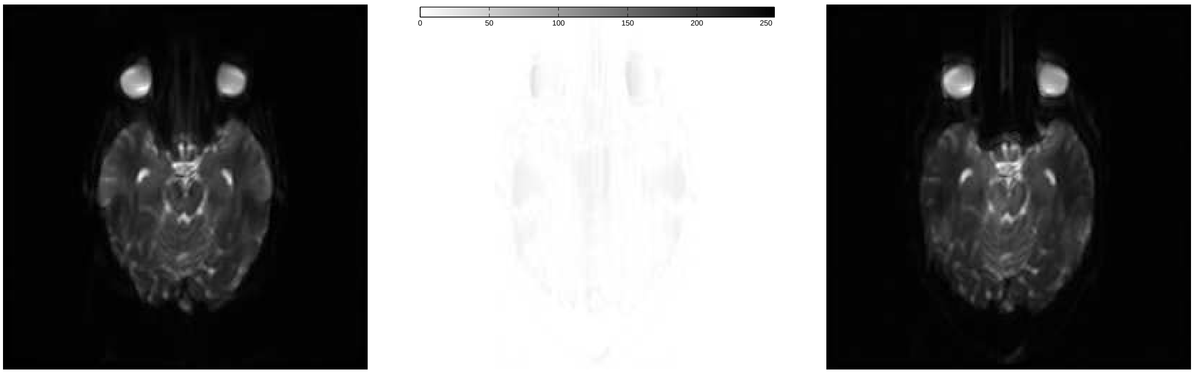


(d) $\mathcal{I}_1$ slice 19 initial - transversal (e) Difference slice 19 initial - transversal (f) $\mathcal{I}_2$ slice 19 initial - transversal

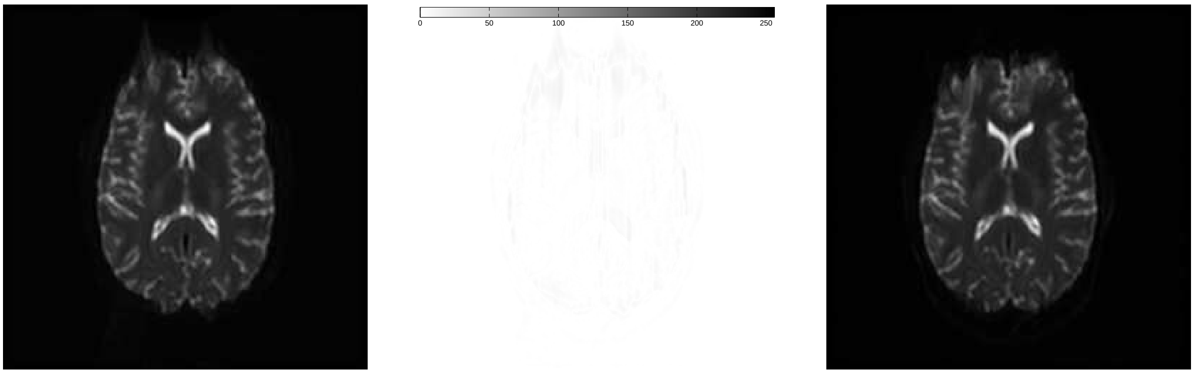


(g) $\mathcal{I}_1$ slice 27 initial - transversal (h) Difference slice 27 initial - transversal (i) $\mathcal{I}_2$ slice 27 initial - transversal

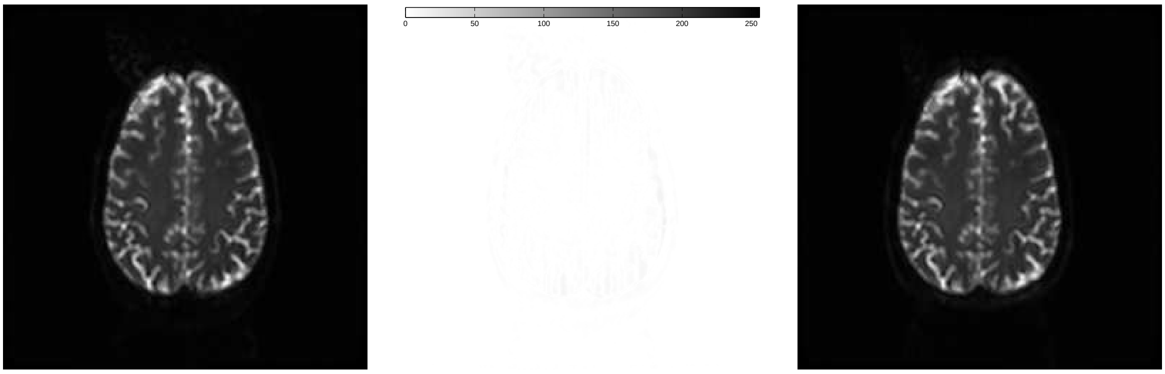
Figure A.12: Initial transversal slices. Left Column shows the app images, right the apa images and the middle the difference images.



(a) $\mathcal{I}_1$ slice 11 final - transversal (b) Difference slice 11 final - transversal (c) $\mathcal{I}_2$ slice 11 final - transversal



(d) $\mathcal{I}_1$ slice 19 final - transversal (e) Difference slice 19 final - transversal (f) $\mathcal{I}_2$ slice 19 final - transversal



(g) $\mathcal{I}_1$ slice 27 final - transversal (h) Difference slice 27 final - transversal (i) $\mathcal{I}_2$ slice 27 final - transversal

Figure A.13: Final transversal slices. Left Column shows the app images, right the apa images and the middle the difference images.

Bibliography

[1] J.L.R. Andersson, C. Hutton, J. Ashburner, R. Turner, and K. Friston. Modeling geometric deformations in EPI time series. *Neuroimage*, 13(5):903–919, 2001.

[2] J.L.R. Andersson, S. Skare, and J. Ashburner. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage*, 20(2):870–888, 2003.

[3] M. Burger. *Vorlesungsskriptum Mathematische Bildverarbeitung.* Westfäklische-Wilhelms-Universität Münster, 2007.

[4] F. Büther, M. Dawood, L. Stegger, F. Wübbeling, M. Schäfers, O. Schober, and K.P. Schäfers. List Mode-Driven Cardiac and Respiratory Gating in PET. *Journal of Nuclear Medicine*, 50(5):674, 2009.

[5] H. Chang and J.M. Fitzpatrick. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. *IEEE Transactions on Medical Imaging*, 11(3):319–329, 1992.

[6] M. Dawood, F. Büther, X. Jiang, and K.P. Schäfers. Respiratory motion correction in 3-D PET data with advanced optical flow algorithms. *IEEE Transactions on Medical Imaging*, 27(8):1164–1175, 2008.

[7] M. Deppe, C. Kellinghaus, T. Duning, G. Moddel, S. Mohammadi, K. Deppe, H. Schiffbauer, H. Kugel, SS Keller, EB Ringelstein, et al. Nerve fiber impairment of anterior thalamocortical circuitry in juvenile myoclonic epilepsy. *Neurology*, 71(24):1981, 2008.

[8] H. Devlin, I. Tracey, H. Johansen-Berg, and S. Clare. *Introduction to fMRI.* FMRIB Centre, 2010.

[9] S. Haber E. Heldmann and J. Modersitzky. A Framework for image-based constrained registration with an application to local rigidity. *Linear Algebra and its Applications*, 2007.

[10] L.C. Evans. *Partial Differential Equations.* American Mathematical Society, 1998.

[11] B. Fischer and J. Modersitzki. Ill-posed medicine—an introduction to image registration. *Inverse Problems*, 24:034008, 2008.

[12] E. Haber and J. Modersitzki. Numerical methods for volume preserving image registration. *Inverse Problems*, 20:1621–1638, 2004.

[13] E. Haber and J. Modersitzki. Image registration with guaranteed displacement regularity. *International Journal of Computer Vision*, 71(3):361–372, 2007.

[14] C. Hutton, A. Bork, O. Josephs, R. Deichmann, J. Ashburner, and R. Turner. Image distortion correction in fMRI: a quantitative evaluation. *Neuroimage*, 16(1):217–240, 2002.

[15] P. Jezzard and R.S. Balaban. Correction for geometric distortion in echo planar images from B0 field variations. *Magnetic Resonance in Medicine*, 34(1):65–73, 1995.

[16] F. Lamare, MJ Ledesma Carbayo, T. Cresson, G. Kontaxakis, A. Santos, C.C. Le Rest, AJ Reader, and D. Visvikis. List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Physics in medicine and biology*, 52(17):5187–5204, 2007.

[17] Z.P. Liang and P.C. Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Approach.* IEEE Press New York, 2000.

[18] J. Modersitzki. *Numerical methods for image registration.* Oxford University Press, USA, 2004.

[19] J. Modersitzki. *FAIR: flexible algorithms for image registration.* SIAM, 2009.

[20] P.S. Morgan, R.W. Bowtell, D.J.O. McIntyre, and B.S. Worthington. Correction of spatial distortion in EPI due to inhomogeneous static magnetic fields using the reversed gradient method. *Journal of Magnetic Resonance Imaging*, 19(4):499–507, 2004.

[21] J. Nocedal and S.J. Wright. *Numerical optimization.* Springer, 2000.

[22] J. Olesch, L. Ruthotto, H. Kugel, Stefan Skare, B. Fischer, and C. H. Wolters. A variational approach for the correction of field-inhomogeneities in epi sequences. In *Proc. of the SPIE Medical Imaging Conference, San Diego, USA*, 2010.

[23] N. Papenberg. *Ein genereller Registrierungsansatz mit Anwendung in der navigierten Leberchirurgie.* PhD thesis, Universität zu Lübeck, 2008.

[24] C. Modersitzky J. Pöschl and O. Scherzer. A variational setting for volume constrained image registration. *Industrial Geometry*, 2009.

[25] T. Rehman, E. Haber, G. Pryor, J. Melonakos, and A. Tannenbaum. 3D nonrigid registration via optimal mass transport on the GPU. *Medical Image Analysis*, 2008.

[26] S. Skare and J.L.R. Andersson. Correction of MR image distortions induced by metallic objects using a 3D cubic B-spline basis set: Application to stereotactic surgical planning. *Magnetic Resonance in Medicine*, 54(1):169–181, 2005.

[27] Stephen M. Smith, Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E.J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, Marilena De Luca, Ivana Drobnjak, David E. Flitney, Rami K. Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J. Michael Brady, and Paul M. Matthews. Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23(Supplement 1):S208–S219, 2004.

[28] R. Tao, P.T. Fletcher, S. Gerber, and R.T. Whitaker. A Variational image-based approach to the correction of susceptibility artefacts in the alignment of diffusion weighted and structural MRI. In *Proceedings of the 21st International Conference on Information Processing in Medical Imaging*, page 675. Springer, 2009.

[29] D. Weishaupt, V.D. Köchli, and B. Marincek. *Wie funktioniert MRI?: Eine Einführung in Physik und Funktionsweise der Magnetresonanzbildgebung.* Springer, 2009.

[30] N. Weiskopf, U. Klose, N. Birbaumer, and K. Mathiak. Single-shot compensation of image distortions and BOLD contrast optimization using multi-echo EPI for real-time fMRI. *Neuroimage*, 24(4):1068–1079, 2005.

[31] M.N. Wernick and J.N. Aarsvold. *Emission Tomography - The Fundamentals of PET and SPECT.* Elsevier, 2004.

[32] C.H. Wolters. *Influence of Tissue Conductivity Inhomogeneity and Anisotropy on EEG/MEG based Source Localization in the Human Brain.* PhD thesis, University of Leipzig, 2003.

[33] M. Wu, L.C. Chang, L. Walker, H. Lemaitre, A.S. Barnett, S. Marenco, and C. Pierpaoli. Comparison of EPI distortion correction methods in diffusion tensor MRI using a novel framework. In *Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II*, page 329. Springer, 2008.

[34] L. Zhu, S. Haker, and A. Tannenbaum. Mass preserving registration for heart MR images. *Lecture Notes in Computer Science*, 3750:147, 2005.

Erklärung der Eigenständigkeit

Hiermit versichere ich, Lars Ruthotto, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe.
Gedanklich, inhaltlich oder wörtlich Übernommenes habe ich durch Angabe von Herkunft und Text oder Anmerkung belegt bzw. kenntlich gemacht. Dies gilt in gleicher Weise für Bilder, Tabellen und Skizzen, die nicht von mir selbst erstellt wurden.

Münster, 05.03.2010