



# Autoadaptive motion modelling for MR-based respiratory motion estimation



Christian F. Baumgartner<sup>a,\*</sup>, Christoph Kolbitsch<sup>a</sup>, Jamie R. McClelland<sup>b</sup>, Daniel Rueckert<sup>c</sup>, Andrew P. King<sup>a</sup>

<sup>a</sup> Division of Imaging Sciences and Biomedical Engineering, King's College London, London, UK

<sup>b</sup> Centre for Medical Image Computing, University College London, London, UK

<sup>c</sup> Biomedical Image Analysis Group, Department of Computing, Imperial College London, London, UK

## ARTICLE INFO

### Article history:

Received 1 June 2015

Revised 22 April 2016

Accepted 7 June 2016

Available online 9 June 2016

### Keywords:

MR-guided interventions

Respiratory motion correction

Motion modelling

Manifold learning

Manifold alignment

## ABSTRACT

Respiratory motion poses significant challenges in image-guided interventions. In emerging treatments such as MR-guided HIFU or MR-guided radiotherapy, it may cause significant misalignments between interventional road maps obtained pre-procedure and the anatomy during the treatment, and may affect intra-procedural imaging such as MR-thermometry. Patient specific respiratory motion models provide a solution to this problem. They establish a correspondence between the patient motion and simpler surrogate data which can be acquired easily during the treatment. Patient motion can then be estimated during the treatment by acquiring only the simpler surrogate data.

In the majority of classical motion modelling approaches once the correspondence between the surrogate data and the patient motion is established it cannot be changed unless the model is recalibrated. However, breathing patterns are known to significantly change in the time frame of MR-guided interventions. Thus, the classical motion modelling approach may yield inaccurate motion estimations when the relation between the motion and the surrogate data changes over the duration of the treatment and frequent recalibration may not be feasible.

We propose a novel methodology for motion modelling which has the ability to automatically adapt to new breathing patterns. This is achieved by choosing the surrogate data in such a way that it can be used to estimate the current motion in 3D as well as to update the motion model. In particular, in this work, we use 2D MR slices from different slice positions to build as well as to apply the motion model. We implemented such an autoadaptive motion model by extending our previous work on manifold alignment.

We demonstrate a proof-of-principle of the proposed technique on cardiac gated data of the thorax and evaluate its adaptive behaviour on realistic synthetic data containing two breathing types generated from 6 volunteers, and real data from 4 volunteers. On synthetic data the autoadaptive motion model yielded 21.45% more accurate motion estimations compared to a non-adaptive motion model 10 min after a change in breathing pattern. On real data we demonstrated the method's ability to maintain motion estimation accuracy despite a drift in the respiratory baseline. Due to the cardiac gating of the imaging data, the method is currently limited to one update per heart beat and the calibration requires approximately 12 min of scanning. Furthermore, the method has a prediction latency of 800 ms. These limitations may be overcome in future work by altering the acquisition protocol.

© 2016 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Recent advances in magnetic resonance (MR) compatible materials and the development of fast parallel computational tech-

niques now allow an increasing range of interventions to be guided by MR images in real-time. MR-guided high intensity focused ultrasound (MRg-HIFU) has been successfully applied to treat a range of conditions such as uterine fibroids, and prostate and liver cancers in patients where invasive therapy is not possible (Tempny et al., 2011; Foley et al., 2013). In MRg-HIFU, targets are identified in MR images and a computer controlled transducer is moved

\* Corresponding author.

E-mail address: [christian.baumgartner@kcl.ac.uk](mailto:christian.baumgartner@kcl.ac.uk) (C.F. Baumgartner).

to sequentially ablate them. Furthermore, MR thermometry can be used to monitor temperature elevation of the tissue and MR imaging can be used after the ablation to evaluate the success of the treatment (Hynynen et al., 1996). Similarly, the recent development of integrated MR linear accelerators shows great potential for accurate guidance of radiotherapy (RT) treatments (Raaymakers et al., 2009) using MR imaging. In MRg-RT, magnetic resonance imaging is not only used to accurately identify and track the target but also to prevent the irradiation of healthy tissue in organs at risk (Crijns et al., 2012).

For treatments targeting organs affected by breathing motion such as the lungs, the liver, the kidneys or the heart, accurate knowledge of the respiratory motion is essential. Apart from ensuring the irradiation or ablation of the intended target and sparing of the organs at risk, knowledge of respiratory motion is also crucial to correct for motion-induced image-artefacts and for adjusting accumulated dose calculations such as temperature maps in MRg-HIFU or dose simulations in MRg-RT (De Senneville et al., 2015). Respiratory motion correction is complicated by the fact that breathing, though approximately periodic, may exhibit large variations within a single breathing cycle (intra-cycle variation or hysteresis) or across breathing cycles (inter-cycle variation) (Blackall et al., 2006; Kini et al., 2003; McClelland et al., 2011). For example, abdominal organs are known to undergo continuous drifts for long treatment durations (von Siebenthal et al., 2007; Arnold et al., 2011) and the motion observed in the lungs may significantly change over the duration of a treatment session due to changing breathing patterns (Blackall et al., 2006; King et al., 2009; Kini et al., 2003).

Respiratory motion correction can be roughly divided into two classes of techniques: Tracking methods, which attempt to track a single (or a few) targets, such as tumours or fibroids, and motion estimation techniques which aim to provide dense motion fields for an entire region of interest. Both can be achieved in real-time by either *directly* measuring the motion in the images, or *indirectly*, using respiratory motion modelling. In this paper, we provide the proof-of-principle for a novel autoadaptive motion modelling framework for MR-guided interventions which can provide dense three-dimensional motion estimates in the entire region of interest while automatically adapting to changes in respiratory patterns such as drift.

The foremost objective in image guided treatments is to keep the radiation beam or transducer aligned with the moving target. Methods to track targets in real-time using MR imaging information have been proposed by a number of authors for MRg-HIFU (Ries et al., 2010; Zachiu et al., 2015) and MRg-RT (Crijns et al., 2012; Brix et al., 2014). Some tracking solutions not requiring any imaging data have been proposed as well (Sawant et al., 2009). While tracking has been used for estimating the motion of single targets, many interventions may benefit from richer motion information. In radiotherapy it is desirable to model the motion of the organs at risk as well as the target in order to avoid their irradiation (Crijns et al., 2012). In MRg-HIFU, MR thermometry is used to monitor the temperature of the sonicated tissue in order to detect when a lethal thermal dose has been delivered (Zachiu et al., 2015). Respiratory motion may distort the temperature measurements and ideally the entire organ's motion should be densely estimated in three dimensions to correct for this motion (Zachiu et al., 2015; Rijkhorst et al., 2011; Ries et al., 2010; Arnold et al., 2011). MR is a good candidate to provide dense motion estimates during such treatments and has been used to this end. For example, (Ries et al., 2010) demonstrated a combination of 2D MR in-plane imaging and 1D through-plane tracking using an MR pencil beam navigator. Zachiu et al. (2015) proposed 2D MR imaging with intermittent adjustment using 3D MR imaging. De Senneville et al. (2015) proposed a general real-time 2D motion estimation tech-

nique for MRg-HIFU and MRg-RT. Current MR technology does not allow imaging in 3D directly with sufficient temporal and spatial resolution. It has been shown, however, that high-quality 3D motion estimations of the whole thorax can be obtained making use of sequentially acquired 2D MR data from different imaging planes (Würslin et al., 2013; Dikaos et al., 2012; von Siebenthal et al., 2007; Baumgartner et al., 2014a).

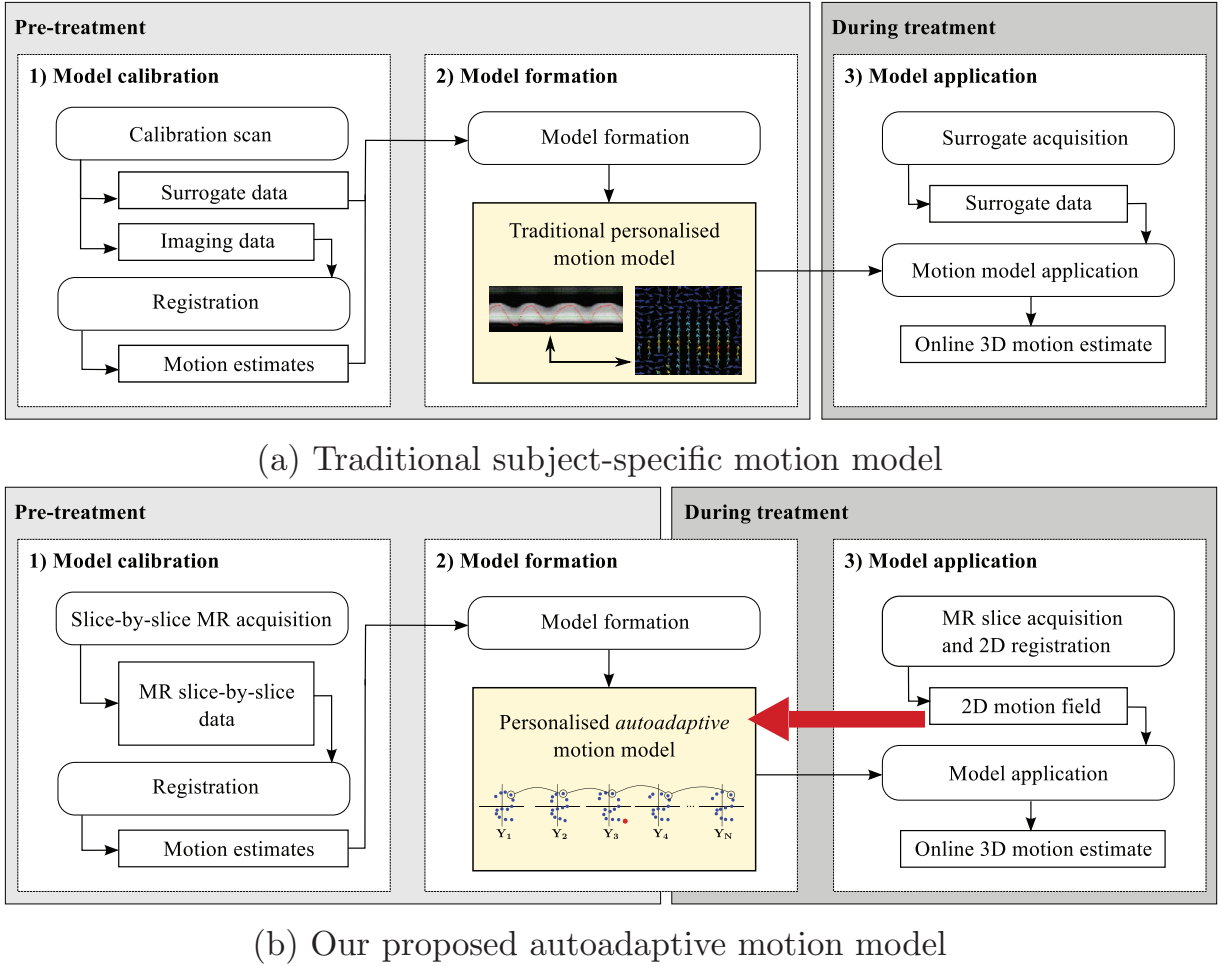
### 1.1. Motion modelling

Motion models offer a solution for indirect estimation of respiratory motion. A motion model is built by relating the patient-specific breathing motion to a simpler respiratory surrogate signal before the treatment. During the treatment, motion estimates can be obtained using only the surrogate signal. Patient-specific motion models have been proposed extensively for motion correction in radiotherapy (Seppenwoolde et al., 2002; Schweikard et al., 2000; 2005; Hoogeman et al., 2009; Cho et al., 2010; Isaksson et al., 2005) and to a lesser degree also in MR-guided HIFU (Rijkhorst et al., 2010; 2011). A comprehensive review of respiratory motion modelling can be found in McClelland et al. (2013).

Traditionally, motion models consist of three distinct stages as illustrated in Fig. 1a: before the treatment, in a *model calibration* step, typically imaging data are acquired along with some simpler surrogate data and image registration techniques are used to extract motion estimates from the imaging data. Surrogate data are often one-dimensional signals derived, for example, by tracking infra-red markers on the patient's chest (Schweikard et al., 2000) or from a spirometer (Low et al., 2005). However, due to increasing requirements for precise motion estimation, recently there has been a trend in motion modelling towards the use of more complex surrogate signals, which offer the chance to capture more respiratory motion variabilities. Examples include chest surface data (Fassi et al., 2014), real-time ultrasound images (Peressutti et al., 2013; 2012), or real-time 2D MR slices (King et al., 2012). Next, in the *model formation* stage, the motion estimates are related to the surrogate data using an appropriate correspondence model. Then, during the treatment (i.e. in the *application phase*), only the surrogate data are continually acquired and motion estimates are derived by using them as input to the correspondence model.

An underlying assumption of the majority of traditional motion models is that the nature of the relationship between the surrogate data and the motion (i.e. the correspondence model) remains constant. For long treatment durations it is possible for the breathing motion to undergo significant changes, for example due to varying degrees of relaxation of the patient during the procedure, because of pain or discomfort experienced (King et al., 2012; Hoogeman et al., 2009) or because of organ drift (Arnold et al., 2011; von Siebenthal et al., 2007). In the traditional motion model paradigm the model is formed before the treatment and has no ability to adapt to changing breathing patterns.

In response to this problem, a number of papers have proposed *adaptive* motion modelling techniques. A common approach is to correct for changing breathing patterns through occasional intra-fractional imaging. For example, in stereotactic x-ray guided radiotherapy systems, the 3D target location can be intermittently obtained every 1–6 min using intra-fractional imaging along with a surrogate signal value (Seppenwoolde et al., 2007; Hoogeman et al., 2009). This data can then be used to recalibrate the model on a first-in-first-out basis (Schweikard et al., 2000; Cho et al., 2010). Isaksson et al. (2005) employed an adaptive motion model based on neural networks in which the model didn't require recalibration. Instead, the weights of the neural network were adjusted based on frequent intra-fractional x-ray imaging. Some radiotherapy systems use the distance between the actual tumour position and the position predicted by the system as a quality measure.



**Fig. 1.** Schematic representation of (a) the traditional subject-specific motion model paradigm and (b) our autoadaptive subject-specific motion model allowing for continuous adaptivity to changing breathing patterns. The red arrow indicates our proposed change to the motion model paradigm allowing for new surrogate/calibration data to be incorporated into the motion model without interrupting the application phase. In this new paradigm the model is initially formed pre-treatment, but is updated continually during the treatment.

When this distance goes above a given threshold, the model application phase is interrupted and the current motion model is discarded, new calibration and surrogate data are acquired, and a new motion model is built (Hoogeman et al., 2009). A similar approach was used by King et al. (2012) for motion correction in a simultaneous PET/MR system. An adaptive motion modelling approach not requiring intra-fractional imaging was proposed by Fassi et al. (2014), who accounted for respiratory baseline drift in x-ray guided radiotherapy by registering a chest surface mesh obtained during the treatment to the chest surface extracted from the 4D CT planning scan. To the best of our knowledge, no adaptive motion modelling techniques have been proposed for MR-guided treatments.

In this paper, we propose a novel *autoadaptive* motion model for MR-guided interventions, which is automatically updated each time surrogate data is acquired. This is achieved by altering the traditional motion model paradigm as shown in Fig. 1b. In our proposed framework, both the calibration as well as the surrogate data are 2D MR slices acquired from variable imaging planes. Since the model calibration and surrogate data are of the same type, the surrogate data acquired in the application phase can be fed back into the model formation phase as the treatment goes on, allowing a continuous updating of the model. In Fig. 1b this update process is indicated by the red feedback arrow. This allows the model to maintain motion estimation accuracy despite gradual changes in

the breathing motion and to adapt to previously unseen breathing patterns. Such a framework has potential application in all MR-guided interventions, in particular in MRg-RT and MRg-HIFU. Note that motion estimates derived from 2D MR data have the potential to more accurately reflect (in-plane) motion than 3D MR data due to their superior image quality (Würslin et al., 2013; Dikaos et al., 2012; Baumgartner et al., 2014a). In the proposed framework, the function of the motion model is to relate 2D MR motion surrogates to dense 3D motion estimates.

We demonstrate how such a motion model can be implemented based on the concepts of manifold learning (ML) and manifold alignment (MA).

## 1.2. Manifold learning and manifold alignment

Time series of medical imaging data, such as the 2D motion fields which form the calibration and update data in the proposed motion model, are often inherently high-dimensional. In recent years manifold learning was shown to be useful in the analysis of motion in such data, either directly on image intensities (Wachinger et al., 2011; Fischer et al., 2014) or on motion fields (Souvenir et al., 2006), making use of the fact that similar points in the low-dimensional space correspond to similar motion states. Applications include the extraction of respiratory gating navigators

from MR and ultrasound images (Wachinger et al., 2011) and the derivation of navigators from X-ray fluoroscopy images for motion modelling in image guided minimally invasive surgeries (Fischer et al., 2014).

Manifold alignment techniques establish correspondences between multiple related datasets, which are not directly comparable in high-dimensional space, by aligning the low-dimensional manifold structure. Such approaches allow the identification of similar data points in distinct datasets. Recently a number of works in medical imaging have exploited the potential of such techniques. For example, (Bhatia et al., 2012) applied manifold alignment for the robust region-wise separation of cardiac and respiratory motion from cardiac MR images. Georg et al. (2008) used a basic manifold alignment approach for the gating of lung CT volumes.

In our previous work (Baumgartner et al., 2013; 2014a) we developed the simultaneous groupwise manifold alignment (SGA) technique on which this work is based. The technique was proposed to relate partial image intensity information from coronal high-resolution 2D MR slices based on the groupwise alignment of manifolds derived from data acquired at the different slice positions. This allowed for accurate 4D MR reconstructions which could be used to correct simultaneously acquired PET data for respiratory motion.

In this paper, we extend our previously proposed SGA technique to build an autoadaptive motion model from multiple 2D motion fields derived from sagittal 2D MR slices acquired at different anatomical positions. These 2D motion fields can then be combined to estimate 3D motion in the thorax. To allow this, significant changes to the original SGA methodology were necessary. (1) manifold alignment was performed on motion fields rather than images; (2) the method was extended to use a combination of coronal and sagittal slices; (3) in order to estimate motion at points where the respiratory pattern is not sufficiently sampled, a new interpolation scheme on the manifold was developed. The present work was presented with preliminary results in Baumgartner et al. (2014b). In this work we extend the methodology by incorporating slices of different orientation into the model and include more extensive evaluations on real and synthetic data. The technique is evaluated for its feasibility in MR-guided interventions using real 20 min MR scans of healthy volunteers, and on synthetic MR data with two different breathing types.

## 2. Background

In the following we will briefly review the necessary theory to understand our previous work, as well as the extension of it which will be introduced in Section 3. Our goal in this work, as well as in our previous works (Baumgartner et al., 2013; 2014a), is to obtain correspondences between high-dimensional data obtained from slices at different anatomical positions by finding correspondences between their low-dimensional representations. In Baumgartner et al. (2013; 2014a) the high-dimensional data were 2D MR slices. In this work, as well as in the preliminary version of this work (Baumgartner et al., 2014b), they are motion fields derived from such slices. In the next section, we will give an introduction to locally linear embeddings (LLE) (Roweis and Saul, 2000), which is the manifold learning technique used in this work, and how it can be applied to data obtained from one slice position. Next, we will discuss how correspondences in the low-dimensional embedded space can be established for data obtained from two slice positions (Section 2.2). In Section 2.3 we will then show how the concept can be extended to many slice positions. For an overview of the mathematical notation used in the remainder of this paper refer to Table 1.

**Table 1**

List of frequently used mathematical notations in this paper.

Variable	Size	Description
$b_p^i$	–	The $i$ -th 2D MR image acquired at slice position $p$ .
$c_p^i$	–	2D motion field derived by registering $b_p^i$ to an exhale slice.
$D$	1	The dimensionality of the input data, i.e. the number of pixels in one slice (in Section 2), or the total number of motion components (in Section 3).
$d$	1	The dimensionality to which the data gets reduced in the manifold alignment step.
$\tau_p$	1	Total number of slices acquired from slice position $p$ at a specific time of the application phase.
$x_p^i$	$D \times 1$	Vectorised image $b_p^i$ (in Section 2), or vectorised motion field $c_p^i$ (in Section 3).
$X_p$	$D \times \tau_p$	Matrix containing $\tau_p$ vectorised data points $x_p^i$ .
$y_p^i$	$d \times 1$	Low-dimensional point corresponding to $x_p^i$ .
$Y_p$	$d \times \tau_p$	Matrix containing $\tau_p$ low-dimensional points.
$W_p$	$\tau_p \times \tau_p$	Reconstruction weight matrix in LLE cost function (see Eq. (1)).
$M_p$	$\tau_p \times \tau_p$	Centred version of weight matrix $W_p$ .
$U_{pq}$	$\tau_p \times \tau_q$	Similarity kernel matrix connecting data from slice positions $p$ and $q$ .
$S_p$	–	Sagittal slice position $p$ .
$C$	–	Coronal slice position.
$\mathcal{G}(\cdot, \cdot)$	–	Groupwise embedding of data from two different slice positions.

### 2.1. Manifold learning on one dataset

LLE can be used to reduce the dimensionality of a high-dimensional imaging dataset  $X_p \in \mathbb{R}^{D \times \tau_p}$ . Such a dataset can be derived by vectorising all  $\tau_p$  images  $b_p^i$  acquired at a single slice position  $p$ , or alternatively, by additionally deriving a motion field  $c_p^i$  for each image and vectorising those. Each of the columns  $x_p^i \in \mathbb{R}^D$  of  $X_p$ , can be thought of as a point in  $D$  dimensional space where  $D$  is the number of pixels in the original image  $b_p^i$ , or the number of motion components in  $c_p^i$ .

In LLE, dimensionality reduction is accomplished by first forming a  $k$ -nearest neighbour graph of the data based on the  $\mathcal{L}_2$ -distance between the data points. The key assumption is that the neighbourhood of each point and its nearest neighbours are on a locally linear patch of the manifold and that therefore each point can be reconstructed as a linear combination of its nearest neighbours. The optimal contributions of each point  $j$  to the reconstruction of  $i$  are given by a weight term  $w_p^{ij}$ . The matrix  $W_p$  containing all the weights can be calculated in closed form as described in Roweis and Saul (2000). A  $d$ -dimensional embedding, where  $d \ll D$ , preserving this locally linear structure is given by the  $Y_p \in \mathbb{R}^{d \times \tau_p}$  minimising the following cost function:

$$\phi(Y_p) = \sum_i \|y_p^i - \sum_{j \in \eta(i)} w_p^{ij} y_p^j\|^2 = \text{Tr}(Y_p M_p Y_p^T). \quad (1)$$

Here  $\eta(i)$  is the neighbourhood of the data point  $i$ ,  $\text{Tr}(\cdot)$  is the trace operator, and  $M_p = (I - W_p)^T (I - W_p)$  is the centred weight matrix. This optimisation problem can be solved under the constraint that  $Y_p^T Y_p = I$  by calculating the eigendecomposition of  $M_p$ . The embedding  $Y_p$  is given by the eigenvectors corresponding to the second smallest to  $d + 1$  smallest eigenvalues of  $M$  (Roweis and Saul, 2000).

### 2.2. Simultaneous embedding of two datasets

Separate datasets generated by the same mechanics, e.g. respiration, will typically lie on similar manifolds, as is the case, for example, for two datasets  $X_p$  and  $X_q$  acquired from two different anatomical positions  $p$  and  $q$ . It has been shown by Wachinger et al. (2011) that this holds for 2D MR data from neighbouring,



as well as distant slice positions and for slice positions of different orientations. This can be explained by the observation that the manifold of data from each slice position is defined by the principal modes of variation, which depend on the respiratory motion common to all slice positions rather than the absolute appearance of the slices. This knowledge can be used to identify corresponding data points from the two datasets. In our case this means finding corresponding data acquired from different anatomical positions but with similar respiratory phases. Unfortunately, generally embeddings obtained from different datasets are not aligned in the low-dimensional space as they may vary due to flipping or rotations of the eigenvectors, and slight variations in the manifold structure. One approach to find aligned manifold embeddings  $Y_p$ ,  $Y_q$  of two high-dimensional datasets is to embed them simultaneously. The cost function of LLE lends itself ideally to be extended to two datasets. The problem of finding a simultaneous embedding can be written as the following minimisation problem

$$\underset{Y_p, Y_q}{\operatorname{argmin}} \phi(Y_p) + \phi(Y_q) + \mu \cdot \psi(Y_p, Y_q), \quad (2)$$

where  $\phi$  is the embedding error within the respective datasets  $p$  and  $q$  (intra-dataset cost functions) as given by Eq. (1), and  $\psi$  is the embedding error *between* the two datasets (inter-dataset cost function). This term ensures that corresponding points will be embedded close to each other. Note that typically no correspondences between the datasets are known *a priori*. Rather, corresponding data points must be identified at runtime. The parameter  $\mu$  regulates the influence of the inter-dataset cost function  $\psi$  on the embedding.

The cost function  $\psi$  can be defined as follows

$$\psi(Y_p, Y_q) = \sum_{i,j} (y_p^i - y_q^j)^2 u_{pq}^{ij}, \quad (3)$$

where

$$u_{pq}^{ij} = K(x_p^i, x_q^j),$$

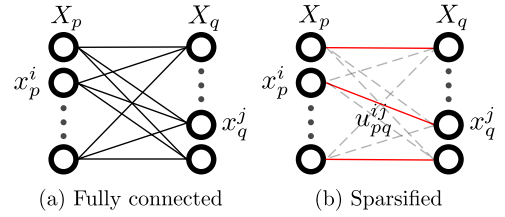
is a (non-symmetric) similarity kernel of the form

$$K(x_p^i, x_q^j) = \exp \left( -\frac{\epsilon(x_p^i, x_q^j)^2}{2\sigma^2} \right). \quad (4)$$

Above,  $\epsilon(\cdot, \cdot)$  is a distance function which must be defined such that the kernel  $K(\cdot, \cdot)$  will take large values for similar data points and small values for dissimilar data points. The similarity kernel can be written as a matrix  $U_{pq}$  with high values connecting similar images from slice positions  $p$  and  $q$ . In Baumgartner et al. (2013) we used intensity-based distance of slices from neighbouring positions  $p$  and  $q$  to define the distance function  $\epsilon$ . In Baumgartner et al. (2014a) we improved the method by correcting this distance measure for deformations that may occur between slice positions using non-rigid registration. As we will show in Section 3.2, in this work we used a similar kernel to the one described in Baumgartner et al. (2014a) for neighbouring slices of the same orientation, and a novel kernel based on motion similarities in the slice overlap for slices with different orientations (see Section 3.2.2).

Next, the similarity kernel  $U_{pq}$  is sparsified to increase the robustness of the method. We use the Hungarian algorithm (Kuhn, 1955) to identify the optimal one-to-one mapping between the two datasets. That is, the mapping where each data point from dataset  $X_p$  is connected to exactly one data point from  $X_q$  and the sum of the remaining weights  $u_{pq}^{ij}$  is maximised. This is illustrated in Fig. 2. We found in Baumgartner et al. (2013) that this kind of sparsification is more robust than a simple nearest neighbour sparsification.

Using the sparsified kernel the problem in Eq. (2) can then easily be rewritten in matrix form and can be solved as an eigenvalue



**Fig. 2.** Graph sparsification of similarity kernel  $U_{pq}$ . The two columns of circles in (a) and (b) represent all the data points acquired from the respective slice positions  $p$  and  $q$ . (a) shows a fully connected similarity kernel where all of the edges  $u_{pq}^{ij}$  exist. (b) shows the sparsified similarity kernel where each  $x_p^i$  is connected to exactly one  $x_q^j$  and the sum over the remaining edge weights  $u_{pq}^{ij}$  is maximised.

problem analogous to the original LLE algorithm as is described in Baumgartner et al. (2013; 2014a).

### 2.3. Embedding data from many slice positions

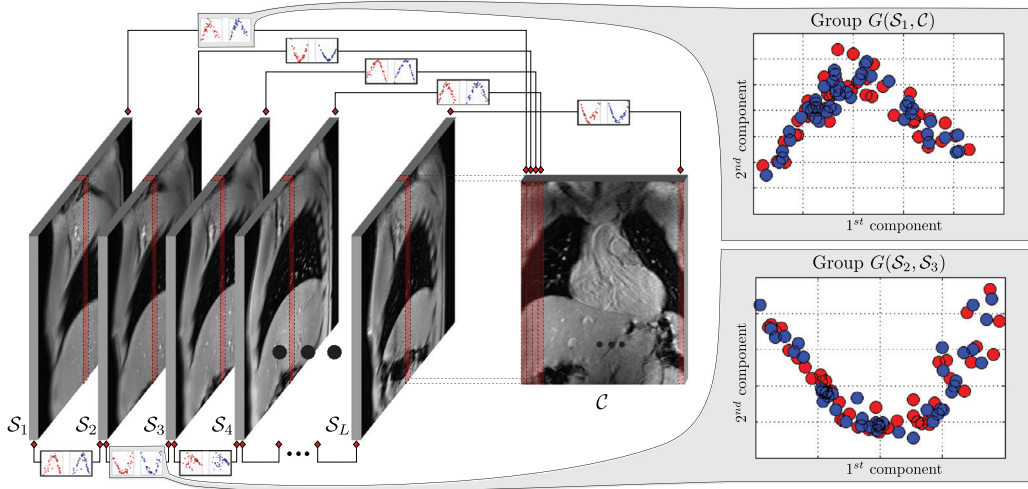
In our previous work we were interested in simultaneously embedding not only two but up to 40 datasets, i.e. the number of slice positions from which our slice-by-slice data originated. It is possible to augment the minimisation problem in Eq. (2) to an arbitrary number of datasets, however, it is not trivial to define the similarity kernel for non-neighbouring slice positions and leaving these kernels undefined leads to an unstable optimisation problem (Baumgartner et al., 2013).

In Baumgartner et al. (2013; 2014a) we proposed to embed the data in overlapping groups of two consisting of data from neighbouring slice positions, which is a much more manageable problem. In order to relate the groups to each other they are chosen so that they share some data. In particular, data from each slice position appears in two different groups. For example, one group may contain data from slice positions 7 and 8, and another data from slice positions 8 and 9. The aligned embeddings of the data gathered from slice position 7 can then be related to the embeddings from slice position 9 by means of the shared data from slice position 8.

## 3. Materials and methods

Simultaneous groupwise manifold alignment was originally proposed for coronal input slices, since the anatomy changes less from slice position to slice position in this plane. In the motion modelling context, however, it is essential that the input data captures as much of the motion as possible. It is well known that respiratory motion is largest in the superior-inferior (S-I) and anterior-posterior (A-P) directions (Seppenwoolde et al., 2002). Therefore, we focused on sagittal input slices in this work. Unfortunately, SGA as described in the previous section is not robust to sagittal input slices because respiratory information often gets lost while propagating from group to group through the body centre, where anatomy changes rapidly from slice position to slice position. Therefore, here we extend the technique to additionally incorporate data acquired from a single coronal slice position to aid this transition through the body centre. Note that the preliminary version of this work (Baumgartner et al., 2014b) included only sagittal slices.

In the following we will describe our proposed method for adaptive motion modelling following the three motion modelling stages outlined in the introduction. We first show how we acquire sagittal and coronal input slices and derive 2D motion estimates from them to train the model. Next, we show how SGA can be extended to use motion fields, rather than images, and how different slice orientations are incorporated into the model. Lastly, we show how the model can be updated during a treatment in the application phase, and how this leads to continuous adaptivity.



**Fig. 3.** Schematic of group connections through simultaneous manifold embeddings. Motion fields from neighbouring and orthogonal slices can be embedded simultaneously using appropriate similarity kernels leading to aligned embeddings. Two close-up views of aligned manifold embeddings, originating from a dataset with 50 motion fields per slice position, are shown on the right hand side.

### 3.1. Calibration scan

The image acquisition scheme of the present work is an extension of the acquisition scheme used in Baumgartner et al. (2014a). We divide the entire region of interest into adjacent sagittal slice positions  $S_1, \dots, S_p$  each spanning 8 mm. Here, the region of interest is the entire thorax including the liver and lungs. Additionally, we also choose one coronal slice position  $C$  to help with the propagation of respiratory information between distant sagittal slices. A schematic of the slice positions is shown in Fig. 3. The coronal slice is chosen such that it coincides with the dome of the left hemi-diaphragm in order to maximise the amount of captured respiratory motion.

We acquire 2D images  $b_p^i$  from these slice positions in a slice-by-slice fashion, iterating through the slice positions, first the sagittals then the coronal, until each slice position is covered  $\tau_p$  times. In order to isolate the respiratory motion, in this study, we acquire only one slice per heart beat at systole. However, in principle, similar data could also be acquired without cardiac gating which would significantly reduce overall acquisition times. The acquisitions were carried out on a Philips Achieva 3T MR scanner using a T1-weighted gradient echo sequence with an acquired in-plane image resolution of  $1.4 \times 1.4 \text{ mm}^2$ , a slice thickness of 8 mm, repetition and echo times (TR and TE) of 3.1 and 1.9 ms, a flip angle (FA) of 30 degrees, and a SENSE-factor of 2. The field of view covering the entire thorax was  $400 \times 370 \text{ mm}^2$ , and each slice took around 180 ms to acquire. To cover the entire thorax typically around 30 sagittal slice positions were needed. Additionally, we acquired exhale slices  $b_p^{(exh)}$  using the same slice-by-slice protocol in a scan consisting of two consecutive breath holds. The volunteers were instructed to try and reproduce the same exhale position as best as they could. Lastly, we acquired a 1D pencil beam navigator immediately before each dynamic image solely for the purpose of validating our method.

In the next step we derive 2D motion fields  $c_p^i$  for each slice position by registering each of the  $\tau_p$  2D images  $b_p^i$  to the corresponding slice  $b_p^{(exh)}$  from the exhale breath hold image. We used the NiftyReg implementation of a non-rigid B-spline registration algorithm with 3 hierarchy levels, a final grid spacing of 15 mm in each direction and no bending energy penalty term (Modat et al., 2010). The vectorised motion fields  $c_p^i$  derived from the slice positions  $S_p$  and  $C$ , respectively, form the datasets  $X_{(p, sag)}$  and  $X_{(cor)}$ .

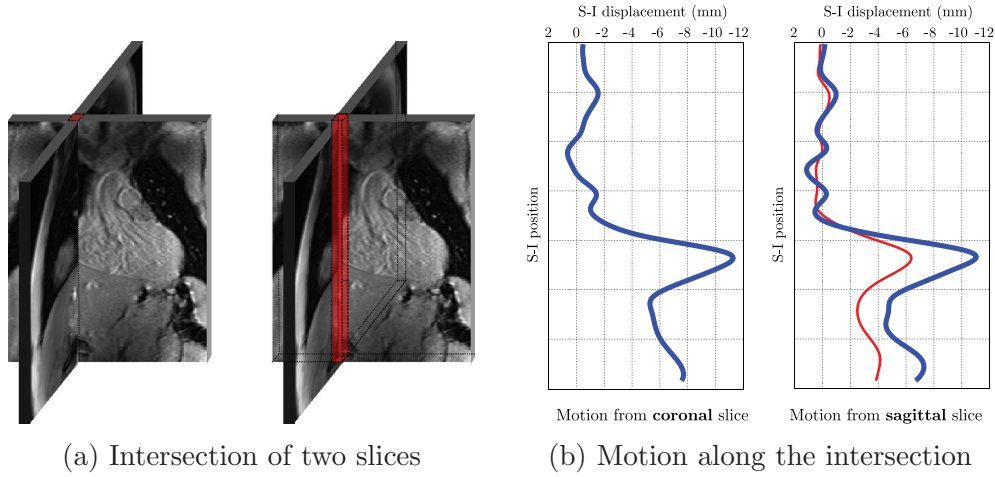
### 3.2. Motion model formation

We propose that a groupwise embedding of all the motion data acquired during the calibration phase can be viewed as a motion model as it contains all respiratory information collected during the calibration and can be applied using new 2D motion information, as will be explained in Section 3.3. Thus, in order to form the motion model we perform an embedding of all the sagittal and coronal slices acquired during the calibration phase in groups of two as shown in Figs. 3 and 5. The embeddings can be performed as described in Sections 2.2 and 2.3. To embed the motion data derived from sagittal and coronal motion fields we need to introduce two significant methodological novelties. First, we need new similarity kernels of the form described in Eq. (4) with which motion fields of slices with the same as well as slices with different orientations can be compared. In particular, we need to define appropriate distance functions  $\epsilon(\cdot, \cdot)$  for both these cases. Secondly, we need a new propagation scheme which allows respiratory information to propagate across the body centre.

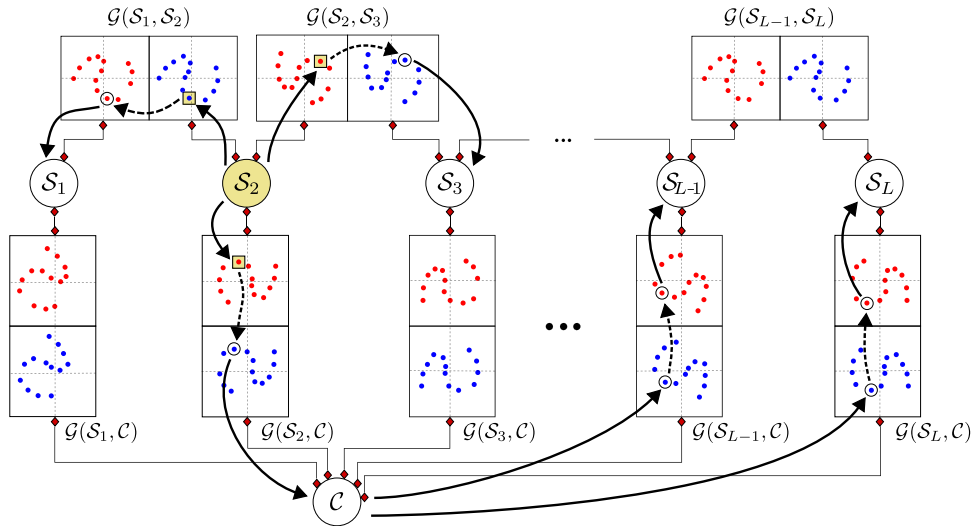
#### 3.2.1. Distance functions for neighbouring slices of the same orientation

We base our choice for neighbouring sagittal motion data on the robust distance measure we proposed in Baumgartner et al. (2014a) for coronal images and adapt it to the scenario of motion fields. For two neighbouring slice positions  $S_p$  and  $S_q$  the distance of data points  $x_{(p, sag)}^i$  and  $x_{(q, sag)}^j$  is assessed based on the  $\mathcal{L}_2$ -distance of the corresponding motion fields  $c_{(p, sag)}^i$  and  $c_{(q, sag)}^j$ . In order to account for the changes in anatomy between sagittal slices we transform one of the motion fields into the coordinate system of the other using transformations  $T_{q \rightarrow p}$ ,  $T_{p \rightarrow q}$  which we obtain by registering the breath hold slices  $b_{(p, sag)}^{(exh)}$  to  $b_{(q, sag)}^{(exh)}$ , and vice versa. As is discussed in more depth in Baumgartner et al. (2014a), transporting motion fields to the new coordinate system is achieved using the method proposed by Rao et al. (2002). To increase robustness we average the results of the comparisons in the spaces of slice position  $S_p$  and  $S_q$ . The final distance measure is defined as

$$\epsilon_{neighb.}(x_{(p, sag)}^i, x_{(q, sag)}^j) = \frac{1}{2} \mathcal{L}_2(c_{(q, sag)}^j, T_{p \rightarrow q}(c_{(p, sag)}^i)) + \frac{1}{2} \mathcal{L}_2(c_{(p, sag)}^i, T_{q \rightarrow p}(c_{(q, sag)}^j)). \quad (5)$$



**Fig. 4.** Derivation of similarity kernel based on motion in slice overlap. (a) Illustration of the overlap of two orthogonal slices, (b) S-I motion components derived from the overlapping area from a coronal (left) and sagittal (right) slice position. We show two possible S-I motion components originating from the sagittal slice: One that is similar to the one derived from the coronal slice (blue) and hence corresponds to a similar motion state, and one that is dissimilar (red) and consequently corresponds to a different motion state. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Schematic of the connection of slice positions by means of pairwise embedding and propagation of respiratory information through the manifolds. Assuming a new input slice at  $S_2$  the neighbouring groups can be directly updated as indicated by the squares with yellow background. Then through a combination of nearest neighbour searches (dotted arrows) and group transitions based on shared data (solid arrows) low-dimensional coordinates, which correspond to the respiratory state, can be propagated to all remaining slice positions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A possible source of errors are structures appearing and disappearing from the plane due to through-plane motion. In Baumgartner et al. (2014a) we showed that, despite such effects, including the transformations  $T_{q \rightarrow p}$ ,  $T_{p \rightarrow q}$  significantly improves the matching accuracy compared to the simple  $\mathcal{L}_2$ -distance between images. Since the changes from sagittal slice position to sagittal slice position can be even larger than for coronal slices, we expect this effect to be more pronounced for slices of this orientation. Note that we used normalised cross correlation in the preliminary version of this work (Baumgartner et al., 2014b). However, in this work we found the  $\mathcal{L}_2$ -distance to be a more robust measure in this context.

### 3.2.2. Distance function for slices with different orientation

To define a distance function for two slices acquired from a sagittal slice position  $S_p$  and a coronal slice position  $C$  we use the fact that such slices have an overlap and thus visualise the same anatomy in the overlapping region as is illustrated in the example in Fig. 4a. Motion estimates derived from two such slices share the S-I motion component along the slice overlap. If  $\sigma_{(p, \text{sag})}^i$  is the

S-I motion in the overlapping region originating from the  $i$ -th acquired sagittal slice at  $S_p$  and  $\sigma_{(cor)}^j$  the motion in the overlapping region from the  $j$ -th acquired coronal slice, we define the distance function as

$$\epsilon_{orthog.}(\mathbf{x}_{(p, \text{sag})}^i, \mathbf{x}_{(cor)}^j) = \mathcal{L}_2(\sigma_{(p, \text{sag})}^i, \sigma_{(cor)}^j). \quad (6)$$

To illustrate this we show examples of S-I line motions  $\sigma_{(p, \text{sag})}^i, \sigma_{(cor)}^j$  originating from sagittal and coronal slices in Fig. 4b. The left hand side shows S-I motion extracted along the intersection (highlighted in Fig. 4a) from a coronal slice and the curves on the right hand side show two possible S-I motions extracted from the same region from the sagittal slice position. The blue curve shows a good match in respiratory position of the sagittal to the coronal slice and will lead to a low distance in Eq. (6). Conversely, the motion in the sagittal slice from which the red curve was extracted has a higher distance to the coronal slice motion and thus corresponds to a different respiratory state.



### 3.2.3. Group connectivity and propagation of respiratory information

By using the distance measures defined in Eqs. (5) and (6) we are able to simultaneously embed any two neighbouring sagittal slice positions  $S_p$  and  $S_q$  and any overlapping sagittal and coronal slice positions  $S_p$  and  $C$ . This is achieved by converting the distances into similarities using Eq. (4) and then solving the general optimisation in Eq. (2) to obtain an embedding.

Similar to our previous work this allows us to embed data from all acquired slice positions in overlapping groups of two. In this work, we again connect all neighbouring sagittal slice positions by embedding them in overlapping groups. However, in addition, we align the data from each sagittal slice position together with the data from the coronal slice position. This is illustrated in Fig. 3. By embedding the data in this way, the 2D motion fields from all slice positions are embedded in three groups, with the exception of  $S_1$  and  $S_L$  which don't have a left-hand or right-hand neighbour, respectively. For example, data from slice position  $S_3$  is embedded in the groups  $\mathcal{G}(S_2, S_3)$ ,  $\mathcal{G}(S_3, S_4)$  and  $\mathcal{G}(S_3, C)$ .

Note that the data within each group are aligned, as is illustrated by the close-up views of  $\mathcal{G}(S_1, C)$  and  $\mathcal{G}(S_2, S_3)$  in Fig. 3.

As opposed to our earlier works (Baumgartner et al., 2014a; 2013) there now no longer is just one path from each slice position to each other slice position. Rather, the different slice positions are now connected by a network of groups as is illustrated in Fig. 5.

This is a crucial element of our proposed technique which allows propagation of respiratory information in the form of low-dimensional embedded coordinates from slice position to slice position without having to go through difficult areas such as the body centre where there are larger anatomical differences between adjacent slices.

In the following section we will outline how low-dimensional coordinates obtained from a 2D input motion field can be propagated from slice position to slice position.

### 3.3. Model updating and adaptivity

After the calibration scan and model formation phase the model is ready to be applied. During the application phase slices can be acquired in the same slice-by-slice fashion as described in Section 3.1. That means each input image is again acquired at a different slice position and can have sagittal or coronal orientation.

From each of these slices we can derive a new 2D motion estimate, embed this motion estimate in the groups containing data from this slice position and reconstruct a pseudo 3D motion estimate by looking up corresponding 2D motions from all other slice positions. Note that the resulting 3D motion fields will lack the left-right (L-R) motion component. The new 2D motion, as well as being used as the surrogate input to the motion model, is retained in the manifold embeddings of the appropriate groups. This leads to the desired autoadaptivity. Each of these steps will be explained in detail below. We will illustrate the process using the example shown in Fig. 5.

#### 3.3.1. Obtaining a 2D update motion estimate

In a first step, the most recently acquired image  $b_p^{(new)}$  is registered in 2D to the corresponding breath hold slice  $b_p^{(exh)}$  in order to form the current update motion field  $c_p^{(new)}$ . That is, the motion field  $c_p^{(new)}$  obtained in this way acts as the surrogate data for the motion model application. We used the same registration parameters as in the initial calibration (see Section 3.1). On a workstation with 8 cores clocked at 2.7 GHz this operation took around 500 ms. In the example in Fig. 5, we assumed that the newest slice was acquired at slice position  $S_2$ , which is highlighted in yellow. Note that only the registrations from  $b_p^{(new)}$  to  $b_p^{(exh)}$  have to be performed during the application phase. The registrations across slice

positions (i.e.  $b_p^{(exh)}$  to  $b_q^{(exh)}$ ), which are required for the registration based similarity kernel described in Section 3.2.1, only need to be performed once during the model formation.

#### 3.3.2. Estimating current 3D motion

In order to estimate 3D motion from partial motion information provided by the single input 2D motion field  $c_p^{(new)}$ , the motion from the newly acquired slice must be related to that from all other slice positions. First, all groups which contain data from the slice position at which the new slice was acquired must be recalculated. If, as in our example, the current update slice was acquired at slice position  $S_2$ , the groups  $\mathcal{G}(S_1, S_2)$ ,  $\mathcal{G}(S_2, S_3)$  and  $\mathcal{G}(S_2, C)$  must be re-evaluated (see Figs. 3 and 5). To achieve this the dataset  $X_2$  is simply augmented by the new entry and the respective embeddings are recalculated. Updating just a few groups is relatively fast and on average took fewer than 100 ms in our single thread MATLAB implementation.

As is shown in Fig. 5, the new motion field now has a corresponding low-dimensional point in each of the low-dimensional embeddings which include dataset  $X_2$ . These points are highlighted by squares with yellow backgrounds in Fig. 5. The coordinates of these low-dimensional embedded points are propagated from group to group following the shortest path, i.e. using the path requiring the fewest group transitions. This is done by making use of the fact that the groups share datasets, and that the datasets within a group are aligned. This effectively means that neighbouring slices are updated through the sagittal-to-sagittal groups and further away sagittal slice positions are connected through the coronal slice. Other methods for choosing the update paths taking into account the quality of the embedding were also investigated and may lead to small improvements. However, in this work for simplicity we confine our analysis to the shortest path method.

Following our example, first the nearest neighbours of low-dimensional points corresponding to data from  $S_1$ ,  $S_3$  and  $C$  are found in the respective groups. The nearest neighbour operation is indicated by the dotted arrows in Fig. 5 and the nearest points are indicated by circles. The high-dimensional motion fields corresponding to the circled points in  $\mathcal{G}(S_1, S_2)$  and  $\mathcal{G}(S_2, S_3)$  are at the same respiratory position as the input slice. Note that in the example in Fig. 5, only one nearest neighbour is shown per slice position. In reality,  $\kappa$  nearest neighbours are identified at this stage and the corresponding 2D motion fields are interpolated.

Next, motion fields from all other sagittal slice positions are chosen by using the coronal slice. The nearest low-dimensional neighbour of the input point in group  $\mathcal{G}(S_2, C)$  is then transported to all other groups containing the coronal slice because that same point exists in all other groups. Note that only the closest neighbour is transported across groups. From there the corresponding points from the sagittal motion datasets are again found by looking up the  $\kappa$  nearest neighbours.

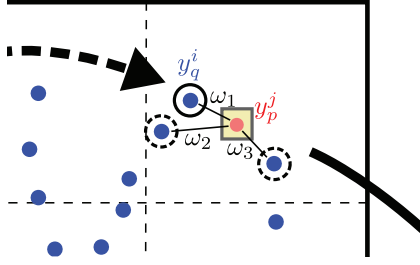
At the end of this process,  $\kappa$  2D motion fields have been identified for each slice position. In the following section, it will be described how these 2D motion fields can be combined to arrive at an interpolated motion estimate for each slice position, and how these partial 2D motion estimates can then be stacked into a full pseudo 3D motion field.

#### 3.3.3. Interpolating motion fields on the manifold and 3D reconstruction

If the motion model has not yet fully sampled all the possible motion states of the new breathing pattern, it is important that it has the ability to interpolate between the motion states which are already there.

In order to estimate the 2D motion field for a slice position,  $\kappa$  nearest neighbours are identified for each slice position as described in the previous section. The estimated motion field is then





**Fig. 6.** Close-up view of the right-hand side of group  $\mathcal{G}(S_2, S_3)$  in Fig. 5. Here  $p = S_2$  and  $q = S_3$ . In addition to the nearest neighbour in the manifold embedding  $y_p^j$  (blue dot in solid circle), the figure shows the other  $\kappa - 1$  nearest neighbours (blue dots in dotted circles), the original point belonging to the manifold of  $S_2$ , i.e.  $y_q^j$  (red dot in yellow box), and the distances  $\omega_i$  from which the similarities  $s_i$  are derived. The similarities are then used to form a weighted average of the corresponding motion fields. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

given as a weighted average of the  $\kappa$  motion fields corresponding to those nearest neighbours. That is, the estimated motion field for a slice position  $q$  is given by

$$c_q^{(est)} = \frac{\sum_{i \in \eta(y_p^j)} s_i c_q^i}{\sum_{i \in \eta(y_p^j)} s_i}, \quad (7)$$

where  $\eta(y_p^j)$  are the  $\kappa$  nearest neighbours on the manifold of slice position  $q$  to the low-dimensional point  $y_p^j$  from a slice position  $p$  which is sharing a group with  $q$ . Furthermore,  $s_i = \frac{1}{\omega_i}$ , where  $\omega_i$  is distance of each neighbour to  $y_p^j$  in the manifold embedding. This process is illustrated in the close-up of  $\mathcal{G}(S_2, S_3)$  which is shown in Fig. 6. Note that the distances  $\omega_i$  could also be used to estimate how well the manifold is sampled around a given motion state which could be used to derive a confidence measure for motion estimation. However, this is not investigated further in this paper.

The  $c_p^{(est)}$  for all sagittal slice positions  $S_1, \dots, S_L$  are then stacked into a pseudo 3D motion field, i.e. a dense 3D motion field lacking the L-R component. This 3D motion estimate is the output of the motion model given the 2D surrogate image  $b_p^{(new)}$  as input. Note that the coronal motion field from slice position  $\mathcal{C}$  is currently only used for the propagation of manifold coordinates but not for the reconstruction. This means the motion field from the coronal slice position will not be part of the volume.

### 3.3.4. Updating the model and adaptivity

The mechanism of embedding the new slice motion field into the corresponding groups automatically updates the model. The new motion fields, after being used to stack a 3D motion field, stay in the model and may be used themselves in the future for new motion estimations.

In this manner, as the application phase goes on, more and more data is added to the model making it adaptive. In the case of respiratory drift or changes in the breathing pattern the model does not lose its validity but rather incorporates these new motion patterns.

## 4. Experiments and results

In order to validate our proposed autoadaptive motion model (AAMM) technique we compared it to two versions of the method, each with one of our major novelties removed: AAMM without the autoadaptivity, and AAMM without the incorporation of slices of different orientations in the groupwise manifold alignment step. That is, we compared the following techniques:

- AAMM: The proposed autoadaptive motion modelling method as described in Section 3.
- AAMM (no adapt.): The proposed method without the adaptivity. This means that after each update step we discarded the most recently added 2D motion field again.
- SGA: The proposed AAMM method but without using the coronal input slices. Essentially, this is our simultaneous groupwise manifold alignment technique (Baumgartner et al., 2014a) extended to use sagittal motion fields instead of coronal images. The adaptivity was implemented in exactly the same way as for AAMM with the sole exception that there were no coronal input slices.

The experiments in this section aim to answer the following main research questions:

1. How does autoadaptivity affect the motion estimates after a short calibration phase with a constant breathing pattern?
2. Can the autoadaptive motion model adapt to a previously unseen breathing pattern?
3. Can the method be applied using real MR data?

In order to pursue these questions, we evaluate the three methods described above on synthetic data derived from 6 volunteer scans and on real data acquired from 4 volunteers. In Experiment 1 (Section 4.2.2), synthetic data representing normal free breathing is generated to answer the first research question. In Experiment 2 (Section 4.2.3), additionally, synthetic data which corresponds to a deep breathing pattern is generated in order to investigate the second of the above questions. Lastly, in Experiment 3, the algorithms are evaluated on real volunteer scans acquired over 20 min. Using this data we seek to answer the third research question and investigate the method's feasibility in a real MR-guided scenario. Furthermore, it is investigated how the methods respond to natural, gradual changes in the breathing pattern which may not have been observed during model calibration and whether the proposed technique can maintain motion estimation accuracy in such circumstances.

Note that no comparison of AAMM to any other state-of-the-art motion modelling techniques was performed. All motion models from the literature follow the traditional motion modelling paradigm (see Fig. 1) and could not be built using the 2D slice-by-slice data used in this work. Thus an evaluation on equal terms was not feasible.

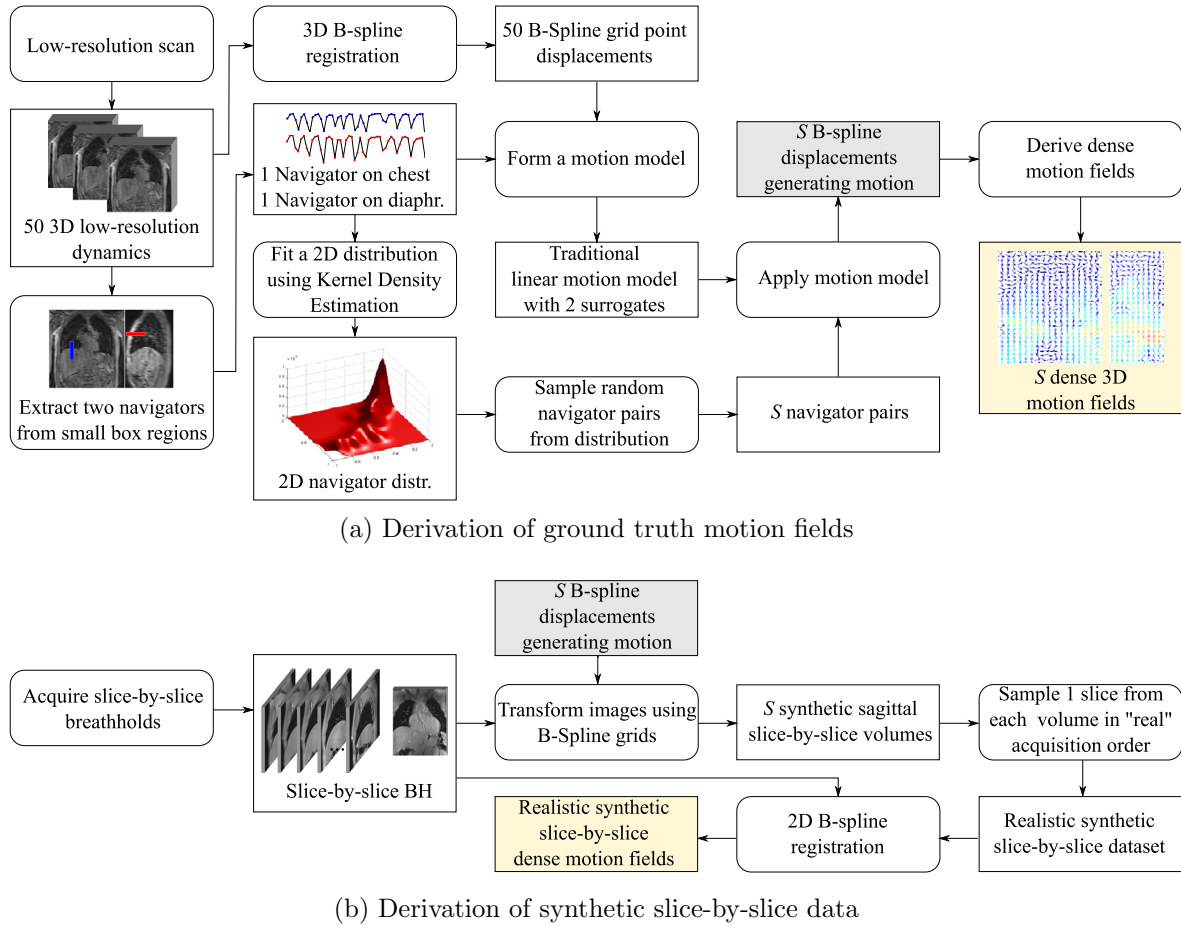
### 4.1. Parameter choices

We chose the free parameters of the investigated techniques based on our previous experience with SGA. In Baumgartner et al. (2014a) we investigated the optimal values for the reduced dimensionality  $d$ , and the weighting parameter  $\mu$  (see Eq. (2)). Furthermore, we found that the method is not significantly affected by the choices of the kernel shape parameter  $\sigma$  (see Eq. (4)) and the number of nearest neighbours  $k$  used in the LLE cost function (see Eq. (1)). Based those findings, here we chose the following parameters for all of the methods:  $\sigma = 0.5$ ,  $\mu = 0.25$ ,  $d = 3$ .

In Baumgartner et al. (2014a) we set the parameter  $k$  to half the number of acquired slices per slice position. In this work, the number of 2D motion fields increased steadily as the model was applied. Consequently, we continually adapted the parameter  $k$  to the current data size. That is, we set

$$k = \text{round}(\tau_p/2),$$

where  $\tau_p$  is the number of 2D motion fields per slice position currently part of the respective groups.



**Fig. 7.** Generation of synthetic slice-by-slice data. (a) Generation of ground truth motion fields from a 3D low-resolution MR scan, (b) generation of synthetic slice-by-slice data by applying the ground truth motion to slice data acquired at end-exhale. The procedure was performed twice, once for normal breathing data and once for deep breathing data.

## 4.2. Experiments on synthetic data

To quantitatively assess our method we generated very realistic synthetic 2D motion fields containing two breathing types by mimicking an actual slice-by-slice acquisition process as will be described in Section 4.2.1. We separately generated two synthetic datasets with two different breathing types: normal breathing, and deep breathing. The normal breathing data was used in Experiment 1 (Section 4.2.2), and a combination of both breathing types was used for Experiment 2 (Section 4.2.3).

### 4.2.1. Generation of realistic synthetic data

The method to generate synthetic data in this work differs from the approach we took in our previous works (Baumgartner et al., 2014b; 2013; 2014a). Previously, we directly transformed a breath hold volume using motion fields derived from low-resolution volumes. However, using this method resulted in a dataset where all slice positions have an identical distribution of respiratory states, which is unrealistic and oversimplifies the problem of finding matching motion states from different slice positions. Here, we aimed to generate 50 synthetic slices per slice position, such that no respiratory state was exactly repeated in the whole dataset.

The underlying idea of our data generation framework was to first build a simple linear subject specific motion model based on two 1D navigators and 3D motion fields derived from short dynamic low-resolution 3D MR scans under different breathing modes. Note that the motion model used to generate the synthetic data is distinct from the autoadaptive motion model we are

proposing in this work. By generating random samples of synthetic navigator values and using them as input to the motion model we obtained synthetic, but realistic, respiratory motion deformations. Note that the number of random samples that can be generated is not limited and can be freely chosen. The respiratory motion deformations, on one hand, served as a ground truth for our experiments, and on the other hand, were used to generate synthetic slice-by-slice data by transforming a slice-by-slice breath hold scan. This approach had two main advantages: 1) more realistic sampling of respiratory positions, 2) even though only a 50 s scan was used as input, an arbitrary number of 2D slices could be generated using this generating framework.

In the following we explain each step in detail. The generation of the synthetic data is summarised in Fig. 7. We split the description of the generation into two parts:

1. The generation of realistic ground truth motion (see Fig. 7a).
2. The generation of synthetic slice-by-slice images and the derivation of slice-by-slice motion fields from them (see Fig. 7b).

For the generation of the ground truth data, in a first step we acquired two sets of 50 3D low-resolution MR volumes on a Philips Achieva 3T MR system using a cardiac-triggered T1-weighted gradient echo sequence with an acquired image resolution of  $1.5 \times 5 \times 4.1 \text{ mm}^3$  (S-I, A-P, L-R), an acquisition time of approximately 600 ms per volume, a SENSE-factor of 2 in A-P and a SENSE-factor of 4 in L-R, TR/TE = 3.3 ms/0.9 ms, a FA of  $10^\circ$ , and a field of view of  $500 \times 450 \times 245 \text{ mm}^3$  covering the entire thorax. The highest

resolution was chosen in the S-I direction, where most respiratory motion occurs (Seppenwoolde et al., 2002). For the first set of images the volunteers were instructed to breathe freely (i.e. normal breathing). For the second set of images we instructed the volunteers to take slow, deep breaths (i.e. deep breathing).

From these two sets of volumes we generated two separate synthetic datasets; a normal breathing and a deep breathing one. To this end we derived 50 B-spline grid displacements for each breathing type by registering the volumes to an exhale volume chosen manually from the set of normal breathing images. The registration was performed using NiftyReg (Modat et al., 2010) using the same parameters as for the real data, that is, 3 hierarchy levels, a final grid spacing of 15 mm in each direction and no bending energy penalty term. Furthermore, we extracted two series of 50 navigator signals  $s_1, s_2$  from the images by measuring the displacements of small rectangular regions on the dome of the left hemidiaphragm and the anterior chest wall (Savill et al., 2011). We chose two signals to increase the amount of respiratory variabilities captured in the resulting model. Next, we formed a motion model for each breathing type by fitting a linear function of the time series of navigator signal values to the displacements of each B-spline grid point (McClelland et al., 2013), i.e.

$$\mathbf{v}(\mathbf{t}) = \alpha_1(\mathbf{t}) + \alpha_2(\mathbf{t})s_1 + \alpha_3(\mathbf{t})s_2, \quad (8)$$

where  $\alpha_1, \alpha_2, \alpha_3$  are the parameters of the motion model and  $\mathbf{v}(\mathbf{t})$  are the grid displacements at grid location  $\mathbf{t}$ .

In the next step, we fitted a 2D distribution to the navigator signal values for each breathing mode using kernel density estimation (Rosenblatt et al., 1956). We then sampled  $S$  random navigator value pairs  $\tilde{s}_1, \tilde{s}_2$  from this distribution. In the final synthetic datasets each 2D slice was associated with a ground truth 3D motion field. Hence, we needed to draw as many synthetic navigator values as the total number of synthetic 2D slices in each dataset. Since we aimed to generate 50 slices per slice position, we needed to sample  $S = L \cdot 50$  navigator values, where  $L$  is the total number of slice positions in the synthetic sequence. Next, by substituting the sampled values  $\tilde{s}_1, \tilde{s}_2$  into Eq. (8) we obtained  $S$  synthetic B-spline transformations per breathing mode, which were then used for the generation of the ground truth 3D motion fields as well as for the generation of the synthetic slice-by-slice data. The ground truth motion fields were derived simply by interpolating a dense motion field using the voxel sizes of the slice-by-slice breath hold volumes described below.

In the following, we will describe how the synthetic slice-by-slice data was generated (see Fig. 7b). In addition to the low-resolution volumes, we also acquired all sagittal slice positions and one coronal slice position in two exhale breath hold acquisitions using the same protocol as for the real data, which was described in Section 3.1. The breath hold data was then transformed using the synthetic B-spline grid displacements. This led to a sequence of  $S$  synthetic slice-by-slice volumes for each of the breathing types. In the real acquisitions at each time point we can observe only one slice position, and hence, we sampled only one slice from each of these volumes, and discarded the rest. However, note that each slice was still associated with a 3D ground truth motion field (see Fig. 7a). The sampled slices were chosen according to the acquisition order of the real data described in Section 3.1. This sampled data constituted the synthetic slice-by-slice dataset and was the synthetic equivalent to the data obtained from a real slice-by-slice scan. As part of the model calibration phase, the slice-by-slice image data was then registered in 2D to the corresponding breath hold slices in order to obtain slice-by-slice 2D motion fields which are the input to the proposed AAMM. The parameters used for the registration were the same as in Section 3.1.

Note that in order to make the acquisition of 3D dynamic volumes feasible, compromises had to be made in the image quality.

The relatively low-resolutions in A-P and L-R directions, and the high SENSE factors led to artefacts and blurring of certain structures. Furthermore, the residual motion during the volume acquisition may cause slight blurring, especially during deep breathing, which may in turn lead to underestimation of the motion close to end-inhale. Nevertheless, we found that the simulated 2D images and motion fields reasonably approximated a real 2D MR acquisition.

#### 4.2.2. Experiment 1: synthetic training adaptivity

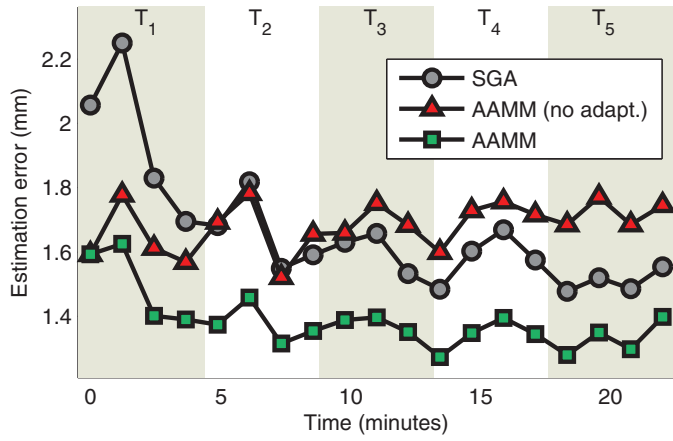
In this section, we investigate the autoadaptive behaviour of our proposed method in the presence of an approximately constant breathing pattern. We quantitatively assessed the motion estimation accuracy using the three compared models on the synthetic, normal-breathing slice-by-slice data which was generated using the technique described in the previous section. Note that the resulting data mimics an acquisition of around 20–25 min. However, it can only reflect breathing patterns observed in the 50 s normal breathing dynamic 3D MR scan. On average L-R motion accounted for 15.85% of the total motion, and the A-P and S-I accounted for 20.44% and 63.71%, respectively.

The motion estimation accuracy was quantitatively assessed using the three compared models on the synthetic slice-by-slice data. Each of the three stages of motion modelling shown in Fig. 1 was performed, i.e. model calibration, model formation and model application. The synthetic data generation can be seen as a synthetic model calibration stage yielding slice-by-slice motion fields. In the next step, we formed the model by embedding a subset of the synthetic data using the three compared methods. We used 10 slices from each slice position for the initial formation of the model. Obtaining this amount of data in a real (cardiac-gated) scan would take approximately 5 min. We then applied the motion model by continually adding all remaining slices one after the other, and at each time step evaluated the accuracy of the estimated motion against the 3D ground truth motion field corresponding to the newest update slice.

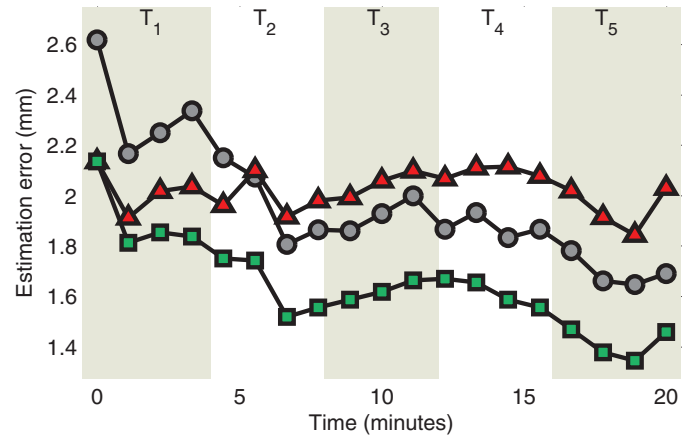
In Fig. 8 we show the resulting motion estimation error curves for all of the volunteers during a synthetic application phase. The evolution of the errors is shown over the duration of the application phase, which is the time it would take to acquire and add the remaining slices in a real scenario. Here we assumed an acquisition frequency of one slice per second which corresponds to a heart rate of 60 beats per minute. Each point in Fig. 8 represents the mean error obtained over a time interval of  $2L$  update slices, i.e. the time taken to acquire each slice position twice.

In order to quantitatively evaluate the 3D motion estimation errors and the adaptivity of the compared techniques, we split the application phase into 5 time periods  $T_1, \dots, T_5$ , of equal length. Those are highlighted in Fig. 8. The mean 3D motion estimation errors in the corresponding time intervals using the three methods for all 6 volunteers can be found in Table 1 in the supplementary materials.

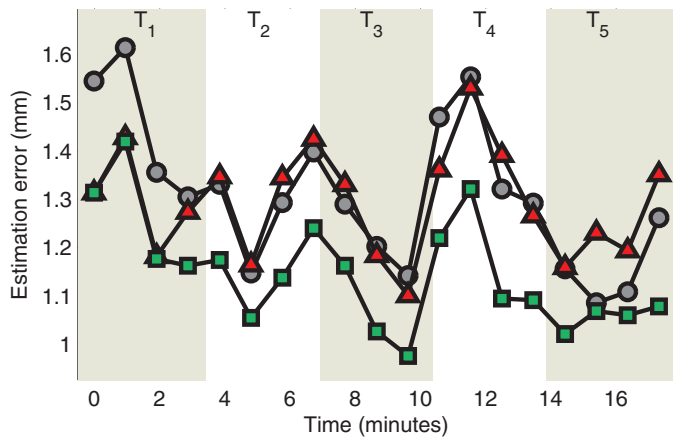
For all volunteers the AAMM technique significantly ( $p < 0.01$ ) outperformed the other two methods in all of the intervals as can be seen by comparing to the error curves shown in Fig. 8 and the figures in Table 1 in the supplementary materials. Significance was assessed using a 1-tailed Wilcoxon signed rank test since the error distributions were generally not symmetric. The estimation errors for AAMM and its non-adaptive counterpart, AAMM (no adapt.), were similar in the beginning of the application phase, but as anticipated, as the application phase went on, the AAMM technique continually improved its accuracy by incorporating more and more data into the model. On average the motion estimation of AAMM improved by 22.94% in  $T_5$  with respect to its non-adaptive counterpart. However, the method has already significantly adapted to



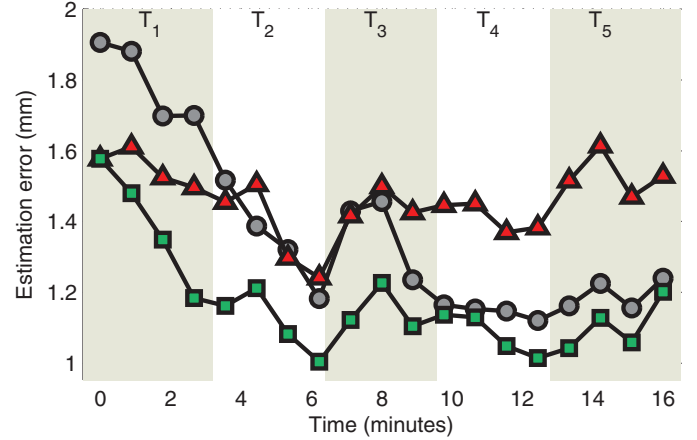
(a) Volunteer A



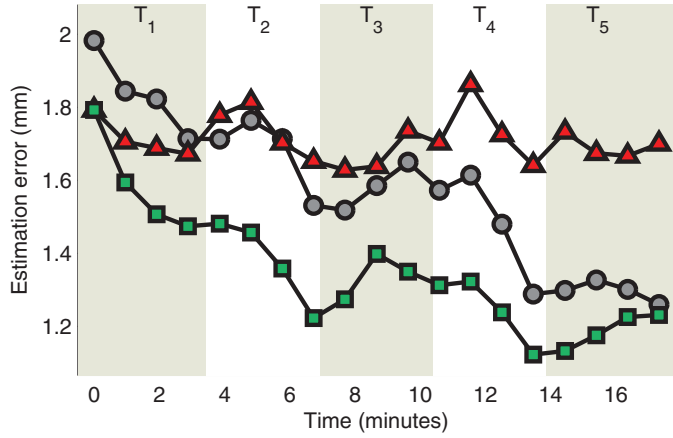
(b) Volunteer B



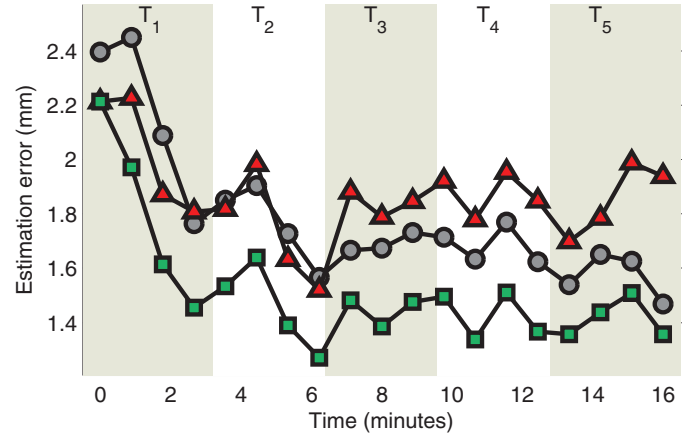
(c) Volunteer C



(d) Volunteer D



(e) Volunteer E



(f) Volunteer F

**Fig. 8.** Results of synthetic validation on normal breathing data. The figure shows average 3D motion estimation errors in mm for all volunteers over the entire duration of the synthetic application phase.

the breathing pattern in  $T_2$ , i.e. after between 3 and 7 min of imaging, where motion estimations were on average 16.87% more accurate than at the beginning of the adaptation phase. By visually inspecting the curves for AAMM in Fig. 8 it can be seen that for many volunteers (in particular volunteers A, D, E, and F) the error curves start to flatten approximately around the 7 min mark. From this it can be concluded that a longer calibration scan of around 12 min would be optimal, that is the 5 min that

were used for calibration in this experiment plus 7 min worth of data added during the application phase. Note that this time could be significantly reduced if a non-cardiac-gated sequence was used.

The AAMM technique also consistently performed better than SGA, i.e. the version without coronal slices. This shows that the addition of data from a coronal slice position in the manifold alignment step improves the 3D motion estimation accuracy.



A fraction of the remaining errors was due to the fact that the technique currently cannot estimate L-R motion. In this experiment on synthetic data the L-R motion was responsible for on average 46.53% of the remaining motion estimation error of the AAMM technique, or on average 0.63 mm. The motion estimation error did not vary depending on the position of the surrogate slice.

Note that the error curves did not necessarily steadily decrease over the entire period of time, but exhibited some variations usually affecting all methods equally. See for example  $T_3$  and  $T_4$  of volunteer C (Fig. 8c). The motion estimation error tends to be smaller for exhale motion states than for inhale motion states, since the motions involved are smaller. The variations in the error can be explained by differences in the frequency of occurrence of exhale or inhale states. For example, there were a large amount of inhale motion states around the 11 min mark of volunteer C, and a large amount of exhale states around the 9 min mark. The variations in the error between volunteers can be explained by the fact that the synthetic data was derived from real volunteers scans and some volunteers naturally had larger or more complicated motion patterns.

#### 4.2.3. Experiment 2: synthetic adaptivity to a new breathing pattern

In Experiment 1, we investigated how the autoadaptive technique behaves if more data of the same breathing pattern is added. However, the data used in that experiment does not reflect any of the long term changes which may occur in real data, such as drift or changes in breathing mode. In order to investigate if the model can adapt to previously unseen breathing patterns, for this experiment, a second synthetic dataset was generated using a 50 s dynamic 3D MR scan performed under deep breathing as input as was described in Section 4.2.1. That scan was performed immediately after the 50 s free breathing scan, but the volunteers were instructed to take deep quiet breaths. For all volunteers this resulted in synthetic data with significantly longer respiratory cycles and significantly larger displacements of the anatomy. The average magnitude of the motion varied significantly between volunteers. On average, L-R motion accounted for 17.59% of the total motion, and the A-P and S-I accounted for 23.76% and 58.65%, respectively.

In order to investigate how the examined methods would react to this new deep breathing pattern, in a first step the models were calibrated and formed by using all time points of the normal breathing data. This means that the models had largely adapted to the normal breathing pattern. Note that the state of the models was the same as for the last time point of the AAMM technique in Fig. 8. In a next step, the motion models were applied using the synthetic deep breathing data. That is, the 2D deep breathing motion fields were added to the model one-by-one, and the motion estimation error was evaluated exactly as in Experiment 1. The resulting error curves are shown in Fig. 9, where each point corresponds to an average over 2L motion estimates. In order to assess the performance of the models quantitatively, the errors for each subject were averaged within 5 time intervals of equal length. The quantitative error figures for all volunteers are given in Table 2 in the supplementary materials.

As before the AAMM method significantly outperformed the two other techniques for all volunteers and for all time intervals. As expected, the estimation errors for AAMM and its non-adaptive counterpart AAMM (no adapt.) started at similar values in  $T_1$ , but AAMM led to improved motion estimates the more data of the new breathing type was added to the model. Already in time-interval  $T_2$ , AAMM led to significant average improvements of 21.45% over AAMM (no adapt.). In  $T_5$ , the average improvements amounted to 27.10%. As before AAMM also performed significantly better than the SGA technique, which is due to the additional ro-

bustness added by the coronal slice employed in the AAMM technique.

For all examined methods the motion estimation errors were significantly larger for the deep breathing pattern than for the normal breathing pattern in Experiment 1. This is due to the fact that the deep breathing data contained much larger motion amplitudes. The variations between the subjects are due to the fact that the extent of the deep breathing motion varied from volunteer to volunteer. The average error over all subjects due to the missing motion estimates in the L-R direction amounted to 42.15% of the motion estimation error of AAMM, or 2.28 mm. As before, the motion estimation error did not vary depending on the position of the surrogate slice.

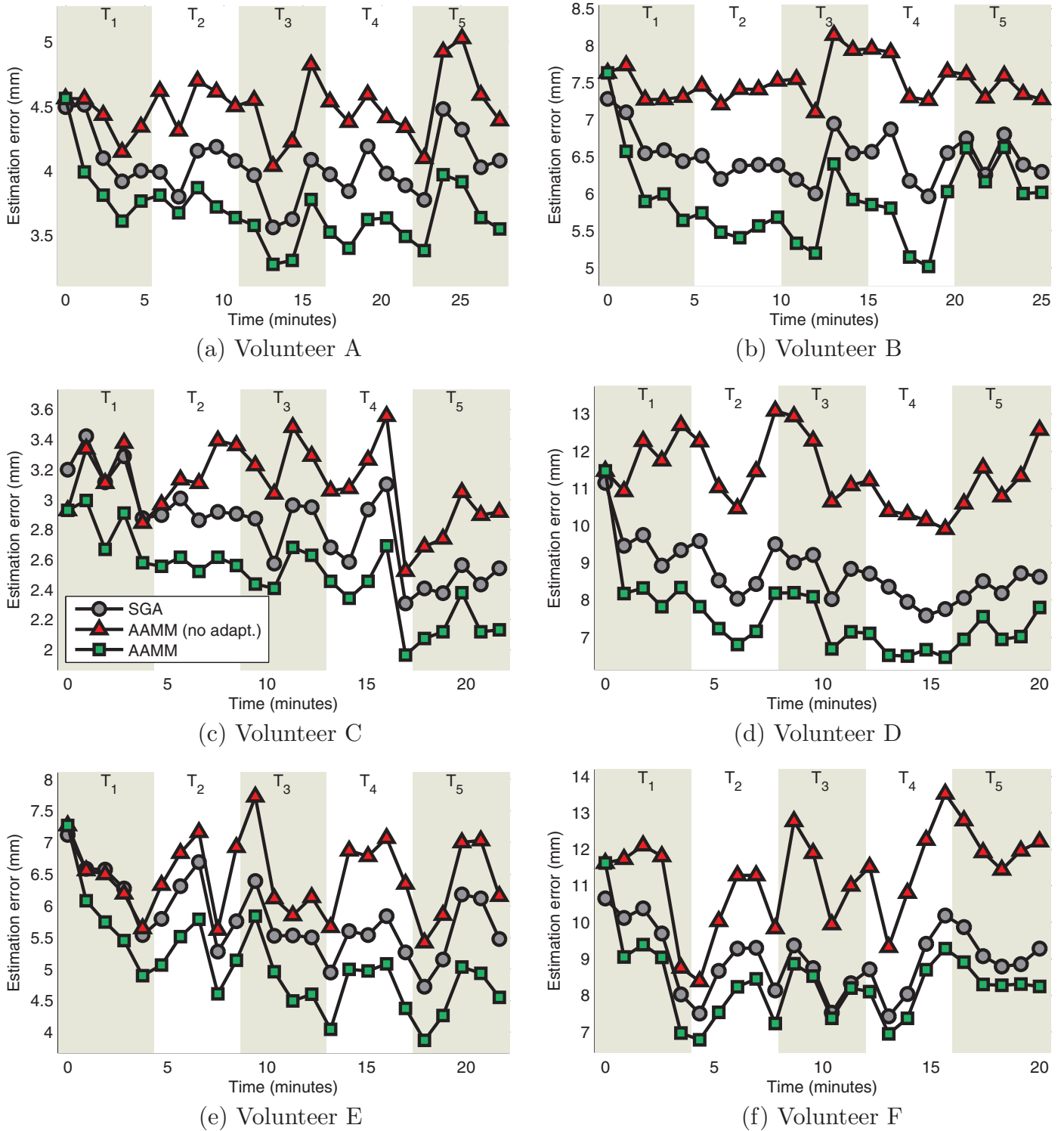
#### 4.3. Experiment 3: adaptivity on real data

For the experiments on real data we acquired real dynamic slice-by-slice data and a slice-by-slice breath hold volume as described in Section 3.1. In order to validate the model we acquired the data for the calibration and model formation, and for the model application in one long scan. Overall, we acquired each slice position 40 times which typically resulted in an approximately 20 min scan. Additionally, we acquired a 1D pencil beam navigator signal from the left hemi-diaphragm immediately before the acquisition of each 2D slice, which we used to validate the accuracy of the motion estimations, but not for any part of the motion modelling framework.

As in Experiment 1, we formed the three models on the motion fields derived from the first 10 slices acquired from each slice position. During the model application phase we then added the remainder of the slices one by one and estimated a 3D motion field for each of the input slices. Note that according to our findings in Experiment 1, ideally the model should be trained with 12 min of slice-by-slice data, or 24 slices per slice position in order to guarantee that the models have been trained to convergence. However, since this would not leave enough data to adequately study the adaptivity, we chose to use only the first 10 slices of each slice position and leave 30 slices per slice position to investigate the methods' behaviour during the model application phase. However, as a consequence we are underestimating the accuracy of the non-adaptive model AAMM (no adapt.).

Because, for the real data, no ground truth motion was available, we instead transformed the slice-by-slice breath hold volume using each estimated 3D motion field and extracted a 1D navigator value from a rectangular region of interest on the dome of the left hemi-diaphragm, i.e. approximately the same location from which the real pencil beam navigator was acquired. Note that for both the pencil beam navigator and the signal estimated from the reconstructed volumes, the displacements in millimetres are known, however, the two signals are offset by an unknown value from each other. For visualisation of the navigator curves in Fig. 11, we corrected the curves estimated by AAMM and pencil beam navigator curves for this shift by subtracting the mean of each from itself. For the quantitative evaluation, we chose to report the normalised cross correlation (NCC) between the signals because this measure is invariant to such offsets. For perfect motion estimation the extracted navigator signal should be strongly correlated with the pencil beam navigator. In reality, however, this correlation depends on the accuracy of the estimated 3D motion.

In order to quantitatively assess the adaptivity of the compared methods we measured the NCC of the estimated navigator signal with the pencil beam navigator over fixed intervals. In Fig. 10 we show the progression of this correlation for all four volunteers. For a robust estimation of the NCC we chose intervals of

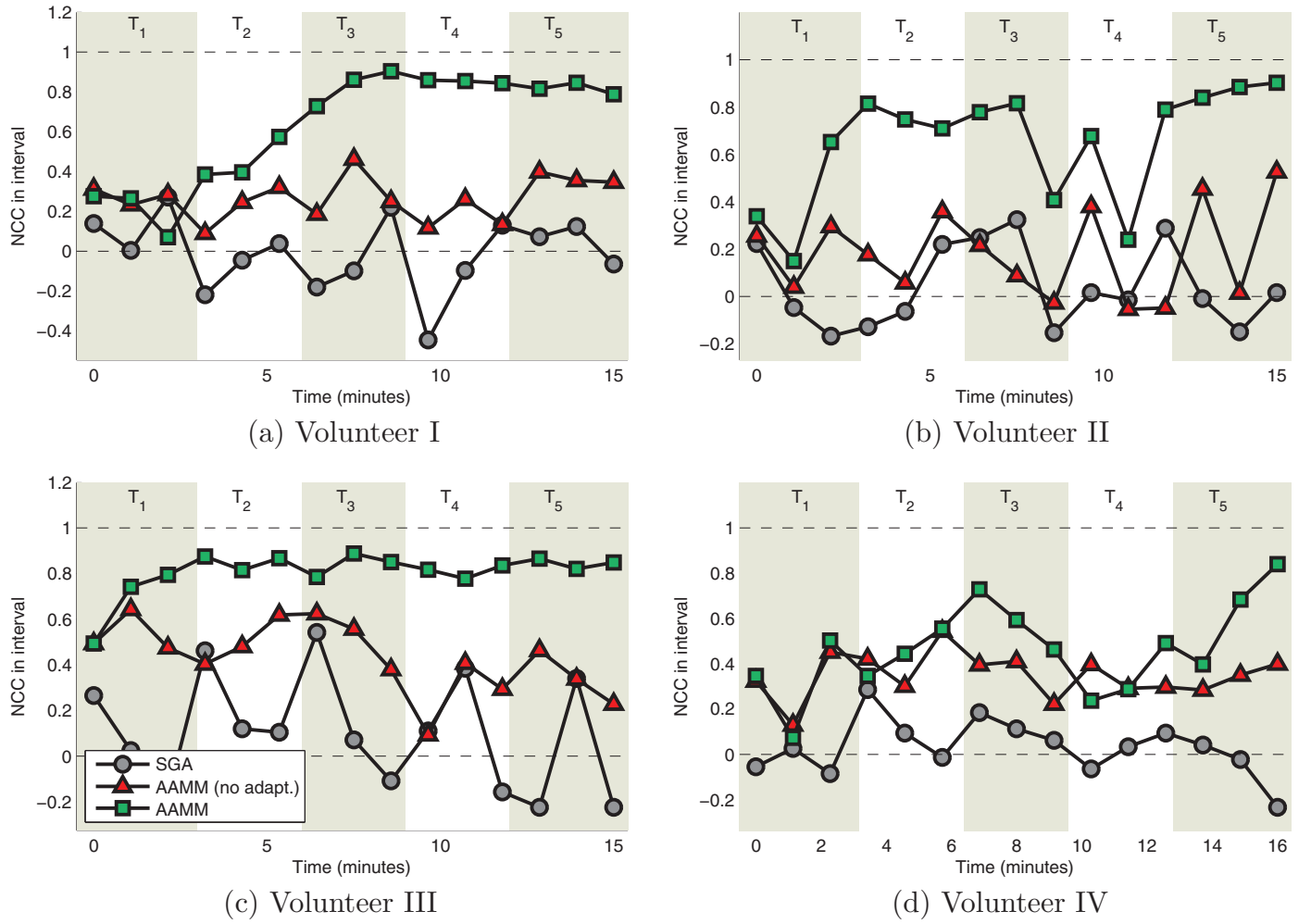


**Fig. 9.** Average 3D motion estimation errors in mm for all volunteers over the entire duration of the synthetic application phase when the breathing type was changed to deep breathing. The entire interval shown contains only deep breathing.

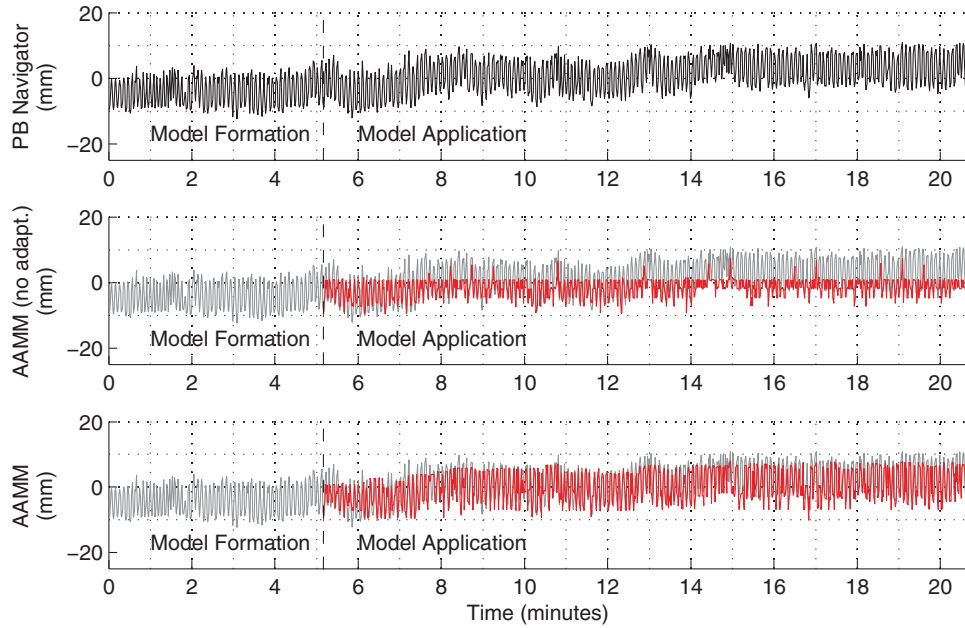
twice the number of slice positions, i.e.  $2L$ , to calculate each error point.

As for the synthetic data we divided the entire application phase into 5 larger time intervals  $T_1, \dots, T_5$ . Table 3 in the supplemental materials contains the NCC between the pencil beam navigator and the retrospectively derived navigator signal for all volunteers over the entire duration of these periods.

AAMM outperformed the other two methods for most time intervals. Furthermore, as for the synthetic data, it could again be observed that the motion estimation accuracy, as measured by NCC, improved over the duration of the application phase. Note that the data in this experiment was derived from a relatively long scan, where natural changes in respiration patterns are very likely to happen due to relaxation or, occasionally, due to the volunteer falling asleep in the scanner. We observed that motion estima-



**Fig. 10.** NCC of pencil beam navigator with a navigator signal derived from volume deformed using a 3D motion estimation provided by each of the three investigated methods over time.



**Fig. 11.** Example of the pencil beam navigator signal acquired for validation for volunteer III (top row) and navigator estimations produced by AAMM (no adapt.) (middle row) and AAMM (bottom row). The original pencil beam navigator is underlaid in grey for comparison. We show the entire time interval including the approximately 5 min of the calibration scan, and the approximately 15 min of model application. All signals have been normalised by subtracting the mean signal value such that the pencil beam navigator signal and the estimated signals can be visually compared.

tion accuracy sometimes dropped due to such changes. For example, volunteer II (see Fig. 10b) started taking deep breaths around  $T_3$ , but then returned to his previous breathing pattern. Had he continued breathing deeply, presumably our model would have adapted to that pattern. Note that for volunteer III the NCC quickly approaches its maximum for AAMM, but continually decreases for its non-adaptive counterpart. AAMM manages to maintain the motion estimation accuracy for the remainder of the session. By examining the original pencil beam navigator signal shown in the top row of Fig. 11, it can be seen that the subject exhibited a significant drift of close to 10 mm in the respiration base level throughout the imaging session. By comparing this signal to the estimated signals by AAMM and AAMM (no adapt.) it can be observed that AAMM manages to follow this drift whilst AAMM (no adapt.) cannot adapt its range of motion predictions.

In contrast to the synthetic experiments, in the experiments on real data, SGA consistently performed worse than the other examined methods. We found that on real data SGA was fundamentally not robust to the sagittal input slices. We observed that respiratory information often failed to propagate through the body middle, that is for example, an input slice from the left body half would often fail to properly estimate motion in the right body half and vice versa. Incorporating data from coronal slice positions in the manifold alignment step of AAMM effectively solved this problem.

## 5. Discussion

We have proposed a novel motion modelling framework which enables accurate 3D motion estimations over extended periods of time in the scenario of MR-guided interventions. This is achieved by using partial motion information, i.e. 2D motion fields estimated from MR slices, to form as well as to apply the motion model. In contrast to 3D MR imaging, 2D MR slices can be acquired close to real-time and offer better in-plane image quality. By acquiring 2D MR data from variable imaging planes we are able to combine the image quality of 2D MR with the coverage of 3D MR. The fact that the calibration and surrogate data are of the same type inherently enables the proposed motion model to automatically adapt to changing breathing patterns without the need to rebuild the model during the application phase.

The vast majority of motion models in the literature cannot adapt to changing breathing patterns and need to be rebuilt entirely if the correlation between the surrogate and the motion data loses validity (McClelland et al., 2013). A small number of papers such as Schweikard et al. (2005) and Cho et al. (2010) proposed adaptive techniques, which have the ability to update the model using intra-fractional imaging. These models can only update the model intermittently whereas the proposed AAMM framework can make use of all the surrogate data acquired to update the model continuously. Adaptive approaches not requiring intra-fractional imaging such as Fassi et al. (2014) can also continually adapt, but, nevertheless, may lose validity if the change in respiratory pattern is more complicated than a simple drift, such as, for example, a change to a new breathing pattern.

We implemented the autoadaptive motion model by extending our previously proposed simultaneous groupwise manifold alignment (SGA) technique to use 2D motion fields as input. This approach has two important limitations: Through-plane motion may distort the motion estimations and motion in the direction orthogonal to the slices cannot be estimated. Park et al. (2012) found that liver tumour motion is smallest in the L-R direction with a magnitude of 3.0 mm on average. In comparison, the average motions in the S-I and A-P directions amount to 17.9 mm and 5.1 mm, respectively. Similarly, Seppenwoolde et al. (2002) found that the S-I motion of lung tumours was 12 mm on average in the lower lobes,

while A-P and L-R motion was 2.2 mm and 1.2 mm on average. Hence, in order to minimise the effects of through-plane motion we derive the motion from sagittal input slices. The large differences in the appearance of sagittal slices acquired from different locations necessitated the incorporation of coronal images from a single slice position. In order to combine sagittal and coronal data we substantially expanded the methodology of SGA to arrive at the proposed AAMM technique.

We demonstrated a proof-of-principle of our proposed motion modelling framework and validated it on realistic synthetic and real data. Our experiments show that the autoadaptive motion model is able to adapt to novel breathing patterns and can thus produce significantly better 3D motion estimations over the duration of an MR-guided treatment compared to its non-adaptive counterpart. Furthermore, the experiments show that the incorporation of data from a single coronal slice position leads to significant improvements in motion estimation. Note that we did not compare the performance of AAMM to traditional motion models from the literature. Our proposed method follows a new paradigm using 2D MR data in all stages of the model, which is conceptually different from the classical motion model paradigm. Hence, it was not possible to compare against existing techniques using the same data. In the synthetic experiments we could have trained a classical motion model on the 3D ground truth motion fields. However, a comparison to such a model would not have been on equal terms because the synthetic 2D motion fields were derived through an additional registration step.

The proposed technique offers a novel way of performing adaptive motion modelling, however, the method has only been evaluated on healthy volunteers and there are a number of challenges which need to be addressed before the technique will be ready for use in a clinical system. In the following we will discuss a number of limitations and possible extensions of the technique.

The current implementation of the proposed autoadaptive framework suffers from significant latency. In the current system, for each update, the following steps need to be performed: 2D MR image acquisitions and reconstruction ( $\sim 200$  ms), 2D registration ( $\sim 500$  ms) and the groupwise embedding and lookup ( $\sim 100$  ms). Consequently, the motion modelling system in its present form has a latency of around 800 ms, which would be unacceptable in a clinical scenario and would increase the motion prediction errors. However, we would like to stress that the focus of this paper is not an efficient implementation but rather a proof-of-principle of autoadaptive motion modelling. The large latency is not an inherent drawback of our proposed method. Rather, it is a result of the acquisition sequence and computational techniques used in this work. In the context of MR-guided interventions, Ries et al. (2010) have demonstrated that motion estimations can be obtained from 2D MR images with latencies less than 114 ms, and De Senneville et al. (2015) has proposed a framework which can provide motion estimations from 2D MR data with a latency of only 80 ms. Significant speed improvements could also be achieved for the groupwise manifold alignment with a more efficient parallel implementation, as the groups containing the aligned manifold embeddings can be processed independently of each other. In this manner it would be possible to reduce the computational times of the manifold alignment to just a few milliseconds by using a parallel implementation and a modern GPU (graphics processing unit) or multi-core work-station. It is therefore believed that an optimised version of the proposed autoadaptive motion modelling system may be able to run with update latencies close to 100 ms. The remaining latency could be addressed by combining the method with a motion prediction technique such as Kalman filtering (Sharp et al., 2004; Ries et al., 2010).

In the present work the update frequency is limited by the cardiac gating of the images to around 1 Hz, based on a typical heart



rate of 60 beats per minute. The reason cardiac gated images were employed was to isolate the respiratory motion for this study. The cardiac gating, however, is not an essential part of the technique and could be easily dropped if the region of interest excluded the heart as in the scenario of a MR-guided HIFU of the liver. For example, De Senneville et al. (2015) acquired 2D MR slices of the liver with a imaging frame-rate of 10 Hz. In the experiments in this paper it was found that the proposed system can significantly improve the motion estimation accuracy by close to 20% in less than 10 min. Potentially, however, much faster adaptivity could be achieved. Based on a hypothetical update frame-rate of 10 Hz, the same improvements in the motion estimation in the whole thorax could be achieved in as little as 1 min.

Currently the motion is estimated only from sagittal 2D MR slices. This orientation was chosen to minimise the effects of through-plane motion. However, the remaining L-R motion may cause artefacts in the registration step of the calibration phase. Furthermore, the 3D motion estimations are obtained by simply stacking the 2D motion fields, which leads to motion fields lacking the L-R component. In the evaluation on synthetic data the missing L-R component accounted for over 40% of the remaining motion estimation error. This is a significant drawback of the presented method and extending it to account also for through-plane motion will be a priority in future work. Note that currently the coronal motion fields are used only in the groupwise manifold alignment step but are not part of the final 3D motion estimations. It may be possible to mitigate the through-plane motion effects by also using motion information derived from one or potentially several coronal slice positions in the 3D motion estimation step.

Lastly, in the autoadaptive motion model in its present form all the data added is retained. The rationale behind this is that, in this manner, the model can go back to breathing patterns which were observed before a change occurred. A patient may, for example, go back and forth between a calm and a nervous breathing pattern as a result of certain actions of the surgeon or the progress of the treatment. However, the larger the model grows the more memory is used to store the 2D motion fields and the more computationally expensive it becomes to evaluate the updated group embeddings. It may therefore make sense to implement a “ring buffer” approach, where older data is discarded as new data is added to the model. An interesting future direction would be to automatically determine which data is essential to model certain breathing types and selectively delete data which is unlikely to be used again.

## 6. Conclusion

Modelling respiratory motion from MR data may provide a solution for correcting MR-guided treatments for respiratory motion. In particular it can provide intra-procedure 3D motion estimations in MR-guided interventions such as MRg-HIFU or MRg-RT. However, such treatments are typically performed in a time frame in which respiratory motion patterns are known to change, causing conventional motion models to lose their validity. This work demonstrates a proof-of-principle for a novel autoadaptive motion modelling framework which is calibrated and applied using the same type of data, i.e. 2D MR slices acquired from variable imaging planes. This allows the proposed motion model to continually adapt every time a new 2D slice is acquired and used to estimate 3D motion. A number of challenges must be addressed before the method can be applied in a clinical setting, in particular the long calibration time and large latency of 800 ms. Nevertheless, our novel motion modelling paradigm provides an important stepping stone which may allow lengthy MR-guided treatments to go on uninterrupted whilst the model continually maintains its ability to provide accu-

rate, up-to-date 3D motion estimations, despite changing breathing patterns.

## Data download

The real 2D MR data acquired for the evaluation of our proposed framework in Section 4.3 are freely available and can be downloaded from <https://zenodo.org/record/55345> under the Creative Commons Attribution license.

## Acknowledgements

This work was funded by EPSRC programme grant EP/H046410/1. This research was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.media.2016.06.005](https://doi.org/10.1016/j.media.2016.06.005)

## References

- Arnold, P., Preiswerk, F., Fasel, B., Salomir, R., Scheffler, K., Cattin, P., 2011. 3D organ motion prediction for MR-guided high intensity focused ultrasound. In: Proc. MICCAI. Springer, pp. 623–630.
- Baumgartner, C., Kolbitsch, C., Balfour, D., Marsden, P., McClelland, J., Rueckert, D., King, A., 2014a. High-resolution dynamic MR imaging of the thorax for respiratory motion correction of PET using groupwise manifold alignment. *Med. Image Anal.* 18 (7), 939–952.
- Baumgartner, C., Kolbitsch, C., McClelland, J., Rueckert, D., King, A., 2013. Groupwise simultaneous manifold alignment for high-resolution dynamic MR imaging of respiratory motion. In: Proc. IPMI. Springer, pp. 232–243.
- Baumgartner, C., Kolbitsch, C., McClelland, J., Rueckert, D., King, A., 2014b. Autoadaptive motion modelling. In: Proc. IEEE ISBI, pp. 457–460.
- Bhatia, K., Rao, A., Price, A., Wolz, R., Hajnal, J., Rueckert, D., 2012. Hierarchical manifold learning. In: Proc. MICCAI. Springer, pp. 512–519.
- Blackall, J., Ahmad, S., Miquel, M., McClelland, J., Landau, D., Hawkes, D., 2006. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys. Med. Biol.* 51, 4147–4169.
- Brix, L., Ringgaard, S., Sørensen, T., Poulsen, P., 2014. Three-dimensional liver motion tracking using real-time two-dimensional MRI. *Med. Phys.* 41 (4), 042302.
- Cho, B., Poulsen, P., Keall, P., 2010. Real-time tumor tracking using sequential kv imaging combined with respiratory monitoring: a general framework applicable to commonly used IGRT systems. *Phys. Med. Biol.* 55 (12), 3299.
- Crijns, S., Raaymakers, B., Lagendijk, J., 2012. Proof of concept of MRI-guided tracked radiation delivery: tracking one-dimensional motion. *Phys. Med. Biol.* 57 (23), 7863.
- De Senneville, B., El Hamidi, A., Moonen, C., 2015. A direct PCA-based approach for real-time description of physiological organ deformations. *IEEE Trans. Med. Imaging.* 34 (4), 974–982.
- Dikaio, N., Izquierdo-Garcia, D., Graves, M., Mani, V., Fayad, Z., Fryer, T., 2012. MRI-based motion correction of thoracic PET: initial comparison of acquisition protocols and correction strategies suitable for simultaneous PET/MRI systems. *Eur. radiol.* 22 (2), 439–446.
- Fassi, A., Schaefer, J., Fernandes, M., Riboldi, M., Sarrut, D., Baroni, G., 2014. Tumor tracking method based on a deformable 4D CT breathing motion model driven by an external surface surrogate. *Int. J. Radiat. Oncol.* 88 (1), 182–188.
- Fischer, P., Pohl, T., Hornegger, J., 2014. Real-time respiratory signal extraction from X-ray sequences using incremental manifold learning. In: Proc. IEEE ISBI, pp. 915–918.
- Foley, J., Eames, M., Snell, J., Hananel, A., Kassell, N., Aubry, J.-F., 2013. Image-guided focused ultrasound: state of the technology and the challenges that lie ahead. *Imaging Med.* 5 (4), 357–370.
- Georg, M., Souvenier, R., Hope, A., Pless, R., 2008. Manifold learning for 4D CT reconstruction of the lung. In: Proc. IEEE CVPRW, pp. 1–8.
- Hoogeman, M., Prévost, J.-B., Nuytens, J., Pöll, J., Levendag, P., Heijmen, B., 2009. Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. *Int. J. Radiat. Oncol.* 74 (1), 297–303.
- Hynynen, K., Freund, W., Cline, H., Chung, A., Watkins, R., Vetro, J., Jolesz, F., 1996. A clinical, noninvasive, MR imaging-monitored ultrasound surgery method. *Radiographics* 16 (1), 185–195.
- Isaksson, M., Jaldén, J., Murphy, M., 2005. On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications. *Med. Phys.* 32 (12), 3801–3809.

- King, A., Buerger, C., Tsoumpas, C., Marsden, P., Schaeffter, T., 2012. Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. *Med. Image Anal.* 16 (1), 252–264.
- King, A., Rhode, K., Razavi, R., Schaeffter, T., 2009. An adaptive and predictive respiratory motion model for image-guided interventions: theory and first clinical application. *IEEE Trans. Med. Imaging* 28 (12), 2020–2032.
- Kini, V., Vedam, S., Keall, P., Patil, S., Chen, C., Mohan, R., 2003. Patient training in respiratory-gated radiotherapy. *Med. Dosim.* 28 (1), 7–11.
- Kuhn, H., 1955. The Hungarian method for the assignment problem. *Nav. Res. Logist. Q.* 2 (1–2), 83–97.
- Low, D., Parikh, P., Lu, W., Dempsey, J., Wahab, S., Hubenschmidt, J., Nystrom, M., Handoko, M., Bradley, J., 2005. Novel breathing motion model for radiotherapy. *Int. J. Radiat. Oncol.* 63 (3), 921–929.
- McClelland, J., Hawkes, D., Schaeffter, T., King, A., 2013. Respiratory motion models: a review. *Med. Image Anal.* 17 (1), 19–42.
- McClelland, J., Hughes, S., Modat, M., Qureshi, A., Ahmad, S., Landau, D., Ourselin, S., Hawkes, D., 2011. Inter-fraction variations in respiratory motion models. *Phys. Med. Biol.* 56 (1), 251–272.
- Modat, M., Ridgway, G., Taylor, Z., Lehmann, M., Barnes, J., Hawkes, D., Fox, N., Ourselin, S., 2010. Fast free-form deformation using graphics processing units. *Comput. Meth. Prog. Bio.* 98 (3), 278–284.
- Park, J., Park, S., Kim, J., Yoon, S., Song, S., Liu, Z., Song, B., Kauwelo, K., Webster, M., Sandhu, A., et al., 2012. Liver motion during cone beam computed tomography guided stereotactic body radiation therapy. *Med. Phys.* 39 (10), 6431–6442.
- Peressutti, D., Penney, G., Housden, R., Kolbitsch, C., Gomez, A., Rijkhorst, E.-J., Barratt, D., Rhode, K., King, A., 2013. A novel bayesian respiratory motion model to estimate and resolve uncertainty in image-guided cardiac interventions. *Med. Image Anal.* 17 (4), 488–502.
- Peressutti, D., Rijkhorst, E., Barratt, D., Penney, G., King, A., 2012. Estimating and resolving uncertainty in cardiac respiratory motion modelling. In: *Proc. IEEE ISBI*, pp. 262–265.
- Raaymakers, B., Lagendijk, J., Overweg, J., Kok, J., Raaijmakers, A., Kerkhof, E., van der Put, R., Meijnsing, I., Crijns, S., Benedosso, F., et al., 2009. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. *Phys. Med. Biol.* 54 (12), N229.
- Rao, A., Sanchez-Ortiz, G., Chandrashekar, R., Lorenzo-Valdés, M., Mohiaddin, R., Rueckert, D., 2002. Comparison of cardiac motion across subjects using non-rigid registration. In: *Proc. MICCAI*. Springer, pp. 722–729.
- Ries, M., De Senneville, B., Roujol, S., Berber, Y., Quesson, B., Moonen, C., 2010. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magn. Reson. Med.* 64 (6), 1704–1712.
- Rijkhorst, E.-J., Heanes, D., Odille, F., Hawkes, D., Barratt, D., 2010. Simulating dynamic ultrasound using MR-derived motion models to assess respiratory synchronisation for image-guided liver interventions. In: *Proc. IPCAI*. Springer, pp. 113–123.
- Rijkhorst, E.-J., Rivens, I., Haar, G., Hawkes, D., Barratt, D., 2011. Effects of respiratory liver motion on heating for gated and model-based motion-compensated high-intensity focused ultrasound ablation. In: *Proc. MICCAI*. Springer-Verlag, pp. 605–612.
- Rosenblatt, M., 1956. Remarks on some nonparametric estimates of a density function. *Ann. Math. Stat.* 27 (3), 832–837.
- Roweis, S., Saul, L., 2000. Nonlinear dimensionality reduction by locally linear embedding. *Science* 290 (5500), 2323–2326.
- Savill, F., Schaeffter, T., King, A., 2011. Assessment of input signal positioning for cardiac respiratory motion models during different breathing patterns. In: *Proc. ISBI*, pp. 1698–1701.
- Sawant, A., Smith, R., Venkat, R., Santanam, L., Cho, B., Poulsen, P., Cattell, H., Newell, L., Parikh, P., Keall, P., 2009. Toward submillimeter accuracy in the management of intrafraction motion: the integration of real-time internal position monitoring and multileaf collimator target tracking. *Int. J. Radiat. Oncol.* 74 (2), 575–582.
- Schweikard, A., Glosser, G., Bodduluri, M., Murphy, M., Adler, J., 2000. Robotic motion compensation for respiratory movement during radiosurgery. *Comput. Aided Surg.* 5 (4), 263–277.
- Schweikard, A., Shiomi, H., Adler, J., 2005. Respiration tracking in radiosurgery without fiducials. *Int. J. Med. Rob. Comp.* 1 (2), 19–27.
- Seppenwoolde, Y., Berbeco, R., Nishioka, S., Shirato, H., Heijmen, B., 2007. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Med. Phys.* 34 (7), 2774–2784.
- Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., van Herk, M., Lebesque, J., Miyasaka, K., 2002. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int. J. Radiat. Oncol.* 53 (4), 822–834.
- Sharp, G.C., Jiang, S.B., Shimizu, S., Shirato, H., 2004. Prediction of respiratory tumour motion for real-time image-guided radiotherapy. *Phys. Med. Biol.* 49 (3), 425.
- Souvenir, R., Zhang, Q., Pless, R., 2006. Image manifold interpolation using free-form deformations. In: *Proc. IEEE Image Proc.*, pp. 1437–1440.
- Tempany, C., McDannold, N., Hynynen, K., Jolesz, F., 2011. Focused ultrasound surgery in oncology: overview and principles. *Radiology* 259 (1), 39–56.
- von Siebenthal, M., Székely, G., Gamper, U., Boesiger, P., Lomax, A., Cattin, P., 2007. 4D MR imaging of respiratory organ motion and its variability. *Phys. Med. Biol.* 52 (6), 1547–1564.
- Wachinger, C., Yigitsoy, M., Rijkhorst, E., Navab, N., 2011. Manifold learning for image-based breathing gating in ultrasound and MRI. *Med. Image Anal.* 16 (4), 806–818.
- Würslin, C., Schmidt, H., Martirosian, P., Brendle, C., Boss, A., Schwenzer, N., Stegger, L., 2013. Respiratory motion correction in oncologic PET using T1-weighted MR imaging on a simultaneous whole-body PET/MR system. *J. Nucl. Med.* 54 (3), 464–471.
- Zachiu, C., de Senneville, B., Moonen, C., Ries, M., 2015. A framework for the correction of slow physiological drifts during MR-guided HIFU therapies: proof of concept. *Med. Phys.* 42 (7), 4137–4148.