*DOI:*
10.1016/j.media.2011.02.009

*Document Version*
Peer reviewed version

# Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation

Christian Buerger, Tobias Schaeffter, Andrew P. King

*Division of Imaging Sciences, King's College, 4th Floor Lambeth Wing, St. Thomas' Hospital, London SE1 7EH, UK*

**Abstract**

Non-rigid image registration techniques are commonly used to estimate complex tissue deformations in medical imaging. A range of non-rigid registration algorithms have been proposed, but they typically have high computational complexity. To reduce this complexity, combinations of multiple less complex deformations have been proposed such as hierarchical techniques which successively split the non-rigid registration problem into multiple locally rigid or affine components. However, to date the splitting has been regular and the underlying image content has not been considered in the splitting process. This can lead to errors and artefacts in the resulting motion fields. In this paper, we propose three novel adaptive splitting techniques, an image-based, a similarity-based, and a motion-based technique within a hierarchical framework which attempt to process regions of similar motion and/or image structure in single registration components. We evaluate our technique on free-breathing whole-chest 3D MRI data from 10 volunteers and two publicly available CT datasets. We demonstrate a reduction in registration error of up to 49.1% over a non-adaptive technique and compare our results with a commonly used free-form registration algorithm.

*Keywords:*
Non-rigid image registration, locally affine, hierarchical, adaptive, respiratory motion estimation

*Email address:* christian.buerger@kcl.ac.uk, tobias.schaeffter@kcl.ac.uk, andrew.king@kcl.ac.uk (Christian Buerger, Tobias Schaeffter, Andrew P. King)

## 1. Introduction

Respiratory motion estimation is essential for a variety of applications in medical imaging such as lung ventilation imaging (Guerrero et al., 2006), functional lung imaging (Reinhardt et al., 2007) or estimating the biomechanical properties of the lung (Sundarama and Gee, 2005). In addition, motion estimation is an important prerequisite for motion modelling approaches which have been proposed for a range of applications. For example, McClelland et al. (2006) modelled lung motion for radiotherapy treatment planning, Manke et al. (2003) applied motion correction techniques for cardiac magnetic resonance angiography, and King et al. (2009) applied a motion model derived from magnetic resonance images (MRI) for image-guided cardiac catheterizations.

Tissue motion caused by respiration is usually non-rigid, especially when considering whole-thorax deformation rather than single organ motion. Various non-rigid image-based registration techniques have been proposed and they have been widely applied to estimate tissue deformations. Elastic matching approaches model tissue motion as a deformation of an elastic material (Gee and Bajcsy, 1999) and can be described by the Navier linear partial differential equation (PDE) (Rohr, 2000). Similarly, fluid registrations model deformation as a viscous fluid, and can be described by the Navier-Stokes equation (Crum et al., 2005). Optical flow methods describe motion by assuming a constant brightness constraint of moving voxels (Horn and Schunk, 1981). Another common approach is to estimate the deformation at a set of control points followed by spline-based interpolations of a deformation field in between these points. Thin-plate splines (TPS) (Bookstein, 1989), for instance, model the deformation field as a thin metal plate which is anchored to the set of control points and bends in between. TPS, however, are computatinally expensive since each control point contributes to the global interpolation result. In comparison, B-Splines are locally controlled and consequently computationally less expensive and have been widely used, for example in the free-form deformation (FFD) algorithm described by Rueckert et al. (1999) for estimating breast deformation, or in the inverse consistent approach of Cao et al. (2009).

These techniques have proven to be successful in capturing tissue deformations but at a high computational cost. Different approaches that attempt to reduce execution times have been proposed, such as efficient linear programming (Glocker et al., 2008) or GPU-based implementations (Modat et al.,

2009). However, with potential increases in image resolution in the future, there will always be a need for algorithms that can accurately capture non-rigid motion with a lower computational complexity, as opposed to a faster implementation. One approach for reducing computational complexity is to decompose the non-rigid registration problem into multiple locally rigid or affine registration problems which subsequently are combined to form an overall non-rigid deformation. Related literature in this field includes Little et al. (1996), who first suggested the incorporation of rigid body motion in a smooth deformation field. Zhuang et al. (2008) proposed locally affine registrations of individual heart chambers in an atlas-based segmentation framework. Similarly, Commowick et al. (2008) applied locally affine registrations on pre-defined segmented brain regions. These techniques, however, are not always feasible since some degree of prior image segmentation is required.

A more generic polyrigid or polyaffine approach was proposed by Arsigny et al. (2005, 2009). This approach does not require pre-defined segmentations but is based on a distribution of anchor points, each associated with a rigid or affine registration. Martin-Fernandez et al. (2009) extended this work by incorporating rigid body structures into the framework. The registration accuracy of such techniques, however, depends on the choice of location for the anchor points.

A different approach, based on a hierarchical registration scheme, was proposed by Likar and Pernus (1999, 2001) to register skeletal muscle fibres. The image being registered was successively split into rectangular blocks, each associated with an affine registration. The initial implementation in 2D was extended to 3D and further improved by Andronache et al. (2005, 2008). Both approaches used (a) regular splitting leading to sub-blocks of identical size in the registration hierarchy and (b) TPS interpolation to determine a smooth deformation in between registered blocks.

In this paper we describe a hierarchical local affine registration scheme similar to Likar and Pernus (1999, 2001); Andronache et al. (2005, 2008). Our main contribution is to propose novel adaptive splitting techniques which attempt to divide the image being registered into regions of similar motion and/or image structure. The aims of the adaptive splitting are that (a) regions with similar deformations are processed in single rather than multiple registration blocks and (b) large deformations can be more accurately captured in the registration hierarchy. In addition, we reduce the computational cost of our algorithm by incorporating a multilevel B-Spline interpolation scheme rather than using TPS interpolation. We apply our approach to tho-

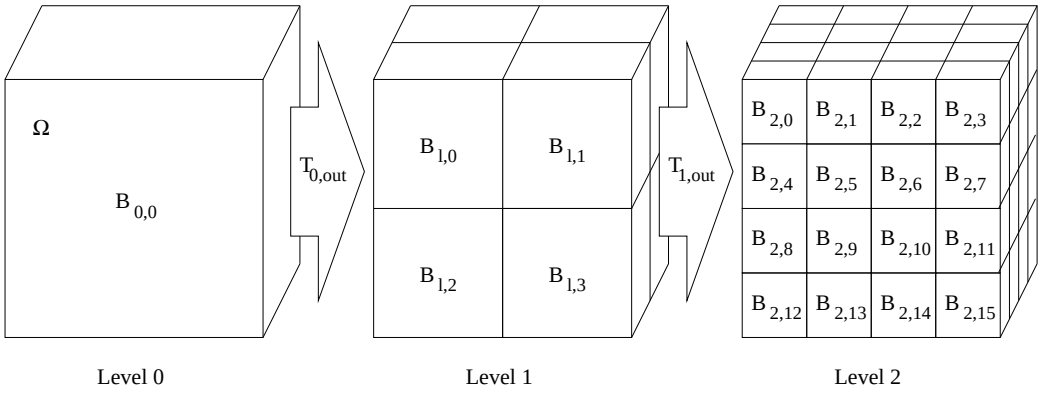



Figure 1: Hierarchical local affine registration for 3 registration levels using regular splitting to illustrate the hierarchical principle. Starting with an affine registration in $B_{0,0}$ equal to the initial region of interest $\Omega$, a hierarchical splitting scheme is applied. At each subsequent level, L, the registration result is refined by splitting each parent block $B_{L,b}$ into sub-blocks called children, in which local registrations are performed. The registration result of each level is formed by interpolating between the registrations for the sub-blocks.

racic respiratory motion estimation from free-breathing 3D MRI data and also on two publicly available CT datasets. We compare our results to a commonly used FFD algorithm based on B-Splines (Rueckert et al., 1999). We have previously presented preliminary results of our work in Buerger et al. (2010). Here, we present refinements to the adaptive splitting and interpolation schemes together with more thorough validation.

## 2. Method

### *2.1. Overview*

We register a floating image $I_F$ with a reference image $I_R$. The hierarchical nature of the local affine registration is illustrated in Fig. 1, which uses regular non-adaptive splitting to illustrate the general principle of the hierarchy. We define a number of levels within the hierarchy that determine when the registrations of local blocks will occur. Starting at level $L = 0$, within the region of interest $\Omega$, we perform an initial affine registration $A$ to capture global deformations. The resulting deformation $u_0(\vec{s})$ for each site $\vec{s} \in \Omega$ is determined by the transformation parameters of this affine registration. The overall output transformation is then given by $T_{0,out}(\vec{s}) = \vec{s} + u_0(\vec{s})$. Subsequently, $\Omega$ is split into 8 sub-blocks of equal size leading to the next

hierarchical level, $L = 1$. In each sub-block $B_{1,b}$, a local affine registration is performed using the transformation of the previous level $T_{0,out}(\vec{s})$ as a starting estimate. The resulting transformation $T_1(\vec{s})$ at level $L = 1$ is formed by interpolating between the local affine transformations. $T_1(\vec{s})$ is combined with the previous transformation $T_{0,out}(\vec{s})$ to form the overall output transformation $T_{1,out}(\vec{s})$ which is used as the starting estimate for all local affine registrations performed at the next finer level, $L = 2$. This process is repeated until the stopping conditions are met.

The following sections describe the stages involved in our hierarchical local affine registration scheme in detail. These are illustrated in Fig. 2 in which each box is labelled with the section number that describes it. Section 2.2 describes the adaptive splitting procedure. We describe the subsequent local affine registrations in Section 2.3. In Section 2.4 we describe the use of a multilevel B-Spline interpolation scheme to combine the affine transformations to produce a smooth output transformation after each level in the hierarchy. In Section 2.5 we outline our approach for avoiding folding artefacts in the motion fields. The deformations of all hierarchical levels are combined in Section 2.6. We describe the stopping conditions that terminate the registration process in Section 2.7.

### *2.2. Adaptive splitting techniques*

In the previous work of Likar and Pernus (1999, 2001) and Andronache et al. (2005, 2008), each block (*parent*) was split into sub-blocks (*children*) of equal size, i.e. a block's centre point was used as its splitting point ($\vec{SP}$). With this regular splitting strategy, however, the underlying image content is not considered, and the split may cause regions with similar underlying motion to be divided into different blocks which can cause registration artefacts at the boundaries between the blocks. Our technique attempts to avoid such artefacts by choosing splitting points based on the content of the images being registered. This adaptive strategy, however, might introduce registration blocks with large or low aspect ratios. Consequently, before splitting a block into its sub-blocks, we check the block's aspect ratio. We try to maintain approximately square shaped blocks by avoiding any splitting along an axis where the aspect ratio is greater than a threshold $\alpha$ or lower than $1/\alpha$. A threshold of $\alpha = 16 : 9$ was empirically determined.

We propose and compare three novel splitting techniques to determine the $\vec{SP}$ of each block. All three approaches define a cost term based on

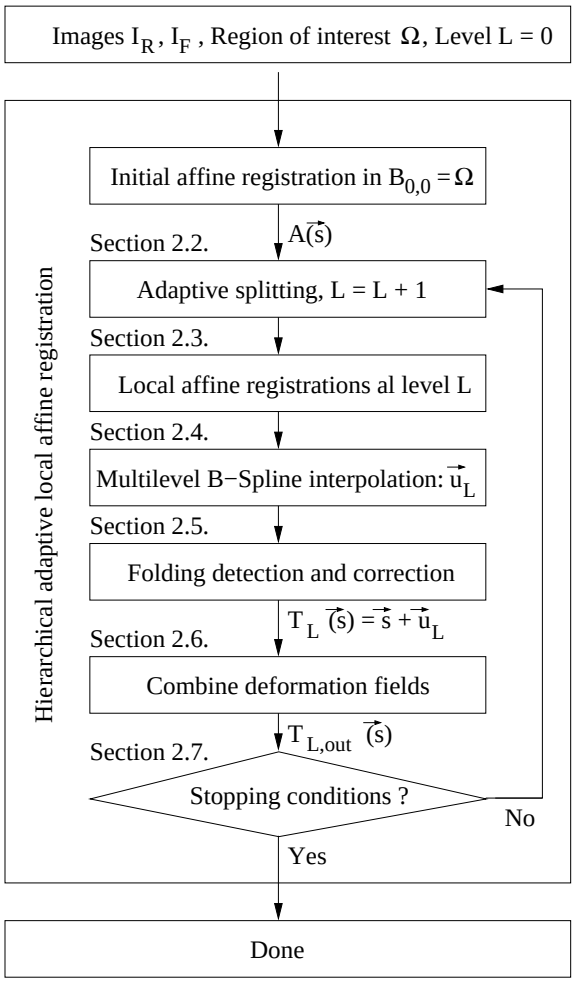



Figure 2: Flow chart of our registration hierarchy. Global deformation in the region of interest $\Omega$ is modelled by an initial affine registration $A$ at level $L = 0$. Residual local deformations are estimated by adaptively splitting $\Omega$ into sub-blocks leading to the next registration level $L = 1$. After local affine registrations in blocks $B_1 = \{B_{1,b}\}$, we apply a multilevel B-Spline approach to produce an overall smooth deformation field, $T_1$. Subsequently, we detect and correct for folding artefacts and combine the deformations of all hierarchical levels. This splitting/registration process is repeated until the stopping conditions are met.

the content of the images, then sum this cost term in a region of interest determined by the splitting point. The position of the splitting point is optimized to minimize the summed cost term. Fig. 3 illustrates this process using a simulated cost image represented by three Gaussian distributions at

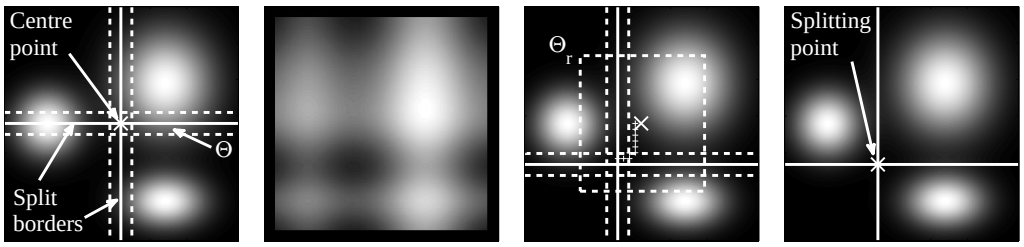



Figure 3: Adaptive splitting technique demonstrated on a simulated 2D cost term image. (a) The cost term image contains three Gaussian distributions of varying locations and standard deviations. Within a region $\Theta(\vec{SP})$ (bounded by the dotted lines) around the splitting borders located around the block's centre, the cost $C$ is computed as the sum of the underlying cost image. (b) Image showing cost $C$ when using each pixel in the image as splitting point. (c) A steepest gradient descent optimizer iteratively finds the optimal splitting point limited by the search range $\Theta_r$ by minimizing $C$. (d) The final splitting point $\vec{SP}$ represents the point at which the cost term within $\Theta(\vec{SP})$ is minimized.

different locations and with different standard deviations. The cost term is computed over a dilated region $\Theta(\vec{SP})$ around the splitting borders (three in 3D, two in 2D) initially passing through the block's centre (Fig. 3.a). We compute the cost term within $\Theta(\vec{SP})$ and optimize the position of $\vec{SP}$ to minimize the summed cost term using a steepest gradient descent optimizer. The optimization scheme is limited to the search range $\Theta_r$ (Fig. 3.c) to avoid the creation of children with a very high or low aspect ratio. Without limiting the optimization to $\Theta_r$, the splitting point might even be placed on one of the block's corner points possibly leading to indefinite registrations in that block since a child block could be identical to its parent block. Our experiments revealed that a $\Theta_r$ of 80% of the block size is an appropriate trade-off between registration accuracy and execution time.

In the following, we describe our three cost terms to determine $\vec{SP}$ based on this optimization scheme.

Our first cost term works in *image space* and considers the image content of the current transformation of the floating image, $T_{L,out}(I_F)$, and the reference image, $I_R$. This approach is motivated by avoiding splits through regions with rich structural information since image regions that are rich in structural information dominate the registration result rather than homogeneous image regions. The structural information in an image is estimated by

the sum of the magnitudes of the image gradients of $T_{L,out}(I_F)$ and $I_R$:

$$C_{L,I}(\vec{s}) = \sum_{\vec{s} \in \Theta(\vec{SP})} \left( ||\nabla I_R(\vec{s})||^2 + ||\nabla [T_{L,out}(I_F)](\vec{s})||^2 \right), \tag{1}$$

where the image gradients $\nabla T_{L,out}(I_F)$ and $\nabla I_R$ are computed by convolving $T_{L,out}(I_F)$ and $I_R$ with a pair of Sobel operators.

The second cost term to determine the $\vec{SP}$ uses information in *similarity space*. Here, we try to avoid splitting through regions with large remaining residual deformation, i.e. regions of low similarity. Consequently, a residual deformation tends to be located and processed within a single registration block rather being processed by multiple blocks. We estimate residual deformation by considering the negated similarity between $T_{L,out}(I_F)$ and $I_R$:

$$C_{L,S}(\vec{s}) = -Sim\left(I_R(\vec{s}), T_{L,out}(I_F)(\vec{s})\right), \vec{s} \in \Theta(\vec{SP}). \tag{2}$$

For example, using the negated sum of squared differences (SSD) as similarity measure leads to a cost term equal to the positive SSD within $\Theta(\vec{SP})$.

Our last cost term focuses on *motion space*. Here, we consider the deformation estimated from the previous level $L-1$ and assume that the remaining residual deformation will be large in regions with already large deformations. Therefore, we consider the deformation vector $u_{L-1}(\vec{s}) = (u_{L-1,x}, u_{L-1,y}, u_{L-1,z})^T$ estimated from the previous level and compute the squared sum of its x/y/z components

$$C_{L,M}(\vec{s}) = \sum_{\vec{s} \in \Theta(\vec{SP})} \left( u_{L-1,x}^2(\vec{s}) + u_{L-1,y}^2(\vec{s}) + u_{L-1,z}^2(\vec{s}) \right). \tag{3}$$

### *2.3. Local affine registrations at level L*

Once the splitting point has been determined, we perform local affine registrations within the resulting sub-blocks that do not meet the stopping conditions (described in Section 2.7). Similar to Rueckert et al. (1999), we include a multiresolution approach to speed up the overall registration process. We use the downsampled image data (low image resolution) at low registration levels (large blocks) and the original image data (high image resolution) at high registration levels (small blocks). Note that the choice of image resolution depends on a given registration level and not on a given block size directly. Further details will be provided in Section 3.2.

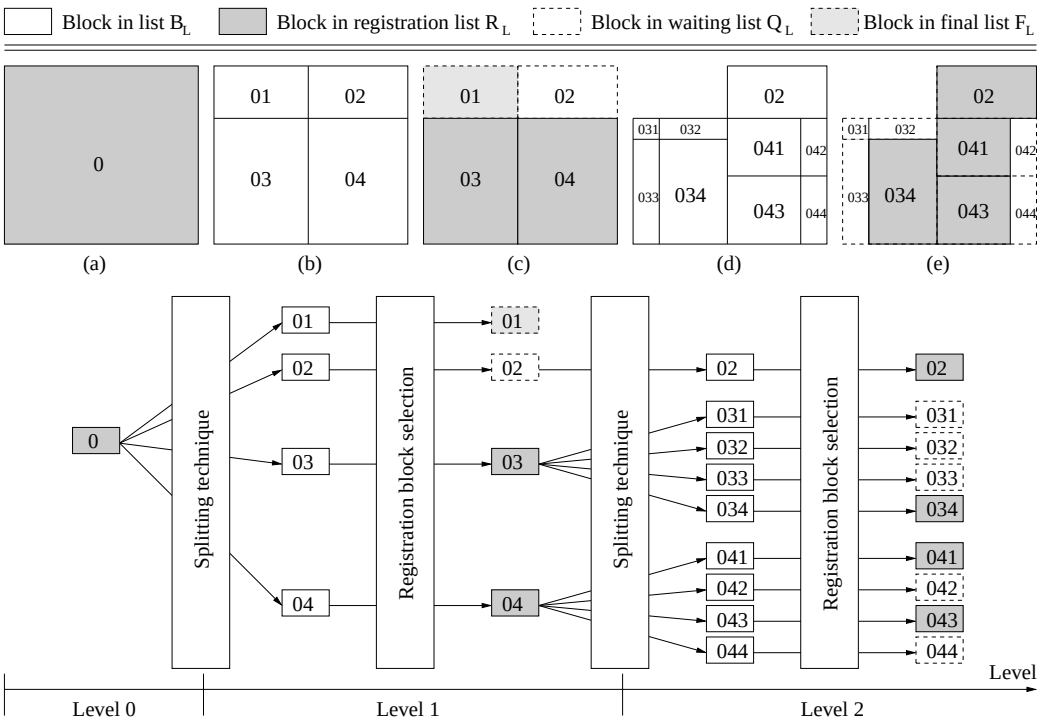



Figure 4: Adaptive splitting/registration scheme illustrated in 2D for one of the three proposed splitting techniques. After registration of the initial block 0 at level $L = 0$ (a), we apply any of our three adaptive splitting techniques (Section 2.2) to determine its children which form the sequence of blocks $B_1 = \{01, 02, 03, 04\}$ at level $L = 1$ (b). If a block in $B_{1,b} \in B_1$ meets a stopping condition (Section 2.7), it is appended to the final list $F_1$ (such as block 01), and no further registration and splitting is performed in $B_{1,b}$. All remaining blocks in $B_1$ are divided into a registration list $R_1 = \{03, 04\}$, and a waiting list $Q_1 = \{02\}$, based on their sizes. All blocks in $R_1$ are registered (c) and split further (d). The new children of level $L = 2$ are combined with the blocks in the waiting list $Q_1$ to form $B_2 = \{02, 031, 032, 033, 034, 041, 042, 043, 044\}$. This combined list of blocks is then divided into a new registration list $R_2 = \{02, 034, 041, 043\}$ and a new waiting list $Q_2 = \{031, 032, 033, 042, 044\}$, according to their sizes (e), and the process is repeated.

Fig. 4 illustrates the registration process. At any level, the new child blocks may vary in size because of the adaptive splitting technique. Since the deformations of large blocks might strongly influence the deformations of smaller neighbouring blocks when the overall motion field is interpolated, we try to ensure that only blocks of similar size are registered at each level. Large blocks are processed before small blocks which means that some smaller

blocks may wait until higher levels before being registered. Using this scheme, we ensure that larger blocks are registered at coarser levels in the hierarchy where they are more capable of capturing large deformations.

At each level $L$, the current list of potential blocks to be registered $B_L$ is divided into three new block lists: $Q_L$ is a waiting list of blocks still to be registered; $F_L$ is the list of final registered blocks in which no further splitting will be performed; and $R_L$ is the list of blocks which will be registered at level $L$ and subsequently be subdivided into further child blocks. Every time new child blocks are created by a splitting, each child is tested to see if it meets the stopping conditions. Those that do meet these conditions are added to the final list $F_L$ and are not processed further. The remaining blocks are either registered at the current level, i.e. added to $R_L$, or added to the waiting list $Q_L$. This choice is made based on the size of each block (its mean edge length). If all blocks are greater than half of the maximum block size, $0.5Max(B_L)$, all blocks in $B_L$ are registered. In other words, if all blocks are of a similar size, all are registered. If one or more blocks in $B_L$ is smaller than $0.5Max(B_L)$, then $B_L$ is divided into a registration list $R_L$ and a waiting list $Q_L$. $R_L$ contains blocks with block size greater than $0.5Max(B_L)$, and all remaining blocks are assigned to $Q_L$ and wait to be processed at higher levels. For example, Fig. 4.(b), shows the blocks $B_1 = \{01, 02, 03, 04\}$ at level $L = 1$ which are the child blocks of block 0 (Fig. 4.(a)). Note here that block 01 has met a stopping condition and is appended to the final list $F_1 = \{01\}$. Since 03 and 04 are significantly larger than 02, $B_1$ is divided into $Q_1 = \{02\}$ and $R_1 = \{03, 04\}$, and only the blocks in $R_1$ are registered (Fig. 4.(c)). Next, the blocks in $R_1$ are split further and combined with $Q_1$ to form a new block list $B_2 = \{02, 031, 032, 033, 034, 041, 042, 043, 044\}$ at level $L = 2$ (Fig. 4.(d)). We check the size of all blocks in $B_2$ and create the registration list $R_2 = \{02, 034, 041, 043\}$ from all large blocks in $B_2$, and all smaller blocks form the waiting list $Q_2 = \{031, 032, 033, 042, 044\}$ (Fig. 4.(e)).

During the hierchical registration process, each block $R_{L,r}$ of the floating image $I_F$ is registered to the complete underlying reference image $I_R$. At this stage, similar to Likar and Pernus (2001), we limit the moving range of each block in order to avoid large local distortions. During registration of a single block, we check if any the block's corner points move outside of a dilation area $D$ corresponding to the block's size. If this is the case, the registration result is limited to this range (Fig. 5).

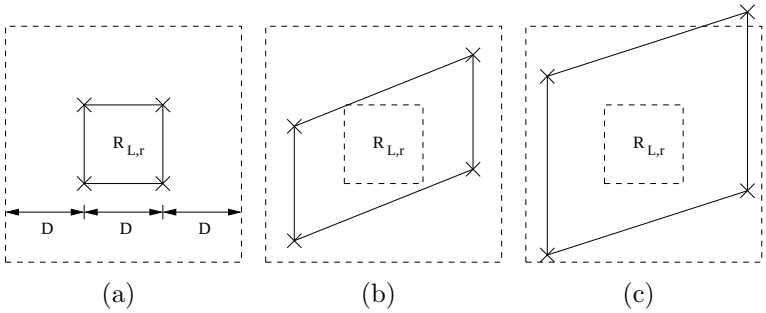



Figure 5: During registration the moving range of each block $R_{L,r}$ is limited to a dilated region $D$ corresponding to the block's size (a). We check if the 8 corner points of the block (here illustrated for the 4 corner points in a 2D example) move outside of this range (b). If this is the case (c) we limit the registration result to this range.

### *2.4. Multilevel B-Spline interpolation*

Because of the independent local affine registrations of all rectangular blocks $R_L$ at level $L$ distributed over $\Omega$, a subsequent interpolation scheme is required to ensure an overall smooth deformation $T_L$. In the original work of Likar and Pernus (1999, 2001) as well as in the extension of Andronache et al. (2005, 2008), the deformation vectors of the blocks' centres were used as anchor points followed by a TPS interpolation to derive an overall smooth deformation field. Although alternative approaches have been proposed which are equally suitable to combine multiple affine components, e.g. the log-euclidean polyaffine framework of Arsigny et al. (2005, 2009), we decided to use the original approach and propose the following modifications.

First, to ensure that rotational, scale and shear transformations of blocks are preserved in the displacements of control points, we define multiple anchor points $\Psi$ for all blocks in $B_L$ (Fig. 6.a). The anchor points of a single block $B_{L,b}$ are the eight corner points (3D) located at a distance of 25% of the block size away from the block's borders. For the anchor points of the registered blocks in $R_L$, we assign the block's deformation vector to each point $\psi(\vec{s}) \in R_{L,r}$ based on the corresponding local affine transformation for $R_{L,r}$. Anchor points located inside of blocks in the waiting list $Q_L$ or in the final list $F_L$ are assigned with zero displacements.

Second, to avoid the computational complexity of TPS interpolation, we use a multilevel B-Spline approach to compute the smooth output transformation $T_L$ as a free-form deformation (FFD) at level $L$ based on all blocks

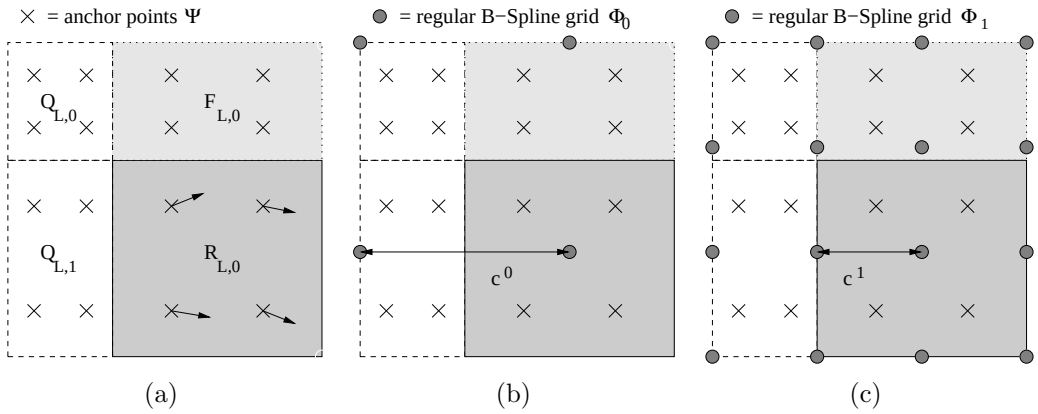



Figure 6: Illustration of the use of multilevel B-Spline interpolation to form a regular grid of control points. (a) Final block $F_L = F_{L,0}$ in which the splitting/registration process was stopped at a previous level (Section 2.7), blocks $R_L$ and $Q_L$, and the anchor points $\Psi = \{\psi(\vec{s})\}$. There is a single large registration block $R_L = R_{L,0}$ and 2 smaller blocks in the waiting list $Q_L = \{Q_{L,0}, Q_{L,1}\}$. The displacements of the anchor points of $R_{L,0}$ are assigned using the local affine registration within $R_{L,0}$ and the anchor points of $Q_L$ and $F_L$ are assigned with zero displacement. (b) Starting at a control point spacing (CPS) $c^0$ corresponding to the mean largest edge length of the registration blocks $R_L$ the deformation vectors at the uniform grid $\Phi_0$ are computed. (c) The deformation that has not been captured by (b) is captured by the next level with CPS $c^1 = c^0/2$. This process is repeated until the CPS $c_L$ at the hierarchical registration level $L$ equal to the mean minimum edge length of $R_L$ is reached.

$B_L$. In our tests multilevel B-Splines resulted in a similar interpolation performance to TPS, but at a much lower computational cost because of the local properties of B-Splines. However, whereas the TPS interpolation schemes of Likar and Pernus (1999, 2001) and Andronache et al. (2005, 2008) did not require control points to be on a regular grid, our B-Spline interpolation does require such a regular grid. Because of the irregular nature of the splitting in our adaptive scheme, we will not necessarily have a regular grid of control points, which presents a difficulty. We overcome this difficulty by making use of the scattered data interpolation technique using multilevel B-Splines (MBS) as proposed by Lee et al. (1997). At each hierarchical registration level $L$, we start with a large regular control point spacing $c^0$ (Fig. 6.b) and stop the MBS interpolation when a minimum control point spacing $c^L$ is reached (Fig. 6.c). The choice of $c^L$ is challenging since $c^L$ (a) has to be small enough so that smaller blocks contribute to the deformation as well as

larger blocks and (b) has to be large enough so that large deformations are not constrained by our folding detection (Section 2.5). Based on our experiments values for $c^0$ equal to the mean of the maximum edge length of any block in $R_L$ and $c^L$ equal to the mean of the minimum edge length in $R_L$ were chosen.

The final result of this process is a set of displacement vectors defined on a regular grid of control points $\Phi_L$. A subsequent B-Spline interpolation allows the determination of a single deformation vector $u(\vec{s})$ for each site $\vec{s} \in \Omega$. The FFD at the current hierarchical registration level $L$ is then defined by

$$T_L(\vec{s}) = \vec{s} + u_L(\vec{s}). \tag{4}$$

*2.5. Folding detection and correction*

Many applications require deformation fields to be smooth and invertible, i.e. to be diffeomorphic, e.g. deformation-based morphometry or statistical shape modelling (Rueckert et al. (2006)). Smoothness for each level $L$ is ensured by the MBS interpolation as described in the previous section. Invertibility, however, is not necessarily ensured, and folding artefacts within the FFD $T_L$ might still be present.

Rueckert et al. (2006) compared two different techniques to avoid folding within a FFD algorithm based on B-Splines. The first method introduced a regularization during the FFD by looking at the determinant of the Jacobian matrix. However, such a regularization is not applicable for our local affine registration scheme since our algorithm is not based on the deformation of control points but on the independent registrations of rectangular blocks. Therefore we cannot employ a regularizer during registration but have to correct for folding artefacts after MBS interpolation. The second method was based on the work of Choi and Lee (2000). They reported that, for each control point of the FFD, folding artefacts were avoided when the x/y/z components of the deformation vector were limited to 40% of the control point spacing (CPS). Rueckert et al. (2006) showed that large deformations can accurately be modeled when using the constraint of Choi and Lee within a CPS hierarchy followed by composing the FFDs as proposed by Hagenlocker and Fujimura (1998).

Therefore, within our hierarchical registration scheme, we employ the work of Choi and Lee to ensure that the resulting overall deformation $T_L$ at level $L$ is diffeomorphic. After the uniform grid $\Phi_L$ with CPS $c_L$ is created by the MBS algorithm, we consider the x/y/z motion components of $u_L(\phi_{i,j,k})$

at each control point $\phi_{i,j,k} \in \Phi_L$ and limit $u_L(\phi_{i,j,k})$ to 40% of $c_L$. If a x/y/z component of $u_L(\phi_{i,j,k})$ moves outside of this range, a scaling factor $\lambda_{L,d}$ is applied to scale down that motion component:

$$\lambda_{L,d} = \begin{cases} 1, & \text{if } ||u_{L,d}|| \leq 0.4c_L, \text{ and} \\ \frac{0.4c_L}{||u_{L,d}||}, & \text{if } ||u_{L,d}|| > 0.4c_L, \end{cases} \tag{5}$$

with $d = x, y, z$. The corrected motion vector at grid point $\phi_{i,j,k}$ is then given by $(\lambda_{L,x}u_{L,x}, \lambda_{L,y}u_{L,y}, \lambda_{L,z}u_{L,z})^T$.

### *2.6. Motion field combination*

After ensuring a diffeomorphic transformation $T_L$ at level $L$, we combine the transformations of all levels up to level $L$ to create an overall output deformation field $T_{L,out}$. We combine our FFDs as follows:

$$T_{L,out}(\vec{s}) = T_L \circ T_{L-1} \circ ... \circ T_1 \circ A(\vec{s}), \tag{6}$$

where $A$ is the initial affine registration at level $L = 0$ and $T_L$ is the FFD of level $L$.

Since the concatenation of diffeomorphic deformations is a diffeomorphism again, we can compute the overall Jacobian $J_{L,out}$ of $T_{L,out}$ by:

$$J_{L,out}(\vec{s}) = J_L \cdot J_{L-1} \cdot ... \cdot J_1 \cdot J_0(\vec{s}), \tag{7}$$

where $J_L$ with $L > 0$ is the Jacobian of the FFD $T_L$ at level $L$ and $J_0$ is the Jacobian of $A$. We ensured in Section 2.5, that the Jacobian of each transformation $T_L$ is greater than 0, i.e. $T_L$ is diffeomorphic. According to (7), the overall Jacobian $J_{L,out}$ of our final deformation field is greater than 0, too. Hence, the overall deformation field $T_{L,out}$ is invertible, and the inverted field $T^{-1}_{L,out}$ can be determined as the inverse concatenation of the inverted level transformations.

### *2.7. Block conditions*

Similar to the starting control point spacing in the hierarchical B-Spline approach of Rueckert et al. (1999), we introduce a *starting condition*. A block $B_{L,b}$ in $B_L$ is only registered if its size (mean edge length) is smaller than the user-defined maximum block size $S_0$. If $B_{L,b}$ is greater than $S_0$, $B_{L,b}$ is split further without being registered. This parameter is analogous to the

maximum control point spacing parameter commonly employed in many free form deformation algorithms.

We also introduce two stopping conditions during the hierarchical registration process: a *local stopping condition* which locally avoids unnecessary registrations and a *global stopping condition* which is independent from local image properties and terminates the overall registration process.

The *local stopping condition* checks the residual deformation in a block based on the similarity measure. If the similarity is sufficiently high, the hierarchical splitting in that block is stopped. Similar to Likar and Pernus (2001), we compute the normalized similarity between the reference image $I_R$ and the transformed floating image $T_{L,out}(I_F)$ for each block $B_{L,b} \in B_L$:

$$Sim_{L,b} = Sim\left(I_R(\vec{s}), T_{L,out}(I_F)(\vec{s})\right), \vec{s} \in B_{L,b}. \tag{8}$$

If $Sim_{L,b}$ is greater than a threshold which is chosen according to the initial global similarity between $I_R$ and $I_F$, the hierarchical splitting process is terminated and no further children of $B_{L,b}$ are created:

$$Sim_{L,b} > \lambda_t \cdot Sim_{0,0}, \tag{9}$$

where $Sim_{0,0}$ is the initial similarity in $\Omega$ and $\lambda_t$ is the threshold. The precise value of $\lambda_t$ will depend on the similarity measure used, and details will be provided in Section 3.2.

The *global stopping condition* determines the overall registration precision. The user can specify a minimum allowed block size $S_1$. When the size of block $B_{L,b}$ is smaller than $S_1$, we do not register $B_{L,b}$, exclude the block from the block list $B_L$, and place it in the final list $F_L$. This parameter is analogous to the minimum control point spacing in freeform deformation algorithms. The overall hierarchical registration process is finished when no blocks are left to be registered.

## 3. Experiments and Results

Our main validation was performed on dynamic 3D MRI data (Section 3.1.1). These experiments compared the different splitting techniques and are described in Sections 3.2.1 and 3.3.1. To allow for more extensive comparison with other algorithms, we performed extra validation on the publicly available POPI CT dataset (Section 3.1.2) and we participated in the EMPIRE10 registration challenge also based on CT data (Section 3.1.3). We describe these experiments and their motivation in Sections 3.2.2/3.3.2 and Sections 3.2.3/3.3.3.

### *3.1. Materials*

#### *3.1.1. MRI data*

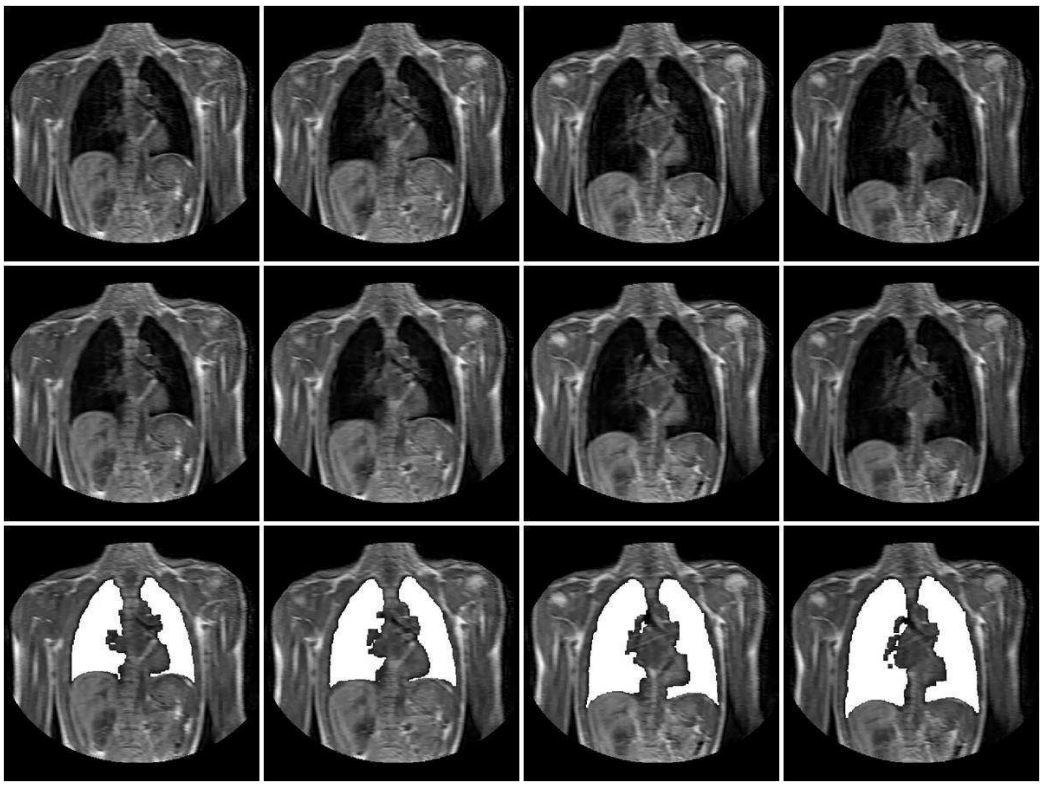
10 volunteers were scanned on a 1.5T Philips Achieva MRI scanner using a 32 channel coil. An ECG triggered dynamic 3D acquisition, with SENSE-factor of 2 in Anterior-Posterior (AP) and SENSE-factor of 4 in Right-Left (RL) direction was applied. With $TR/TE = 3.3ms/0.9ms$ and a flip angle of 10° we acquired whole-thorax volume with a field of view (FOV) of $500 \times 450 \times 245 mmm^3$. With an acquired image resolution of $1.5 \times 4.1 \times 5mm^3$ (Feet-Head (FH), RL, AP) and a reconstructed image resolution of $1.5 \times 1.5 \times 5mm^3$ we acquired 35 three-dimensional image frames at a time resolution of $0.7s$ per frame. All volunteers were instructed to breathe deeply during imaging to maximize the magnitude of the motion caused by respiration.

For each volunteer, 10 dynamic images distributed over a range of respiratory positions were selected and sorted from end-exhale to end-inhale according to the FH translation of the right hemi-diaphragm (Fig. 7, top row). This allowed us to compare registration results for small (dynamic 1) and large deformations (dynamic 10).

In order to ensure a relatively high temporal resolution of $0.7s$ per image frame, we needed to select a relatively low spatial resolution. This leads to noise and Gibbs ringing artefacts due to small readout matrices in the RL and AP directions (see acquired image resolution). Therefore, as a preprocessing step we performed a smoothing operation to minimize these artefacts. We want to smooth homogeneous image regions with no structural information but avoid any smoothing between different tissue types, i.e. to preserve edges. To this end, anisotropic diffusion filtering as proposed by Perona and Malik (1990) was applied. This filter performs an iterative smoothing operation on an image $I$ at each voxel $\vec{s} = (x, y, z)$, using a smoothing weight determined by a conduction coefficient function $c(\vec{s})$ which depends on the underlying image gradient $\nabla I(\vec{s})$. Filtering examples are shown in Fig. 7, middle row.

#### *3.1.2. POPI CT data*

We applied extra registrations using the publicly available point-validated pixel-based breathing model (POPI-model, Vandemeulebroucke et al. (2007)) which contains 10 thoracic CT images from a single individual covering the respiratory cycle, from end-exhale to end-inhale. This dataset also provides



(a) (b) (c) (d)

Figure 7: Acquired dynamic 3D MRI data from *volunteer 1* in coronal view at four different respiratory phases, from end-exhale (a) to end-inhale (d). Top row: Original acquired data. Middle row: Image data after anisotropic diffusion filtering. Bottom Row: An automated histogram-based threshold technique was applied to segment the lungs.

400 landmarks (40 in each phase) which have been identified by medical experts in order to assess motion field accuracies.

### *3.1.3. EMPIRE10*

We also participated in the registration challenge "Evaluation of methods for pulmonary image registration" (EMPIRE10[1]) (Murphy et al. (2010)), allowing a comparison of our approach with more than 20 competing groups working in the field of medical image registration. Participants apply their

[1] `http://empire10.isi.uu.nl/`

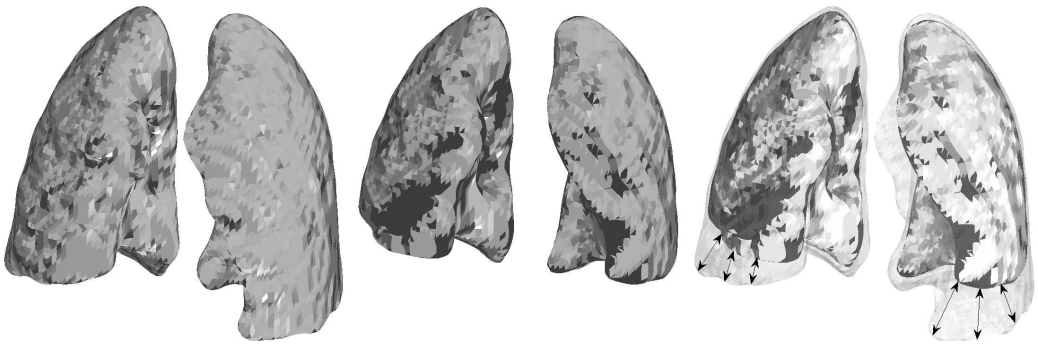
approaches to 30 lung-CT image pairs, and the estimated motion fields are evaluated using the same criteria to allow adequate performance comparisons.

### *3.2. Experiments*

#### *3.2.1. MRI data*

We applied our hierarchical registration algorithm to register each selected dynamic MRI image with a floating image at the end-exhale position by applying 4 different splitting techniques (Section 2.2): regular non-adaptive splitting (NA), adaptive splitting based on image gradients (GRAD), adaptive splitting based on the similarity measure (SIM), and adaptive splitting based on previously estimated motion fields (MOT). This led to 400 registrations over all volunteers. We used a steepest gradient descent optimizer for the local affine registrations and the negated sum of squared differences (SSD) as a similarity measure (both for each individual affine registration and for SIM as the adaptive splitting technique). Since we tested our technique on intra-modal registration using free-breathing MRI data, a mono-modal similarity measure was appropriate. We used an empirically chosen similarity threshold of $\lambda_t = 0.05$ as the *local stopping condition* and applied the registrations within a rectangular region of interest $\Omega$ of $450 \times 420 \times 225 mm^3$ which was manually selected around the chest of each volunteer to exclude the background. With respect to the reconstructed image resolution of our acquired data ($1.5 \times 1.5 \times 5 mm^3$), we downsampled the image data using $5 \times 5 \times 1$ and $3 \times 3 \times 1$ Gaussian kernels leading to 3 image resolution levels. With a typical number of 20 registration levels for the adaptive techniques, we used the lowest resolution data for the registration levels 0-6 (resolution level 2), the medium resolution data for level 7-13 (resolution level 1), and the original image data from level 14 onwards (resolution level 0). With a typical number of 5 registration levels for the regular splitting technique, we used the lowest resolution data for level 0, the medium resolution data for level 1, and the original data from level 2 onwards.

To evaluate registration accuracy, we performed a comparison with the commonly used free-form deformation (FFD) algorithm based on B-Splines as proposed by Rueckert et al. (1999). We applied the FFD algorithm for all volunteers. Large deformations as well as small deformations were captured by using a composition of FFDs, i.e. the control point spacing was varied in a coarse to fine manner. For our technique, we specified a maximum



(a) (b) (c)

Figure 8: (a) Deformed lung $T_{L,out}(I_{F,Lung})$ of the lung segmentation $I_{F,Lung}$ of the floating image. (b) Lung segmentation $I_{R,Lung}$ of the reference image. (c) As a measure of registration accuracy we compute the shortest distance from each surface point of $I_{R,Lung}$ to the surface of $T_{L,out}(I_{F,Lung})$.

block size $S_0$ equal to the size of the whole FOV as the *starting condition*. For an adequate comparison between our technique and FFD we empirically specified a minimum block size $S_1 = 10mm$ as the *global stopping condition* for our approach and a minimum control point spacing of $10mm$ for the FFD algorithm. We evaluated registration accuracies using the following two methods.

*Method 1* compares the overall amount of misalignment between the transformed floating image $T_{L,out}(I_F)$ and the reference image $I_R$. We concentrated on the deformation of the lungs and considered a manually selected rectangular block around the lungs of the end-inhale image (dynamic 10). As a measure of registration accuracy, we computed the SSD between $T_{L,out}(I_F)$ and $I_R$ within this region of interest.

*Method 2* computes the surface distances between the lung segmentation of $I_R$ and the warped lung segmentation of $I_F$ (Fig. 8). First, we automatically segmented the lungs in all dynamic images. This was done by modelling and interpolating the 3D histogram of each image with a mixture of Gaussian distributions using the Expectation-Maximization (EM) algorithm (Dempster et al., 1977). Lung tissue and background were then separated from the remaining thoracic tissue using a threshold corresponding to the first minimum in the interpolated histogram. A subsequent region growing technique was applied to exclude the background, and a morphological operation (opening/closing) to close small holes in the segmentation result. Segmenta-

tion examples are overlaid onto the images of Fig. 7, bottom row. Next, we applied the hierarchical registration result to warp the lung segmentation of the floating image $I_{F,Lung}$. Finally, we extracted the surface vertices of the lung segmentations using *isosurface* in MATLAB and computed the shortest distances between each surface point of $I_{R,Lung}$ and the lung surface of $T_{L,out}(I_{F,Lung})$. As a measure of registration accuracy, we computed the root mean squared error (RMSE) over all these surface distances.

### *3.2.2. POPI CT data*

We registered the reference end-inhale phase with the remaining 9 phases of the POPI model. We used normalized cross correlation (NCC) as similarity measure and an empirically chosen similarity threshold of $\lambda_t = 0.99$. We used the complete field of view as starting block size $S_0$ and a minimum block size of $S_1 = 5mm$. With a typical number of registration levels of 15 for the adaptive techniques, we downsampled the images by a $6 \times 6 \times 3$ kernel for levels 0 to 3, by a $4 \times 4 \times 2$ kernel for levels 4 to 7 and by a $2 \times 2 \times 1$ kernel for levels 8 to 11. From level 12 on, the original image resolution was used. When using NA as splitting technique we observed a typical number of 5 registration levels and downsampled the images by a $6 \times 6 \times 3$ kernel for level 0, by a $4 \times 4 \times 2$ kernel for levels 1 to 2 and by a $2 \times 2 \times 1$ kernel for levels 3 to 4. From level 5 on, the original image resolution was used.

We assessed motion field accuracy by computing the distances between the transformed landmarks of the reference frame and the landmarks of the other phases (target registration errors, $TRE$).

### *3.2.3. EMPIRE10*

We registered all 30 image pairs provided in the challenge. Since this dataset requires to estimate small as well as large deformations, we used the complete field of view as starting block size $S_0$ and chose the minimum block size to be $S_1 = 5mm$. We used NCC as similarity measure and an empirically chosen similarity threshold of $\lambda_t = 0.98$. With a typical number of 17 registration levels over all 30 registration pairs, we downsampled the images by a factor 8 for levels 0 to 4, by factor 4 kernel for levels 5 to 9 and by factor 2 for levels 10 to 15. From level 16 on, the original image resolution was used.

Motion field accuracy was assessed by lung boundary and fissure alignment, by computing the $TRE$ based on manually selected landmarks, and

by performing a singularity check in all deformation fields.

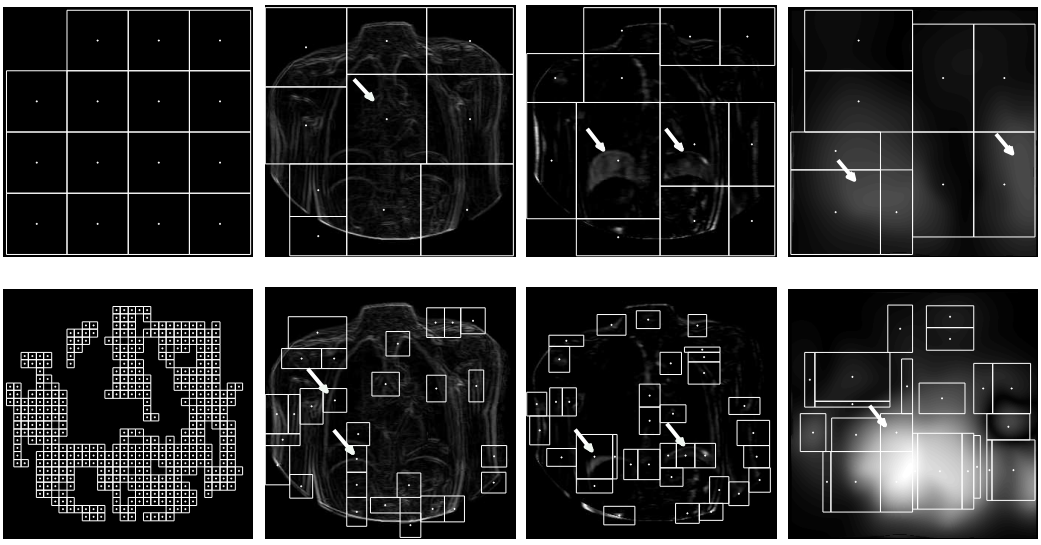
### *3.3. Results*

#### *3.3.1. MRI data*

Fig. 9 illustrates the determination of the optimal splitting points for our different splitting techniques. In this 2D registration example we used a floating image $I_F$ produced from Fig. 7.a and a reference image $I_R$ produced from Fig. 7.d. The registration blocks $R_2$ at registration level $L = 2$ (top row) and the blocks $R_5$ at level $L = 5$ (bottom row) are overlaid onto the images that were used to estimate the optimal splitting points. For non-adaptive splitting (Fig. 9.a), no image information is used to determine the splitting points and each block is successively split into sub-blocks of equal size. Our three adaptive techniques consider the cost functions as described in Section 2.2 to adaptively determine the optimal splitting points. As can be seen in Fig. 9.b, GRAD considers the sum of gradient images of $I_R$ and the transform of $I_F$ and intends to avoid splitting through regions with large amounts of structural information (indicated with arrow). When using SIM (Fig. 9.c) and the negative SSD as similarity measure, the squared difference (SD) image between $I_R$ and the transform of $I_F$ is considered and the splitting point is optimized by avoiding splitting through regions with low similarity (here regions with large values in the SD image), e.g. the residual deformation of both the left and the right hemi-diaphragms. This allows us to process single large deformations in single registration blocks. Our last adaptive technique, MOT (Fig. 9.d), uses the squared sum of motion components and tends to keep large registration blocks around regions with previously estimated large deformations. Note that we used the same stopping conditions for all the splitting techniques shown in Fig. 9.

Fig. 10 shows the first 11 registration levels for a 2D registration example produced from the images from Fig. 7.a and Fig. 7.d (*volunteer 1*) using SIM as the splitting technique. SIM focuses on regions with remaining residual deformations. As can be observed at levels 8 and 9, the overall registration is refined by locating registration blocks around the remaining deformation, e.g. of the diaphragm. The final motion field is overlaid onto the warped floating image.

Fig. 11 illustrates *Method 1* and *Method 2* for a 3D registration example for *volunteer 2*. This example registered the end-inhale reference image $I_R$



(a) (b) (c) (d)

Figure 9: Comparison of the regular splitting techique and the three proposed splitting techniques in the hierarchical registration framework, here in 2D. The blocks shown are the registration blocks at level $L = 2$ (top) and at level $L = 5$ (bottom). (a) Non-adaptive splitting (NA). The underlying image content is not considered, and each block is split into its child components of equal size. (b) Adaptive splitting based on image gradients (GRAD). GRAD attempts to avoid splittings through regions showing high image gradients. The underlying image shows the sum of image gradients. (c) Adaptive splitting based on the underlying image similarity (SIM). SIM attempts to avoid splitting though regions of low image similarity. We used the negative SSD as similarity measure and the underlying image shows the negative squared difference image. (d) Adaptive splitting based on motion fields (MOT). MOT attempts to avoid splitting through regions showing previously estimated large deformations. The underlying image shows the absolute motion estimated at the previous level $L - 1 = 1$ (top) and $L - 1 = 4$ (bottom). Note that the stopping conditions (Section 2.7) are the same for all splitting techniques (a) - (d). While the missing blocks in (a) are excluded because they met the local stopping condition, the missing blocks in (b) - (d) are excluded because they either met the local stopping condition, or they were too small and are waiting in the waiting list $Q_L$ to be processed at higher registration levels.

(dynamic 10, Fig. 11.a) with the end-exhale floating image $I_F$ using our approach with the proposed regular and adaptive splitting techniques (Fig. 11.b - e) and using the FFD algorithm (Fig. 11.f). In Fig. 11, top row, the final transformations $T_{L,out}(I_F)$ are shown. Based on the squared difference images (Fig. 11, middle row), *Method 1* computes the SSD between

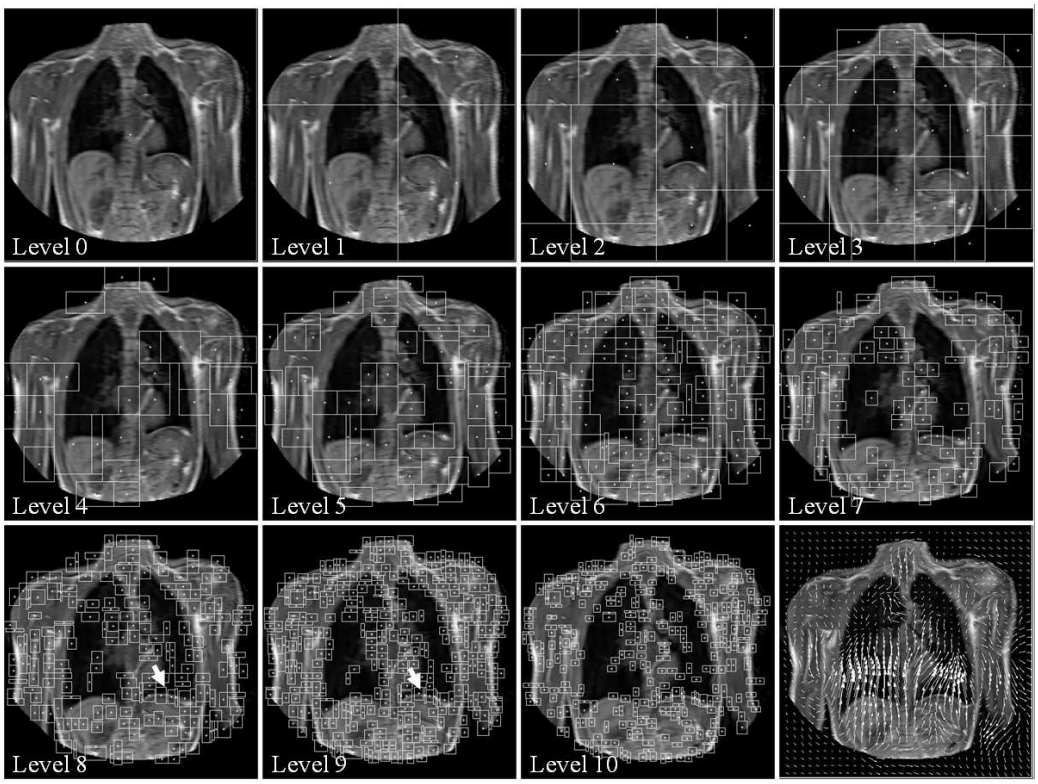



Figure 10: 2D registration example between the floating image $I_F$ from Fig. 7.a and the reference image $I_R$ from Fig. 7.d (*volunteer 1*) using a minimum block size of $10mm$ and an adaptive splitting technique based on the similarity measure (SIM). The registration hierarchy is shown for registration levels $L = 0$ to $L = 10$. As indicated in the images of level 8 and 9, this splitting technique aligns registration blocks around regions with remaining deformation, e.g. the residual deformation of the diaphragm indicated with arrows. The final motion field is shown and overlaid onto the warped exhale image. Note that motion fields appearing outside of the body are a result of motion field interpolation due to a lack of structural information.

$I_R$ and $T_{L,out}(I_F)$. In Fig. 11, bottom row, each surface point of the lung segmentation of $I_R$ is colour encoded according to the shortest distance to the warped surface of the corresponding lung segmentation of $I_F$. These surface distances are used to compute RMSE according to *Method 2*. When using regular splitting (NA), maximum distances greater than $40mm$ can be observed around the diaphragm. Clear overall improvement is produced by all adaptive techniques, GRAD, SIM and MOT, which achieved comparable

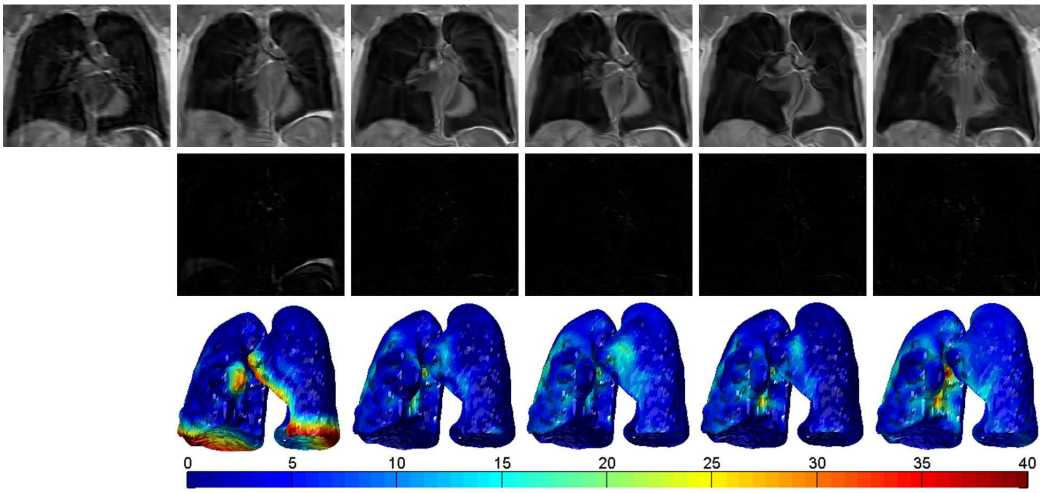



Figure 11: 3D registration example for *volunteer 2*. Top row: Coronal view of a warped floating image $T_{L,out}(I_F)$ at end-exhale position after alignment with a reference image $I_R$ at end-inhale position shown in (a). We compared registration accuracies between (b) non-adaptive splitting (NA), (c) gradient-based splitting (GRAD), (d) similarity-based splitting (SIM), (e) motion-based splitting (MOT), and (f) the free-form deformation (FFD) algorithm as proposed by Rueckert et al. (1999). Middle row: *Method 1*, coronal view of the squared difference images between $T_{L,out}(I_R)$ and $I_F$ within a manually selected region of interest around the lungs. Bottom row: *Method 2*, each point of the lung surface of $I_R$ is colour encoded using the shortest distance to the lung surface of $T_{L,out}(I_F)$. As can be observed in errors around the diaphragm, NA performs least well. Registration accuracy was clearly improved by all adaptive techniques (GRAD, SIM and MOT) with comparable results to FFD.

results to the FFD algorithm.

Fig. 12 shows the results of *Method 1* and *Method 2* for the same *volunteer 2* but now for all 10 dynamics, from small deformations (dynamic 1) to large deformations (dynamic 10). In addition, we investigated local variations in registration accuracy by dividing the manually selected region of interest around the end-inhale lungs into 3 segments. The corresponding values for SSD (*Method 1*) and RMSE (*Method 2*) were separately computed for the top, middle, and bottom lung regions. For small deformations, i.e. dynamics 1 - 6 and the top and middle lung regions, the performances of our different splitting techniques were quite similar and comparable to the registration results of the FFD algorithm. However, a clear difference in registration

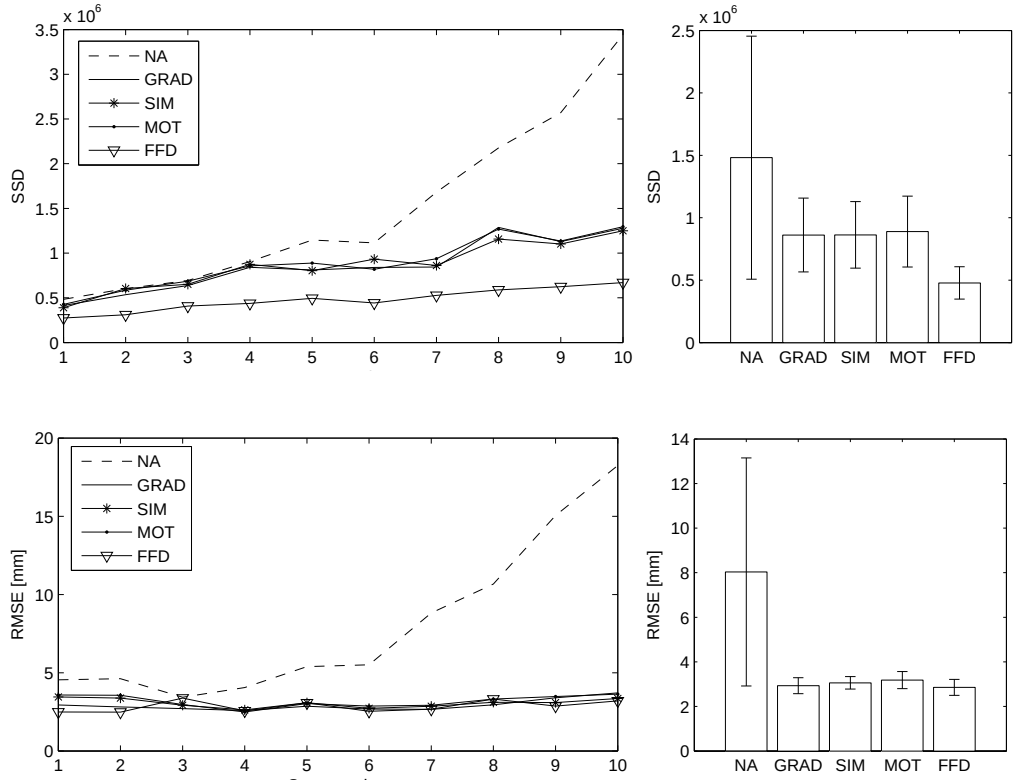



Figure 12: *Method 1* and *Method 2* for *volunteer 2* within a manually selected region of interest around the lungs. We distinguish between small and large deformations (1) by sorting all dynamics from end-exhale (dynamic 1) to end-inhale (dynamic 10) and (2) by considering 3 different lung segments (top, middle and bottom). Here, we show the results for the bottom region of the lung where the largest deformations are expected. Top: *Method 1*, SSD between the reference image $I_R$ and the warped floating image, $T_{L,out}(I_F)$. Bottom: *Method 2*, RMSE between the lung surface of $I_R$ and the lung surface of $T_{L,out}(I_F)$.

accuracy can be observed when considering large deformations, i.e. dynamics 7 - 10 and the bottom region of the lung (Fig. 12). While the FFD algorithm shows the most stable and overall most accurate results ($SSD = 4.8 \times 10^5 \pm 1.3 \times 10^5$, $RMSE = 2.9 \pm 0.4mm$), a clear decrease in registration accuracy can be observed for NA ($SSD = 14.8 \times 10^5 \pm 9.7 \times 10^5$, $RMSE = 8.0 \pm 5.1mm$). In comparison, all adaptive splitting techniques, GRAD ($SSD = 8.6 \times 10^5 \pm 3.0 \times 10^5$, $RMSE = 2.9 \pm 0.4mm$), SIM ($SSD = 8.6 \times 10^5 \pm 2.7 \times 10^5$, $RMSE = 3.0 \pm 0.3mm$) and MOT ($SSD = 8.9 \times 10^5 \pm 2.8 \times 10^5$, $RMSE =$

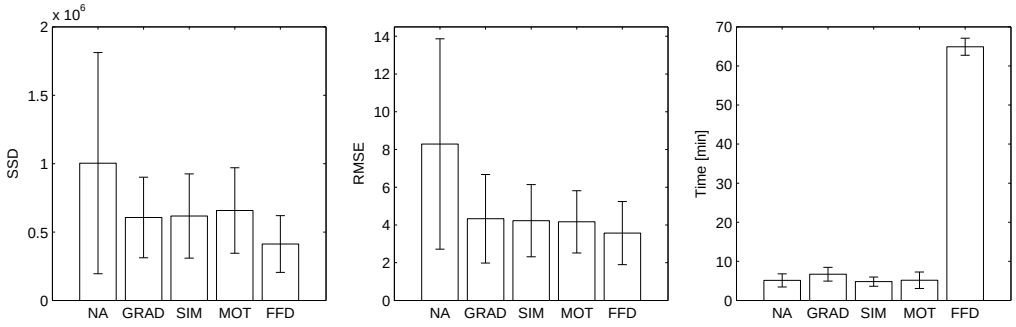



Figure 13: Results for all volunteers. (a) *Method 1*, SSD between the reference image $I_R$ and the warped floating image, $T_{L,out}(I_F)$. (b) *Method 2*, RMSE between the lung surface of $I_R$ and the lung surface of $T_{L,out}(I_F)$. As can be observed especially for large deformations, our adaptive techniques clearly improved the registration accuracy over non-adaptive/regular splitting. (c) Computational times.

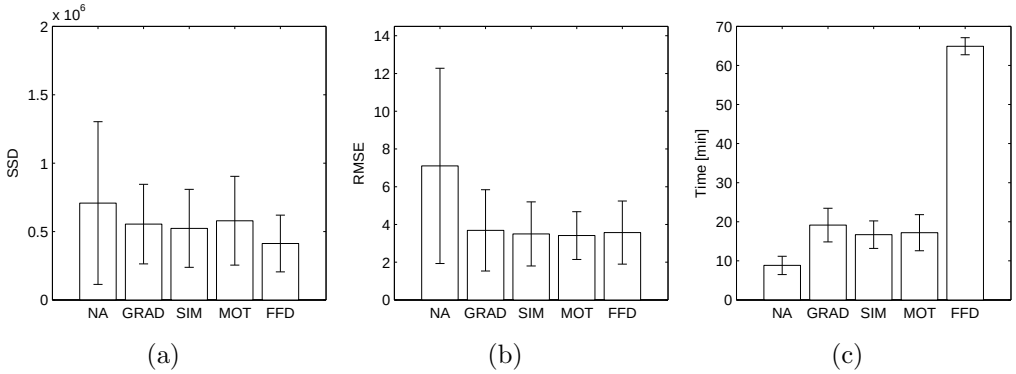
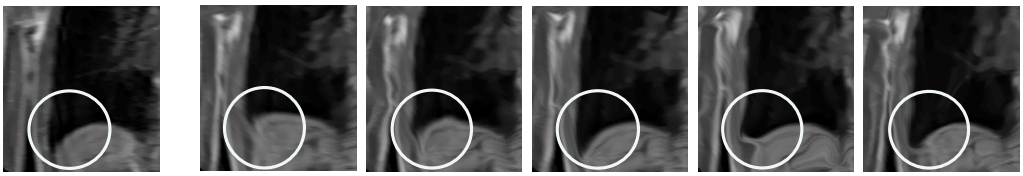



Figure 14: Results for all volunteers. (a) and (b) show again the results of *Method 1* and *Method 2*, now with a minimum block size of $5mm$. All splitting techniques were improved, and we achieved even better results for SIM and MOT compared to FFD. (c) Computational times.

$3.2 \pm 0.4mm$) were more capable of capturing large deformations, resulting in more stable registration results. Note here that the results for dynamic 10 correspond to the example illustrated in Fig. 11.

The overall registration performance of our hierarchical approach and of the FFD algorithm for all 10 volunteers (10 registrations each) is shown in Fig. 13.a,b. Although our approach incorporates a multilevel B-Spline



(a) (b) (c) (d) (e) (f)

Figure 15: 2D MRI registration example between an exhale phase and the inhale-phase shown in (a). (b-f) Warped exhale images to align with (a). While the regular splitting technique (NA) performs worse (b), all three adaptive techniques, GRAD (c), SIM (d) and MOT (e), improve the registration result. SIM shows best performance and compares well to FFD (f).

interpolation rather than TPS, an alternative stopping condition in the splitting process and multiple control points within each block, our non-adaptive technique can be viewed as a modified version of the algorithm described by Andronache et al. (2008), which represents the current state of the art in hierarchical locally affine image registrations. Considering the largest deformations, i.e. the lower third of the lung, we observed registration accuracies of $SSD = 10.0 \times 10^5 \pm 8.1 \times 10^5$ and $RMSE = 8.3 \pm 5.6mm$ for NA. Our adaptive techniques (GRAD, SIM, MOT) achieved improvements of 39.6% (GRAD), 38.5% (SIM), and 34.5% (MOT) over the non-adaptive approach using *Method 1*, and 47.9% (GRAD), 49.1% (SIM), and 49.8% (MOT) using *Method 2*.

We also tested the effect of a further decrease in minimum block size from $10mm$ to $5mm$ for our technique, while the mininum CPS remained unchanged ($10mm$) for FFD. While the computational times were still clearly lower than FFD, we achieved additional improvement in registration accuracy for all splitting techniques (Fig. 14), and SIM and MOT showed even better performance than FFD for *Method 2*.

Our results indicate that the accuracies achieved by our adaptive approaches, GRAD, SIM and MOT, are comparable with the "gold standard" of the FFD approach. However, small local differences in registration performance cannot be distinguished by our global measures, *Method 1* and *Method 2*. Fig. 15 illustrates some local differences and shows that all adaptive techniques perform better than NA, with SIM performing the best. We further evaluate local performances in the next section.

The computational complexity of our algorithm is significantly lower since

we decompose the non-rigid registration problem into locally independent affine components. In comparison to Andronache et al. (2005, 2008), we additionally reduced complexity by (a) applying a multilevel B-Spline interpolation rather than TPS and (b) including a multi resolution approach as incorporated in Rueckert et al. (1999). We ran all registrations on a workstation with 2 Six-Core processors, Intel Xeon X5670, 2.93GHz, and 96GB Memory. The mean computational time for each registration (3D dynamic to 3D dynamic) was $64.9 \pm 2.2min$ when applying a CPU implementation of the FFD algorithm (Fig. 13.c). In comparison our CPU implementation of the hierarchical local affine approach achieved computational times of $5.1 \pm 1.7min$ (NA), $6.7 \pm 1.8min$ (GRAD), $4.8 \pm 1.2min$ (SIM) and $5.2 \pm 2.1min$ (MOT). This represents speedup factors of 12.7 (NA), 9.7 (GRAD), 13.5 (SIM) and 12.6 (MOT).

### *3.3.2. POPI CT data*

Considering all 9 registrations of the POPI model, we computed the $TRE$ over all landmarks and achieved a $TRE$ (mean $\pm$ standard deviation) of $1.2 \pm 0.9mm$ for NA, $2.0 \pm 1.8mm$ for GRAD, $1.2 \pm 1.0mm$ for SIM, and $1.2 \pm 0.9mm$ for MOT. For comparison, the FFD approach has been reported to achieve a $TRE$ of $1.0 \pm 0.5mm$. Note here that NA performs well because the overall motion to be captured is quite small ($3.3 \pm 2mm$ as mean original landmark distances). However, as shown in Section 3.3.1, NA fails when large deformations have to be captured. We ran all registrations on the same machine as in Section 3.3.1 and achieved mean execution times of $6.8min$ for NA, $9.0min$ for GRAD, $5.27min$ for SIM and $12.5min$ for MOT. These results, combined with the MRI results from Section 3.3.1, suggest that SIM is the best adaptive splitting technique in our hierarchical framework

### *3.3.3. EMPIRE10*

Since Section 3.3.1 and Section 3.3.2 indicate that SIM is the adaptive splitting technique that produces the fastest and most accurate registration results, we applied our registration algorithm using SIM to all 30 registration pairs. Over all registrations, we achieved an error of lung boundary and fissure alignment of 0.01% and 1.27%, respectively, a $TRE$ of $1.91mm$, and all our motion fields avoided singularities (0.00% error). We ran all registration on the same machine as used for the results being reported in Section

3.3.1, and our CPU-based implementation achieved a mean execution time of $13.2min$.

## 4. Discussion

In this paper, we presented a fast and robust non-rigid registration algorithm based on a composition of hierarchical local affine registrations. The original work of Likar and Pernus (1999, 2001) and Andronache et al. (2005, 2008) was improved in three main ways. First, we introduced adaptive splitting techniques which consider the underlying image content and attempt to process regions of similar motion and/or image structure in single registration components. Second, we introduced a waiting-registration scheme. While all blocks of a single registration level have the same size in NA, our adaptive techniques introduce blocks of varying size. Since large blocks are registered before small blocks, the average block size per registration level decreases slower over the increasing registration levels, so that the block size at a single registration level is significantly larger (Fig. 9.b-d) compared to NA (Fig. 9.a). As a result, the search range of each registration block is increased and large deformations at high registration levels can better be captured. Third, we reduced computational cost by using a locally-based B-Spline interpolation scheme.

Although the overall results of GRAD, SIM and MOT show similar performance when we applied the different splitting techniques to the POPI-model, differences in execution times were observed. In addition, visual inspection revealed some local anatomically unrealistic motion field distortions using GRAD and MOT in the MRI dataset (Fig. 15). More realistic and stable results were observed with SIM. Therefore, the similarity-based splitting technique seems to be the best splitting technique in our hierarchical framework. For SIM, the combination of adaptive splitting and the waiting-registration scheme led to improvements in registration accuracy of up to 38.5%/49.1% (*Method 1*/*Method 2*), compared to regular non-adaptive splitting.

We also compared our technique to a commonly used free-form deformation algorithm (Rueckert et al., 1999). While this FFD algorithm showed a registration accuracy of $RMSE = 3.6mm$ for the largest deformations, our hierarchical adaptive approach showed slightly less accurate results ($RMSE = 4.2mm$). However, we demonstrated a significant speedup factor of up to 13.5 in execution time. Furthermore, by reducing the minimum block size in our

hierarchical technique we achieved similar results to FFD and still achieved a speedup factor of 3.9.

In addition, we participated in the EMPIRE10 registration challenge (using SIM as the splitting technique). Compared to other participating groups in the challenge, the main advantages of our approach are: (i) with an average execution time of $13.2min$ our method was one of the fastest CPU implementations reported; (ii) since we use the complete field of view as a starting block size, our method is robust against large deformations allowing us to accurately align the lung and fissure boundaries (0.01%/1.27% error) without having to choose registration parameters to optimize performance; and (iii) the generated motion fields are singularity-free (0.00% error). A possible weakness of our approach is an accurate alignment of fine image structures at high registration levels which led to a $TRE$ of $1.9mm$. Some competing groups achieved $TRE$ values of less than $1mm$. The constraints of affine registrations, even when applied in very small blocks, appear to limit our technique's ability to capture deformations with very high (i.e. sub-millimetre) accuracy. In addition, the robustness of these small-block registrations may be limited by the small number of contributing voxels in such blocks. In future work, we will address this issue, e.g. by considering overlapping blocks at high registration levels, or an alternative approach to describe the level deformations.

We currently use the sum of squared differences and normalized cross correlation as similarity measures. However, the general framework we have described allows other, more sophisticated, similarity measures to be incorporated easily. It should be noted, though, that some similarity measures, such as mutual information (MI), can be difficult to estimate in small image regions such as our registration blocks. Different techniques have been proposed to deal with this. In Likar and Pernus (2001) a combined global/local MI was used by including a prior joint probability, and in Andronache et al. (2008) "pseudo-modal" cross-correlation was used as a replacement measure when block size became too small. We plan to build on these ideas to develop our algorithm into a more general approach that can be used for inter-modal registration problems.

In future work we also plan to investigate how our algorithm can be extended to incorporate temporal constraints on the registration results of dynamic images acquired sequentially. This could include, for example, a "trajectory" constraint between images on the registration results of sub-blocks, similar to that recently proposed in Castillo et al. (2010), or an

analysis of the degree of periodicity of the motion of different sub-blocks. The ultimate aim would be to register all dynamic images simultaneously rather than sequentially, to ensure smooth variation in motion between images. We also intend to apply these approaches to allow discontinuous tissue deformations, such as the sliding motion of the lung-chest interface, e.g. by adaptively employing smaller control point spacings in such regions to model "quasi"-discontinuous deformations.

Our algorithm is capable of estimating whole-thorax deformations from free-breathing 3D MRI. Since the computational complexity is significantly reduced compared to common non-rigid registration algorithms, our approach shows potential for practical applications such as our intended application of motion compensated PET reconstruction within hybrid PET-MRI imaging.

## 5. Acknowledgement

The work presented in this paper is part of the Hybrid PET-MRI image project (HyperImage, http://p101002.typo3server.info/index.php?id=hybrid-pet-mr) which is supported by the European Union under the 7th framework program (201651).

## 6. Vitae

**Christian Buerger** (Feb 1981) received the degree "Diplom-Ingenieur" in Electrical Engineering and Information Technology from the RWTH Aachen University, Germany, in 2008. He joined the Division of Imaging Sciences, Kings College London, in 2008 as a PhD student. His work concentrates on a combination of image acquisition and image registration techniques for optimal motion field estimations from 4D-MRI.

**Andrew P. King** (Sept 1967) received a BSc. (Hons) degree in Computer Science from Manchester University in 1989, an MSc. (with distinction) in Cognition, Computing and Psychology from Warwick University in 1992, and a PhD degree in Computer Science from Warwick University in 1997. He has seven years experience of postdoctoral research with the Computational Imaging Sciences Group and the Division of Imaging Sciences at King's College London, working mainly on registration, image-guided interventions and soft-tissue modelling. From 2001-2005 he worked as an Assistant Professor in the Computer Science department at Mekelle University in Northern Ethiopia.

**Tobias Schaeffter** (Sept 1967) received the degree "Diplom-Ingenieur" in Electrical Engineering from the Technical University Berlin in 1993 and did his PhD in MRI at University Bremen (Prof. Leibfritz) in 1996. Afterwards, he worked ten years as a Principal Scientist at the Philips Research Laboratories in Hamburg. In April 2006, he took up the post as the Philip Harris Professor of Imaging Sciences at King's College London. The major direction of his research is the investigation of new acquisition and reconstruction techniques for cardiovascular and quantitative MRI. Furthermore, he works on multi-modality imaging approaches for interventional procedures.

## 7. Download

A download link and usage instructions for our registration algorithm are available at https://www.isd.kcl.ac.uk/internal/hyperimage/.

## 8. References


Andronache, A., Cattin, P., Szekely, G., 2005. Adaptive subdivision for hierarchical non-rigid registration of multi-modal images using mutual information, in: Proc. of MICCAI.

Andronache, A., von Siebenthal, M., Szkely, G., Cattin, P., 2008. Non-rigid registration of multi-modal images using both mutual information and cross-correlation. Medical Image Analysis 12, 3–15.

Arsigny, V., Commowick, O., Ayache, N., Pennec, X., 2009. A fast and log-euclidean polyaffine framework for locally linear registration. Journal of Mathematical Imaging and Vision 33, 222–238.

Arsigny, V., Pennec, X., Ayache, N., 2005. Polyrigid and polyaffine transformations: A novel geometrical tool to deal with non-rigid deformations - application to the registration of histological slices. Medical Image Analysis 9, 507–523.

Bookstein, F., 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11, 567–585.

Buerger, C., Schaeffter, T., King, A., 2010. Hierarchical adaptive local affine registration for respiratory motion estimation from 3-D MRI, in: Proc. of International Symposium on Biomedical Imaging - ISBI.

Cao, K., Christensen, G., Ding, K., Reinhardt, J., 2009. Intensity-and-landmark-driven, inverse consistent, b-spline registration and analysis for lung imagery, in: The Second International Workshop on Pulmonary Image Analysis, Proc. of Medical Image Computing and Computer-Assisted Intervention.

Castillo, E., Castillo, R., Martinez, J., Shenoy, M., Guerrero, T., 2010. Four-dimensional deformable image registration using trajectory modeling. Physics in Medicine and Biology 55, 305–327.

Choi, Y., Lee, S., 2000. Injectivity conditions of 2d and 3d uniform cubic b-spline functions. Graphical Models 62, 411–427.

Commowick, O., Arsigny, V., Isambert, A., Costa, J., Dhermain, F., Bidault, F., Bondiau, P.Y., Ayache, N., Malandain, G., 2008. An efficient locally affine framework for the smooth registration of anatomical structures. Medical Image Analysis 12, 427–441.

Crum, W., Tanner, C., Hawkes, D., 2005. Anisotropic multi-scale fluid registration: evaluation in magnetic resonance breast imaging. Physics in Medicine and Biology 50, 5153–5174.

Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistican Society 39, 1–38.

Gee, J., Bajcsy, R., 1999. Elastic matching: Continuum mechanical and probabilistic analysis, in: Proc. of Brain Warping, Academic Press. pp. 183–197.

Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N., 2008. Dense image registration through MRFs and efficient linear programming. Medical Image Analysis 12, 731–741.

Guerrero, T., Sanders, K., Castillo, E., Zhang, Y., Bidaut, L., Pan, T., Komaki, R., 2006. Dynamic ventilation imaging from four-dimensional computed tomography. Physics in Medicine and Biology 51, 777–791.

Hagenlocker, M., Fujimura, K., 1998. CFFD: a tool for designing flexible shapes. The Visual Computer 14, 271–287.

Horn, B., Schunk, B., 1981. Determining optical flow. Artificial Intelligence 17, 185–203.

King, A., Boubertakh, R., Rhode, K., Ma, Y., Chinchapatnam, P., Gao, G., Tangcharoen, T., Ginks, M., Cooklin, M., Gill, J., Hawkes, D., Razavi, R., Schaeffter, T., 2009. A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. Medical Image Analysis 13, 419–431.

Lee, S., Wolberg, G., Shin, S., 1997. Scattered data interpolation with multilevel B-splines. IEEE Transactions on Visualization and Computer Graphics 3, 228–244.

Likar, B., Pernus, F., 1999. Registration of serial transverse sections of muscle fibers. Cytometry 37, 93–106.

Likar, B., Pernus, F., 2001. A hierarchical approach to elastic registration based on mutual information. Image and Vision Computing 19, 33–44.

Little, J., Hill, D., Hawkes, D., 1996. Deformations incorporating rigid structures, in: Proc. of the Workshop on Mathematical Methods in Biomedical Image Analysis, pp. 104–113.

Manke, D., Nehrke, K., Boernert, P., 2003. Novel prospective respiratory motion correction approach for free-breathing coronary MR angiography using a patient-adapted affine motion model. Magnetic Resonance in Medicine 50, 122–131.

Martin-Fernandez, M., Martin-Fernandez, M., Alberola-Lopez, C., 2009. A log-euclidean polyaffine registration for articulated structures in medical images, in: Proc. of MICCAI, pp. 156–164.

McClelland, J., Blackall, J., Tarte, S., Chandler, A., Hughes, S., Ahmad, S., Landau, D., Hawkes, D., 2006. A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. Medical Physics 33, 3348–3358.

Modat, M., Ridgway, G., Taylor, Z., Lehmann, M., Barnes, J., Hawkes, D., Fox, N., Ourselin, S., 2009. Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine In Press.

Murphy, K., van Ginneken, B., Reinhardt, J., Kabus, S., Ding, K., Deng, X., Pluim, J., 2010. Evaluation of methods for pulmonary image registration: The EMPIRE10 study. Grand Challenges in Medical Image Analysis 2010, 11–22.

Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. IEEE Transactions On Pattern Analysis And Machine Intelligence 12, 629–639.

Reinhardt, J., Christensen, G., Hoffman, E., Ding, K., Cao, K., 2007. Registration-derived estimates of local lung expansion as surrogates for regional ventilation, in: Proc. of Information Processing in Medical Imaging, pp. 763–774.

Rohr, K., 2000. Elastic registration of multimodal medical images: a survey. KI-Kuenstliche Intelligenz , 11–17.

Rueckert, D., Aljabar, P., Heckemann, R., Hajnal, J., Hammers, A., 2006. Diffeomorphic registration using B-splines, in: Proc. of Medical Image Computing and Computer-Assisted Intervention, pp. 702–709.

Rueckert, D., Sonoda, L., Hazes, C., Hill, D., Leach, M., Hawkes, D., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging 18, 712–721.

Sundarama, T., Gee, J., 2005. Towards a model of lung biomechanics: pulmonary kinematics via registration of serial lung images. Medical Image Analysis 9, 524–537.

Vandemeulebroucke, J., Sarrut, D., P., Clarysse, 2007. The popi-model, a point-validated pixel-based breathing thorax model, in: Proc of the XVth International Conference on the Use of Computers in Radiation Therapy (ICCR).

Zhuang, X., Rhode, K., Arridge, S., Razavi, R., Hill, D., Hawkes, D., Ourselin, S., 2008. An atlas-based segmentation propagation framework using locally affine registration - application to automatic whole heart segmentation, in: Proc. of Medical Image Computing and Computer-Assisted Intervention, pp. 425–433.