



Temporal sparse free-form deformations

Wenzhe Shi^a, Martin Jantsch^a, Paul Aljabar^{c,a}, Luis Pizarro^{a,e}, Wenjia Bai^a, Haiyan Wang^a, Declan O'Regan^d, Xiahai Zhuang^{b,*}, Daniel Rueckert^{a,*}

^a Biomedical Image Analysis Group, Imperial College London, UK

^b Shanghai Advanced Research Institute, Chinese Academy of Sciences, China

^c Imaging Sciences and Biomedical Engineering, King's College London, UK

^d Institute of Clinical Science, Imperial College London, UK

^e Escuela de Ingeniera Informtica, Universidad Diego Portales, Chile

ARTICLE INFO

Article history:

Available online 16 May 2013

Keywords:

Free-form deformation

Sparse

Registration

Cardiac

Temporal

ABSTRACT

FFD represent a widely used model for the non-rigid registration of medical images. The balance between robustness to noise and accuracy in modelling localised motion is typically controlled by the control point grid spacing and the amount of regularisation. More recently, TFFD have been proposed which extend the FFD approach in order to recover smooth motion from temporal image sequences. In this paper, we revisit the classic FFD approach and propose a sparse representation using the principles of compressed sensing. The sparse representation can model both global and local motion accurately and robustly. We view the registration as a deformation reconstruction problem. The deformation is reconstructed from a pair of images (or image sequences) with a sparsity constraint applied to the parametric space. Specifically, we introduce sparsity into the deformation via L_1 regularisation, and apply a bending energy regularisation between neighbouring control points within each level to encourage a grouped sparse solution. We further extend the sparsity constraint to the temporal domain and propose a TSFFD which can capture fine local details such as motion discontinuities in both space and time without sacrificing robustness. We demonstrate the capabilities of the proposed framework to accurately estimate deformations in dynamic 2D and 3D image sequences. Compared to the classic FFD and TFFD approach, a significant increase in registration accuracy can be observed in natural images as well as in cardiac images.

Crown Copyright © 2013 Published by Elsevier B.V. All rights reserved.

1. Introduction

Image registration is one of the fundamental tasks in medical image analysis. The classic FFD registration approach (Rueckert et al., 1999) is widely used for medical image registration. Several improvements of the method have been proposed in the past, including a diffeomorphic variant of FFD to generate biologically more plausible deformations (Rueckert et al., 2006; De Craene et al., 2012) which ensures a one-to-one mapping; a TFFD model to allow 4D deformations (De Craene et al., 2010; Metz et al., 2011); an analytical expression for the mutual information based similarity metric to accelerate the optimisation (Modat et al., 2010) as well as different optimisation strategies (Glocker et al., 2008; Klein et al., 2010). Despite its popularity, little effort has been devoted to improve the accuracy of the formulation compared to other registration methods such as optical flow (Horn and Schunck, 1981; Mémin and Pérez, 2002; Roth and Black,

2005; Sun et al., 2010) and the Demons approach (Thirion, 1998; Vercauteren et al., 2008).

In this paper, we address one main difficulty of the classic FFD approach, namely the conflict between the robustness of the registration and the ability to model highly-localised and potentially discontinuous deformations. We refer to accuracy as the ability to reconstruct the highly-localised and potentially discontinuous deformation with as little error as possible and the robustness as the ability to recover and reconstruct such deformation in the presence of the noise in the images. The trade-off between accuracy and robustness stems from the fact that the FFD approach uses a smooth B-spline basis to model the contribution of each control point to the deformation. To model global and smooth deformations a coarse control point spacing is typically used. To allow very localised deformations a finer control point spacing is required. However, this can render the FFD-based registration less robust as the model has far more degrees of freedom which must be optimised. A conventional approach to address this issue uses a coarse-to-fine approach in which the initial coarse control point mesh is successively subdivided to produce finer control point meshes (Rueckert et al., 1999).

* Corresponding authors.

E-mail addresses: zhuangxh@sari.ac.cn (X. Zhuang), d.rueckert@imperial.ac.uk (D. Rueckert).

The standard smoothness constraints for non-rigid registration methods (Rueckert et al., 1999; Horn and Schunck, 1981; Thirion, 1998) assume that the deformation within a neighbourhood changes only gradually since the underlying deformation itself is smooth. Combining the implicit smoothness of the B-spline basis and the explicit smoothness constraint in the regularisation, leads to FFD-based registration results with smooth deformations.

As mentioned above, the control point grid spacing has a significant impact on the ability to capture motion discontinuities robustly. Previous research focussed on the adaptive parameterisation of the B-spline control point grid (Schnabel et al., 2001; Rohlfing and Maurer, 2001; Hansen et al., 2008) that has been driven by the intensity information in the images. In addition, Kumar et al. (2006) proposed to use shape prior information and a shape dependent basis function to account for boundaries between foreground and background. An improved model should enable more control points to be placed in an area in which more flexibility for the modelling of deformations is required. Xie and Farin (2004) proposed to use a distance measure to determine the need for localised motion after initial registration using a coarse control point grid spacing.

Many other approaches to image registration have been proposed that aim to overcome the conflict between robustness and accuracy in the motion estimation, in particular in the field of optical flow estimation (Mémmin and Pérez, 2002; Roth and Black, 2005; Sun et al., 2010). More recently, sparse coding methods have been proposed to evaluate the patch similarity between two images (Roozgard et al., 2011) and to constrain the transformation (Shen and Wu, 2010). We address the trade-off between robustness and accuracy for FFD-based registration by using sparsity constraints as an additional regularization term.

FFD-based methods are now one of the most widely used methods in recent cardiac motion tracking challenges (Tobon-Gomez et al., 2012; De Craene et al., 2013; Piella et al., 2013; Heyde et al., 2013). Among these approaches TFFD-based methods either based on velocity fields (De Craene et al., 2010; De Craene et al., 2012; Piella et al., 2013) or based on displacement fields (Metz et al., 2011) have attracted increasing attention due to their ability to extract consistent temporal motion pattern, which is important for estimation of cyclic motion such as the cardiac motion.

1.1. Overview and contributions

In this paper, we introduce a sparse representation for FFD to estimate the registration transformation. Our approach is inspired by the work in Roozgard et al. (2011) and Shen and Wu (2010). Our proposed model uses standard smoothness constraints and only imposes one additional assumption on the conventional FFD, namely that the deformation is sparse in the parametric space. Thus we call our method SFFD. The assumption of sparsity of the deformation is generally true because the deformation between images is usually less discontinuous than the actual images themselves. In this work, we use a multi-level FFD model to represent the deformations in a parametric form. In this multi-level FFD each level consists of a B-spline control point mesh with increasingly finer resolution. As can be seen from Fig. 3, a dense motion field can be represented by a sparse multi-level FFD in parametric space. Based on this assumption, we formulate the registration of two images using a sparse multi-level FFD representation of the control points. We introduce a regularisation term to impose smoothness at each level and a sparsity term to enforce coupled multi-level sparsity. A preliminary version of this approach was published at MICCAI 2012 (Shi et al., 2012a).

In addition to the approach initially published in (Shi et al., 2012a), we extend the SFFD to the TSFFD based on the periodic TFFD (Metz et al., 2011) and apply it to the cardiac motion estimation

problem. TFFD refers to the implementation of (Metz et al., 2011) in the following sections. The extension to 4D enables us to better model the cyclic motion. It is well known that the diastasis and atrial systole phases, which are the final two phases of the diastole, occupy a disproportionately large time window in the cardiac cycle (Bray, 1999). Left ventricular myocardium undergoes little motion under these two phases. This means that the contraction of the heart occurs in a relatively short time. Taken that into account, the sparsity assumption is true in both the temporal and spatial domain in the context of the application to cardiac motion estimation.

The novelty and contributions of this paper are the introduction of a sparsity model that reduces the influence of the *a priori* selection of an appropriate control point grid spacing. Furthermore, the approach reduces the conflict between global smoothness (robustness) and the local level of detail of the transformation (accuracy) by optimising the different levels of the FFD coarser than the image resolution level simultaneously with a sparsity constraint. These advantages allow the robust estimation of deformation fields in the presence of highly localised or discontinuous deformations in both time and space. We refer to this new approach as SFFD in 2D/3D and TSFFD in 4D. In the evaluation, we demonstrate that the proposed method can consistently capture all motion with high accuracy in both 2D + t cases and 3D + t cases.

2. Datasets

In this work, we have evaluated the proposed SFFD and TSFFD against the classic FFD and TFFD model on four different datasets. The datasets we have used for evaluation include:

- The Middlebury benchmark dataset (Baker et al., 2007) (which is a standard dataset widely used in computer vision);
- 2D cardiac MR images with synthetic smooth motion. In the following this dataset will be referred to as *Cardiac A*;
- 2D cardiac MR images with synthetic discontinuous motion. In the following this dataset will be referred to as *Cardiac B*;
- 10 synthetic but highly realistic 3D cardiac ultrasound image sequences which have also been used in a recent cardiac motion tracking challenge (De Craene et al., 2013). In the following this dataset will be referred to as *Cardiac C*.

Examples of the datasets are presented in Fig. 2. For basic benchmarking we have used greyscale images with two frames only from the Middlebury benchmark dataset (Baker et al., 2007) for which the ground-truth flow is available. The image sets are the RubberWhale, Hydragea, Dimetrodon, Grove2, Grove3 and Venus. The Middlebury benchmark dataset contains deformations with multiple independently moving rigid objects and background. For this dataset the ground truth deformation between each pair of images is available. The ground truth is derived from the camera position and parameters. There is a relatively large diversity in motion discontinuities present in this dataset. Due to the above reasons, it is used as a primary benchmark to test and develop the SFFD registration algorithm. However, Middlebury benchmark has certain characteristics which do not hold in medical imaging applications. Thus, we use three additional synthetic medical image datasets with dense motion ground truth to further evaluate the performance of the methods.

In addition, we have tested our approach using 2D cardiac MR images. Synthetic deformation fields have been generated as proposed in Pizarro et al. (2011), with displacements in each segment following sinusoidal waves with gradually changing amplitudes, wavelengths and phases in both x- and y-direction. This produces smooth deformation fields with discontinuities between different

segments. We generated two groups of 10 synthetic datasets each. The first group (*Cardiac A*) contains smooth motions created with a single sinusoidal function over the whole image domain. The second group (*Cardiac B*) contains discontinuous motions created with multiple sinusoidal functions fused together at different regions of the images. An example of this can be seen in Fig. 1b.

We have also used 10 synthetic cardiac ultrasound image sequences from the second cardiac motion tracking challenge De Craene et al. (2013) (*Cardiac C*). The synthetic image sequences proposed in the challenge combine an electro-mechanical model provided by Sermesant et al. (2012) with an ultrasound imaging model from Gao et al. (2009). The challenge provides 10 highly realistic sequences spanning different values of the global conductivity, global contractility, and electrical delays parameters. In the dataset, a single ultrasound probe design was considered. Scatterers were randomly placed in the myocardial geometry and moved along the cardiac cycle according to the result of the mechanical simulation. Details of the dataset can be found in De Craene et al. (2013). The image resolution is approximately $(0.34 \times 0.34 \times 0.34 \text{ mm})$ and the image sequence contains 23 frames. Ground truth deformation is provided as a series of volumetric meshes. The nodes of the meshes are landmarks covering the myocardium and give the ground truth deformation.

In all datasets, the accuracy is measured by the mean and SD of the distance between the ground truth motion and the reconstructed motion. In the dataset *Cardiac C*, we only evaluate the accuracy in points inside the imaging area.

3. Classic free-form deformation model

In the classic FFD registration (Rueckert et al., 1999), a non-rigid deformation $\mathbf{h} = [XYZ]^T$ is represented using a B-spline model in which the deformation is parameterised using a set of control points $\Phi = [UVW]^T$ such that

$$\mathbf{h} = \begin{bmatrix} B & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & B \end{bmatrix} \Phi, \quad (1)$$

where B denotes the matrix of the B-spline basis functions. The B-spline basis function is a 3D-tensor product of 1D B-splines. Details can be found in Appendix A.

To find the optimal deformation between two images, the registration minimises an energy functional E written as a function of Φ , which is typically a combination of two terms $E(\Phi) := E_D(I_r, I_s \circ \mathbf{h}) + E_R(\Phi)$. The term E_D is a data constraint measuring the similarity between the target image I_r and the transformed source image $I_s \circ \mathbf{h}$. The term E_R is a regularisation constraint that enforces a smooth transformation. In this classic FFD approach, the energy function is typically minimised using gradient descent approaches (Modat et al., 2010; Klein et al., 2010) or discrete optimisation approaches (Glocker et al., 2008).

4. Sparse free-form deformation model

To be able to deal with large, global deformations and to improve the robustness, the classic FFD registration uses a multi-level approach: First, the optimal registration parameters are determined for a control point grid with large spacing. The grid is then successively subdivided to capture local deformations (Rueckert et al., 1999). This requires an a priori choice of the control point grid multi-resolution schedule. Furthermore, each level is optimised separately and once a level has been optimised it is no longer updated, leading to suboptimal registration results as can be seen in Fig. 4. It was suggested in Shen and Wu (2010) that a realistic transformation can be easily embedded into a sparse representation. We postulate that an automatic selection of control points across different levels can be achieved by optimising all FFD levels simultaneously while using a sparsity constraint.

4.1. Sparse representation of transformation

In this section, we propose estimating the displacement \mathbf{h} with a sparse representation of the control points Φ . We use a multi-level FFD representation (Schnabel et al., 2001), $\Phi = [U^1 \dots U^m V^1 \dots V^m W^1 \dots W^m]^T$ where m denotes the number of levels, as it is well suited for sparse representations. Accordingly, we utilise a multi-level B-splines basis $B = [B^1 \dots B^m]$. The

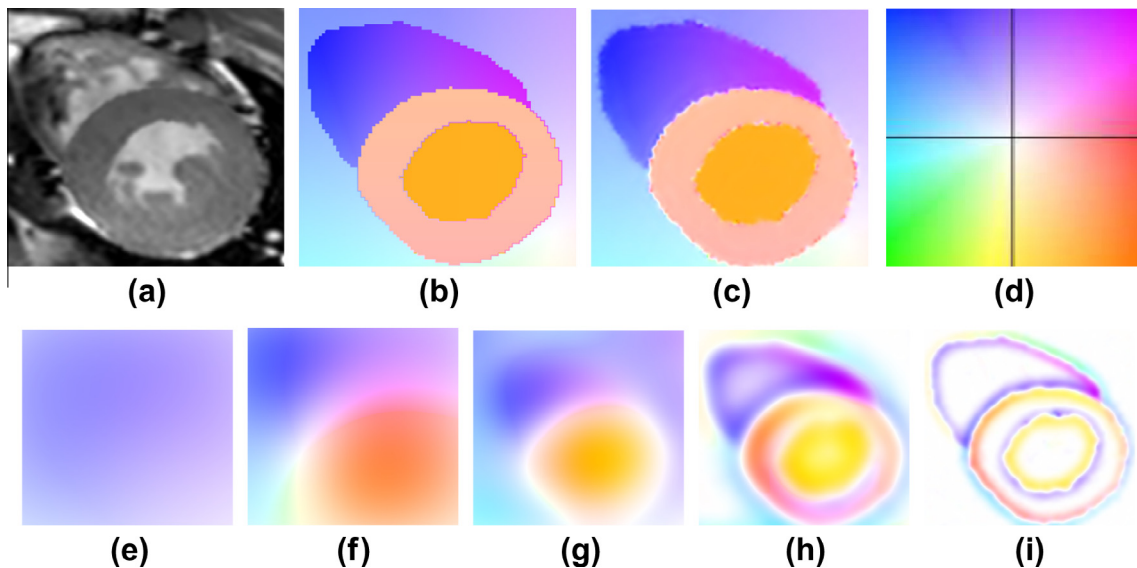


Fig. 1. Visualisation of the sparse representation of a dense FFD using the colour scheme from (Baker et al., 2007): (a) Target image; (b) synthetic ground truth motion; (c) reconstructed motion as the sum of (e) to (i); (d) colour scheme from (Baker et al., 2007); and (e–i) motion from FFD of different levels (coarse to fine). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

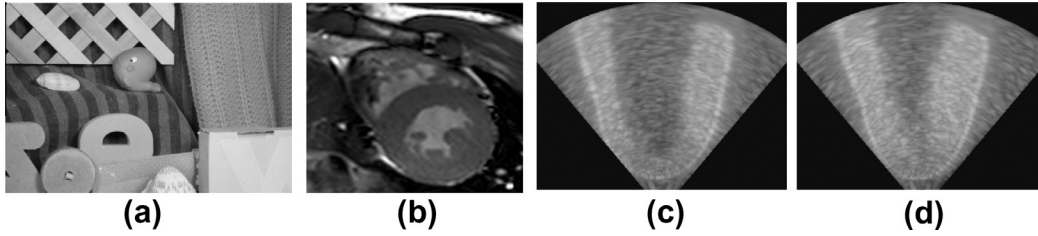


Fig. 2. This figure shows the example images of the datasets. (a) Image from the Middlebury dataset; (b) 2D cardiac MR image; (c) 3D cardiac ultrasound image at end diastolic phase; and (d) 3D cardiac ultrasound image at end systolic phase.

displacement \mathbf{h} is computed as in Eq. (1) with the above redefinitions of Φ and B using the following equation:

$$\mathbf{h} = \begin{bmatrix} B^1 \dots B^m & 0 & 0 \\ 0 & B^1 \dots B^m & 0 \\ 0 & 0 & B^1 \dots B^m \end{bmatrix} \Phi, \quad (2)$$

The multi-level FFD is illustrated in Fig. 1. Our assumption is that a typical FFD with dense displacement \mathbf{h} can be sparse in its parametric representation Φ . The control points at the finest level are only activated around the motion boundaries. The sparsity of the parametric representation can be confirmed by Fig. 3 where the histogram of the coefficients in the parametric space is plotted.

Basis pursuit denoising (Donoho and Huo, 2001) is a mathematical optimisation that balances the trade-off between sparsity and reconstruction fidelity. In the context of image registration, the problem can be formulated as:

$$\arg \min_{\Phi} E(\Phi) := \|I_r - I_s \circ \mathbf{h}\|_2^2 + \|\Phi\|_1, \quad (3)$$

Here the first term corresponds to the SSD between the target image and the transformed source image and acts as similarity measure. The second term enforces the sparsity of the solution Φ by using the L_1 -norm. The L_1 -norm is used to enforce sparsity because it is convex and has many favourable theoretical properties

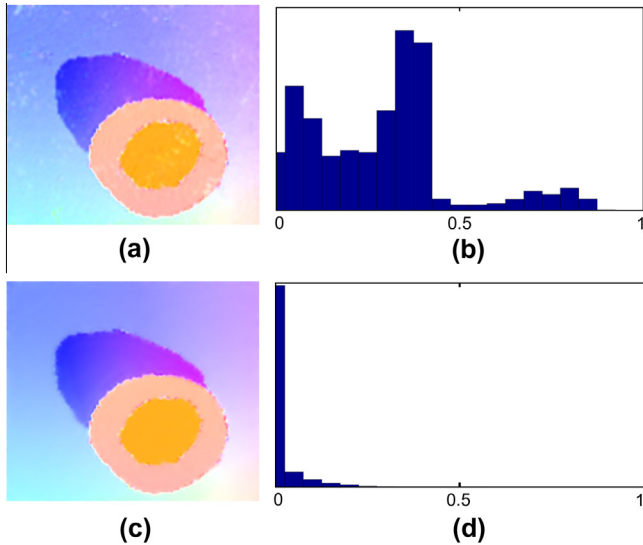


Fig. 3. This figure shows the histogram of coefficients in parametric space from the result of the classic FFD and the SFFD. The result are obtained from the best reconstructed motion from different methods of the Cardiac B dataset. The ground truth is shown in Fig. 1b. (a) Reconstructed motion from FFD; (b) histogram of the magnitude of the coefficients from FFD; (c) reconstructed motion from SFFD (same as Fig. 1c); and (d) histogram of the magnitude of the coefficients from SFFD. For the histograms, the horizontal axis shows the control point value in mm and the vertical axis shows the number of control points with that binned control point value. The width of the bins is 0.05 mm.

(Donoho and Huo, 2001; Roozgard et al., 2011; Shen and Wu, 2010). In general, an arbitrary (dis) similarity measure can be utilised in the data term $E_D(I_r, I_s \circ \mathbf{h})$, including information theoretic measures such as MI or its normalised counterparts NMI (Studholme et al., 1999).

Following these principles, we formulate a novel registration approach, namely the SFFD model, as

$$\arg \min_{\Phi} E(\Phi) := E_D(I_r, I_s \circ \mathbf{h}) + \lambda_R \sum_{i \in [0, m]} E_R(\Phi^i) + \lambda_S \|\Phi\|_1, \quad (4)$$

with constants $\lambda_R, \lambda_S \in \mathbb{R}^+$ weighting the regularisation term and the sparsity term, respectively. Note that the regularisation term imposes smoothness at each level of the multi-level FFD independently, while the sparsity term enforces coupled multi-level sparsity. This allows us to actively determine the importance of the control points across all levels in a joint manner, not independently as in the classic FFD framework. In Yuan and Lin (2005), the authors proposed a grouped LASSO based on a voxel-based sparse classifier using L_1 -norm regularised linear regression model (Tibshirani, 1996). Friedman et al. (2010) extended the grouped LASSO to sparse grouped LASSO by using both L_1 and L_2 terms. This yields sparsity at both the group and individual feature levels. In our work, similarly, the L_1 -norm regularisation encourages sparsity and the smoothness regularisation at each level in the parametric space encourages grouped zero and non-zero components. The combination of both terms encourages the sparse non-zero components to be coupled and results in a grouping effect.

The above strategy can be used to estimate deformation fields robustly and to preserve motion discontinuities, as it will be seen in the experimental validation. We optimise Eq. (4) using the interior point method of Kim et al. (2007) that uses a log barrier function to make the sparsity term differentiable. The parameter λ_S is normalised between the data and the sparsity terms using the finite convergence to zero property. That is, for the L_1 -regularised least squares problem, convergence is achieved for a finite value λ_{max} of λ_S . The value of λ_{max} can be determined using Eq. (4) in Kim et al. (2007). In our experiments, we use:

$$\lambda_S^N = \lambda_S / \lambda_{max}, \quad (5)$$

where λ_S^N denotes the normalised sparsity parameter with respect to the finite value λ_{max} . λ_S^N is a user specified value while λ_{max} and λ_S are automatically determined by the optimisation.

In addition, we use a preconditioning scheme for improving the efficiency of the optimisation of the registration problem as proposed in Zikic et al. (2011). The preconditioner is theoretically justified especially for high-dimensional registration problems in Zikic et al. (2011). In our case, it avoids that the gradient terms corresponding to the coarse resolution control points dominate the optimisation and reduce the influence of the image gradient. For completeness, the reader will find the derivation of the derivatives of the similarity measures and the regularisation terms with respect to ϕ_i in Appendix A.

5. Temporal sparse free-form deformation model

To extend the classic FFD registration method, as described in Section 3 to model the cyclic deformation $h = [X_0 \dots X_{N_t} \ Y_0 \dots Y_{N_t} \ Z_0 \dots Z_{N_t}]^T$ in an image sequence with N_t temporal frames denoted by I_l^i for $l = 1, \dots, N_t$, we construct a set of 4D control points with coefficients $\Phi = [U_0^1 \dots U_{t_1}^1 \dots U_0^m \dots U_{t_m}^m \ V_0^1 \dots V_{t_1}^1 \dots V_0^m \dots V_{t_m}^m \ W_0^1 \dots W_{t_1}^1 \dots W_0^m \dots W_{t_m}^m]^T$. We have m 4D control point meshes, and each one with temporal resolution t_k . The first and last spatial control points are defined to be direct neighbours to enforce cyclic motion (Metz et al., 2011). When registering cardiac sequences the periodicity of the model ensures a temporally smooth motion. For other cases this neighbourhood definition can be easily modified. This lead to the following equation:

$$\mathbf{h} = \begin{bmatrix} B_0^1 \dots B_{t_1}^1 \dots B_0^m \dots B_{t_m}^m & 0 & 0 \\ 0 & B_0^1 \dots B_{t_1}^1 \dots B_0^m \dots B_{t_m}^m & 0 \\ 0 & 0 & B_0^1 \dots B_{t_1}^1 \dots B_0^m \dots B_{t_m}^m \end{bmatrix} \Phi, \quad (6)$$

The matrix B in Eq. (6) now consists of B-spline basis functions defined by a 4D-tensor product. Note that in this definition, there is no component of the deformation in the temporal direction. This means we use a 4D mesh of control points and each control point is a 3D displacement vector. There are cases where this might not be true, e.g. because of temporal misalignment due to variability in the cardiac cycle. In these cases, the approach of (Perperidis et al., 2005) can be used to perform an interleaved spatial and temporal optimisation to correct 4D displacement field.

The TFFD registration method also uses a multi-level approach as described in Section 4, but with a constant temporal spacing. However, to make use of the temporal sparsity of the motion, our TSFFD method uses multiple grid spacings for all four dimensions. To be able to make use of the full parametric space, we need to subdivide the spatial and temporal grid spacing independently. Then combinations between all spatial resolutions and every temporal resolution should be used. This is necessary because the cardiac contraction occurs during a very short period of the cardiac cycle and there might be some temporally consistent but spatially discontinuous motion (for example the sliding motion that occurs at the pericardium) over the whole cardiac cycle. However, in practice, it is not computational feasible and very memory intensive. Thus, we choose to subdivide the spatial control point mesh between levels.

The definition of the control points Φ and the B-spline basis matrix for the TFFD model is straightforward. Following the defini-

tion in Section 4.1 the new energy function for the TSFFD model can be written as

$$\arg \min_{\Phi} E(\Phi) := \sum_{l \in [1, N_t]} E_D(I_l^i, I_s \circ \mathbf{h}) + \lambda_R \sum_{l \in [0, t]} \sum_{i \in [1, m]} E_R(\Phi^{i,l}) + \lambda_S \|\Phi\|_1, \quad (7)$$

where I_s is a chosen reference frame in the sequence. We are using a cyclic definition of the control points so any frame may be chosen as reference frame. As in the SFFD method, the first regularisation imposes smoothness on each level independently. But this explicit smoothness constraint only affects the spatial component while the L_1 term enforces 4D multi-level sparsity. The optimisation of Eq. (7) is performed in the same way as for the SFFD method.

6. Results

6.1. Implementation details

During the optimisation of the classic FFD, we use seven different levels of image resolution. In the coarsest level each image is subsampled by a factor of 64 in each dimension. In addition to the control points in the image domain, we use phantom control points outside the image domain to ensure that the free-form deformation is also defined at the boundaries of the image. In the paper, we refer to the control point grid size excluding the phantom control points. We use cubic B-spline interpolation for the subsampling as evidence suggests that cubic B-spline based interpolation is superior to linear interpolation (Sun et al., 2010; Modat et al., 2010). We use SSD as the similarity measure. Before optimisation, we normalise the SSD to range between 0 and 1. The Middlebury benchmark is an optical flow based benchmark where the assumption of constant intensity generally holds. The assumption holds for the 2D cardiac MR synthetic dataset (*Cardiac A* and *Cardiac B*) since no noise is introduced in our case. The constant intensity assumption makes SSD a preferable candidate as the similarity measure in these cases.

For the 3D cardiac ultrasound image sequences (*Cardiac C*), we are modelling the motion within a single sequence. Although the presence of speckle noise in the ultrasound image sequence invalidates the constant intensity assumption, we have shown in our experiment that by simply using a coarser finest grid spacing SSD can still perform well.

One of the most crucial parameters of the classic FFD registration is the control point grid spacing. We create a spacing at each image pyramid level by subdividing the FFD from the previous

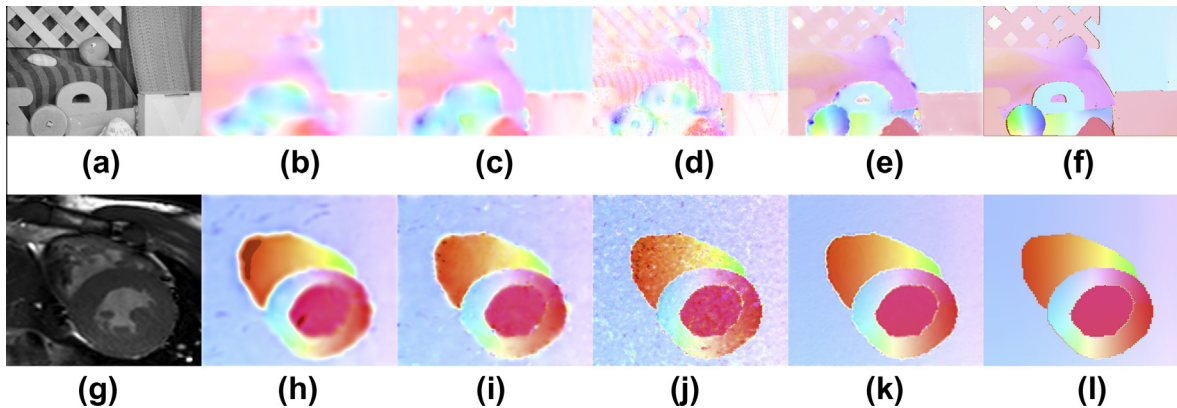


Fig. 4. Visual comparison between the classic FFD and the proposed SFFD using the colour scheme from (Baker et al., 2007). The colour range corresponds to different direction and magnitude of the displacement. (a) and (g) Show the rubber whale image from the Middlebury dataset and a frame from the cardiac B experiment, respectively; (f) and (l) display the ground truth transformations with noticeable motion discontinuities; (b–d) and (h–j) show the estimated motion with the classic FFD approach for 512/8 mm, 256/4 mm, and 64/1 mm control point spacing; (e) and (k) exhibit the estimated motion with the proposed SFFD approach where $\lambda_S^N = 0.04$.

level. For the Middlebury dataset as well as *Cardiac A* and *Cardiac B* datasets, we have evaluated different initial control point spacings at the coarsest level varying from 512 mm to 64 mm while the control point spacings at the final level varies from 8 mm to 1 mm. The spacing at the coarser level is subdivided by a factor of 2 to create the finer level. For the SFFD, we use a multi-level FFD with the coarsest level having a control point grid spacing of 64 mm and finest level having a spacing of 1 mm. The number of control points of SFFD at coarsest level is 4×4 . During the optimisation, SFFD control point resolutions are only activated on levels which are coarser than the current image resolution. The coefficient for the smoothness penalty is set to 0.01 as in (Rueckert et al., 1999). The number of image pyramid levels and the smoothness penalty λ_R are the same for both methods.

For cardiac C dataset, we use four different levels of image resolution. The first image of the ED phase is used as the reference image. In the coarsest level each image is subsampled by a factor of sixteen in each dimension. In the finest level, each image is subsampled by a factor of two in each dimension to reduce the computational expense. In this dataset, we use relatively coarser

control point spacing to deal with noise in the speckle pattern of the images. For the TFFD and FFD approaches, we have evaluated different initial control point spacings at the coarsest level varying from 256 mm to 32 mm while the control point spacings at the final level varies from 32 mm to 4 mm. For the TFFD registration, the temporal spacing is set to $1/N_t$ of the whole image sequence where N_t is the number of frames and not subdivided between levels. For the TSFFD and SFFD, we use a multi-level TFFD with the coarsest level having a control point grid with 64 mm spatial spacing and 0.25 temporal spacing and finest level with 2 mm spatial spacing and $1/N_t$ temporal spacing. The number of control points of at coarsest level is $3 \times 3 \times 3$ and $3 \times 3 \times 3 \times 5$ for SFFD and TSFFD respectively. The temporal control point mesh spacing is linearly interpolated from coarsest level to finest level. The remaining parameters are the same as the Cardiac A and Cardiac B datasets.

Finally, we used student *t*-test to test whether there are significant difference between each category with different parameters. The difference between the errors of each category are assumed to follow a Gaussian distribution.

6.2. Effects of the sparsity constraint

In this section, we evaluate the effects of the sparsity constraint in the Middlebury dataset. The Middlebury benchmark provides accurate motion ground truth from real images and contains both discontinuous motion and smooth motion. We first apply the classic FFD approach to the data. It is noticeable in Fig. 5 that the accuracy is different from image to image. From our observation, the optimal initial grid spacing depends highly on the level of motion discontinuity of the data. We applied a comparable SFFD approach with largest grid spacing of 64 mm to the benchmark. The SFFD approach shares the same number of image resolutions, smoothness parameter and similarity measure with the classic FFD. In Fig. 5, the performance of different sparsity constraints is similar across the different images. There is a significant increase in accuracy when the sparsity parameter is first introduced and the accuracy remains stable until around $\lambda_S^N = 0.16$ with a sharp decrease towards the final point of convergence. The curve of the accuracy has no obvious bias towards different degrees of discontinuity in the dataset.

To observe the effect of the sparsity parameter, we plot the L_1 -norm of the SFFD parametric basis against the reconstructed motion displacement in Fig. 6. The L_1 -norm of the parametric basis decreases dramatically when the sparsity parameter is introduced. It finally reaches zero when the finite convergence condition is met. On the other hand, the L_1 -norm of the displacement is stable without significant decrease until the parameter becomes too large, e.g., in the range 0.32–0.64 in Fig. 6. There is in general a difference of a magnitude between the L_1 -norm of the displacement and the L_1 -norm of the parametric basis when the sparsity parameter is introduced. This shows that the displacement of the FFD can be represented by a sparse representation in the parametric space if the sparsity constraint is enforced.

In the SFFD framework, one user-defined parameter is the coarsest grid spacing in the multi-level FFD representation. We have examined the effect of different coarsest grid spacing. The results are presented in Table 1. The registration accuracy increases slightly with grid spacing from 32 mm to 64 mm. It remains constant for initial spacing between 64 mm and 512 mm. Overall, the coarsest spacing does not have a huge impact on the registration accuracy. From our experience, it seems beneficial that the control point's B-spline basis in the coarsest level influence the whole images. This allows the model to capture global smooth non-rigid motion easily. This is because at the coarsest level the local support is in effect a global support and thus allows the description of global motion. This in turn will help to sparsify the

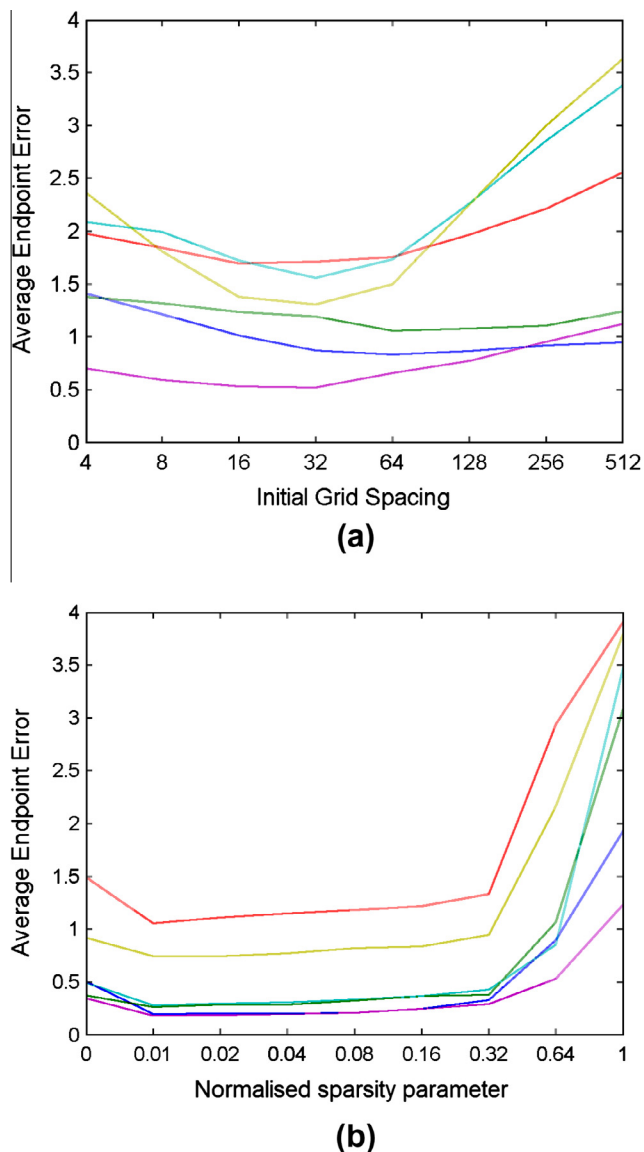


Fig. 5. The accuracy of different parameters. Each line corresponds to a pair of 2D images in the Middlebury benchmark. (a) The influence of different initial grid spacing to the classic FFD. (b) The influence of different normalised sparsity parameter λ_S^N to the SFFD.

transformation at fine levels in terms of coefficients. The finest level is usually in the order of the voxel size in order to capture localised deformation. Finally, the average computational times on Middlebury dataset are 138 s and 147 s for FFD and SFFD respectively.

6.3. Evaluation

In this section, we evaluated our methods using the parameters discussed in Section 6.1. For the classic FFD approach, different spacing leads to significantly different results as shown in Tables 2 and 3. Moreover, different datasets require different initial spacings to achieve the best performance. It can be seen from the tables that the SFFD approach is robust against the choice of λ_s^N compared to the choice of the control point spacing of the FFD. There is little need to adjust λ_s^N across datasets to achieve very good performance for each individual dataset. The result suggests that a $\lambda_s^N \approx 0.04$ yields consistently good performance. SFFD with $\lambda_s^N = 0$ gives

slightly different results from the FFD. This is due to different optimisation strategies: The SFFD is optimised simultaneously across all levels while the FFD is optimised using a coarse to fine strategy.

The mean accuracy results of the SFFD approach without sparsity constraint (where $\lambda_s^N = 0$) is consistently worse than results with different degrees of sparsity constraint. Thus, in all cases the approach benefits from the sparsity constraint. Moreover, the SFFD framework exhibits a significant improvement against the best results from the classic FFD method. An increasing ability to capture local and discontinuous motion while maintaining robustness over smooth regions can be also observed in the visual comparison in Fig. 4.

Finally, in the 3D cardiac ultrasound image sequences, we evaluated TSFFD framework against TFFD. The results are presented in Table 4. The average computational times on this dataset are 68 min, 120 min, 249 min and 584 min for FFD, SFFD, TFFD and TSFFD respectively over the whole cardiac cycle. We observe an improvement from the TFFD approach against the FFD, 0.65 mm vs 1.07 mm, in the results. We believe this improvement stems from the additional temporal consistency and the constraints of periodicity. Second, the TSFFD approach performed significantly better than TFFD, 0.46 mm vs 0.65 mm. Fig. 7 shows the registration errors for case 1 at frame 8 where the error peaks. In Table 5, we presented the result from TSFFD case by case. The results are comparable to the best results in the recent cardiac motion tracking challenge (Piella et al., 2013; Heyde et al., 2013; Somphone et al., 2013; Alessandrini et al., 2013). It is noticeable that in this paper, we did not take any specific measure to explicitly deal with the geometry of the heart as in (Heyde et al., 2013) or the speckle pattern as in (Piella et al., 2013). In addition, our result requires no myocardial segmentation in contrast to (Somphone et al., 2013) and estimated motion in both LV and RV compare to (Heyde et al., 2013).

7. Discussion and conclusions

In this paper, we have developed a SFFD model for registration which addresses some of the most important short-comings of the original FFD registration model. Control points across different FFD levels are optimised simultaneously using a sparse representation. Compared to the classic FFD, the SFFD requires less parameter tuning across different datasets. The user no longer needs to choose an appropriate control point spacing a priori. Our experiments have shown a consistent improvement compared to the original FFD approach when motion ground truth is available.

We have extended the proposed method to a TSFFD model and evaluated our method on the synthetic but highly realistic 4D cardiac ultrasound images (De Craene et al., 2013) from the MICCAI 2012 cardiac motion tracking challenge. Compared to the TFFD and other challengers, our method demonstrated a significantly better tracking accuracy.

During this work, we constrain the sparsity of the coefficients in the parametric space by using L_1 norm as an approximation to the L_0 norm. The advantage of this approach is that it can be easily optimised. The disadvantage is that it penalises large control points values. This might be problematic when trying to recover large motions. For the TSFFD, during the optimisation, we coupled the levels between spatial and temporal control points, e.g., the temporal and spatial resolution increases simultaneously from coarse to fine levels. This is not desirable but as a compromise in terms of the computational cost. Finally, due to the analytical optimisation (Modat et al., 2010), the current framework only works with SSD and NMI. It might be difficult to extend our method to other similarity measures efficiently, for example cross-correlation. Future works will address the above problems. We have achieved a grouped sparse solution using both the sparsity and smoothness

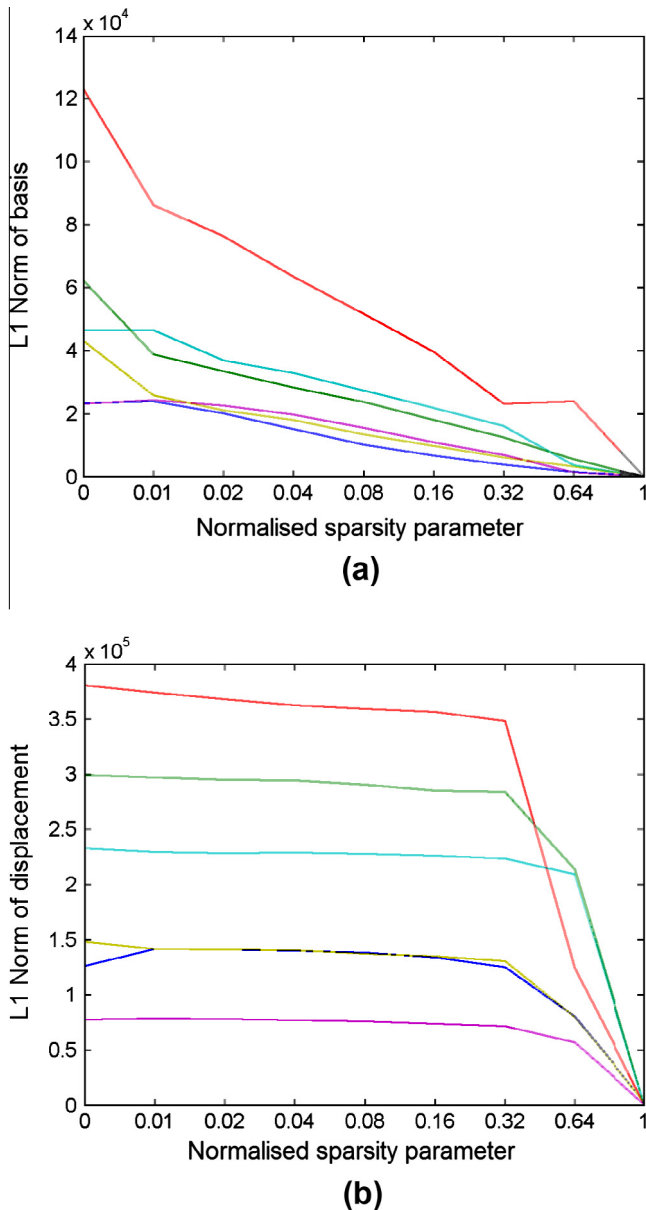


Fig. 6. (a) The L_1 -norm of the parametric basis of the SFFD with respect to different normalised sparsity parameter λ_s^N . (b) The L_1 -norm of the displacement at the finest FFD level of the SFFD with respect to different normalised sparsity parameter λ_s^N .

Table 1

Mean and SD of accuracies from the SFFD with different coarsest and finest spacing when normalised sparsity parameter $\lambda_s^N = 0.04$.

| Spacing | 32/1 mm | 64/1 mm | 128/1 mm | 256/1 mm | 512/1 mm |
|----------|------------|------------|------------|-------------|------------|
| Accuracy | 0.36(0.23) | 0.34(0.27) | 0.35(0.28) | 0.34 (0.27) | 0.34(0.27) |

Table 2

Mean and SD of accuracies from the classic FFD with different initial and final spacing and the SFFD with different normalised sparsity parameter λ_s^N . Cardiac A and B refer to the 2D cardiac MR image with synthetic smooth motion and discontinuous motion respectively. The units of the accuracy metrics are mm. For the results, the bold font indicates the best result and * indicates statistically significantly different from the best result of the classic FFD using t-test with p-value 0.05.

| FFD | 64/1 mm | 128/2 mm | 256/4 mm | 512/8 mm |
|------------|-------------|-------------------|---------------------|-------------------|
| Middlebury | 0.79(0.30) | 0.70(0.22) | 0.66(0.27) | 0.74(0.33) |
| Cardiac A | 0.03(0.02)* | 0.03(0.03)* | 0.02(0.01) | 0.02(0.01) |
| Cardiac B | 0.06(0.04)* | 0.05(0.01) | 0.05(0.02) | 0.06(0.02)* |
| SFFD | 0.00 | 0.01 | 0.04 | 0.16 |
| Middlebury | 0.78(0.30) | 0.34(0.28)* | 0.34 (0.27)* | 0.36(0.30)* |
| Cardiac A | 0.02(0.01) | 0.01(0.00) | 0.01(0.00)* | 0.01(0.00)* |
| Cardiac B | 0.06(0.03) | 0.02(0.00)* | 0.02(0.01)* | 0.02(0.01)* |

Table 3

Mean and SD of accuracies from the classic FFD with different initial and final spacing and the SFFD with different normalised sparsity parameter λ_s^N . Cardiac C refers to 3D cardiac ultrasound image sequences. The units of the accuracy metrics are mm. For the results, the bold font indicates the best result and * indicates statistically significantly different from the best result of the classic FFD using t-test with p-value 0.05.

| FFD | 32/4 mm | 64/8 mm | 128/16 mm | 256/32 mm |
|-----------|-------------|-------------|---------------------|------------|
| Cardiac C | 1.57(1.12)* | 1.13(1.04) | 1.07 (1.07) | 1.10(1.09) |
| SFFD | 0.00 | 0.01 | 0.04 | 0.16 |
| Cardiac C | 1.04(0.95) | 0.85(0.80)* | 0.84 (0.81)* | 0.92(0.83) |

Table 4

Mean and SD of accuracies from the TFFD with different initial and final spacing and the TSFFD with different normalised sparsity parameter λ_s^N . Cardiac C refers to the 3D cardiac ultrasound image sequences. The units for the registration error are in mm. For the results, the bold font indicates the best result and * indicates statistically significantly different from the best result of the TFFD using t-test with p-value 0.05.

| TFFD | 32/4 mm | 64/8 mm | 128/16 mm | 256/32 mm |
|-----------|-------------|-------------|---------------------|-------------|
| Cardiac C | 1.54(1.70)* | 0.71(0.70) | 0.65 (0.71) | 0.74(0.80) |
| TSFFD | 0.00 | 0.01 | 0.04 | 0.16 |
| Cardiac C | 0.53(0.58) | 0.47(0.54)* | 0.46 (0.53)* | 0.46(0.52)* |

constraints. In this paper, we choose the smoothness parameter according to (Rueckert et al., 1999). It might be interesting to further investigate the interaction between the two regularization terms in the future. We can also consider to use dictionaries of the deformation as the basis function to improve sparsity and use adaptive weighting of the sparsity to provide more efficient optimisation.

During the evaluation, we used datasets with dense motion ground truth. The results indicated that the proposed methods has a good ability to recover dense motion very close to ground truth. In the future, we would like to investigate if this improvements can translate to improvements in different medical applications such as respiratory motion reconstruction (Castillo et al., 2009) and brain segmentation (Klein et al., 2009).

There is a wide range of applications of the SFFD and TSFFD. As a general registration method, it can be used to improve the accu-

racy in registration based cardiac motion tracking (De Craene et al., 2010; Metz et al., 2011; De Craene et al., 2012; Shi et al., 2012b) and segmentation (Zhuang et al., 2010; Schaerer et al., 2010; Shi et al., 2011) frameworks. Moreover, the discontinuous motion embedded in the finest level of the parametric space can be used as extra information for segmentation problems. It will help segmentation algorithm to distinguish the boundary between low contrast organs like myocardium and liver. Due to the high accuracy of the reconstructed deformation, our method might be a good candidate to recover deformation from longitudinal analyses where accurate voxel to voxel correspondence between different time points are needed. Finally, our method can be easily applied to other regions where the motion is discontinuous, for example, in lung and abdominal images.

Appendix A. derivation of energy gradient

Let $\mathbf{x} = (x, y, z)^T$ stand for a world coordinate location and let $\mathbf{u} = (u, v, w)$ denote a lattice coordinate for a FFD. The FFD lattice represents the locations of the control point vectors $\Phi_{l,m,n}$. The dimensions of the FFD lattice are U, V and W in the principal directions, we can index the control point vectors so that the subscripts l, m and n each range from 0 to $U - 1, V - 1$ and $W - 1$ respectively. The displacement $\mathbf{h}(\mathbf{x})$ at world location \mathbf{x} with corresponding lattice coordinates \mathbf{u} is given by

$$\mathbf{h}(\mathbf{x}) = \sum_{l=0}^{U-1} \sum_{m=0}^{V-1} \sum_{n=0}^{W-1} B_l(u) B_m(v) B_n(w) \Phi_{l,m,n},$$

where B_l, B_m and B_n are univariate B-Spline kernels centred at l, m and n . The transformation \mathbf{T} is obtained by adding the displacement:

$$\mathbf{T}(\mathbf{x}) = \mathbf{x} + \mathbf{h}(\mathbf{x}),$$

so that the Jacobian of \mathbf{T} and the Jacobian of $\mathbf{h} = (\mathbf{h}_x, \mathbf{h}_y, \mathbf{h}_z)$ are related by

$$\frac{\partial \mathbf{T}}{\partial \mathbf{x}} = \mathbf{I} + \frac{\partial \mathbf{h}}{\partial \mathbf{x}},$$

where \mathbf{I} is the identity matrix. We will restrict J to denote the Jacobian of \mathbf{T} so that, in the world coordinate frame, the Jacobian of the transformation is

$$J = \frac{\partial \mathbf{T}}{\partial \mathbf{x}} = \mathbf{I} + \begin{pmatrix} \frac{\partial \mathbf{h}_x}{\partial x} & \frac{\partial \mathbf{h}_x}{\partial y} & \frac{\partial \mathbf{h}_x}{\partial z} \\ \frac{\partial \mathbf{h}_y}{\partial x} & \frac{\partial \mathbf{h}_y}{\partial y} & \frac{\partial \mathbf{h}_y}{\partial z} \\ \frac{\partial \mathbf{h}_z}{\partial x} & \frac{\partial \mathbf{h}_z}{\partial y} & \frac{\partial \mathbf{h}_z}{\partial z} \end{pmatrix}. \quad (\text{A.1})$$

We assume that the origin of the world and lattice frames coincide so that the conversion between world and lattice coordinates can be achieved by matrix multiplication. Let M be the matrix to convert world coordinates to lattice coordinates, i.e. $\mathbf{u} = M\mathbf{x}$.

The number of degrees of freedom in the FFD is N where $N = 3UVW$. Let $\phi = (\phi_0, \phi_1, \dots, \phi_{N-1})^T$ represent the parameter set of the FFD obtained by concatenating all the control point vectors $\Phi_{l,m,n}$ into a single vector. These can be easily extended to multi-level FFD as described in the main sections.

The energy function to be optimised consists of three parts in a standard formulation (Rueckert et al., 1999):

$$E = E_D + \lambda_r E_{\text{smooth}} + \lambda_v E_{\text{vol}}$$

The similarity energy, E_D is considered here to be the SSD (Kybic and Unser, 2003) although other options are possible:

$$E_D = \int_{\Omega} D(I_r, I_s, \mathbf{T}, \mathbf{x}) d\mathbf{x}$$

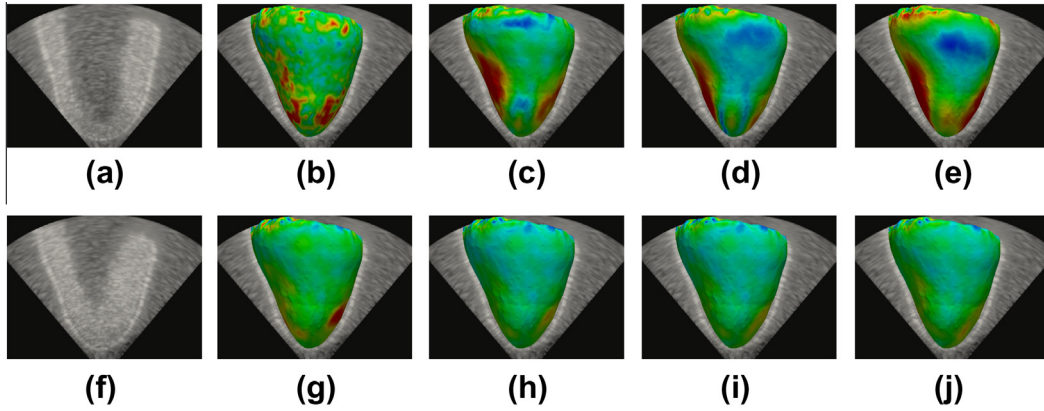


Fig. 7. Visualisation of the registration error for case 1 (lateral view) at frame 8 from cardiac C dataset. The colour overlay on the deformed ground-truth mesh represents the magnitude of the registration error (mm). (a) and (f) Show the ultrasound image from frame 1 and 8, respectively; (b–e) show the error map from the TFFD approach with 32/4 mm, 64/8 mm, 128/16 mm and 256/32 mm control point spacing and (g–j) show the error map from the TFFD approach with $\lambda_s^N = 0.01$, $\lambda_s^N = 0.04$ and $\lambda_s^N = 0.16$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5

Mean and SD of accuracies from the TSFFD with $\lambda_s^N = 0.04$. The results concern the Cardiac C data and the numbering refers to the original numbering in the challenge. The units of the accuracy are mm

| ID | 01 | 08 | 12 | 20 | 22 | 28 | 36 | 44 | 60 | 88 |
|------|------|------|------|------|------|------|------|------|------|------|
| Mean | 0.71 | 0.37 | 0.39 | 0.39 | 0.66 | 0.39 | 0.38 | 0.47 | 0.42 | 0.38 |
| SD | 0.78 | 0.37 | 0.41 | 0.38 | 0.74 | 0.40 | 0.39 | 0.49 | 0.41 | 0.55 |

where Ω is a fixed region of interest and $D(I_r, I_s, \mathbf{T}, \mathbf{x}) = (I_r(\mathbf{x}) - I_s(\mathbf{T}(\mathbf{x})))^2$ is the SSD between target image I_r and source image I_s after applying transformation \mathbf{T} at point \mathbf{x} . We require the gradient of this energy E_D with respect to the deformation parameters and we assume that $D(I_r, I_s, \mathbf{T}, \mathbf{x})$, as a function of the components of ϕ and \mathbf{x} , satisfies the conditions for differentiation under the integral sign. This means that we can write

$$\frac{\partial E_D}{\partial \phi} = \int_{\Omega} \frac{\partial D(I_r, I_s, \mathbf{T}, \mathbf{x})}{\partial \phi} d\mathbf{x} \quad (\text{A.2})$$

The derivative of the integrand is readily found in Kybic and Unser (2003)

$$\begin{aligned} \frac{\partial}{\partial \phi_i} (I_r(\mathbf{x}) - I_s(\mathbf{T}(\mathbf{x})))^2 &= \frac{\partial (I_r(\mathbf{x}) - I_s(\mathbf{T}(\mathbf{x})))^2}{\partial I_s(\mathbf{T}(\mathbf{x}))} \frac{\partial I_s(\mathbf{T}(\mathbf{x}))}{\partial \mathbf{T}(\mathbf{x})} \frac{\partial \mathbf{T}(\mathbf{x})}{\partial \phi_i} \\ &= 2(I_r(\mathbf{x}) - I_s(\mathbf{T}(\mathbf{x}))) \nabla I_s \circ T|_{\mathbf{x}} \frac{\partial \mathbf{T}(\mathbf{x})}{\partial \phi_i}. \end{aligned} \quad (\text{A.3})$$

The smooth preservation energy is a spatially integrated function of the bending energy:

$$E_{\text{smooth}} = \frac{1}{|\Omega|} \int_{\Omega} \left[\left(\frac{\partial^2 \mathbf{T}}{\partial x^2} \right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial y^2} \right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial z^2} \right)^2 \right] d\mathbf{x} \quad (\text{A.4})$$

$$+ \frac{1}{|\Omega|} \int_{\Omega} \left[2 \left(\frac{\partial^2 \mathbf{T}}{\partial xy} \right)^2 + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial xz} \right)^2 + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial yz} \right)^2 \right] d\mathbf{x} \quad (\text{A.5})$$

where $|\Omega|$ is the size of the region. We require the gradient of this energy with respect to the deformation parameters. Abbreviating Eq. (A.5) as $E_{\text{smooth}} = \frac{1}{|\Omega|} \int_{\Omega} [A^2 + B^2 + C^2 + 2D^2 + 2E^2 + 2F^2] d\mathbf{x}$, the derivative of the penalty term with respect to a deformation parameter ϕ_i involves a sum of derivatives each of which can be obtained using the chain rule, e.g. $\frac{\partial (A^2)}{\partial \phi_i} = 2A \frac{\partial A}{\partial \phi_i}$. The same chain rule can be applied to each and every component of the Eq. (A.5).

Finally, the volume preservation energy is a spatially integrated function of the determinant of the Jacobian, J :

$$E_{\text{vol}} = \int_{\Omega} (\log |J|)^2 d\mathbf{x}$$

where $|J|$ denotes the determinant of J . We require the gradient of this energy with respect to the deformation parameters and we assume that $(\log |J|)^2$, as a function of the components of ϕ and \mathbf{x} , satisfies the conditions for differentiation under the integral sign. This means that we can write

$$\frac{\partial E_{\text{vol}}}{\partial \phi} = \int_{\Omega} \frac{\partial \log |J|}{\partial \phi} d\mathbf{x} \quad (\text{A.6})$$

Differentiating the integrand, we obtain

$$\frac{\partial (\log |J|)^2}{\partial \phi} = \frac{2 \log |J|}{|J|} \frac{\partial |J|}{\partial \phi}. \quad (\text{A.7})$$

$|J|$ and its logarithm are readily calculated so we require an expression for the gradient $\frac{\partial |J|}{\partial \phi}$. The i th component of the $\frac{\partial |J|}{\partial \phi}$ is given by

$$\frac{\partial |J|}{\partial \phi_i} = \text{tr} \left(\text{adj}(J) \frac{\partial J}{\partial \phi_i} \right)$$

for $i = 0 \dots N - 1$, where $\text{tr}(\cdot)$ denotes the trace and $\text{adj}(\cdot)$ denotes the adjugate. We expand $\frac{\partial J}{\partial \phi_i}$ to give

$$\begin{aligned} \frac{\partial J}{\partial \phi_i} &= \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{T}}{\partial \mathbf{x}} \right) = \frac{\partial}{\partial \phi_i} \left(I + \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right) = \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right) \\ &= \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right) \end{aligned} \quad (\text{A.8})$$

We have

$$\frac{\partial \mathbf{u}}{\partial \mathbf{x}} = M,$$

where M is the constant world to lattice transformation matrix, so, after applying the product rule to Eq. (A.8), we obtain

$$\frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right) = \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{u}} \right) M = \frac{\partial^2 \mathbf{h}}{\partial \phi_i \partial \mathbf{u}} M$$

The term $\frac{\partial \mathbf{h}}{\partial \mathbf{u}}$ is the Jacobian of the deformation in the frame of the FFD lattice. We can re-write $\frac{\partial \mathbf{h}}{\partial \mathbf{u}}$ as three columns to obtain

$$\frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{u}} \right) = \left[\frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial u} \right) \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial v} \right) \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial w} \right) \right] \quad (\text{A.9})$$

The effect of differentiating with respect to the i th parameter ϕ_i depends on whether it corresponds to a displacement in the u , v or w direction of the lattice. The parameter vector ϕ is obtained by concatenation of the control point vectors $\Phi_{l,m,n}$ so that the parameters ϕ_i correspond to u -displacements if $i = 0 \bmod 3$, v -displacements if $i = 1 \bmod 3$ and w -displacements if $i = 2 \bmod 3$.

Assume, for example, that parameter ϕ_i corresponds to the first (u) component of the specific control point vector $\Phi_{\lambda,\mu,\nu}$. Consider the first column of the matrix on the right of Eq. (A.9):

$$\frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{u}} \right) = \frac{\partial}{\partial \phi_i} \sum_{l,m,n} \frac{dB_l(u)}{du} B_m(v) B_n(w) \Phi_{l,m,n} = \frac{\partial}{\partial \phi_i} \left[\frac{dB_l(u)}{du} B_\mu(v) B_\nu(w) \Phi_{\lambda,\mu,\nu} \right] = \begin{pmatrix} \frac{dB_l(u)}{du} B_\mu(v) B_\nu(w) \\ 0 \\ 0 \end{pmatrix}$$

The second and third columns can be found in a similar way to give

$$\frac{\partial}{\partial \phi_i} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{u}} \right) = \begin{pmatrix} \frac{dB_l(u)}{du} B_\mu(v) B_\nu(w) & B_l(u) \frac{dB_\mu(v)}{dv} B_\nu(w) & B_l(u) B_\mu(v) \frac{dB_\nu(w)}{dw} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The cases where \mathbf{h}_i corresponds to a second (v) or third (w) component of $\Phi_{\lambda,\mu,\nu}$ can be found in a similar fashion. The resulting expressions can be substituted back via Eqs. A.9, A.8 and A.7 to provide the value for the gradient of the volume preservation energy term in Eq. (A.6).

Finally, we precondition the gradient of the energy term using the principle of the elongation preconditioning scheme from Zikic et al. (2011):

$$f(\Phi) = \frac{1}{\|\Delta E(\Phi)\| + \epsilon} \Delta E(\Phi) \quad (\text{A.10})$$

where $f(\Phi)$ is the final driving force of the optimisation at control point Φ .

References

- Alessandrini, M., Liebgott, H., Barbosa, D., Bernard, O., 2013. Monogenic phase based optical flow computation for myocardial motion analysis in 3d echocardiography. *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM)*, 1–1.
- Baker, S., Scharstein, D., Lewis, J., Roth, S., Black, M., Szeliski, R., 2007. A database and evaluation methodology for optical flow. In: *International Conference on Computer Vision (ICCV)*, pp. 1–8.
- Bray, J., 1999. *Lecture Notes on Human Physiology*. Wiley-Blackwell.
- Castillo, R., Castillo, E., Guerra, R., Johnson, V.E., McPhail, T., Garg, A.K., Guerrero, T., 2009. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Physics in Medicine and Biology* 54, 1849.
- De Craene, M., Allain, P., Gao, H., Prakosa, A., Marchesseau, S., Hilpert, L., Somphone, O., Delingette, H., Makram-Ebeid, S., Villain, N., D'hooge, J., Sermesant, M., Saloux, E., 2013. Synthetic and phantom setups for the second cardiac motion analysis challenge. *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM)*, 1–1.
- De Craene, M., Piella, G., Camara, O., Duchateau, N., Silva, E., Doltra, A., Dhooge, J., Brugada, J., Sitges, M., Frangi, A., 2012. Temporal diffeomorphic free-form deformation: application to motion and strain estimation from 3D echocardiography. *Medical Image Analysis* 16, 427–450.
- De Craene, M., Piella, G., Duchateau, N., Silva, E., Doltra, A., Gao, H., Dhooge, J., Camara, O., Brugada, J., Sitges, M., et al., 2010. Temporal diffeomorphic free-form deformation for strain quantification in 3D-US images. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 1–8.
- Donoho, D., Huo, X., 2001. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory* 47, 2845–2862.
- Friedman, J., Hastie, T., Tibshirani, R., 2010. A note on the group lasso and a sparse group lasso. *arXiv*.
- Gao, H., Choi, H., Claus, P., Boonen, S., Jaecques, S., van Lenthe, G., Van Der Perre, G., Lauriks, W., D'hooge, J., 2009. A fast convolution-based methodology to simulate 2-d/3-d cardiac ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 56, 404–409.
- 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.
- Hansen, M., Larsen, R., Glocker, B., Navab, N., 2008. Adaptive parametrization of multivariate B-splines for image registration. In: *Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8.
- Heyde, B., Barbosa, D., Claus, P., Maes, F., Dhooge, J., 2013. Three-dimensional cardiac motion estimation based on non-rigid image registration using a novel transformation model adapted to the heart. *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM)*, 1–1.
- Horn, B., Schunck, B., 1981. Determining optical flow. *Artificial Intelligence* 17, 185–203.
- Kim, S., Koh, K., Lustig, M., Boyd, S., Gorinevsky, D., 2007. An interior-point method for large-scale L1-regularized least squares. *IEEE Journal of Selected Topics in Signal Processing* 1, 606–617.
- Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., et al., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46, 786.
- Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J., 2010. Elastix: a toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging* 29, 196–205.
- Kumar, D., Geng, X., Hoffman, E.A., Christensen, G.E., 2006. Bicir: boundary-constrained inverse consistent image registration using web-splines. In: *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 68–68.
- Kybic, J., Unser, M., 2003. Fast parametric elastic image registration. *IEEE Transactions on Image Processing* 12, 1427–1442.
- Mémin, E., Pérez, P., 2002. Hierarchical estimation and segmentation of dense motion fields. *International Journal of Computer Vision* 46, 129–155.
- Metz, C., Klein, S., Schaap, M., Van Walsum, T., Niessen, W., 2011. Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. *Medical Image Analysis* 15, 238–249.
- 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. *Computer Methods and Programs in Biomedicine* 98, 278–284.
- Perperidis, D., Mohiaddin, R., Rueckert, D., 2005. Spatio-temporal free-form registration of cardiac MR image sequences. *Medical Image Analysis* 9, 441–456.
- Piella, G., Porras, A.R., De Craene, M., Duchateau, N., Frangi, A.F., 2013. Temporal diffeomorphic free form deformation to quantify changes induced by left and right bundle branch block and pacing. *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM)*, 1–1.
- Pizarro, L., Delpiano, J., Aljabar, P., Ruiz-del Solar, J., Rueckert, D., 2011. Towards dense motion estimation in light and electron microscopy. In: *International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 1939–1942.
- Rohlfing, T., Maurer, C., 2001. Intensity-based non-rigid registration using adaptive multilevel free-form deformation with an incompressibility constraint. In: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 111–119.
- Roosgard, A., Barzigar, N., Cheng, S., Verma, P., 2011. Dense image registration using sparse coding and belief propagation. In: *International Conference on Signal Processing and Communication Systems*, pp. 1–5.
- Roth, S., Black, M., 2005. On the spatial statistics of optical flow. In: *International Conference on Computer Vision (ICCV)*, pp. 42–49.
- Rueckert, D., Aljabar, P., Heckemann, R., Hajnal, J., Hammers, A., 2006. Diffeomorphic registration using B-splines. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 702–709.
- Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D., 1999. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging* 18, 712–721.
- Schaerer, J., Casta, C., Pousin, J., Clarysse, P., 2010. A dynamic elastic model for segmentation and tracking of the heart in MR image sequences. *Medical Image Analysis* 14, 738–749.
- Schnabel, J., Rueckert, D., Quist, M., Blackall, J., Castellano-Smith, A., Hartkens, T., Penney, G., Hall, W., Liu, H., Truweit, C. et al., 2001. A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations. In: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 573–581.
- Sermesant, M., Chabiniok, R., Chinchapatnam, P., Mansi, T., Billet, F., Moireau, P., Peyrat, J., Wong, K., Relan, J., Rhode, K., et al., 2012. Patient-specific electromechanical models of the heart for the prediction of pacing acute effects in crt: a preliminary clinical validation. *Medical image analysis* 16, 201–215.
- Shen, X., Wu, Y., 2010. Sparsity model for robust optical flow estimation at motion discontinuities. In: *Computer Vision and Pattern Recognition (CVPR)*, pp. 2456–2463.
- Shi, W., Zhuang, X., Pizarro, L., Bai, W., Wang, H., Tung, K., Edwards, P., Rueckert, D., 2012a. Registration using sparse free-form deformations. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 659–666.
- Shi, W., Zhuang, X., Wang, H., Duckett, S., Luong, D., Tobon-Gomez, C., Tung, K., Edwards, P., Rhode, K., Razavi, R., et al., 2012b. A comprehensive cardiac motion estimation framework using both untagged and 3-d tagged mr images based on nonrigid registration. *IEEE Transactions on Medical Imaging* 31, 1263–1275.
- Shi, W., Zhuang, X., Wang, H., Duckett, S., O'Regan, D., Edwards, P., Ourselin, S., Rueckert, D., 2011. Automatic segmentation of different pathologies from cardiac cine MRI using registration and multiple component EM estimation. In: *International Conference on Functional Imaging and Modeling of the Heart (FIMH)*, pp. 163–170.
- Somphone, O., Dufour, C., Mory, B., Hilpert, L., Makram-Ebeid, S., Villain, N., De Craene, M., Allain, P., Saloux, E., 2013. Motion estimation in 3d echocardiography using smooth field registration. *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM)*, 1–1.

- Studholme, C., Hill, D., Hawkes, D., et al., 1999. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition* 32, 71–86.
- Sun, D., Roth, S., Black, M., 2010. Secrets of optical flow estimation and their principles. In: *Computer Vision and Pattern Recognition (CVPR)*, pp. 2432–2439.
- Thirion, J., 1998. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis* 2, 243–260.
- Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 267–288.
- Tobon-Gomez, C., De Craene, M., Dahl, A., Kapetanakis, S., Carr-White, G., Lutz, A., Rasche, V., Etyngier, P., Kozerke, S., Schaeffter, T., et al., 2012. A multimodal database for the 1st cardiac motion analysis challenge. *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM)*, 33–44.
- Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2008. Symmetric log-domain diffeomorphic registration: a demons-based approach. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 754–761.
- Xie, Z., Farin, G.E., 2004. Image registration using hierarchical b-splines. *IEEE Transactions on Visualization and Computer Graphics* 10, 85–94.
- Yuan, M., Lin, Y., 2005. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68, 49–67.
- Zhuang, X., Rhode, K., Razavi, R., Hawkes, D., Ourselin, S., 2010. A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. *IEEE Transactions on Medical Imaging* 29, 1612–1625.
- Zikic, D., Baust, M., Kamen, A., Navab, N., 2011. A general preconditioning scheme for difference measures in deformable registration. In: *IEEE International Conference on Computer Vision (ICCV)*. IEEE, pp. 49–56.