Model-based Regularisation for Respiratory Motion Estimation with Sparse Features in Image-guided Interventions

Matthias Wilms, In Young Ha, Heinz Handels and Mattias P. Heinrich

Institute of Medical Informatics, University of Lübeck, Lübeck, Germany
`wilms@imi.uni-luebeck.de`

Abstract. Intra-interventional respiratory motion estimation has become vital for image-guided interventions, especially radiation therapy. While real-time tracking of highly discriminative landmarks like tumours and markers is possible with classic approaches (e.g. template matching), their robustness decreases when used with non-ionising imaging (4D MRI or US). Furthermore, they ignore the motion of neighbouring structures. We address these challenges by dividing the computation of dense deformable registration in two phases: First, a low-parametric full domain patient-specific motion model is learnt. Second, a sparse subset of feature locations is used to track motion locally, while the global motion patterns are constrained by the learnt model. In contrast to previous work, we optimise both objectives (local similarity and globally smooth motion) jointly using a coupled convex energy minimisation. This improves the tracking robustness and leads to a more accurate global motion estimation. The algorithm is computationally efficient and significantly outperforms classic template matching-based dense field estimation in 12 of 14 challenging 4D MRI and 4D ultrasound sequences.

1 Introduction

Effective respiratory motion compensation is a key factor for successful non-invasive radiotherapy or High Intensity Focused Ultrasound (HIFU) treatments of thoracic and abdominal tumours. Recent advances in imaging technologies have led to the integration of 4D (3D+t) ultrasound (US)- and magnetic resonance imaging (MRI)-guidance into HIFU and radiation therapy [1–3]. Compared to traditional motion management using external breathing signals or X-ray projections, this opens up new possibilities for accurate intra-fraction motion estimation. However, the intra-interventional images need to be processed in real-time to control the treatment beam, which excludes the use of accurate but computationally demanding deformable image registration approaches.

Most published methods for online respiratory motion estimation based on temporal MRI or US have limitations compared to traditional deformable image registration algorithms (see [4] for an overview on US tracking). A common shortcoming is the use of template matching [5–7], which only focuses on a direct tracking of the tumour or sparse landmarks (markers, vessels, . . .) and ignores the spatial regularity of organ motion. These approaches achieve high computational speed with good accuracy for the specified template location, but do not provide an estimate of the motion of other structures, which requires dense estimation of displacements (e.g. using [8]). Dense displacement fields for the full patient body can then be reconstructed from the sparse motion vectors using trained motion models [9, 7, 10–12]. So far, sparse feature point matching and dense motion field reconstruction have been often treated as separate disconnected tasks (see e.g. [7, 10]). We, however, think that prior knowledge about respiratory motion should and can be incorporated into the sparse motion estimation step for improved robustness and accuracy without substantially increasing the computation time.

We propose a novel, robust, and efficient model-based method for online respiratory motion estimation in image-guided interventions that jointly combines local similarity-based block matching for sparse feature points with a global patient-specific statistical motion model for regularisation. The resulting minimisation problem is efficiently solved using ideas from discrete coupled convex optimisation for image registration [13]. This enables the use of very sparsely distributed feature vectors and achieves highly accurate dense displacement fields for complex respiratory motion (including a natural handling of sliding motion). Our approach is (to our knowledge) the first non-linear respiratory motion estimation approach that jointly optimises sparse feature point matching and model-based regularisation with computationally fast discrete optimisation techniques. In previous work on the joint use of image data and model-based regularisation, authors either use all the image data available [8, 14] instead of sparse features, perform gradient descent-based optimisation [14], only estimate affine transformations [14], and/or only compute 2D motion vectors [8].

2   Method

Although being independent of the imaging modality used, we will describe our method in an MRI-guided radiotherapy scenario for ease of understanding. Modern integrated MRI linear accelerators are able to acquire (multiple) 2D slices of the moving patient anatomy in real-time during the treatment [1].

Given a static 3D reference image $I_R : \Omega \to \Omega$ ($\Omega \subset \mathbb{R}^3$) depicting the region of interest at a reference time point, our goal is to determine a transformation $\varphi_t = Id + u_t : \Omega \to \Omega$ that describes the deformation of the structures in I_R at treatment time t based on the 2D or 3D moving image frame(s) $I_{M,t} : \Omega \to \mathbb{R}$ provided by the treatment system. Here, u_t represents a dense displacement field.

For computational efficiency, we initially restrict the motion estimation process to a sparse set of N feature points $\Omega_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, within the reference
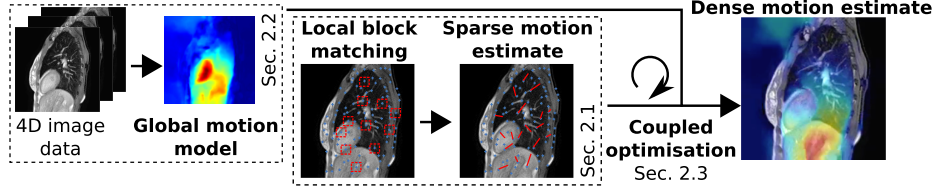
Fig. 1. Graphical overview of the proposed model-based method for respiratory motion estimation that combines local block matching with a global motion model.

image. Our method aims to find an optimal sparse displacement field $\tilde{u}_t$ defined at these feature points, which minimises a cost function $E(\tilde{u}_t)$:

$$E(\tilde{u}_t) = \sum_{\Omega_N} \mathcal{D}(I_R, I_{M,t}, \tilde{u}_t) + \alpha \mathcal{R}(\tilde{u}_t) \ . \tag{1}$$

Here, $\mathcal{D}$ quantifies the point-wise (dis)similarity between $I_R(\mathbf{x})$ and $I_{R,t}(\mathbf{x} + \tilde{u}_t(\mathbf{x}))$ around locations $\mathbf{x}$ and $\mathcal{R}$ is a regularisation term, which penalises deviations of $\tilde{u}_t$ from plausible solutions and is weighted by α.

In this work, $\mathcal{D}$ is based on the self-similarity context descriptor (SSC, cf. Sec. 2.1) [15] and $\mathcal{R}$ is derived from a patient-specific motion model (cf. Sec. 2.2), which is used both for regularisation and the final reconstruction of the dense displacement field u_t given a sparse estimate $\tilde{u}_t$. Minimising the joint cost function Eq. 1 is difficult due to its non-linear dependency on $\tilde{u}_t$. We, therefore, propose an efficient coupled convex discrete optimisation approach (cf. Sec. 2.3), which alternately optimises over the dissimilarity distribution of the local sparse block-matching and the global model-based regularisation (see Fig. 1 for a graphical overview).

2.1 Sparse Feature Point Detection and Similarity-driven Block Matching

The feature points Ω_N are automatically selected in the reference image I_R using the Harris/Foerstner corner detector [16] (alternatively, manually defined landmarks could also be employed). The tracking of Ω_N is based on the self-similarity context descriptor (SSC) [15], which has been chosen for its insensitivity to local changes in image contrast and to image noise as these effects regularly degrade the quality of interventional images. Furthermore, based on quantised SSC descriptors SSC_R and $SSC_{M,t}$ it allows the definition of a L_1 metric [15]

$$\mathcal{D}(\mathbf{x}_i, \mathbf{y}_{i,t}) = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} \Xi\{SSC_R(\mathbf{x}_i + p) \oplus SSC_{M,t}(\mathbf{y}_{i,t} + p)\} \ . \tag{2}$$

$\mathcal{D}$ assesses the similarity of the image contents at feature point $\mathbf{x}_i$ in I_R and its potentially corresponding location $\mathbf{y}_{i,t} = \mathbf{x}_i + \mathbf{d}_{i,t}$ in $I_{M,t}$, which can be efficiently computed in Hamming space using an XOR operator $\oplus$ followed by

a bit count $\varXi$. $\mathbf{d}_{i,t}$ denotes a displacement vector out of a predefined set of 3D displacements $\mathcal{L}$ (chosen according to the expected motion magnitude). Here, $\mathcal{P}$ is a local 3D block around each location $\mathbf{x}_i$ or $\mathbf{y}_i$ for which a block sum is formed. Using an unrestricted block-matching (minimising Eq. 1 with $\alpha = 0$), the optimal displacement $\hat{\mathbf{d}}_{i,t}$ could be directly obtained, resulting in a sparse displacement field $\tilde{u}_t$. This outcome might, however, be highly irregular.

2.2   Patient-specific Motion Model Building

In addition to the reference image I_R, we expect a patient-specific dynamic 4D MRI data set $\{I_j\}_{j \in \{1,\dots,M\}}$ covering a small number of breathing cycles to be available prior to the intervention to build a statistical motion model. In practise, this data could be acquired during a short set-up phase.

First, all M images $I_j \in \Omega \to \mathbb{R}$ are nonlinearly registered to the reference image I_R, resulting in a set of displacement fields $\{u_j\}_{j \in \{1,\dots,M\}}$. While being independent of the registration approach in principle, we chose the fast *deeds* algorithm [17], as it has demonstrated high accuracy in respiratory motion estimation tasks, and is able to correctly handle sliding motion.

Second, a principal components analysis (PCA) is applied to the vectorised displacement fields $\mathbf{u}_j \in \mathbb{R}^{3V}$ (V denotes the number of image voxels) to obtain a low-parametric representation of the space of plausible displacement fields. PCA is a widely used technique for respiratory motion modelling [9, 7, 10, 11] and can be performed using an eigendecomposition of the sample covariance matrix

$$\mathbf{C} = \frac{1}{M} \sum_{j=1}^{M} (\mathbf{u}_j - \bar{\mathbf{u}})(\mathbf{u}_j - \bar{\mathbf{u}})^T = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T \ , \ \text{with } \bar{\mathbf{u}} = \frac{1}{M} \sum_{j=1}^{M} \mathbf{u}_j \ . \qquad (3)$$

The columns of the orthonormal matrix $\mathbf{P} \in \mathbb{R}^{3V \times 3V}$ are the eigenvectors of $\mathbf{C}$ and the diagonal elements of diagonal matrix $\mathbf{\Lambda} = diag(\lambda_1, \dots, \lambda_{3V}) \in \mathbb{R}^{3V \times 3V}$ are the corresponding eigenvalues in descending order. Aiming at a low-parametric representation of the space of plausible displacement fields, only the eigenvectors with the k largest eigenvalues that explain a certain percentage of variance (here: 95%) are retained. Displacement fields belonging to the space spanned by a reduced $\mathbf{P}_k \in \mathbb{R}^{3V \times k}$ can be generated by $\mathbf{u} = \bar{\mathbf{u}} + \mathbf{P}_k \mathbf{\Sigma_k} \mathbf{b}$, with weights $\mathbf{b} \in \mathbb{R}^k$ and diagonal matrix $\mathbf{\Sigma}_k = diag(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k})$. We aim to find an optimal weight vector $\mathbf{b}$ that reconstructs a dense displacement field vector $\mathbf{u}$ based on the sparse (vectorised) displacements $\tilde{\mathbf{u}}_t$. This can be achieved by minimising the ridge regression-like cost function [7, 12]:

$$E(\mathbf{b}) = \|\tilde{\mathbf{P}}_k \mathbf{\Sigma}_k \mathbf{b} - (\tilde{\mathbf{u}}_t - \bar{\mathbf{u}})\|_2^2 + \eta \|\mathbf{b}\|_2^2 \ . \qquad (4)$$

Here, matrix $\tilde{\mathbf{P}}_k \in \mathbb{R}^{3N \times k}$ only contains the $3N$ elements of the k eigenvectors that correspond to the elements present in the sparse displacement field $\tilde{u}_t/\tilde{\mathbf{u}}_t$. The regularised least-squares cost (Eq. 4) balances a close estimate of the observed sparse motion and deviations from the mean motion due to noise and has been frequently used to reconstruct dense displacement fields (e.g. in [7, 10, 11]).

2.3 Coupled Convex Optimisation of Model-based Regularisation

The displacement vectors $\hat{\mathbf{d}}_{i,t}$ obtained independent of each other using an unconstrained block-matching search (Eq. 1 with $\alpha = 0$) will contain erroneous estimates for challenging data. The ridge regression (Eq. 4 with $\eta > 0$) can dampen these errors but may reduce the overall accuracy of the densely reconstructed field. It will therefore be advantageous to minimise the block-matching dissimilarity and the model-penalty jointly (see Eq. 1). Due to the nonlinear dependency on $\tilde{u}_t$, directly minimising Eq. 1 is difficult. Following [13], a good approximation to the global optimum can be obtained in few iterations by adding a coupling term $\|\tilde{u}_t - \tilde{v}_t\|_2^2$ to Eq. 1 and introducing an auxiliary vector $\tilde{v}_t$:

$$E(\tilde{u}_t, \tilde{v}_t) = \sum_{\Omega_N} \mathcal{D}(I_R, I_{M,t}, \tilde{u}_t) + \theta\|\tilde{u}_t - \tilde{v}_t\|_2^2 + \alpha\mathcal{R}(\tilde{v}_t) \tag{5}$$

The optimisation of Eq. 5 is initialised with results of the unconstrained (here: $\theta=0$) block-matching search $\tilde{u}_t$, which is used to estimate a first regularised sparse field $\tilde{v}_t$ by projecting $\tilde{u}_t$ to the space spanned by the model using Eq. 4. The weighting parameter α in Eq. 5 is implicitly set through the number of eigenvectors k used to form $\mathbf{P}_k$. This alternating scheme is iterated using a series of increasing values of θ. During this process, the two objectives are encouraged to converge to a similar optimum, while updated estimates of $\tilde{u}_t$ (including the non-zero coupling term) make use of the full distribution of block-matching dissimilarities $\mathcal{D}$ of Eq. 2 (Eq. 2 only has to be computed once). In contrast to [13], which used an unspecific Gaussian regularisation, our approach elegantly incorporates both local uncertainty information from sparse feature points and a global domain-specific motion model. This enables us to estimate complex dense motion very efficiently and avoid the negative influence of errors from an unconstrained block-matching. Note, that this method will correctly estimate sliding motion, if it was present in the training data. We have used 6 iterations of the optimisation scheme in our experiments with $\theta = \{0.5, 1.5, 2.5, 10, 50, 100\}$.

Furthermore, prior knowledge of temporally smooth motion can be included by adding a second regularisation term $\beta\|\tilde{u}_t - \tilde{u}_{t-1}\|_2^2$ to Eq. 5 that penalises deviations of motion vectors compared to the previous frame in the sequence. The weighting parameter β should be chosen according to the expected inter-frame differences of the motion and the image noise level (here: $\beta = \{0, 0, 0, 0, 5, 10\}$).

3 Experiments and Results

Experiments on 5 thoracic/abdominal 4D MRI data sets and 9 liver 4D US data sets are performed to show the benefits of our new model-based respiratory motion estimation approach when compared to separate block matching and dense displacement field reconstruction. The 14 4D data sets are used to mimic the online motion estimation process in MR/US-guided treatment scenarios. A subset of each data set is used to train a patient-specific motion model while the remaining images serve as intra-interventional data.

Table 1. Mean estimation errors with respect to the ground-truth displacement fields obtained for the different approaches applied on the 4D MRI and US data. Results are given as mean±standard deviation in mm over all patients and frames included in each collection. The first row gives the error for all body voxels while the second row lists only the errors at the feature point locations. For the MRI data, results at voxels with large mean motion (>80th percentile of each case) are given for comparison (3rd row).

| Data	Motion	GT recon.	BM $\eta = 0$	BM $\eta > 0$	Model -based	Model & temporal
4D MRI experiments						
All voxels	1.76±0.39	0.74±0.26	1.76±0.38	1.09±0.31	0.87±0.26	**0.87±0.26**
Feature pts.	1.70±0.40	0.57±0.14	1.28±0.31	0.90±0.34	0.69±0.15	**0.69±0.15**
Large mot.	4.48±1.21	1.31±0.56	3.12±0.75	2.01±0.48	1.64±0.66	**1.66±0.67**
4D US experiments						
All voxels	4.22±1.14	1.07±0.27	7.29±1.59	2.82±0.89	2.20±0.81	**1.86±0.64**
Feature pts.	3.82±1.40	0.66±0.28	4.19±1.29	2.48±1.03	1.77±0.84	**1.38±0.57**

Data: The 4D MRI data collection contains 3 data sets from our own fund and 2 publicly available data sets [11]. Each sequence consists of $157 - 200$ 3D images (see Fig. 1 for example slices) acquired with a temporal resolution of 200 $- 500$ ms, an isotropic in-plane spatial resolution of $1.2 - 3.9$ mm, and an inter-slice distance of $5 - 10$ mm. The 4D US data collection used here is a subset (data sets SMT-01 $- 09$) of the CLUST challenge data [4][1]. Each data set consists of 96 $- 97$ 3D frames acquired at 8 Hz with an isotropic spatial resolution of 0.70 mm.

Experimental design: The first image in each MRI/US sequence is chosen as the reference image for an inter-sequence registration using *deeds* (cf. Sec. 2.2). The first third (MRI)/half (US) of the resulting displacement fields are used to build the patient-specific motion models. The remaining fields serve as ground-truth data for the quantitative evaluation. Their accuracy was evaluated on a subset of the data with a small number of manually defined landmarks. Landmark errors were in the range of 1 mm (MRI, in-plane)/$1 - 2$ mm (US) for landmarks with an average mean motion of 4 mm (MRI)/6 mm (US).

Our model-based algorithm is then employed to estimate the motion between the reference image and each image not used for model formation based on $250 - 300$ (MRI)/$70 - 80$ (US) automatically determined feature points (cf. Sec. 2.1). For the MRI data, feature point selection and block matching are restricted to $5 - 10$ equidistantly spaced 2D slices to simulate the sparse data acquired by an MR scanner during the treatment. We quantitatively assess the estimation accuracy by computing mean vector differences between the estimated fields and the ground-truth fields for all inner-body voxels/feature point locations. Due to the large inter-slice distances out-of-plane motion is ignored for MRI data sets.

Results: The results of our experiments are summarised in Tab. 1. In addition to two versions of our algorithm (model-based regularisation and model-based regularisation + temporal constraint (cf. Sec. 2.3)), Tab. 1 list results for an

---

[1] We thank the CLUST challenge organisers for providing the US data.
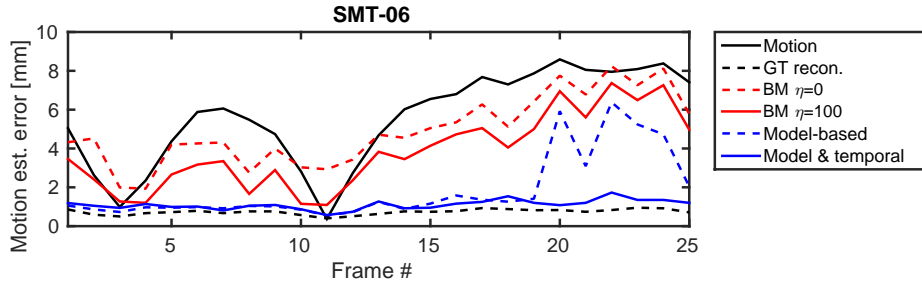
Fig. 2. Mean motion estimation errors obtained for different frames of the SMT-06 US data set at the feature point locations. The advantage of the model-based regularisation with additional temporal constraint is clearly seen starting from frame 20.

unrestricted block matching followed by a dense field reconstruction (BM) with $\eta = 0$/optimised $\eta > 0$ (cf. Eq. 4, [7]). The η parameter controls the amount of regularisation during least-squares fitting and was patient-specifically optimised with respect to the mean error over all frames. For comparison, the motion to be compensated and the error obtained by performing an optimal model-based least-squares reconstruction of the dense ground-truth field (GT recon.) is given.

From Tab. 1 and Fig. 2, it can be seen that the unrestricted BM with $\eta = 0$ leads to unsatisfactory results due to large outliers present in the sparse field. Their effect is substantially reduced by using $\eta > 0$ for dense field reconstruction. Including our proposed coupled optimisation with model-based regularisation outperforms both BM approaches ($\eta = 0$ & $\eta > 0$) in a statistically significant way (paired t-test, $p < 0.05$) in 86% of the cases (12 of 14). The differences in Tab. 1 between BM ($\eta = 0$ & $\eta > 0$) and model-based regularisation for the US experiments are also statistically significant, whereas the differences for the MRI experiments and $\eta > 0$ are not. Tab. 1 and Fig. 2 also show the advantages of incorporating the temporal constraint into the optimisation for the US experiments due to the low image quality that leads to severe inter-frame differences. Computationally, our approach needs 0.5–4 s to process each frame on a six-core Xeon CPU. Most of the time is spent for the BM and the SSC descriptor calculations, which could be easily transferred to the GPU with substantial speed-up, whereas the overhead for the coupled optimisation is minimal.

4   Conclusion

In this work, a novel model-based method for online respiratory motion estimation in image-guided interventions has been presented. The approach combines local similarity-based block matching for sparse feature points and a global motion model for regularisation. The resulting cost function is efficiently minimised by using a coupled convex discrete optimisation scheme. Our experiments show that this approach significantly outperforms decoupled template matching and dense motion field reconstruction methods implemented in the same framework.

The evaluation in this paper serves as a first proof-of-concept and further experiments on additional data would strengthen our findings. However, we expect the relative performances of the different approaches to remain the same. Future work will also include the integration of more sophisticated feature selection approaches, which might further improve the estimation accuracy.

References

1. Lagendijk, J.J.W., Raaymakers, B.W., den Berg, C.A.T.V., Moerland, M.A., Philippens, M.E., van Vulpen, M.: MR guidance in radiotherapy. Phys Med Biol **59**(21) (2014) R349–R369
2. de Senneville, B.D., Ries, M., Bartels, L.W., Moonen, C.T.W.: MRI-Guided High-Intensity Focused Ultrasound Sonication of Liver and Kidney. In: Interventional Magnetic Resonance Imaging. Springer (2012) 349–366
3. Fontanarosa, D., van der Meer, S., Bamber, J., Harris, E., OShea, T., Verhaegen, F.: Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management. Phys Med Biol **60**(3) (2015) R77–R114
4. Luca, V.D., Benz, T., Kondo, S., et. al.: The 2014 liver ultrasound tracking benchmark. Phys Med Biol **60**(14) (2015) 5571–5599
5. Cervino, L.I., Du, J., Jiang, S.B.: MRI-guided tumor tracking in lung cancer radiotherapy. Phys Med Biol **56**(13) (2011) 3773–3785
6. Brix, L., Ringgaard, S., Sorensen, T.S., Poulsen, P.R.: Three-dimensional liver motion tracking using real-time two-dimensional MRI. Med Phys **41**(4) (2014) 042302
7. Preiswerk, F., Luca, V.D., Arnold, P., Celicanin, Z., Petrusca, L., Tanner, C., Bieri, O., Salomir, R., Cattin, P.C.: Model-guided respiratory organ motion prediction of the liver from 2D ultrasound. Med Image Anal **18**(5) (2014) 740–751
8. de Senneville, B.D., Hamidi, A.E., Moonen, C.: A Direct PCA-Based Approach for Real-Time Description of Physiological Organ Deformations. IEEE Trans Med Imag **34**(4) (2015) 974–982
9. McClelland, J., Hawkes, D., Schaeffter, T., King, A.: Respiratory motion models: A review. Med Image Anal **17**(1) (2013) 19–42
10. Klinder, T., Lorenz, C.: Respiratory motion compensation for image-guided bronchoscopy using a general motion model. In: ISBI 2012. (2012) pp. 960–963
11. Boye, D., Samei, G., Schmidt, J., Szkely, G., Tanner, C.: Population based modeling of respiratory lung motion and prediction from partial information. In: SPIE Medical Imaging 2013. Volume 8669. (2013) pp. 86690U–86690U–7
12. Wulff, J., Black, M.J.: Efficient sparse-to-dense optical flow estimation using a learned basis and layers. In: CVPR 2015. (2015) pp. 120–130
13. Heinrich, M.P., Papież, B.W., Schnabel, J.A., Handels, H.: Non-parametric Discrete Registration with Convex Optimisation. In Ourselin, S., Modat, M., eds.: WBIR 2014. Volume 8545 of LNCS., Springer, Heidelberg (2014) pp. 51–61
14. Peressutti, D., Penney, G.P., Housden, R.J., Kolbitsch, C., Gomez, A., Rijkhorst, E.J., Barratt, D.C., Rhode, K.S., King, A.P.: A novel bayesian respiratory motion

model to estimate and resolve uncertainty in image-guided cardiac interventions. Med Image Anal **17**(4) (2013) 488–502

15. Heinrich, M.P., Handels, H., Simpson, I.J.A.: Estimating Large Lung Motion in COPD Patients by Symmetric Regularised Correspondence Fields. In Navab, N., Hornegger, J., Wells, M.W., Frangi, F.A., eds.: MICCAI 2015, Part II. Volume 9350 of LNCS., Springer International Publishing (2015) pp. 338–345

16. Rohr, K.: On 3D differential operators for detecting point landmarks. Image and Vision Computing **15**(3) (1997) 219–233

17. Heinrich, M.P., Jenkinson, M., Brady, M., Schnabel, J.A.: MRF-Based Deformable Registration and Ventilation Estimation of Lung CT. IEEE Trans Med Imag **32**(7) (2013) 1239–1248