

# Simultaneous Data Volume Reconstruction and Pose Estimation from Slice Samples

Manfred Georg\*, Richard Souvenir†, Andrew Hope‡, Robert Pless\*

Washington University in St. Louis

\* Saint Louis, MO 63130 USA

{mgeorg,pless}@cse.wustl.edu

† UNC Charlotte

Charlotte, NC 28223 USA

souvenir@uncc.edu

‡ University of Toronto

Toronto, ON, M5G 2M9, Canada

Andrew.Hope@rmp.uhn.on.ca

## Abstract

Modeling the dynamics of heart and lung tissue is challenging because the tissue deforms between data acquisitions. To reconstruct complete volumes, sample data captured at different times and locations must be combined. This paper presents a novel end-to-end, data driven framework for the complete reconstruction of deforming tissue volumes. This framework is a joint optimization over an undeformed tissue volume, a deformation map that describes tissue motion for given pose parameters (i.e. breathing and heartbeat), and an estimate of those parameters for each data acquisition. Tissue motion is modeled by deforming a reference volume with a cubic B-spline free form deformation, and we use Isomap to derive initial estimates of the pose of sample data. An iterative method is used to simultaneously solve for the reference volume and deformation map while updating the pose estimates. This same process is demonstrated on 4D CT lung data and heart/lung MR data.

## 1. Introduction

Understanding respiratory and cardiac tissue motion is a key pre-cursor to many diagnostic and treatment protocols. Reconstructing a 4D model of how a tissue volume deforms is a challenging computational problem, that requires the joint optimization of three parts:

1. *Reference volume estimation* – Estimating the undeformed appearance of the data volume.
2. *Parameterized deformation map generation* – Estimating the deformation of the reference volume due to tissue motion.
3. *Deformable-pose estimation* – Estimating, for each time sample, the motion parameter of the deformation map at the time of sampling.

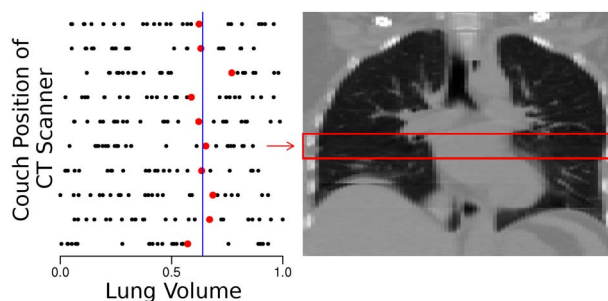


Figure 1. Previous 4D CT lung reconstruction methods pick the slab in each scanner position with lung volume closest to a target value (shown as vertical line). Each slab only covers part of the lung (one such slab is outlined in red). Discontinuity errors can be seen at the boundaries of scanner positions.

Common approaches to this problem requires additional data in the form of either (a) a known reference volume, or (b) feature detectors that reliably find and track points to estimate the deformation. In this paper, we propose a completely data driven method that is successful without either of these.

The key contribution that allows this approach is a recently proposed algorithm for manifold learning called Isomap [13], and more recent specializations to allow parameterization of medical images [4]. This family of algorithms takes in an unsorted collection of data volumes that vary due to an unknown deformation and gives a perceptual organization to those data volumes, *i.e.* sorting them so that similar data volumes have similar parameters. This provides a novel bootstrapping approach focusing on the third piece of the optimization procedure by automating a mechanism to estimate the pose of each data volume, without specifying either the reference volume or what type of deformation there is.

We demonstrate this algorithm in two different scenarios: reconstructing the deformation field of 2D slices captured in a cine-MRI, and reconstructing 4D deformation

fields in lung CT. The lung CT acquisition process only captures a small slab of data at a time, making it difficult to acquire a full reference volume.

In 4D computed tomography (4DCT) of the lung [5, 6], a part of the lung about 1 inch in height is repeatedly imaged while the patient breaths. The patient is then moved 1 inch and a set of images at the next position is acquired. Each image is ordered by the breath cycle phase at which it was acquired. A lung volume at a specific breath phase can be created by taking a sample from each scanner position at the same breath phase. This is shown in figure 1. Our approach to this problem differs in that it directly deconstructs the samples into a reference image and a deformation map, which allows us to create a lung image at any breath position, even when a gap in the data occurs.

## 2. Previous Work

Previous work in the analysis of dynamic medical imagery includes studies of 4D CT data and cardiac MRI. B-spline models of tissue deformation have been fit to tagged MR images [7, 14], and untagged MR images [2]. In 4D CT of the lung a B-spline deformable model has been used to find the deformation between 3D reconstructions at approximate maximum inhale and approximately maximum exhale [9]. More recent work aims to minimize the artifacts caused in the 3D reconstruction at each breath phase; when data is missing in a particular slab, optic flow is used to reconstruct a 4D CT data set by interpolating data from existing measurements to exactly the predefined respiratory phase at a particular phase. This helps for some kinds of artifacts but not artifacts caused by vertical motion that causes tissue to leave the (one inch high) data acquisition volume [3]. Some 4D CT studies of the lung have relied on a given reference image which is acquired by scanning the patient during a breath hold [15, 8]. A lung deformation model can then be created by registering the sample images from different parts of the breath to the breath hold image and fitting a B-spline to the deformation. Our method does not require a separately acquired reference image but can compute this image.

In the computer vision community, recent work on holistic analysis of sets of images that vary due to a few parameters is often based on manifold learning. Approaches include using manifold constraints to regularize segmentation of cardiac MR images [16], and to learn the low-dimensional parameterization of a cardiac MR set [11], in order to find nearby neighbors for effective image interpolation. However, neither approach solves for an explicit reference volume or deformation map. In the general case of an object undergoing translational movement manifold pursuit can be used [10].

## 3. Modeling Deformations of Tissue Volumes

Medical images of moving tissue can be represented as a function  $I(\vec{x}, \vec{\theta})$  which captures the intensity of tissue in the image at all spatial locations  $\vec{x}$  and poses  $\vec{\theta}$ . This function should match the intensity values of the image samples  $S_i(\vec{x})$  at their respective pose  $\vec{\theta}_i$ . To constrain the data volume  $I$  we assume that it is created from an underlying reference volume  $I_r$  which accounts for the appearance of the structure in all spatial dimensions and a set of free form deformations  $\vec{f}(\vec{x}, \vec{\theta})$  parameterized by pose  $\vec{\theta}$  which account for the motion of the structure during physiological activity. We use cubic B-splines as our free form deformation for simplicity of computation and differentiability and for the finite support of its basis functions. The equation for a B-spline deformation with two input dimensions is computed as

$$\vec{f}(x, \theta) = \sum_{a=1}^4 \sum_{b=1}^4 A_a(x) B_b(\theta) \vec{w}_{ab} \quad (1)$$

where  $A_a$  and  $B_b$  are cubic basis functions with weights  $\vec{w}_{ab}$ . The deformation takes as input the spatial position and pose of the structure, and outputs the deformed position of the structure in a reference coordinate system. The deformed image  $I_d$  at a pose  $\vec{\theta}$  is calculated by indexing the reference image  $I_r$  with the deformation map.

$$I_d(\vec{x}, \vec{\theta}) = I_r(\vec{f}(\vec{x}, \vec{\theta})) \quad (2)$$

## 4. Model Optimization

The model is defined by three sets of parameters which correspond to the three kinds of inference problem.

1. The values of the voxels in the reference volume  $I_r$ .
2. The parameters of the deformation map  $\vec{w}_{\vec{a}}$ .
3. The pose estimates  $\vec{\theta}_i$  for each sample.

The cost of the model is defined as the reconstruction error of the samples.

$$C = \sum_{S_i \in \mathcal{S}} \int (I_d(\vec{x}, \vec{\theta}_i) - S_i(\vec{x}))^2 dx \quad (3)$$

The integral is defined over all the pixels of the sample image  $S_i$ , which may only be part of the range of the model.

We use a standard non-linear solver L-BFGS-B based on a Quasi-Newton approach to solve for the minimum of the cost function [1]. In each iteration we obtain a new set of parameters for the deformation map and pose estimates from the L-BFGS-B algorithm. we compute the cost function and gradient of the cost function with respect to pose estimates and deformation parameters and provide these values to the L-BFGS-B algorithm. Additionally, if the cost function has

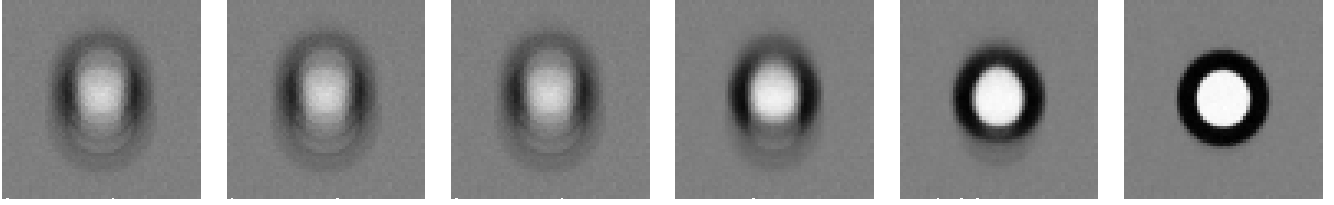


Figure 2. Reference images of synthetic data in iteration 1, 10, 50, 100, 400, and at convergence (1307).

gone down since the last iteration then the reference image is recomputed by deforming the samples images into the coordinate system of the reference image using the current deformation map.

#### 4.1. Basis Weight Derivatives

For clarity we present the following equations for the case with exactly one spatial dimension and one pose dimension. The gradient of the cost function with respect to a basis function weight can be computed analytically.

$$\begin{aligned} \frac{\delta}{\delta w_{\alpha\beta}} C &= \frac{\delta}{\delta w_{\alpha\beta}} \sum_{S_i \in \mathbb{S}} \int (I_d(x, \theta_i) - S_i(x))^2 dx \\ &= 2 \sum_{S_i \in \mathbb{S}} \int (I_d(x, \theta_i) - S_i(x)) \frac{\delta}{\delta w_{\alpha\beta}} I_d(x, \theta_i) dx \end{aligned} \quad (4)$$

The derivative of the deformed image with respect to the weight of a basis function can be calculated using the chain rule on equation 2.

$$\begin{aligned} \frac{\delta}{\delta w_{\alpha\beta}} I_d(x, \theta_i) &= \frac{\delta}{\delta w_{\alpha\beta}} I_r(f(x, \theta)) \\ &= \frac{\delta I_r(f(x, \theta))}{\delta f(x, \theta)} \frac{\delta}{\delta w_{\alpha\beta}} f(x, \theta) \end{aligned} \quad (5)$$

The generalized form of the chain rule is required when the model has more than one spatial dimension; however, this value is still easily computed. The first term of the product can be calculated directly as an image derivative of the deformed image. The second part can be calculated from the B-spline using equation 1.

$$\begin{aligned} \frac{\delta}{\delta w_{\alpha\beta}} f(x, \theta) &= \frac{\delta}{\delta w_{\alpha\beta}} \sum_{a=1}^4 \sum_{b=1}^4 A_a(x) B_b(\theta) w_{ab} \\ &= A_\alpha(x) B_\beta(\theta) \end{aligned} \quad (6)$$

This just picks the basis function which corresponds to the weight  $w_{\alpha\beta}$ .

#### 4.2. Pose Estimate Derivatives

The gradient of the cost function with respect to the pose  $\theta_k$  of a sample can also be calculated analytically. Similarly to equation 4, the derivative of the cost function can be

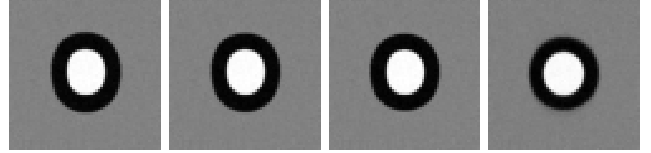


Figure 3. Final reference images as the deformation is constrained to be closer to the identity at the central pose.

calculated; however, only one sample will be non-zero.

$$\frac{\delta}{\delta \theta_k} C = 2 \int (I_d(x, \theta_k) - S_k(x)) \frac{\delta}{\delta \theta_k} I_d(x, \theta_k) dx \quad (7)$$

The partial derivative of the B-spline function with respect to  $\theta_k$  requires taking the derivative of some of the basis functions of the B-spline.

$$\frac{\delta}{\delta \theta_k} f(x, \theta_k) = \sum_{a=1}^4 \sum_{b=1}^4 A_a(x) B'_b(\theta_k) w_{ab} \quad (8)$$

In which  $B'_b$  is the derivative of  $B_b$ .

#### 4.3. Computing Reference Images

An optimal reference image can be computed given a deformation map and pose estimates. Each sample image can be deformed into the coordinate system of the reference image, using the inverse of the B-spline deformation. The mean of these deformed sample images is the best possible reference image for the given deformation and pose estimates. Figure 2 shows the evolution of the reference image as the optimization iterates.

#### 4.4. Initial Conditions

The identity transformation is used as an initial guess for the deformation. The initial pose estimates can be measured directly using a physical device such as a belt around the patients abdomen, or can be computed from the image data using manifold learning techniques [13]. Assuming the identity transform as an initial deformation leads to an initial reference image which is the average of all the samples.

#### 4.5. Anchoring the Reference Image

As described so far, the optimization problem is unconstrained, for example, a shift of the reference image is

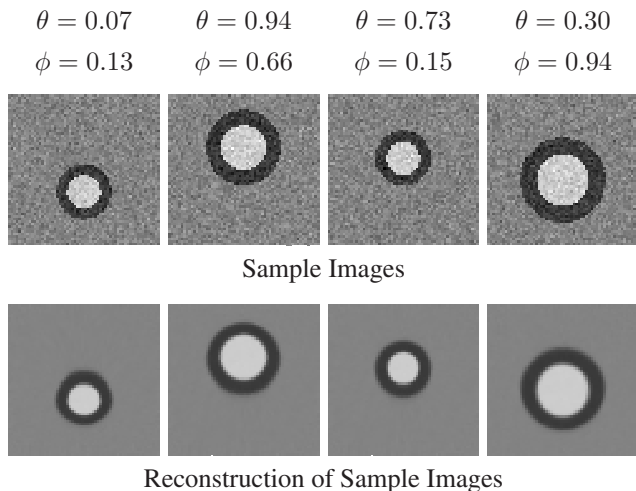


Figure 4. Image samples and their reconstructions. The reconstructions contain less noise, since they combine the data from many samples.

equivalent to a shift of the deformation map. Furthermore, there is no guarantee that the reference image even resembles the tissue. This problem can be solved by constraining the reference image to be equal to the tissue volume at a particular pose. This is accomplished by penalizing deviation ( $l_2$  norm) of the deformation map from the identity deformation at the central pose. A large scaling factor on this penalty term can lead to slow convergence and may adversely effect pose estimates if they are allowed to change during optimization. Figure 3 shows a sequence of final reference images for increasing penalty terms; notice that the reference images become rounder and more like the sample images as the level of constraint on the deformation is increased.

#### 4.6. Finding Deformations in Low Contrast Regions

In low contrast regions, any deformation map will yield similar images, making it difficult to determine the appropriate deformation map. However, tissue normally remains together as it moves, and hence has a similar direction of movement at nearby points. By applying a smoothness constraint we can determine the most likely deformation map even in low contrast regions, and otherwise overcome noise in our images. We enforce a smoothness constraint by applying a penalty term on the square of all second derivatives of the B-spline deformation with respect to both its spatial parameters and pose parameters.

Different deformations or penalty terms can induce properties such as volume preservation. However, for efficient optimization, the deformation must be differentiable. A useful application of penalty terms could be in distinguishing rigid and semi-rigid bodies [12]. In general, we have found that a coarser grid of control points is more effective

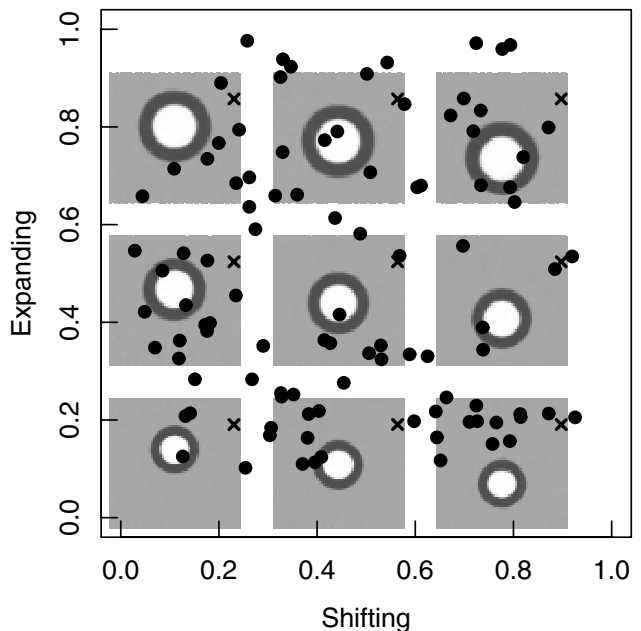


Figure 5. Pose estimates (black dots) for synthetic data. Reconstructed images of specific poses marked with a  $\times$  are shown.

at producing a smoother deformation than the use of a smoothness penalty term.

## 5. Sample Applications

In this section we show results of our algorithm for three data sets: a synthetic data set, a 4D CT lung data set, and a 2D heart MRI data set.

### 5.1. Synthetic Data

We create a synthetic dataset of the heart consisting of two concentric circles which we have deformed in two ways, shifting (to simulate the effects of breathing on the heart) and expansion (to simulate a heartbeat). Each of the 100 sample images is 60x60 pixels and has Gaussian noise added. Figure 4 shows some sample images and their reconstructions. Figure 5 displays the data volume of this synthetic data. Each black dot represents the location in parameter space of one of the 2D sample images used as input to the algorithm. At the poses marked with  $\times$  we draw the cross section of the data volume  $I(*, *, \theta, \phi)$  computed by our algorithm. The generating size and positions were chosen randomly and were not used by the algorithm. Instead, initial poses were estimated using Isomap to extract two parameters and employing Independent Component Analysis to rotate the axes to a more meaningful position.

B-spline control points were placed on a 10 pixel grid (6 spans in both x and y coordinates) with 5 spans over the shifting ( $\theta$ ) parameter and 5 spans over the expansion ( $\phi$ )



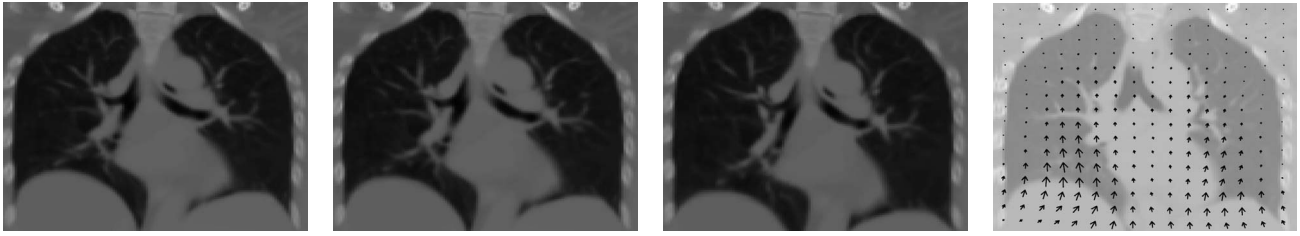


Figure 6. Lung reconstructions arranged from exhalation to inhalation. Right: vectors of tissue movement during breathing.

parameter for a total of 10368 parameters. Optimization took about 3 minutes on an AMD Athlon 64 3800.

### 5.2. 4D CT lung

4D CT scans are created by scanning a patient's lung multiple times as they breathe. A 16 slice scanner in cine-mode is used to acquire 25 data volumes each approximately one inch in height of the same location on the patient. The patient is then moved one inch to begin capturing the next slab.

Initial estimates of the lung volume were provided by a belt around the patient's abdomen which measured chest circumference. Control points were placed every 8 voxels in all spatial dimensions, and 4 spans were used to model breathing, producing a total of 57,330 B-spline parameters. The reference image has dimension 56 by 80 by 152 (680,960 voxels). And each of 250 samples has a single breath parameter. So 57,580 parameters are explicitly optimized for in the solver. Optimization took about 3 hours on an AMD Athlon 64 3800. Figure 6 shows a coronal cross section of the lung at three stages during exhalation and a vector map showing tissue motion during breathing. Lung CT reconstructions are possible even without belt measurements, by using an Isomap based pose estimation method [4].

To measure the quality of the deformation model we determine how well the model is able to register slabs with similar pose estimates. For each of the 250 sample slabs we deform the sample slab with pose estimate most similar to it using the deformation map to account for motion due to breathing. The average voxel-wise difference between the sample and its deformed nearest neighbor is found to be 15.8 Hounsfield Units (HU) with a standard deviation of 4.0 HU. In comparison, a direct difference between the sample and its nearest neighbor (without deformation) is 16.2 HU with standard deviation 4.3 HU. Although the improvement is small, a two tailed paired t-test over all 250 samples shows that the difference in mean between the two distributions is significant (p-value  $2.1 \cdot 10^{-4}$ ).

### 5.3. MRI heart/lung

Data was captured as 193 frames of an un-gated cardiac MR during free breathing conditions. Initial pose estimates

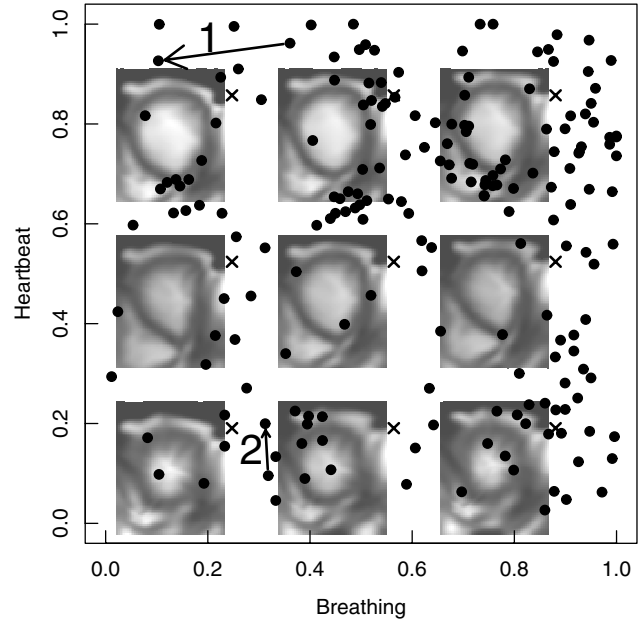


Figure 7. Large dots show the pose estimates of the MRI heart/lung dataset. Reconstructed images are shown at points marked with  $\times$ . The arrow 1 and 2 correspond to pose changes along the breathing and heartbeat axes respectively: the motion corresponding to these pose changes is shown in figure 8.

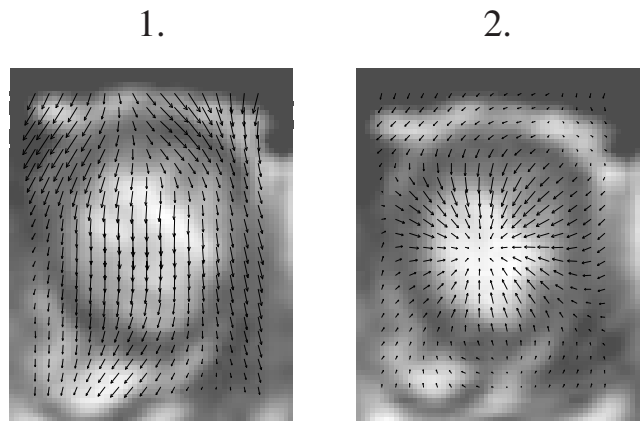


Figure 8. The motion corresponding to a change in pose along arrow 1 and 2 in figure 7.

were computed using Isomap with a modification for embedding onto a cylinder. Results are shown in figure 7. The vertical translation of the heart due to breathing is visible along the x-axis, while the cyclic heartbeat is visible along the y-axis. Figure 8 shows the deformation field that corresponds to motion along the arrows (marked 1 and 2) in figure 7.

Since the images vary in part due to blood flow that appears in some parts of the heartbeat cycle, the deformation field does not always correspond to tissue motion. Integrating a reference image which is able to account for such changes is a key direction of future work.

## 6. Conclusion

We have presented a framework for modeling moving tissue in medical images that can be applied to many different imaging modalities. By, fully modeling the deformation using a reference image and deformation map, we provide a method capable of simultaneously (1) estimating the undeformed appearance of the data volume, (2) estimating arbitrary deformations, and (3) estimating the pose of a data sample. In the future, we hope to use our framework to address some of the shortcomings of current reconstruction methods and provide high-fidelity medical data reconstruction, specifically for 4D CT of lung tissue.

Looking forward, a key future problem in this domain is the verification of automatically generated tissue motion models. The availability of comparative data sets, with known ground truth based on physical markers or phantoms with known motion would allow better quantitative analysis.

## References

- [1] R. H. Byrd, P. Lu, and J. Nocedal. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific and Statistical Computing*, 16(5):1190–1208, 1995.
- [2] R. Chandrashekhara, R. Mohiaddin, and D. Rueckert. Comparison of Cardiac Motion Fields from Tagged and Untagged MR Images Using Nonrigid Registration. In *FIMH*, pages 425–433, 2005.
- [3] J. Ehrhardt, R. Werner, D. Säring, T. Frenzel, W. Lu, D. Low, and H. Handels. An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing. *Med Phys*, 34(2):711–21, 2007.
- [4] M. Georg, J. J. Cannon, A. J. Hope, W. Lu, D. A. Low, and R. B. Pless. Automating 4D CT Reconstruction Using Manifold Learning. In *American Radium Society Annual Meeting*, May 2007.
- [5] H. Handels, R. Werner, T. Frenzel, D. Säring, W. Lu, D. Low, and J. Ehrhardt. Generation of 4D CT Image Data and Analysis of Lung Tumour Mobility during the Breathing Cycle. *Stud Health Technol Inform*, 124, 2006.
- [6] A. J. Hope, M. Georg, J. J. Cannon, J. Hubenschmidt, W. Lu, D. A. Low, and R. B. Pless. Applications of Manifold Learning Techniques in 4D-CT reconstruction. In *International Conference on the use of Computers in Radiation Therapy*, June 2007. Reviewer’s Choice.
- [7] J. Huang, D. Abendschein, V. G. Dávila-Román, and A. A. Amini. Spatio-Temporal Tracking of Myocardial Deformation with a 4D B-Spline Model from Tagged MRI. *IEEE Trans. Med. Imaging*, 18(10):957–972, 1999.
- [8] J. R. McClelland, J. M. Blackall, S. Tarte, A. C. Chandler, S. Hughes, S. Ahmad, D. B. Landau, and D. J. Hawkes. A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. *Medical Physics*, 33(9):3348–3358, September 2006.
- [9] E. Schreiber, G. Chen, and L. Xing. Image Interpolation in 4D CT using a BSpline Deformable Registration Model. *Int J Radiat Oncol Biol Phys*, 2006.
- [10] A. Shashua, A. Levin, and S. Avidan. Manifold pursuit: A new approach to appearance based recognition. In *ICPR*, pages 590–594, 2002.
- [11] R. Souvenir, Q. Zhang, and R. Pless. Image Manifold Interpolation using Free-Form Deformations. In *IEEE International Conference on Image Processing*, pages 1437–1440, October 2006.
- [12] M. Staring, S. Klein, and J. P. W. Pluim. A rigidity penalty term for nonrigid registration. *Medical Physics*, 34(11):4098–4108, November 2007.
- [13] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.
- [14] N. Tustison, V. Dvila-Romn, and A. Amini. Myocardial kinematics from tagged mri based on a 4-d b-spline model. *IEEE Trans Biomed Eng*, 50(8):1038–40, 2003.
- [15] S. Xu, R. Taylor, G. Fichtinger, and K. Cleary. Lung Deformation Estimation and Four-dimensional CT Lung Reconstruction. *Academic Radiology*, 13(9):1082–1092, September 2006.
- [16] Q. Zhang, R. Souvenir, and R. Pless. On Manifold Structure of Cardiac MRI Data: Application to Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*, 1:1092–1098, 2006.