Combining Image Registration, Respiratory Motion Modelling, and Motion Compensated Image Reconstruction

Jamie R. McClelland, Benjamin A. S. Champion, and David J. Hawkes

Centre for Medical Image Computing, University College London, London, UK

Abstract. Respiratory motion models relate the motion of the internal anatomy, which can be difficult to directly measure during image guided interventions or image acquisitions, to easily acquired respiratory surrogate signal(s), such as the motion of the skin surface. The motion models are usually built in two steps: 1) determine the motion from some prior imaging data, e.g. using image registration, 2) fit a correspondence model relating the motion to the surrogate signal(s). In this paper we present a generalized framework for combining the image registration and correspondence model fitting steps into a single optimization. Not only does this give a more theoretically efficient and robust approach to building the motion model, but it also enables the use of 'partial' imaging data such as individual MR slices or CBCT projections, where it is not possible to determine the full 3D motion from a single image. The framework can also incorporate motion compensated image reconstruction by iterating between model fitting and image reconstruction. This means it is possible to estimate both the motion and the motion compensated reconstruction just from the partial imaging data and a respiratory surrogate signal.

We have used a simple 2D 'lung-like' software phantom to demonstrate a proof of principle of our framework, for both simulated 'thick-slice' data and projection data, representing MR and CBCT data respectively. We have implemented the framework using a simple demons like registration algorithm and a linear correspondence model relating the motion to two surrogate signals.

1 Introduction

Respiratory motion is often a problem when acquiring images or planning and guiding interventions (e.g. surgery, radiotherapy) in the abdomen and thorax. It can cause artefacts in reconstructed images, limiting their utility, and can cause misalignment between the planned intervention and the moving anatomy, limiting the accuracy of the guidance and leading to uncertainties in the delivered treatment. One solution to the problem of respiratory motion, which has been proposed for a wide range of different applications, is the use of respiratory motion models [1].

There are three elements to these motion models: the motion of the organ/anatomy of interest, respiratory surrogate data, and a correspondence model. If the motion of

adfa, p. 1, 2011.

interest is known during the procedure (the image acquisition or the image guided intervention) then it can be corrected for, e.g. by performing a motion compensated image reconstruction [2,3], or by 'animating' the treatment/intervention plan to follow the motion [4]. However, it is usually very difficult or impossible to directly measure the full motion of interest during the procedure due to limitations of the imaging equipment and/or impositions made by the intervention. In contrast, the respiratory surrogate data should be easy to acquire during the procedure, but cannot be used directly to compensate for the motion. The respiratory surrogate data is usually one or more simple 1D signals, such as the displacement of the skin surface or diaphragm, or the tidal volume measured with spirometry. The correspondence model relates the surrogate signal(s) to the motion of interest, and is fitted to some prior data before the start of the procedure (Fig 1). During the procedure the model is then used to estimate the motion of interest from measurements of the surrogate signal(s).

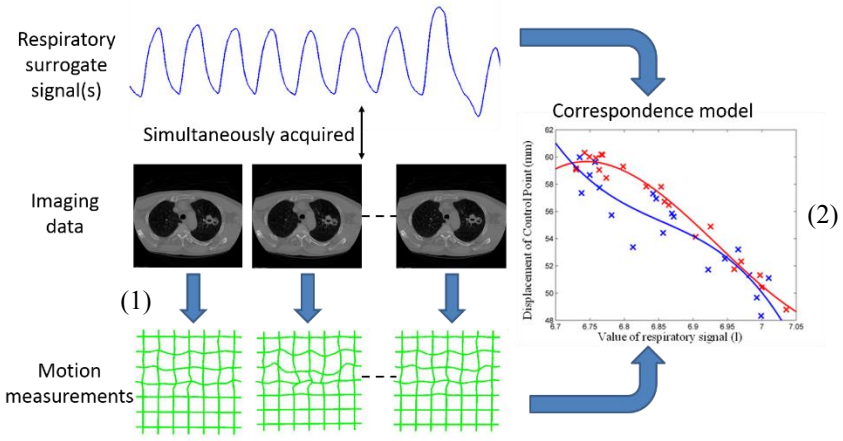

Fig. 1. An illustration of how a respiratory motion model is typically built. Respiratory surrogate data is simultaneously acquired with imaging data. Image registration is then used to determine the motion from the imaging data (1). Once the motion has been determined a correspondence model is fitted which relates the motion to the surrogate data (2).

The prior data used to fit the correspondence models consists of simultaneously acquired surrogate signal(s) and imaging data. The motion models are usually constructed in two distinct steps: 1) determine the respiratory motion from the imaging data using image registration, and 2) fit the correspondence model relating the surrogate signal(s) to the motion determined in step 1 (Fig 1). In this paper we propose a new generalized framework that combines the image registration and the model fitting into a single optimization. Not only does this give a more theoretically efficient and robust approach to determining and modelling the motion, but it also enables the use of 'partial' imaging data such as CBCT projections, or individual MR slices, where it is not possible to determine the full 3D motion from a single image. Additionally, it is straightforward to incorporate motion compensated image reconstruction into this framework using an iterative scheme.

This means the framework is particularly well suited to motion compensated image reconstruction applications, as the motion model can be fitted to the 'unreconstructed' partial imaging data, rather than requiring some previously reconstructed imaging data for fitting the motion model. E.g., some authors have proposed building a motion model from 4DCT data, and using this motion model to motion compensate a CBCT scan acquired at a later time [2]. This can lead to problems when there are differences in the motion and/or anatomy between the 4DCT data and the CBCT scan (as is often seen during the course of radiotherapy). Using the framework proposed here it is possible to build the motion model directly from the CBCT projection data, and then use the model to motion compensate the CBCT reconstruction [5]. No prior 4DCT data is required.

It should be noted that a few publications that could be said to already fit into the framework proposed here, e.g. [5,6]. However, these publications have all been focused on a particular application, registration algorithm, and motion model, whereas the framework presented here is a generalized framework that can be applied to a wide range of applications, registration algorithms, and motion models.

2 Theory and Methodology

2.1 Respiratory Motion Models

To build a respiratory motion model imaging data must be simultaneously acquired with one or more respiratory surrogate signals. The imaging data consists of N_t distinct images, $\mathbf{I}_1 \cdots \mathbf{I}_{N_t}$, each acquired at a different time-point and representing a different respiratory state. These should cover at least one breath cycle, although sometimes several breath cycles will be imaged so that inter-cycle variations can be observed and modelled. A reference-state image, $\mathbf{I}_0$, is also required. This could be one of the images already acquired or it could be some other image, e.g. a high quality breath-hold image.

The typical approach to building respiratory motion models is to first determine the motion from the image data using image registration, and to then fit a correspondence model that relates the motion to the surrogate signals using a methods such as linear least squares [1]. To determine the respiratory motion image registration is performed between $\mathbf{I}_0$ and each of the other images, $\mathbf{I}_n$, where $n = 1 \ldots N_t$. Each registration is usually performed independently. The motion at time-point n can then be represented by the motion parameters, $\mathbf{M}_n$, which describe the spatial transformation resulting from the registration to image $\mathbf{I}_n$.

A respiratory correspondence model is parameterized by a set of model parameters, $\mathbf{R}$, and relates $\mathbf{M}_n$ to the surrogate signals, $\mathbf{S}_n = s_{1,n} \cdots s_{N_s,n}$, where N_s is the number of surrogate signals measured at each time-point, n, i.e. $\mathbf{M}_n$ is a function of $\mathbf{S}_n$ and $\mathbf{R}$:

$$\mathbf{M}_n = f(\mathbf{S}_n, \mathbf{R})$$

E.g. a linear correspondence model relating the motion to two surrogate signals can be parameterized by 3 model parameters (for each motion parameter):

$$\mathbf{M}_n = \mathbf{R}_1 s_{1,n} + \mathbf{R}_2 s_{2,n} + \mathbf{R}_3$$

Many correspondence models have been proposed in the literature, as detailed in [1]. Any of these models, including linear, polynomial, and B-spline models, can be used within the framework presented below.

2.2 Combining Image Registration and Respiratory Motion Modelling

The framework presented in this paper does not follow the typical approach described above. Instead, it directly optimizes the correspondence model parameters on all the image data simultaneously, such that the motion estimated by the correspondence model can be used to transform the reference-state image to best match the image data.

Most image registration algorithms attempt to optimize the value of a cost function which expresses how good the registration is:

$$C(\mathbf{I}_1, \mathbf{I}_2, \mathbf{M})$$

This is a function of the motion parameters, $\mathbf{M}$, and two images, $\mathbf{I}_1$ and $\mathbf{I}_2$, where $\mathbf{I}_2$ is usually the result of transforming another image according to the motion parameters:

$$\mathbf{I}_2 = T(\mathbf{I}_0, \mathbf{M})$$

T is the function that applies the transformation parameterized by $\mathbf{M}$ to the image $\mathbf{I}_0$. Note, in this work $\mathbf{I}_0$ is the moving image, as this is more convenient when using partial imaging data. The cost function, C, usually consists of one or more similarity terms and zero or more constraint terms. A common approach is to calculate the gradient of C with respect to the motion parameters:

$$\frac{\partial C}{\partial \mathbf{M}}$$

and then use an optimization method such as gradient descent or conjugate gradient to find the optimal values of $\mathbf{M}$ that give the best value of C.

From the respiratory correspondence model it is possible to calculate the gradient of the motion parameters, $\mathbf{M}$, with respect to the model parameters, $\mathbf{R}$, e.g. for the linear correspondence model above:

$$\frac{\partial \mathbf{M}}{\partial \mathbf{R}_1} = s_1, \qquad \frac{\partial \mathbf{M}}{\partial \mathbf{R}_2} = s_2, \qquad \frac{\partial \mathbf{M}}{\partial \mathbf{R}_3} = 1$$

The chain rule can be used to find the gradient of the cost function with respect to the correspondence model parameters, i.e.:

$$\frac{\partial C}{\partial \mathbf{R}} = \frac{\partial C}{\partial \mathbf{M}} \frac{\partial \mathbf{M}}{\partial \mathbf{R}}$$

and this can be used to directly optimize the model parameters to give the best value of the cost function on the image data.

The model parameters should be optimized on all the image data simultaneously, so the total cost function over all images needs to be calculated:

$$C_{total} = \sum_{n=1}^{N_t} C_n$$

$$\text{where } C_n = C\left(\mathbf{I}_n, \mathbf{I}_{T_n}, \mathbf{M}_n\right),\ \mathbf{I}_{T_n} = T(\mathbf{I}_0, \mathbf{M}_n),\ \mathbf{M}_n = f(\mathbf{S}_n, \mathbf{R})$$

$\mathbf{S}_n$ is the values of the surrogate signals at time n, and $\mathbf{R}$ is the current model parameters. Likewise, the gradient of the total cost function over all images needs to be calculated:

$$\frac{\partial C_{total}}{\partial \mathbf{R}} = \sum_{n=1}^{N_t} \frac{\partial C_n}{\partial \mathbf{R}}$$

where $\frac{\partial C_n}{\partial \mathbf{R}}$ is the gradient of the cost function with respect to the model parameters calculated using image, $\mathbf{I}_n$, and surrogate signals, $\mathbf{S}_n$. The gradient of the total cost function can then be used to find the optimal values of $\mathbf{R}$ using the same optimization method as used for standard image registration between two images, e.g. gradient descent.

2.3 Using Partial Imaging Data

With the framework described above it is possible to use 'partial' imaging data, such as single slices or projections, instead of full images, providing the total partial image data still sufficiently samples the motion over the region of interest. When using partial imaging data it is necessary to model the image acquisition/reconstruction process:

$$\mathbf{P}_n = A_n(\mathbf{I}_n) + \boldsymbol{\varepsilon}_n$$

where $\mathbf{P}_n$ is the partial imaging data at time n, $\mathbf{I}_n$ is the full image at time n, $\boldsymbol{\varepsilon}_n$ is the imaging noise, and A_n is the function which simulates the image acquisition at time n. E.g. for projection data A_n would be the forward-projection operator and for slice data A_n would be the slice selection profile. Using A_n the cost function can be calculated for each partial image, $\mathbf{P}_n$:

$$C_n = C\left(\mathbf{P}_n, A_n\left(\mathbf{I}_{T_n}\right), \mathbf{M}_n\right)$$

To calculate $\frac{\partial C}{\partial \mathbf{M}}$ the difference image between the two input images is often required. When using partial imaging data it is necessary to transform the difference image from the space of the partial images, $\mathbf{P}_n$, into the space of the full images, $\mathbf{I}_n$, as the spatial transform parameterized by $\mathbf{M}_n$ is defined is the space of the full images. This is done using the adjoint of the function A_n, written as A_n^*, e.g. the adjoint of the forward projection operator is the back projection operator.

For example, if using Sum of Squared Differences (SSD) as the cost function and the linear correspondence model with two surrogate signals from the previous examples, then:

$$C_n = \left\| \mathbf{P}_n - A_n\left(\mathbf{I}_{T_n}\right) \right\|_2^2$$

$$\frac{\partial C_n}{\partial \mathbf{R}_1} = \left(-2A_n^* \left(\mathbf{P}_n - A_n\left(\mathbf{I}_{T_n}\right)\right) \nabla \mathbf{I}_{T_n}\right) s_1$$

$$\frac{\partial C_n}{\partial \mathbf{R}_2} = \left(-2A_n^* \left(\mathbf{P}_n - A_n\left(\mathbf{I}_{T_n}\right)\right) \nabla \mathbf{I}_{T_n}\right) s_2$$

$$\frac{\partial C_n}{\partial \mathbf{R}_3} = -2A_n^* \left(\mathbf{P}_n - A_n\left(\mathbf{I}_{T_n}\right)\right) \nabla \mathbf{I}_{T_n}$$

2.4 Incorporating Motion Compensated Image Reconstruction

The framework as described above assumes we have a full reference-state image, $\mathbf{I}_0$, available. However, the framework can easily be combined with motion compensated image reconstruction in an iterative approach, meaning a prior reference-state image is not required.

A number of publications have described how motion compensated image reconstructions can be performed for different imaging modalities, providing the motion is known [2,3]. The motion compensated image reconstruction can be combined into the framework described above by iterating between the image reconstruction and model fitting [5]. Firstly, a standard (non-motion compensated) reconstruction is performed from the partial imaging data. The result will contain blurring and other artefacts caused by motion, but can be used as an initial estimate of $\mathbf{I}_0$. This is used to fit the model parameters, $\mathbf{R}$, as described above. The fitted motion model is then used to perform a motion compensated reconstruction and obtain a better estimate of $\mathbf{I}_0$. The process then continues to iterate between fitting $\mathbf{R}$ using the most recent $\mathbf{I}_0$, and performing a motion compensated image reconstruction of $\mathbf{I}_0$ using the most recent values of $\mathbf{R}$.

As the model fitting framework described above is itself iterative, and will likely be performed using a multi-resolution scheme (as such schemes are commonly employed in image registration algorithms), there are different options for how often to perform the motion compensated image reconstruction: 1) every time $\frac{\partial C}{\partial \mathbf{R}}$ is recalculated, 2) after fitting the model parameters at the current resolution level, 3) after fully fitting the model parameters at all resolution levels.

3 Phantom Experiments

In order to demonstrate our framework we have performed a number of experiments using a simple 2D 'lung-like' software phantom of size 128 x 128 pixels (Fig 2a). To 'animate' the phantom we used a linear correspondence model with two surrogate signals. The first signal, s_1, was the displacement of the skin surface measured from a real patient, and the second signal, s_2, was the temporal gradient of the first signal (Fig 2b),

i.e. the 2nd signal is a derived surrogate signal [1]. The motion parameters, **M**, represented a deformation field with a 2D vector at each pixel. As the reference-state image corresponded to surrogate values of 0, the 'offset' model parameter, $\mathbf{R}_3$, was not required, so there were 4 model parameters for each pixel (the x and y components of $\mathbf{R}_1$ and $\mathbf{R}_2$). We manually defined the true correspondence models parameters, $\mathbf{R}^{\text{true}}$, so that they were smoothly varying over the phantom (Fig 2c) and produced plausible looking respiratory motion that includes non-linear deformation and both intra- and inter-cycle variation (the motion is different during inhalation and exhalation and during different breath cycles).

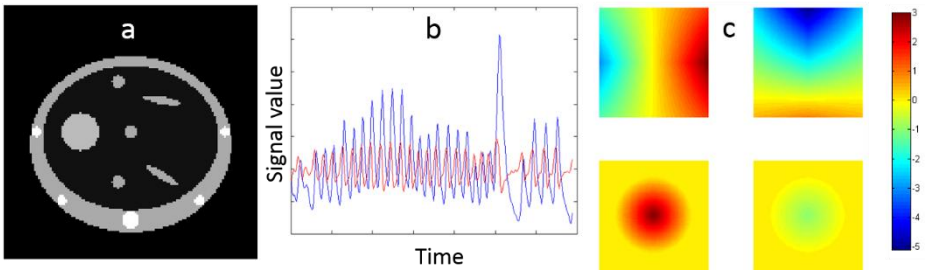

Fig. 2. a) 2D 'lung-like' software phantom. b) surrogate signals used to animate the phantom, s_1 is plotted in blue and s_2 is plotted in red. c) top-left: x component of $\mathbf{R}_1^{\text{true}}$, top-right: y component of $\mathbf{R}_1^{\text{true}}$, bottom-left: x component of $\mathbf{R}_2^{\text{true}}$, bottom-right: y component of $\mathbf{R}_2^{\text{true}}$

We then simulated the acquisition of 3 different types of imaging data (Fig 3):

1. Full 2D images representing 4DCT like data. 13 images were simulated, covering one complete breath cycle. 5% Gaussian noise was added to the images. Fig 3a.
2. 1D slices representing MR like data. The slices were 5 pixels wide and a Gaussian slice profile was used. The slice spacing was 1 pixel (i.e. overlapping slices), and slices were acquired from each location in both the x and y directions. Each slices was acquired 3 times, giving a total of 768 slices. 5% Gaussian noise was added to the slices. Fig 3b.
3. 1D projections representing CBCT projection like data. Projections were acquired with an angular spacing of 1° and a total of 360 projections. 1% Gaussian noise was added to the projection data. Fig 3 c.

To fit the correspondence model parameters, $\mathbf{R}^{\text{fit}}$, to the imaging data we implemented a simple 'demons-like' registration algorithm. No constraint term was used but $\frac{\partial C}{\partial \mathbf{R}}$ was smoothed with a Gaussian filter (standard deviation = 5 pixels) prior to updating the model parameters. SSD was used as the similarity measure. The optimization was done using gradient descent and a multi-resolution scheme with 3 resolution levels, and stopped when the cost function improved by less than 0.1%. For the slice and projection data the model fitting was performed both using the original reference-state image used to simulate the image acquisition, and using motion compensated image reconstruction from the partial data. The motion compensated reconstruction was updated after fitting the model parameters at the each resolution level (option 2 in section 2.4).

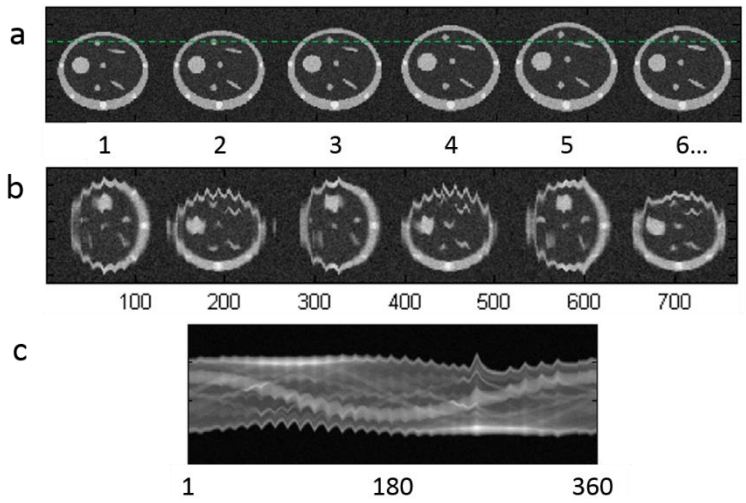

Fig. 3. Simulated imaging data. a) full 2D image (6 of 13 images shown), b) 1D slices, c) 1D projections. Image number is plotted on the x axis.

4 Results

Table 1 gives the mean, standard deviation, and maximum of the absolute difference between $\mathbf{R}^{true}$ and $\mathbf{R}^{fit}$ for each of the fitted models, as well as the absolute values of $\mathbf{R}^{true}$ for comparison. The summary statistics are calculated over all 4 model parameters for all pixels inside the phantom. To assess how well the original motion can be reproduced by the fitted models we also calculated the Euclidean distance between $\mathbf{M}^{true}$ and $\mathbf{M}^{fit}$ (the pixel displacements generated by $\mathbf{R}^{true}$ and $\mathbf{R}^{fit}$ respectively). Table 2 gives the mean, standard deviation, and maximum values calculated over all pixels inside the phantom and all time-points used in the simulated image acquisitions. The values for $\mathbf{M}^{true}$ are also given to indicate how much motion occurred (these values are from the slice acquisition as this had the most time-points, but the values for the other acquisitions are similar).

Table 1. Summary statistics for the absolute differences between $\mathbf{R}^{true}$ and $\mathbf{R}^{fit}$ for each of the fitted models, and for absolute values of $\mathbf{R}^{true}$. F: full 2D images, S: 1D slices, S-MCR: 1D slices using motion compensated image reconstruction, P: 1D projections, P-MCR: 1D projections using motion compensated image reconstruction.

	$\mathbf{R}^{true}$	$\mathbf{R}^{true}$ - $\mathbf{R}^{fit}$ for model fitted to:				
		F	S	S-MCR	P	P-MCR
Mean	0.89	0.07	0.17	0.23	0.22	0.33
Std. Dev.	0.80	0.10	0.24	0.31	0.26	0.30
Max.	3.87	0.69	1.53	1.98	1.37	1.56

Table 2. Summary statistics for the Euclidean distance between $\mathbf{M}^{\text{true}}$ and $\mathbf{M}^{\text{fit}}$ for each of the fitted models, and the Euclidean distance of $\mathbf{M}^{\text{true}}$. F: full 2D images, S: 1D slices, S-MCR: 1D slices using motion compensated image reconstruction, P: 1D projections, P-MCR: 1D projections using motion compensated image reconstruction. All values are in pixels.

	$\|\mathbf{M}^{\text{true}}\|_2$	$\left\|\mathbf{M}^{\text{true}} - \mathbf{M}^{\text{fit}}\right\|_2$ for model fitted to:				
		F	S	S-MCR	P	P-MCR
Mean	1.65	0.10	0.22	0.30	0.33	0.49
Std. Dev.	1.60	0.10	0.23	0.28	0.38	0.43
Max.	19.90	0.78	9.75	10.50	6.69	6.49

It can be seen from Tables 1 and 2 that when using the full images the models can be fitted very well and the motion can be reproduced very accurately. The results are worse for the partial imaging data, as may be expected, but the models are still fitted well and the majority of the motion is reproduced accurately. The mean values are lower for the slice data, but the maximum values are lower for the projection data. The results are worse when performing motion compensated reconstructions than when using the original reference-state image, as would be expected. However, the model parameters and the motion can still be recovered reasonably well, even when the reference-state image is not available.

Figure 4 shows the motion compensated reconstructions from the slice and projection data using no motion (i.e. a standard reconstruction), $\mathbf{M}^{\text{fit}}$, and $\mathbf{M}^{\text{true}}$. It can be seen that the motion compensated reconstructions are clearly superior to the non-motion compensated reconstructions. The reconstructions using $\mathbf{M}^{\text{fit}}$ are very similar to those using $\mathbf{M}^{\text{true}}$, indicating the fitted models can adequately reproduce the motion for the purpose of performing motion compensated image reconstructions.

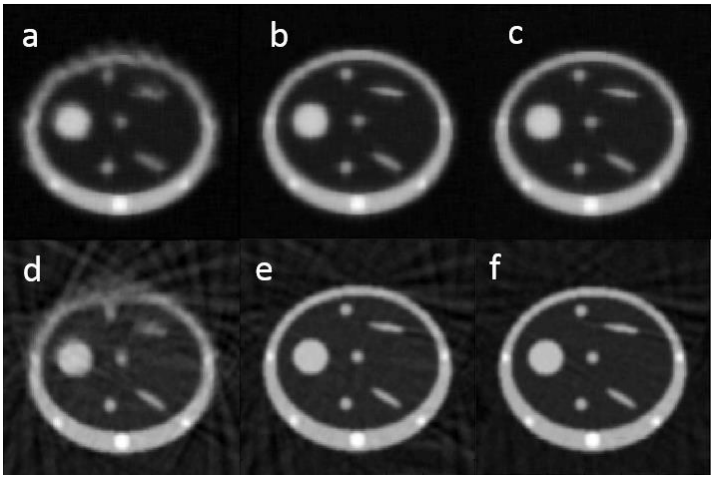

Fig. 4. Motion compensated reconstructions from slice data (top) and projection data (bottom) using (a, d) no motion compensation, (b, e) $\mathbf{M}^{\text{fit}}$, and (c, f) $\mathbf{M}^{\text{true}}$.

5 Conclusions

In this paper we have presented a general framework that can be used to combine image registration, respiratory motion modelling, and motion compensated image reconstruction. This framework can be used with any registration algorithm that uses the gradient of the cost function to optimize the registration and any respiratory correspondence model where the gradient of the motion with respect to the model parameters can be calculated. This includes most registration algorithms that are commonly used, and all correspondence models that we are aware of in the literature.

To demonstrate this framework we implemented it using a 'demons-like' registration algorithm and a linear correspondence model using two surrogate signals. We used a simple 2D software phantom to simulate the acquisition of full images, slice image data, and projection image data. For all types of image data the framework was able to fit the model directly to the image data and to give a good estimate of the true motion. We also showed that motion compensated image reconstruction can be included in the framework using an iterative approach, and that we were able to reconstruct images that were very similar to those obtained when the true motion was used.

On-going work includes implementing the framework with an efficient open-source registration package based on the B-spline registration algorithm (NiftyReg[1]), and thoroughly validating the framework using a variety of real clinical data, including MR, CBCT projections, and 4DCT. The framework can then be applied to a wide range of medical imaging applications that are affected by respiratory motion. In each of these applications it will be necessary to investigate the ideal combination of image data, registration algorithm (and settings), respiratory surrogate signals and correspondence model, and image reconstruction algorithm (if required).

References

1. McClelland, J.R., Hawkes, D.J., Schaeffter, T., King, A.P.: Respiratory motion models: A review. Medical Image Analysis 17, 19-42 (2013)
2. Rit, S., Wolthaus, J.W.H., van Herk, M., Sonke, J.-J.: On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion. Medical Physics 36, 2283-2296 (2009)
3. Batchelor, P.G., Atkinson, D., Irarrazaval, P., Hill, D.L.G., Hajnal, J., Larkman, D.: Matrix description of general motion correction applied to multishot images. Magnetic Resonance in Medicine 54, 1273-1280 (2005)
4. Schweikard, A., Glosser, G., Bodduluri, M., Murphy, M.J., Adler, J.R.: Robotic motion compensation for respiratory movement during radiosurgery. Computer Aided Surgery 5, 263–277 (2000)
5. Martin, J., McClelland, J., Yip, C., Thomas, C., Hartill, C., Ahmad, S., O'Brien, R., Meir, I., Landau, D., Hawkes, D.: Building motion models of lung tumours from cone-bean CT for radiotherapy applications. Physics in Medicine and Biology 58, 1809-1822 (2013)
6. Hinkle, J., Szegedi, M., Wang, B., Salter, B., Joshi, S.: 4D CT image reconstruction with diffeomorphic motion model. Medical Image Analysis 16, 1307-1316 (2012)

[1] http://sourceforge.net/projects/niftyreg/