Respiratory Motion Modelling

Mingzhou Hu

April 2019

Abstract: In this project, 10 different respiratory motion models are built and evaluated, which associate the motion of the internal anatomy to easily measured surrogate signals. Five techniques, including visual assessment, deformation field error, landmark error, AIC/BIC and uncertainty estimation are involved to evaluate the performance of these models. The limitations and future work will be discussed at the end of the report.

1. Introduction

Respiratory motion often causes problems when acquiring images and guiding interventions, such as surgery and radiotherapy in the abdominal and thoracic organs [1]. During image acquisition, it could cause artefacts and thus lowers the quality of reconstructed images. While during image guided interventions, it could lead errors and uncertainties in the exact location of the anatomy, and hence lowers the accuracy of guidance. Massive approaches have been proposed to solve problems, including breath-holding, respiratory gating and motion tracking within X-ray. However, all three solutions mentioned above have disadvantages. Breath-holding usually limits acquisition or intervention time to under 30 seconds, which is insufficient for many cases. Respiratory gating only acquires or uses imaging data in a limited region, which significantly increases the time for acquisition or intervention. Using X-ray to implement motion tracking is an invasive approach and motion information is only available for a limited region with landmarks but not for the whole region of interest [2].

Due to the drawbacks of the approaches mentioned above, respiratory motion models have been proposed to use correspondence models to relate the motion of the internal anatomy to easily measured respiratory surrogate signals [1], such as the motion of skin surface and diaphragm. Fig 1.1 and Fig 1.2 show the illustrations of how a correspondence model is typically build and used. Before treatment, at least one respiratory surrogate signals are collected simultaneously with imaging data. Then the internal motion is measured from the imaging data, such as using deformable image registration. Lastly, a correspondence model is fitted, which associates to internal motion to surrogate signals (Fig 1.1). During treatment, the correspondence model is responsible for estimating the internal motion corresponding to measured surrogate signals (Fig 1.2). These motion estimations could potentially be used for gated and tracked treatments guidance.
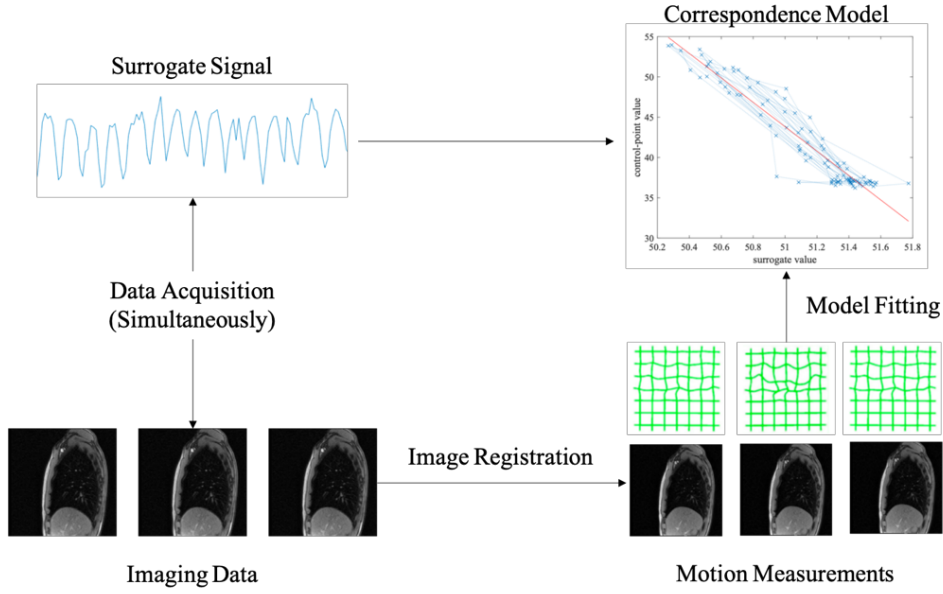
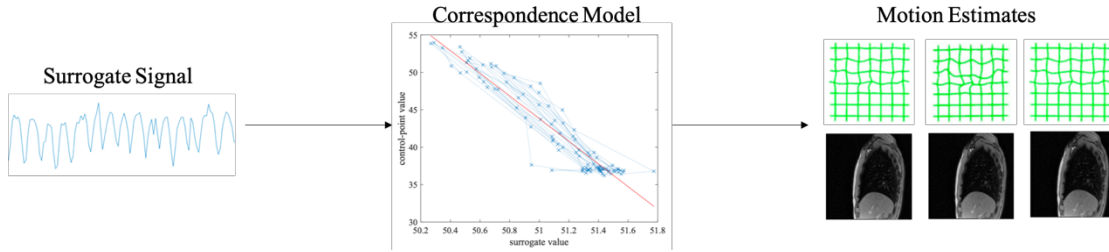*Figure 1.1: Correspondence model building, modified from [3, Fig 1.1]*

Figure 1.2: Correspondence model usage, modified from [3, Fig 1.2]

2. Data and Models

This chapter will introduce the relevant data given by a healthy volunteer, containing sagittal MR images, registration results, anatomical landmarks and deformation field mask. 10 different respiratory motion models and modelling processes will be introduced in this chapter as well.

2.1 Data

In this project, 2D sagittal cine-MR images are used to measure internal motion and acquire surrogate signals. Since the imaging data is 2D, just 2D motion can be modelled. While it is not problematic in this case, as generally most of the motion of the lung is in Anterior-Posterior (AP) and Superior-Inferior (SI) directions, which is the reason to use sagittal slices. Additionally, 2D data is a necessary compromise for capturing the

motion with sufficiently high spatial and temporal resolution. The surrogate signals are generated from the motion of the skin surface of the lung as seen in the MRI data. In this case, 1500 images in *nifti* format are provided and they are captured at about 3fps with an in-slice resolution of 2mm×2mm and a slice thickness of 10mm.

The internal motion is estimated using B-spline deformable image registration which allows for sliding between two regions. The first region represents the inner organs, such as lung and liver, and the second represents the outer anatomy, such as rib cage and muscles. A series of deformations of each region are stored in a regular Control Point (CP) grid. In this case, totally 3000 registration results images for two regions are provided. The purpose of registration is to find the mapping from a target image (each image in the time series) to a source image (*0007.nii.gz*) to wrap the source image for which aligning better with the target image. For the sliding registration, the signed distance map *segmentation/0007_sdt.nii.gz is also needed* to internally determine which B-spline transformation to use.

The information of four anatomical landmarks in blood vessels in each MR image is stored as *landmark_pos_phys.mat.* These landmarks are used to evaluate the registration and motion models. Deformation field mask (*0007_mask.nii.gz*) in the source image domain is also needed when computing the deformation field error to ignore pixels outside the lung.

2.2 Models

The data mentioned above are used to build 10 different respiratory motion models:

1. Linear model with one surrogate signal: $x = c_1 s + c_0$.
2. 2nd order polynomial model with one surrogate signal: $x = c_2 s^2 + c_1 s + c_0$.
3. 3rd order polynomial model with one surrogate signal: $x = c_3 s^3 + c_2 s^2 + c_1 s + c_0$.
4. Linear model with two surrogate signals: $x = c_2 s_2 + c_1 s_1 + c_0$.
5. Polynomial model with two surrogate signals: $x = c_2 s_2{}^2 + c_1 s_1 + c_0$.
6. Linear model with one surrogate signal and the signal gradient: $x = c_2 g + c_1 s + c_0$.
7. Polynomial model with one surrogate signal and the signal gradient: $x = c_2 g^2 + c_1 s + c_0$.
8. Linear separate models for inhalation and exhalation: $\begin{cases} x_{in} = c_1 s_{in} + c_0 \\ x_{ex} = c_1 s_{ex} + c_0 \end{cases}$.
9. 2nd order separate models for inhalation and exhalation: $\begin{cases} x_{in} = c_2 s_{in}{}^2 + c_1 s_{in} + c_0 \\ x_{ex} = c_2 s_{ex} + c_1 s_{ex} + c_0 \end{cases}$.
10. 3rd order separate models for inhalation and exhalation:
$\begin{cases} x_{in} = c_3 s_{in}{}^3 + c_2 s_{in}{}^2 + c_1 s_{in} + c_0 \\ x_{ex} = c_3 s_{ex}{}^3 + c_2 s_{ex}{}^2 + c_1 s_{ex} + c_0 \end{cases}$.

where x is the deformation described at each control point, s_n are the surrogate signals values, g is the signal gradient value and c_n are the model coefficients which need to be fitted to the data.

In this project, two surrogate signals are generated for building correspondence models. The first surrogate signal is generated by finding the AP location of the skin surface at SI= 150 pixels in each of the images, as seen in Fig 2.1. The location is determined by finding the first point in the image corresponding to an intensity of 20 when coming from the anterior side. The second surrogate signal is generated by the same mean as mentioned before but at SI=75 pixels on the skin surface, as seen in Fig 2.2.
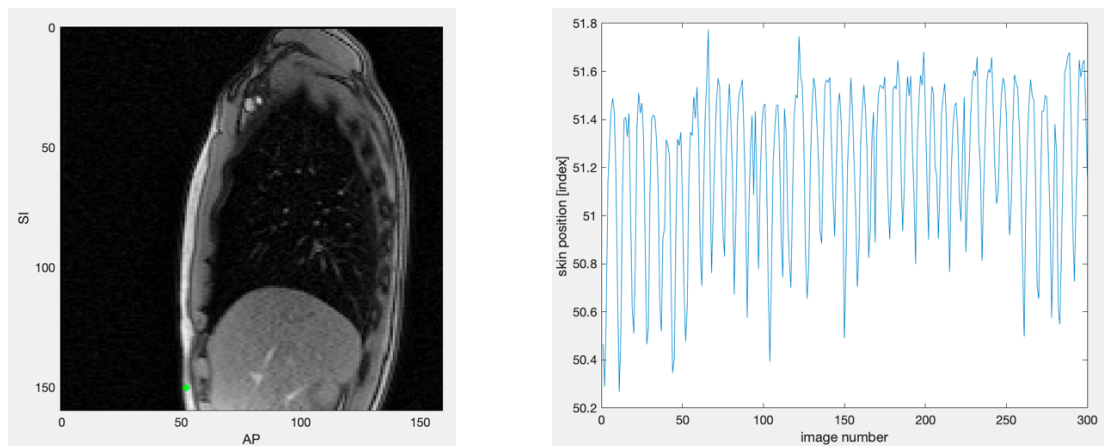
Figure 2.1: Location of skin surface at SI=150 and surrogate signal
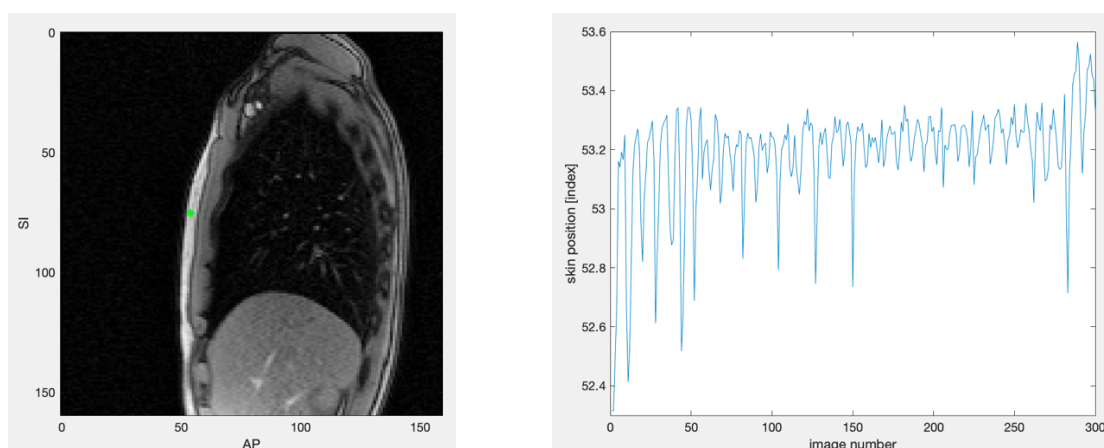
Figure 2.1: Location of skin surface at SI=75 and surrogate signal

The gradients of surrogate signal are determined by computing the difference of surrogate signal values between adjacent images. The respiratory phases are determined by finding the local maximum and minimum of the surrogate signal. The local maximums generally represent the end-exhalation which can be regarded as the peaks of the surrogate signals in Fig 2.1 and Fig 2.2, and the gradient values change from positive to negative. On the contrary, the local minimums represent the end-inhalation which can be regarded as the troughs of the surrogate signals in Fig 2.1 and Fig 2.2, and the gradient values change from negative to positive.

3. Results and Evaluation

In this project, the first 100 registration results are used to fit the models, while the remaining 1400 registration results are regarded as test images to evaluate the motion models.

3.1 Results

The SI deformation for CP 44,38 of region 1 and the fitting results to this CP deformation for model 1, 2, 3, 4, 7 and 8 are showed in Fig 3.1-3.4. We can see from Fig 3.1 that the respiratory motion for model 1, 2 and 3 has linear, quadratic and cubic relationships with the surrogate signal respectively. From Fig 3.2, we know that the respiratory motion for model 4 has linear relationship with both of two surrogate signals. From Fig 3.3, we can see that the respiratory motion for model 7 has linear and quadratic relationships with the surrogate signals and the gradient respectively. It is obviously from Fig 3.5 that both of inhalation and exhalation have linear relationships with the surrogate signals for model 8.
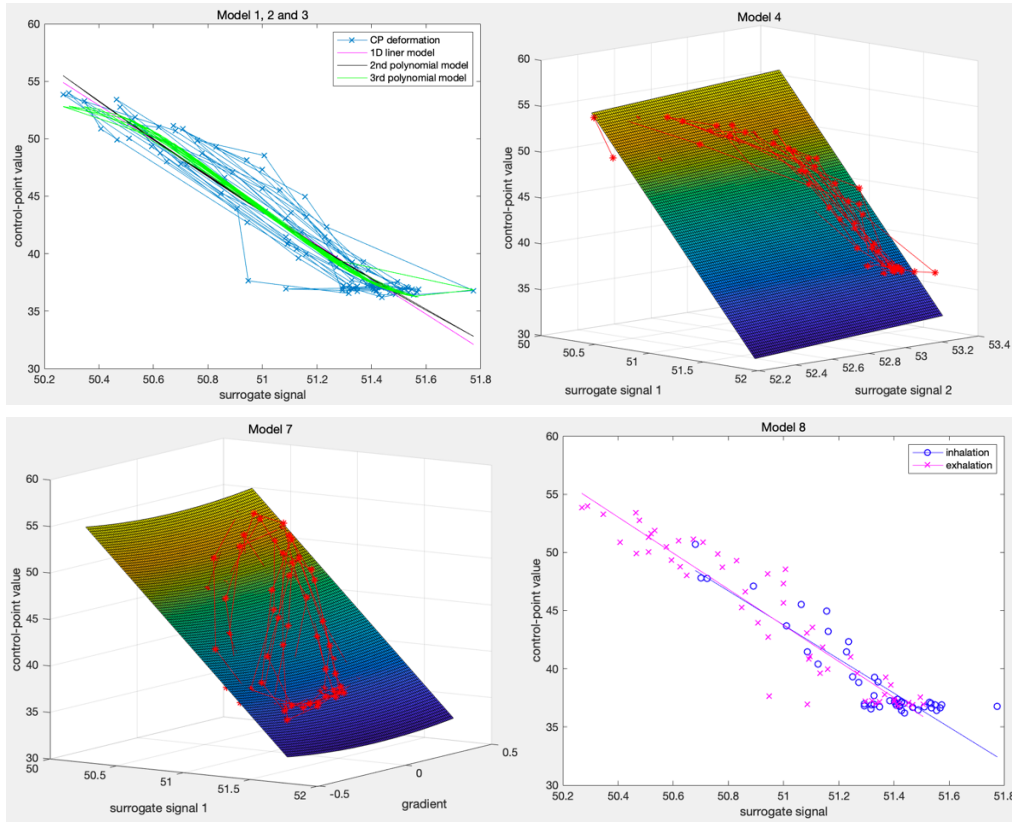

Figure 3.1-3.4: Plots of fitting results for CP 44,38 for model 1, 2, 3, 4, 7 and 8

Once the model fittings for CP 44,38 are successful, the separate models for every CP deformation will be fitted and the residual fitting error for CP 44,38 and over all CP deformations will be computed, as seen in Table 3.1. We can see from the table below that the polynomial models have smaller residual fitting errors than linear models and

the separate models for inhalation and exhalation have smaller residual fitting errors than non-separate models. Therefore, in terms of RSS, polynomial and separate models are better than linear and non-separate models respectively.

Model (refer to 2.2)	RSS for CP 44,38	RSS for all CPs
1 (1D linear)	328.7474 (worst)	2.7207×10^6
2 (1D 2nd)	325.2705	2.6749×10^6
3 (1D 3rd)	271.6202	2.5533×10^6
4 (2D linear)	324.8580	2.6740×10^6
5 (2D polynomial)	324.9140	2.6741×10^6
6 (2D linear)	328.6826	2.6502×10^6
7 (2D polynomial)	325.4646	2.7086×10^6
8 (separation linear)	326.4324	NaN
9 (separation 2nd)	299.2970	NaN
10 (separation 3rd)	265.1597 (best)	NaN

Table 3.1: RSS for CP 44,38 and all CPs

3.2 Evaluation

The model estimates are assessed by five following techniques: visual assessment, deformation field error, landmark error, AIC/BIC and uncertainty estimation.

Visual Assessment: Fig 3.5 and Fig 3.6 show the visual assessment and residual motion for model 2, 5 and 7. We can see from these two figures that there is no significant difference between registration image and deformed test images, and the residual motion for the three models is also really similar. Actually, all 10 models in this project have the similar results in terms of visual assessment and residual motion. Hence, we plot the residual motion for the first 7 models to process more accurate analysis, as seen in Fig 3.7. We can see from Fig 3.7 that the first 3 models have smaller residual motion than the last 4 models, hence, 1D models have better performance than 2D models in this case. The reason which leads to the results may that the 2D models overfit the data.
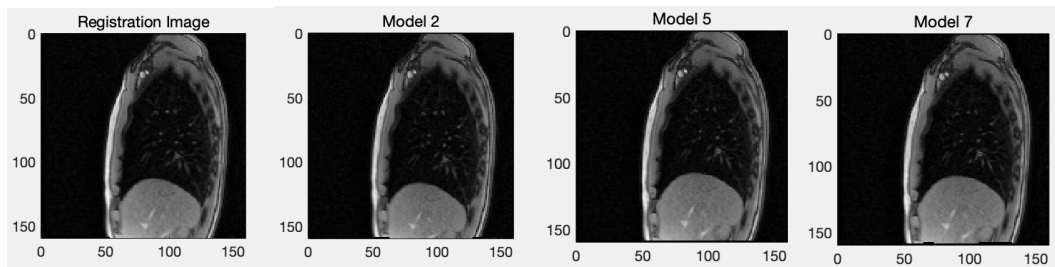

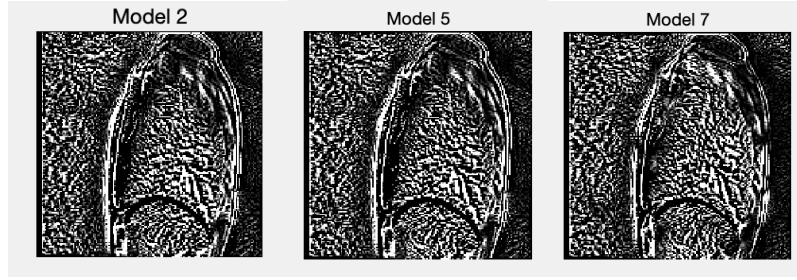*Figure 3.5: Visual assessment for model 2, 5 and 7*

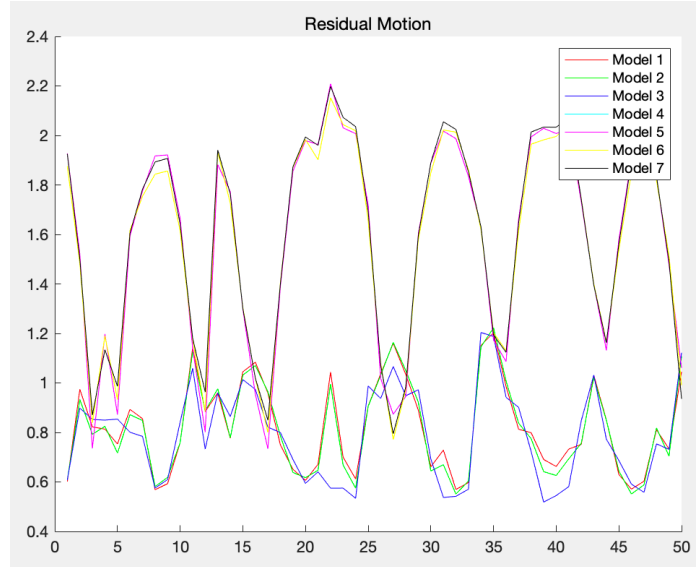*Figure 3.6: Residual motion for model 2, 5 and 7*

Figure 3.7: Plot of residual motion for the first 7 models

Deformation field error: The DFE is the difference between the deformation fields caused by the estimated transformation and by the registration result for the target image respectively. The plot of DFE for the first 50 test images for the first 7 models is shown in Fig 3.8. It is obviously from Fig 3.7 and Fig 3.8 that the trend for DFE is really similar to the trend for residual motion. It means that, in this case, 1D models have better performance than 2D models as well. Table 3.2 presents the same results as we mentioned above, which shows the mean of DFE for 1400 test images. To be more specific, the 1D 3rd polynomial model is the best model and the 2D linear model with the surrogate signal and the gradient is the worst model in this case. The reason which leads to the results may that the 2D models overfit the data.
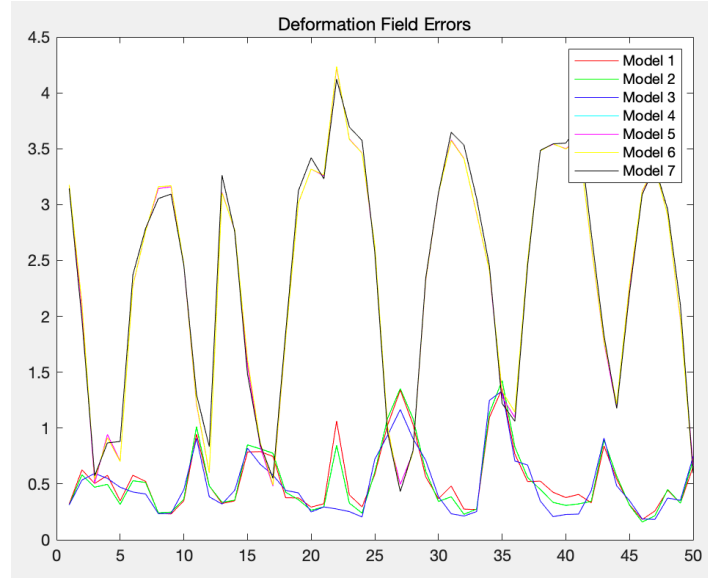
Figure 3.8: Plot of DFE for the first 7 models

Deformation Field Error			
Model (refer to 2.2)	AP (mean)	SI (mean)	L2 norm (mean)
1 (1D linear)	0.3207	1.1018	1.1557
2 (1D 2nd)	0.3232	0.9928	1.0535
3 (1D 3rd)	0.3124	0.8284	0.9015 (best)
4 (2D linear)	0.3135	1.0978	1.1492
5 (2D polynomial)	0.3133	1.0977	1.1490
6 (2D linear)	0.3726	1.1104	1.1800 (worst)
7 (2D polynomial)	0.3439	1.1184	1.1799

Table 3.2: The mean of DFE for all test images

Landmark Error: The landmark error is the difference between the location of truth landmark and the location of estimated landmark. Fig 3.8 shows two examples of locations estimates of four landmarks for model 2 and 3. The red points are truth landmarks and the blue points are estimated landmarks. From Fig 3.8, we could know that the location estimates of landmarks are accurate. The mean of errors for four landmarks estimates for all test images for the first 7 models are presented in Table 3.3-3.6. In terms of first three landmarks, 1D 3rd order model has the smallest L2 norm error, hence it is the best model, and 2D liner model with the surrogate signal and the gradient is the worst model is the worst model in this case. As for the last landmark, 1D 3rd order model is still the best model, while the worst model is 2D linear model with two surrogate signals.
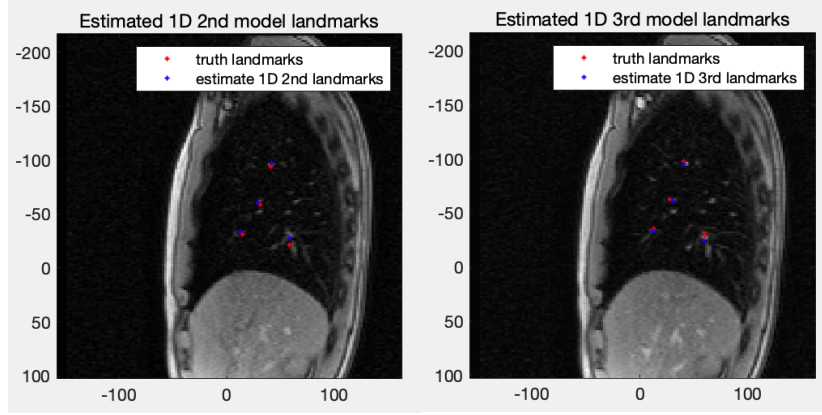
Figure 3.8: Locations estimates of landmarks for model 2 and 3

Landmark 1			
Model (refer to 2.2)	AP (mean)	SI (mean)	L2 norm (mean)
1 (1D linear)	0.3162	2.4221	2.4440
2 (1D 2nd)	0.2577	2.4557	2.4732
3 (1D 3rd)	0.5702	1.5372	1.6182 (best)
4 (2D linear)	0.2647	2.4617	2.4921
5 (2D polynomial)	0.2669	2.4606	2.4916
6 (2D linear)	0.5382	2.4283	2.5549 (worst)
7 (2D polynomial)	0.3305	2.4135	2.4386

Table 3.3: The mean of errors for landmark 1 for all test images

Landmark 2			
Model (refer to 2.2)	AP (mean)	SI (mean)	L2 norm (mean)
1 (1D linear)	0.2442	1.9318	1.9484
2 (1D 2nd)	0.1037	1.8305	1.8368
3 (1D 3rd)	0.1857	1.0094	1.0491 (best)
4 (2D linear)	0.2573	1.9777	2.0183
5 (2D polynomial)	0.2605	1.9768	2.0184
6 (2D linear)	0.5019	1.9716	2.1187 (worst)
7 (2D polynomial)	0.2500	1.9229	1.9404

Table 3.4: The mean of errors for landmark 2 for all test images

Landmark 3			
Model (refer to 2.2)	AP (mean)	SI (mean)	L2 norm (mean)
1 (1D linear)	0.2084	1.3893	1.4137
2 (1D 2nd)	0.3134	1.1734	1.2221
3 (1D 3rd)	0.6130	0.7549	1.0910 (best)
4 (2D linear)	0.2192	1.4297	1.4692
5 (2D polynomial)	0.2210	1.4291	1.4692
6 (2D linear)	0.5273	1.4394	1.6207 (worst)
7 (2D polynomial)	0.2342	1.3793	1.4108

Table 3.5: The mean of errors for landmark 3 for all test images

Landmark 4			
Model (refer to 2.2)	AP (mean)	SI (mean)	L2 norm (mean)
1 (1D linear)	0.9687	4.0469	4.1696
2 (1D 2nd)	0.7724	3.8710	3.9537
3 (1D 3rd)	0.8907	2.3697	2.6616 (best)
4 (2D linear)	0.8648	4.1431	4.2628 (worst)
5 (2D polynomial)	0.8676	4.1415	4.2624
6 (2D linear)	0.9410	4.0542	4.2475
7 (2D polynomial)	0.9788	4.0219	4.1465

Table 3.6: The mean of errors for landmark 4 for all test images

AIC and BIC: AIC is Akaike Information Criterion and BIC is Bayesian Information Criterion, which are computed and for both AP and SI directions for two regions in this case. The results of the mean of AIC and BIC for all CPs for the first 7 models are presented in Table 3.7 and 3.8. We find that it is very hard to determine the best and worst models through the two tables below. For example, in terms of AP, model 5 is the best model for both two regions, while for SI, model 5 is the worst model for region 1. It is strange that either for AP or for SI, model 3 is the worst model for region 2, which may result from the penalty for increased parameters.

AIC				
	Region1		Region2	
Model (refer to 2.2)	AP (mean)	SI (mean)	AP (mean)	SI (mean)
1 (1D linear)	-3229.1	-2715.0	-1299.6	-1202.8
2 (1D 2nd)	-3099.7	-2961.9	-1026.4	-1213.5
3 (1D 3rd)	-3069.2	-2911.2	-885.1	-991.5
4 (2D linear)	-3247.1	-2689.3	-1329.2	-1209.7
5 (2D polynomial)	-3247.7	-2688.3	-1329.7	-1209.9
6 (2D linear)	-3029.7	-2750.2	-1262.8	-1148.8
7 (2D polynomial)	-2940.3	-2707.4	-1293.0	-1197.7

Table 3.7: The mean of AIC for all CPs

BIC				
	Region1		Region2	
Model (refer to 2.2)	AP (mean)	SI (mean)	AP (mean)	SI (mean)
1 (1D linear)	-3213.3	-2699.3	-1283.9	-1197.0
2 (1D 2nd)	-3078.7	-2940.9	-1005.5	-1192.5
3 (1D 3rd)	-3043.0	-2885.0	-858.9	-965.3
4 (2D linear)	-3220.9	-2663.1	-1303.0	-1183.5
5 (2D polynomial)	-3221.5	-2662.1	-1303.5	-1183.7
6 (2D linear)	-3003.5	-2724.0	-1236.6	-1222.6
7 (2D polynomial)	-2914.0	-2681.2	-1266.8	-1171.5

Table 3.8: The mean of BIC for all CPs

Uncertainty Estimation: Wild bootstrapping is used to estimate the distribution of model parameters and their uncertainty for model 1, 2 and 3, because wild bootstrapping is suitable for the data with different variances. Fig 3.9-3.11 show the results of parameters uncertainty for model 1-3. From the results, we could know that the parameters uncertainty is higher when the order of the model is higher, hence, in this case, the linear model can be regarded as the best model.
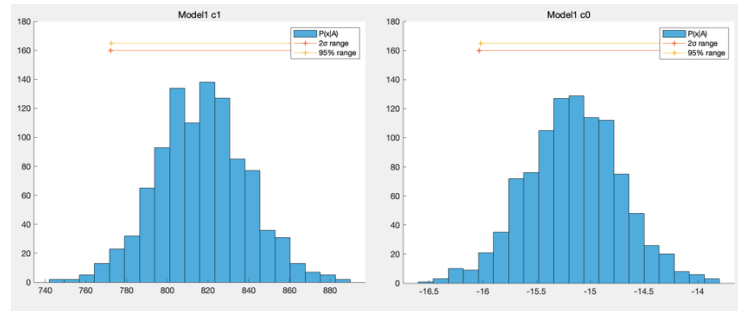

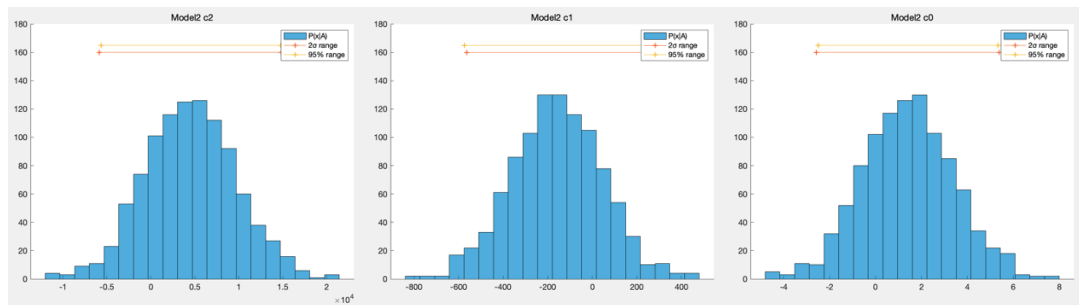*Figure 3.9: Parameters uncertainty for model 1*

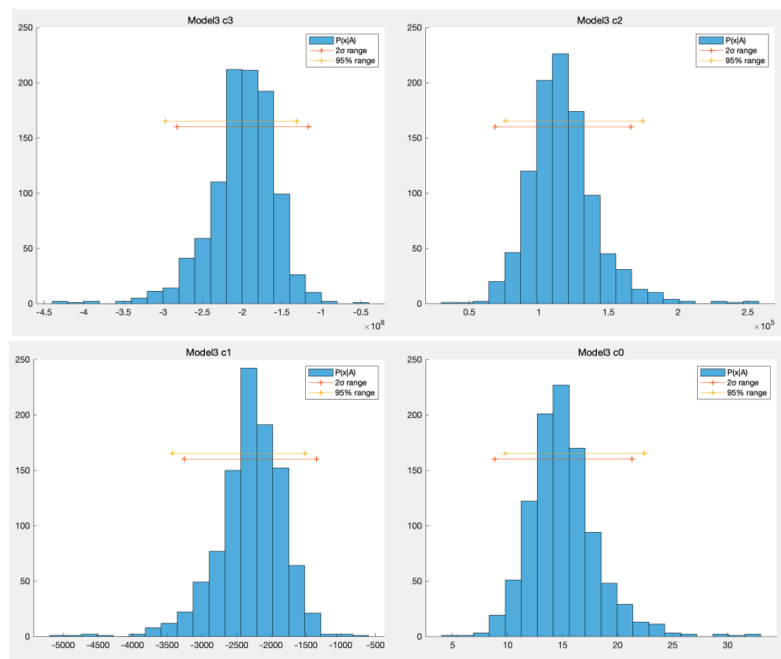*Figure 3.10: Parameters uncertainty for model 2*

Figure 3.11: Parameters uncertainty for model 3

4. Discussion

This project constructs 10 different respiratory motion models and uses 5 techniques evaluate the models. Firstly, we use visual assessment and residual motion to do the evaluation and find that 2D models perform better than 1D models. Visual assessment is the simplest method while the evaluation results are difficult to analyse. Secondly, DEF is used to do the evaluation and we find that the results are similar to using residual motion. Thirdly, we use landmark error to do the evaluation and find the 1D 3rd order model gives the best performance. Landmark error is easy to analyse while the experiment is difficult to process in practice. Then AIC and BIC are used to do the evaluation, but it is hard to determine the best model from the results. Lastly, we use uncertainty estimation to do the evaluation and find that linear model is the best model, because it has the least parameters.

Overall, the results of the different assessment techniques are generally same. However, we still need more methods to do the evaluation to obtain better results. In this project, the separate models for inhalation and exhalation are just built but are not evaluated, because it is difficult to do that in a limited time. In future work, I will complete the evaluation of separate models. Also, I will build more respiratory motion models and evaluate them.

Reference

[1] McClelland, J.R., Champion, B.A. and Hawkes, D.J., 2014, July. Combining image registration, respiratory motion modelling, and motion compensated image reconstruction. In *International Workshop on Biomedical Image Registration* (pp. 103-113). Springer, Cham.

[2] McClelland, J.R., Hawkes, D.J., Schaeffter, T. and King, A.P., 2013. Respiratory motion models: a review. *Medical image analysis*, *17*(1), pp.19-42.

[3] McClelland, J., 2013. Estimating internal respiratory motion from respiratory surrogate signals using correspondence models. In *4D Modeling and Estimation of Respiratory Motion for Radiation Therapy* (pp. 187-213). Springer, Berlin, Heidelberg.

[4] Eiben, B., Tran, E.H., Menten, M.J., Oelfke, U., Hawkes, D.J. and McClelland, J.R., 2018, June. Statistical Motion Mask and Sliding Registration. In *International Workshop on Biomedical Image Registration* (pp. 13-23). Springer, Cham.

[5] Wilms, M., Werner, R., Ehrhardt, J., Schmidt-Richberg, A., Schlemmer, H.P. and Handels, H., 2014. Multivariate regression approaches for surrogate-based diffeomorphic estimation of respiratory motion in radiation therapy. *Physics in Medicine & Biology*, *59*(5), p.1147.

[6] McClelland, J.R., Hughes, S., Modat, M., Qureshi, A., Ahmad, S., Landau, D.B., Ourselin, S. and Hawkes, D.J., 2010. Inter-fraction variations in respiratory motion models. *Physics in Medicine & Biology, 56*(1), p.251.