# Surrogate driven respiratory motion model derived from CBCT projection data

*Adeyemi Akintonde*

A dissertation submitted in partial fulfillment

of the requirements for the degree of

**Doctor of Philosophy**

of

**University College London**.

Department of Medical Physics & Biomedical Engineering

University College London

April 29, 2021

I, Adeyemi Akintonde, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

# Abstract

Cone Beam Computed Tomography (CBCT) is the most common imaging method for Image Guided Radiation Therapy (IGRT). However due to the slow rotating gantry, the image quality of CBCT can be adversely affected by respiratory motion, as it blurs the tumour and nearby organs at risk (OARs), which makes visualization of organ boundaries difficult, in particular for organs in the thoracic region. Currently one approach to tackle the problem of respiratory motion is the use of respiratory motion model to compensate for the motion during CBCT image reconstruction. The overall goal of this work is to estimate the 3D motion, including the breath-to-breath variability, on the day of treatment directly from the CBCT projection data, without requiring any external devices. The work presented here consist of two main parts: firstly, we introduce a novel data driven method based on Principal Component Analysis PCA, with the goal to extract a surrogate signal related to the internal anatomy from the CBCT projections. Secondly, using the extracted signals, we use surrogate-driven respiratory motion models to estimate the patient's 3D respiratory motion. We utilized a recently developed generalized framework that unifies image registration and correspondence model fitting into a single optimization. This enables the model to be fitted directly to unsorted/unreconstructed

data (CBCT projection data), thereby allowing an estimate of the patient's respiratory motion on the day of treatment. To evaluate our methods, we have used an anthropomorphic software phantom combined with CBCT projection simulations. We have also tested the proposed method on clinical data with promising results obtained.

# Impact statement

CBCT imaging is routinely used in most modern radiotherapy clinical institutions as part of the treatment regime. The CBCT image is used to align the patient on the day of treatment in order to facilitate accurate dose delivery and can potentially be used for adaptive radiotherapy. However, image quality of CBCT is poor due to the intrinsic nature of the acquisition.

The work presented in this thesis develops and demonstrates methods for building respiratory motion models from CBCT data and subsequent correction for the motion. A particular advantage of the proposed methods is that it only needs the CBCT projection data from standard 3D-CBCT scans, with no extra external signal nor prior motion model from 4D-CT or other data required. This makes the method relatively straightforward to translate to clinical practice and has the added benefit that additional cost/equipment is not needed. The results presented on the clinical data are promising, and are the first results demonstrated on real clinical data of a method that can account for breath-to-breath variation and does not rely on respiratory sorted 4D-CT. With further development the work presented in this thesis could provide accurate motion information and ultimately lead to more accurate and more effective radiotherapy treatments for lung tumours.

During my PhD I have presented my research at a number of international conferences, including the International Conference on the use of Computers in Radiation Therapy (ICCR), the European Society for Radiotherapy and Oncology, and the IEEE Medical Imaging Conference (MIC). I have also been awarded the UCL Institute of Healthcare Engineering (IHE) bursary to attend international conferences. My work has also attracted the interest of Elekta, one of the largest manufacturers of Radiotherapy equipment, who are interested in funding a studentship at UCL to continue working in this area.

Outside my field, I have engaged about the wider context of my work with a range of audiences including school children, and the general public at the Bloomsbury science festival.

*Dedicated*

*to*

*My Mother*

*For her endless love, support and encouragement*

# Acknowledgements

I would like to take this opportunity to thank all those who have contributed to this Ph.D. research, directly or indirectly.

First and foremost, I would like to thank my supervisors Professor Jamie McClelland and Professor Kris Thielemans whose enthusiasm has been a source of inspiration. It has been a privilege to have supervisors who always keeps their door open and are eagerly willing to discuss the work.

A special thank you to Simon Rit, Pauline Mouches, Björn Eiben and Ottavia Bertolli for their contribution to this work. In addition, I would also like to thank Helen Grimes and Syed Moinuddin at University College London Hospital for providing the clinical data used in this study.

These acknowledgements would not be complete without mentioning my research colleagues: Elena Tran, Ludovica Brusaferri, Elise Emond, Alexander Whitehead, and Danny Ramasawmy. It was a great pleasure working with them and I appreciate their ideas, help and good humour. A special thank you goes to Björn Eiben who is also there to help with difficult questions.

Finally, I must express my very profound gratitude to my family, and Lola, Anto, Fredi, Bolu and Lanre for providing me with unfailing support and continuous

encouragement throughout my years of study and through the process of researching and writing this thesis.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| CBCT | Cone Beam Computed Tomography |
| WHO | World Health Organisation |
| IMRT | Intensity Modulated Radiation Therapy |
| IMPT | Intensity Modulated Proton Therapy |
| IGRT | Image Guided Radiation Therapy |
| MRI | Magnetic Resonance Imaging |
| Linac | Linear Accelerator |
| MCIR | Motion Compensated Image Reconstruction |
| 4D-CBCT | Four-Dimensional Cone Beam Computed Tomography |
| CT | Computed Tomography |
| HU | Hounsfield Units |
| PET | Positron Emission Tomography |
| TPS | Treatment Planning Software |
| ICRU | International Commission on Radiation Units and Measurements |
| GTV | Gross Tumour Volume |
| CTV | Clinical Tumour Volume |

| | |
|---|---|
| ITV | Internal Target Volume |
| PTV | Planning Treatment Volume |
| MidP | Mid-Position |
| SI | Superior-Inferior |
| AP | Anterior-Posterior |
| ABC | Active Breathing Control |
| 4D-CT | Four-Dimensional Computed Tomography |
| kV | Kilo-Voltage |
| FBP | Filter Back Projection |
| FDK | Feldkamp-Davis-Kress |
| MI | Mutual Information |
| NMI | Normalised Mutual Information |
| CC | Cross Correlation |
| SSD | Sum of Squared Differences |
| GD | Gradient Descent |
| CG | Conjugate Gradient |
| DIR | Deformable Image Registration |
| DVF | Deformation Vector Field |
| RPM | Real-time Position Management |
| AS | Amsterdam Shroud |
| TNLM | Temporal Non-Local Mean |
| CNR | Contrast to Noise Ratio |

| | |
|---|---|
| PICCS | Prior Image Constrained Compressed Sensing |
| ROOSTER | RecOnstructiOn using Spatial and Temporal Regularization |
| MC-FDK | Motion Compensated Feldkamp-Davis-Kress |
| PC | Principal Component |
| OAR | Organ At Risk |
| MM-FD | Motion Modelling and Free-form Deformation |
| VPD | Volume Percentage Difference |
| COMS | Centre Of Mass Shift |
| SMEIR | Simultaneous motion estimation and image reconstruction |
| SART | Simultaneous Algebraic Reconstruction Technique |
| MC-SART | Motion-Compensated Simultaneous Algebraic Reconstruction Technique |
| FOV | Field of View |
| PSNR | Peak Signal to Noise Ratio |
| GPU | Graphic Processing Unit |
| DL | Deep Learning |
| FFD | Free Form Deformation |

# List of Publication

A. Akintonde, O. Bertolli, E. Chandy, G. Helen, S. Moinuddin, R.A. Sharma, J.R. McClelland, K. Thielemans. (2019). "Data-Driven region of interest respiratory surrogate signal extraction from Cone Beam CT Projections" **Under revision**

A. Akintonde, H. Grimes, H. Moinuddin, R.A. Sharma, J.R. McClelland, K. Thielemans. (2019). "Data driven region of interest respiratory surrogate signal extraction from CBCT data." Radiotherapy and oncology, 2019-04, Vol.133, p.1139-1140.

AC. Whitehead, E.C. Emond, N. Efthimiou, A. Akintonde, S.W. Wollenweber, C.W. Stearns, B.F. Hutton, J.R. McClelland, K. Thielemans. (2019) "Impact of Time-of-Flight on Respiratory Motion Modelling using Non-Attenuation-Corrected PET". 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)

A. Akintonde, R.A. Sharma, P. Mouches, C. Mory, S. Rit, K. Thielemans, J.R. McClelland. (2019) "Respiratory motion model derived from CBCT projection data". 19th International Conference on the Use of Computers in Radiation Therapy

A. Akintonde, J.R McClelland, H. Grimes, S. Moinuddin, R.A. Sharma, S. Rit, K. Thielemans, (2017). "Data Driven Cone Beam CT Motion Management for Radiotherapy Application". 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)

# Chapter 1

# Introduction

Cancer is one of the leading causes of death worldwide according to the report of the World Health Organisation (WHO) [4]. There was an estimated 17 million cases of cancer worldwide as of 2018, and 9.6 million deaths due to cancer. There are more than 360,000 new cases of cancer in the UK every year [5]. Three main methods of cancer treatment exists, surgery, chemotherapy and radiotherapy. 45% of patients diagnosed with cancer have surgery to remove the tumour as part of their primary cancer treatment, while 27% of patients diagnosed with cancer in England during 2013-2014 had radiotherapy as part of their primary cancer treatment [6]. The aim of radiotherapy is to deliver the optimum radiation dose to kill cancerous tissue while minimising the dose to the healthy tissues. Lung cancer is the $3^{rd}$ most common cancer in the UK, it accounts for 13% of all new cancer cases. Lung cancer does not usually cause noticeable symptoms until it spread through the lungs, as a result the outlook for the condition is not as good as many other types of cancer [7].

With treatment techniques such as Intensity Modulated Radiation Therapy (IMRT), and Intensity Modulated Proton Therapy (IMPT), it is now possible to

deliver a more precise dose distribution to complex tumour shapes close to critical organs. However accurate patient positioning and beam placement is critical for achieving the desired treatment goal. In order to ensure adequate patient positioning, IGRT was developed. This attempts to correct for any changes in patient position and target localization before each treatment session. Currently, IGRT can be performed using many systems and techniques, including, Magnetic Resonance Imaging (MRI), ultrasound, portal imaging, and CBCT. CBCT is one of the most common IGRT techniques used in clinical practice [8]. Most modern linear accelerators (Linac) have a CBCT mounted 90° to the linac beam modulator. With highly conformal treatment techniques such as IMRT and IMPT, an advanced imaging modality is required for a precise localization of the target and OARs. CBCT enables us to correct for changes of the target position prior to treatment and allows monitoring of complex changes of the patient and tumour.

CBCT images suffer from respiratory motion artifacts such as blur, as a result of the relatively long acquisition time (approx 1min). This limits soft tissue visualization and localization accuracy, particularly in thoracic and abdominal region. To take advantage of the higher conformity of radiation dose with these new treatment techniques, methods to deal with uncertainties due to respiratory motion are desirable. In this work, we have developed a surrogate driven respiratory motion model derived from CBCT projection data. The derived motion model can be used to correct some of the artifacts caused by respiratory motion in CBCT images. The model can be used to produce sharper (less blurry) images of the anatomy in the average position, this is particularly useful for patient position verification. In ad-

dition, the derived motion model can be used to estimate the daily motion, which can be used to assess if planned treatment is suitable and/or to inform adaptive RT, where treatment is adapted to account for changes to the patient. The work presented here consists of two main parts. Initially we developed a novel data driven respiratory motion signal extraction technique. This generates a signal from a region of interest within the thoracic cavity that exhibits respiratory motion, such as the diaphragm and tumour in the CBCT projections. The method was evaluated using anthropomorphic computerised XCAT phantom. Using the extracted signal we used surrogate driven respiratory motion models to estimate the patient's respiratory motion prior to treatment based on a generalized framework that combines image registration and respiratory motion model fitting into a single optimisation. This potentially allows for the estimation of the 3D organ motion during CBCT acquisition, using this we performed Motion Compensated Image Reconstruction (MCIR) in order to account for and, reduce motion artifacts and improve the image quality of CBCT images.

There now follows a brief overview of the contents of the rest of the chapters in this thesis:

**Chapter 2:** Provides background information and relevant literature relating to the work discussed in this thesis. It also contains a brief summary of the incidence of lung cancer, the problem of respiratory motion in radiotherapy, methods of imaging respiratory motion, image registration and modelling respiratory motion.

**Chapter 3:** Presents a data driven method for extracting a respiratory surrogate signal from CBCT projection data. The aim is to obtain a signal that is correlated to

the internal motion of the anatomy, which can be used to drive a correspondence respiratory motion model. In addition a secondary aim of the work presented here was to produce a signal that is also suitable for respiratory-correlated Four-Dimensional Cone Beam Computed Tomography (4D-CBCT) reconstruction.

**Chapter 4:** Describes the extensions and modifications made to the generalised surrogate-driven motion modelling framework so that it can be applied to CBCT projection data. It then presents experiments and results that evaluate the CBCT-based motion models using an anthropomorphic software phantom.

**Chapter 5:** Extends the work in the previous chapter by including MCIR, with the motion model fitting in an iterative approach. This means that an existing 'reference' image is not required thus the method can be applied to real clinical data. This chapter presents experimental results of applying the iterative motion model fitting and MCIR to both the same simulated data as used in chapter 4 and real clinical data-sets.

**Chapter 6:** Summarises and discusses the findings and conclusions from this research. We also discuss the future work that could be performed based on the findings from this study.

# Chapter 2

# Background

In this chapter we give a brief description of the background related to this thesis. We cover general topics such as lung cancer, means of treatment with radiotherapy, the problem with respiratory motion and the associated treatment methods, including the different modalities used over the course of a radiotherapy treatment. We also introduce concepts such as image reconstruction and image registration methods that are related to the topic. Also included are the relevant literatures related to the work conducted in chapters 3, 4 and 5.

## 2.1 Lung Cancer

There were approximately 130 new lung cancer cases every day between 2015 to 2017 in the UK [9]. Lung cancer incidence is strongly related to age, with the highest incidence rates being in older people. On average, between 2015 and 2017, each year more than 4 in 10 cases were in people aged 75 and above [10]. Lung cancer usually has a poor prognosis because most of the patients present with advanced or metastatic disease at the time of diagnosis [11]. The treatment of lung cancer pa-

tient's is dependent on the histological diagnosis, the patient's fitness, the stage of the cancer and the patient's choice. Treatment can include a combination of surgery, radiotherapy and chemotherapy.

## 2.2 Radiotherapy

The primary aim of radiotherapy is to treat cancerous tissue with ionizing radiation. An effective radiotherapy treatment aims at delivering the optimum radiation dose to the tumour while minimizing the dose delivered to normal healthy cells. The overall goal is to kill the cancer cells by damaging their DNA with radiation. Various types of radiotherapy treatment method exist, such as external beam therapy, brachytherapy and stereotactic radiosurgery. The main delivery system for external beam treatment is a Linac as shown in figure 2.1. In a Linac, the ionizing beam is produced by steering the electron beam into a target where the electrons undergo bremsstrahlung [1], which creates high energy photons.

The Linac rotates around the patient and delivers radiation to the tumour. Various types of treatment delivery with the Linac exists, the most common types being 3D conformal radiotherapy and IMRT. Both techniques can form highly conformal dose distributions, where the high-dose region closely matches the shape of the tumour, and there is a sharp fall off in dose outside the tumour to minimise dose to surrounding OARs. Therefore, accurate verification of the tumour target is vital.

A crucial factor that determines a successful treatment is the ability to spare healthy tissue while delivering the maximum dose to the tumour. However, pro-

[1] Bremsstrahlung radiation is any radiation produced due to the deceleration of a charged particle passing through matter in the proximity of the strong electric fields of the atomic nuclei.

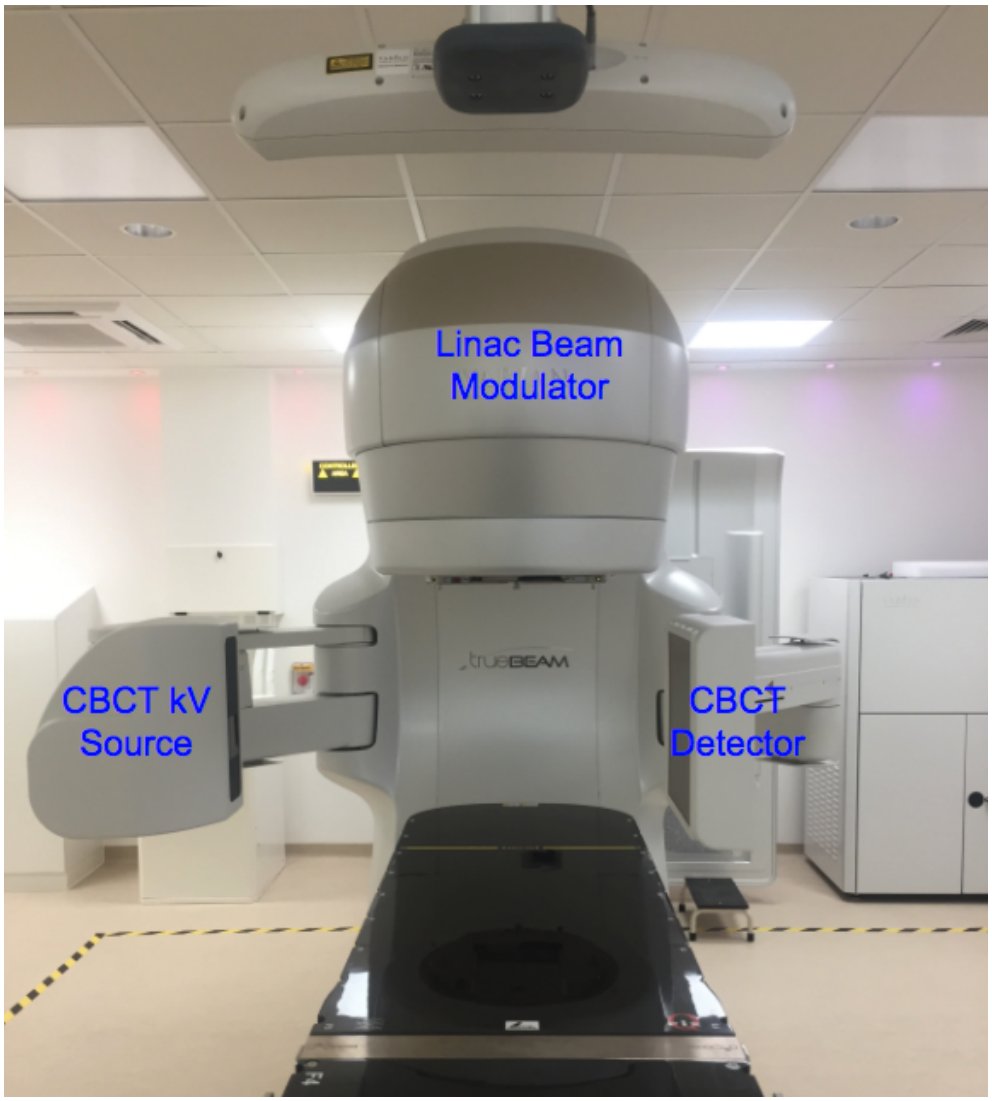


**Figure 2.1:** A modern Linac

ducing such a treatment plan is complicated. For instance, the tumour can be intertwined with healthy tissue. The clinical practice for planning and delivering radiotherapy treatment in most institutions is to acquire anatomical patient information from Computed Tomography (CT). The CT numbers in Hounsfield Units (HU) provide tissue electron density information which is used for dose calculation. The delineation of the tumour and the OARs are performed by the clinician. MRI and Positron Emission Tomography (PET) can aid with the delineation, which

can sometimes be difficult on CT images alone. The 3D images and the delineation are then used by the Treatment Planning Software (TPS) to produce an optimised treatment.

Geometrical errors between the planned treatment from the CT scan and the treatment delivered can occur. These errors can occur when defining the target volume, creating the treatment plan or differences in the patient's positioning on the day of treatment. To compensate for these geometrical errors a 'safety' margin is typically added around the tumour volume delineated on the CT scan during planning. A clinical guideline for target volumes in radiotherapy has been recommended by the International Commission on Radiation Units and Measurements (ICRU) 83 [12].

- Gross Tumour Volume (GTV): This is the tumour volume that can be seen via imaging.

- Clinical Tumour Volume (CTV): This is the GTV with an additional margin that account for microscopic disease spread not visible with imaging.

- Internal Target Volume (ITV): This consists of an internal margin added to the CTV to account for internal physiological movement and variations in shape, size, and position of the CTV.

- Planning Treatment Volume (PTV): This is a geometric concept designed to ensure that the radiation dose is actually delivered to the CTV. The aim is to ensure that $> 95\%$ of the prescribed dose is delivered to 99% of the CTV [13]. The PTV allows for the uncertainties due to inter-fraction changes to patient

and the setup errors.

## 2.3 Image Guided Radiation Therapy

IGRT uses medical imaging to improve the precision and accuracy of radiotherapy [14]. IGRT aims to detect the tumour position immediately before treatment, and allows for the adaptation of the radiotherapy plan in case the target position changes compared with the planned position. In many cases radiotherapy treatment is based on treatment plan constructed approximately 1-2 weeks before the start of the treatment. Moreover, the treatment is delivered in multiple fractions on different days. As a result, it is vital to align the patient with the treatment machine to ensure that the planned treatment is delivered. The alignment can be performed using skin markers (tattoos), however the internal anatomy can move relative to the skin due to setup errors and anatomical/physiological changes. Therefore, IGRT is used to improve the delivery of the treatment plan.

There are several commonly used IGRT techniques that can be deployed to increase the accuracy of treatment delivery, ranging from standard portal imaging to CBCT. Portal imaging uses the treatment beam and a flat panel detector for imaging. It relies on using the bony anatomical structure for the alignment as it is difficult to see soft tissue in the 2D projections. CBCT on the other hand allows for 3D volume reconstruction, therefore it is possible to setup the alignment on the soft tissue, *i.e.* tumour. Nevertheless, CBCT still has poor soft-tissue contrast hence the recent development of MR-linacs. MR-linacs can potentially provide the best image guidance for radiotherapy treatment, especially for tumours in/near soft tissue, but

require expensive specialist equipment. 3D IGRT such as CBCT and MR-linac potentially allows for adaptive radiotherapy. As treatment delivery equipment becomes more complex, enabling escalating the dose delivered, the need for precise target definition is important more than ever.

Currently, CBCT is the most common method used for IGRT. It is relatively cheap compare to MR-linac. CBCT imaging can be used to identify anatomical changes such as changes in the shape, size and to some extent the magnitude of the motion during the course of radiotherapy treatment. This can be used to assess if the plan is still suitable and/or facilitate adaptive radiotherapy. However, CBCT still has its associated disadvantages. Due to the relative long acquisition time CBCT images can be corrupted with motion blur artifacts, which can lead to inaccurate measurements of the size, shape and the extent of the motion.

## 2.4 Respiratory Motion

Respiration is a quasi-periodic process of inhalation followed by an exhalation. Oxygen is delivered to the cells from the external environment and carbon dioxide is transported in the opposite direction. The most dominant muscle involved is the diaphragm. During inhalation, the diaphragm contracts so that the lung volume expands by pulling in the inferior direction. The inter-costal muscles contract during inhalation, pulling the ribs anteriorly and superiorly, therefore the lateral and anterior posterior diameter of the thorax is increased [15]. There are two main types of breathing, diaphragmatic and costal breathing. In general the diaphragmatic breathing move the thoracic region in the Superior-Inferior (SI) direction, whereas the

costal breathing move the thoracic region in the Anterior-Posterior (AP) direction. A typical patients breathing pattern consists of a combination of both.

### 2.4.1 The problem of respiratory motion in imaging

In general, the largest range of respiratory motion is in the SI direction, where it can range between 6-34mm [16]. Large variation in respiratory motion can also be found between individuals [17]. Typically, the extent of the lung tumour motion in the lower lobe tends to be largest. However, other studies have also showed cases where large tumour motion had been observed in the middle/upper lobe [18, 19]. While respiratory motion often follows a repetitive pattern, it is typically irregular. The diaphragm and inter-costal muscles affect the motion, and their contribution can vary from breath to breath, and hence so can the motion. This can be affected by physiological factors such as exertion, anxiousness, and changes to pose, and also by natural breath-to-breath variations (which can be larger for some individuals than others). And in addition, there can be a difference between the patient breathing pattern from one day to another (*i.e.* contribution of diaphragm and inter-costal muscles) but also "base-line-shifts" caused by anatomical and physiological changes [20]. Also breathing motion can exhibit intra-cycle variation (also called hysteresis) whereby the motion follows a different path during inhalation and exhalation

Respiratory motion can cause motion induced artifacts during image reconstruction which can lead to distortion of the target volume and incorrect volumetric and positional information. Respiratory motion artifacts such as blurring not only

affect the delineation of tissue, but it can also affect the dose calculation accuracy.

### 2.4.2 Treating tumours affected by respiratory motion

Various methods have been developed to account for respiratory motion in radiotherapy. Breath-hold technique can be used for treating thoracic tumours, since it significantly reduces respiratory motion and changes in the internal anatomy. Several methods of breath-hold technique have been implemented, such as the Active Breathing Control (ABC). This is a device that helps facilitate reproducible breath-holds without requiring the patient to reach maximum inspiration capacity. Another approach is to ask the patient to hold their breath at some point in the breathing cycle. During a breath-hold, the beam is turned on, and dose is delivered to the tumour [21]. However, all breath-hold methods increase treatment time, and it can be dependent on the ability of the patient to hold their breath, or training of additional staff might be required to operate devices such as the ABC system.

There are two main treatment planning approaches used to account for respiratory motion: ITV approach and the mid-position (MidP) concept. Both of these approaches require Four-Dimensional Computed Tomography (4D-CT) images (see section 2.5.3.2 ). For the ITV based approach, the GTV is extended such that it covers the whole respiratory tumour motion in all phases of the 4D-CT [22]. For the MidP position approach an image is created by deforming all the phases of the 4D-CT to produce a single image representing the time-weighted position of all the images [23]. However, both of these approaches assume there is no breath-to-breath and day-to-day variations.

Techniques that monitor the respiratory motion of the patient in real time are also available. These includes respiratory gating and real-time tumour tracking. Respiratory gating involves the delivery of radiation within a particular part of the patient's breathing cycle, commonly referred to as the "gate". While for the real-time tumour tracking, the tumour motion is followed by dynamic movement of the treatment field. Respiratory gating increases the treatment time, but real-time tumour tracking can be more complex and need more accurate information to guide the treatment which typically involves knowing the exact position of the tumour at every point in time. Nevertheless, both techniques require accurate information on the position of the tumour during treatment, and currently this generally involves invasively placing a marker in or close to the tumour. Alternatively, new advanced techniques such as MR-linacs can also be used to facilitate gated and tracked treatments and this is an active area of research [24, 25]

## 2.5 CT & CBCT Imaging

In this section we discuss some fundamental concepts of tomographic imaging with CT and CBCT. We cover topics from the projection acquisition to the different image reconstruction techniques such as analytical and iterative reconstructions. We also discuss different methods of tackling the problem of respiratory motion in CT imaging, and we give a brief overview of motion mitigating methods in CBCT.

### 2.5.1 Projection

Image projection, also known as the Radon transform, is a linear transformation. The projection of an image $f(x,y)$ can be computed by summing all the intensity

values of the pixels in the image along that specific direction [26]. In 2D, we can write the equation of a line as given in equation (2.1), where $\rho$ is the smallest distance of the line to the origin, and $\theta$ is the angle. Image projection can be represented in the form of equation (2.2) [27]. This is the line integral or ray sum through the image $f(x,y)$ whereby the position of the line is determined by the parameters $\rho$ and $\theta$. The $\delta$ is the Dirac function. Shown in figure 2.2 is a projection of a 2D image at a specific angle $\theta$.

$$\rho = x\cos\theta + y\sin\theta \tag{2.1}$$

$$G(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(\rho - x\cos\theta - y\sin\theta)dxdy \tag{2.2}$$

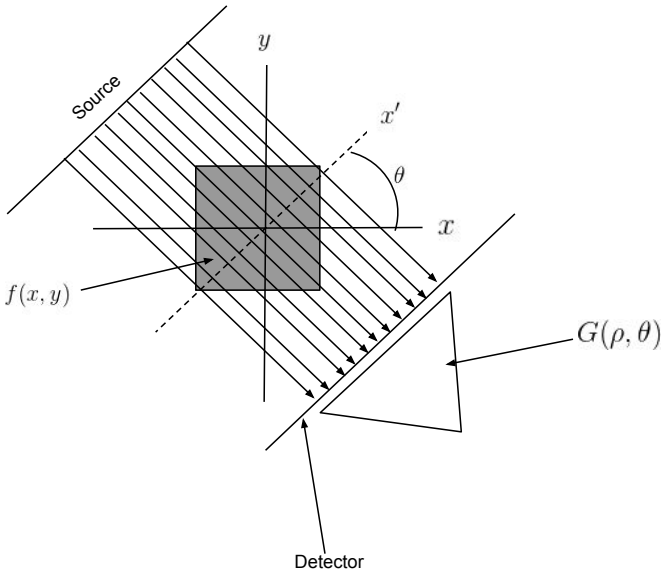


**Figure 2.2:** Schematic layout of a single projection of a 2D image at a specific angle

### 2.5.2 Image Reconstruction in Tomography

Image reconstruction is a mathematical technique used to create tomographic images from projection data that have been acquired at various angles around the patient. There are two main types of image reconstruction methods: analytical and iterative. Analytical approaches are based on Filtered Back Projection (FBP) and are widely used clinically. Iterative methods are more complex but can offer better image quality. The type of image reconstruction algorithm used has crucial implications for the image quality. In this section we will discuss the FBP and FDK (Feldkamp, Davis, and Kress) algorithms, which are the most common analytical reconstruction techniques used for tomography in medical imaging. We will also give a brief description of the iterative reconstruction approach.

#### 2.5.2.1 Analytical reconstruction

One of the most widely used analytical image reconstruction techniques is FBP. This method assumes that the x-rays travel in straight lines, the x-ray photons all have the same energy, and the x-ray intensity attenuates exponentially in the object (Beer's Law). FBP is closely related to simple Back Projection (BP). In BP the projection is redistributed evenly along the original path in the image space. The attenuation value is divided by the number of image pixels along the direction of the projection, and the mean attenuation value obtained is assigned to these pixels. This is repeated for each angle. However, performing a simple back projection will result in a blurred version of the object, since all the projections are redistributed along the original path. To reduce the effect of the blur on the image, a filter is introduced,

the projections are first filtered using a high pass filter such as a ramp filter before back-projection (FBP). This filter has the effect of removing low frequency signals from the image, therefore contrasting features are highlighted.

For 3D-CBCT reconstruction, the most common technique used is the FDK algorithm [28]. The FDK algorithm provides an approximate reconstruction and it performs well only with a small cone angle. It has some distortion in the non central axis, and some image quality artifacts. Image artifacts are more pronounced as a function of distance from the central slice and the image quality degrades as the cone angle increases. Nevertheless, the algorithm is practical and robust to use. A schematic diagram of a CBCT reconstruction from Beaudry et al. [1] is shown in figure 2.3, where the projection is defined by the projection angle and the detector coordinates $(a,b)$, whereas the beam is defined by the fan angle $\gamma$ and the cone angle $\kappa$.


**Figure 2.3:** The cone beam geometry showing the variables used in the image reconstruction (adapted from Beaudry et al. [1])

The FDK algorithm can be summarised as:

1. Pre-scale the projections by the cosine of the angle between the ray and the central ray of the projection.

$$w(a,b) = \cos\gamma\cos\kappa \tag{2.3}$$

where $\gamma$ and $\kappa$ are the fan angle and the cone angle respectively, and $a$ and $b$ are the detector coordinates.

2. Apply a row-wise (1D) ramp filter $g$ to the pre-scaled data

$$\tilde{p}(\beta,a,b) = (w(a,b)\cdot p(\beta,a,b)) * g \tag{2.4}$$

3. The filtered and pre-scaled projections are then back projected onto the volume of reconstruction

$$V_{FDK}(x,y,z) = \int^{2\pi} \frac{R^2}{U(x,y,\beta)^2}\tilde{p}(\beta,a,b)d\beta \tag{2.5}$$

where $U(x,y,\beta)^2$ is dependent on the distance between the source and the reconstructed voxel projected onto the central ray

For further detail relating to CBCT reconstruction, the interested reader is recommended to read the original paper [28] and other good summaries of this approach, for instance [1, 29, 30].

#### 2.5.2.2 Iterative Reconstruction

Recently due to higher computational speed, iterative reconstruction methods are getting more attention in medical imaging. This technique can be thought of as a method of solving a system of equations, allowing for some noise in the data. Iterative reconstruction discretises the image into pixels and treats each pixel value independently as an unknown. We can then describe the imaging process using a system of equations based on the geometry and physics. We can use an iterative algorithm to solve the system. For CBCT the relationship between the image and the projections is a linear equation. A system of linear equations can be thought of in a matrix form $AX = B$, whereby each element in $X$ is a voxel in the reconstructed 3D volume, and each element in $B$ is a pixel in a 2D projection. $A$ consist of the matrix with coefficients that relate pixels in $X$ and $B$. $A$ is a large but sparse matrix, as each pixel is only affected by the voxels that intersect a line from the source to that specific pixel.

Iterative reconstruction methods consist of 3 major steps as shown in figure 2.4 [2]. We initially start with an arbitrary estimation of the image, we forward project the current estimate, we compare the projection from the estimate with the measured raw projection data. The comparison (the difference between the estimated and measured projections) is back-projected and used to update the current image, thereby creating a new estimate. We iterate these steps and update the current estimate at each iteration. We stop once the maximum number of iterations has been achieved or when there is no improvement when we compare the estimated and measured projections.


**Figure 2.4:** The 3 main steps involved in iterative reconstruction method (adapted from Beister et al [2])

#### 2.5.2.3 Deep Learning Methods

Recently Deep Learning (DL) techniques are being applied to medical image reconstruction, this is partly inspired by the success of DL in computer vision problem and medical image analysis. DL techniques for image reconstruction are data driven, where training datasets are used to tune parametric reconstruction algorithms that approximate the solution.

Some of the method are closely related to the conventional methods of image reconstruction. DL is used as a reconstruction step by combining it with a conventional image reconstruction technique to recover the missing details in the input signal or to improve the resulting image. With this technique there are two main approaches in dealing with the problem, performing the DL method in the pre-processing domain or in the post-processing domain. In the pre-processing domain, various studies [31,32] have used a DL model to estimate missing parts of the

signal that were not acquired in order to improve the input to a conventional image reconstruction technique. The incomplete data could for instance be due to a limited Field of View (FOV), sub-sampling or a low dose acquisition. In the case of the post-processing, the main objective is to learn the mapping between the low-quality reconstructed image and its high-quality counterpart. For instance, Jin et al. [33] enhanced FBP results on a sparse-view CT through subsequent filtering by a U-Net to reduce artifacts. Similarly, Chen et al. [34] used an auto-encoder to improve FBP results on a limited angle CT projection. However, the combination of conventional reconstruction method and DL methods has potential downsides. For example, the performance of DL post-processing methods can be impacted by the results of the conventional methods. An additional limitation of DL methods is the need for a good quality input images to train the method. However, CBCT/4D-CBCT images suffer from various artifacts [35]. The DL methods described here are out-of-scope for the work presented in this thesis.

### 2.5.3 Computed Tomography for Radiotherapy

CT imaging relies on X-ray projection images. The projection images are acquired at different angles with the source and the detector rotating around the patient. A typical CT scanner uses fan beam geometry, where the beam diverges from the source. The patient is typically scanned in the SI direction producing a helical scan. With helical scan, the patient moves through the scanner. The x-ray beam traces a helical path. Helical acquisition scan takes approximately 20sec.

Organs that move with high enough speed can have their position change sig-

nificantly over the course of one CT scan. This can be the case for lung tumours, and they can be challenging to image accurately. In such cases the tumour often display distortions due to the interplay of their motion with that of a CT scanner and treatment couch. There are different CT scanning techniques that attempt to account for the patient's respiratory motion, a few are summarised below.

#### 2.5.3.1 Breath hold CT

Breath hold CT refers to the acquisition of a standard helical CT scan while the patient holds their breath. While giving better image quality in most cases, the anatomy of the patient during a breath hold scan will be in a different position when compared to normal breathing [36]. Moreover, studies have shown that it is difficult for some patients to sustain a breath-hold during the scan [37], causing artifacts in the images due to the patient struggling to hold their breath for the duration of the scan.

#### 2.5.3.2 Four Dimensional CT

Four Dimensional CT (4D-CT), otherwise known as respiratory correlated CT, is a technique that attempts to overcome some of the limitations of breath hold CT. 4D-CT attempts to image the patient anatomy at different respiratory phases during free breathing. The CT images can be acquired in either helical or cine mode [2]. In both cases each slice is imaged multiple times over a complete breath cycle, but different slices are acquired during different breath cycles. The CT data is acquired simultaneously with a respiratory surrogate signal that is correlated with the patient

[2] A cine scan acquires data continuously at the same position for a duration of time. Multiple cine scans are required to cover a larger scan area.

motion. The slices are then retrospectively sorted using the surrogate signal to form coherent volumes.

4D-CT images are subject to artifacts resulting from aspects of the 4D acquisition and processing [38]. These are caused by inter-cycle variations in the breathing motion, *i.e.* the anatomy can be in different positions during different breath cycles, even when at exactly the same phase according to the surrogate signal. This causes artifacts in the images, as the anatomy does not match between slices acquired during different breath cycles. These artifacts can lead to systematic errors in the patient contouring and dose calculations. Studies have shown discontinuities in the diaphragm and heart in 4D-CT images [39]. Artifacts in 4D-CT are very common, a study found that at least one artifact appeared in 4D-CT images for 45 out of 50 patients [40].

### 2.5.4 Cone Beam Computed Tomography (CBCT)

CBCT equipment consists of a kilo-voltage (kV) x-ray source and a flat-panel detector. In radiotherapy, both are mounted on a Linac as seen in figure 2.1. The kV source is positioned orthogonal to the Linac head. Multiple projections with the kV source are acquired as the Linac gantry rotates around the patient. The projection images are reconstructed to obtain a volumetric 3D image of the patient. Figure 2.5 shows the typical approach of acquiring multiple projections of a patient on a CBCT device. The causes of CBCT image artifacts include respiratory motion, radiation scatter and the beam hardening effect. [41].

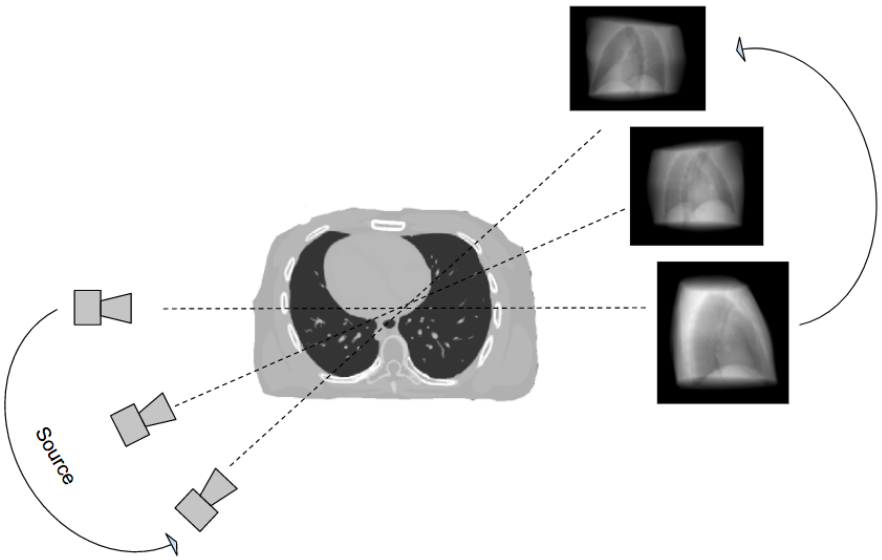


**Figure 2.5:** CBCT acquisition method

#### 2.5.4.1 Methods for handling respiratory motion in CBCT

Motion artifacts are a limiting factor in the use of reconstructed CBCT images. This is mainly due to the mathematical model of CT reconstruction, which assumes that the object is static. To reduce the artifacts induced by respiratory motion, respiratory motion correction techniques are used. In general there are three main methods used to account for respiratory motion in CBCT. Here, we give a brief summary of these methods, however we expand on these topics in chapters 3, 4, and 5 respectively.

The first approach is gated CBCT. This approach is not commercially available or widely used clinically. Here acquisition of the CBCT occurs within a given respiratory gate, for instance the end expiration gate. Therefore the reconstructed CBCT images are less susceptible to respiratory motion artifacts [42]. However, a possible limitation of this approach is the use of an external marker on the abdomen which is dependent on correlation with the internal anatomy, and in addition, gated

CBCT scan can lead to longer acquisition time of approx 5min. Moreover, this approach does not capture the extent of the motion.

The second method is respiratory correlated CBCT [43] or 4D-CBCT. This involves binning the cone beam projections into different bins which correspond to a specific respiratory phase. Each subset of the projections is then used to reconstruct a 3D-CBCT image that represents a specific phase. Therefore, a 4D image that represents the entire respiratory cycle can be obtained. However, due to the insufficient angular sampling of the projections in each respiratory phase, and the longer acquisition time (approx 4min) high-frequency streak artifacts are present in the reconstructed images [44].

The final approach is respiratory motion compensated CBCT. Here an estimate of the patient motion is made during the CBCT acquisition and the estimated motion is used to correct for motion during the image reconstruction. This method has advantages compared to the previous method due to the fact that reconstructed images can be obtained without streak artifacts [45], since all the projections are used in the reconstruction. If the motion estimate is accurate, the CBCT images would have less motion artifacts. However the primary challenge of this approach is the estimation of the 3D-motion of the patient that is required for respiratory motion compensation. Different methods of estimating the motion that have been used in the literature will be reviewed in section 2.8.

## 2.6 Image Registration

Medical image registration is the process of aligning two or more images. The goal of an image registration method is to find the optimal transformation that best aligns the structures of interest in the input images. Image registration is a crucial step for image analysis in which valuable information is conveyed in more than one image. It serves as a fundamental basis for procedures such as IGRT. Image registration enables the estimation of motion between two images, thus it can improve the correspondence of information in multi-modality imaging, allowing additional information to be obtained for normal tissue and tumour definition.

An image registration method at its simplest can be divided into three main components: (1) the transformation between two images (often called the source and the target image) (2) the cost function incorporating a measure of similarity between the source and the target image, this is composed of at least one similarity measure, that measures how similar the images are, and possibly also penalty/constraint terms to penalise undesirable transformations (this is important for Deformable Image Registration (DIR) but not usually required for rigid/affine registration). (3) an optimisation method in order to improve the cost function to reach its optimum value by changing the transformation. We will now briefly discuss these 3 components.

### 2.6.1 Geometric Transformation Models

The geometric transformation model defines the type of transformation that is permitted and how they are parameterised. The number of Degrees Of Freedom (DOF)

in a transformation corresponds to the number of parameters used to define the transformation. In general, the more the degrees of freedom, the more complex the transformation can be represented. Geometric transformation can be categorized into two main types: linear and non-linear transformations.

Rigid transformation is an example of a linear transformation. The aim of rigid registration is to find the six DOF (3 rotations and 3 translations) of transformation that maps the point in the source image into the corresponding point in the target image. This transformation model does not allow any structures within the image to change shape or size. Affine transformation is an extension of the rigid model that includes 12 DOF (3 translation, 3 rotation, 3 scaling and 3 shearing). Affine transformation allows changes to both the shape and size of all the structures in the image. However, these changes are global across the entire image. Both these transformations are generally not suitable for describing the deformation that can occur in soft tissue due to respiration, but they are a starting point for more complex non-rigid transformation.

Non-linear transformations have the capacity for modelling local deformations. It includes a wide range of transformations, spanning from transformations with few DOF to transformations which have a separate displacement for each voxel. There are 2 main types of non-linear transformation model: parametric and non-parametric models. Parametric models parameterise the transform using a mathematical model, whereas non-parametric uses a Deformation Vector Field (DVF) which directly describes the transformation at each voxel. The most commonly used parametric transformation model is a cubic B-spline, and this is the one used

in this thesis. Registrations using a B-spline transformation model were originally proposed for registration of MRI images [46].

The B-spline transformation model parameterises the transformation using the displacements of regularly spaced control points which are typically separated by several voxels. The displacement of points between the control points is computed by using B-spline interpolation. B-spline functions are compactly supported, *i.e.* a control point only influence the transformation in the local region around the control point. B-spline transformation models can accurately describe local deformations and reduce the computational complexity as a result of having lower number of DOF when compared to non-parametric transformation model such as DVF.

### 2.6.2 Cost Function

Once the transformation model has been chosen, a function that will quantify how good the registration is after the transformation has been applied is required. This is done by defining a cost function which attempts to get the best alignment between the two images. A cost function consists of a similarity measure and zero or more constraint/penalty terms. The choice of similarity measure is dependent on issues such as, if the images were acquired with different modalities or imaging parameters.

Voxel-based similarity measures compute the similarities of the images at every voxel. Numerous voxel-based similarity measures have been developed for different applications. One of the most common voxel-based similarity measures used

for lung CB(CT) images is Sum of Squared Differences (SSD).

$$SSD(A,B) = \sum_{n}^{N_v}((A(x_n) - B(x_n))^2 \tag{2.6}$$

where $A$ and $B$ are the images with $N_v$ voxels. $x_n$ is the coordinate of the $n^{th}$ voxel. SSD assumes that the images are identical once registered, and so is applicable when the only differences between the images are due to motion. It also assumes that the intensity values in both images are the same.

Another similarity measure that can be used is Normalised Cross Correlation (NCC). This algorithm assumes that there is a linear relationship between the intensities in the two images. NCC is advantageous in the case of registering different modalities, such as CBCT and CT image. NCC can be computed as follows:

$$NCC(A,B) = \frac{1}{N_v \sigma_A \sigma_B} \sum_{n}^{N_v}(A(x_n) - \bar{A})(B(x_n) - \bar{B})) \tag{2.7}$$

where $\bar{A}$ and $\bar{B}$ the mean intensities and $\sigma_A, \sigma_B$ the standard deviations for the two respective images.

One of the most widely used similarity measure in medical image registration is Normalised Mutual information (NMI). NMI can be used for multi-modal registrations as it does not assume any specific predefined relationship between the image intensities but rather measures the amount of information shared between them. NMI can be fined as:

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)} \tag{2.8}$$

where $H(A)$ and $H(B)$ represents the marginal entropy's of images A and B respectively. $H(A,B)$ is the joint entropy of the two images

It is worth noting that SSD is a dissimilarity measure, a larger value indicate that the images are less similar, whereas for NCC and NMI a larger value indicate that the images are more similar. Therefore, when optimising the registration we seek to minimise the SSD, but maximise the NCC and NMI.

Image registration is an ill-posed problem, it is important to prevent non-plausible deformations, caused by non-smooth and/or folding transformations. There are many ways of constraining the registration, such as using finer/coarser B-spline grids. Different penalty terms penalise different types of transformation, e.g. diffusion encourages smooth deformations, log of Jacobian penalises volume changes, bending energy encourages the first derivative of the transformation to be smooth [47].

### 2.6.3 Iterative Optimisation Methods

The iterative optimisation method tries to obtain the parameters of the transformation model that give the best value of the similarity measure. The choice of optimisation method is dependent on the transformation model and the similarity measure being used. The most commonly used optimisation techniques are gradient based methods, which includes Gradient Descent (GD) and Conjugate Gradient (CG). GD and CG optimisation techniques are often used with B-spline transformation models. At each iteration the gradient of the cost function with respect to the transformation parameters is computed. GD updates the transformation by adding

the gradient (scaled by a step-size) so as to improve the cost function. The CG method has faster convergence compared to GD. CG uses the gradient from previous iterations together with the current gradient to determine the direction of the update [48].

## 2.7 Modelling Respiratory Motion

### 2.7.1 General Concept

Motion modelling offers a potential solution to some of the problems introduced by respiratory motion. The goal of a respiratory motion model is to model the relationship between the motion of internal organs and a surrogate signal, such as the displacement of the skin surface.

There are four main components that form a motion model: the surrogate data, the geometric transformation (see 2.6.1), the correspondence model that relates the parameters of the transformation to the surrogate signal and finally the fitting method used to fit the correspondence model. These components determine the quality of the motion model. Hereafter we discuss the general usage of motion models and we also give a brief description of the main components.

### 2.7.2 Use of motion models

Motion models are primarily used when it is not possible to measure the motion of interest directly with adequate temporal resolution, but surrogate data that are closely related to the motion of interest can easily be acquired at high temporal resolution. Motion models are typically based on imaging data that captures the

internal motion. The surrogate signal is acquired at the same time as the imaging data and the motion of interest. The model can then be used to estimate the internal motion from the surrogate signals.

### 2.7.3 Surrogate Signals

As previously mentioned a surrogate signal that is closely related to the motion of interest is required to build a motion model. The surrogate signal can be either a 1D signal , 2D data, or even 3D images. Various types of surrogate signals exist. A spirometer is commonly used as a source of surrogate for respiratory motion models [49]. However due to instrumentation errors, there can be a considerable drift in the spirometry signal obtained. One of the most common methods of acquiring a respiratory surrogate signal is by tracking the motion of points on the surface of the chest or abdomen. This is typically performed using an optical tracking system such as the Varian Real-time Position Management (RPM) system, or a laser based tracking system can also be used. Higher dimensional data can also be used as surrogate signals for motion models, such as surface information and imaging. [50] Various studies have also proposed using higher dimensional surrogate data as a source from which to derive simpler surrogate signals such as the use of CBCT projections as surrogates [51].

Surrogate signals can be used to sort images to drive motion model in two main ways: (1) phase binning, here an image is assigned to a bin according to the phase of the breathing signal at the moment the image was generated. (2) amplitude binning the images are grouped according to the amplitude of the corresponding

breathing signal. Both approaches has associated advantage and limitation, some studies [36, 52] have demonstrated that amplitude based sorting method is more accurately related to the target position, whereas phase sorting performed well for patients with more reproducible breathing pattern.

To model both intra-cycle variation and inter-cycle variation, a common approach when using a simple 1D position tracker is to use the current value of the signal and the time-derivative of the signal as proposed by Low et al [53]. In addition, the more signals used, the more the variation that can be modelled. However, using more signals can cause overfitting, since the more the degrees of freedom that the model has, the more data is required to fit the correspondence model.

### 2.7.4 Correspondence models

There are two main types of models that approximate the relationship between the surrogate data and the motion of interest: direct correspondence and indirect correspondence. For the direct correspondence, the model estimates the motion as a direct function of the surrogate as seen in figure 2.6. This can also be written as:

$$M = \alpha(s) \tag{2.9}$$

where $s$ is the surrogate data (which may be a one or higher dimensional signals such as displacement of a point or surface), $\alpha$ the direct correspondence model and $M$ a vector of transformation parameters.

For an indirect model, surrogate data is not acquired during the acquisition. The motion is estimated from the imaging data, and a statistical model is used to

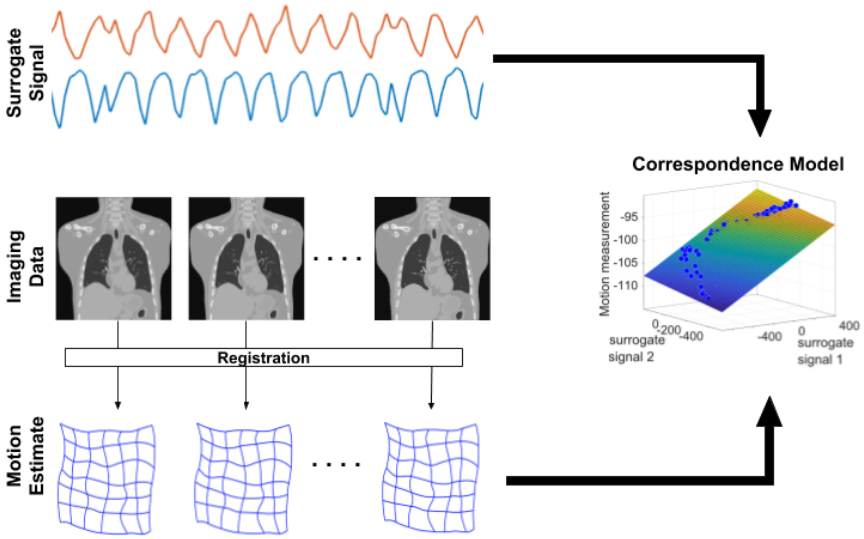


**Figure 2.6:** An illustration of a typical respiratory motion model. Respiratory surrogate signals are acquired simultaneously with imaging data, and image registration is used to estimate the motion from the imaging data. Finally the correspondence motion model approximates the relationship between the surrogate data and the motion.

build the motion estimate, e.g using PCA [54]. The weights of the Principal Component (PC) can be used to estimate the respiratory motion. During the model application (see figure 2.7) the PC weights are optimised to find the best match between the measured surrogate data and the estimate of the surrogate data generated by the motion model [3]. The direct correspondence approach was used in this thesis. Here, one of the most common models used is the linear correspondence model, mainly due to its simplicity where the transformation parameters are modelled as a linear combination of the surrogate signal. Other correspondence models used include polynomial and B-spline models.

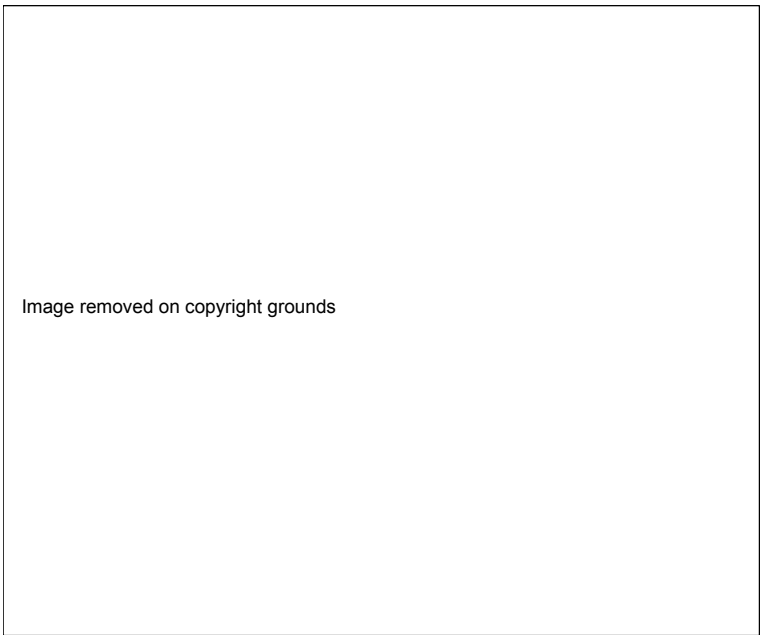


**Figure 2.7:** An illustration of the indirect correspondence model during the model application. Here the internal variables are the PC weights. [3]

## 2.8 Literature Review

In this section we will give a detailed review of the literature relating to the work presented in chapters 3, 4 and 5. We will discuss relevant literature relating to respiratory surrogate signal extraction suitable for building motion models and sorting 4D-CBCT projection data. Initially we will discuss literature relating to external methods of extracting respiratory surrogate signals information in radiotherapy, and then we will discuss data driven methods of extracting surrogate signals from CBCT projection data. We will also discuss relevant studies relating to the mitigation of the effect of motion artifacts in CBCT images. We will review methods such as phase-correlated CBCT and discuss other methods of correcting for motion artifacts namely algorithmic approaches and the use of motion models to correct for CBCT motion artifacts.

### 2.8.1 Surrogate Signals from External Device

In this section we discuss methods found in literature that have been used for respiratory surrogate signal extraction using an external means. Many methods of generating external surrogate signals for motion models and sorting 4D volumes have been proposed. The most common approach are spirometers [55], surface markers [56] (e.g. RPM), and 3D surface information [50, 57]. An example of a technique used to monitor patients breathing motion during radiotherapy is a spirometer. Spirometers measure airflow. The airflow during inhalation or exhalation is measured using an apparatus attached to the mouth of the patient. The change in air volume is then inferred from the integrated airflow signal over time [55]. This approach has been used to estimate the change in lung volume [58, 59].

Another external method commonly used for estimating breathing motion is optical tracking. Optical tracking uses a tracking device to detect the position of an external marker positioned on the body of the patient, e.g. abdominal region for lung motion. Ford et al. [56] was one of the first studies to use the RPM, by Varian Medical Systems (Palo Alto, CA) to evaluate respiratory motion during gated radiotherapy. The RPM works by monitoring the position of a marker placed on the surface of the patient's chest using an infrared source and an infrared video camera. Similarly, Mageras et al. [60], used the RPM system to assess motion during gated radiation therapy. They performed their study on 6 lung cancer patients. They validated their study using fluoroscopy to show that the external monitoring technique correlated well with the diaphragm motion.

Dong et al. [61] used fiducial markers placed on the patient's body to predict

tumour location in lung cancer, they developed a method of automatically identifying locations on the patient's surface with the optimal prediction power for tumour motion. The main assumption was based on the fact that there was a linear relationship between the tumour motion and the surface marker position.

All these techniques are based on the assumption that the signal from the external device is related to the internal motion, however this is not always the case, as various studies have shown. Ann et al. [57] evaluated the ability of external skin markers to predict tumour motion by comparing fluoroscopic imaging. In their study 34 sites in lungs and 14 sites in the diaphragm were investigated. They found a high correlation between the external marker and the tumour motion in only 44% of the patients studied.

Hoisak et al. [49] performed a study to evaluate the correlation between lung motion information and the abdominal displacement with tumour motion with X-ray fluoroscopy. The two measurements were acquired simultaneously. The correlation between the signals obtained by the two methods ranged from 0.99 - 0.39. Similarly, Gierga et al. [62], studied the correlation between internal and external markers for tumours in the abdominal region. They found the motion of the tumour was well correlated to the external markers, however they also discovered that large underlying tumour motion ranging from 2 to 9 mm can occur between internal and external marker motion. Yan et al. [63], also evaluated the correlation between external and internal markers for lung treatment. For this work they used multiple external markers. They found a wide rage of variation between the correlation of the external and internal signals.

Studies have shown that external signals can be related to the internal motion, but questions remain over how strong and reliable these relationships are. However, such devices are still commonly used due to their simplicity and relatively low cost.

### 2.8.2 Data Driven Methods

Many data driven methods have been proposed for motion extraction, with the signal extracted from the imaging data. Data driven techniques are advantageous because they deal with some of the challenges highlighted above for the external based methods. Signals derived from the data are related to the data itself, thus potentially could be more related to the internal motion of the patient. Additionally there is no equipment cost, and no extra setup time required.

Dhou et al. [64] developed a method using local intensity feature tracking. This method works by extracting a respiratory signal based on feature points that exhibit respiratory movement in the projections. Feature points extracted in the first CBCT projection were tracked in subsequent projections to determine a trajectory that represents the entire breathing motion of the patient, the method was able to detect respiratory signal in all the projections, however the average run time of the process was approx 26min. Similarly Park et al. [65] proposed an image registration based technique to sort CBCT projection data into different bins, based on the related motion information. Their method combines the intensity based feature point detection and trajectory tracking using random sample consensus. Feature points observed between adjacent projections were selected based on image intensity. Then the trajectories of the feature points were tracked using an image registra-

tion algorithm. Respiratory signals were acquired from the analysis of the obtained trajectories through periodicity analysis of the transitional patterns of the feature point positions. They tested their method with 2 phantom data and 6 patient data. Their method was compared to the Amsterdam Shroud (AS) (see below), Fourier based method and the local intensity feature tracking method. The obtained respiratory signals were used to perform phase correlated 4D-CBCT and their method produced better image quality than the other techniques based on the Peak Signal to Noise Ratio (PSNR).

Probably the most well-known method of motion extraction for CBCT projection data is the technique developed by Zijp et al. [66], termed the AS method. We give a detailed description here. This technique assumes that there is a high contrast structure that relates to breathing motion. This is usually provided by the interface between the diaphragm and thoracic cavity. During respiratory motion the diaphragm mainly has a translational motion along the SI direction. Therefore the diaphragm is ideal for extracting information related to respiratory motion. A detailed diagram of the 3 main steps involved in the AS technique is illustrated in figure 2.8.

The first step applies a derivative filter along the SI direction, this highlights structures with sharp edges such as the diaphragm as seen in figure 2.8B. The resulting image is summed along the horizontal direction to obtain a 1D projection. These steps are repeated for all the CBCT projections. The resulting 1D projections from all the CBCT projections can be depicted as a 2D image called the “AS” image. The region in the AS image that corresponds to the diaphragm is normally

the high contrast region. An example of this region is shown in figure 2.8C. Each column represents a specific projection at a particular time point. The horizontal axis of the "AS" image represents time.

Finally, to extract a respiratory signal from the AS image shown in 2.8C, each column is aligned to the next column, such that the SSD of all pixel values is minimized. The number of pixels that each column has to be shifted by for an optimal fit represents the respiration signal. Though the AS approach enables the extraction of a respiratory surrogate signal from CBCT projections, a limitation of this method is that it is highly dependent on the presence of a high contrast structure (such as the diaphragm) within the FOV of the CBCT projection data. In addition, the AS method suffers from baseline drift in the extracted signal due to error accumulation between successive time points in the Shroud image, and the motion of the contrasting structures in the 2D image due the rotation of the detector. A majority of the methods are designed to generate a phase signal to be used for binning CBCT projections, therefore a drift in the signal is not a problem.

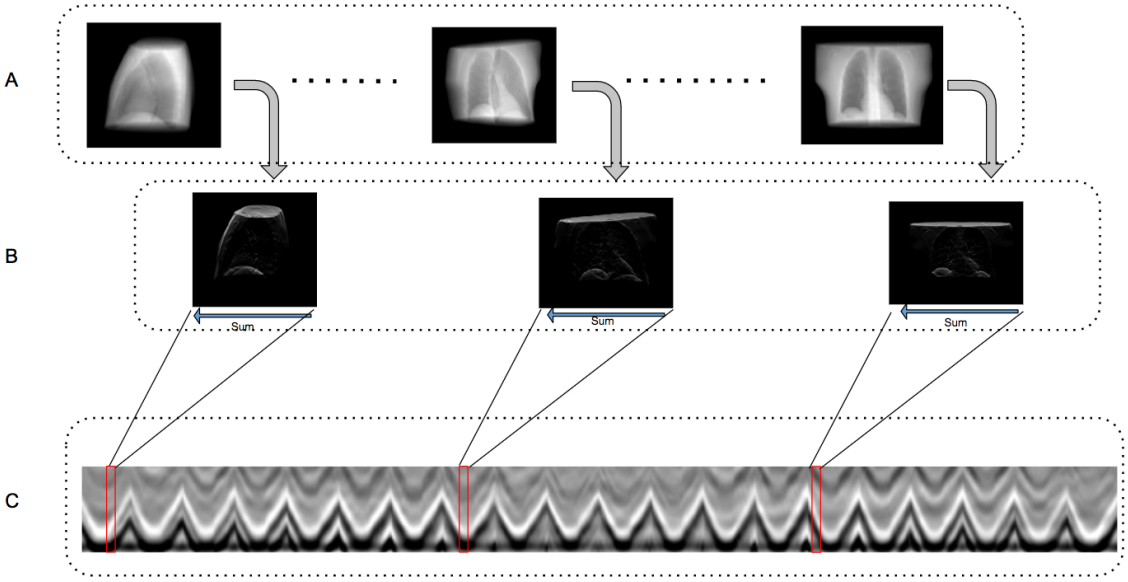


**Figure 2.8:** The AS method. **A)** The CBCT projections. **B)** The derivative of the CBCT projections along the SI direction. **C)** The Shroud image, each column represents the sum of the image in *B* summed along the horizontal axis

Chao et al. [67] proposed a method of extracting respiratory signals from CBCT based on the AS technique. An adaptive normalisation filtering was applied to the AS image to further increase the weak oscillating structure locally. From the corrected AS image, the respiratory signal was obtained using a two-step optimisation technique in order to reveal the large-scale regularity of the breathing signals. Using their proposed method, they were able to obtain less noisy breathing signals for 5 patients. They compared their method to the original AS method by evaluating the breaths per minute (bpm), and they found that the new algorithm outperformed the original AS technique for all patients by 8.5% to 30% when the waveform acquired by an air bellows belt was used as reference. Similarly Yan et al. [68] proposed a method termed Local Principal Component Analysis (LPCA) using the AS image. It works by applying PCA to a subset of the AS data-set in a

sliding window manner (see section 3.2.2). By doing so, this reduces the effect of the detector motion on the extracted signal. The proposed method showed improved performance when compared to other techniques such as the AS and the intensity analysis method, however for certain patient cases the method failed when there is insufficient contrast within the projections in particular for half-fan scan mode.

Kavanagh et al. [69] developed another method analyzing the variation in pixel values between adjacent projections by developing a pixel value summation and then applying a high-pass filter. The comparison of their method to the AS method yielded good visual comparison. They concluded that their new technique generates a suitable breathing signal for projection sorting, however further analysis on patient data and more quantitative analysis to other methods found in the literature is required.

The primary interest in this work is in relation to respiratory surrogate signal extraction from CBCT data. Nevertheless, data driven methods for non-CBCT data also exist (*e.g.* PET). Thielemans et al. [70] proposed a technique of using PCA to extract motion information from gated PET data. Their technique worked based on extracting the first PC from the data which represents the largest amount of variation, in this case the largest variation in the data is the respiratory signal. Their method was tested on clinical PET data and CINE CT images.

### 2.8.3 Phase Correlated CBCT

An initial solution to account for respiratory motion is respiratory correlated CBCT also known as 4D-CBCT. This method involves binning the cone beam projections

depending on their position in the respiratory cycle based on a respiratory signal. The subset of the cone beam projections is then used to reconstruct a 3D-CBCT image, each image representing a particular phase of the respiratory cycle, therefore we obtain a 4D-CBCT image of the entire respiratory cycle. Sonke et al. [43] first presented this solution on a CBCT scanner integrated with a Linac. In their study, they retrospectively sorted the projections from patients and phantom data respectively using the AS signal. [66]. Motion artifacts present in the 3D-CBCT were substantially reduced in the 4D-CBCT, such that the shape of moving structures could be identified more accurately. Dietrich et al. [71] also presented a similar solution. In their work they tested the method using phantom and patient data-sets, in this case the breathing phase was detected by an external gating system. The reconstructed 4D-CBCT images were compared to 4D-CT therefore allowing for re-positioning of the patient for inter-fraction set-up motion. Kriminski et al. [72] performed a similar study, however they used a mobile CBCT consisting of a C-arm to image a moving phantom. They found that images reconstructed from the phase binning of the CBCT projections showed a significant improvement when compared to the reconstructed image from the complete projection data. Li et al. [73] also performed a similar study, they used an internal planted fiducial marker as the respiratory signal, and they found that it was possible to reconstruct artifact (blur, view aliasing) free 4D-CBCT images.

Respiratory correlated CBCT is one of the main techniques developed to account for the effects of respiratory motion on a CBCT mounted on a Linac. However, this method has its associated limitations. The reduced number of projections

that contributes to the reconstruction of each respiratory phase causes view aliasing artifacts which degrades the image quality [44]. The artifacts can be reduced by performing multiple gantry rotations, or slowing down the gantry speed, such as with the case of Sonke et al. [43], whereby a 4min CBCT acquisition scan was acquired. However, the increase in imaging acquisition time leads to increased dose and potentially introduces greater positioning uncertainty, due to patient motion over the long acquisition duration. Moreover, respiratory correlated CBCT also assumes that the patient's respiratory pattern is regular. There have been various studies in the literature that demonstrate otherwise [74, 75]. In addition, studies have also shown that respiratory correlated CBCT can underestimate lung target motion during radiotherapy [76], since large variation in respiration can occur from breath to breath and from day to day.

### 2.8.4 Algebraic approach to the removal of motion artifacts in CBCT

Algebraic reconstruction techniques have been proposed to improve image quality while reducing the number of projections required to reconstruct images of acceptable quality. Several methods adopt optimisation-based approaches that use compressed sensing (including total variation (TV) regularization) to reconstruct images from sparsely angular sampled data. These methods have the potential to reduce patient dose due to their ability to use fewer projections.

Jia et al. [77] proposed two novel solution for 4D-CBCT reconstruction using a temporal non-local mean (TNLM) energy term approach. The method involves

a 4D-CBCT reconstruction algorithm and an enhancement algorithm through temporal regularization. Different phases of the 4D-CBCT were reconstructed or enhanced at the same time by minimising a data fidelity term and the TNLM term. The TNLM approach was used to account for the temporal redundancy of the 4D-CBCT images. Both approaches generated better image quality visually when compared to the FDK approach. The reconstruction approach improved contrast-to-noise ratio (CNR) by a factor of 2.56–3.13 and the enhancement method increased the CNR by 2.75–3.33 times relative to the FDK results.

Leng et al. [78] proposed an image reconstruction method based on prior image constrained compressed sensing (PICCS) [79] to achieve streak artifact-free images in 4D-CBCT. PICCS allows accurate reconstruction of a volume using undersampled numbers of projections (approx. 10–20 projections) without artifacts. The prior image used was the volume reconstructed using the entire CBCT projections without phase binning. They tested their approach on simulated data, and qualitative results showed improvement in the image quality of the PICCS approach relative to the standard FDK reconstruction method.

Mory et al. [80] proposed the 4D RecOnstructiOn using Spatial and Temporal Regularization (4D ROOSTER) method. This method was originally proposed for C-arm system during intra-cardiac interventions. There are two main parts to the algorithm, it alternates between a reconstruction and regularization in order to reconstruct the heart from a single 10sec acquisition. The reconstruction step is based on CG and the regularisation consists of four main components: enforcing positivity, averaging along time outside a motion mask that contains the heart and

vessels, 3D spatial total variation minimization, and 1D temporal total variation minimization. Mory et al. [81] later built upon the existing ROOSTER and is called motion- aware ROOSTER (MA-ROOSTER). It performs temporal regularization along curved trajectories, following the motion estimated on a prior 4D-CT scan. The MA-ROOSTER was compared to the ROOSTER, motion-compensated FDK (MC-FDK), and respiration-correlated method, using CBCT acquisitions from both physical phantom and patient data. Qualitatively the result for the MA-ROOSTER yielded streak-free reconstruction comparable to the MC-FDK results.

### 2.8.5 Motion model approach to the removal of motion artifacts in CBCT

Motion models can be used to improve the quality of CBCT images by parameterising respiratory motion in terms of a surrogate signal and the model can be used to estimate anatomical changes (see section 2.7 ). This information can then be used to improve image reconstruction. Various techniques has been proposed to model respiratory induced motion [3]. Normally a 3D data + time datasets that describes the motion of the body such as a 4D-CT is required for constructing a motion model.

Zhang et al. [82] proposed a PCA based motion model, whereby deformation fields were generated from 4D-CT volumes and signal from the diaphragm was used as surrogate. PCA was performed to parameterize the 3D deformation field in terms of the diaphragm motion. They found that the first two PCs were adequate to accurately describe the organ motion based on data from 4 patients. In their subsequent study, Zhang et al. [83] used their method to correct motion artifacts

in 1min CBCT scans. The CBCT were respiratory sorted into multiple bins and the 3D images from each phase were deformed to an end-exhale image. Similar to their previous study PCA was used to parameterize the 3D deformation field in terms of the diaphragm motion. The obtained model was used to deform each of the phase sorted CBCT volumes to a chosen reference volume, and the average of the images resulted in a single CBCT image. They evaluated their proposed method on two patient cases and simulated data. They concluded that blurring and streaking artifacts are visibly reduced with motion correction. However, the use of respiratory sorted CBCT images to build the model is affected by the strong streaking artifacts observed in the reconstruction. The artifacts will be different in each phase image thereby influencing the non-rigid registration.

Rit et al. [44] proposed a method of compensating for respiratory motion of 3D-CBCT image using a prior motion model estimated from a 4D planning CT. The respiratory motion model was estimated from 4D-CT using an optical flow method to estimate the 3D DVFs from the End Exhale (EE) image to the other images of the 4D-CT. The method for the motion compensated algorithm back-projects the projection data into the subject space to produce a tomographic image. However, unlike the FDK approach which back-projects along straight lines of x-ray acquisition, the algorithm used here back-projects along straight lines, and then warps the back-projections. Their method assumes that the respiratory motion during the CBCT acquisition is identical to that of the 4D planning CT. They found that noise and view-aliasing artifacts were lower on the motion compensated CBCT images with 1 min scan than on respiration-correlated CBCT images with 4min scan.

Riblett et al [84] also proposed a similar method by combining group-wise DIR to build a motion model and MCIR. 4D-CBCT reconstructions were registered to pre-selected respiratory phase reference images (end inhalation) to model respiratory motion. The resulting transformations were used to deform projection data during the FDK back-projection operation to create MCIR similar to Rit et al [44]. The improvement in image quality was assessed by evaluating the reduction in aliasing artifacts, image noise reduction, contrast from implanted markers and tissue inter-phase sharpness, which was defined as the slope of a sigmoid curve fit to a moving tissue boundary. Overall, they found that their method had improved image quality when compared to 4D-CBCT images, and reductions in view aliasing artifacts were obtained for clinical data.

Vandemeulebroucke et al [85] proposed a method of estimating respiratory motion from CBCT projections by including prior information from 4D-CT. The model was fitted to the CBCT projection data by optimising the model parameters such that the measured CBCT projections matched the simulated projections using NCC as the similarity measure. They found that their proposed method was sufficient to estimate the internal patient motion from CBCT projections. A limitation of this approach is that it assumes the model from 4D-CT is still valid at the time of the CBCT acquisition, thus would not account for day-to-day variations in the model.

Zhang et al. [86] proposed a technique to estimate 4D-CBCT using prior information and limited angle projections. They proposed a motion modeling and free-form deformation (MM-FD) technique. In the MM-FD technique, the DVFs

were estimated using a motion model derived from planning 4D-CT using PCA. The motion model parameters were optimised by matching the projections of the deformed volumes to the input projections, then the estimated DVFs were fine-tuned using a FD model based on data fidelity constraint and deformation energy minimization. The proposed method was able to estimate CBCT images accurately. However this study was limited to simulated idealized scatter-free imaging data. The MM-FD technique does not work well with real clinical imaging data, which contains scatter, beam-hardening effects and energy spectrum mismatches between CT and CBCT. In their follow up study, Zhang et al. [87] also investigated clinical feasibility using phantom and patient datasets, by modifying the previous MM-FD technique by introducing normalised cross correlation (CC) as a similarity metric. The accuracy of this technique was assessed by comparing the estimated anatomical structure to the Ground Truth (GT) reference 4D-CBCT. The Volume Percentage Difference (VPD) and the Center Of Mass Shift (COMS) of the tumour volume were used as an evaluation metric. They found that their new approach was able to estimate 4D-CBCT images with geometrical accuracy of the tumour within 10% VPD and 2 mm COMS.

Sauppe et al. [88] proposed a method of reducing respiratory induced motion blur in motion compensated 4D-CBCT images called phase-to-amplitude resampling (PTAR). The method uses a phase-gating approach to estimate the initial DVF and then a phase-adapted amplitude gating method is applied, here the number and distribution of bins is defined based on the mean amplitude of the corresponding phase bin. The proposed method produced better image quality when compared to

conventional phase gated images, especially for irregular breathing patterns for both simulated and patient data-set.

Numerous studies have shown success with building adequate respiratory motion models derived from prior CT. Nonetheless, models derived from 4D-CT may not provide appropriate information for CBCT reconstruction due to changes in breathing patterns and anatomy between acquiring treatment simulation images and treatment delivery as previously discovered by [89–91]. Thus it is preferable to derive a motion model based on CBCT projections acquired with the patient in treatment position. However, these approaches typically estimate the patient's motion from the 4D-CBCT data, which are often corrupted with view aliasing artifacts, and this can affect the accuracy of the motion estimated.

Wang et al. [92] developed a method of improving the image quality of 4D-CBCT images by proposing a simultaneous image reconstruction and motion estimation approach termed simultaneous motion estimation and image reconstruction (SMEIR). The method works by alternating between a MCIR based on Simultaneous Algebraic Reconstruction Technique (SART) [93] and a motion model estimation to obtain an optimal DVF. The CBCT projections were binned into multiple breathing phases, using these images and the measured projections. SMEIR extracted DVFs between the breathing phases by solving an optical flow-based optimization problem [94]. The performance of the proposed method was evaluated on a 4D phantom and patient data. The quality of reconstructed 4D images and the accuracy of tumour motion were assessed by comparison with conventional 4D-CBCT images. Patient breathing motion irregularity was found to be one of the

potential limitations of this method since this causes the number of projections in each bin to vary greatly, thus leading to streaking artifacts.

Recently Guo et al. [95] proposed a motion compensated CBCT method based on SART [93], termed the Motion-Compensated Simultaneous Algebraic Reconstruction Technique (MC-SART), which is capable of reconstructing high quality images and motion models from cone beam projections and respiratory surrogate measurements. This method derived DVFs from respiratory sorted CBCT reconstructed images and a motion model was estimated using DIR between the reconstructed bins. MC-SART uses an iteratively updated motion model to incorporate motion compensation into the forward and back projection operators of the image reconstruction. Qualitative image quality improvement was obtained when compared to other state-of-the-art algorithms. However, this method was only tested on one CBCT acquisition.

### 2.8.6 Summary

Multiple methods of determining respiratory related signals exist, each with their own advantages and disadvantages. External based methods do not require any surgery but the motion acquired externally may not have a strong relationship with the motion of the internal structure of the anatomy. Also some of these techniques require training the users, reduce time efficiency due to the set-up time required, and require an extra apparatus at extra financial cost. These types of techniques could also introduce additional setup errors into the radiotherapy process. Data driven methods avoid the problems of the external methods but have intrinsic disadvan-

tages as well. For instance, most methods are highly dependent on the presence of a high contrast structure (such as the diaphragm) within the FOV of the CBCT projection data. Potentially this type of method would not be sufficient for acquiring a respiratory signal under certain clinical acquisitions whereby the diaphragm is not present within the FOV. In addition methods such as the AS method suffer from signal drift, which is not a problem for phase sorting, however it is a potential problem for motion modelling. In chapter 3 we propose a new data driven respiratory motion extraction technique from CBCT projection data to address the limitations of the previous techniques. In particular, our technique can be applied to different regions of the projection data, thus it does not require a high contrast structure (such as the diaphragm) to be present in the FOV of the projections.

Numerous methods have been proposed in the literature to account for respiratory motion in CBCT reconstructed images. However, these techniques have their associated limitations. For instance, respiratory correlated approaches assume reproducible motion during every respiratory cycle, so therefore cannot be used to model breath-to-breath variations, and often contain artifacts caused by these variations. In addition, these methods require long acquisition times, and as the images are formed from data acquired during different breath cycles they may not give a good representation of the true motion and its variability. In addition, the insufficient angular sampling of the cone-beam projection during respiratory phase binning causes view-aliasing artifacts such as high frequency streaks. Methods have been proposed for improving the quality of respiratory correlated images, such as reduction in the gantry speed during CBCT projection acquisitions. However, the

increase in imaging time could increase patient discomfort and introduces greater uncertainty due to patient motion over the longer scan duration.

Even though motion models derived from 4D-CT have been used widely in the literature, the clinical utility of such methods may be limited by the accuracy of the prior model [74]. The anatomy between the time of 4D-CT acquisition and the day of treatment might have changed, thus making the motion model inaccurate, and in addition, the robustness of the method could be limited due to inaccuracies of the prior model as a result of artifacts in the 4D planning CT or mis-registration [96] [44]. Furthermore, changes in the breathing pattern between planning CT acquisition and treatment day is a another cause for concern.

In addition, motion models derived from 4D-CBCT also have their associated limitations such as view aliasing artifacts in the reconstructed images which can affect the accuracy of the motion estimates. Despite the success of modern motion compensation algorithms, some limitations potentially remain regarding the accuracy of the obtained DVFs. Many of the proposed algorithms assume regular breathing patterns and are unable to account for inter-cycle variations in the data used to calculate the motion model.

In chapters 4 and 5 we build upon a recently proposed framework that unifies image registration and motion model fitting into a single optimisation [97]. This method enables the model to be fitted directly to the CBCT projections, and does not require the data to be sorted and reconstructed into full 3D volumes. Since the model is fitted to all of the projections simultaneously, the full 3D motion can be estimated. MCIR can also be performed by using an iterative scheme. Finally, no

prior model estimate is required, thus the obtained model is only as a result of the CBCT data on the day of treatment.

Chapter 3

# Data-driven ROI respiratory signal extraction from CBCT Projections

## 3.1 Aim of this chapter

This chapter presents a method to generate a data driven respiratory motion signal from CBCT projection data that could be potentially used as a surrogate for respiratory motion model fitting. Our method is based on reducing the dimensionality of the data using PCA.

PCA can be applied to datasets which contain a large amount of variability, such as the case of clinical CBCT projection data. However, there are still challenges related to using PCA as a means of extracting motion information from CBCT projection data. Within a given CBCT projection, there are two types of motion, (1) Detector induced motion, and (2) Respiratory induced motion. The signals generated are affected by both types of motion. The respiratory motion is the main component that's of interest to us, therefore when using PCA, different techniques had to be introduced to reduce the effect of detector motion on the extracted signal.

In this work we propose a new data driven respiratory motion extraction technique from CBCT projection data. Our primary goal is to extract a signal that can be used as input for a respiratory motion model. In contrast to normal 4D-CBCT which uses the phase of the respiratory signal, motion models require a signal that correlates well with the internal motion of the patient. A secondary aim of this work is to produce a signal that remains suitable (for either motion models or 4D-CBCT) when the diaphragm is not in the FOV of the projection. By combining several ideas from the literature, our method aims to overcome the problems of the current techniques.

We first use the Region of Interest (ROI) enhancement method previously introduced by Martin et al. [98] to enhance the contrast and to reduce sensitivity to structures moving across the FOV of the projections. We then use PCA in a sliding window manner as used in Yan et al. [68], on the enhanced data. Our technique can be applied to different regions of the projection data, allowing signals to be generated from specific structures such as the diaphragm. In this chapter, we present the method, test it extensively on simulated data and provide preliminary results on patient datasets.

## 3.2 Method

Our method can be divided into 4 major steps: (1) use an ROI enhancement method to select a region from the projection data [98]; (2) extract respiratory related information using PCA with a sliding window approach, inspired by Thielemans et al. [70] and Yan et al . [68]; (3) combine signals from the sliding window after correcting for the arbitrary sign flip resulting of using PCA independently on subsets of the data; (4) determine the overall sign of the extracted signal. We now give more details about the steps involved in our proposed method.

### 3.2.1 ROI Enhancement

Measured CBCT data gives projections through the whole body. The enhanced ROI method [98] aims to obtain approximate projections of the ROI only, by reducing the influence of other body parts being visible in the enhanced projections. In addition, by cropping the enhanced ROI image to the mask of the ROI projections we minimise the effect of detector rotation which causes the ROI to move in the pro-

jections, if it is not positioned at the centre of rotation. Figure 3.2 shows a graphical representation of the ROI enhancement method. The projection data (A in figure 3.2) is reconstructed into a 3D volume (B) using a standard approach, e.g. FDK. The ROI is defined by the user and a non-ROI volume (C) and ROI volume (D) are created by setting all voxels inside/outside the ROI to the same intensity as air. Projections are simulated through these volumes using the same geometry as the original projections to generate non-ROI projections (E) and ROI projections (F). The non-ROI projections are subtracted from the original projections to produce the difference projections (G). The ROI appears enhanced in the difference projections, but they can also contain artifacts due to motion outside the ROI and the limited FOV of the reconstructed volume. To minimise the impact of these artifacts the difference projections are masked by the ROI projections to produce the enhanced ROI projections (H). To further reduce the influence of the detector motion when the ROI is not positioned at the centre of rotation, we cropped the enhanced projections (H) to the mask of the ROI (F). Shown in the right of figure 3.2 is the cropped region of the enhanced ROI for the tumour.

An additional step is added to the previously developed ROI enhancement method [98] when applied to truncated data, *i.e.* where part of the anatomy is imaged by projections that cover $< 180$ degrees. This part of the anatomy is missing from the reconstructed image, and therefore also from the simulated non-ROI projections. As a consequence this anatomy remains in the enhanced ROI projections. To help mitigate this problem, a global linear correction (*i.e.* scaling and offset) is applied to the non-ROI projections by performing a least squares fit between

the non-ROI projections and the original projections (excluding pixels in the ROI). While this global linear correction does not correct for truncation, it is a simple and fast method that reduces the effect of truncation on the enhanced ROI projections. The effect of applying the global linear correction is shown in figure 3.1

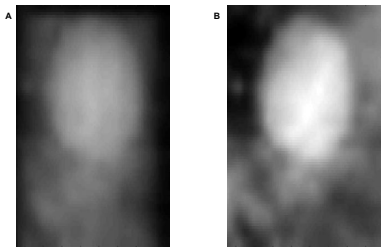


**Figure 3.1:** The effect of applying a global linear correction. (A) The enhanced projection when no correction was applied. (B) The enhanced projection after applying a global linear correction

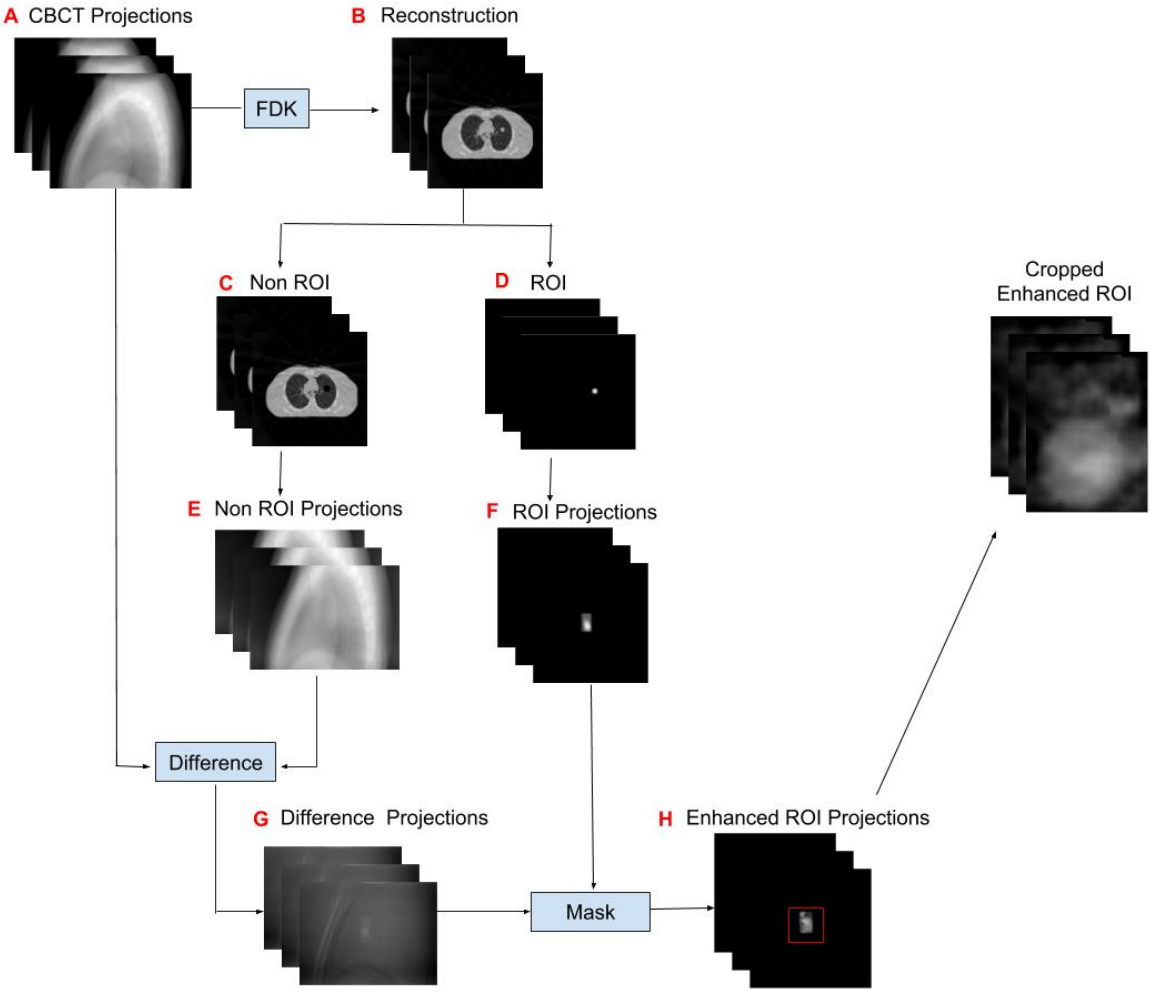


**Figure 3.2:** Flowchart of the region of interest enhancement technique

### 3.2.2 PCA using a Sliding Window

PCA seeks to reduce the dimensionality of given data where there is a large amount of interrelated variables [99]. Similar to Thielemans et al. [70] PCA was used to describe dynamic data $d_i$ as a linear expansion of $K$ orthogonal PCs. Here it is applied on the enhanced ROI projections. This can be expressed as in equation 3.1, where $d_i$ is the data at time point $i$, $\bar{d}$ is the mean of the data over all time points used for PCA, $P_k$ the PCs with the same dimension as $d_i$, and $w_{ik}$ the weight factors. Since $P_k$ are orthogonal basis vectors, the weight factors $w_{ik}$ can be computed as in equation 3.2 :

$$d_i \approx \bar{d} + \sum_{k=1}^{K} w_{ik} P_k \tag{3.1}$$

$$w_{ik} = P_k.(d_i - \bar{d}) \tag{3.2}$$

PCA has been used to extract a respiratory signal from PET data [70] using all projection data over the whole acquisition. Note that in PET the detectors are stationary. In contrast, an intrinsic property of CBCT projection data is the movement of structures across the FOV due to the detector motion. Although the ROI enhancement method reduces the influence of detector motion, it can still have an impact when moving structures move in front of or behind the ROI. To further reduce the influence of detector motion, we introduce a sliding window technique, as in the LPCA method [68]. By selecting a subset of the data-set prior to computing PCA, the approach reduces any variation in the selected data that is due to detector mo-

tion. Therefore if respiratory motion is present in the selected data $s_i$, it is expected that the motion information will be contained in the weight factor that corresponds to the $1^{st}$ PC.

In order to apply PCA, the enhanced ROI projection from each time point is rearranged into a 1D vector, and these are concatenated into a 2D matrix as shown in figure 3.3(1). Each column here represents a specific time point. We compute PCA on a subset of the data by choosing $n$ time points (corresponding to $n$ number of projections). We then repeatedly shift the window by one time point and compute PCA on each subset, as illustrated by the different colour windows shown in figure 3.3(1). For each window, we assume that the 1st PC and its corresponding weight factor represent the respiratory motion information. An illustration of the 1st PC and its associated weight factors are shown in figure 3.3(2) for the first four windows.

The sign of the respiratory signal extracted by PCA cannot be unequivocally obtained since the result of the product $P_k w_{ik}$ will not change if at the same time, $w_{ik} \rightarrow -w_{ik}$ and $P_k \rightarrow -P_k$. Since we are computing PCA independently on each window, each PC and weight will have an arbitrary sign, as illustrated in figure 3.3(3).

Since the sign of the signals extracted by PCA using the sliding window method are uncorrelated, we devised a method of correcting for this by using the Euclidean distance between adjacent PCs. For instance in equation 3.3, in which $P_1^x$ represents the 1st PC for the $x^{th}$ window, then $P_1^{x+1}$ is the first PC for the next window. We can determine which PCs require a sign change by enforcing adjacent PCs to be similar:

$$\mathrm{P}_1^x = \begin{cases} P_1^x & \text{if} \quad \left\|P_1^x + P_1^{x+1}\right\| > \left\|P_1^x - P_1^{x+1}\right\| \\ -P_1^x & \text{otherwise} \end{cases} \tag{3.3}$$

Performing this results in the weights shown in figure 3.3(4). To obtain a single signal that represents the respiratory motion, we calculate the average at each time point as represented by the black respiratory curve in 3.3(5) .

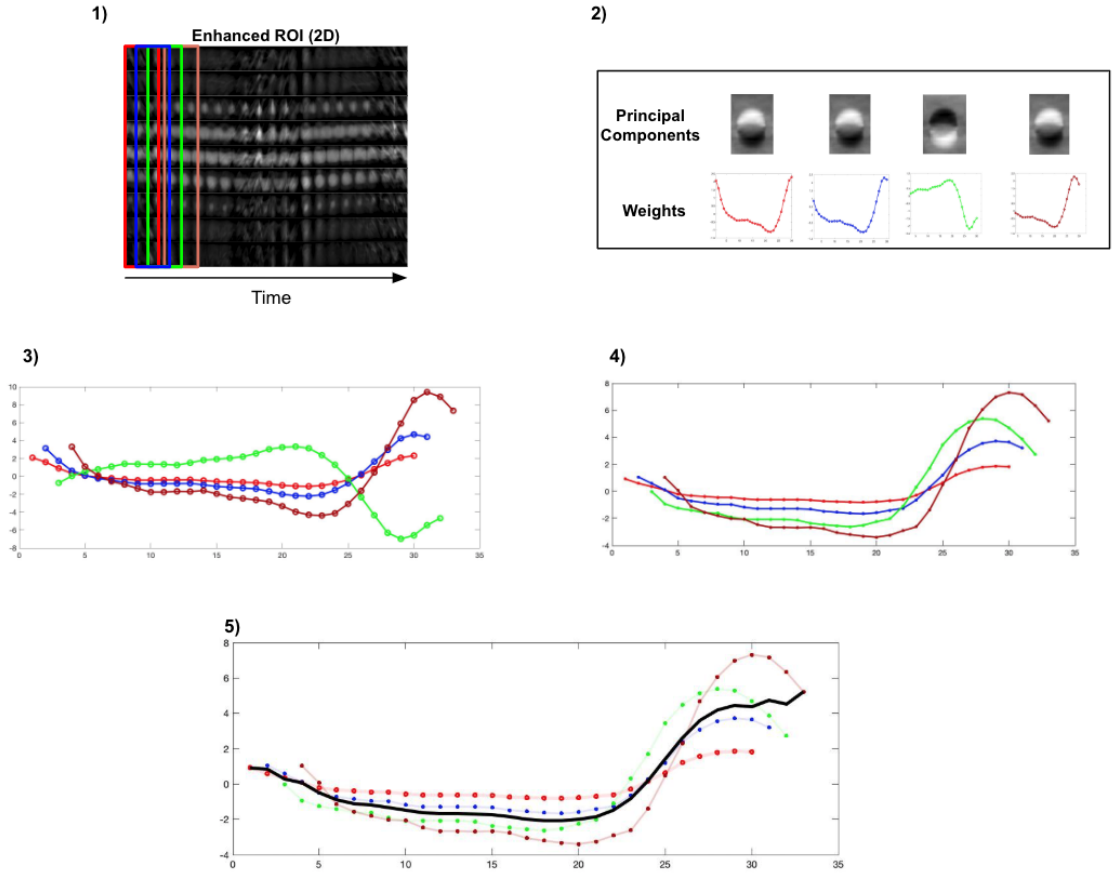


**Figure 3.3:** Illustration of the proposed sliding window method based on the first 4 windows. 1) The enhanced ROI, each colour represents a particular window; 2) the corresponding first PC coefficients and their associated weights for each colour/window; 3) the weight factors; 4) the weight factors after applying sign flip correction; 5) the black curve shows the average weight at each time point.

### 3.2.3 Overall Respiratory Signal Sign Determination

As discussed above, the sign of the respiratory signal obtained via PCA is ambiguous. As a consequence the black curve shown in figure 3.3(5), which represents the respiratory motion is also ambiguous. A new method was required to determine the true sign of the respiratory trace. We implemented a method based on a previous work in PET [100]. The method assumes that respiratory motion predominantly occurs along the SI direction [101]. Bertolli et al. [100] derived an alternative signal with a known sign from the dynamic data $d_i$ that is utilised for PCA. A signal was computed as shown in equation 3.4.

$$r_i = g_z.(d_i - \bar{d}) \tag{3.4}$$

where $g_z$ is the gradient along the SI direction of the mean of the data $\bar{d}$. The description of the temporal change in the projection from its mean resembles the result expected from PCA (see figure 3.4), as shown in equation 3.2, however with a fixed sign. Bertolli et al. therefore flipped the sign of the PCA trace to be positively correlated with $r$. For this study we used the signal obtained from equation 3.4 to flip the sign of the signal obtained from our proposed method.

### 3.2.4 Evaluation experiments

#### 3.2.4.1 Simulations

To evaluate the performance of the methods introduced in this work, we tested our approach with simulated phantom data using the XCAT computerized phantom [102], a software package that is commonly used to create anatomical models

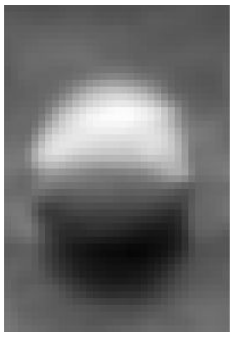

**Figure 3.4:** $g_z$, the gradient along the SI direction of the mean of the data $\bar{d}$. This resembles the result expected for the PC as shown in figure 3.3

for medical imaging research. XCAT provides an approximate representation of the human anatomy. XCAT uses non-uniform rational B-spline surfaces to model realistic and detailed human anatomical structures. It is also capable of simulating realistic 4D respiratory motion. We used XCAT to simulate attenuation images of the thoracic cavity of the human body where attenuation was computed for photons of 140 keV as appropriate for CBCT. The output of the XCAT software is a 4D data-set where the 4th dimension represents time, *i.e* a series of 3D volumes, that describes the respiratory motion. The XCAT software takes two input parameters to define the motion: the translation of the diaphragm and the AP translation of the thoracic region.

In this study, the XCAT phantom motion simulations were based on respiratory signals extracted from cine MRI data from a lung cancer patient. The respiratory signals were acquired from a local region of the diaphragm and the skin surface sections respectively. The volumes simulated from the XCAT had $512 \times 512 \times 160$ voxels with a spatial resolution of $0.878 \times 0.878 \times 2.0 mm^3$, the same as reconstructions from a clinical CBCT scanner. The number of volumes and their temporal resolution varied based on the CBCT system and the type of scan being simulated.

The acquisition rate for the 3D-CBCT scan on the Varian OBI was 11.38fps, while for 4D-CBCT acquisition it was 15fps [103]. The acquisition rate for the 3D-CBCT scan on the Varian true-beam was 14.9fps, while for 4D-CBCT it was 7fps.

CBCT projections were then simulated from the 4D XCAT images using RTK [104], an open source toolkit for CBCT reconstruction. The CBCT geometry and other simulation parameters were based on real CBCT scans and the patient data available for this study. We simulated CBCT projections based on the true-beam and on-board Imaging (OBI) systems (both from Varian Medical Systems, Palo Alto, CA) respectively. The detectors for both systems contain $1024 \times 768$ pixels of size $0.388 \times 0.388 mm^2$. Different types of CBCT acquisitions were simulated in order to replicate different clinical scenarios, as detailed in table 3.1. At our clinical institution, acquisitions using the OBI system are typically centred on the tumour (at the treatment iso-centre), and so the projection data is often truncated and some of the anatomy is outside the FOV of the reconstruction. However, acquisitions using the true-beam system are typically centred close to the centre of the patient, and therefore the reconstructions include all of the patient's anatomy. Therefore, we simulated both situations of patient-scanner alignment seen in the real scans at our institution.

Simulations 1-3 were based on the OBI system, whereas simulations 4-6 were based on the true-beam configuration. Simulations 1-2 and 4-5 were all based on 3D-CBCT scans with a duration of 60 seconds, whereas simulations 3 and 6 were based on 4D-CBCT scans with durations of 180 sec and 114 sec for simulation 3 (OBI) and simulation 6 (true-beam) respectively. In addition, simulations 3 and 6

had tumours in the upper-lung, so the diaphragm was not present in the projection data, whereas the other simulations had mid-lung tumours, so the diaphragm was present. Simulations 2 and 5 used a 30 mm diameter tumour, whereas the others used a 15 mm diameter tumour.

**Table 3.1:** The acquisition parameters for the CBCT were: Source to Iso-centre Distance = 1000mm, Source to Detector Distance = 1536mm, and a total rotational angle of 360 degrees.

| | No.of Projections | Acquisition Duration (s) | Acquisition Type (s) | Tumour Diameter (mm) | Tumour Position |
|---|---|---|---|---|---|
| **Varian OBI** | | | | | |
| Simulation 1 | 683 | 60 | 3D-CBCT | 15 | Middle ** |
| Simulation 2 | 683 | 60 | 3D-CBCT | 30 | Middle ** |
| Simulation 3$^{+}$ | 2700 | 180 | 4D-CBCT | 15 | Upper * |
| **Varian true-Beam** | | | | | |
| Simulation 4 | 894 | 60 | 3D-CBCT | 15 | Middle * |
| Simulation 5 | 894 | 60 | 3D-CBCT | 30 | Middle * |
| Simulation 6$^{+}$ | 1008 | 144 | 4D-CBCT | 15 | Upper ** |

** Right lung.

* Left lung.

$^{+}$ No diaphragm.

### 3.2.5 Clinical Data

We tested our method on three clinical CBCT projection data-sets from real lung cancer patients. The first data-set was acquired using the OBI system, and consisted of 683 projections covering 360°acquired at 11.38 fps. The second and third data-sets were acquired from the same patient on two different days using the true-beam system, and consisted of 894 projections covering 360°acquired at 14.9 fps. As for the simulated data, the OBI acquisition was centred on the tumour, resulting in truncated projections and part of the anatomy was missing from the FOV of the

reconstruction. The true-beam acquisition was centred on the centre of the patient (and also employed a half-fan scan mode), so included all of the patient's anatomy in the reconstruction, however this also limits the anatomy that appears in every projection.

### 3.2.6 Signal Extraction

For simulations 1-2 and 4-5 two signals were generated by placing the ROI on the diaphragm (D-ROI) and the tumour (T-ROI) respectively. For simulations 3 and 6 only the T-ROI signals were generated since there was no diaphragm in the projections. For the real data-sets, signals were generated by defining the ROI on the diaphragm. To evaluate the quality of results generated from the proposed method, we compared the respiratory signals extracted from the simulated data to the tumour SI motion signal from the 4D XCAT. The tumour SI motion signal was obtained using the centre of mass (COM) of the XCAT tumour volumes.

As previously mentioned the method proposed in this study relies on a user-defined window size for the sliding window. This was determined based on 4 preliminary simulated CBCT acquisitions with different real patient's respiratory traces as the input motion signal. For each simulation we tested our method using different window sizes and compared the result obtained from the T-ROI to the tumour GT motion signal, and the result from the D-ROI was compared to the diaphragm GT motion signal. An average of the correlation coefficient for each window size across the 4 different simulations is shown in figure 3.5. The optimum window size was 30 projections, corresponding to 2.65 seconds. This window size was therefore

used for all further results in this paper.

To compare our proposed method with the commonly used AS method [66], signals were generated using the implementation in RTK. In RTK's implementation an upper bound and lower bound region of the AS image need to be specified. For the simulations where the diaphragm was present, we specified the bound region to correspond to the positions of the diaphragm. For the other simulations, we selected bounds that corresponded to the tumour region. As previously mentioned, there can be a drift in the AS signal caused by error accumulation. This affected some of the AS signals generated in this study. If the AS signal is being used to determine the respiratory phase, e.g. for binning 4D-CBCT data, then the signal drift is not a problem, but if the AS signal is being compared to a known input signal as here, this drift unfairly impacts the results. Therefore, we attempted to correct for the drift by fitting a $2^{nd}$ order polynomial to the signal and subtracted it from the AS signal. We call this signal the "Modified AS" (M-AS). An illustration of the difference between the AS and the M-AS is shown below in figure 3.6.

Hereafter all signals shown in this work have been normalised such that the signals have a mean of 0 and a standard deviation of 1.

### 3.2.7 Analysis

For the simulations, the tumour SI motion extracted from the XCAT volumes were correlated with the T-ROI, the D-ROI and the M-AS signals respectively. The comparison was performed using the CC between the signals.

Respiratory correlated 4D-CBCT reconstruction was performed based on the

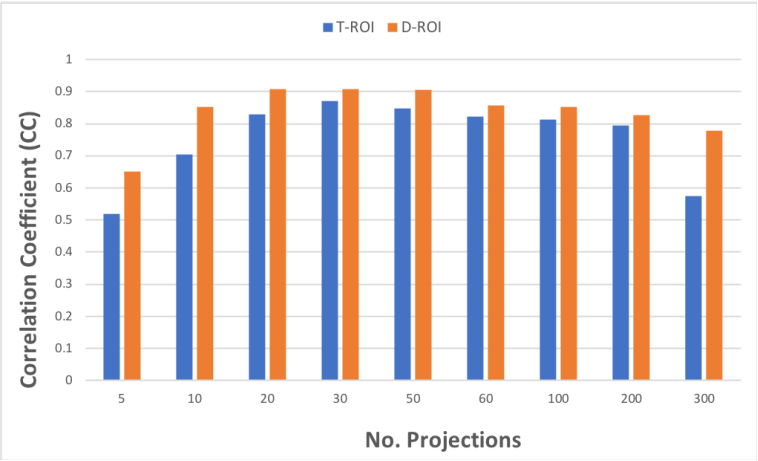


**Figure 3.5:** Average of the correlation coefficient when compared to the input signal motion for different window sizes from 4 different preliminary simulated CBCT data. The results shown here are when either the T-ROI and D-ROI were used for signal extraction.

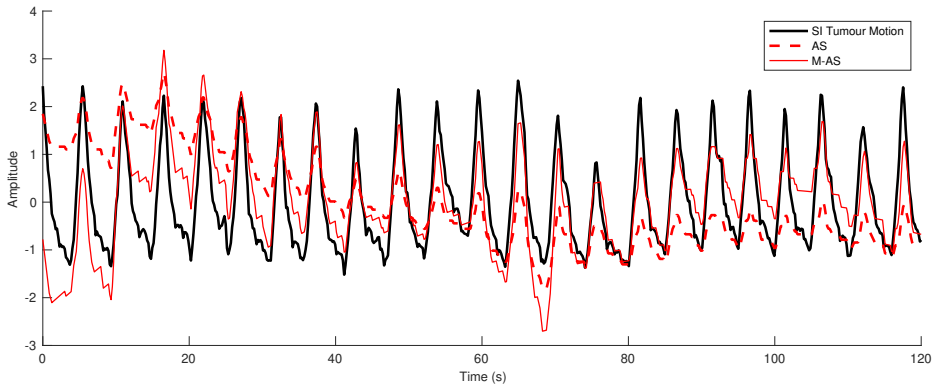


**Figure 3.6:** Illustration of the difference between the tumour SI Motion, the AS, and the M-AS signals

acquisitions from simulations 3 and 6. The projections were binned into 10 bins, according to the respiratory phase values determined from the breathing signals. The phase of the signals was calculated using a saw-tooth waveform as described in Chinneck et al. [105]. We then reconstructed each phase image individually using openRTK. The obtained 4D-CBCT reconstructions from the different signals were assessed qualitatively and by comparing the reconstructed images and a line profile through the tumour. In addition, the accuracy of the phase bin assignment for the T-

ROI and M-AS signals was evaluated using equation 3.5, which computes the Root Mean Squared Difference (RMSD) in the bin assigned by the evaluated signal and bin assigned by the tumour SI motion, taking phase wrap-around into account.

$$\text{bin}_{\text{RMSD}} = \sqrt{\frac{1}{N}\sum_{t} min[(g_t^{SI} - g_t^f)^2, (g_t^{SI} - (g_t^f - N))^2, (g_t^{SI} - (g_t^f + N))^2]} \tag{3.5}$$

where $N$ is the numbers of bins, $g_t^{SI}$ is the bin value that corresponds to the SI tumour motion signal at time point $t$ and $g_t^f$ is the bin value that corresponds to the T-ROI or the AS signal respectively at time point $t$

For the clinical data, a clinician manually identified the diaphragm location in the projections for the true-beam data. The peak of the diaphragm in the ipsilateral lung was identified. For time points where the peak diaphragm was unidentifiable from the projections a point was selected that was closest to the peak of the previous projection. The manually extracted signal was compared to the D-ROI and the M-AS signals respectively. However, for the OBI patient data, it was difficult to identify the peak of the diaphragm in the ipsilateral lung due to the poor image quality of the projections and radiation scatter. Therefore, the D-ROI and the M-AS signals from the patient's projections were compared in order to indicate if the signal from the D-ROI is plausible when compared to the M-AS. It was not possible to use the T-ROI from the clinical data-sets since there was no clear tumour in the reconstructed images and in addition for the true-beam acquisition, the tumour was not in every projection image.

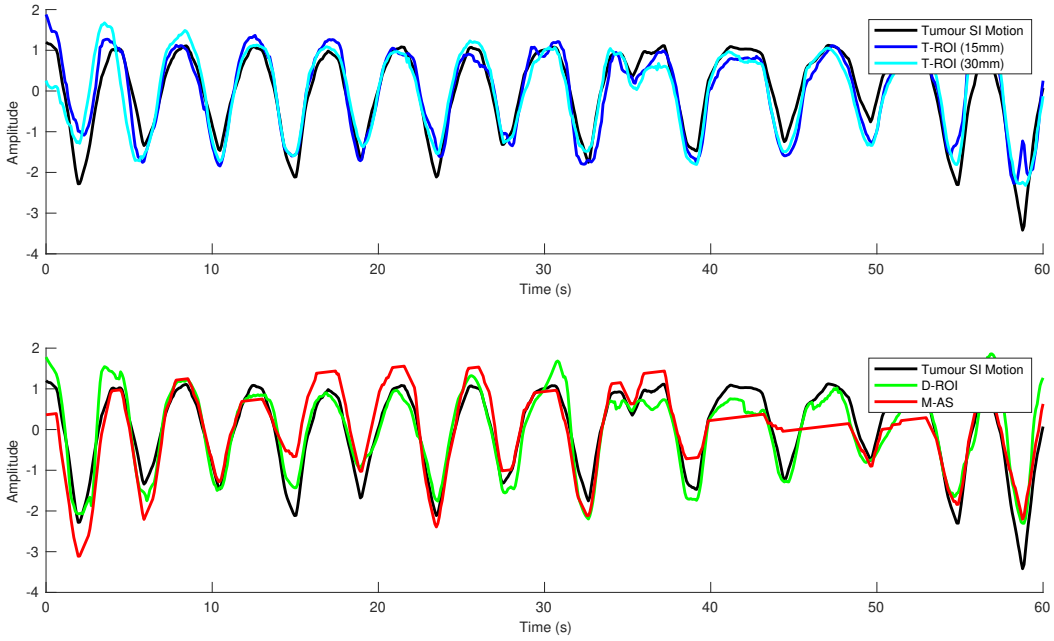


**Figure 3.7:** Respiratory traces obtained from the Varian OBI imaging system acquisitions (simulations 1-2). *Top* shows the comparison between the tumour SI motion signal and the two T-ROIs (15mm and 30mm) respectively. *Bottom* shows the comparison between the D-ROI signal and the M-AS signal relative to the tumour SI motion signal

## 3.3 Results

Figure 3.7 shows the results obtained for simulations 1-2, with the 3D-CBCT acquisitions on the OBI imaging system. The CC between the T-ROIs (15mm and 30mm) and the tumour SI motion were 0.911 and 0.890 respectively, while the comparison of the D-ROI signal to the tumour SI motion yielded a CC of 0.810. The CC between AS and tumour SI and CC between M-AS and tumour SI were 0.573 and 0.737 respectively. Simulations 1 and 2 are exactly the same with the exception of the tumour sizes, so therefore the results for the D-ROI and AS/M-AS signals were the same for both simulations. This was also the case for simulations 4-5.

Results obtained for simulations 4-5 with the 3D-CBCT true-beam imaging

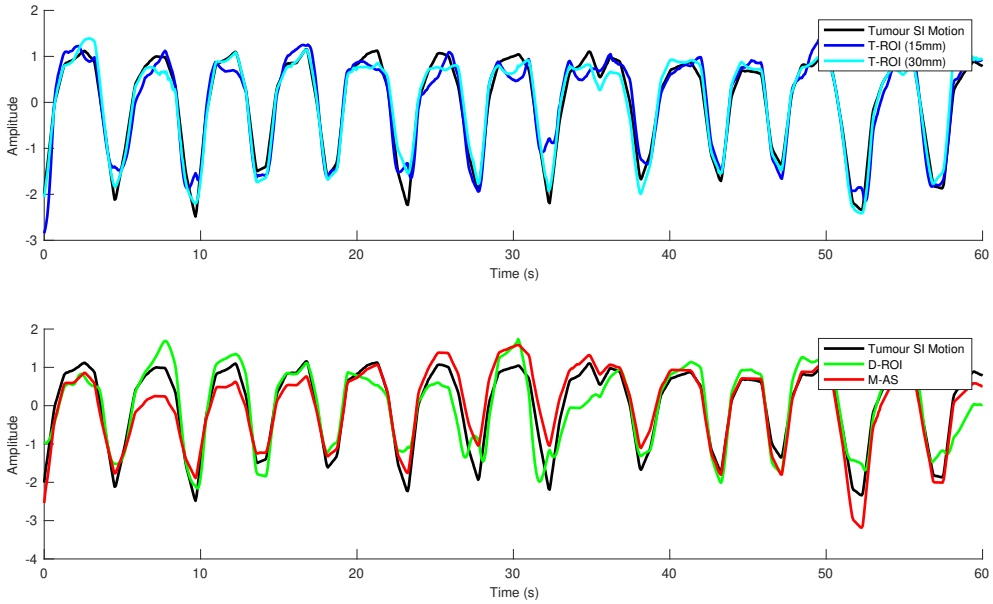


**Figure 3.8:** Respiratory traces obtained from the Varian true-beam imaging system acquisitions (simulations 4-5). *Top* shows the comparison between the SI tumour motion signal and the two T-ROIs (15mm and 30mm) respectively. *Bottom* shows the comparison between the tumour SI motion, the D-ROI signal and the M-AS signal.

system acquisitions are shown in figure 3.8. The CC comparison between the tumour SI motion and the T-ROIs (15mm and 30mm) were 0.957 and 0.977 respectively, while for the diaphragm, the AS, M-AS and the D-ROI had a CC of 0.890, 0.893 and 0.898 respectively when compared to the tumour SI motion.

Results for the 4D-CBCT acquisitions from both the true-beam imaging system and the OBI imaging system are shown in figure 3.9. As previously mentioned, for these acquisitions the diaphragm was not present in the FOV of the CBCT projections. CC of 0.045 and 0.065 were obtained for the AS and M-AS respectively, while for the T-ROI a CC of 0.880 was obtained. Similarly for the true-beam acquisition (simulation 6), a CC of 0.009 and 0.142 were obtained for the AS and M-AS respectively when compared to the tumour SI motion signal, while for the T-ROI

a CC of 0.879 was obtained. All the correlation results for all the simulations are illustrated in table 3.2.

Figure 3.10 shows the sagittal slices from the end-exhale (EE) and end-inhale (EI) images from the 4D-CBCT reconstruction, obtained using the SI tumour motion signal, the T-ROI signal and the M-AS signal for simulation 3. Figure 3.11 shows similar results for simulation 6. Line profiles through the tumour are also shown in figures 3.10 and 3.11. The $bin_{RMSD}$ for simulation 3 was 0.195 and 0.487 for the T-ROI and M-AS respectively, while for simulation 6 the $bin_{RMSD}$ was 0.141 and 0.397 for the T-ROI and the M-AS respectively.

Respiratory signals obtained from the clinical CBCT projections from 3 different acquisitions are shown in figure 3.12. For patient 1 a CC of 0.710 was obtained between the D-ROI and the M-AS. For patient 2 (acquisition 1) a CC of 0.45, 0.841 and 0.860 were obtained for the AS, M-AS and the D-ROI when compared to the manually extracted signal respectively. While for patient 2 (Acquisition 2) a CC of 0.62, 0.840 and 0.740 were obtained for the AS, M-AS and the D-ROI when compared to the manually extracted signal respectively. Results from the clinical data-sets are also shown in table 3.2

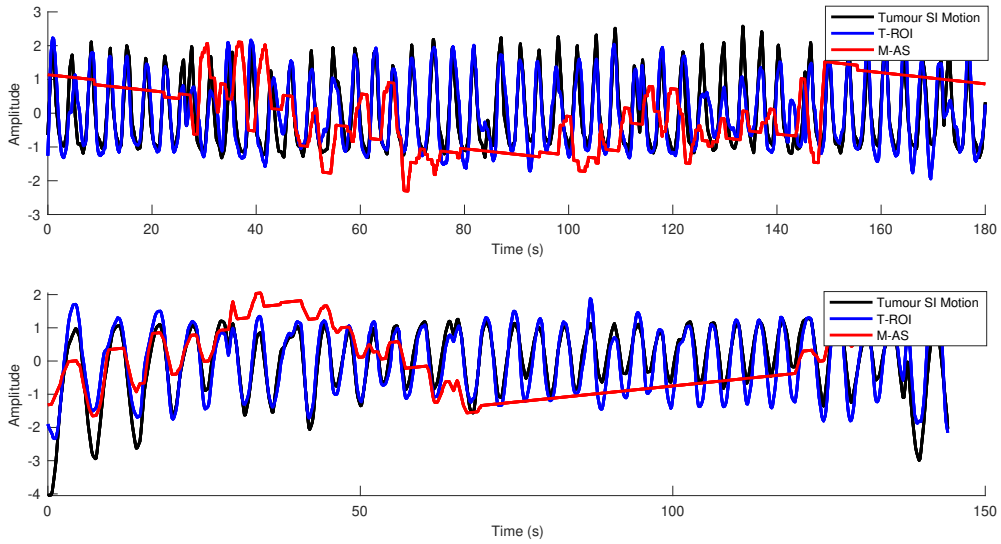


**Figure 3.9:** Respiratory traces obtained for the simulations when the diaphragm was absent in the FOV of the projection data for the 4D-CBCT acquisitions. Shown is the comparison of the tumour SI motion, the T-ROI signal and the M-AS signal. *Top* the resulting traces obtained from simulation 3. *Bottom*, shows the respiratory traces from simulation 6.

**Table 3.2:** Correlation coefficients obtained from the simulated and the clinical data-sets. For the simulations, the signals from the T-ROI, D-ROI, AS, and the M-AS were compared to the tumour SI motion signal, while for the clinical data the AS, M-AS and the D-ROI were compared to the manually extracted diaphragm peak position signal.

| | | | | T-ROI | |
|---|---|---|---|---|---|
| | AS | M-AS | D-ROI | 15mm | 30mm |
| **Varian OBI** * | | | | | |
| 3D-CBCT | 0.573 | 0.737 | 0.810 | 0.911 | 0.890 |
| 4D-CBCT | 0.045 | 0.065 | | 0.880 | |
| **Varian true-Beam*** | | | | | |
| 3D-CBCT | 0.890 | 0.898 | 0.898 | 0.957 | 0.977 |
| 4D-CBCT | 0.009 | 0.142 | | 0.879 | |
| **Patient 2**** | | | | | |
| Acquisition 1 | 0.45 | 0.841 | 0.860 | | |
| Acquisition 2 | 0.62 | 0.840 | 0.740 | | |

* Simulations.

** Clinical Data.

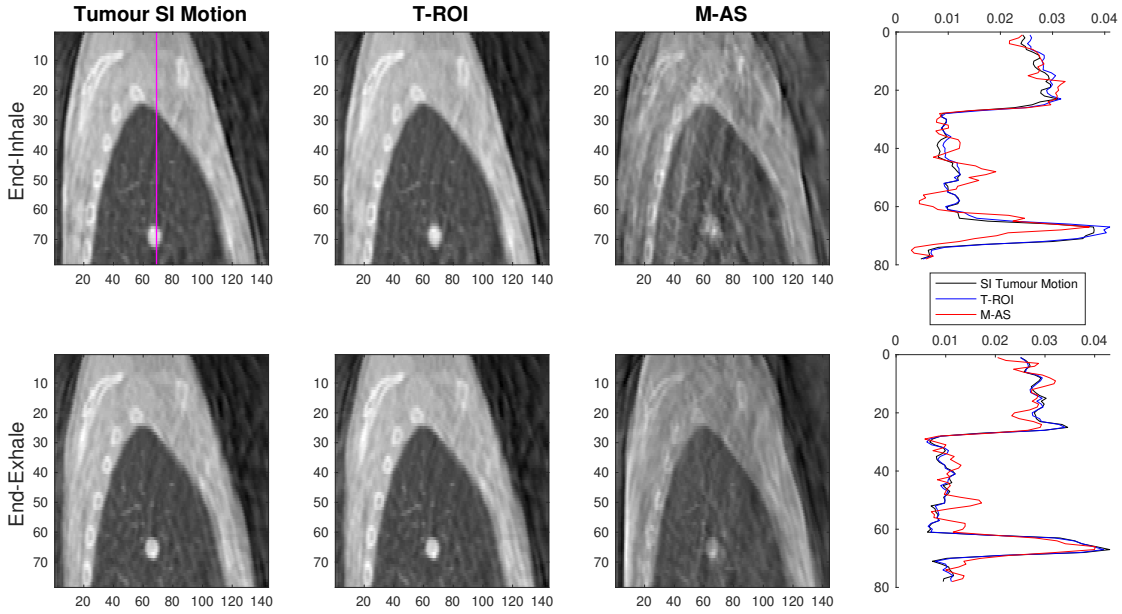


**Figure 3.10:** Sagittal planes of the images reconstructed from respiratory sorted phase bins using the tumour SI motion, T-ROI, and the M-AS signals respectively from simulation 3. Shown in the top row are the sagittal EI images, while the bottom row shows the EE images. The last column shows the line profile through the tumour for each corresponding signal used to bin the projections.

## 3.4 Discussion

Figure 3.5 indicates that our proposed method is not sensitive to the actual size used for the sliding window, as long as it is not too short compared to the breathing cycle nor too long compared to the detector motion. This suggests that a fixed value can be used for all or most patients as used here, although it would be possible to update the window size after a first pass.

The respiratory signals extracted for all the simulations presented in this study using our proposed method showed better correlation with the tumour SI motion when compared to AS and M-AS as illustrated by the results in table 3.2. We only compared our results to the tumour SI motion because the two input parameters (diaphragm and AP translation) that define the motion in the XCAT phantom were

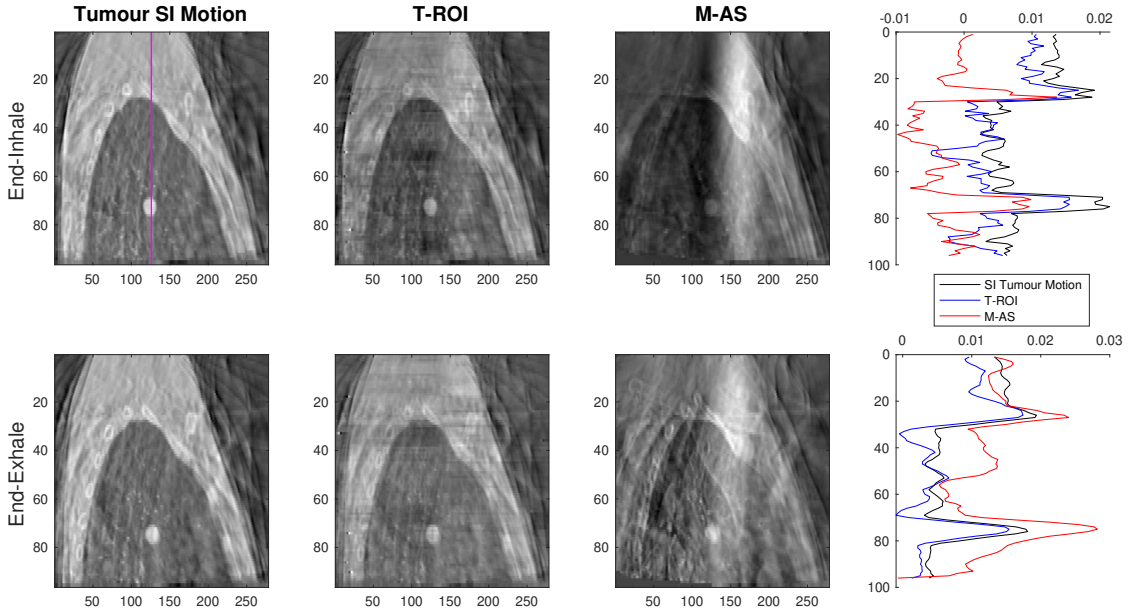


**Figure 3.11:** Sagittal planes of the images reconstructed from respiratory sorted phase bins using the tumour SI motion, T-ROI, and the M-AS signals respectively from simulation 6. Shown in the top row are the sagittal EI images, while the bottom row shows the EE images. The last column shows the line profile through the tumour for each corresponding signal used to bin the projections.

completely correlated. Therefore, the same result would be obtained for the tumour SI motion and the tumour AP motion from the XCAT. Overall the highest CC of 0.977 was obtained for the Varian true-beam acquisition (simulation 4) with the 30mm tumour, while for the M-AS the highest CC was 0.893 for the same simulation when compared to the tumour SI motion. The lowest CC for our method was 0.810 for the OBI acquisition (simulation 1), while for the M-AS method the lowest correlation coefficient was 0.065 for the 4D-CBCT OBI acquisition (simulation 3). Visual inspection of the traces confirmed the correlation results. Respiratory signals obtained from the T-ROIs comparison to the tumour SI motion signals as shown at the *top* of figures 3.7 and 3.8 indicated that our method is invariant to tumour sizes when applied to a T-ROI.

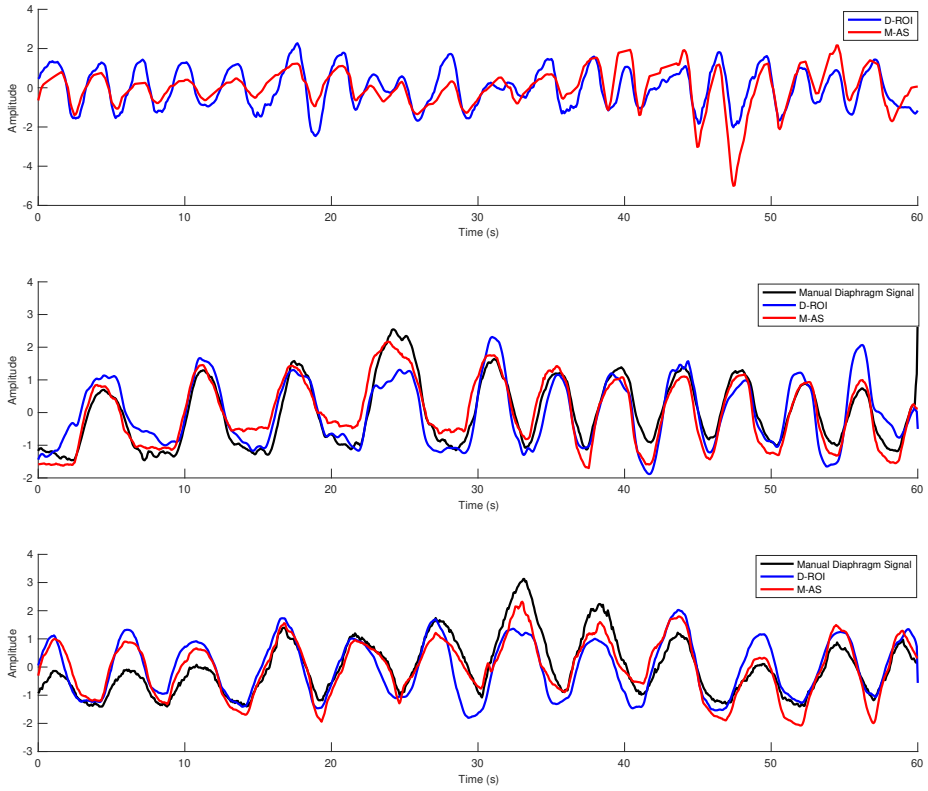


**Figure 3.12:** Respiratory signals obtained from the clinical CBCT projections from 3 different acquisitions. *Top* shows the D-ROI and the M-AS from patient 1 acquired on a Varian OBI imaging system. *Middle and bottom*, shows the manually extracted diaphragm signal relative to the D-ROI and the M-AS from patient 2, The CBCT projections were acquired on different dates on a Varian true-beam imaging system.

As previously observed in the literature [106], the AS method produced poor results when relying on poorer contrast features such as the tumour, as other parts of the anatomy can obstruct their view in some projections. Our results suggest that our method is robust to changes in FOV, making the method suitable independent of what region of the thorax is being treated. The proposed method does not require the diaphragm in the FOV. However, it does require a structure that is moving with the patient's respiration.

It is worth noting that the extracted signal is not expected to be stable over multiple fractions since we can have changes in the anatomy and variability in the patient's breathing from one fraction to another. In addition if the ROI is placed in a different location this could potentially impact the result obtained.

The 4D-CBCT reconstructed images obtained from using the SI tumour signal and the T-ROI signal are of similar visual quality, and show a similar amount of motion for the tumour. However, the reconstruction images obtained from using the M-AS signal appear blurrier due to the incorrect binning of some projections, and the tumour motion appears reduced, as expected from the observed CCs. The profile comparisons shown in figures 3.10 and 3.11 also illustrate the difference in image quality between the two sets of reconstructions for both imaging systems respectively. Similar trends were observed for both the true-beam and the OBI simulations, however the 4D-CBCT reconstructions from the true-beam showed poorer image quality and profiles when compared to the OBI reconstruction. A potential reason could be that more projections were acquired for the OBI 4D-CBCT simulations (2700 projections for OBI, 1008 projections for the true-beam 4D-CBCT simulations), thus angular under-sampling can be seen in the reconstructed images. Overall the obtained signal obtained with the method proposed can be used for 4D-CBCT phase sorting as demonstrated in figures 3.10 and 3.11, and in addition, it should give a reliable amplitude that can be used for either amplitude sorting of 4D-CBCT or amplitude-based motion models.

Results obtained from patient 1 are limited as there was no manually extracted signal present for comparison and the acquired data was not suitable for 4D- CBCT.

Nevertheless, we compared the signals extracted from the D-ROI and the M-AS, which yielded a CC of 0.71, this illustrates a degree of consistency between the two different methods. For both acquisitions from patient 2, reasonable results were obtained for the M-AS method and our proposed method when compared to the manually extracted signal. In acquisition 1, the proposed method performed marginally better than the M-AS, with the D-ROI obtaining a CC of 0.860 while the M-AS obtained a CC 0.841. For acquisition 2 the M-AS performed better with a CC of 0.840 while the D-ROI obtained a CC of 0.740.

When performing the ROI enhancement method, it is difficult to completely omit the influence of structures moving across the ROI projection due to the intrinsic nature of the CBCT projection data. There are two issues that cause other structures to be present in the enhanced projections. Firstly, truncated projections lead to parts of the anatomy to be missing from the reconstruction, and hence they are missing from the forward projections and consequently present in the enhanced ROI projections. Truncation is an issue for the CBCT acquired on the OBI since the centre of rotation of the CBCT projection is positioned at the centre of the tumour. For the true-beam acquisitions, the centre of rotation is positioned relatively close to the centre of the body, thus the full anatomy is in the reconstructed image. The current method described in section 3.2.1 is simple but there were still some structures remaining. Secondly, the ROI enhancement method assumes that all of the anatomy outside the ROI does not move, which is not the case. When structures outside the ROI move they will appear blurred in the non-ROI image. The non-ROI projections are generated from the static non-ROI image, so the structures do not move in the

non-ROI projections. This causes a mismatch between the non-ROI projections and original projections, and leads to artifacts in the difference projections. The effect of the artifacts is mitigated by masking the difference projections to the ROI, but when moving structures pass directly in-front-of or behind the ROI they can still impact the extracted signal. While this indicates that the enhancement method could be further improved, our current evaluation suggests that it is largely sufficient for producing input data for PCA. It would be possible to improve the ROI enhancement technique. In particular, the truncation problem could potentially be addressed by performing an iterative reconstruction instead of the FDK reconstruction technique currently used. However iterative reconstructions increase computation time, hence affecting the clinical applicability of the method.

The proposed method fails when the sliding window technique was excluded in the method. Both the proposed methods in this paper and the LPCA method [68] utilise PCA in a sliding window manner. However there are significant differences between the two approaches. Firstly, the pre-processing step for the LPCA method is primarily dependent on the AS technique and a method of subtracting the background from the AS image. In contrast, the method proposed in this paper is based on using an enhanced version of the actual projection data directly. Our proposed method also provides the user with more flexibility in specifying what region of the body to focus on to extract a respiratory signal. A further difference between the methods is that the LPCA method relies on computing the first 5 PCs, and then chooses the "best" PC based on the PC of the previous window. In our case, we only compute the 1st PC which represents the largest variation in the pre-processed

data. Comparing our method to the LPCA technique is beyond the scope of this study and remains for a future study.

A potential limitation of the proposed method in the study arises when selecting the ROI. The ROI selection is limited to the area of the reconstructed image whereby the ROI would be present in all the forward projections. For the OBI simulations, the centre of rotation is positioned at the tumour location, thus this region would be present in all projections. In our current approach, the ROIs have been selected manually, however future work would include using the PTV (with expansion in the SI direction, to cover the range of the tumour movement). However for the true-beam acquisitions, where there is a larger detector offset and the centre of rotation of the projections is positioned close to the centre of the patient, this limits the selected ROI to be close to the centre of the patient. This was the case for the two acquisitions for patient 2. Nevertheless, we were able to obtain a meaningful signal when compared to the manually extracted signal since part of the diaphragm was close to the centre of rotation. This could be accounted for by generating two signals from different lungs respectively depending on which lung(s) are visible in the FOV, and then using a combination of both signals. The current computational time of the method proposed is approx 30min, however this can be improved upon by using a more dedicated hardware such as a Graphic Processing Unit (GPU). In addition, it is not possible to apply the proposed method “on the fly” during CBCT projection acquisition since the reconstructed volume is required for our approach, whereas for technique such as the AS method, it is possible to extract the respiratory signal “on the fly”.

## 3.5 Conclusion

In this work we have introduced a novel method of extracting respiratory motion from CBCT projection data, with the primary aim of extracting a signal that is better correlated to the internal motion, and hence suitable for building motion models (see chapter 4) as well as performing 4D-CBCT reconstructions. The method allows selecting a ROI for generating the signal, allowing targeting a specific region such as the diaphragm or a tumour. We have shown on simulated data that our method can produce a signal from both the diaphragm and the tumour that is strongly correlated with the tumour SI motion, and that the method is robust to tumour size and location. We have also demonstrated that the method can be successfully applied to real clinical data.

Chapter 4

# Deriving a respiratory motion model from CBCT projection data

## 4.1 Aim of this chapter

Respiratory induced organ motion is one of the primary sources of tumour localisation uncertainty, which can cause error and uncertainty during radiotherapy. 4D-CT can be used to measure motion such that it can be taken into account during treatment planning, but the magnitude and pattern of breathing can vary from day to day, thus the 4D-CT may not give a good representation of the motion at the time of treatment. 4D-CBCT can be used to monitor motion just prior to treatment delivery, but this requires a longer acquisition (compared to 3D-CBCT) and hence more dose to the patient. In addition, 4D-CBCT can suffer from poor image quality, and therefore may not give a good representation of the true motion and its variability. Patient specific motion models can potentially provide accurate estimates of the respiratory motion, including modelling breath to breath variability. McClelland et al. [97] proposed a framework unifying DIR and motion modelling into a single optimisation. This framework can fit a motion model directly to 'unsorted' dynamic data, *e.g.* multi-slice data (from CT or MR) or projection data (from CBCT). Previously published results demonstrated good performance for CT and MR slice data. Here we adapt the framework to work with CBCT projection data, and investigate the optimal parameters to use and thoroughly evaluated the results (using phantom data)

The method described in McClelland et al. [97] enables the model to be fitted directly to the frames of the CBCT projections, and does not require the data to be first sorted into full 3D volumes. The model can also account for breath-to-breath variability, therefore this approach does not assume reproducible breathing. Since

the model is fitted to all of the frames of the projection image data simultaneously, the full 3D motion can be estimated. MCIR can also be performed based on the estimated motion. Finally, no prior model estimate is required, therefore, the obtained model is only as a result of the CBCT data available on the day of treatment.

It is very difficult to evaluate the motion model accuracy using real patient data as the true motion is unknown. Therefore, in this chapter an anthropomorphic software phantom (XCAT) [102] was used to simulate CBCT projection data and evaluate the method as a proof-of-concept. The next chapter details some amendments required for the method to be applicable to real patient data and provides a proof-of-concept.

## 4.2 Methodology

In this section we discuss the generalized framework unifying image registration and respiratory motion models proposed by McClelland et al. [97], and the adjustments required for it to be applicable for CBCT data.

As discussed previously in section 2.7, a correspondence model is typically used to describe the relationship between surrogate signals and the internal motion of the anatomy. Once the model coefficients are estimated, the model itself can be used to provide a DVF at any time point using surrogate signals acquired during image acquisition. A typical example of a correspondence model is illustrated in figure 2.6. Ideally, 3D dynamic images would be acquired along with a surrogate signal simultaneously. In that case, registration can be used to determine the geometric transformation to warp the reference-static image $I_0$ to be aligned with the

dynamic image $I_t$. The reference static image can be one of the dynamic images or a breath-hold image. The static image is registered to the dynamic images so that the static image is deformed to appear like the dynamic images. Then a correspondence model is fitted between the transformations obtained from each time point and the corresponding surrogate signals.

However, due to the limitations of scanner technology, it is difficult or even impossible to acquire 3D dynamic image data at sufficient time resolution to determine the motion of interest. In particular, the acquisition of a complete set of CBCT projections requires approximately 60 seconds for 3D-CBCT, since the system must perform one complete turn around the patient to cover all projection angles.

The proposed framework in [97] allows the model to be fitted to partial image data such as CBCT projections and it does not require any sorting of the projections. The framework directly optimises the model parameters on all the dynamic projections simultaneously, it transforms a reference static image $I_0$ such that the projections of the deformed images match the measured projections. The cost function $C_total$ to be minimised on all the dynamic projections in order to find the optimum parameters is given by equation :

$$C_{total} = \sum_{t=1}^{Ni} C_t \tag{4.1}$$

$$C_t(P_t, I_0, M_t) = Sim(P_t, A_t(T(I_0, M_t))) + Con(M_t) \tag{4.2}$$

$$M_t = R_1 s_{t,1} + R_2 s_{t,2} + R_3 \tag{4.3}$$

where $N_i$ is the number of projections. $P_t$ is the measured projection, $A_t$ is the projection operator, $M_t$ is the motion parameter which is a Free Form Deformation (FFD) with b-splines, $T(I_0, M_t)$ is the reference static image $I_0$ deformed with the motion parameters $M_t$. $A_t(T(I_0, M_t)$ denotes the simulated projection at time $t$. The FFD is a function of the model parameters $R$ and the surrogate at time $t$, where we used a linear model of two surrogate signals eq. 4.3. Finally, the *Sim* and *Con* are similarity and constraint functions respectively.

We can use a gradient based technique to find the minimum of the cost function. Calculating the gradient of the cost function, $C_{total}$, with respect to the model parameters $R$:

$$\frac{\partial C_{total}}{\partial R} = \sum_{t=1}^{Ni} \frac{\partial C_t}{\partial R} \tag{4.4}$$

The chain rule can be used to calculate the gradient for each projection, $\frac{\partial C_t}{\partial R}$

$$\frac{\partial C_t}{\partial R} = \frac{\partial M_t}{\partial R} \frac{\partial C_t}{\partial M_t} \tag{4.5}$$

This means that to optimise the model parameters on the projection data directly, the gradient of the cost function with respect to the motion parameters can be calculated from the gradient of the similarity measure and the constraint term with respect to $M_t$, and the gradient of the motion $M_t$ with respect to the model parameters $R$. For the model of eq. 4.3:

$$\frac{\partial M_t}{\partial R_1} = s_{t,1}, \quad \frac{\partial M_t}{\partial R_2} = s_{t,2}, \quad \frac{\partial M_t}{\partial R_3} = 1 \tag{4.6}$$

To compute the gradient of cost function with respect to motion parameters, $\frac{\partial C_t}{\partial M_t}$, we note that the similarity *Sim* is defined in the space of the partial image data but the transformation parameterized by $M_t$ is defined in the space of the full images. Therefore, when calculating, $\frac{\partial C_t}{\partial M_t}$, it is necessary to transform the gradient of *Sim* from partial image space into full image space. This is done using the adjoint of the image acquisition function $A_t^*$ i.e. the adjoint of the forward projection operator, which is the back projection operator. Quoting from [97]:

$$\frac{\partial C_t}{\partial M_t} = \frac{\partial Sim}{\partial M_t} + \frac{\partial Con}{\partial M_t} \tag{4.7}$$

$$\frac{\partial Sim}{\partial M_t} = \frac{\partial I_{T_t}}{\partial M_t} \frac{\partial Sim}{\partial I_{T_t}} \tag{4.8}$$

$$\frac{\partial Sim}{\partial I_{T_t}} = A_t^* \left( \frac{\partial Sim}{\partial P_{A_t}} \right) \tag{4.9}$$

*where* $I_{T_t} = T(I_0, M_t)$*, and* $P_{A_t} = A_t(I_{T_t})$*.* $\frac{\partial Sim}{\partial P_{A_t}}$ *is the gradient of the similarity measure with respect to the simulated partial image data,* $\frac{\partial Sim}{\partial I_{T_t}}$ *is the gradient of the similarity measure with respect to the transformed reference state image, and* $\frac{\partial I_{T_t}}{\partial M_t}$ *is the gradient of the transformed reference-state image with respect to the motion parameters, e.g. for B-spline transformations this is the spatial gradient of the reference-state image transformed by* $M_t$ *and convolved with the B-spline kernel.*

[97].

Theoretically the framework can work with any type of data, providing $A_t$ and $A_t^*$ are known. The initial implementation used for the framework only worked with volume and slice data (where $A_t$ was a point-spread-function of the slice data). That initial implementation was extended to work with CBCT projection data by a Masters student from CREATIS. OpenRTK was used to implement the forward and back-projection [107]. However, the different orientations used by the motion modelling framework (which was based on Nifty-Reg and used nifti images so assumed the standard nifti RAS orientation) and OpenRTK (which uses a scanner based coordinate system [108]) caused complications meaning the code did not work correctly when using a multi-resolution approach.

Hereafter are the contributions of the author of this thesis. The coordinate system complication was addressed by consistently converting between Nifty-Reg and OpenRTK coordinate systems. In order to do this, we assumed that the patient was always scanned in a supine position (which lung RT patients are) this meant the scanner based coordinate system was equivalent to LSA patient based coordinate system. In addition the initial work of the Masters student only included limited validation using an existing motion model from a 4D-CT. Here we have performed a thorough validation using the XCAT phantom including tuning of the hyper-parameters.

The ability to use the motion model for MCIR was also implemented based on the FDK algorithm. The estimated motion from the model was used to warp the back-projection step of the FDK before performing a weighted average of the

deformed back projections to obtain a MCIR. Each individual back-projection was deformed using a different DVF. When performing the MCIR the dynamic data must be deformed into the space of $I_0$. The inverse of the estimated transformations can be used. However, this can be computationally expensive. Instead the adjoint function of the transformation is used *i.e.* "push-interpolation" instead of the usual "pull-interpolation". This can potentially lead to voxels with no intensity information (*i.e.* "holes") in the deformed image, however if the holes are not present in the same location for all dynamic images, the holes will not affect the final MCIR. A visual representation of the push/pull interpolation is shown in figure 4.1.

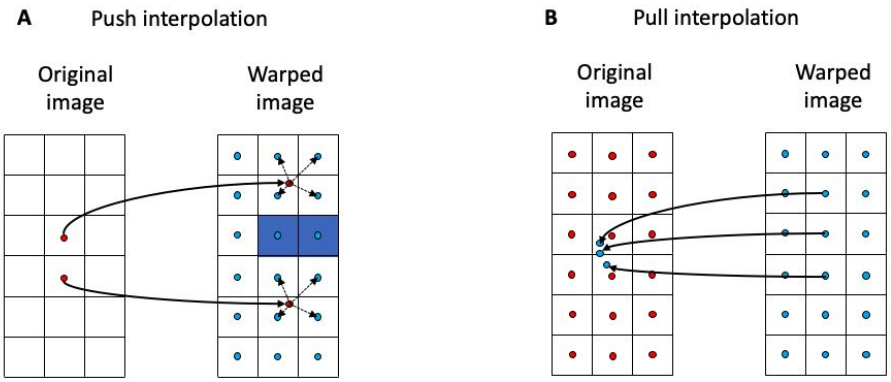


**Figure 4.1:** Push/Pull interpolation. (A) push-interpolation, here holes can occur in the deformed image, marked in blue. (B) pull-interpolation mitigates this problem since intensity values at locations that do not coincide with pixel coordinates can be obtained from the original image using interpolation

.

The framework was implemented using the open source NiftyReg software. This uses a B-spline transformation model and CG optimisation in a multi-resolution approach. For all the experiments performed SSD as been used as the similarity measure, and a linear correspondence model with two surrogate signals relating the motion to both the value and temporal derivative of the surrogate signal was used. This correspondence model has been popular in the literature due to its

simplicity and ability to model both intra- and inter-cycle variations [53, 109, 110]. It is worth noting that the reference $I_0$ volume does not represent the average position of the anatomy during acquisition, therefore an extra constant offset model parameter was included in the linear model to account for the difference between the $I_0$ and the average position of the anatomy. This is achieved by including an extra surrogate signal with a constant value of 1 for every time point. Hence, when the motion parameters are computed from the linear combination of the model parameters weighted by the surrogate signal for every time point, the constant signal enables the offset between the reference volume $I_0$ and the average position to be accounted for. If the offset term is not included, then the reference image position has to corresponds to the (0,0) value of the breathing signal. However the reference image that was used in this chapter does not correspond to the average position, so therefore the constant term was included. For the experiments in this study, the GT motion and reference-state image are known, so the results of the motion model fitting and the MCIR can be quantitatively assessed.

### 4.2.1 Hyper-parameter Optimisation

Here we discuss the hyper-parameters that were tuned in order to obtain a model with optimal performance. We define the optimum performance as the ability of the respiratory correspondence model generated to estimate the motion of the patient while also taking into account computational efficiency such as computational time. The 4 main hyper-parameters that were investigated were: Control Point Grid (CPG) spacing, the number of respiratory correspondence model fitting itera-

tions, number of the spatial multi-resolution levels, and the numbers of projections required to fit the model. Below is a brief introduction to each of these hyper-parameters and their effect on the resulting model fitting.

The balance between accuracy and robustness is partly controlled by the CPG spacing for the B-Spline basis. In order to be able to model a localised deformation a fine CPG spacing is typically used. However, to model a global and smooth deformation a coarse control point spacing would be beneficial.

The spatial multi-resolution can be described as a pyramid set of successively smoothed and down-sampled images of the original image and the transformation. The model fitting is initially performed at the coarsest level. Once the maximum number of iterations have been used (or once there is no more improvement), the next level of the pyramid is initiated, finishing with the original (finest) resolution level. In addition, when moving to a finer resolution level the transformation needs to be up-sampled, so that the CPG spacing relative to the voxel size is the same at each level. Usually the amount of data decreases two-fold in each dimension at each successive step. Performing the model fitting process at multiple scales ensures that large scale features are aligned with the finer structures being registered during later levels. This approach aims to avoid that the optimisation is trapped in a local minimum. It can also help to speed up the computational time by reducing the amount of data used in earlier iterations.

Another parameter is the maximum number of iterations used at every level of the pyramid. Lower number of iterations can lead to under-fitting, while higher iteration number can increase the computational time. We will therefore investigate

the optimum balance between the number of iterations and an adequate estimate from the model.

CBCT uses approximately between 630 and 1008 projections (true-beam) depending on the type of acquisition, we hypothesise that we may be able to still fit an accurate model with substantially fewer projections. Therefore, we investigate if reducing the number of projections still results in a reasonable fit. We disregarded every other projection, etc. retaining $1/2$, $1/3$, etc.. of the total projections. This was to retain a consistent angular sampling so that the motion was sampled from different directions as well as over different breath cycles (thus sampling breath-to-breath variability.)

## 4.3 Experiments

In this section we describe the experiments used to tune and evaluate our proposed method. Firstly the optimum hyper-parameters to fit the model and the choice of surrogate signal to use were investigated. Secondly the proposed method was tested on different simulations based on clinical acquisition scenarios. Finally using the motion estimated from the simulations MCIRs were performed and compared to other reconstruction methods.

### 4.3.1 Simulations

The accuracy of the developed method in this study was evaluated with different patient scenarios and scanning parameters using datasets from the XCAT computerized phantom [102]. We used XCAT to simulate attenuation images of the thoracic cavity of the human body as previously discussed in section 3.2.4.1 .

#### 4.3.1.1 Experimental setup

The XCAT phantom software can additionally output a DVF defining the motion of each voxel in the phantom. The XCAT DVFs (which we will refer to as XDVFs) are similar to DVFs calculated with DIR techniques, with the very important discrepancy being that the XDVFs are determined directly from the motion of the XCAT NURBS surfaces, and do not rely on any type of DIR. Consequently, the XDVF can be used to determine the location and deformation of every voxel in the phantom throughout respiration without the uncertainty introduced by using DIR-derived DVFs.

Though XCAT provides reproducible and realistic simulated motion, it has its limitations. For instance, XCAT can generate inconsistent and non-invertible deformations whereby structures can move through each other. Eiben et al. [25] proposed a post processing technique for the XCAT DVF, that corrects and inverts the XCAT-DVFs while preserving sliding motion between organs. We have incorporated that technique into our approach in order to obtain modified XDVFs (M-XDVFs) suitable for validation.

We briefly summarise the approach here. There are 3 main steps involved in the post processing. Firstly the XCAT volume with a tumour is generated while XDVFs for the entire time series are also generated. Secondly the XDVFs are processed for consistency and inverted. They were smoothed such that structures no longer moved through each other, but in such a way as to preserve sliding motion along the chest wall. Finally, a post-processed XCAT time series is obtained by deforming the reference XCAT volume with a tumour using the M-XDVFs. An illustration of

the post-processing steps is shown in figure 4.2.

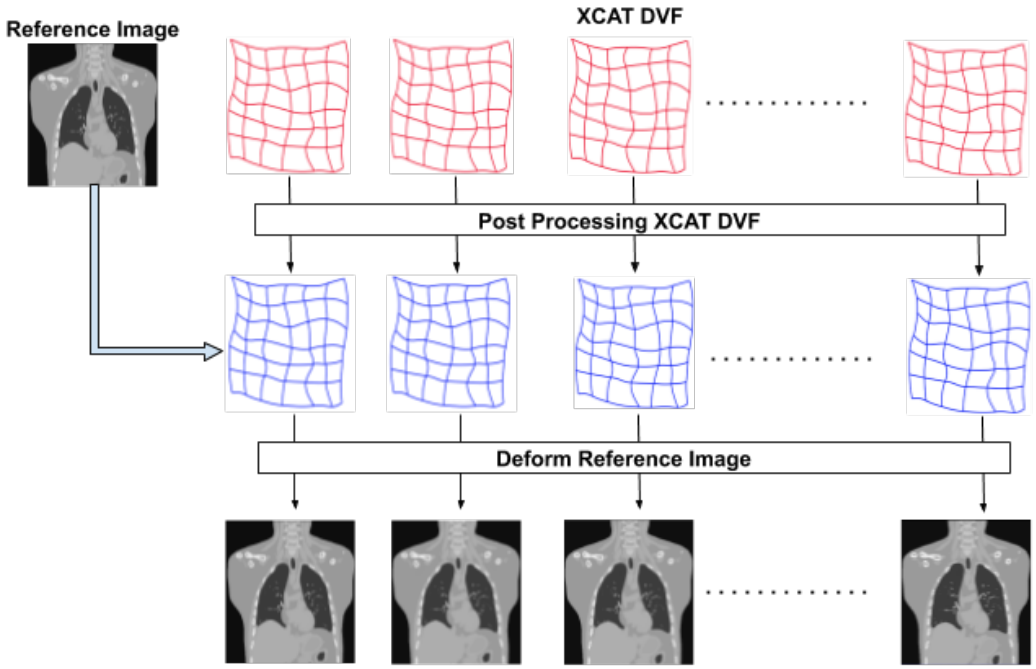


**Figure 4.2:** An illustration of the XDVFs post processing method that corrects and inverts the XCAT-DVFs while preserving sliding motion between organs.

In this study, the XCAT phantom motion simulations were based on respiratory signals extracted from cine MRI data from a lung cancer patient. The respiratory signals were acquired from a local region of the diaphragm and the skin surface sections respectively (*i.e.* diaphragm and skin-AP signals). The phantom datasets and the experiments performed with them mimicked the real datasets and experiments as closely as possible. The volumes simulated from the XCAT had $256 \times 256 \times 180$ voxels with a spatial resolution of $1.757 \times 1.757 \times 2.0 mm^3$. The number of volumes and their temporal resolution varied based on the CBCT system and the type of scan being simulated (see table 4.1).

CBCT projections were then simulated from the post-processed 4D-XCAT volumes illustrated in figure 4.2 using openRTK [104]. The CBCT geometry and the

simulation parameters were based on real CBCT scans and the patient data available for this study. We simulated CBCT projections based on the true-beam and on-board Imaging (OBI) systems (both from Varian Medical Systems, Palo Alto, CA) respectively. The detector for both systems contains $1024 \times 768$ pixels of size $0.38 \times 0.38 mm^2$. This was scaled down by a factor of 4 in each direction to reduce computational time, giving $256 \times 192$ pixels of size $1.55 \times 1.55 mm^2$. Different types of CBCT acquisitions were simulated in order to replicate different clinical scenarios, as detailed in table 4.1. The acquisition parameters for the CBCT were: Source to iso-centre distance of 1000mm, source to detector distance of 1536mm. Simulation 1 was based on the OBI system, whereas simulations 2-5 were based on the true-beam configuration. As previously described in section 3.2.4.1, at our clinical institution, acquisitions using the OBI system are typically centred on the tumour (at the treatment iso-centre), and therefore the projection data is often truncated and some of the anatomy is outside the FOV of the reconstruction. However, acquisitions using the true-beam system are typically centred close to the centre of the patient, and therefore the reconstructions include all of the patient's anatomy.

**Table 4.1:** Parameters for the XCAT simulations used in this chapter.

| | Tumour size(mm) | No.of Projections | Acquisition Duration (s) | Acquisition type | Tumour Position |
|---|---|---|---|---|---|
| Simulation 1 | 20 | 683 | 60 | 3D-CBCT | Middle ** |
| Simulation 2 | 25 | 1008 | 144 | 4D-CBCT | Upper ** |
| Simulation 3 | 15 | 1008 | 144 | 4D-CBCT | Lower * |
| Simulation 4 | 25 | 1008 | 144 | 4D-CBCT | Upper * |
| Simulation 5 | 30 | 1008 | 144 | 4D-CBCT | Lower ** |

* Left lung.

** Right lung.

Simulation 1 was based on 3D-CBCT scan with a duration of 60 seconds. Simulations 2-5 were all based on 4D-CBCT acquisition with a scan duration of 144 seconds. In addition, simulation 1 had a tumour of size 20mm diameter in the mid-lung position, simulations 2 and 4 had tumours in the upper-lung of size 25mm diameter, whereas simulations 3 and 5 had the tumour positioned in the lower lobe of the lung with 15mm and 30mm diameter tumour sizes respectively. The simulations performed in this study are summarised in table 4.1, and the tumour positions relative to the body for all simulations are also shown in figure 4.3. We have chosen these simulations to analyse the different effects that the motion model has for different regions of the lung, hence why we have positioned the tumour in different parts of the lungs, similarly we have used different tumour sizes for each simulation in order to investigate how the size of the tumour in each region is affected by the motion estimated by the model.

The simulation parameters presented in table 4.1 were based on the acquisition parameters from our clinical institution. At the start of this project only 3D OBI data were acquired at our clinical institution, hence why the hyper parameters of the model were optimised based on the OBI configuration (simulation 1). Acquisitions from the true-beam machine were later acquired on the clinical system, as a result simulations 2-5 were based on true-beam acquisition. Additionally we have simulated 4D-CBCT acquisitions for simulations 2-5 in order to fairly compare our proposed method to the standard respiratory correlated 4D-CBCT images.

The phantom data-set from simulation 1 was used to optimise the hyper parameters of the model fitting as described in section 4.2.1. We used the GT respiratory

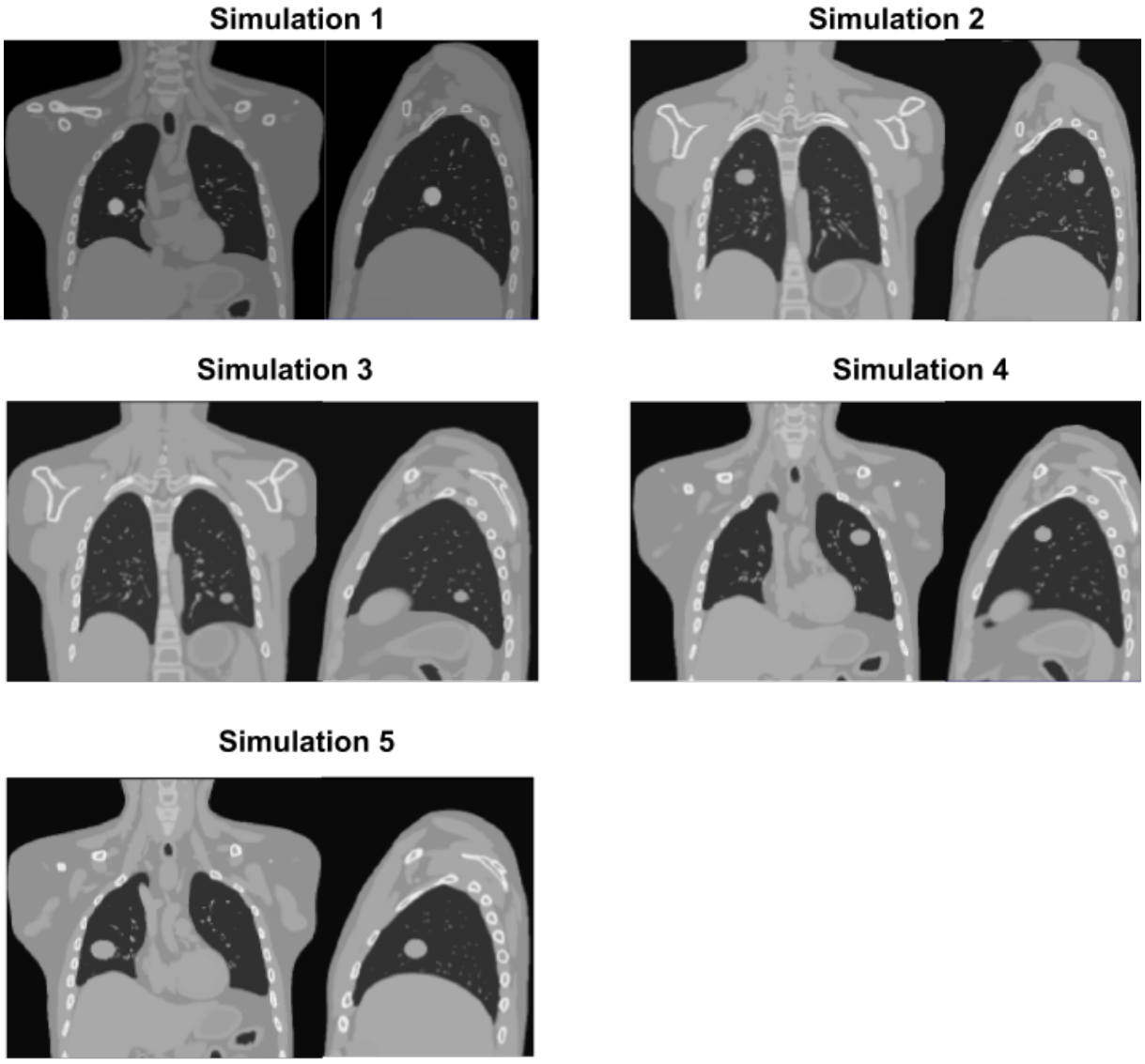


**Figure 4.3:** The coronal and sagittal slices of the XCAT volumes showing the relative position of the tumour for each simulation.

signals (the two input parameters to define the motion: the translation of the diaphragm and the anterior-posterior translation of the thoracic region as previously described in section 3.2.4.1) as the surrogate signals. Here we optimised the CPG spacing, the number of spatial multi-resolution levels, the number of model fitting iterations, and the number of projections required to fit the model respectively. For the CPG spacing optimisation the same spacing was used in all directions. We in-

vestigated CPG spacings of: 2, 5, 8, 10, and 13 voxels. We initialised the hyper parameters with the default setting of 3 multi-resolution levels with the maximum number of iterations to fit the model to 300 per level, while varying the CPG spacing. The results for the model while tuning the hyper parameters for simulation 1 were evaluated using the Deformation Field Error (DFE) (see section 4.3.2) between the model estimated DVF and the M-XDVF. The mean DFE was computed from each time point and these were averaged over all time points. A body mask was also used when calculating the DFE in order to ignore voxels outside the body. Based on the results obtained, an optimum value for the CPG spacing was chosen and this value was used henceforth for future experiments. Furthermore the results from simulations 2-5 were evaluated using DFE, DICE coefficient, and the Centre of Mass Distance (COMD). Further information on the evaluation metrics is provided in section 4.3.2.

5 different multi resolution levels (1 level, 2 levels, ..., 5 levels) were investigated while fixing the CPG spacing (based on the value found in the previous experiment) and the maximum number of iterations to a constant value to fit the model (300 max fit iterations). The optimum number of levels was determined based on the mean DFE. Based on this the optimum value was used henceforth.

Using the default number of iterations to fit the model (300 iterations) allowed the fitting to converge, *i.e.* it stopped before 300 iterations were used, while 50 iterations caused the fitting to stop early, but preliminary experiments indicated most of the improvement in the cost function occurred during the first 50 iterations, so we hypothesised that stopping after 50 iterations would reduce the computation

time but only have a minor impact on the results. Therefore we investigated setting the maximum number of fitting iterations to 50 and to 300. The total number of projections used to fit the model was varied from using the full data to $1/10$ (1, $1/2$, ..., $1/10$) of the datasets. The mean DFE and the computational time was used to determine the optimum values required to fit a motion model.

We investigated the effect of the choice of surrogate signal on the model estimated motion. The method described in section 3.2.6 was used to generate a surrogate signal from the D-ROI and the T-ROI respectively for simulation 1. A correspondence model was built using the GT signal, the D-ROI signal and the T-ROI signal respectively as surrogates to drive the model fitting, while using the optimum parameters as discussed above.

From the optimum hyper parameters obtained for the model based on simulation 1 and using the optimum surrogate signal, we used these parameters on the remaining simulations to produce a respiratory motion model estimate. For simulations 2-5, the mean DFE between the M-XDVF and the correspondence model estimate was evaluated, the COMD (see section 4.3.2) between the GT tumour volume and the model estimated tumour volume was also calculated. In addition the DICE coefficient was obtained for the tumour in the GT volumes and the model estimated volumes.

MCIR was obtained for simulations 2-5. The image quality of the MCIR tests if the accuracy of the model is sufficient to compensate for the motion during the acquisition and hence improve the quality of the reconstructed image. The MCIR was compared to different volumes: (1) $I_{staticRecon}$, this is a reconstruction from the sim-

ulated projections through a single static XCAT volume (at the reference position) and it is used so as to isolate and study the effects of motion on the reconstructed image (by completely removing the effects of motion). (2) a MCIR with the known GT motion used to produce the volumes initially. (3) a standard FDK reconstruction (no motion model applied). (4) a 4D-CBCT reconstruction. The comparisons between the different volumes were performed using the Root Mean Squared Error (RMSE) and PSNR, these terms are described in section 4.3.2. In addition line profiles through important structures such as the tumour and diaphragm were also assessed. Image quality was also evaluated qualitatively by assessing the visibility of structures within the volume.

### 4.3.2 Evaluation

To assess the accuracy of the estimate from the model, we performed quantitative and qualitative analysis using different metrics. Since the GT DVFs were available, the results of all the phantom experiments were assessed by calculating the Displacement Field Error (DFE), defined as the 3D Euclidean distance at each pixel between the GT displacement field and the displacement field resulting from the fitted motion model. The DFE was calculated at every pixel inside the body of the deformed phantom for every time-point. Pixels outside the phantom were ignored as they do not contain any image data to guide the model fitting.

A tumour volume for each time point can be simulated by using the M-XDVF to deform a reference tumour XCAT 3D volume. In addition, the DVF obtained from the fitted model can also be used to deform a reference tumour XCAT 3D vol-

ume. These were converted into binary volumes. Then we can assess the ability of the model to estimate the tumour position at each time point using 2 methods, (1) by calculating the centroid of the tumour in each image and measuring the 3D Euclidean distance between them (COMD), and, (2) calculating the DICE coefficient between the respective tumour volumes.

$$COMD = \left\| c(Y) - c(\hat{Y}) \right\| \tag{4.10}$$

$$DICE = \frac{2|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \tag{4.11}$$

where Y is the GT volume of the tumour and $\hat{Y}$ is the model estimated volume for the tumour. In equation 4.10, $c(Y)$ and $c(\hat{Y})$ represent the centre of mass coordinate of volumes Y and $\hat{Y}$ respectively and $\|\cdot\|$ is the Euclidean distance.

To qualitatively evaluate the ability of the fitted model to model respiratory motion, we used a colour overlay technique. We used the GT volumes and the estimated volumes using the model at each time point respectively and superimposed both images. The blue and red colour shows regions of discrepancy between the GT volume and the volume as estimated from the fitted model and this is done at every time point to produce a movie.

We employed different metrics to assess the quality of the MCIR resulting from the fitted model. The Root Mean Square Error (RMSE) for two arbitrary images can be defined as in equation 4.12. The volumes used in this study were from the XCAT phantom where the voxel values represent the attenuation coefficient. Note that this

has not been converted to HU.

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(A_i - B_i)^2} \tag{4.12}$$

where $A_i$ is the voxel value of the $I_{staticRecon}$ and $B_i$ is the voxel value of the volume being assessed (which is not always the MCIR volume) and $n$ the number of voxels.

In addition, to further quantify the accuracy of the reconstructed images, we used the PSNR defined as:

$$PSNR = 10Log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{4.13}$$

where $MAX$ is maximum possible voxel value of the volume, and $MSE$ is the mean squared error, *i.e.* MSE between $I_{staticRecon}$ and the volume to be assessed.

## 4.4 Results

### 4.4.1 Hyper-Parameter Optimisation

DFE results obtained for the hyper-parameter optimisation of the model are summarised in figure 4.4. The mean DFEs for the different CPG spacings are shown in table 4.2. The lowest DFE was obtained for a CPG spacing of 8 voxels, here we obtained a mean value of 0.83 and a std of 0.44, with a maximum DFE of 2.40mm. Figure 4.4A shows a graphical representation of the DFEs. The highest mean DFE value was obtained when a CPG spacing of 5 voxels was used. Meanwhile when larger CPG spacings such as (8,10,13) voxels were used lower DFE was observed. When 10 voxels were used as the CPG spacing the DFE was slightly higher than that

of 8 voxels, which was not expected. Nevertheless the DFE value obtained when 10 voxels was used was still considerable lower than when smaller CPG spacings were used. These results indicated that coarser CPG spacing generally produced relatively lower DFE errors when compared to using finer CPG spacing. In the case of using no motion model the mean DFE obtained was 2.40mm with a maximum DFE value of 7.56mm. Based on these results, we used a CPG spacing of 8 voxels for the rest of the hyper-parameter optimisation since this produced the lowest CPG spacing. Colour overlay comparing the GT image and the model estimated image for the 2 voxels and 8 voxels CPG spacings are shown in figure 4.5. The corresponding movie for these images over all time-points is in the attached supplementary material. (*CPG_Spacing_2pixels_vs_8pixels.mp*4 - *First row shows the coronal slice comparison between the GT and the model estimate when using a CPG spacing of 2 voxels, while the second row shows the coronal slice comparison between the GT and the model estimate when using a CPG spacing of 8 voxels. Based on the colour overlay shown in the middle column, both CPG spacings were able to model the motion accurately however the colour overlay for the CPG spacing of 8 voxels showed lower amount of mismatch between the GT and the model estimate indicating that it performed better relative to when using a CPG of 2 voxels*)

The DFEs obtained for the optimisation of the multi-resolution levels are shown in table 4.3. 5 different numbers of resolution levels were investigated in total. Using 1 resolution level resulted in the highest mean DFE of 1.42mm with a std of 1.04mm and the maximum DFE of 4.63mm, while using 3 resolution levels yielded the lowest mean DFE of 0.83mm with a std of 0.44mm and a maximum

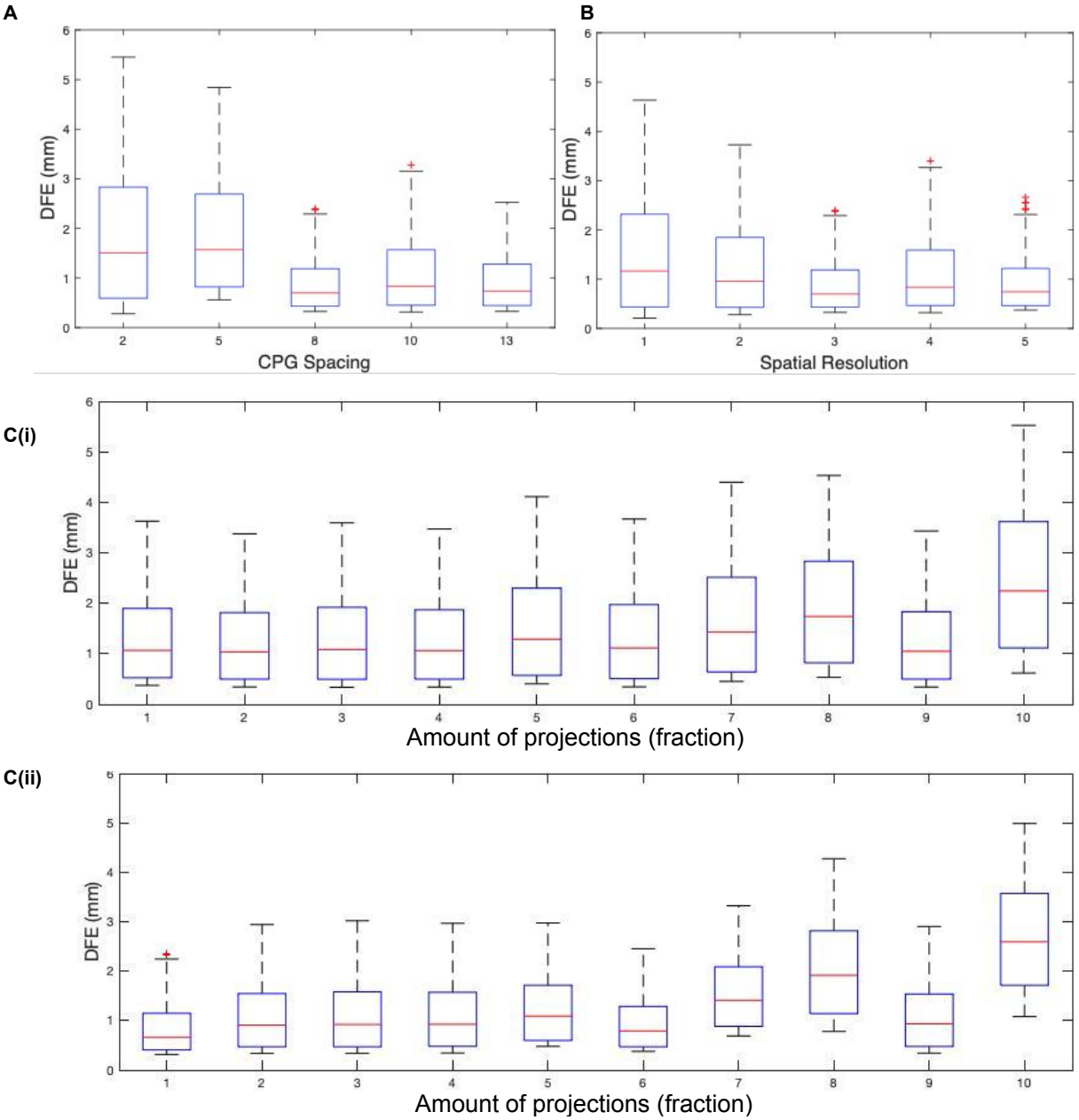


**Figure 4.4:** Box plot of the DFEs in mm for the different hyper-parameters investigated. (*A*) the DFEs obtained for the different CPG spacing, (*B*) the DFEs obtained for the different spatial resolution. (*C(i)*) the DFEs obtained while varying the number of projections required to fit the model using 50 iterations. (*C(ii)*) the DFEs obtained while varying the number of projections required to fit the model using 300 iterations.

DFE of 2.40mm. Using 4 resolution levels yielded a mean DFE of 1.06mm with a maximum of 3.40mm and a std of 0.66. The gradual reduction of the DFE as the number of multi-resolution levels was increased (with the exception of using 4

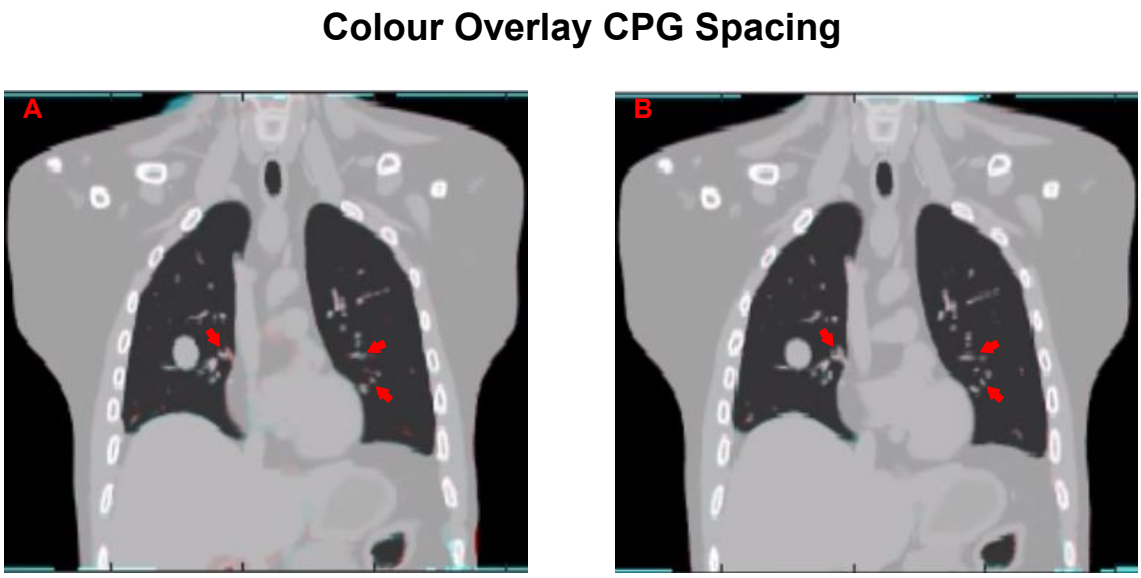


**Figure 4.5:** Coronal slice of the colour overlay images between the GT volume and the model estimated volume at EE, the red and blue colours show regions where there is discrepancy (also highlighted by the red arrows) between the GT volume and the model estimated volume. (*A*) 2 voxels CPG spacing; (*B*) 8 voxels CPG spacing.

**Table 4.2:** DFEs (mm) for the different CPG spacings used to build the correspondence model, while the other hyper-parameters were: multi-resolution levels = 3, maximum iterations to fit the model = 300.

| CPG spacing (voxels) | mean | max | min | std |
|---|---|---|---|---|
| 2 | 1.76 | 5.45 | 0.28 | 1.23 |
| 5 | 1.79 | 4.84 | 0.56 | 0.99 |
| 8 | 0.83 | 2.40 | 0.33 | 0.44 |
| 10 | 1.04 | 3.28 | 0.32 | 0.64 |
| 13 | 0.89 | 2.53 | 0.33 | 0.48 |

resolution levels) suggests that the optimisation did result in local minima if only 1 or 2 levels were used, but did not when 3 or more were used. Colour overlay comparisons of the model estimate and the GT image of a coronal slice between 3 spatial resolution levels are shown in figure 4.6. The corresponding movie for these images over all time-points can be found in the attached supplementary material. (*Res_level_1_vs_3_level.mp4 - First row shows the coronal slice comparison*

*between the GT and the model estimate when the number of multi-resolution levels was set to 1, while the second row shows the corresponding image when using 3 resolution levels. Looking at the colour overlay over all time points, both cases were able to model the motion accurately, but less discrepancy between the GT and the model estimate can be observed for 3 resolution levels, thus supporting the quantitative analysis.*) Figure 4.4B shows the DFEs for the different number of spatial resolution levels used. 3 spatial resolution levels were used for subsequent experiments since this yielded the lowest DFE value.

**Table 4.3:** DFEs (mm) for the different numbers of spatial resolution levels to build the correspondence model, while the other hyper-parameters were: CPG spacing = 8 voxels, maximum iteration to fit the model = 300.

| Spatial Resolution | mean | max | min | std |
|---|---|---|---|---|
| 1 | 1.42 | 4.63 | 0.21 | 1.04 |
| 2 | 1.18 | 3.73 | 0.28 | 0.79 |
| 3 | 0.83 | 2.40 | 0.33 | 0.44 |
| 4 | 1.06 | 3.40 | 0.32 | 0.66 |
| 5 | 0.87 | 2.66 | 0.37 | 0.46 |

DFEs obtained when using different number of projections to fit the model are summarised in tables 4.4 and 4.5 respectively. The maximum number of respiratory correspondence model fitting iterations was set to 50 and 300 respectively. For the 50 iterations setting, the largest mean DFE of 2.40mm was obtained when using $1/10$ of the projection data, while the lowest DFE of 1.18mm was obtained for $1/2$. Using $1/4$ of the data yielded a mean DFE of 1.21mm. Similarly for the 300 iterations setting, the largest DFE of 2.66mm obtained when using $1/10$ of the data, while the lowest DFE of 1.03mm was obtained for $1/2$ and using $1/4$ of the data yielded a mean DFE of 1.05mm. Overall from the observation of figure 4.4 in most

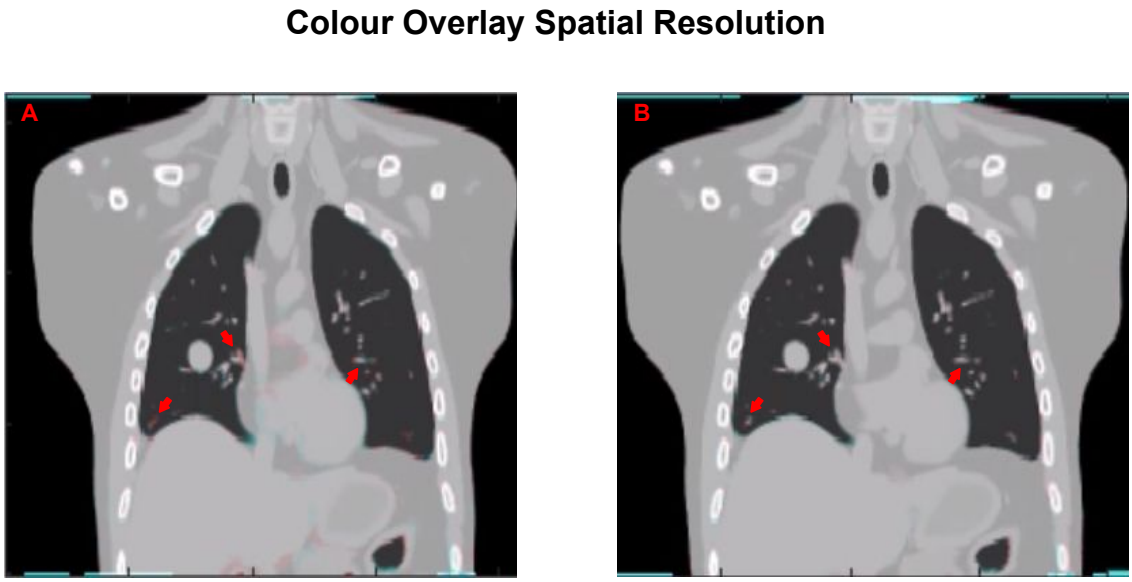


**Figure 4.6:** Coronal slice of the colour overlay images between the GT volume and the model estimated volume for different numbers of spatial resolution levels. The red and blue colours show regions where there is discrepancy (also highlighted by the red arrows) between the GT volume and the model estimated volume. (*A*) using 1 spatial resolution level. (*B*) using 3 spatial resolution level.

cases the DFEs were lower when a higher iteration value of 300 was used compared to using 50 iterations, however this is at a cost of increased computational time. (see figures 4.4 and 4.5)

There is a clear trend when using $1/7$ of the data or less, the accuracy decreases as less data is used, with the exception of using $1/9$ of the data. Currently this result is difficult to explain and requires further investigation, nevertheless this further highlights the importance of using more datasets to find the optimal parameters. We chose $1/4$ with 50 iterations as the optimum parameter to use, since the computational time was lower when using 50 iterations, and the DFE of $1/4$ also yielded a relatively lower DFE. Qualitative results also support this suggesting that there is not a substantial increase in the accuracy when using 50 or 300 maximum fitting iterations. Figure 4.7 shows the colour overlay comparison be-

tween the GT image and the model estimated image when using different number of projections to fit the model and using different maximum number of respiratory correspondence model fitting iterations. The corresponding movie for these images over all time-points can be found in the attached supplementary material. (*No_iteration_No_of_projections.mp4 - Shows the comparison between the GT and the model estimate when using 50 and 300 iterations respectively, and while using the full data set and 1/4 of the number of projections to fit the model. From the colour overlay images shown in the middle column, no major difference can be observed when using the full data set or 1/4 of the data to fit the model, and a similar trend was observed when using either 50 or 300 iterations to fit the model, no major changes were observed.*).

**Table 4.4:** DFEs (mm) for the different number of projections used to build the correspondence model, while the other hyper-parameters were: CPG spacing = 8 voxels, spatial resolution = 3 levels , maximum iterations to fit the model = 50.

| Amount of data (fraction) | mean | max | min | std | time (hr) |
|---|---|---|---|---|---|
| 1 | 1.25 | 3.63 | 0.37 | 0.75 | 2.72 |
| 2 | 1.18 | 3.38 | 0.34 | 0.71 | 1.36 |
| 3 | 1.24 | 3.59 | 0.34 | 0.76 | 0.90 |
| 4 | 1.21 | 3.47 | 0.34 | 0.73 | 0.68 |
| 5 | 1.47 | 4.11 | 0.40 | 0.91 | 0.50 |
| 6 | 1.27 | 3.67 | 0.35 | 0.78 | 0.43 |
| 7 | 1.62 | 4.40 | 0.45 | 0.98 | 0.36 |
| 8 | 1.86 | 4.53 | 0.53 | 1.03 | 0.35 |
| 9 | 1.19 | 3.43 | 0.34 | 0.72 | 0.30 |
| 10 | 2.40 | 5.52 | 0.62 | 1.27 | 0.28 |

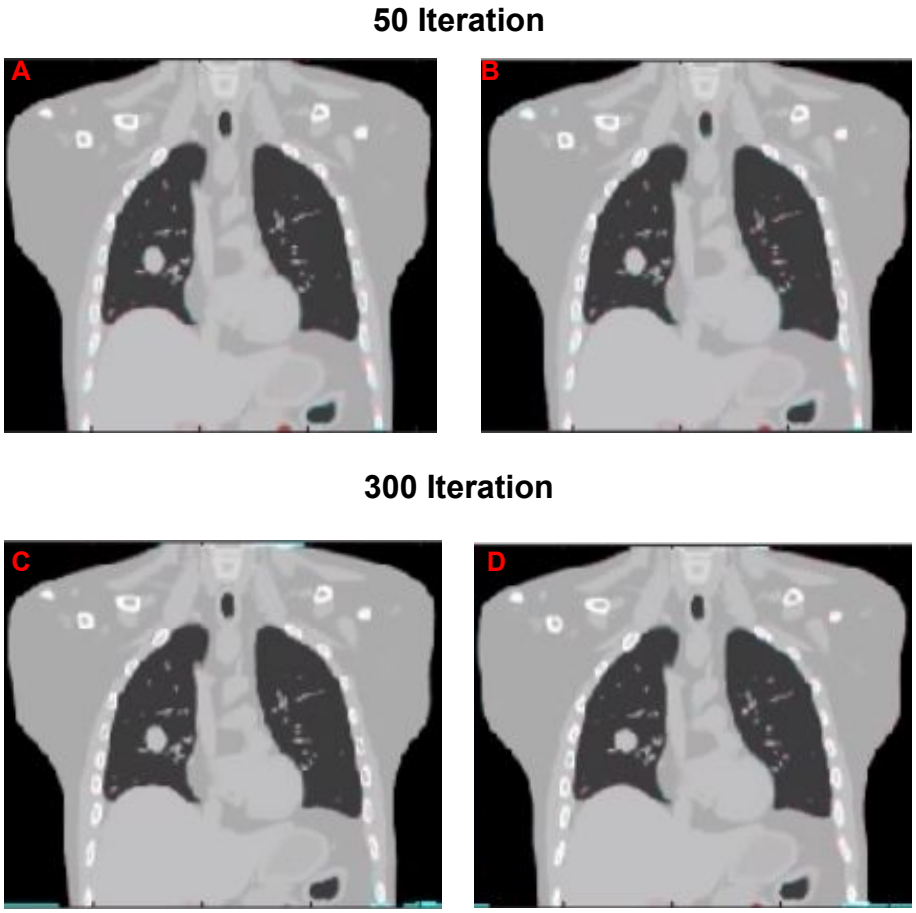


**Figure 4.7:** Coronal slice of the colour overlay image between the GT volume and the model estimated volume. Top row shows the result when 50 maximum fitting iterations were used to fit the correspondence model, while bottom row shows the comparison when 300 maximum fitting iterations was used. *(A and C)* shows the comparison obtained when using the full data to fit the model. *(B and D)* shows the comparison when $1/4$ of the data were used to fit the model respectively.

**Table 4.5:** DFEs (mm) for the different number of projections used to build the correspondence model, while the other hyper-parameters were: CPG spacing = 8 voxels, spatial resolution = 3 levels , maximum iterations to fit the model = 300

| Amount of data (fraction) | mean | max | min | std | time (hr) |
|---|---|---|---|---|---|
| 1 | 0.80 | 2.34 | 0.31 | 0.43 | 14.35 |
| 2 | 1.03 | 2.94 | 0.34 | 0.58 | 4.00 |
| 3 | 1.05 | 3.02 | 0.34 | 0.60 | 3.6 |
| 4 | 1.05 | 2.97 | 0.34 | 0.59 | 1.13 |
| 5 | 1.19 | 2.97 | 0.48 | 0.59 | 1.85 |
| 6 | 0.90 | 2.45 | 0.38 | 0.45 | 2.53 |
| 7 | 1.51 | 3.32 | 0.69 | 0.63 | 0.62 |
| 8 | 2.01 | 4.28 | 0.78 | 0.87 | 1.83 |
| 9 | 1.03 | 2.90 | 0.34 | 0.58 | 0.90 |
| 10 | 2.66 | 4.99 | 1.08 | 0.96 | 0.32 |

### 4.4.2 Choice of Surrogate Signals

For the experiments in this section the optimum hyper-parameters (CPG spacing = 8 voxels, number of spatial multi-resolution levels = 3, number of iterations = 50, number of projections to fit the model = 1/4) determined based on the results from section 4.4.1. We evaluated the impact of using the signals developed in chapter 3 compared to the GT signals. Figure 4.8 shows the box-plot of the DFEs obtained when the GT signal, the D-ROI signal, and the T-ROI signal were used as the surrogate signal to drive the model fitting respectively. A summary of the results is shown in table 4.6. In addition colour overlays of coronal slice images as estimated by the model and the corresponding GT images are shown in figure 4.9 for the 3 different types of surrogate signals used. The corresponding movie for these images over all time-points can be found in the attached supplementary material. (*Surrogate_signals_comparisons.mp4 - The first row shows the coronal slice through the GT volume, the difference colour overlay, and the estimate from the model when the GT signal was used as surrogate to build the motion model. The second and third rows show the corresponding images through the estimated volumes when using the D-ROI and T-ROI signals respectively to fit the model. From observing the colour overlay images in the middle column, the overall motion appears to have been modelled reasonably well, with the diaphragm and other lung structures moving up and down with the breathing. However, the mismatch between the GT image and the model estimated image when using the GT signal was similar to the discrepancy seen when using the D-ROI signal as surrogate to drive the model. In addition, the colour overlay image between the GT image and the model estimated image for the*

*T-ROI showed larger difference between the two, for instance for some time points large discrepancy between the GT and the estimated motion can be observed.*) The results for the COMD of the tumour volume for the signals (GT, D-ROI and T-ROI) used to build the model relative to the GT tumour volume are shown in table 4.7. Overall the DFE obtained using the D-ROI was lower when compared to using T-ROI. This is not greatly surprising since the D-ROI signal had a better correlation to the GT signal relative to the T-ROI, partly due to the fact that the diaphragm region has a higher contrast thus it was easier for the method described in chapter 3 to identify the signal. Nevertheless the mean DFE obtained for the D-ROI and T-ROI were still in the same range and lower than when no motion model was used, therefore suggesting that using either signal as surrogate should produce a suitable estimate of the motion.

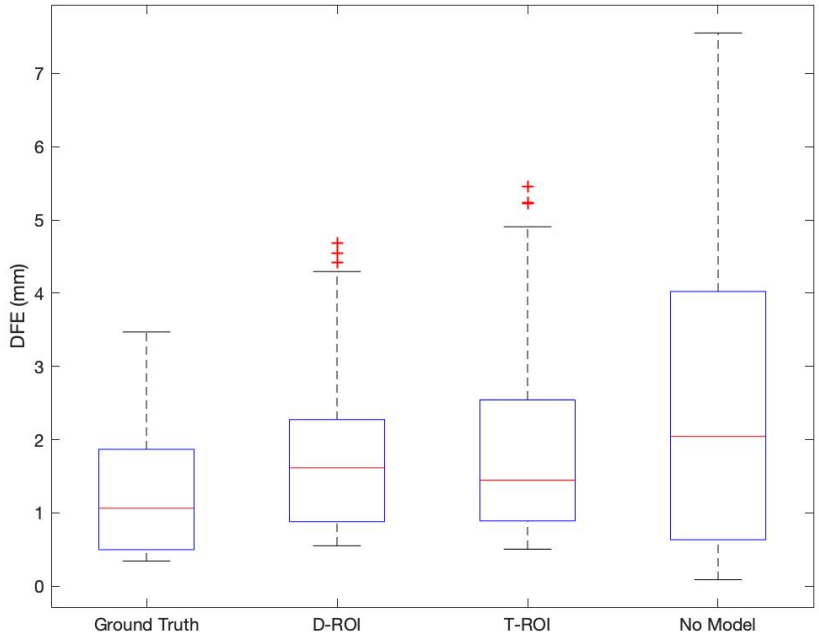


**Figure 4.8:** Box plot of the DFEs for the different surrogate signals used to obtain a correspondence model

**Table 4.6:** DFEs in mm when using different surrogate signals to drive the correspondence model.

| DFE (mm) | mean | max | min | std |
|---|---|---|---|---|
| GT | 1.21 | 3.47 | 0.34 | 0.73 |
| D-ROI | 1.64 | 4.68 | 0.55 | 0.50 |
| T-ROI | 1.74 | 5.46 | 0.50 | 1.00 |
| No Model | 2.40 | 7.56 | 0.08 | 1.83 |

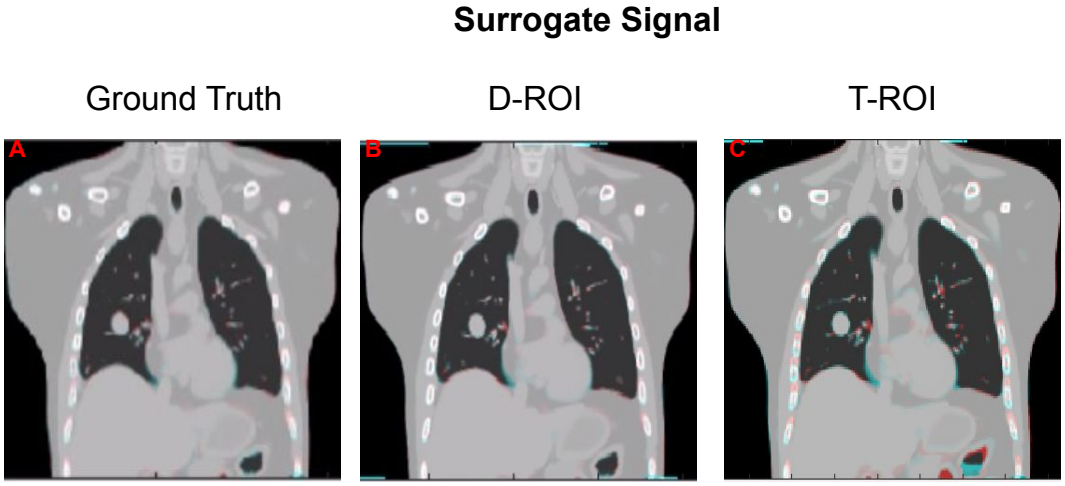


**Figure 4.9:** Coronal slice of the colour overlay image between the GT volume and the model estimated volume. *A* colour overlay comparison obtained when the GT signal was used as surrogate signal. *B* colour overlay comparison obtained when the D-ROI signal was used as surrogate signal. *C* colour overlay comparison obtained when the T-ROI signal was used as surrogate signal.

**Table 4.7:** The distance (mm) between the centroids of two volumes (GT volume and model estimated volume). Shown are the mean and the std based on the different type of signals used to build the model.

| | GT | D-ROI | T-ROI | No Model |
|---|---|---|---|---|
| COMD | 0.67($\pm$0.39) | 1.45($\pm$1.29) | 1.69($\pm$1.34) | 4.72($\pm$3.74) |

Using the optimum hyper parameters obtained previously and the signal from the D-ROI as surrogate, we fitted a motion model to 4 simulations (simulations 2-5) described in table 4.1. The resulting DFEs are shown in table 4.8. The results of the DICE coefficients between the GT tumour volume and the model estimated tumour

volume are shown in table 4.9. The COMD between the GT volume relative to the case of no motion model and when a motion model was applied are shown in table 4.10.

**Table 4.8:** The DFEs in mm obtained when the XDVFs were compared to the model estimates for different simulations

| DFE (mm) | Motion model | No Motion Model |
|---|---|---|
| Simulation 2 | 1.97 ($\pm$0.75) | 2.53 ($\pm$1.79) |
| Simulation 3 | 1.91 ($\pm$0.63) | 2.53 ($\pm$1.79) |
| Simulation 4 | 1.82 ($\pm$0.85) | 2.26 ($\pm$1.91) |
| Simulation 5 | 1.82 ($\pm$0.75) | 2.26 ($\pm$1.91) |

**Table 4.9:** The DICE coefficients obtained from the tumour volume when comparing the GT volumes to the model estimated volumes for different simulations

| DICE | mean | max | min | std |
|---|---|---|---|---|
| Simulation 2 | 0.98 | 0.99 | 0.93 | 0.01 |
| Simulation 3 | 0.87 | 0.98 | 0.66 | 0.06 |
| Simulation 4 | 0.95 | 0.99 | 0.85 | 0.02 |
| Simulation 5 | 0.92 | 0.98 | 0.72 | 0.05 |

**Table 4.10:** The COMD (mm) for the case of motion model and when no motion model was applied for simulations 2-5.

| COMD (mm) | Motion model | No Motion model |
|---|---|---|
| Simulation 2 | 0.32($\pm$0.27) | 1.21($\pm$1.01) |
| Simulation 3 | 1.28($\pm$0.93) | 4.75($\pm$4.05) |
| Simulation 4 | 0.60($\pm$0.44) | 1.53($\pm$1.58) |
| Simulation 5 | 1.67($\pm$1.23) | 4.56 ($\pm$4.17) |

The mean DFE for all simulations was $<$2mm with simulation 2 having the highest mean DFE of 1.97mm. Simulations 2 and 3 had the same underlying GT motion, with the exception of different tumour sizes and the tumour being in a different location in the lung, in addition the FOV of the projection acquisition for simulation 3 has been lowered slightly since the tumour was in the lower part of the

lung, these could explain the slight difference in DFE obtained between simulations 2 and 3. The mean DFEs obtained for simulations 4 and 5 were identical and both simulations had the same underlying motion, even though the acquisition FOV has been adjusted slightly to account for the fact that the tumour was in the lower region of the lung for simulation 5, and the tumour in both simulations 4 and 5 had different volumes and in different region of the lungs. In addition, the DFE values obtained when the model was applied to simulations 2-5 were all lower than when no motion model was applied. In all cases the COMDs were lower when a motion model was applied. The relatively lower COMD value obtained for simulations 2 and 4 for the motion model and no motion model cases were due to the position of the tumour being in the upper lobe of the lung where there is less respiratory motion.

Figure 4.10 shows colour overlays of EE and EI images as estimated by the model compared to the corresponding GT volume for simulation 2, also shown is the relative tumour centroid position in the AP and SI directions for the GT and the model estimated tumour volumes. Similar illustration for simulations 3, 4 and 5 are shown in figures 4.11, 4.12 and 4.13 respectively. The corresponding movie for these images over all time-points comparing the GT motion and the model estimated motion can be found in the attached supplementary material. (***Simulation_2.mp4*** - *Shows the movie of the visual comparison between the GT and the model estimate for simulation 2. Shown are the coronal and sagittal slices for the GT and the model estimate, and the colour overlay between the two is shown in the middle. In general, the motion appears to have been modelled reasonably well, with the diaphragm and other lung structures moving up and down with the breathing. However, discrepan-*

*cies can be seen in particular for deep breaths in the lower part of the lung where more motion is observed. Similar finding were observed for* ***Simulation_3.mp4***, ***Simulation_4.mp4***, *and* ***Simulation_5.mp4***)

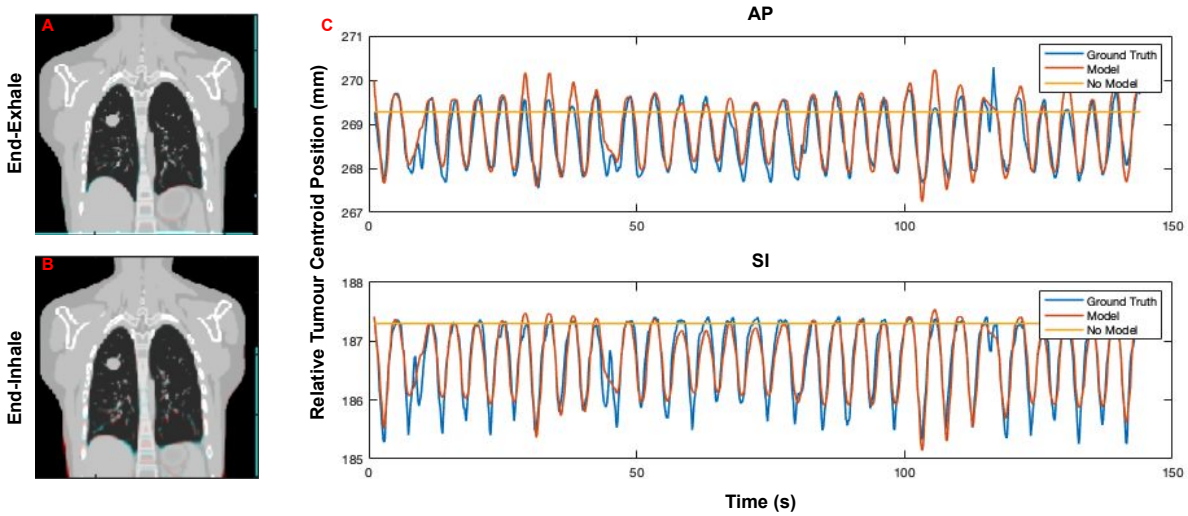


**Figure 4.10:** (*A and B*) shows the colour overlay of the estimated and the GT volume at EE and EI for simulation 2. (*C*) comparison of the tumour centroid position between the model and the case where no motion model was applied relative to the GT motion for the AP and SI directions respectively

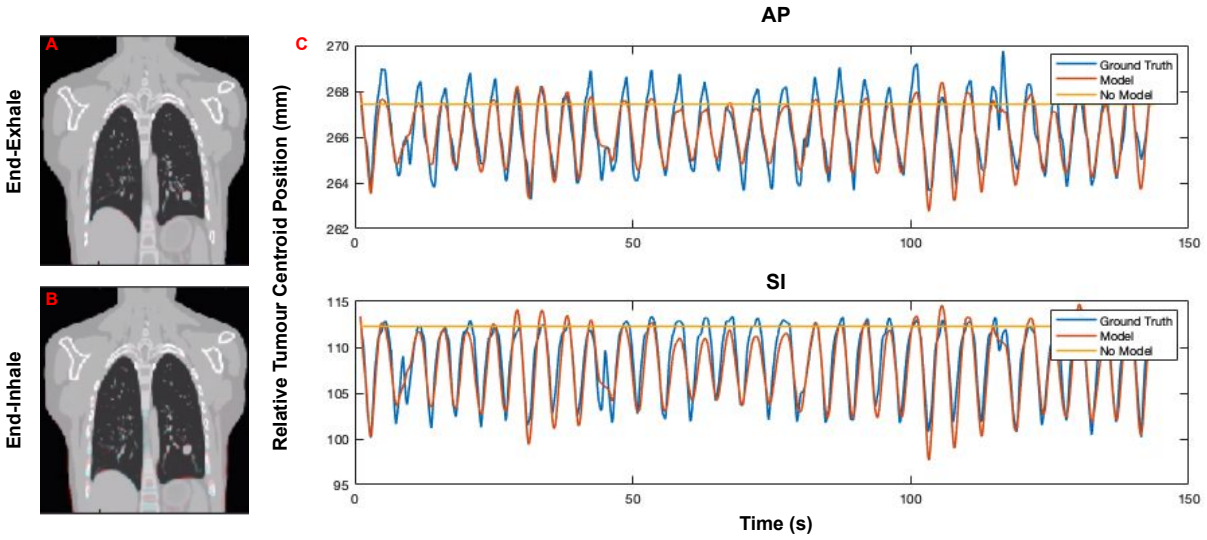


**Figure 4.11:** (*A and B*) shows the colour overlay of the estimated and the GT volume at EE and EI for simulation 3. (*C*) comparison of the tumour centroid position between the model and the case where no motion model was applied relative to the GT motion for the AP and SI directions respectively

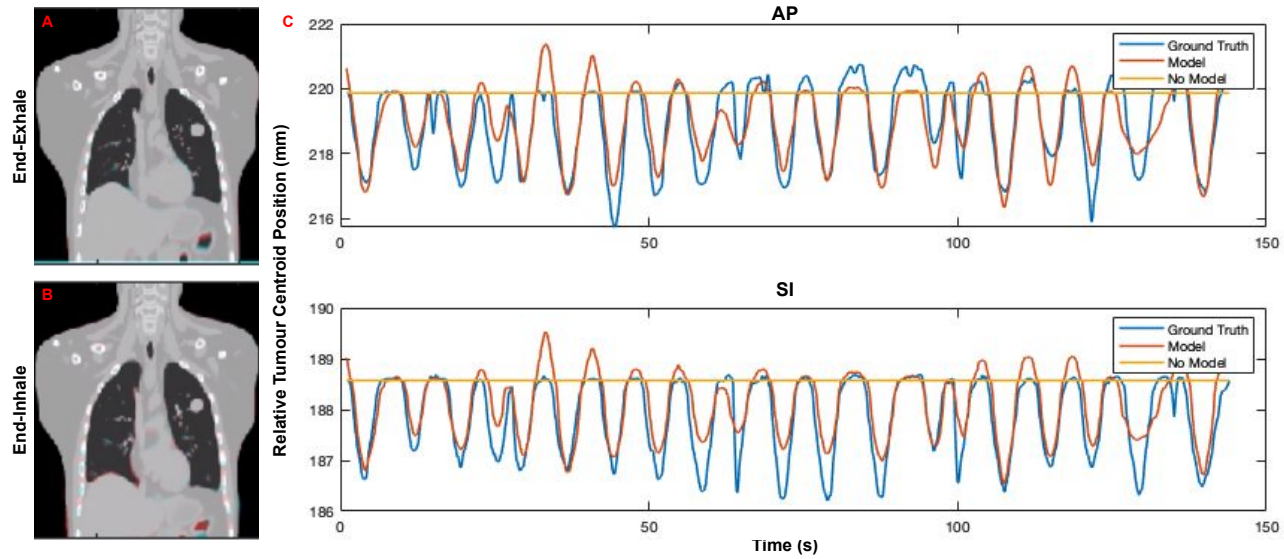


**Figure 4.12:** (*A and B*) shows the colour overlay of the estimated and the GT volume at EE and EI for simulation 4. (*C*) comparison of the tumour centroid position between the model and the case where no motion model was applied relative to the GT motion for the AP and SI directions respectively

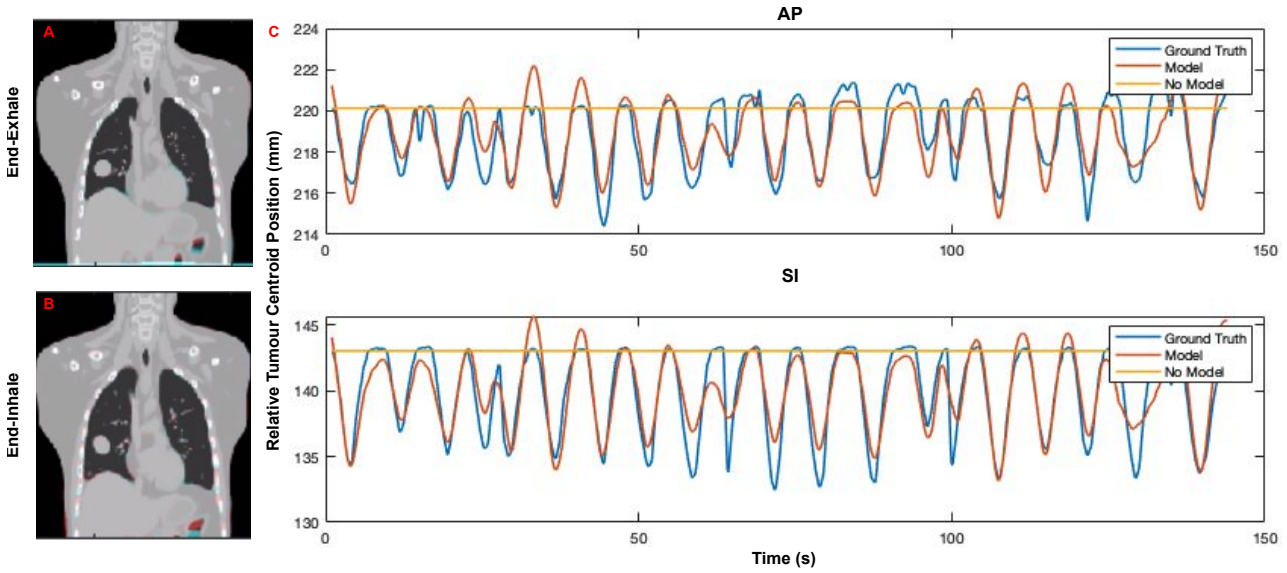


**Figure 4.13:** (*A and B*) shows the colour overlay of the estimated and the GT volume at EE and EI for simulation 5. (*C*) comparison of the tumour centroid position between the model and the case where no motion model was applied relative to the GT motion for the AP and SI directions respectively

### 4.4.3 Image Quality Assessment

Quantitative results of the reconstructed image quality are shown in table 4.11. The

$I_{staticRecon}$ volume was compared to the GT-MCIR volume, MCIR volume, the re-

construction with no motion model applied, and the 4D-CBCT volume respectively. The evaluation was performed using RMSE and PSNR. Relative to the $I_{staticRecon}$ reconstruction, for all the simulations the RMSE was low for the GT-MCIR volume, and a gradual increase in RMSE value was observed for the MCIR from the proposed method, followed by the no motion model reconstruction, and then finally the 4D-CBCT. The PSNR was highest for the GT-MCIR volumes when compared to the $I_{staticRecon}$, while a gradual reduction in PSNR was observed for the MCIR, no motion model volume and the 4D-CBCT volume respectively. These results were as expected. Since the GT-MCIR represents the ideal motion estimate, the results are expected to match the $I_{staticRecon}$ relative to the other methods, nevertheless the results of the MCIR were also comparable to that of the GT-MCIR, which suggests that the mode used to perform the MCIR was sufficient to compensate for the majority of the motion in the reconstruction. In addition, the MCIR results have accounted for most of the impact of the motion seen in the case of no motion model. The PSNR was higher for the no motion model case when compared to the 4D-CBCT, this could potentially be due to aliasing artifacts observed in the 4D-CBCT reconstructions. Qualitative assessment of the image quality also supports the quantitative results.

Qualitative image quality assessment for simulation 2 is shown in figure 4.14. Coronal slice comparisons between the $I_{staticRecon}$ reconstruction, the GT-MCIR, the MCIR, the no motion model and the 4D-CBCT are shown in figure 4.14A, while the corresponding images for the sagittal slice are shown in figure 4.14B. The black arrow illustrates a region of visual discrepancies between the reconstructed images.

Figure 4.14C shows the line profile through the diaphragm based on the red line shown in figure 4.14A, while figure 4.14D shows the line profile through the tumour as illustrated by the red line in figure 4.14B. Similar results for simulations 3-5 are illustrated in figures 4.15, 4.16 and 4.17 respectively. Overall the visual observation showed improvement in image quality of the structures within the lungs for the GT-MCIR and the MCIR respectively, these improvements could not be observed for the no motion model and the 4D-CBCT in most cases where motion blur or aliasing artifacts corrupt the images, similarly the diaphragm appears sharper for the GT-MCIR and the MCIR relative to the no motion model reconstructions and the 4D-CBCT images. The line profiles of the diaphragm and the tumour also confirmed the visual observation, with the line profile of the $I_{staticRecon}$, GT-MCIR, and the MCIR showing similar sharpness at the diaphragm boundaries.

**Table 4.11:** The quantitative analysis obtained for the simulated data for different reconstructions when compared to the static reconstruction volume.

| | Metrics | GT-MCIR | MCIR | No Motion Model | 4D-CBCT |
|---|---|---|---|---|---|
| Simulation 2 | RMSE | $6.9 \times 10^{-4}$ | $7.7 \times 10^{-4}$ | $1.0 \times 10^{-3}$ | $1.7 \times 10^{-3}$ |
| | PSNR | 63.20 | 62.24 | 59.67 | 55.13 |
| Simulation 3 | RMSE | $6.0 \times 10^{-4}$ | $7.0 \times 10^{-4}$ | $1.2 \times 10^{-3}$ | $2.4 \times 10^{-3}$ |
| | PSNR | 63.24 | 62.95 | 57.82 | 52.24 |
| Simulation 4 | RMSE | $5.2 \times 10^{-4}$ | $5.4 \times 10^{-4}$ | $7.7 \times 10^{-4}$ | $9.2 \times 10^{-3}$ |
| | PSNR | 65.72 | 65.39 | 62.32 | 40.67 |
| Simulation 5 | RMSE | $3.9 \times 10^{-4}$ | $5.3 \times 10^{-4}$ | $7.6 \times 10^{-4}$ | $9.2 \times 10^{-4}$ |
| | PSNR | 68.03 | 65.47 | 62.29 | 40.71 |

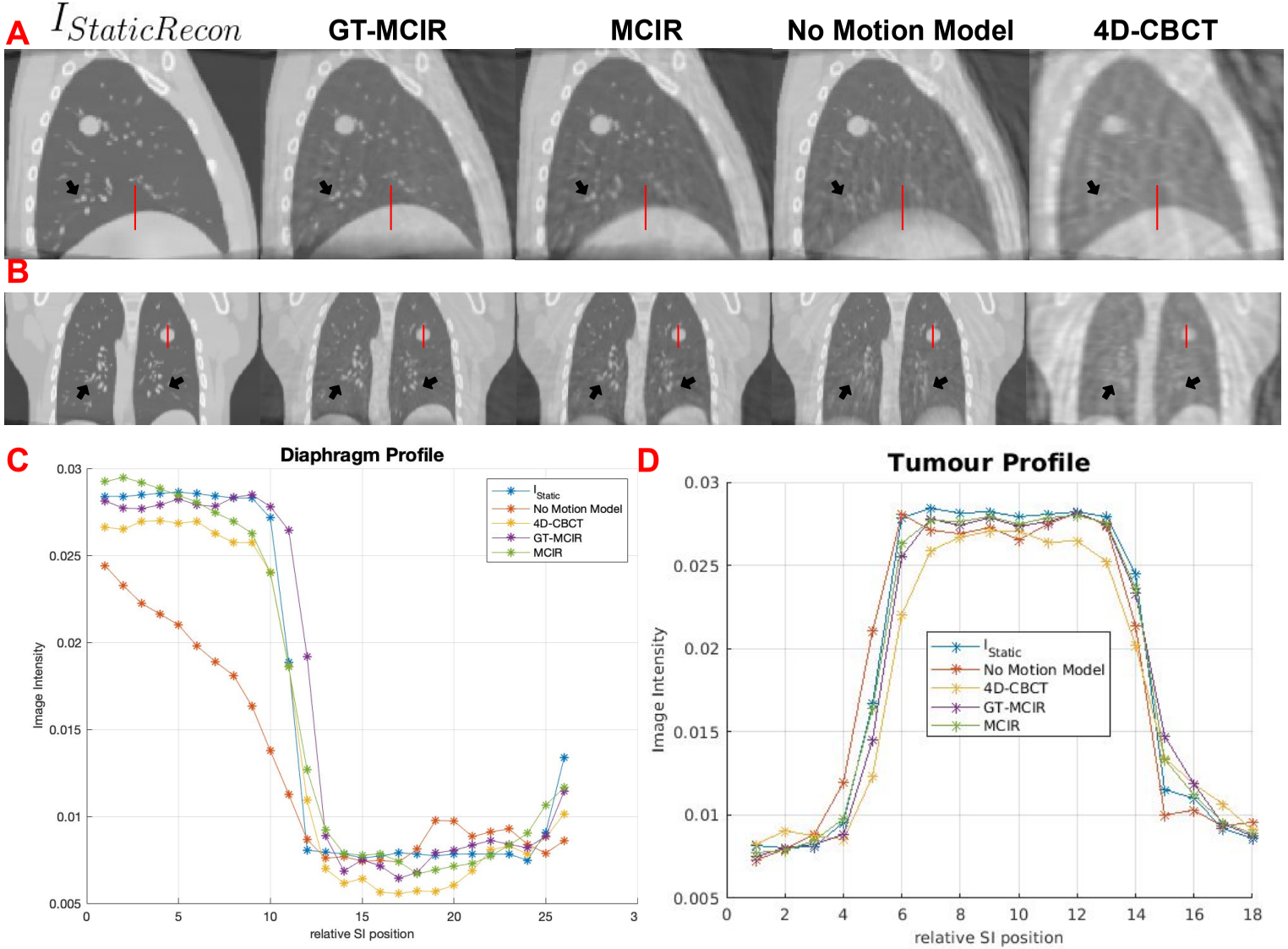


**Figure 4.14:** Comparison of the reconstructed images for simulation 2. *A and B* shows the coronal and sagittal slices of the reconstructions from the $I_{staticRecon}$ image, the GT-MCIR, the MCIR, the no motion model reconstruction and a 4D-CBCT reconstruction. The black arrow shows a region of discrepancies between the images. *C and D* shows the line profile through the diaphragm and tumour region as illustrated by the red lines in *A and B*

.

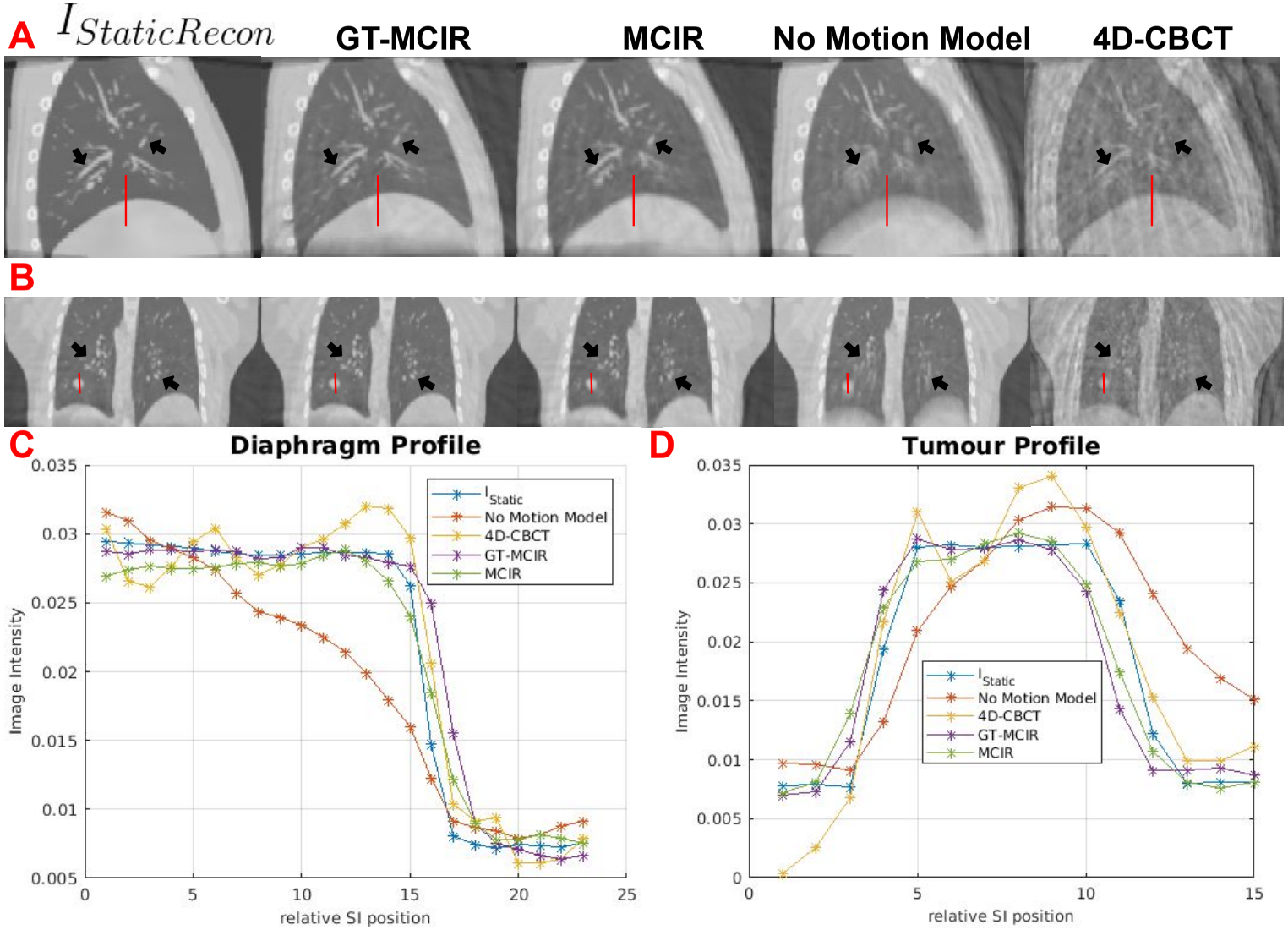


**Figure 4.15:** Comparison of the reconstructed images for simulation 3. *A and B* shows the coronal and sagittal slices of the reconstructions from the $I_{staticRecon}$ image, the GT-MCIR, the MCIR, the no motion model reconstruction and a 4D-CBCT reconstruction. The black arrow shows a region of discrepancies between the images. *C and D* shows the line profile through the diaphragm and tumour region as illustrated by the red lines in *A and B*

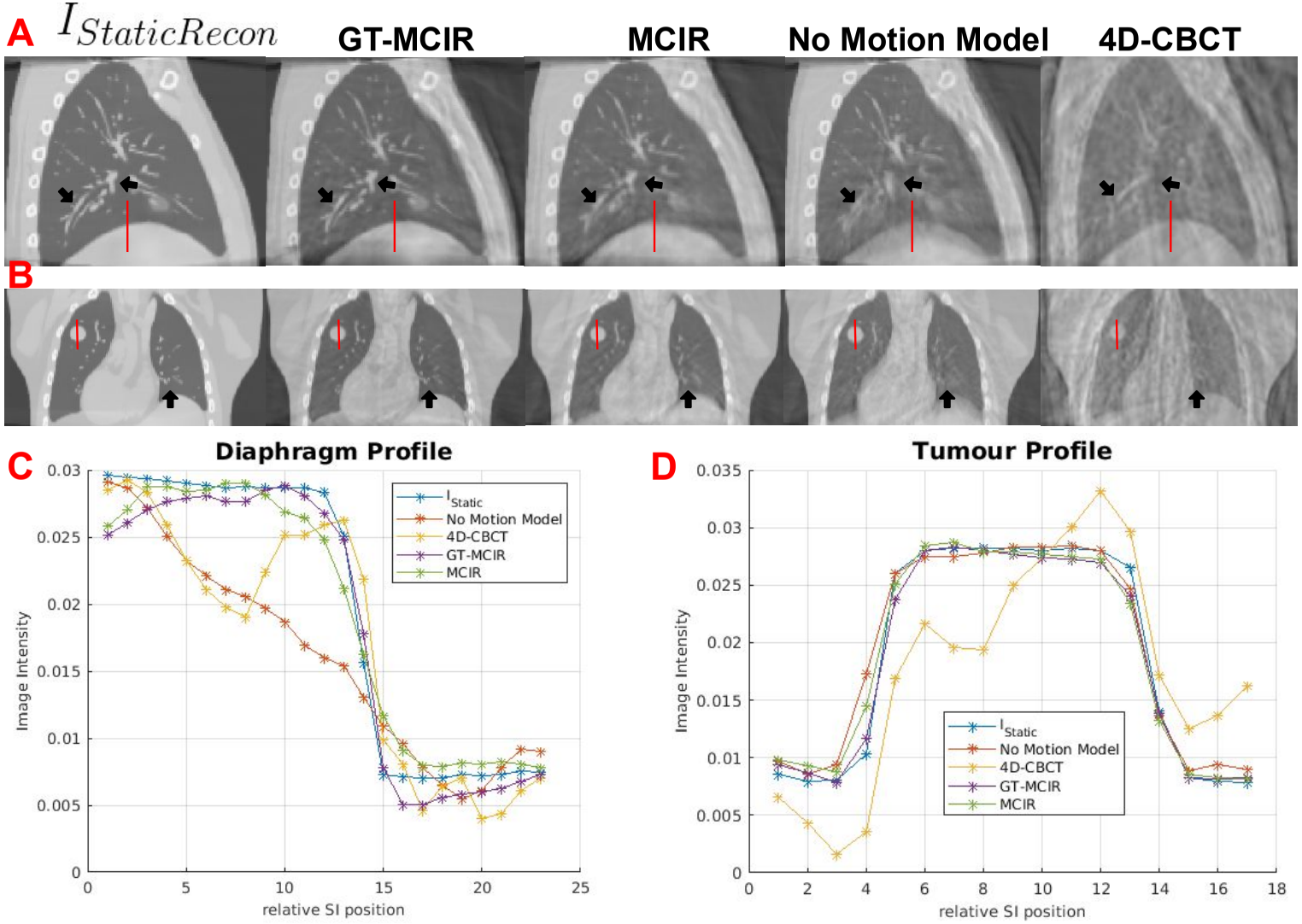


**Figure 4.16:** Comparison of the reconstructed images for simulation 4. *A and B* shows the coronal and sagittal slices of the reconstructions from the $I_{staticRecon}$ image, the GT-MCIR, the MCIR, the no motion model reconstruction and a 4D-CBCT reconstruction. The black arrow shows a region of discrepancies between the images. *C and D* shows the line profile through the diaphragm and tumour region as illustrated by the red lines in *A and B*

.

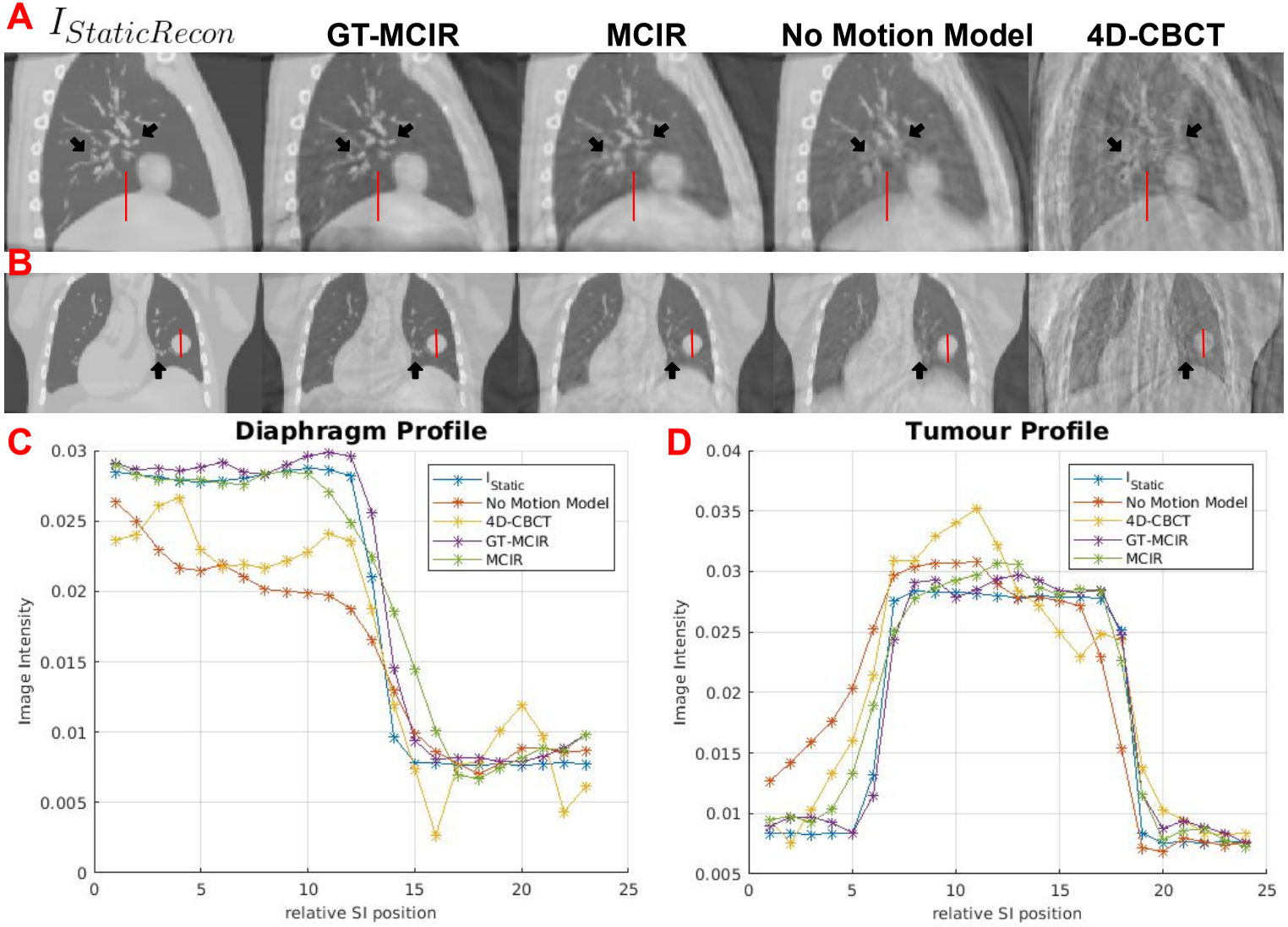


**Figure 4.17:** Comparison of the reconstructed images for simulation 5. *A and B* shows the coronal and sagittal slices of the reconstructions from the $I_{staticRecon}$ image, the GT-MCIR, the MCIR, the no motion model reconstruction and a 4D-CBCT reconstruction. The black arrow shows a region of discrepancies between the images. *C and D* shows the line profile through the diaphragm and tumour region as illustrated by the red lines in *A and B*

## 4.5 Discussion

The optimisation of the hyper-parameters of the model was performed in order to find the optimum values that maintain for computational efficiency while also producing an adequate motion estimate. The optimum parameters were found to be (CPG spacing = 8 voxels, number of spatial multi-resolution levels = 3, maximum number of iterations to fit the model = 50, number of projections required to fit the model = 1/4). The criteria for choosing the hyper-parameters were based on the DFE when the GT motion was compared to the model estimated motion and the computational time. The results of the DFEs for all the hyper-parameters tested are shown in figure 4.4. Overall the mean DFEs obtained based on the optimised hyper parameters were less then when no motion model was applied.

The GT signals were used for the hyper-parameter optimisation because these signals would give the optimum results that represent the actual motion observed in the simulations. However, this was not feasible in a clinical setting since these signals are not available. Thus, we investigated the use of the D-ROI and T-ROI signals developed in chapter 3, and the results were compared to the GT signal. The DFEs between the GT motion and the model estimated motion when the individual surrogate signals were used are shown in table 4.6. As expected using the GT signal yielded the lowest mean DFE, this is probably due to the GT surrogate signal being a closer representation of the GT motion relative to the derived D-ROI and T-ROI signals. In addition, the GT signal used to build the model consist of 2 independent signals (*i.e.* diaphragm and skin-AP), whereas for the D-ROI and T-ROI the signal and their temporal derivative was used. The mean DFE obtained for the D-ROI

was lower than that obtained for the T-ROI. A reason for this could be down to the method derived in chapter 3. In this case the method was able to identify the moving structure of the diaphragm better due to its higher contrast region compared to that of the tumour region. Nevertheless, the results obtained using the three different surrogate signals were less than when no motion model was applied. Even though the T-ROI resulted in the highest DFE, the result is still plausible, which suggests it can be used to obtain an adequate motion estimate. Overall the results suggest that either the D-ROI or the T-ROI can be used.

Reduction in motion blur (diaphragm, tumour in lower lobe and blood vessels ) can be visibly seen in all the images for our proposed MCIR technique relative to the current standards such as the no motion reconstructed images and the 4D-CBCT reconstruction based on the results shown in figures 4.14, 4.15, 4.16 and 4.17 respectively.

Quantitative results also support the visual assessments, as shown in table 4.11. For all simulations the RMSE values were lower when the $I_{StaticRecon}$ volume was compared with the MCIR from our proposed method, relative to the no motion model reconstruction and a 4D-CBCT volume. Higher PSNR was obtained for our MCIR volumes relative to the no motion model reconstruction and 4D-CBCT volume. These results suggest that the obtained MCIR through our proposed method was able to produce better reconstructed images in terms of the image quality when compared to the standard approaches (no motion reconstruction and 4D-CBCT). Furthermore, the results obtained for the GT-MCIR represent the "best case scenario" whereby the GT motion was used in performing the MCIR. The GT-MCIR

serves as an indication of what the MCIR via our proposed method should look like if we could perfectly recover the motion. As expected in all cases the GT-MCIR outperformed our proposed MCIR and the other methods, however quantitative results obtained for the GT-MCIR volumes and the MCIR volumes were relatively similar, visual assessments also supports this with the diaphragm, blood vessels and tumour volume appearing significantly less blurred when compared to the no motion reconstruction and the 4D-CBCT images. These results suggest that the proposed approach was able to produce adequate MCIR similar to the "best case scenario".

The intensity profiles through the diaphragm and the tumour also support the visual observations for all simulations. The line profile of the diaphragm for all simulations showed relatively sharp edges for the $I_{StaticRecon}$, the GT-MCIR, the MCIR and in some cases for the 4D-CBCT. However, the profile from the diaphragm in the case of no motion model reconstruction were not as sharp, mainly due to motion blur artifact, similar observations apply to the line profile through the tumour region, with the exception of tumours in the upper lobe where less respiratory induced motion occurs.

The 4D-CBCT reconstructions showed adequate sharpness along the edges of the diaphragm region for simulations 2-5 when compared to the $I_{StaticRecon}$ as illustrated in figures 4.14, 4.15, 4.16 and 4.17, nevertheless prominent view aliasing artifacts can be observed in the images. The quantitative analysis also supports this observation with the PSNR of the 4D-CBCT relative to the $I_{StaticRecon}$ reconstruction being lower for all simulations when compared to the GT-MCIR, MCIR or the no motion model reconstruction.

We have used the XCAT phantom in this work to evaluate the accuracy of the motion estimated since the GT motion is known, however there are some associated limitations to using a digital phantom study. For instance the tissues in the phantom are modelled homogeneously (the same CT number is assigned to the same organ), which is not necessarily the case for patients [111]. Also the simulated projections using openRTK are different to real projection data. The current implementation of openRTK does not model artifacts such as beam hardening and scatter.

In this study we have used 4 simulations with 2 different GT motions to evaluate the proposed method. More simulations with more variable GT motion should be performed in order to further assess the accuracy of the estimated motion from the model. Ideally more simulations are required for Varian OBI and true-beam data respectively. A subset of these data (giving good representation of different types of data) should be used to tune the hyper-parameters and then more thorough validation (as performed for simulations 2-5) should be conducted using all data. Nevertheless, the initial results presented in this chapter are promising and show the approach can produce good results in a clinically realistic setting. Rather than spending more time evaluating the approach from this chapter on more datasets, instead it was decided to investigate modifications to the method so that it can be suitable for real patient data (see chapter 5).

The purpose of the method derived in chapter 3 was to develop a method of extracting surrogate signals that closely relate to the internal motion of the anatomy. The results from using the signals from the T-ROI and the D-ROI suggest that these signals are adequate to be used to drive a respiratory motion model. Though we

have used a data driven approach here to derive the surrogate signal, the use of other methods such as the RPM signal should also be investigated, in order to see how the resulting motion estimation performs.

It is important to note that there is a trade-off between computational time and accuracy. Reducing the amount of data (1/4) used to fit the model also reduced the computational time drastically while giving an adequate estimate of the motion when compared to using all of the data to fit the model. Even though the optimum maximum fit iterations was found to be 50, this does not necessarily indicate that this number of iterations would produce the best motion estimate, however partly due to the computational efficiency this value was chosen, and the results using only 50 maximum fit iterations generated adequate motion estimates for all experiments, thus suggesting that the value was sufficient. The average computation time for simulations 2-5 was approximately 2hrs, and this is not clinically feasible. Computational time could be improved upon by using a more advanced implementation of the motion modelling algorithm and dedicated hardware such as a GPU.

The aim of the initial experiments was not necessarily to find the best parameters, but rather sufficient parameters that gave suitable results. More experiments with more parameter values (and more data) are needed to really determine the optimum parameters. In addition, this should also be done with real data, since the optimum parameters for real data may be different. However this is difficult as the true motion is not known.

## 4.6 Conclusion

The proposed method was tested on simulated data-set with known GT. For all the simulated studies we were able to recover the motion and they were assessed both quantitatively and qualitatively. In addition, the MCIR volumes obtained based on the estimated motion showed an improvement in image quality relative to when no motion model was applied (*i.e* standard CBCT reconstruction technique and 4D-CBCT phase binning reconstruction) both quantitatively and qualitatively. So far the method has been applied to simulated data-sets where a reference static image is required. The next chapter details how we have expanded on this study to be applicable for clinical data.

Chapter 5

# Iterative respiratory motion estimation and motion compensated image reconstruction

## 5.1 Aim of this chapter

The method described in chapter 4 is not easily applicable to clinical data. The method requires a static volume. There are two main potential approaches to overcome this: either obtaining a 4D-CT image or from the CBCT data itself.

In the first approach, one of the phases of a 4D-CT, (e.g used for planning) could be used as a reference volume for the motion model fitting. The 4D-CT will provide a relatively high-quality reference static image for the motion model, and in addition the 4D-CT data would contain all the anatomy in the FOV which is required when simulating projection through 4D-CT image. However, this approach has its associated disadvantages. Firstly, anatomical changes can occur during radiotherapy treatment, such as previously blocked regions of the lung re-inflating, which will be difficult to account for by deforming the existing CT. Secondly, due to differences in imaging devices, simulated projections through an existing CT will appear different to the real CBCT projections due to scatter and beam hardening [41].

Another approach is to use a MCIR volume from the CBCT as the static image. However, MCIR requires motion estimates. This chapter proposes an iterative approach for performing the respiratory correspondence model fitting and MCIR in an iterative manner (termed MM-MCIR). The method described here is only applicable to non-truncated data-sets such as the ones acquired on the true-beam system at our clinical institution with the full anatomy in the reconstruction FOV. The reason for this will be discussed in section 5.5. We first tested this approach on simulated data, then the method was applied to clinically acquired data-sets.

## 5.2 Methodology

In this section we give a brief description of our approach. The correspondence model fitting approach used here was initiated by first performing a standard FDK reconstruction with an initial estimate of no motion. The result will contain blurring and/or other motion artifacts, but this is used as an initial estimate of $I_0$ to fit the model parameters. The fitted model is then used to perform MCIR (see section 4.2), updating the estimate of $I_0$. The process continues to iterate between fitting the model and reconstructing $I_0$ until there is no more improvement or the maximum number of iterations is reached. For each iteration the output motion models and images were saved. The method is run for a fixed number of iterations. Alternatively a cost function (such as used for the model fitting) could be used to determine if there is improvement after each iteration of model fitting and MCIR.

## 5.3 Experiments

We tested the proposed method on the simulated datasets from simulations 2-5 from the previous chapter (see section 4.3.1.1). These simulations were based on true-beam acquisition, thus the full anatomy is in the reconstruction FOV. Since simulation 1 was based on Varian OBI acquisition with limited FOV of the anatomy in the reconstructed volume, this data-set could not be used for the proposed technique. The same hyper-parameters for the motion model estimation were used as in the previous experiments (CPG spacing = 8 voxels, spatial multi-resolution level = 3, numbers of iteration to fit the model = 50, number of projections to fit the model = 1/4).

In addition, we tested our method on two clinical CBCT projection data-sets from a real lung cancer patient. It is the same true-beam data as described in chapter 3. The data-sets were acquired from the same patient on two different days using the true-beam Varian system, and consisted of 894 projections covering 360° acquired at 14.9 fps. The true-beam acquisition was centred on the centre of the patient (and also employed a half-fan scan mode), so included all of the patient's anatomy in the reconstruction. The same hyper-parameters as for the simulated data were used, with the following exceptions:

- all of the available data was used to fit the model
- we fine-tuned the CPG spacing used to fit the model, while keeping the remaining hyper-parameters constant.

The CPG spacing was fine-tuned using the first real data, and the result was then applied to the second real data-set. As in the previous chapter, the D-ROI signal obtained from chapter 3 for the clinical data was used as the surrogate signal to drive the model.

### 5.3.1 Evaluation

For the simulated data-sets, the $I_0$ used to fit the model is in the average position, while the GT XCAT DVFs were obtained relative to a EE volume. Also a constant offset term is not used in the model, as the $I_0$ used here corresponds to the average position. In order to perform a fair comparison between the MM-MCIR and the $I_{staticRecon}$ both volumes need to be in the same position. $I_{staticRecon}$ in the average position was obtained by determining the respiratory motion trajectory based on

the diaphragm and the AP motion signals used to drive the XCAT simulation. The volume at the time point that corresponds to the average position was determined by choosing the volume that corresponds to the (0,0) input signals. This position is highlighted by the red point in the respiratory motion trajectory as shown in figure 5.1. The corresponding volume for this point is also illustrated in figure 5.1 as $I_{staticRecon[223]}$ [1], also shown is the $I_{staticRecon[1]}$ [2] used in the previous chapter and an example of the MCIR obtained when applying the MM-MCIR approach. The blue line illustrates the alignment of the diaphragm position between the three volumes. As expected the MCIR volume and the $I_{staticRecon[223]}$ volume are in a similar position.

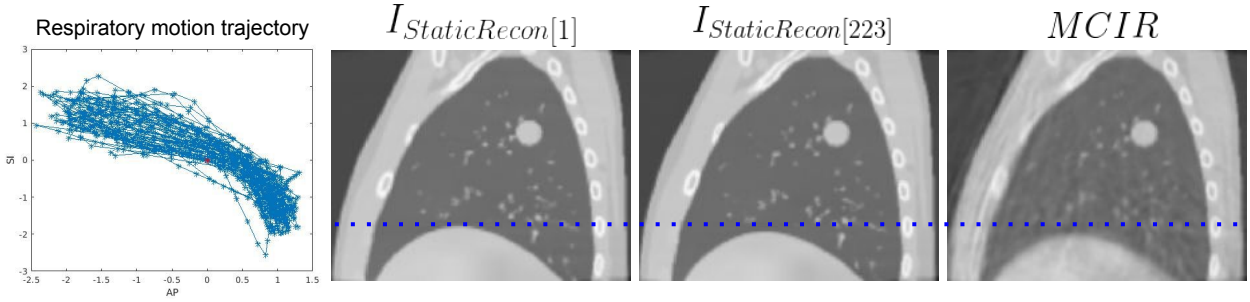


**Figure 5.1:** Illustration of the respiratory motion trajectory from the GT XCAT simulation and the corresponding $I_{staticRecon}$ positions and the MCIR volume obtained via MM-MCIR. The blue line in the volumes illustrates the alignment of the diaphragm position.

The image quality of the MCIR volumes were assessed both qualitatively and quantitatively relative to the $I_{staticRecon}$ volume, and the standard FDK reconstruction with no motion model. The respective volumes were assessed qualitatively by comparing the image quality of the different structures such as the diaphragm, tumour and the different vessels within the lungs based on visual assessment. We also

[1] This volume is the $223^{rd}$ XCAT volume that corresponds to the average position
[2] This volume is the $1^{st}$ XCAT volume that corresponds to the EE position

compared the intensity line profile through the diaphragm and the tumour for each volume. Quantitatively we used the RMSE and the PSNR (as described in section 4.3.2) to assess the quality of the images with respect to the $I_{staticRecon}$ volume.

Since there was no $I_{StaticRecon}$ available for the clinical data, the MM-MCIR volumes were assessed qualitatively relative to the standard FDK reconstruction. We evaluated the image quality of the different structures within the body such as the sharpness of the diaphragm and the vessels within the lungs. A line profile along the diaphragm region for the MCIR volumes and the standard reconstruction were plotted to illustrate the relative sharpness.

## 5.4 Results

### 5.4.1 Simulations

Quantitative results for simulations 2-5 are summarised in table 5.1. Shown are the results for the comparison of the $I_{StaticRecon}$ image to the standard reconstruction (no motion model) image and the MM-MCIR obtained after iterations 1 and 5 respectively. In all cases the RMSE were lower for the MM-MCIR volumes relative to the no motion model volume, which indicated better similarity between the MM-MCIR volumes and the $I_{StaticRecon}$. The largest difference for the RMSE value was obtained for simulation 2, with a value of $5.08 \times 10^{-3}$ for the no motion model, and $4.70 \times 10^{-4}$ for the MM-MCIR after both iterations 1 and 5. In all cases the PSNR obtained were higher for the two MM-MCIR volumes relative to the no motion model volume, again this indicated an improvement in the image quality for the two MM-MCIR volumes. In all cases only relatively small differences can be

observed between the results obtained by MM-MCIR after iterations 1 and 5 respectively. The results of MM-MCIR for simulations 2-5 are shown in figure 5.2. For some of the numerical results shown in table 5.1, iteration 5 was slightly worse than iteration 1. The difference is small, however it is consistent across all simulations. Overall the largest improvement in image quality can be seen after iteration 1, in some cases there was still smaller improvement after iteration 2, however it is possible that further iterations can make the result slightly worse as the method starts to over-fit to artifacts in the MCIRs.

**Table 5.1:** The quantitative analysis obtained for the simulated data. The $I_{StaticRecon}$ was evaluated against the no motion model reconstruction, the first iteration and the fifth iteration of MM-MCIR respectively.

| | Metrics | No Motion Model | Iteration 1 | Iteration 5 |
|---|---|---|---|---|
| Simulation 2 | RMSE | $5.80 \times 10^{-4}$ | $4.70 \times 10^{-4}$ | $4.70 \times 10^{-4}$ |
| | PSNR | 64.73 | 66.40 | 66.48 |
| Simulation 3 | RMSE | $5.78 \times 10^{-4}$ | $5.12 \times 10^{-4}$ | $5.01 \times 10^{-4}$ |
| | PSNR | 64.77 | 65.81 | 65.92 |
| Simulation 4 | RMSE | $6.67 \times 10^{-4}$ | $6.17 \times 10^{-4}$ | $6.42 \times 10^{-4}$ |
| | PSNR | 63.53 | 64.20 | 63.84 |
| Simulation 5 | RMSE | $6.67 \times 10^{-4}$ | $6.24 \times 10^{-4}$ | $6.42 \times 10^{-4}$ |
| | PSNR | 63.52 | 64.08 | 63.84 |

Qualitative evaluation for simulations 3 and 5 are shown in figures 5.3 and 5.4, while the results for simulations 2 and 4 can be found in appendix A. The first column shows a slice through the sagittal plane of the $I_{StaticRecon}$ volume, no motion model reconstruction, and the MM-MCIR volume after iterations 1 and 5 respectively. For each reconstructed volume, the corresponding coronal slices are shown in the second column. The red arrows shown in the images highlight regions of discrepancies between the respective slices through the volume as an indicator of the

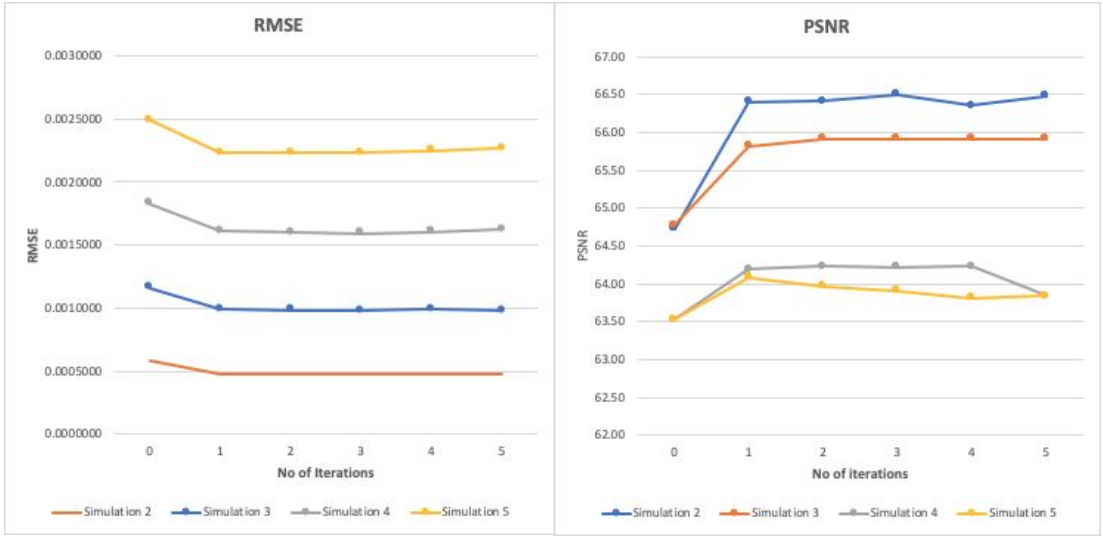


**Figure 5.2:** Results of the RMSE, and PNSR for simulations 2-5 and the different numbers of iterations used.

improvement in image quality. Based on visual assessment, the improvement of the structures indicated by the red arrows can be seen for the MM-MCIR reconstructions relative to the no motion model case when compared to the $I_{StaticRecon}$ image. Less motion blur artifacts can be observed for the MM-MCIR images, in particular in the lower part of the lung where large motion is typically expected. However it is difficult to distinguish any visual improvement between the MM-MCIR images from iterations 1 and 5.

Figures 5.5 and 5.6 for simulations 3 and 5 illustrate the intensity line profiles through the diaphragm and the tumour region respectively. The relative sharpness of the boundaries is an indicator of improved image quality. This can be observed for all cases. The line profiles obtained for the case of no motion model and the MM-MCIR after iterations 1 and 5 were not as sharp as the $I_{StaticRecon}$ profile as expected. Nevertheless, improved sharpness was observed for the MM-MCIR after iterations 1 and 5 when compared to the case of no motion model, thus suggesting

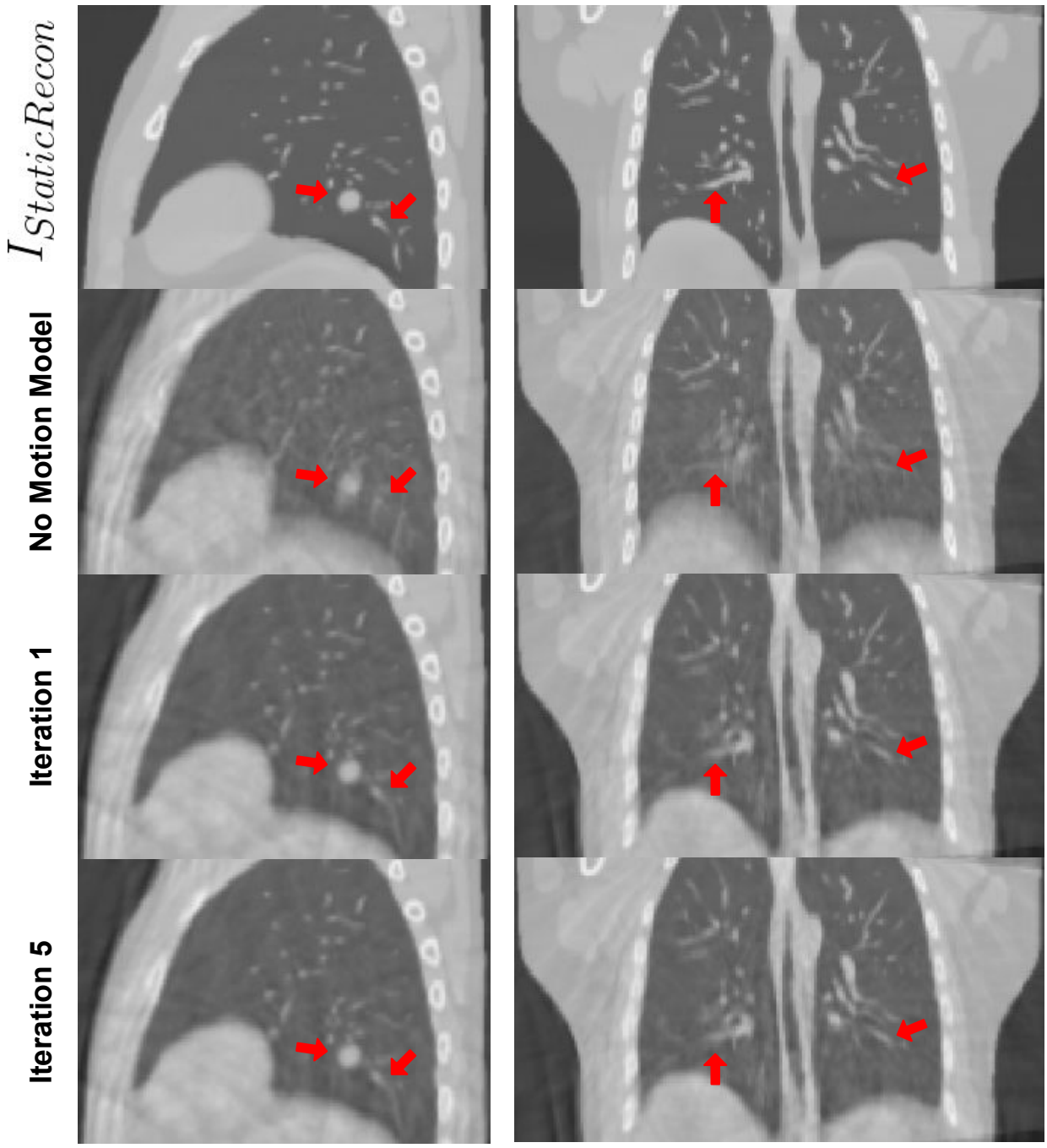


**Figure 5.3:** Qualitative results shown for simulation 3. The first column shows the sagittal slice through the reconstruction images for the $I_{StaticRecon}$, no motion model, the MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations respectively, while the second column shows the coronal slices through the same corresponding volumes. The arrows shown in red highlight regions of discrepancies between the different volumes.

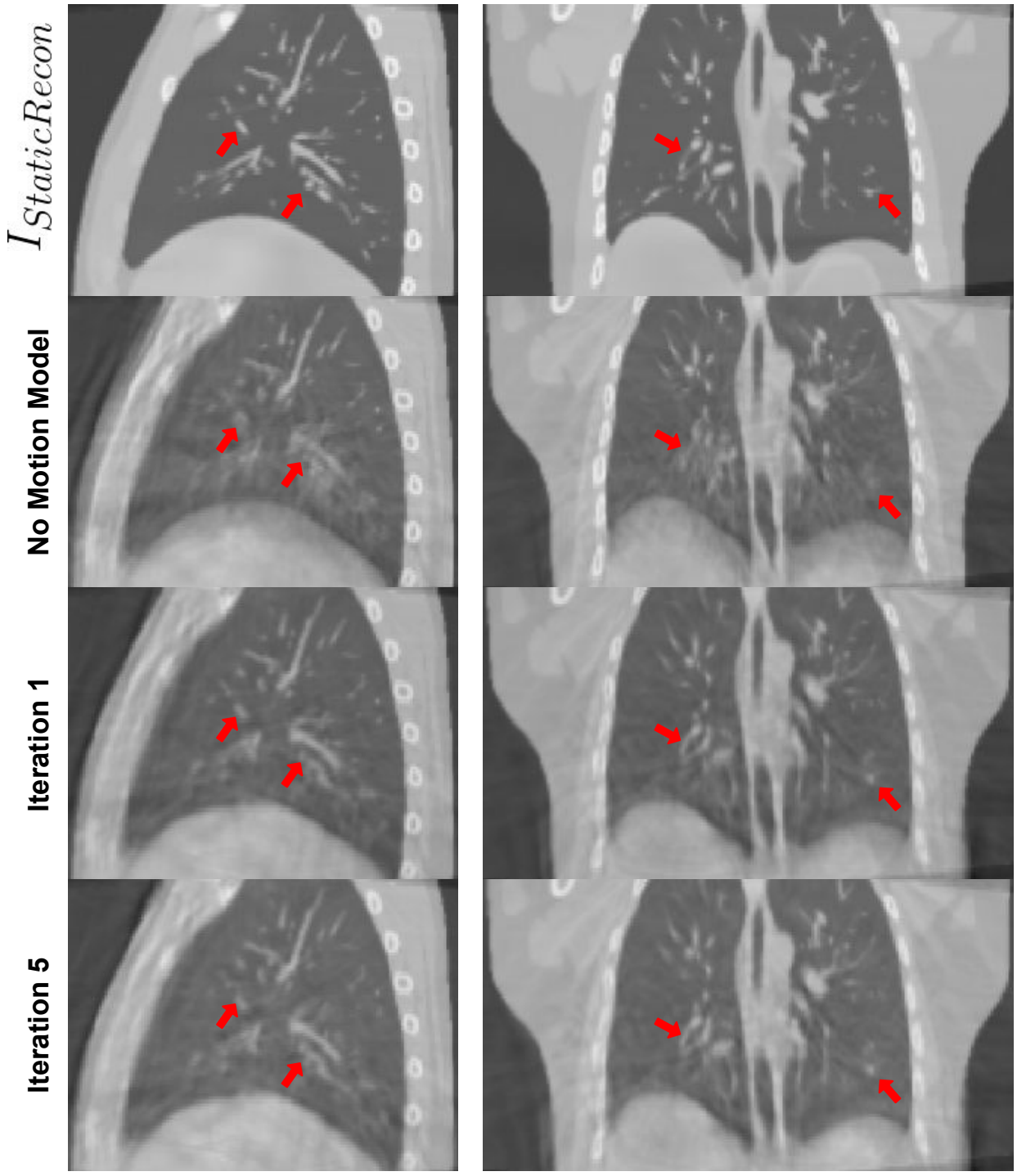


**Figure 5.4:** Qualitative results shown for simulation 5. The first column shows the sagittal slice through the reconstruction images for the $I_{StaticRecon}$, no motion model, the MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations respectively, while the second column shows the coronal slices through the same corresponding volumes. The arrows shown in red highlight regions of discrepancies between the different volumes.

that MM-MCIR performance is better than the standard approach of CBCT reconstruction. The corresponding results for simulations 2 and 4 are shown in appendix A. Overall for the case of the tumour profiles, a relatively small difference was observed between all the reconstructed images, in particularly for simulations 2 and 4 with tumours in the upper lobe of the lung. Nevertheless, for simulations 3 and 5 with tumours in the lower lobe less motion blur of the tumour can be observed.

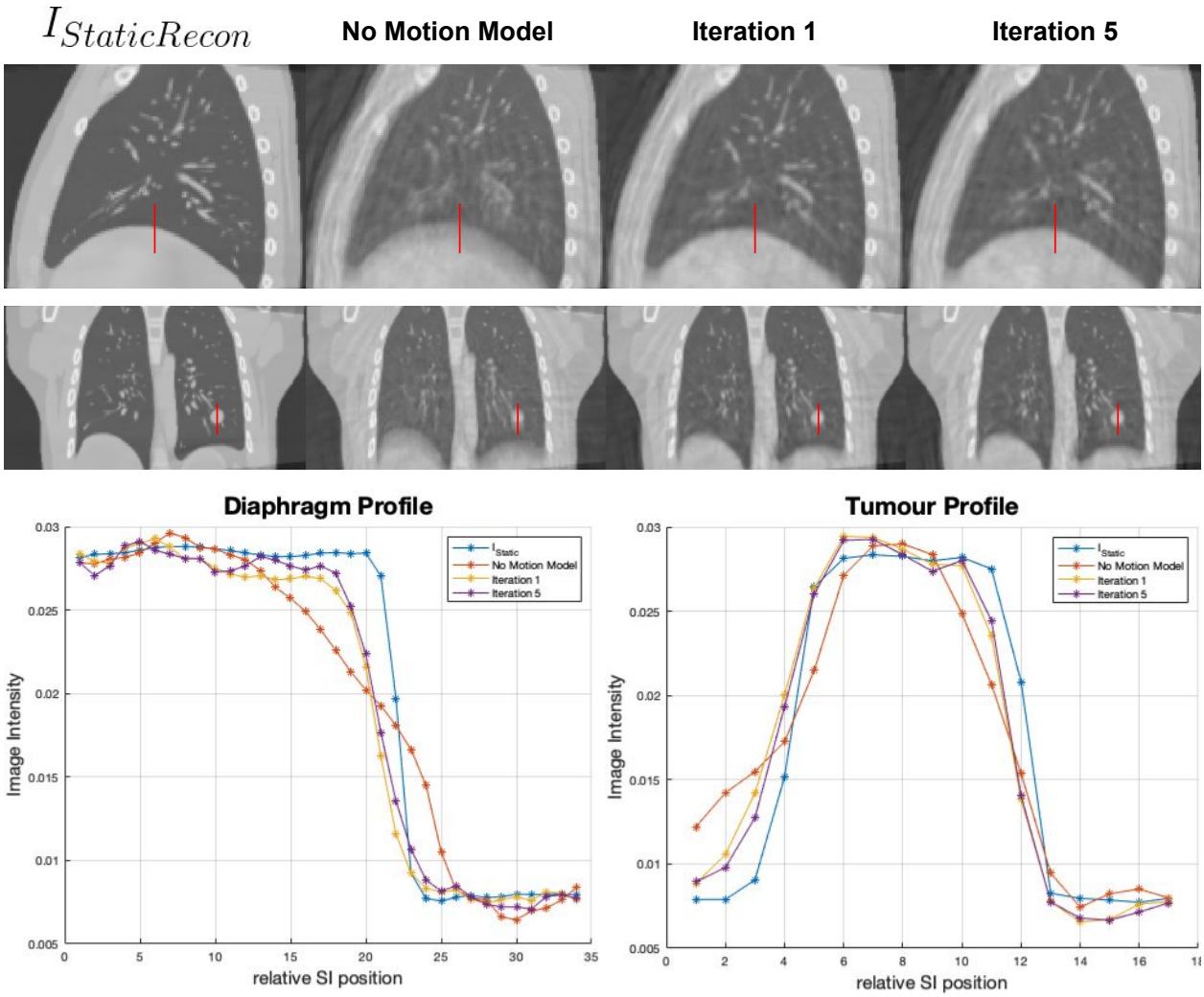


**Figure 5.5:** Results obtained for simulation 3. Shown are the slices through the reconstruction images for the $I_{StaticRecon}$, no motion model, MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations. The coronal slices are shown in the first row, while the corresponding sagittal slices for the volumes are shown in the second row. The diaphragm profile shows the intensity profile through the diaphragm region in the first row as highlighted by the red line, while the tumour profile shows the intensity profile through the tumour region as highlighted by the red line in the second row.

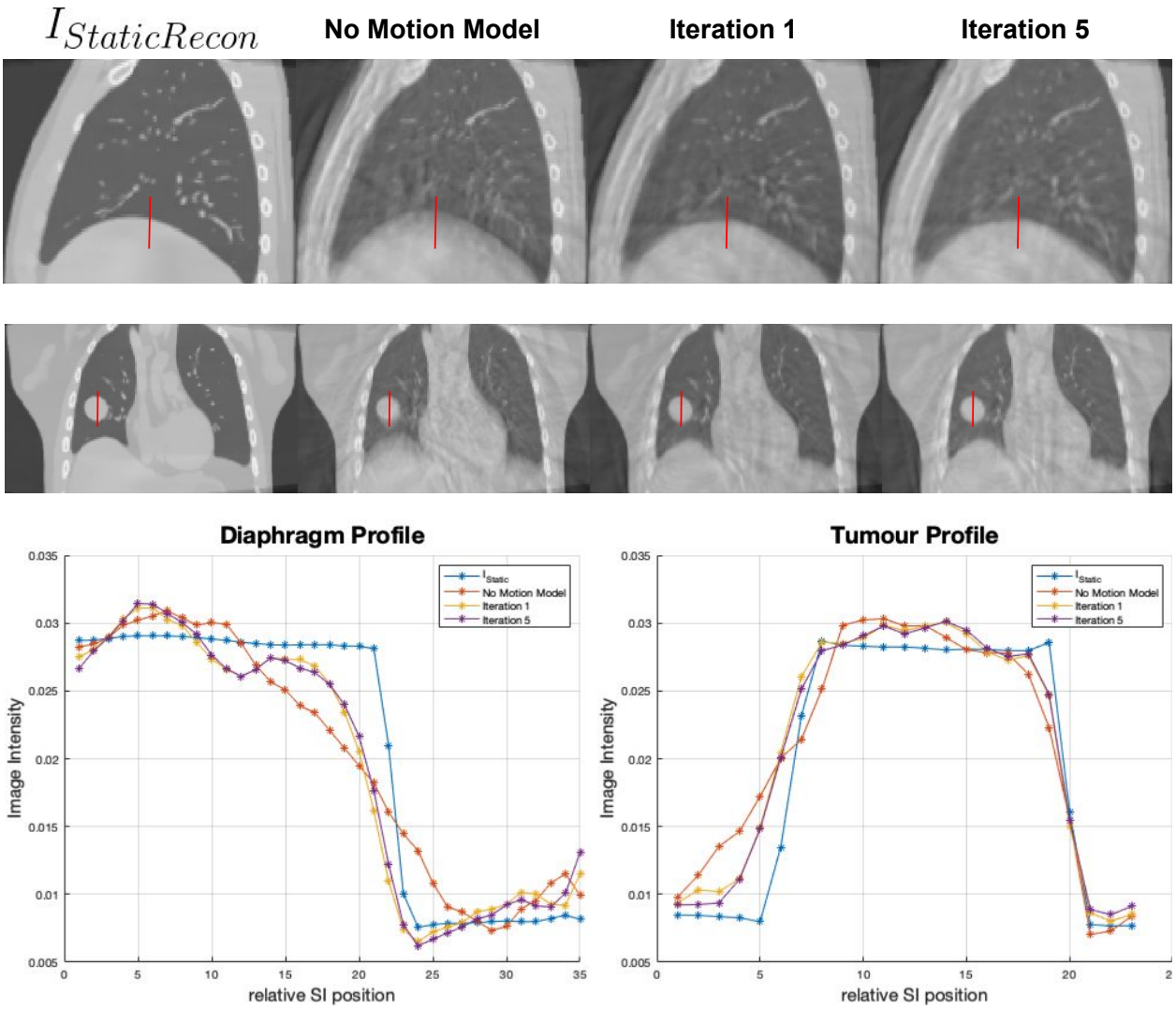


**Figure 5.6:** Results obtained for simulation 5. Shown are the slices through the reconstruction images for the $I_{StaticRecon}$, no motion model, the MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations. The coronal slices are shown in the first row, while the corresponding sagittal slices for the volumes are shown in the second row. The diaphragm profile shows the intensity profile through the diaphragm region in the first row as highlighted by the red line, while the tumour profile shows the intensity profile through the tumour region as highlighted by the red line in the second row.

### 5.4.2 Clinical Data

Figure 5.7 shows the sagittal slice through the MCIR images while varying the CPG spacing (in voxels) for patient 1 acquisition 1. The first row shows the images

obtained from the no motion model reconstructions, while the second and third rows show the images obtained after the first and fifth iterations of MM-MCIR respectively. From the visual observation the images displayed on the first row are identical as expected. However, for the second and third rows, the image quality improved as the CPG spacing was increased. The non-physical artifacts seen in the MM-MCIR images occur as a result of using smaller CPG spacing, however the image quality gradually improves when 20, 30 and 40 (voxels) CPG spacing were used respectively. A similar trend was observed for patient 1 acquisition 2 as illustrated in figure 5.8. The use of a coarser CPG spacing resulted in a smoother deformation, thereby limiting the effect of the non-physical artifacts seen in the images.

Figures 5.9 and 5.10 illustrate the qualitative results obtained from the MM-MCIR for patient 1 acquisition 1 and 2 respectively using a CPG of 30 voxels. Shown are the coronal and sagittal slices for the case of no motion model, MM-MCIR after the first iteration and the MM-MCIR after the fifth iteration. The red arrows highlight regions of discrepancy between the different structures seen in each image, such as the improved sharpness of the diaphragm, and the reduction in blur artifacts observed within the vessels for the MM-MCIR images. Based on visual assessment it is difficult to distinguish any difference/improvement in the image quality between the MM-MCIR for iteration 1 and 5. Also shown are the intensity line profiles through the diaphragm region as highlighted by the red line in the coronal slice for each reconstructed image respectively. The line profile through the diaphragm for the MM-MCIR images has a steeper profile than the

no motion model image indicating improved image quality for the MM-MCIR reconstructions. The corresponding movies for both patient acquisition showing the motion output from the model animated with the reference static image are shown in the attached supplementary material. (*Patient_2A_1_iter_vs_5_iter.mp4* and *Patient_2B_1_iter_vs_5_iter.mp4 - In both cases the first row shows the coronal and sagittal slice of the reference static image animated with the output of the model after 1 iteration. The second row shows the corresponding slices of the reference static image animated with the output of the model after iteration 5. In general the motion of the diaphragm follows a breathing pattern, however unexpected left-right motion of the rib cage (left) can also be observed* )

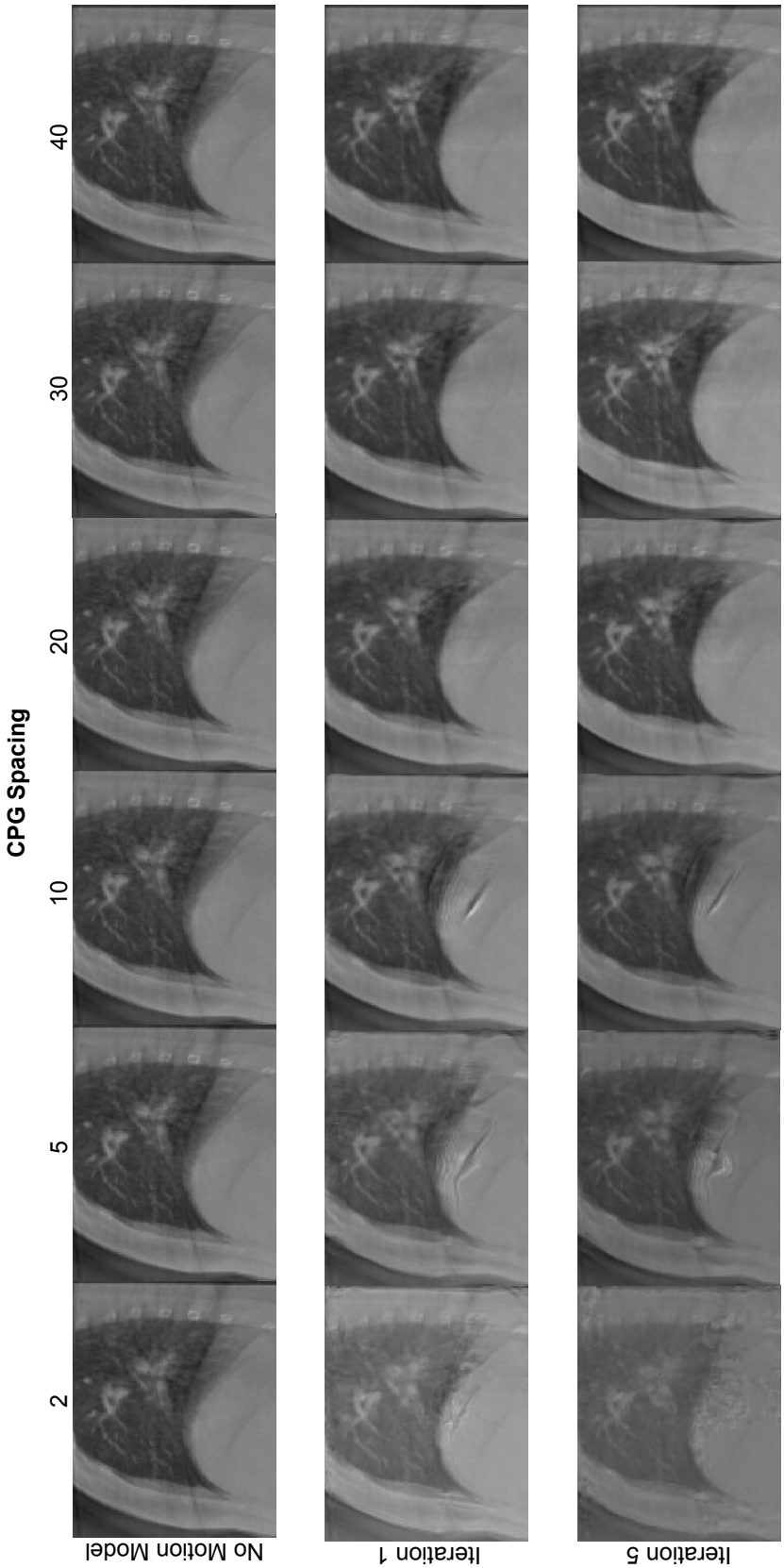


**Figure 5.7:** Shown are the MCIR images obtained with different CPG spacing (voxels) used for the motion model obtained for MM-MCIR iteration 1 and 5 respectively for patient 1 acquisition 1. Also shown is the FDK reconstruction with no motion model used as the reference static image

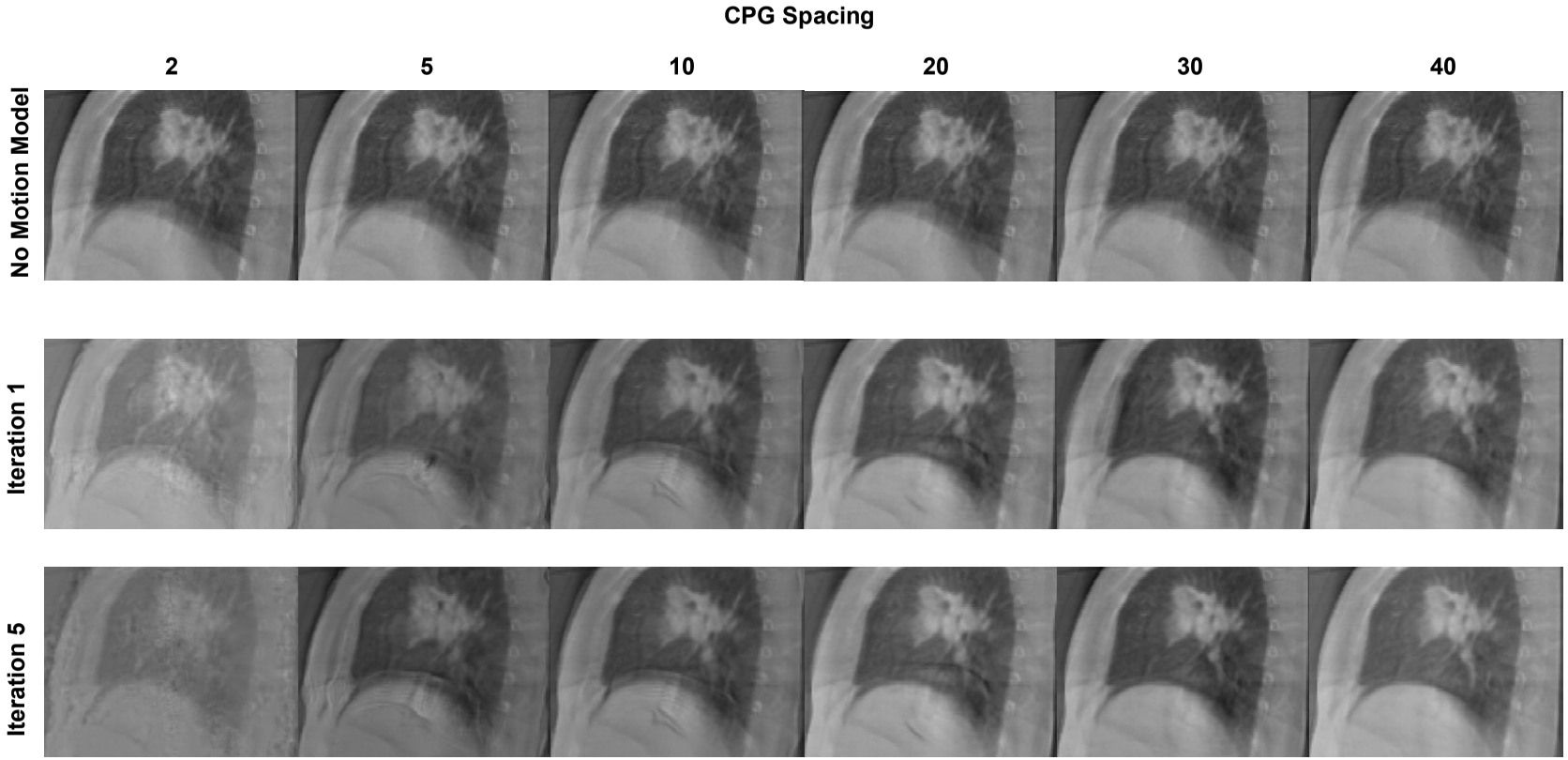


**Figure 5.8:** Shown are the MCIR images obtained with different CPG spacing (voxels) used for the motion model obtained for MM-MCIR iteration 1 and 5 respectively for patient 1 acquisition 2. Also shown is the FDK reconstruction with no motion model used as the reference static image

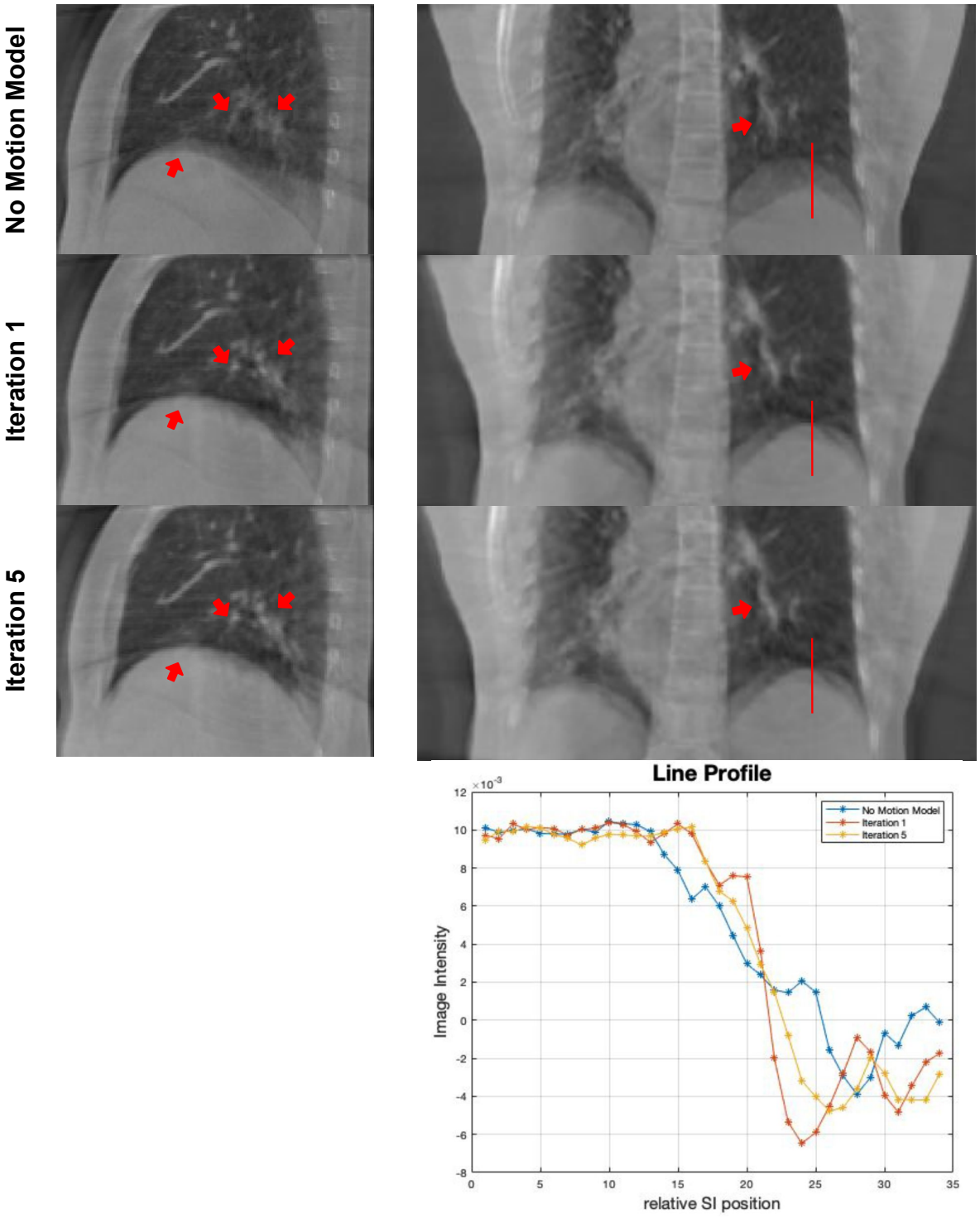


**Figure 5.9:** The image quality assessment between the reconstructions obtained with no motion correction, and MCIR images obtained after iteration 1 and 5 of MM-MCIR respectively for patient 1 acquisition 1. The red arrows highlight regions of difference between the different reconstructed images. The first column shows the sagittal slices for the reconstructed volumes . The second column shows the coronal slices for the reconstructed volumes. The intensity line profile shown in the bottom corresponds to the red line along the diaphragm shown in the second column.

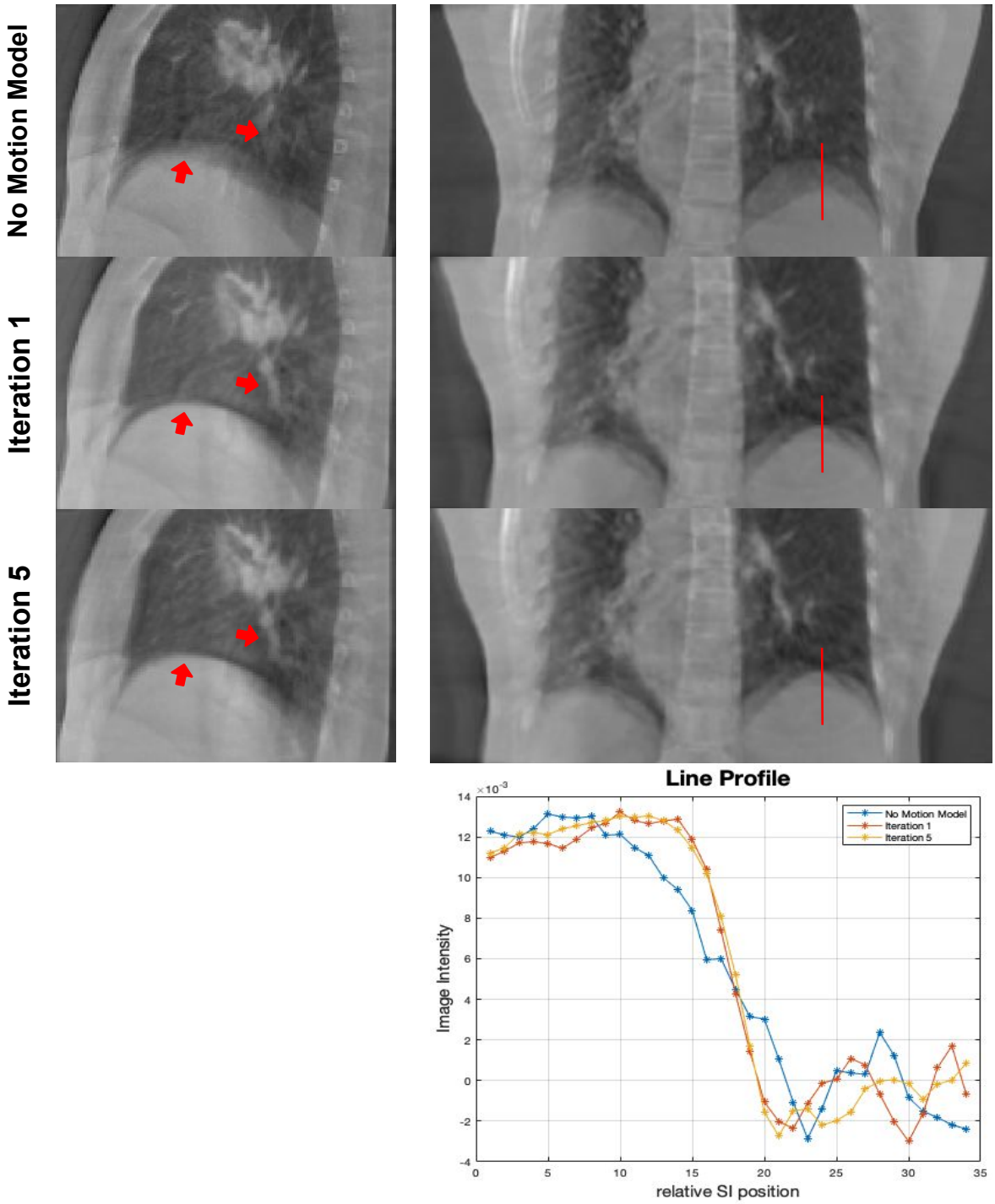


**Figure 5.10:** The image quality assessment between the reconstructions obtained with no motion correction, and MCIR images obtained after iteration 1 and 5 of MM-MCIR respectively for patient 1 acquisition 2. The red arrows highlight regions of difference between the different reconstructed images. The first column shows the sagittal slices for the reconstructed volumes. The second column shows the coronal slices for the reconstructed volumes. The intensity line profile shown in the bottom corresponds to the red line along the diaphragm shown in the second column.

## 5.5 Discussion

Quantitative results for the simulations shown in table 5.1 suggest that there were improvements in the reconstructed images when compared to the case of no motion model. The RMSE values were lower for the MM-MCIR reconstructions relative to the no motion model reconstructions for all cases. The PSNR for the MM-MCIR cases were higher when compared to the no motion model reconstructions for all simulations. The qualitative results also showed improved visual assessments of the reconstructed images for the MM-MCIR images when compared to the no motion model images, as shown in figures 5.3, and 5.4 for simulations 3 and 5 respectively. In both cases, the tumour region appears less blurry in the MM-MCIR images relative to the no motion model case when compared to $I_{StaticRecon}$, also blood vessels and airways appear less blurry and more visible in some cases for the MM-MCIR reconstructions as highlighted by the red arrows. Similar results were obtained for simulations 2 and 5 as shown in figures A.1 and A.3 in appendix A, however less discrepancies in the tumour were observed in these cases as expected, since the tumours were located in the upper lobe of the lung where less respiratory motion is observed. Similarly the intensity line profile comparison between the $I_{StaticRecon}$, the no motion model reconstruction and the MM-MCIR reconstructions also showed similar trends. The relative sharpness of the diaphragm can be observed from the line profiles shown in figures 5.5 and 5.6 for simulations 3 and 5 respectively. The relative sharpness of the diaphragm serves as an indicator of the improvement in the image quality.

For the patient data, we qualitatively assessed the performance of the MM-

MCIR approach by comparing the result to the reconstruction obtained when no motion model was applied. Figures 5.9 and 5.10 illustrate the results obtained. Based on visual assessment, the results showed clear improvements, the vessels were more visible in the MM-MCIR images relative to the case of no motion models as indicated by the red arrows, similarly the diaphragms appeared sharper for the MM-MCIR images. In addition no visual difference was observed between the MM-MCIR for iterations 1 and 5 respectively. Intensity line profiles of the diaphragm region from the reconstructed images for the two clinical acquisitions also support the above observations.

Even though M-XDVF were available for simulations 2-5, calculation of the DFEs is more complicated due to using a different $I_0$ position for the simulation and for the generation of the MM-MCIR. The MM-MCIR approach was based on using standard (no motion) reconstruction. This reconstruction was based on the average position of the acquisition, the resulting motion model was also based on this average position. Whereas the M-XDVFs as obtained from the XCAT phantom were based on computing transformations from the first time point EE position to all subsequent time points. As a result, the DVFs obtained from the model do not correspond to those obtained from the GT motion. Nevertheless, it is possible to calculate the GT DVFs from the average position to each time-point, by inverting and composing transformations, and these could be used to calculate the DFEs. Further analysis on the simulated data would be the subject of future work.

Due to the limited time to complete the work for this chapter, I decided to focus on the results from the clinical data rather than perform the DFE analysis for

the simulated data.

There are some intrinsic intensity differences between the simulated and clinical datasets. For instance in the case of the Varian data-set used in this study, we discovered substantial fluctuations in the image intensity between successive projections in some cases as illustrated in figure 5.11. These fluctuations are not present in the simulated projections from the reference static image, as a result this would likely affect the result of the motion model, since these intensity differences do not correspond to the motion of interest. The comparison of the projections over all time points can be found in the supplementary material. (*DynamicProj_vs_simulatedProjection.mp4 - Shows the comparison between the real projections (left), and the simulated model estimated projections (right). The difference between the two set of projections is shown in the middle. The intensity fluctuation of the real projections can be seen over time.*)

The hyper-parameters used for the simulations in this work were based on the ones from chapter 4. Using these hyper-parameters poorer results were obtained for the patient data. For instance, using CPG spacing from the last chapter gave poor results on the clinical data, hence we investigated using coarser CPG spacing for the clinical data. However based on the visual observation in figures 5.7 and 5.8, a small change in CPG spacing could lead to a large change in the results. Gradual improvement in image quality can be observed for higher CPG spacing, which constrains the motion to be smooth. The reference static volume used in chapters 4 and 5 are of different image quality, and there are intrinsic intensity differences between the real projections and the simulated projections (see figure

5.11 ), these could potentially affect the optimum parameters.

The hyper-parameters used in chapter 4 gave good results on the simulated data for the MM-MCIR approach but they did not on real data. When we analysed the results in figures 5.7 and 5.8 some of the recovered motion was not physically plausible so we decided to investigate if using a coarser CPG spacing (to constrain the motion to be smoother) would lead to better results. We also investigated the reason for the poorer results on the real data and discovered the intensity fluctuations, which are likely having a large impact on the results. Therefore, while more work is certainly required to tune the parameters for real data, it is imperative that the intensity fluctuations are first addressed, as these are likely having a large impact on which parameters work best (and requiring that larger constraints are placed on the models and motion to get plausible results than may be the case if the intensity fluctuations were not present). Similarly, for the patient data an optimisation of the hyper-parameters is necessary in order to determine the most suitable ones to use. Based on the different CPG spacings used in this experiment for the patient data, a CPG spacing of 30 and 40 voxels produced a relatively improved image quality, however this may not be the case once the intensity fluctuation has been resolved.

Further tuning and the optimisation of the hyper-parameters on a larger clinical data-sets should lead to more robust results. The use of a larger CPG spacing for the real data in an attempt to constrain the motion can lead to degradation in the image quality in the rib and lung boundary region as shown in figure 5.10. Using a linear model and B-Spline may not be a perfect representation of the true motion for the clinical data, thus affecting the image quality of the MCIR. Finally CBCT

artifacts such as scatter present in the real data would also impact the accuracy of the estimated model and the MCIR.

A full iteration of the MM-MCIR (fitting the model and then performing MCIR) for the patient data took approximately 5 hours, while for the simulated data it was approximately 2 hours. We decided to use all of the projections when using the real data to try and minimise the impact of the intensity fluctuation, even though this led to an increase in the computation time to fit the models. It is desirable to perform as few iterations as possible while also obtaining good enough improved image quality in the MCIR volume and a good estimate of the motion, these need further investigating in the future. In most cases for the simulated and clinical data-sets no substantial improvement was observed between the resulting reconstructions from MM-MCIR after iteration 1 and iteration 5 respectively (see figure 5.2). Thus performing only one iteration of MM-MCIR could be sufficient since this would also reduce the computation time drastically.

For the simulated data the run time for the model fitting was approximately 50 min, while the MCIR was just over 1 hour. While for the patient data the model fitting took just over 3 hours and the MCIR was approximately 2 hours. The current run-time for the MM-MCIR needs to be dramatically reduced for it to be clinically viable, although there is evidence that this should be achievable by reducing the amount of data required for fitting the model and using more advanced implementations of the motion modelling algorithm and dedicated hardware such as Graphics Processing Units (GPU). Regarding the MCIR, one way to improve on the computational time is to bin the projections that have similar motion rather than using fully

continuous motion estimates, this could enable several projections to be corrected at once, rather than doing each separately, thereby greatly reducing the computational time with minimal impact on the MCIR.

Though we were able to apply the MM-MCIR approach to clinically acquired data, there is still a constraint on the type of data this method could be suitable for in its current state. The CBCT data used in this study were projection images whereby the FDK reconstructions covered the entire trans-axial FOV of the patient. For many clinical applications, the FOV of the CBCT reconstruction is mainly focused on the tumour region, therefore leading to truncated reconstruction, thus this method would have limited use. However at our collaborating clinical institution, as well as some others, there seems to be a recent tendency towards non-truncated data, thus this could be less of an issue in the future. A technique to combat the truncated data-set issue will be discussed in chapter 6.

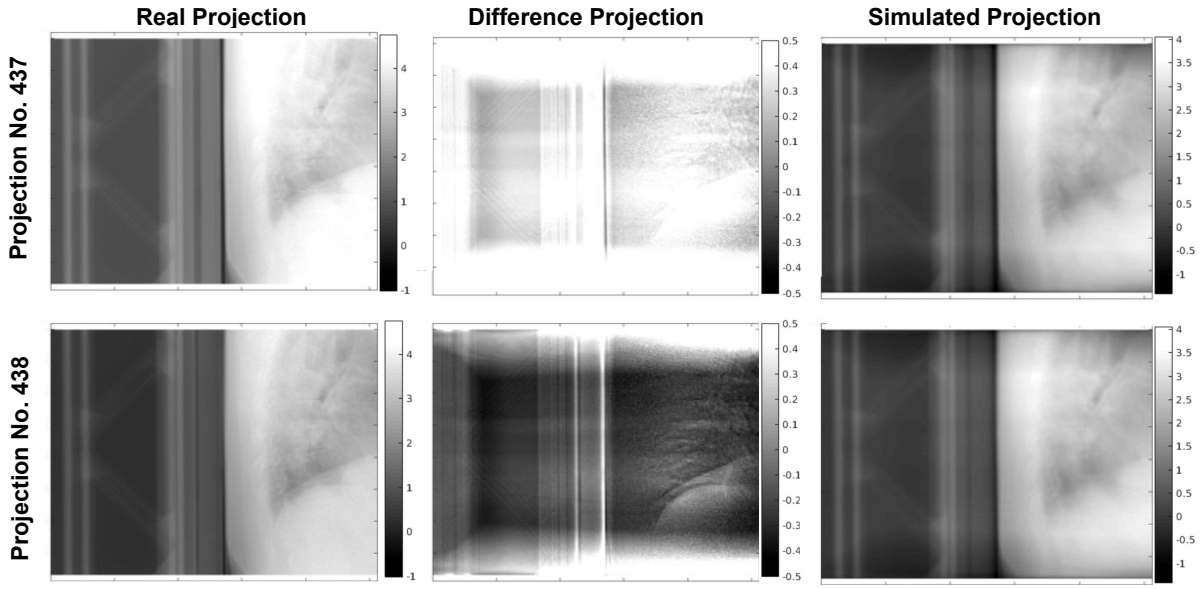


**Figure 5.11:** Illustration of the difference between adjacent projections. The first column shows the real projections for adjacent time points. The second column shows the difference projection between the real projections and the simulated projections. The third shows the simulated projections for adjacent time points.

## 5.6 Conclusion

This chapter extended the correspondence model developed in chapter 4 so that it can be applicable to clinical data. With the true-beam CBCT projection which allows full reconstruction of the patient's FOV, a correspondence motion model was built and MCIR was applied in an iterative approach using the no motion model reconstruction as the initial static volume. Improved image quality was obtained after the first iteration of MM-MCIR. The MM-MCIR approach was tested on both simulated and clinical data. In most cases no substantial quantitative or qualitative improvement in the image quality was observed after the first iteration of MM-MCIR. This implies that only a single iteration of MM-MCIR could be sufficient, but further investigation is required to confirm this conclusion.

We have obtained promising results, even though there are clearly still some errors in the real patient data (which are more evident in the movies). This may be due to the intensity fluctuations. The results presented here are entirely from CBCT projection data from standard 3D-CBCT scan with no external signal, and no previous motion model from 4D-CT or other data. To our knowledge these are the first results on real clinical data to achieve this, and that can account for breath-to-breath variability and do not rely on respiratory sorted 4D-CT.

# Chapter 6

# Summary and Future work

In the work presented in this thesis, we have developed a way for respiratory motion correction of CBCT data, which is solely data-driven and does not need any external monitoring equipment.

Firstly, we introduced a method of extracting a respiratory motion signal from CBCT projection data that correlates to the internal motion. The method allows selecting a ROI for generating the signal, allowing targeting a specific region such as the diaphragm or a tumour. We have shown on simulated data that this method can produce a signal from both the diaphragm and the tumour that is strongly correlated with the internal motion.

In relation to the study presented in chapter 3, future work should involve further analysis with more real data to demonstrate the clinical utility of the proposed method. In this work the clinical data used were from a Varian scanner, it is also important to investigate if the method is applicable to different types of scanner technology with different scan protocols. To truly test the advantage of the proposed technique, the method should be tested on acquisitions whereby there is no

diaphragm in the FOV of the projections.

Validating the method on clinical data is difficult since there is no GT signal. A potential solution to this, is to manually determine the respiratory signal from the projection data by selecting a region that corresponds to respiratory motion from each projection image, as done for one patient, however this approach is time consuming. Another approach would be to use data with implanted markers such that the markers are easy to identify manually. In addition, 4D-CBCT can be performed (as we have performed on simulated data) using the signal and this can be compared to other methods, however qualitative/quantitative assessment on the 4D-CBCT reconstruction can be difficult due to its poor image quality. Another approach to validating the method would be to use the signal to build motion models and then assess the impact of the signal based on the model results.

The correspondence model for CBCT data was extended based on the previous work by the Masters student by thoroughly testing on simulated data and accurately recovering motion (as measured by DFE and COM). MCIR was also implemented using the model estimate (i.e. using "continuous" motion estimates rather than phase-binned). To ensure the method was applicable to patient data, an iterative approach (MM-MCIR) was implemented such that an existing static reference image was not required and a standard reconstruction image was used instead. This method was tested on both simulated and clinical data, with results of MM-MCIR showing improved image quality relative to when no motion model was applied.

In the current work we have used a commonly employed technique of using a single measured signal, and its temporal derivative [3]. However, it would also

be possible to use two measured signals which may be better than using a single signal and its derivative. For CBCT this could be the T-ROI and the D-ROI, D-ROI and RPM signal, or chest and abdominal skin markers could also be used. More experiments on real patient data with real signals need to be performed in order to determine which signal(s) is(are) best to use.

The proposed method of building a correspondence model from CBCT projection data requires further validation. More simulation studies with more complex variable motion/anatomy are required. In addition, a physical phantom could potentially be used to further validate the method proposed, this would enable end-to-end testing of the method, however realistic motion is difficult to achieve with physical phantoms. Furthermore, the motion model can potentially be validated using patient data if two CBCT acquisitions from the same patient are available on the same day. A motion model could be built using the projection data from both acquisitions, and the resulting motion estimate and MCIR could be compared, as well as how well each model (and MCIR) predicts the projections from the other acquisition.

For validating the accuracy of the model over a single CBCT acquisition, the projection data can be used even though it is difficult to estimate the GT motion from the projection data. This is due to difficulties such as low contrast and image quality, structures moving in-front/behind each other within the FOV, and in some cases limited FOV. A potential solution to tackle this issue is to manually identify high contrast structures such as the diaphragm. Implanted markers could be used, however this approach can be invasive and ideally the markers would be removed from the projection images so that they are not unfairly affecting the model fitting.

The unified framework can use different similarity measures (important for using a CT volume as the reference static image. See below), different constraint terms and different transformation models. Future work could investigate if these improve performance. The current implementation of the framework is entirely CPU based, but it could be implemented on GPU (RTK is partially implemented on GPU) which would be expected to significantly reduce the run-time by an order of magnitude. For the framework to fit the current clinical workflow, ideally the computational time needs to be in the order of 1-2 mins. Different approaches to reduce the computational time exist besides implementation on a GPU. As demonstrated on the simulated data, the number of projections used to fit the model can be reduced. In addition, the use of a CT volume as static reference image would remove the need for an iterative MM-MCIR, therefore reducing the computational time. Also using a prior motion model from patient planning or a model generated from a previous fraction to initialise the model fitting and MCIR could potentially reduce the computational time.

As discussed in chapter 5, using the MCIR as the reference static image in an iterative approach requires non-truncated data so that the full anatomy is in the reconstructed FOV and thus will be present in the simulated projections. However, many clinical centres still acquire truncated CBCT data, and the current iterative approach cannot be applied to these datasets. A way of overcoming this issue is to use a 4D-CT volume as the reference image, since the whole of the patient's anatomy is present in the FOV, and the image quality is better than that of conventional 3D-CBCT image reconstruction. Another advantage of using 4D-CT from planning is

that the radiotherapy structures defined on the 4D-CT volume can be animated by the model and propagated onto the model estimates. However this approach has its associated limitations. Firstly the 4D-CT is typically acquired days/weeks prior to treatment, therefore as a result the anatomy could have changed considerably prior to the day of CBCT acquisition, in particular for later fractions when the radiotherapy treatment may have caused anatomical changes. In addition, there are differences between the intensity values of 4D-CT and CBCT, therefore SSD is not suitable as a measure of similarity. A different image similarity metric that is invariant to image intensity such as Local Normalised Cross Correlation (LNCC) could be used, and/or the clinically acquired projections can be modified to be similar to the simulated projections from 4D-CT (either by correcting for effects such as scatter in the real projections, or by simulating those effects in the simulated projections).

How long the estimated model is valid for is another area that requires investigating, although in most cases using a prior model may provide a good initialization. However it is still unclear if it remains valid over a fraction of treatment, and this needs investigating as part of future work. Acquisitions of CBCT projections at the beginning and end of treatment could be used to test if the models are still valid.

There are multiple opportunities for improving the current unified registration and model fitting framework. An important example is to modify the deformation model to handle sliding motion when two parts of the anatomy slide past each other, e.g. the chest wall and the lung during respiration. Most transformation models try to produce a continuous and smooth deformation field. However, when sliding motion occurs the deformation is not continuous, therefore any continuous defor-

mation will inevitably include errors. A number of algorithms have been proposed to address sliding motion [112–114]. However, some of these methods can struggle when there is no clear intensity difference at the sliding boundary (such as between liver and chest wall). Most algorithms for sliding motion require the target image to be segmented. In the case of CBCT the target images are the projection images, so it is not possible to segment the sliding regions. Therefore, Eiben et al. [115] developed an algorithm, which instead only requires the source image to be segmented. In the future this could be incorporated into the motion modelling framework.

The motion model presented in this work could potentially be used for adaptive radiotherapy studies. It could be used for offline adaptation, where the model and the MCIR can be used to give a better estimate of the dose delivered based on the motion during CBCT acquisition. It could also be used for online treatment adaptation, where the model and the MCIR could be used to redefine the target based on the current motion and re-optimise the treatment plan based on the updated target and current motion. The model could potentially be used to guide active treatment for instance gating/tracking treatment, however this would require the use of a different surrogate signal that can be acquired during treatment delivery.

In summary, this PhD thesis has demonstrated the potential for building respiratory motion models from CBCT data. While there is still much work to be done to optimise the results and improve the method so that it is suitable for clinical use, the results presented here are promising, and with further development could provide accurate motion information and ultimately lead to more accurate and more effective radiotherapy treatments for lung tumours.

## Chapter 7

# Appendices

# Appendix A

# Figures

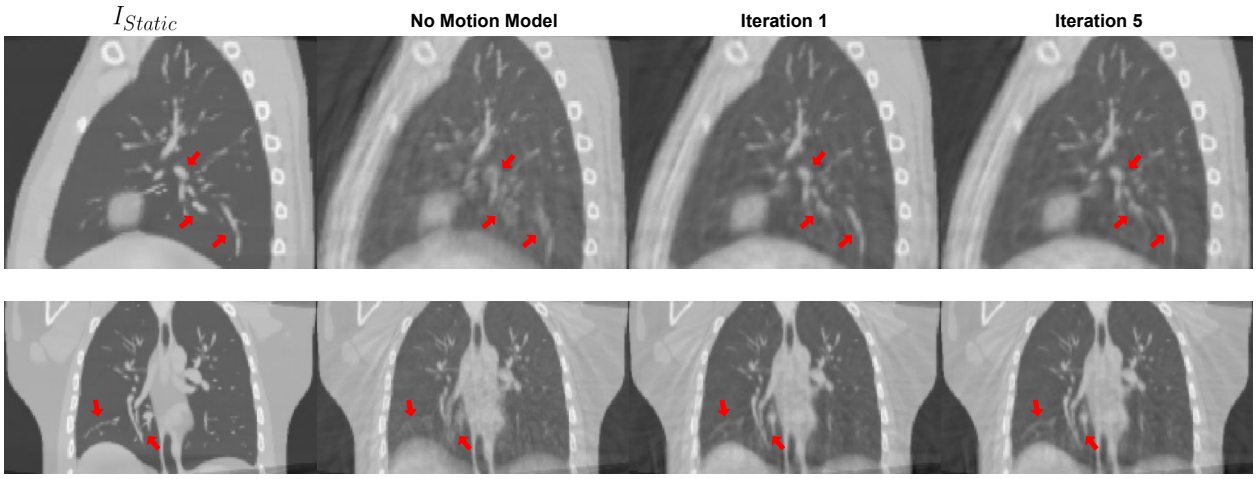


**Figure A.1:** Qualitative results shown for simulation 2. The first row shows the slice through the reconstruction images for $I_{Static}$, no motion model, the MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations respectively, while the $2^{nd}$ row shows the sagittal slices through the same corresponding volumes. The arrows shown in red highlight regions of discrepancies between the different volumes.

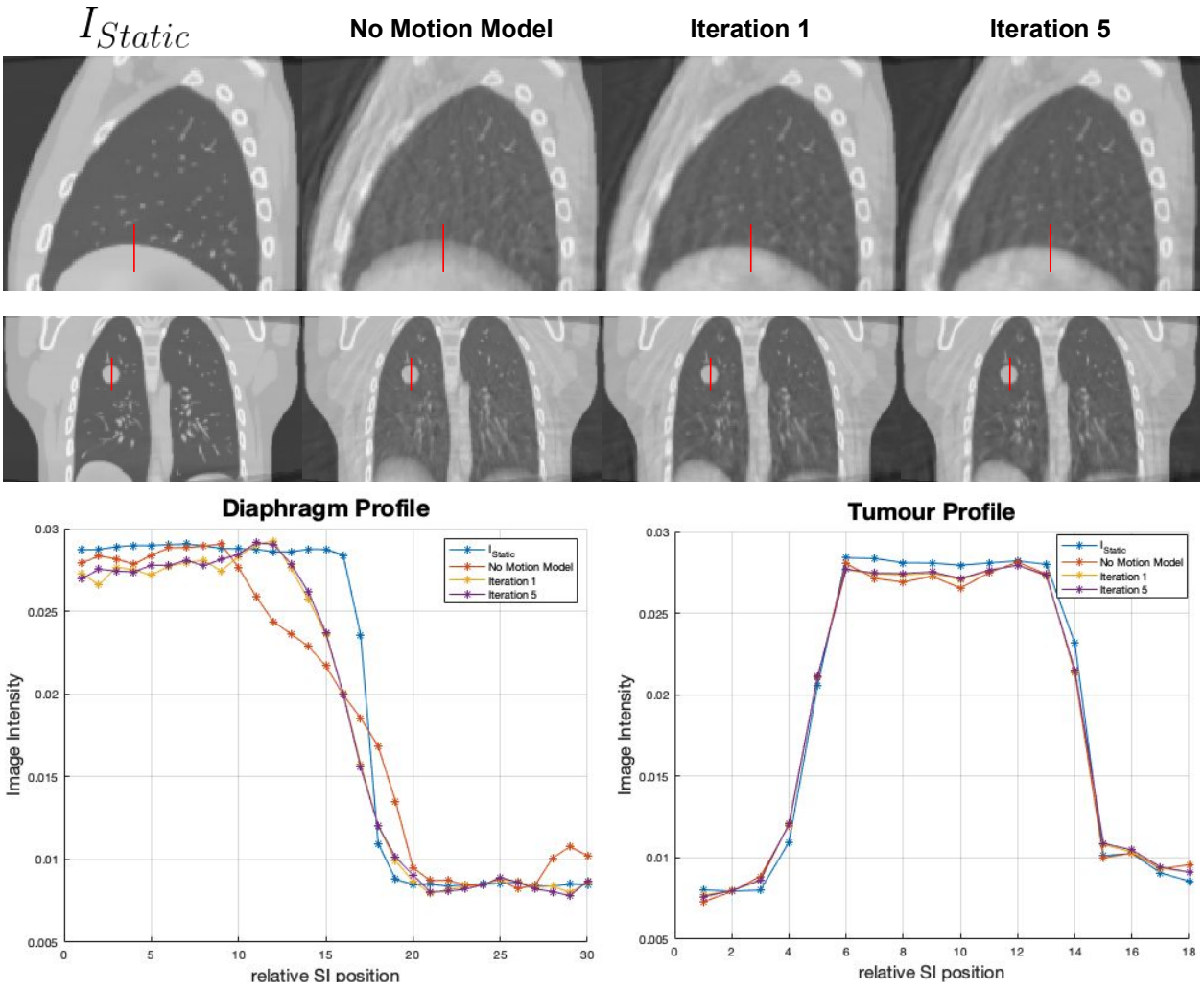


**Figure A.2:** Results obtained for simulation 2. Shown are the slices through the reconstruction images for $I_{Static}$, no motion model, the MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations iterations. The coronal slices are shown in the $1^{st}$ row, while the corresponding sagittal slices for the volumes are shown in the $2^{nd}$ row. The diaphragm profile shows the intensity profile through the diaphragm region in the $1^{st}$ row as highlighted by the red line, while the tumour profile shows the intensity profile through the tumour region as highlighted by the red line in the $2^{nd}$ row.

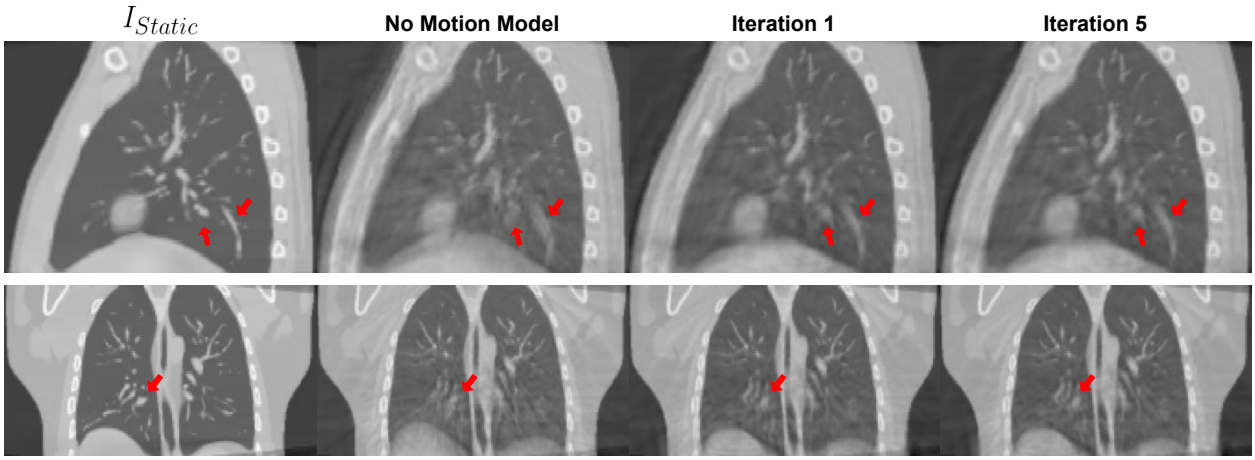


**Figure A.3:** Qualitative results shown for simulation 4. The first row shows the slice through the reconstruction images for $I_{Static}$, no motion model, the MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations respectively, while the $2^{nd}$ row shows the sagittal slices through the same corresponding volumes. The arrows shown in red highlight regions of discrepancies between the different volumes.

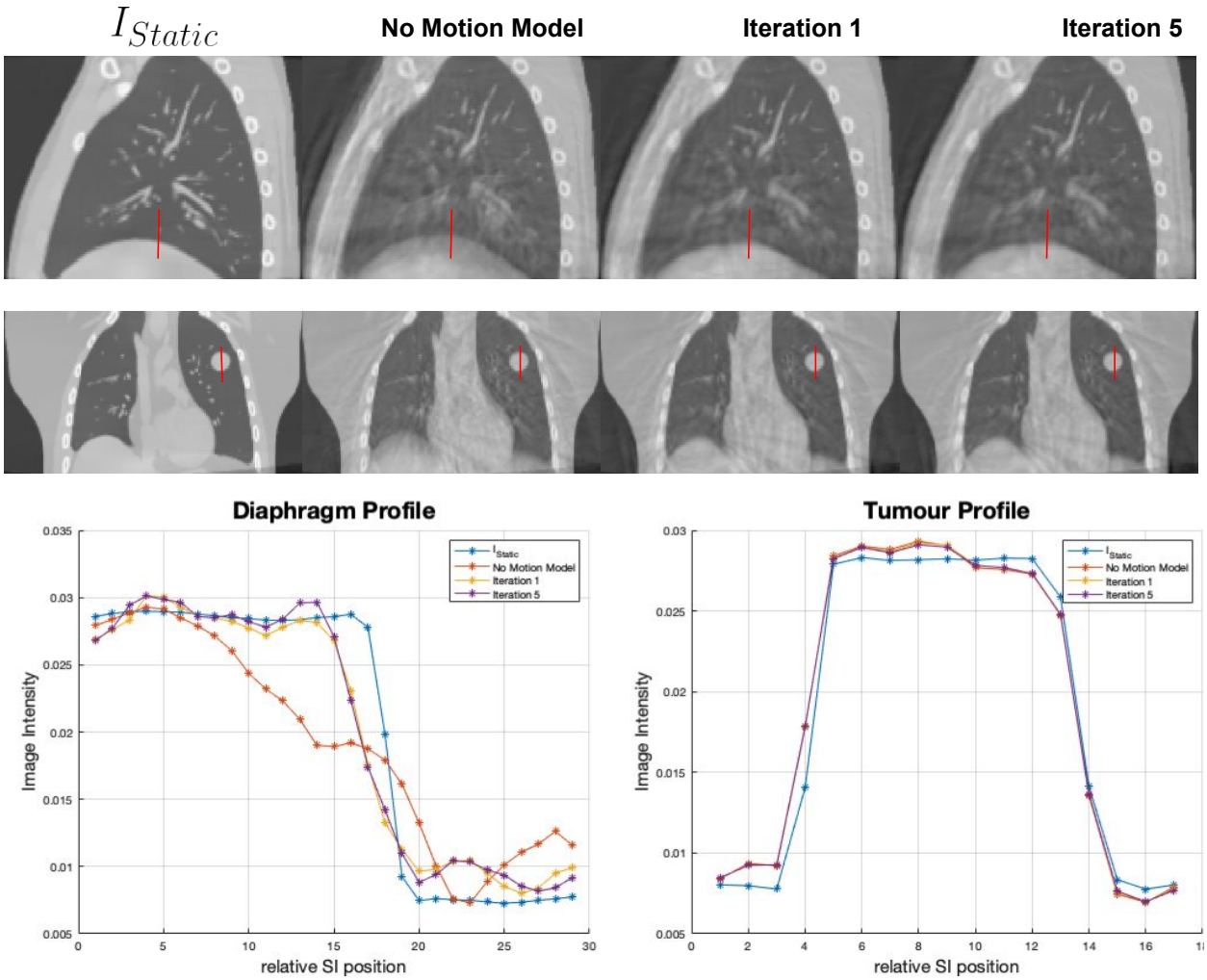


**Figure A.4:** Results obtained for simulation 4. Shown are the slices through the reconstruction images for $I_{Static}$, no motion model, the MM-MCIR after 1 iteration and the MM-MCIR after 5 iterations. The coronal slices are shown in the $1^{st}$ row, while the corresponding sagittal slices for the volumes are shown in the $2^{nd}$ row. The diaphragm profile shows the intensity profile through the diaphragm region in the $1^{st}$ row as highlighted by the red line, while the tumour profile shows the intensity profile through the tumour region as highlighted by the red line in the $2^{nd}$ row.

# Bibliography


[1] Joel Beaudry. *4D Cone-Beam CT Image Reconstruction of Varian TrueBeam v1.6 Projection Images for Clinical Quality Assurance of Stereotactic Ablative Radiotherapy to the Lung by*. PhD thesis, The University of British Columbia (Vancouver), 2015.

[2] Marcel Beister, Daniel Kolditz, and Willi A. Kalender. Iterative reconstruction methods in X-ray CT. *Physica Medica*, 28(2):94–108, 2012.

[3] J. R. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King. Respiratory motion models: A review. *Medical Image Analysis*, 17(1):19–42, 2013.

[4] P. Boyle, B. Levin and I. A. World cancer report 1. *International Agency for Research on Cancer*, 2008.

[5] Cancer Research UK. Cancer Statistics for the UK, 2019.

[6] National Cancer Registration & Analysis Service and Cancer Research UK. "Chemotherapy, Radiotherapy and Tumour Resections in England: 2013-2014" workbook, 2017.

[7] Peter M. Ellis and Rachel Vandermeer. Delays in the diagnosis of lung cancer. *Journal of Thoracic Disease*, 3(3):183–188, 2011.

[8] Kavitha Srnivasan, Mohammad Mohammadi, and Justin Shepherd. Applications of linac-mounted kilovoltage Cone-beam Computed Tomography in modern radiation therapy: A review. *Polish Journal of Radiology*, 79:181–193, 2014.

[9] Cancer Research UK. No Title, 2019.

[10] Office of National Statistics. Cancer registration statistics, England: first release, 2015, 2019.

[11] Thierry Berghmans, Marianne Paesmans, and Jean Paul Sculier. Prognostic factors in stage III non-small cell lung cancer: A review of conventional, metabolic and new biological variables. *Therapeutic Advances in Medical Oncology*, 3(3):127–138, 2011.

[12] Timothy Das, Indra and Andersen, Aaron and Chen, Zhe and Dimofte, Andrea and Glatstein, Eli and Hoisak, Jeremy and Huang, Long and Langer, Mark and Lee, Choonik and Pacella, Matthew and Popple, Richard and Rice, Roger and Smilowitz, Jennifer and Sponseller, Patri. State of Dose Prescription and Compliance to International Standard (ICRU-83) in Intensity Modulated Radiation Therapy among Academic Institutions. *Practical Radiation Oncology*, 7, 2016.

[13] Joep C. Stroom, Hans C.J. De Boer, Henk Huizenga, and Andries G. Visser. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *International Journal of Radiation Oncology Biology Physics*, 43(4):905–919, 1999.

[14] Matthias Guckenberger. Image-guided Radiotherapy Based on Kilovoltage Cone-beam Computed Tomography — A Review of Technology and Clinical Outcome. *European Oncology & Haematology*, 07(02):121, 2011.

[15] J.B West. *Respiratory Physiology: The Essentials*. Williams and Wilkins, seventh edition, 2004.

[16] Erridge SC, Y Seppenwoolde, SH Muller, M van Herk, De Jaeger K, JS Belderbos, LJ Boersma, and JV Lebesque. Portal imaging to assess set-up errors, tumor motion and tumor shrinkage during conformal radiotherapy of non-small cell lung cancer. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, 66(1):75–85, 2003.

[17] K Konno and J Mead. Measurement of the separate volume changes of rib cage and abdomen during breathing. *Journal of applied physiology*, 22(3):407–422, 1967.

[18] EA Barnes, BR Murray, DM Robinson, LJ Underwood, J Hanson, and WH Roa. Dosimetric evaluation of lung tumor immobilization using breath hold at deep inspiration. *International Journal of Radiation Oncology Biology Physics*, 50(4):1091–1098, 2001.

[19] Christian Plathow, Sebastian Ley, Christian Fink, Michael Puderbach, Waldemar Hosch, Astrid Schmähl, Jürgen Debus, and Hans Ulrich Kauczor. Analysis of intrathoracic tumor mobility during whole breathing cycle by dynamic MRI. *International Journal of Radiation Oncology Biology Physics*, 59(4):952–959, 2004.

[20] S. S Vedam, V. R Kini, P. J Keall, V Ramakrishnan, H Mostafavi, and R. Mohan. Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. *Med. Phys.*, 30(4):505–513, 2003.

[21] G Van Goethem, Martin J.J., Dermaut B., Löfgren A, Wibail A., Ververken D., Tack P., Dehaeneg I., Van Zandijcke M., Moonen M., Ceuterick C., De Jonghe P., and Van Broeckhoven C. Recessive POLG mutations presenting with sensory and ataxic neuropathy in compound heterozygote patients with progressive external ophthalmoplegia. *Neuromuscular Disorders*, 13:133–142, 2003.

[22] T. Landberg, J. Chavaudra, J. Dobbs, J. P. Gerard, G. Hanks, J. C. Horiot, K. A. Johansson, T. Möller, J. Purdy, N. Suntharalingam, and H. Svensson. ICRU. Report 62. *Journal of the International Commission on Radiation Units and Measurements*, os32(1), 1999.

[23] Jochem W.H. Wolthaus, Jan Jakob Sonke, Marcel van Herk, José S.A. Belderbos, Maddalena M.G. Rossi, Joos V. Lebesque, and Eugène M.F. Damen. Comparison of Different Strategies to Use Four-Dimensional Computed Tomography in Treatment Planning for Lung Cancer Patients. *Inter-*

*national Journal of Radiation Oncology Biology Physics*, 70(4):1229–1238, 2008.

[24] Elena H. Tran, Björn Eiben, Andreas Wetscherek, Uwe Oelfke, Gustav Meedt, David J. Hawkes, and Jamie R. McClelland. Evaluation of MRI-derived surrogate signals to model respiratory motion. *Biomedical Physics and Engineering Express*, 6(4), 2020.

[25] Bjorn Eiben, Jenny Bertholet, Martin J. Menten, Simeon Nill, Uwe Oelfke, and Jamie R. McClelland. Consistent and invertible deformation vector fields for a breathing anthropomorphic phantom: a post-processing framework for the XCAT phantom. *Phys. Med. Biol.*, in press, 2020.

[26] Gengsheng Lawrence Zeng. *Medical image reconstruction A conceptual tutorial*. Springer, Germany, 2010.

[27] F Natterer. *The Mathematics of Computerized Tomography*. Society for Industrial and Applied Mathematics, 2001.

[28] L A Feldkamp. Practical cone-beam algorithm. *Journal of the Optical Society of America*, 1(6):612–619, 1984.

[29] Henrik Turbell. *Cone-Beam Reconstruction Using Filtered Backprojectionn*. PhD thesis, 2001.

[30] James Martin. *Surrogate-driven motion models from cone-beam CT for motion management in radiotherapy treatments*. PhD thesis, 2014.

[31] Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin Keun Seo. Deep learning for undersampled MRI reconstruction. *arXiv*, 2017.

[32] Kaichao Liang, Yuxiang Xing, Hongkai Yang, and Kejun Kang. Improve angular resolution for sparse-view CT with residual convolutional neural network. (March 2018):55, 2018.

[33] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017.

[34] Hu Chen, Yi Zhang, Weihua Zhang, Peixi Liao, Ke Li, Jiliu Zhou, and Ge Wang. aLow-dose CT via convolutional neural network. *Biomedical Optics Express*, 8(2):679, 2017.

[35] R. Schulze, U. Heil, D. Groß, D. D. Bruellmann, E. Dranischnikow, U. Schwanecke, and E. Schoemer. Artefacts in CBCT: A review. *Dentomaxillofacial Radiology*, 40(5):265–273, 2011.

[36] Eike Rietzel, Tinsu Pan, and George T.Y. Chen. Four-dimensional computed tomography: Image formation and clinical protocol. *Medical Physics*, 32(4):874–889, 2005.

[37] Stephan M. Stockberger, Jeffrey A. Hicklin, Yun Liang, Justin L. Wass, and Walter T. Ambrosius. Spiral CT with ionic and nonionic contrast material:

Evaluation of patient motion and scan quality. *Radiology*, 206(3):631–636, 1998.

[38] Jan Ehrhardt, René Werner, Dennis Säring, Thorsten Frenzel, Wei Lu, Daniel Low, and Heinz Handels. An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing. *Medical Physics*, 34(2):711–721, 2007.

[39] Tinsu Pan, Xiaojun Sun, and Dershan Luo. Improvement of the cine-CT based 4D-CT imaging. *Medical Physics*, 34(11):4499–4503, 2007.

[40] Tokihiro Yamamoto. Retrospective Analysis of Artifacts in Four-Dimensional CT Images of 50 Abdominal and Thoracic Radiotherapy Patients. *International Journal of Radiation Oncology • Biology • Physics*, 72(4):1250 – 1258, 2008.

[41] AnilKumar Nagarajappa, Neha Dwivedi, and Rana Tiwari. Artifacts: The downturn of CBCT image. *Journal of International Society of Preventive and Community Dentistry*, 5(6):440, 2015.

[42] Russell E Kincaid Jr, Ellen D Yorke, Karyn a Goodman, Andreas Rimner, Abraham J Wu, S Gig, and Russell E Kincaid. Investigation of gated cone-beam CT to reduce respiratory motion blurring Investigation of gated cone-beam CT to reduce respiratory motion blurring. *Medical Physics*, 40(4):041717, 2013.

[43] Jan-jakob Sonke, Lambert Zijp, Peter Remeijer, and Marcel Van Herk. Respiratory correlated cone beam CT. *Med. Phys.*, 32(4):1176 – 1186, 2005.

[44] Simon Rit, Jochem Wolthaus, Marcel Van Herk, and Jan Jakob Sonke. On-the-fly motion-compensated cone-beam CT using an a priori motion model. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 36(6):729–736, 2009.

[45] T Li, E Schreibmann, Y Yang, and L Xing. Motion correction for improved target localization with on-board cone-beam computed tomography. *Physics in Medicine and Biology*, 51(2):253–267, jan 2006.

[46] D. Rueckert et al. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imag.*, 18(8):712–21, 1999.

[47] Francisco P.M. Oliveira and João Manuel R.S. Tavares. Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(2):73–93, 2014. PMID: 22435355.

[48] Nicholas J. Tustison, Brian B. Avants, and James C. Gee. Directly manipulated free-form deformation image registration. *IEEE Transactions on Image Processing*, 18(3):624–635, 2009.

[49] Jeremy D.P. Hoisak, Katharina E. Sixel, Romeo Tirona, Patrick C.F. Cheung, and Jean Philippe Pignol. Correlation of lung tumor motion with external sur-

rogate indicators of respiration. *International Journal of Radiation Oncology Biology Physics*, 60(4):1298–1306, 2004.

[50] J. R. McClelland, S. Hughes, M. Modat, A. Qureshi, S. Ahmad, D. B. Landau, S. Ourselin, and D. J. Hawkes. Inter-fraction variations in respiratory motion models. *Physics in Medicine and Biology*, 56(1):251–272, 2011.

[51] Ruijiang Li, John H. Lewis, Xun Jia, Xuejun Gu, Michael Folkerts, Chunhua Men, William Y. Song, and Steve B. Jiang. 3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy. *Medical Physics*, 38(5):2783–2794, 2011.

[52] Z. van Kesteren, A. van der Horst, O. J. Gurney-Champion, I. Bones, D. Tekelenburg, T. Alderliesten, G. van Tienhoven, R. Klaassen, H. W. M. van Laarhoven, and A. Bel. A novel amplitude binning strategy to handle irregular breathing during 4DMRI acquisition: improved imaging for radiotherapy purposes. *Radiation Oncology*, 14(1):80, dec 2019.

[53] Daniel A. Low, Parag J. Parikh, Wei Lu, James F. Dempsey, Sasha H. Wahab, James P. Hubenschmidt, Michelle M. Nystrom, Maureen Handoko, and Jeffrey D. Bradley. Novel breathing motion model for radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 63(3):921–929, 2005.

[54] A. P. King, C. Buerger, C. Tsoumpas, P. K. Marsden, and T. Schaeffter. Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. *Medical Image Analysis*, 16(1):252–264, 2012.

[55] Tiezhi Zhang, Harry Keller, Matthew J. O'Brien, Thomas R. Mackie, and Bhudatt Paliwal. Application of the spirometer in respiratory gated radiotherapy. *Medical Physics*, 30(12):3165–3171, 2003.

[56] E. C. Ford, G. S. Mageras, E. Yorke, K. E. Rosenzweig, R. Wagman, and C. C. Ling. Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging. *International Journal of Radiation Oncology Biology Physics*, 52(2):522–531, 2002.

[57] S. Ahn, B. Yi, Y. Suh, J. Kim, S. Lee, S. Shin, S. Shin, and E. Choi. A feasibility study on the prediction of tumour location in the lung from skin motion. *British Journal of Radiology*, 77(919):588–596, 2004.

[58] John W. Wong, Michael B. Sharpe, David A. Jaffray, Vijay R. Kini, John M. Robertson, Jannifer S. Stromberg, and Alavro A. Martinez. The use of active breathing control (ABC) to reduce margin for breathing motion. *International Journal of Radiation Oncology Biology Physics*, 44(4):911–919, 1999.

[59] Laura A. Dawson, Kristy K. Brock, Sahira Kazanjian, Dwight Fitch, Cornelius J. McGinn, Theodore S. Lawrence, Randall K. Ten Haken, and James Balter. The reproducibility of organ position using active breathing control (ABC) during liver radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 51(5):1410–1421, 2001.

[60] G. S. Mageras, E. Yorke, K. Rosenzweig, L. Braban, E. Keatley, E. Ford, S. A. Leibel, and C. C. Ling. Fluoroscopic evaluation of diaphragmatic

motion reduction with a respiratory gated radiotherapy system. *Journal of applied clinical medical physics / American College of Medical Physics*, 2(4):191–200, 2001.

[61] B Dong, Y J Graves, X Jia, and S B Jiang. Optimal surface marker locations for tumor motion estimation in lung cancer radiotherapy. *Physics in Medicine and Biology*, 57(24):8201–8215, 2012.

[62] David P. Gierga, Johanna Brewer, Gregory C. Sharp, Margrit Betke, Christopher G. Willett, and George T.Y. Chen. The correlation between internal and external markers for abdominal tumors: Implications for respiratory gating. *International Journal of Radiation Oncology Biology Physics*, 61(5):1551–1558, 2005.

[63] Hui Yan, Fang Fang Yin, Guo Pei Zhu, Munther Ajlouni, and Jae Ho Kim. The correlation evaluation of a tumor tracking system using multiple external markers. *Medical Physics*, 33(11):4073–4084, 2006.

[64] Salam Dhou, Yuichi Motai, and Geoffrey D. Hugo. Local intensity feature tracking and motion modeling for respiratory signal extraction in cone beam CT projections. *IEEE Transactions on Biomedical Engineering*, 60(2):333–342, 2013.

[65] Seonyeong Park, Student Member, Siyong Kim, Byongyong Yi, Geoffrey Hugo, H Michael Gach, Yuichi Motai, and Senior Member. A Novel Method

of Cone Beam CT Projection Binning Based on Image Registration. *IEEE Transactions on Medical Imaging*, 36(8):1733–1745, 2017.

[66] L Zijp, Jj Sonke, and M van Herk. Extraction of the respiratory signal from sequential thorax Cone-Beam X-ray images. *International Conference on the Use of Computers in Radiation Therapy*, pages 507–9, 2004.

[67] Ming Chao, Jie Wei, Tianfang Li, Yading Yuan, Kenneth E. Rosenzweig, and Yeh Chi Lo. Robust breathing signal extraction from cone beam CT projections based on adaptive and global optimization techniques. *Physics in Medicine and Biology*, 61(8):3109–3126, 2016.

[68] Hao Yan, Xiaoyu Wang, Wotao Yin, Tinsu Pan, Moiz Ahmad, Xuanqin Mou, Laura Cerviño, Xun Jia, and Steve B Jiang. Extracting respiratory signals from thoracic cone beam CT projections. *Physics in Medicine and Biology*, 58(5):1447–1464, mar 2013.

[69] Anthony Kavanagh, Philip M. Evans, Vibeke N. Hansen, and Steve Webb. Obtaining breathing patterns from any sequential thoracic x-ray image set. *Physics in Medicine and Biology*, 54(16):4879–4888, 2009.

[70] Kris Thielemans, Shailendra Rathore, Fredrik Engbrant, and Pasha Razifar. Device-less gating for PET/CT using PCA. *IEEE Nuclear Science Symposium Conference Record*, pages 3904–3910, 2011.

[71] Lars Dietrich, Siri Jetter, Thomas Tücking, Simeon Nill, and Uwe Oelfke. Linac-integrated 4D cone beam CT: First experimental results. *Physics in Medicine and Biology*, 51(11):2939–2952, 2006.

[72] Sergey Kriminski, Matthias Mitschke, Stephen Sorensen, Nicole M. Wink, Phillip E. Chow, Steven Tenn, and Timothy D. Solberg. Respiratory correlated cone-beam computed tomography on an isocentric C-arm. *Physics in Medicine and Biology*, 50(22):5263–5280, 2005.

[73] Peter Munro and Christopher Mcguinness. Four-dimensional cone-beam computed tomography using an on-board imager. *Med. Phys.*, 33(10):3825 – 3833, 2006.

[74] Yvette Seppenwoolde, Hiroki Shirato, Kei Kitamura, Shinichi Shimizu, Marcel Van Herk, Joos Lebesque, and Kazuo Miyasaka. Precise and Real-Time Measurement of 3D Tumor Motion in. *Int J Radiat Oncol Biol Phys*, 53(4):822–834, 2002.

[75] I Berbeco Ross, Nishioka Seiko, Shirato Hiroki, T Y Chen George, and B Jiang Steve. Residual motion of lung tumours in gated radiotherapy with external respiratory surrogates. 50, 2005.

[76] Elisabeth Steiner, Chun-chien Shieh, Vincent Caillet, Jeremy Booth, Ricky O Brien, Adam Briggs, Nicholas Hardcastle, Dasantha Jayamanne, Kathryn Szymura, Thomas Eade, and Paul Keall. Both four-dimensional computed tomography and four-dimensional cone beam computed tomography under-

predict lung target motion during radiotherapy. *Radiotherapy and Oncology*, 135:65–73, 2019.

[77] X. Jia, Z. Tian, Y. Lou, and S. Jiang. WE-A-301-02: Four-Dimensional Cone Beam CT Reconstruction and Enhancement with a Temporal Non-Local Means Method. *Medical Physics*, 38(6):3796, 2011.

[78] Shuai Leng, Jie Tang, Joseph Zambelli, Brian Nett, Ranjini Tolakanahalli, and Guang Hong Chen. High temporal resolution and streak-free four-dimensional cone-beam computed tomography. *Physics in Medicine and Biology*, 53(20):5653–5673, 2008.

[79] Guang Hong Chen, Jie Tang, and Shuai Leng. Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Medical Physics*, 35(2):660–663, 2008.

[80] Cyril Mory, Vincent Auvray, B. Zhang, Michael Grass, Dirk Schäfer, S. James Chen, John D. Carroll, Simon Rit, Françoise Peyrin, Philippe Douek, and Loïc Boussel. Cardiac C-arm computed tomography using a 3D + time ROI reconstruction method with spatial and temporal regularization. *Medical physics*, 41(2):021903, 2014.

[81] Cyril Mory, Guillaume Janssens, and Simon Rit. Motion-aware temporal regularization for improved 4D cone-beam computed tomography. *Physics in Medicine and Biology*, 61(18):6856–6877, 2016.

[82] Qinghui Zhang, Alex Pevsner, Agung Hertanto, Yu Chi Hu, Kenneth E. Rosenzweig, C. Clifton Ling, and Gig S. Mageras. A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Medical Physics*, 34(12):4772–4781, 2007.

[83] Q. Zhang, S. Kriminski, C. Ling, K. Rosenzweig, and G. Mageras. SU-FF-J-49: Reducting of Motion Artifacts in Cone Beam CT Using a Patient Specific Respiratory Motion Model. *Medical Physics*, 34(6):2379, 2007.

[84] Matthew J. Riblett, Gary E. Christensen, Elisabeth Weiss, and Geoffrey D. Hugo. Data-driven respiratory motion compensation for four-dimensional cone-beam computed tomography (4D-CBCT) using groupwise deformable registration. *Medical Physics*, 45(10):4471–4482, 2018.

[85] Jef Vandemeulebroucke, Jan Kybic, Patrick Clarysse, and David Sarrut. Respiratory motion estimation from cone-beam projections using a prior model. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5762 LNCS(PART 2):365–372, 2009.

[86] You Zhang, Fang Fang Yin, W. Paul Segars, and Lei Ren. A technique for estimating 4D-CBCT using prior knowledge and Limited-angle projections. *Medical Physics*, 40(12), 2013.

[87] You Zhang, Fang Fang Yin, Tinsu Pan, Irina Vergalasova, and Lei Ren. Preliminary clinical evaluation of a 4D-CBCT estimation technique using

prior information and limited-angle projections. *Radiotherapy and Oncology*, 115(1):22–29, 2015.

[88] Sebastian. Motion vector field phase-to-amplitude resampling for 4D motion- compensated cone-beam CT. *Phys. Med. Biol*, 63, 2018.

[89] Paul J Keall, Gig S Mageras, James M Balter, Richard S Emery, Kenneth M Forster, Steve B Jiang, Jeffrey M Kapatoes, Daniel a Low, Martin J Murphy, Brad R Murray, Chester R Ramsey, Marcel B Van Herk, S Sastry Vedam, John W Wong, and Ellen Yorke. *The management of respiratory motion in radiation oncology report of AAPM Task Group 76.*, volume 33. 2006.

[90] Keith R. Britton, George Starkschall, Susan L. Tucker, Tinsu Pan, Christopher Nelson, Joe Y. Chang, James D. Cox, Radhe Mohan, and Ritsuko Komaki. Assessment of Gross Tumor Volume Regression and Motion Changes During Radiotherapy for Non-Small-Cell Lung Cancer as Measured by Four-Dimensional Computed Tomography. *International Journal of Radiation Oncology Biology Physics*, 68(4):1036–1046, 2007.

[91] Kristin J. Redmond, Danny Y. Song, Jana L. Fox, Jessica Zhou, C. Nicole Rosenzweig, and Eric Ford. Respiratory Motion Changes of Lung Tumors Over the Course of Radiation Therapy Based on Respiration-Correlated Four-Dimensional Computed Tomography Scans. *International Journal of Radiation Oncology Biology Physics*, 75(5):1605–1612, 2009.

[92] J. Wang and X. Gu. TH-C-103-02: Simultaneous Motion Estimation and Image Reconstruction (SMEIR) for 4D Cone-Beam CT. *Medical Physics*, 40(6):542, 2013.

[93] Spyros Sioutas, Dimitrios Tsoumakos, Alexandros Panaretos, Giannis Tzimas, Ioannis Karydis, and Dimitrios Tsolis. Sart. *ACM SIGAPP Applied Computing Review*, 12(3):60–74, 2012.

[94] Berthold K. Horn and Brian G. Schunck. Determining Optical Flow. *Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cam bridge*, 17:185–203, 1981.

[95] Minghao Guo, Geraldine Chee, Dylan O'Connell, Salam Dhou, Jie Fu, Kamal Singhrao, Dan Ionascu, Dan Ruan, Percy Lee, Daniel A. Low, Jun Zhao, and John H. Lewis. Reconstruction of a high-quality volumetric image and a respiratory motion model from patient CBCT projections. *Medical Physics*, 46(August):3627–3639, 2019.

[96] Wendy Harris, You Zhang, Fang Fang Yin, and Lei Ren. Estimating 4D-CBCT from prior information and extremely limited angle projections using structural PCA and weighted free-form deformation for lung radiotherapy. *Medical physics*, 44(3):1089–1104, 2017.

[97] Jamie R McClelland, Marc Modat, Simon Arridge, Helen Grimes, Derek D Souza, David Thomas, Dyla O Connell, Daniel A Low, Evangelia Kaza, David J Collins, Martin O Leach, and David J. Hawkes. A generalized

framework unifying image registration and respiratory motion models and incorporating image reconstruction, for partial or full imaging data. *Physics in medicine and biology*, 62(Iccr 2016):4273–4292, 2017.

[98] James Martin, Jamie McClelland, Connie Yip, Christopher Thomas, Clare Hartill, Shahreen Ahmad, Richard O'Brien, Ivan Meir, David Landau, and David Hawkes. Building motion models of lung tumours from cone-beam CT for radiotherapy applications. *Physics in Medicine and Biology*, 58(6):1809–1822, 2013.

[99] Ian T. Jollife and Jorge Cadima. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 2016.

[100] Ottavia Bertolli, Simon Arridge, Scott D. Wollenweber, Charles W. Stearns, Brian F. Hutton, and Kris Thielemans. Sign determination methods for the respiratory signal in data-driven PET gating. *Physics in Medicine and Biology*, 62(8):3204–3220, 2017.

[101] Paul J. Schleyer, Michael J. O'Doherty, and Paul K. Marsden. Extension of a data-driven gating technique to 3D, whole body PET studies. *Physics in Medicine and Biology*, 56(13):3953–3965, 2011.

[102] W. P. Segars, M. Mahesh, T. J. Beck, E. C. Frey, and B. M.W. Tsui. Realistic CT simulation using the 4D XCAT phantom. *Medical Physics*, 35(8):3800–3808, 2008.

[103] Andrew P. Santoso, Kwang H. Song, Yujiao Qin, Stephen J. Gardner, Chang Liu, Indrin J. Chetty, Benjamin Movsas, Munther Ajlouni, and Ning Wen. Evaluation of gantry speed on image quality and imaging dose for 4D cone-beam CT acquisition. *Radiation Oncology*, 11(1):1–11, 2016.

[104] Simon Rit, M Vila Oliva, S Brousmiche, R Labarbe, D Sarrut, and G C Sharp. The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK). *Journal of Physics: Conference Series*, 489:012079, 2014.

[105] C. D. Chinneck, M. McJury, and A. R. Hounsell. The potential for undertaking slow CT using a modern CT scanner. *British Journal of Radiology*, 83(992):687–693, 2010.

[106] Simon Rit, Marcel van Herk, Lambert Zijp, and Jan-Jakob. Quantification of the Variability of Diaphragm Motion and Implications for Treatment Margin Construction. *International Journal of Radiation Oncology*Biology*Physics*, 82(3):e399–e407, mar 2012.

[107] Pauline Mouches. *The adaptation of RTK and NiftyRegResp framework for respiratory motion model using CBCT projection data*. PhD thesis, Centre Régional de Lutte Contre le Cancer Centre Léon Bérard (Lyon), 2017.

[108] IEC. Radiotherapy equipment - coordinates, movements and scales. 1217, 1996.

[109] Dirk Manke, Kay Nehrke, and Peter Börnert. Novel prospective respiratory motion correction approach for free-breathing coronary mr angiography using a patient-adapted affine motion model. *Magnetic resonance in medicine*, 50(1):122—131, July 2003.

[110] Laura I. Cervîo, Alvin K Y Chao, Ajay Sandhu, and Steve B. Jiang. The diaphragm as an anatomic surrogate for lung tumor motion. *Physics in Medicine and Biology*, 54(11):3529–3541, 2009.

[111] Raj K. Panta, Paul Segars, Fang-Fang Yin, and Jing Cai. Establishing a framework to implement 4D XCAT Phantom for 4D radiotherapy research. *J Cancer Res Ther.*, 8(4):565–570, 2012.

[112] Floris F. Berendsen, Alexis N. T. J. Kotte, Max A. Viergever, and Josien P. W. Pluim. Registration of organs with sliding interfaces and changing topologies. *Medical Imaging 2014: Image Processing*, 9034(March 2014):90340E, 2014.

[113] V. Delmon, S. Rit, R. Pinho, and D. Sarrut. Registration of sliding objects using direction dependent B-splines decomposition. *Physics in medicine and biology*, 58(5):1303–1314, 2013.

[114] Bartłomiej W. Papiez, Mattias P. Heinrich, Jérome Fehrenbach, Laurent Risser, and Julia A. Schnabel. An implicit sliding-motion preserving regularisation via bilateral filtering for deformable image registration. *Medical Image Analysis*, 18(8):1299–1311, 2014.

[115] Björn Eiben, Elena H. Tran, Martin J. Menten, Uwe Oelfke, David J. Hawkes, and Jamie R. McClelland. Statistical motion mask and sliding registration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10883 LNCS:13–23, 2018.