# Registration Based Respiratory Motion Models for use in Lung Radiotherapy

James Robert McClelland

A dissertation submitted in partial fulfilment of the requiremnets for the degree of

**Doctor of Philosophy**

of University College London

March 2008

Centre for Medical Image Computing, University College London





ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, James Robert McClelland, 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 thesis.

James Robert McClelland, 31st March 2008

# Abstract

Respiratory motion is a major factor contributing to errors and uncertainties in Radiotherapy (RT) treatment of lung tumours. Knowledge of this motion may improve the planning and delivery of RT treatment for lung cancer patients. This thesis develops and evaluates methods of building patient specific respiratory motion models. These relate the internal motion to respiratory parameters derived from an external surrogate signal that can be measured during data acquisition and treatment delivery. The models offer a number of advantages over current methods of imaging and analysing respiratory motion, in particular their ability to account for variations in the respiratory motion.

Computer Tomography (CT) data is acquired over several respiratory cycles to sample some of the variation in the respiratory motion. B-spline registrations are used to recover the motion and deformation from the CT data. The models are then constructed by fitting functions that relate the registration results to the respiratory parameters. This thesis describes the CT data and respiratory parameters that have been used to construct the motion models. It details the registrations protocols used and evaluates their results. The initial models presented in the thesis relate the registration results to a single parameter, the phase of the respiratory cycle, and average out any variation in the respiratory motion. The later models relate the registration results to two respiratory parameters, with the intention of modelling some of the variation. A number of different functions are assessed for both the single and two parameter models. The results show that the models can predict the respiratory motion in the CT data very accurately (mean error < 1.4 mm). This thesis also discusses some of the uses of the motion models in RT and, in particular, explores the use of the motion models for 'tracking' respiratory motion while delivering intensity modulated RT.

# Acknowledgements

I would like to thank the following people for their help over the last few years. Dave Hawkes, my supervisor, for his continued support and encouragement over the last four years. Adam Chandler, Jane Blackall, Segolene Tarte, and Gang Gao, the other members of the 'lung motion team' from CMIC, and Shahreen Ahmad, Simon Hughes, and David Landau, the members of the 'lung motion team' from Guy's and St. Thomas' Hospitals. You have all provided invaluable help throughout my PhD, and more importantly it has been great fun working with you.

Steve Webb and Dualta McQuaid for their help and input in the IMRT tracking work, and thanks to them and the rest of the 'lung motion team' at the Institute of Cancer Research and Royal Marsden Hospital for some extremely useful and stimulating discussions.

Norman Smith, Ivan Meir, Gideon Hale, James Turner, and everyone else at Vision RT for providing us with their camera system and continually supporting us using (and breaking) it, and of course, for partially funding my PhD.

All of the radiographers and medical physicists at Guy's and St. Thomas' who have spent many hours acquiring images for us, and especially to Matthew Adams the head radiographer, who spent much of his own time operating the scanner for us while we developed our acquisition protocols. And a special thanks to all the patients who volunteered to be part of our various studies.

Everyone else at Centre for Medical Image Computing and the Imaging Sciences Group, KCL, for making me feel welcome over the last four years.

All my friends and family, who make my life outside of research so enjoyable, and finally an extra special thanks to my Mum for looking after me and putting up with me all this time – without your support I'd never have got this (or much else) done.

# Table of contents

# List of figures

# List of tables

# Acronyms

| | |
|---|---|
| 1D | One Dimensional |
| 2D | Two Dimensional |
| 3D | Three Dimensional |
| 4D | Four Dimensional |
| 5D | Five Dimensional |
| CT | Computed Tomography |
| 4DCT | Four Dimensional Computed Tomography |
| MRI | Magnetic Resonance Imaging |
| PET | Positron Emission Tomography |
| CBCT | Cone-Beam Computed Tomography |
| RT | Radiotherapy |
| LINAC | Linear Accelerator |
| OAR | Organs At Risk |
| IMRT | Intensity-Modulated Radiation Therapy |
| MLC | Multi-leaf Collimators |
| DMLC | Dynamic Multi-Leaf Collimator |
| IMB | Intensity Modulated Beam |
| S-I | Superior-Inferior |
| A-P | Anterior-Posterior |
| L-R | Left-Right |
| MV | Mega-voltage |
| kV | Kilo-voltage |
| mm | Millimetre |
| mAs | Milliamp Seconds |
| GHz | Gigahertz |
| IR | Infra-Red |
| PCA | Principal Components Analysis |
| CP | Control Point |
| CPG | Control Point Grid |
| SSD | Sum of Squared Differences |
| CC | Cross Correlation |
| MI | Mutual Information |

| | |
|---|---|
| NMI | Normalised Mutual Information |
| LUL | Left Upper Lobe |
| RUL | Right Upper Lobe |
| RH | Right Hilum |
| RLL | Right Lower Lobe |
| TLE | Target Localisation Error |
| TRE | Target Registration Error |
| TEBR | Target Error Before Registration |
| TME | Target Modelling Error |
| DFE | Deformation Field Error |
| CPU | Computer Processing Unit |
| GPU | Graphics Processing Unit |
| T | Tumour |
| IL | Ipsilateral Lung |
| OT | Other Tissue |
| BH | Breath-Hold |
| FB | Free Breathing |

# List of publications

**Peer-reviewed journal papers**

**2007**

J. R. McClelland, S. Webb, D. M. Binnie, and D. J. Hawkes, "Tracking 'differential organ motion' with a 'breathing' multileaf collimator: magnitude of problem assessed using 4D CT data and a motion-compensation strategy," *Physics in Medicine and Biology* 52, p4805-4826, 2007.

**2006**

J. R. McClelland, A. G. Chandler, J. M. Blackall, S. Tarte, S. Hughes, S. Ahmad, D. Landau, D. J. Hawkes, "A Continuous 4D Motion Model from Multiple Respiratory Cycles for Use in Lung Radiotherapy," *Medical Physics* 33 (9), p3348-3358, 2006.

J. M. Blackall, S. Ahmad, M. E. Miquel, J. R. McClelland, D. B. Landau, D. J. Hawkes, "MRI Based Measurement of Respiratory Motion Variability and Assessment of Imaging Strategies for Radiotherapy Planning," *Physics in Medicine and Biology* 51 (17), p4147-4169, 2006

**2005**

D. J. Hawkes, D. Barratt, J. M. Blackall, C. Chan, P. J. Edwards, K. Rhode, G. P. Penney, J. R. McClelland, D. L. G. Hill, "Tissue Deformation and Shape Models in Image-guided Interventions: A Discussion Paper." *Medical Image Analysis* 8(2), p163-175, 2005.

**Peer-reviewed conference papers**

**2007**

J. R. McClelland, S. Webb, D. J. Hawkes, "Tracking 2D 'differential organ motion' using the DMLC technique and 4DCT motion models for the delivery of IMRT to sites affected by respiratory motion," *Proceedings of the 15th International Conference on the use of Computers in Radiation therapy (ICCR) 2007*, p 123-127, 2007. (Oral presentation)

S. Webb, D. Binnie, J. R. McClelland, "General formulation of the problem of 'differential organ motion', magnitude of problem assessed using 4D CT data and motion-compensation strategy suitable for a 'breathing' multileaf collimator," *Proceedings of the 15th International Conference on the use of Computers in Radiation therapy (ICCR) 2007*, p118-122, 2007.

**2005**

J. R. McClelland, A. G. Chandler, J. M. Blackall, S. Ahmad, D. Landau, D. J. Hawkes, "4D Motion Models over the Respiratory Cycle for use in Lung Cancer Radiotherapy Planning." *Proceedings of SPIE Medical Imaging 2005: Visualization, Image-Guided Procedures, and Display*, 5744, p173-183, 2005. (Oral presentation)

**2004**

J. R. McClelland, A. G. Chandler, J. M. Blackall, S. Ahmad, D. Landau, D. J. Hawkes, "Feasibility study of a novel technique for constructing respiratory motion models, for use in 4D lung cancer radiotherapy planning." *Proceedings of Medical Image Understanding and Analysis (MIUA) 2004*, p220-223, 2004. (Poster presentation)

**Conference abstracts**

**2007**

J. R. McClelland, J. M. Blackall, S. Tarte, S. Hughes, D. J. Hawkes, "Non-rigid registration based respiratory motion models of the lung using two parameters," *Medical Physics* 34, p2516, 2007.

(Oral presentation at the 49th Annual Meeting of the American Association of Physicists in Medicine (AAPM), 2007.)

R. Kashani, J Balter, M. Kessler, M. Hub, L. Dong, L. Zhang, L. Xing, Y Xie, D. Hawkes, J. Schnabel, J. McClelland, S. Joshi, "Objective assessment of deformable image registration in radiotherapy – a multi-institution study," *Medical Physics* 34, p2545, 2007.

K. K. Brock and on behalf of the Deformable Registration Accuracy Consortium, "A multi-institution deformable registration accuracy study," *International Journal of Radiation Oncology, Biology, Physics* 69, pS44, 2007.

**2006**

J. R. McClelland, S. Hughes, S. Ahmad, A. G. Chandler, J. M. Blackall, S. Tarte, D. Landau, D. J. Hawkes, "Comparing Breath-Hold and Free Breathing CT Derived Motion Models for Predicting Respiratory Motion of Lung Cancer and Lymph Nodes," *Radiotherapy and Oncology* 81 S1, pS213, 2006.

(Poster presentation at European Society for Therapeutic Radiology and Oncology (ESTRO) 25, 2006.)

S. Tarte, J. R. McClelland, S. Hughes, J. M. Blackall, D. Landau, D. J. Hawkes, "A Non-Contact Method for the Acquisition of Breathing Signals that Enable Distinction Between Abdominal and Thoracic Breathing," *Radiotherapy and Oncology* 81 S1, pS209, 2006.

S. Ahmad, S. Tarte, D. Withers, D. J. Hawkes, J. M. Blackall, D. Landau, S. Hughes, J. R. McClelland, A. G. Chandler, M. Adams, J. Boutland, "Comparison of Patient-Specific and Generic Internal Target Volume (ITV) Margins for Planning Radical Radiotherapy in Lymph-Node Positive Non-Small Cell Lung Cancer (NSCLC)," *Radiotherapy and Oncology*, 81 S1, pS384, 2006.

J. R. McClelland, A. Chandler, J. M. Blackall, S. Hughes, S. Ahmad, S. Tarte, D. Landau, D. J. Hawkes, "A Novel Technique for Generating a Patient-Specific 4D Model of Respiratory Motion and Its Potential Application to Radiotherapy Planning For Lung Cancer," *Clinical Oncology* 18(2):S3, 2006.

(Oral presentation at Advances in Lung Imaging conference organised by the Institute of Physics and Engineering in Medicine, 2006.)

S. Hughes, S. Ahmad, J. R. McClelland, A. Chandler, J. M. Blackall, S. Tarte, D. J. Hawkes, D. Landau, "The Respiratory Movement of Upper Lobe Tumours and Mediastinal Lymph Node Regions in Patients with NSCLC," *Clinical Oncology* 18(2):S3, 2006.

**2005**

J. R. McClelland, J. M. Blackall, S. Hughes, S. Ahmad, D. Landau, D. J. Hawkes, "4D Models Of Respiratory Motion for Use In RT Planning and Their Advantages Over 4DCT Techniques," *Radiotherapy and Oncology* 76 S2, pS93, 2005.

(Oral presentation at European Society for Therapeutic Radiology and Oncology (ESTRO) 24, 2005.)

S. Ahmad, J. R. McClelland, A. Chandler, J. M. Blackall, S. Hughes, D. Hawkes, D. Landau, "Movement of Tumour And Mediastinal Lymph Nodes With Respiration In Patients With Non Small Cell Lung Cancer (NSCLC)," *Radiotherapy and Oncology* 76 S2, pS114, 2005.

# Chapter 1

# Introduction

Respiratory motion is a significant problem for many medical imaging applications. It can cause artefacts when reconstructing images and can be the source of errors and uncertainties in image directed therapies and image guided interventions. One area where respiratory motion can be particularly problematic, and where it has been the focus of much research in the last few years, is the delivery of Radiotherapy (RT) to sites in the thorax and abdomen, e.g. the lung. This PhD has been primarily focussed on creating models of the respiratory motion for use in planning and delivering RT treatment to lung cancer patients, as this is one of the applications that may benefit the most from accurate and detailed models of the respiratory motion. However, it is expected that many of the findings of this research, and the methodologies developed as part of it, will potentially be of interest and use in a wide range of applications in image acquisition, image guided interventions (of which RT is just one), and studies of lung physiology.

The respiratory motion can be thought of as being 'quasi-periodic'. It is generally thought that the motion during most respiratory cycles is similar (unless the subject is undergoing erratic breathing, such as couching), although it is also known that there can be short and long term variations in the respiratory motion. Furthermore, it is generally thought that the internal motion can be related to "external" respiratory surrogate signals that can be measured during image acquisition and treatment delivery, such as the displacement of the chest or abdomen. There are many different surrogate signals that could be used including internally measured signals such as the position of the diaphragm.

Advances in imaging technology over the last five to ten years, in particular the development of multi-slice Computed Tomography (CT) scanners, have enabled the acquisition of three dimensional (3D) volumetric data at a high enough temporal resolution to image respiratory motion. However, current CT scanner technology is unable to simultaneously scan the entire region of interest, usually the entire thorax for planning lung RT. This means it is necessary to acquire data at different locations (or from different projections) during different respiratory cycles, and to sort the data according to a respiratory parameter (derived from a surrogate signal) into respiratory correlated, or Four Dimensional Computed Tomography (4DCT), volumes. This process relies on there being a relationship with the chosen respiratory parameter, at least for the duration of the image acquisition. Any variation in the motion cannot be accounted for and may produce artefacts in the 4DCT volumes.

This thesis presents a method of modelling the respiratory motion that uses the unsorted 12-16 slice 'sub-volumes', referred to in this thesis as Cine CT volumes (as they are acquired with the scanner in Cine mode) rather than 4DCT volumes that have already been sorted according to a respiratory parameter. This allows variations in the respiratory motion to be measured and accounted for in the models. The models can be used to analyse different relationships with different respiratory parameters and can average out variation that is not explained by the parameters. The models are fully continuous over one or two of the respiratory parameters. The models are also fully continuous in the three spatial dimensions as they are based on 3D non-rigid registrations.

3D image registration is the task of finding correspondence between two different 3D volumes, and is a very useful tool in a large number of medical imaging applications.

When the images are from the same subject image registration can be used to recover the motion and deformation that has occurred between the images. This means that image registration can potentially be used to determine the local 3D motion and deformation that occurs in the lungs and the surrounding anatomy due to respiration.

The main aim of this thesis was to investigate the hypothesis that the respiratory motion of the internal anatomy can be predicted from an external respiratory surrogate signal. This was broken down into three more specific aims:

- Use non-rigid registration to determine the 3D motion that occurs between volumes sampling the respiratory motion.
- Develop motion models that relate the internal motion (encoded by the registration results) to a single respiratory parameter, modelling the 'average' respiratory motion, and investigate the ability of the models to predict the internal motion.
- Further develop the motion models to relate the internal motion to two respiratory parameters, potentially allowing some of the variation to be modelled, and investigate the ability of the models to predict the internal motion.

An additional aim of this thesis was to:

- Demonstrate how the motion models developed above could be used to help account for respiratory motion when delivering RT treatment to lung tumours.

The work presented in this thesis has made a number of contributions that help to address the problem of respiratory motion in lung RT:

- A methodology for constructing patient specific respiratory motion models has been developed. These models relate the internal motion, described by non-rigid registrations, to an external respiratory signal that can be measured during data acquisition and RT treatment.

- It has been demonstrated that the models can accurately predict the internal motion from the external surrogate signal. The final models assessed as part of this thesis were found to have a mean error less than 1.4 mm when assessed using manually identified landmark points.

- It is possible to produce 'artefact free' predicted volumes using the motion models. These volumes do not contain any discontinuities at the boundaries between adjacent couch positions frequently seen in 4DCT volumes, and so should provide a better estimate of the true appearance of the anatomy.

- The performance of different respiratory parameters and modelling functions has been investigated.

- The potential of using the motion models to deliver RT treatment that 'tracks' the respiratory motion has been demonstrated.

- Appropriate data acquisition and non-rigid registration protocols have been developed.

There now follows a brief overview of the contents of each of the chapters in this thesis.

Chapter 2 provides background information and reviews the current literature. It contains sections on: the problem of respiratory motion in lung RT and some potential methods of addressing respiratory motion in lung RT, methods of imaging respiratory motion, respiratory surrogate signals and parameters, image registration, modelling respiratory motion, and possible uses for the motion models outside of lung RT.

Chapter 3 provides a detailed description of the data that has been acquired for the motion models and the methods and protocols that have been used to acquire it. This includes details of the CT data that has been acquired to image the internal motion, and the respiratory surrogate signals and the respiratory parameters that have been calculated from them.

Chapter 4 describes the non-rigid registrations that have been performed to recover the respiratory motion, and evaluates the registration results that are used to construct the motion models in the following chapters.

Chapter 5 describes how to construct motion models that relate the internal motion to a single respiratory parameter, the phase of the respiratory cycle. It evaluates a number of different potential functions that can be used to describe the relationship between the phase and the internal motion. It then assesses the ability of the motion models to predict the Cine CT data. Finally it explains how the model predictions from different couch positions can be combined into a single continuous prediction over the whole region of interest.

Chapter 6 explains how to construct motion models that relate the internal motion to two respiratory parameters. It describes several example functions that can be used in the two parameter models, and evaluates their performance both against the registration results and the Cine CT data.

Chapter 7 describes some of the ways the motion models can be used for lung RT, and details two experiments performed as part of this research that demonstrate the use of the motion models: the comparison of breath-hold and free-breathing motion models, and the delivery of tracked IMRT (Intensity-Modulated Radiation Therapy) treatment.

Chapter 8 summarises and discusses the findings and conclusions from this research and gives pointers to future work

# Chapter 2

# Background and Literature review

## 2.1 Introduction to problem

### 2.1.1 Lung cancer

Lung cancer is the leading cause of cancer-related death worldwide. It has a "death to incidence" ratio of 0.87 and was responsible for 1.18 million deaths worldwide in 2002. For men it is the most common cause of cancer-related death and for women the second most common, behind breast cancer (Parkin et al. 2005). There are approximately 8 male cases for every 5 female cases diagnosed each year, and approximately 50 new cases per hundred thousand population each year. Incidence increases with increasing age, with the most prevalent age group being 75+.The method of treatment for lung

cancer will depend on histological diagnosis, stage of disease, patient fitness, and patient choice. Treatments can involve combinations of surgery, chemotherapy, radiotherapy, and supportive care.

### 2.1.2 Radiotherapy

This research is concerned with Radical Radiotherapy (RT) for non-small cell lung cancer, as this is the form of lung cancer treatment that may benefit most from having greater knowledge of motion and deformation that occurs due to respiration. RT is the use of electromagnetic radiation to kill cancer cells by causing irreparable damage to their DNA. Radical RT involves treating the tumour with high dose radiation, with the intent of curing the patient or providing long term control of the disease. The radiation is delivered in the form of a beam of megavoltage x-rays from a machine known as a Linear Accelerator, or LINAC. Although all of the anatomy in the path of a RT beam will receive some radiation dose, by using beams from two or more different directions it is possible to create regions of high dose around the target. Conformal radiotherapy attempts to make the region of high dose conform to the shape and location of the target (the tumour and other areas requiring RT treatment such as involved mediastinal nodes), so as to limit the dose delivered to Organs At Risk (OARs, e.g. the lungs, the heart, the spinal cord, etc.). One way the beam can be shaped to conform to the tumour is to use Multi-leaf Collimators (MLCs) located in the head of the LINAC. MLCs are a series of opposing leaf pairs (each 5-10mm wide) that can be adjusted to shape the RT beam.

MLCs can also be used to deliver Intensity Modulated Radiotherapy (IMRT). This goes beyond conformal RT by varying the amount of dose delivered at different points in the beam, allowing for much more complex dose distributions. IMRT can be delivered in two ways, using the step-and-shoot method where the beam is switched off while the MLCs move to another position, and using dynamic delivery where the (opening in the) leaves sweep across the beam, with the time between the leading (opening) and trailing (closing) leaves determining the dose delivered at each point (Webb 2000).

### 2.1.3 The problem of respiratory motion

The current clinical practice for planning and delivering RT treatment for lung cancers in most UK institutions is to first acquire a Computed Tomography (CT) scan for planning. This can provide full 3D information on the location of the anatomy, and contains the x-ray attenuation values that are needed for the RT dosimetry calculations. The standard CT scanning protocol used for RT planning acquires the data in helical mode while the patient is free breathing, as they will be during RT treatment. It is hoped that the scan will represent the average location of the tumour and other anatomy. However, several groups have demonstrated that the inter-play that can exist between the motion of the scanner gantry during acquisition and the respiratory motion of the anatomy can cause large errors and artefacts in the planning CT scans (Balter et al. 1996, Rietzel et al. 2005a, Shimizu et al. 2000). This means that not only is the location of the anatomy in the image unlikely to be its average location over the respiratory cycle, but that its size and shape can be completely misrepresented, particularly for mobile lung tumours.

The target and OARs are delineated on the planning scan. As lung tumours may exhibit some respiratory (and possibly cardiac) motion, and there can be errors associated with setting up the RT treatment, it is necessary to add margins which enlarge the region of high dose in order to avoid a geographic miss and ensure the intended target is not under-dosed during the treatment. The use of such margins also increases the dose received by healthy tissue surrounding the intended target, and hence may reduce the dose that can be safely delivered to target.

The planning scan does not provide any information on the target motion that may occur, so standard margins (ICRU report 50 1993, ICRU report 62 1999) need to be added to account for motion and setup errors. However, it has been shown using 4D imaging techniques that the standard margins added do not usually represent the actual respiratory motion very well (Alasti et al. 2006, Allen et al. 2004, Lagerwaard et al. 2001, Underberg et al. 2004). This can lead to both under-dosing of the target and unnecessary dosing of healthy tissue, even for the same patient. There is evidence that the delivery of higher RT doses to the target will result in an improved survival for non-small cell lung cancer patients (Kong et al. 2005), but the amount of dose that can be delivered to the target is limited by the amount that the OARs can tolerate (Graham et al. 1999, Werner-Wasik et al. 2002).

Once the target and the OARs have been delineated candidate plans are generated. Dosimetry calculations are performed to determine the dose received by the target and the OARs, and this information is used to assess the candidate plan and decide if it is acceptable. Dosimetry calculations that are performed on the 'static' 3D planning scans will differ from the actual dose delivered in the presence of respiratory motion (Bortfeld et al. 2004, Flampouri et al. 2006). This may cause the dose delivered to the tumour to be overestimated and the dose delivered to the OARs to underestimated.

For more information see the report of AAPM Task Group 76 ("The management of respiratory motion in radiation oncology", Keall et al. 2006) which gives a good concise description of the different problems that respiratory motion can cause for RT planning and treatment delivery. It also summarises many studies that have measured respiratory motion, and explains and discusses several methods of accommodating or compensating for respiratory motion in RT treatment (see below for further discussion of these topics).

### 2.1.4 Magnitude of the problem

There have been many studies to measure the respiratory motion of lung tumours using several different modalities (see Section 2.2 for a description of the different modalities used to image respiratory motion). The report of AAPM Task Group 76 (Keall et al. 2006) summarises the results of many of these studies. For the majority of these studies (most of which were based on 10-30 patients) the mean tumour motion in at least one direction is greater than 5mm, and the maximum tumour motion is greater than 10mm, with one study reporting tumour motions of up to 50mm (Chen et al. 2001). Similar results have been reported by more recent studies not included in the report referenced above (Britton et al. 2007, Guckenberger et al. 2007, Mori et al. 2007, Sonke et al. 2008).

Overall it can be said that lower lobe tumours are likely to move more than middle lobe tumours, which are likely to move more than upper lobe tumours, but there is considerable variation and there are examples of upper lobe tumours that move more than lower lobe tumours (Seppenwoolde et al. 2002, van Sörnsen de Koste et al. 2003). Likewise, tumours are likely to exhibit more motion in the Superior-Inferior (S-I) direction than in the Anterior-Posterior (A-P) direction, and are likely to exhibit more motion in the A-P direction than the Left-Right (L-R) direction, but tumours that exhibit

considerable (> 5mm) L-R motion and/or negligible S-I motion have been reported (Table I, Keall et al. 2006). It has been shown that there are no statistically significant correlations between the extent of the tumour motion and patient characteristics such as: weight, height, age, stage of disease, tumour size, degree of chest wall motion, or pulmonary function (Giraud et al. 2001, Stevens et al. 2001).

Hysteresis is the term commonly used to mean the respiratory motion follows a different trajectory during inhalation than it does during exhalation. Most studies on tumour motion have reported hysteresis for some patients, and separations of up to 5mm between the inhalation and exhalation trajectories have been observed (Seppenwoolde et al. 2002).

### 2.1.4.1 Inter-cycle variation

It is known from physiological studies of airflow and other measurements (Benchetrit 2000) that although individuals tend to have characteristic breathing patterns, there can also be considerable short and long term inter-cycle (breath to breath) variations. As RT treatment is typically performed over several minutes and is repeated for several fractions over the course of a few weeks, both the short term and long term variations in respiratory motion will be of interest. In addition to the long term variation seen in healthy subjects, further changes to the respiratory motion may be seen in patients due to the effects of the RT treatment and changes to the pathology. There can be variations in both the period of the respiratory cycle and the internal motion path. Generally variations in the internal motion path are of more concern than variations in the period for most RT treatments, although variations in the period may well produce corresponding variations in the internal motion path.

Seppenwoolde et al (2002) studied both short and long term variation in the respiratory motion using implanted gold markers and x-ray imaging. They found that both long and short term shifts and drifts in the average tumour location could occur. However, the shapes of the tumour trajectories were fairly constant throughout the duration of treatment for 19 out of 20 patients. Small short term variations in the trajectories did occur and the end-exhale location was found to be more reproducible than the end-inhale location. Sonke et al. (2008) studied long term variation in 56 patients using 4D Cone-Beam CT. They also found that the shape of the tumour trajectories was stable,

with mean variability not exceeding 1mm, but there were again large baseline variations (shifts in the average tumour location). Britton et al. (2007) studied long term variation in 8 patients using 4DCT and found that large variations in both tumour size and mobility occurred. The plots they produce of the 3D tumour trajectories indicate that they are not as stable as found by Seppenwoolde et al. (2002) and Sonke et al. (2008).

Guckenberger et al. (2007) studied short term variation in 10 patients (with a total of 14 tumours) using 4DCT, acquiring 4 scans over 30 minutes. Note: the patients were immobilised in a stereotactic body frame and abdominal compression was applied to 7 patients in order to limit breathing. They concluded that for most patients short term variation was relatively small, with a drift in the mean tumour location greater than 3 mm and 5 mm observed for five and two of the tumours respectively. One patient with poor pulmonary function displayed a continuous increase in respiratory motion from 17 mm to 28 mm throughout the 30 minutes.

The above studies indicate that inter-cycle variation can have a large effect on respiratory motion, particularly between different fractions of RT treatment. It should be noted that most of these studies have been published recently, and that measuring and analysing inter-cycle variation is a very active area of research. Several groups have proposed the use of audio, visual, or audio-visual feedback to try and reduce the inter-cycle variation (George et al. 2006, Kini et al. 2003, Mageras et al. 2004, Neicu et al. 2006). There have been mixed results from using such devices, but it is generally agreed that with further development such devices will help to regularise the respiration of some patients.

### 2.1.5 Treating tumours affected by respiratory motion

There are a number of approaches to treating lung tumours that are subject to respiratory motion, including: treating at breath hold, accurate target delineations and dose calculations, gated treatment, and tracked treatment.

Treating at breath hold literally stops the respiratory motion, although there may still be some residual motion due to muscle relaxation. Although treating at (deep inspiration) breath hold has been shown to have considerable dosimetric advantages, it is not suitable for all patients, can increase treatment time, and relies on the reproducibility of

the breath hold position (Hanley et al. 1999, Mah et al. 2000). The reproducibility of the breath holds can be improved by use of the Active Breathing Control device which can enforce a breath hold at the same level of inspiration each time (Dawson et al. 2001). When treating at breath hold knowledge of the respiratory motion is not necessary, although an estimation of the breath hold reproducibility will aid in defining appropriate treatment margins.

If the tumour motion can be imaged then a target can be delineated that encompasses the tumour during the entire respiratory cycle, rather than adding standard margins that may not reflect the motion. Alternatively, it has been proposed that if the average tumour position can be located accurately then only relatively small margins are required to account for the respiratory motion (van Herk 2007). Additional margins will still be required to account for variations in the motion and to account for setup errors. Defining the encompassing or average target may involve delineating the tumour in several 3D volumes acquired throughout the respiratory cycle (e.g. a 4DCT dataset, see Section 2.2.2.3). If the displacement of the tumour from one volume to another is known (from registering the volumes, see Section 2.4) then the tumour only needs to be delineated in one volume and can be automatically propagated to the others (Weiss et al. 2008). If the displacement of all of the anatomy is known then it is also possible to calculate dose distributions that account for the effects of respiratory motion (Bortfeld et al. 2004, Flampouri et al. 2006, Heath et al. 2006, Keall et al. 2005, Pagnetti et al. 2004, Rietzel et al. 2005b, Rosu et al. 2005).

The idea of respiratory gating is to deliver the RT treatment when the tumour is at a known location, and to stop the delivery when the tumour moves away from this location (Berbeco et al. 2005, Seppenwoolde et al. 2002, Shirato et al. 2000, Tsunashima et al 2004, Vedam et al 2001, Zhang et al. 2003). The idea is similar to treating at breath hold and offers similar dosimetric benefits, but the patient is freely breathing, so the tumour motion must be followed during treatment. As the patient does not actually stop breathing during gated RT there will be some residual motion when the RT is being delivered. There will be a trade-off between the amount of residual motion and the time that the beam is switched off for (known as the duty cycle). The less residual motion the greater the dosimetric benefits of gated treatment, but the smaller the duty cycle the greater the increase to the treatment time (Vedam et al. 2001). Although there has been lots of research into gated RT there is still debate over its

potential benefits (van Herk 2007) and whether it is ready for clinical use (X. Li and Keall 2007).

Tracked RT treatment can be considered the most ambitious way to treat respiratory motion. The idea of tracked treatment is to follow the tumour motion with the RT beam during treatment, hopefully offering similar dosimetric benefits to gated or breath hold treatment, but with no associated increase in treatment time (Murphy et al. 2004). There are two main methods of delivering tracked treatment that have been proposed in the literature. The first is to robotically mount the LINAC and physically move it so that it follows the tumour (Schweikard et al 2000). The second is to use a Dynamic Multi-Leaf Collimator (DMLC), as used to deliver dynamic IMRT (Intensity Modulated RT), to make the shape of the beam follow the tumour (Keall et al 2001). A DMLC can be used to track the respiratory motion during the delivery of both Conformal RT (Keall et al. 2005) and IMRT (Keall et al. 2001,). Although there are commercial products available based on both of these tracking technologies (Cyberknife with Synchrony Respiratory Tracking System, http://www.accuray.com/Products/Synchrony/index.aspx, Accuray, Sunnyvale, California, USA, and TrackBeam, http://www.initiamed-rt.com/htm/sub-trackbeam.htm, Initia-RT, Petah Tikva, Israel), tracked delivery is still very much a research area, and more work is required both on developing the methodologies and verifying the accuracy of such systems before they can enter routine clinical use (Robotically mounted LINAC: Schweikard et al 2004, Seppenwoolde et al 2007, DMLC tracking: Alasti et al. 2006, McQuaid and Webb 2006, Neicu et al 2003, Papiez et al. 2005).

To plan gated or tracked treatment the respiratory motion of the tumour, and ideally the rest of the anatomy, must be determined prior to treatment. If the residual motion is small for gated treatment it may be possible to accurately estimate the dose distributions based on static anatomy (Flampouri et al 2006), but for tracked treatment the respiratory motion will need to be accounted for if accurate dose distributions are to be calculated (Keall et al. 2005). To deliver gated or tracked RT treatment it is necessary to know the location of the tumour during treatment delivery. This can be achieved by imaging the tumour during treatment, or by constructing a model that relates the internal motion to an external respiratory surrogate signal (see Section 2.3). Imaging the tumour directly during treatment is very difficult so implanted markers are often used to follow its motion, although their use has a number of associated problems (see Section 2.2.1). The use of external surrogate signals also has a number of problems, the most important

being establishing a stable relationship between the internal motion and the surrogate signal, and ensuring that the relationship is valid throughout treatment (see Section 2.3). The Synchrony system used with the Cyberknife (Accuray, Sunnyvale, California, USA) combines an external surrogate signal with infrequent internal imaging. This allows the relationship between the internal motion and the surrogate signal to be assessed and corrected during treatment (Schweikard et al 2000, Seppenwoolde et al. 2007).

However the tumour location is determined there will be some delay associated with the system, meaning that the location that has just been determined is not the tumour's current location, but was its location a short time earlier. Therefore, techniques that can predict the future tumour location (or surrogate values) are likely to be required for gated and tracked treatments to be delivered in a timely manner (Sharp et al. 2004, Vedam et al. 2004).

## 2.2 Imaging Respiratory Motion

In order to study, and ultimately model, the respiratory motion of lung tumours and the surrounding anatomy, it is necessary to acquire images that give us some information about the motion that occurs. Three imaging modalities have been widely used to study respiratory motion in the literature: projection X-ray imaging, Computed Tomography (CT), and Magnetic Resonance Imaging (MRI).

### 2.2.1 Projection X-ray imaging

Dynamic projection x-ray imaging (such as fluoroscopy), from here on simply referred to as x-ray imaging, is a popular choice for imaging respiratory motion of lung tumours and other anatomy as the images can be acquired with a very high temporal resolution (up to 30Hz.). Another advantage of x-ray imaging over other modalities is that it can be acquired during treatment, but x-ray imaging does have some serious limitations.

It can be very difficult to accurately identify the lung tumours directly on the x-ray images, although some groups have managed to do this manually (Hoisak et al. 2004, Tsunashima et al. 2004), and a semi-automatic method has been proposed (Schweikard et al. 2004). Some groups have implanted small radio-opaque markers close to or in the tumour in order to follow its motion (de Mey et al. 2005, Ozhasoglu and Murphy 2002, Schweikard et al. 2000, Shirato et al. 2000). These markers are easily detectable in x-ray images, and can be automatically tracked. Although inserted markers can greatly aid the real-time tracking of tumours, they have a number of disadvantages, including the invasiveness of the insertion procedure, and the possibility of the markers moving away from their intended locations (de Mey et al. 2005, Shirato et al. 2003).

Another disadvantage of x-ray imaging is that it produces 2D projections data, so only 2D motion can be measured from one view. 3D motion can be deduced by obtaining data simultaneously from two or more views, but this is only possible for points which can be easily identified in both x-ray images, such as implanted markers (Schweikard et al. 2000, Shirato et al. 2000). It is very difficult to obtain information on the full 3D anatomy and its motion from X-ray imaging, although recently Zeng et al. (2007a) have

proposed a method that uses slowly rotating x-ray projections (from a Cone-Beam CT system) to estimate the full 3D respiratory motion.

A final disadvantage of x-ray imaging is that it necessitates an extra dose of radiation being delivered to the patient. The amount of dose will depend on the system being used (in particular how many views are acquired simultaneously) and the length of time for which the x-ray imaging systems are switched on. The dose delivered due to continuous x-ray imaging throughout treatment has been estimated to be in the order of 1% of the treatment dose and mostly delivered to the skin (Shirato et al. 2000).

### 2.2.2 CT imaging

Computed Tomography (CT) uses x-ray projection data acquired from rotating views to reconstruct the internal anatomy in 3D. CT scanning is usually performed with a helical acquisition scheme, where the CT detector mounted on a gantry is constantly rotating as the CT couch moves through the CT bore, thus imaging all of the desired anatomy. CT scanners were originally single-slice with a single row of detectors acquiring one slice of the anatomy at a time. When this research commenced multi-slice CT scanners, capable of acquiring 12-16 slices simultaneously (giving a 2-3 cm coverage) with gantry rotation times of approximately 0.5 seconds, were already being used for clinical practice. These scanners could acquire standard 3D volumes much faster than the single-slice scanners, and made acquiring respiratory CT data much more feasible (see below).

Although 3D CT data is extremely useful for planning RT, as already noted standard helical scans can be seriously degraded by respiratory motion, and do not give any information on the motion that occurs. Sections 2.2.2.1 - 2.2.2.3 describe methods of acquiring CT data that gives some information on the respiratory motion. Section 2.2.2.4 discusses 'Cone-Beam' CT and its potential for imaging respiratory motion.

One problem with all CT scanning methods is that they deliver an extra dose of radiation to the patient. The exact amount of dose will depend on which type of CT scan is being performed (and the settings used for it), but even for 4DCT scans (see Section 2.2.2.3), which generally deliver more dose than other types of CT scan, the dose delivered is orders of magnitude smaller than the dose delivered from RT treatment. It

can be argued that the increased accuracy in RT treatment (and corresponding reduction in dose to healthy tissue) made possible by the CT scan(s) more than outweighs the extra dose required by the scan. However, the extra dose delivered by the CT limits the use of CT scanning, even for RT patients.

### 2.2.2.1 Slow CT

Slow CT techniques slow down the CT gantry rotation so that the whole respiratory cycle is captured in one rotation (Lagerwaard et al. 2001, Mori et al. 2006, van Sörnsen de Koste et al. 2003). This will produce a blurry image from which it may be possible to deduce information regarding the extent of motion for high intensity objects surrounded by low intensity backgrounds, as is the case for lung tumours and lung tissue. The method can have problems when the tumours are attached to or located near other high intensity structures such as the chest wall or mediastinum. In addition, the slower gantry rotation used for Slow CT can produce large artefacts due to inter-play effects between the gantry and tumour motion (Mori et al. 2006, Rietzel et al. 2005a). Slow CT scans only give information of the overall extent of the tumour motion so their uses are limited, but they may be a useful aid for target and margin definitions when more advanced techniques are not available.

### 2.2.2.2 CT acquired at breath-hold

Another simple method used for studying respiratory motion is to acquire standard helical CT scans at breath hold (Balter et al. 1996, Mah et al. 2000, Sarrut et al. 2006, Schweikard et al. 2004). Most modern commercial CT scanners are capable of acquiring a full volume covering the entire thorax, at a reasonably high resolution, within the time of one comfortable breath hold (15-20s). The anatomy during deep inspiration/expiration breath hold scans will clearly be in a different location to where it is during normal tidal breathing, although if the treatment is to be delivered at breath hold this is not a concern (Mah et al. 2000). However, it has also been shown that even at normal tidal breath holds, the location of the anatomy can differ from the location during free breathing (Blackall et al. 2006, Rietzel 2005a, and Section 7.3 of this thesis). Despite this breath hold scans can be useful for approximating the extent of respiratory

motion for all of the anatomy (not just the tumour as with Slow CT scans), and can again be an aid to target and margin definitions when free breathing information is not available.

### 2.2.2.3 Four Dimensional CT (4DCT)

Four Dimensional CT (4DCT) attempts to image the patient's anatomy at a number of different respiratory states while the patient is freely breathing. As it is not possible to image the entire region of interest (usually the entire thorax, but sometimes a smaller region around the tumour due to limitations with the scanner, reconstruction software, or patient dose) simultaneously with current CT scanners, it is necessary to acquire data from different parts of the anatomy at different times, and then sort this data into coherent volumes. Current 4DCT is based on the assumption that the respiratory cycle is reproducible, and that a simple respiratory surrogate signal (see Section 2.3) can be used to determine the respiratory state. The CT data and the respiratory surrogate signal are simultaneously acquired. The idea is to acquire CT data at each location being imaged over an entire respiratory cycle. The respiratory signal can then be used to sort the data into CT volumes covering the entire region of interest at different respiratory states. There are three different methods of acquiring 4DCT volumes that have been reported in the literature, which I will refer to as the 'helical' method, the 'stop-and-go cine' method, and the 'continuous cine' method.

The helical 4DCT method is similar to a standard helical CT scan, except that CT couch moves much more slowly so that projection data is acquired over an entire respiratory cycle (plus the gantry rotation time) at each location. This data is then sorted in sinogram space (i.e. the x-ray projection data is sorted) according to the respiratory signal, and CT volumes are reconstructed at the desired respiratory states from the appropriate projections. The spatial and temporal resolution of the data will depend on the scanners capabilities and the settings used. Any desired number of different respiratory states can be reconstructed, although there may be an overlap in the data used to reconstruct different respiratory states depending on the scanner settings, the desired number of respiratory states, and the respiratory period of the patient being scanned. The helical 4DCT method was originally proposed for a single slice scanner (Ford et al. 2003, Vedam et al. 2003). Even though the 4DCT volumes showed a

considerable reduction in motion artefacts compared to standard planning CTs for both phantom and patient data (Vedam et al. 2003), and were successfully used to study tumour motion in patients (Mageras et al. 2004), it was acknowledged that the spatial and temporal resolutions of the reconstructed images were limited due to using a single slice scanner. Keall et al. (2004) described an implementation of the helical method on a multi-slice scanner, and this method has since been adopted commercially and widely used.

The other two methods of acquiring 4DCT data use the scanner in cine mode. In cine mode the CT couch is stationary while data is being acquired. Data is acquired at one couch position, then the scanner stops acquiring data and moves the couch to an adjacent position, where it resumes acquiring data. Enough couch positions are used to fully cover the region of interest. At each couch position a series of Cine CT volumes (small sub-volumes covering a thin slab of data) are acquired. The number of slices and superior-inferior coverage of these volumes is limited by the number and size of the slices that the CT detector can acquire simultaneously. The Cine CT volumes are then sorted according to the respiratory signal, and concatenated to produce 4DCT volumes covering the entire region of interest at the desired respiratory states. At each couch position the Cine CT volumes can be acquired either using the continuous method (Pan et al. 2004) or the stop-and-go method (Low et al. 2003).

When using the continuous method the CT gantry is continuously rotating and acquiring data for the duration of one respiratory cycle (plus the gantry rotation time) at each couch position. The desired number of Cine CT volumes can then be reconstructed for each couch position and sorted into 4DCT volumes. As with the helical method, there may be an overlap in the data used to reconstruct temporally adjacent volumes. Note, the Cine CT volumes are usually reconstructed equally spaced over the acquisition time, but the start of the acquisition is not synchronised to the respiratory cycle so these volumes will actually represent slightly different respiratory states from each couch position. This may result in small 'binning errors' when sorting them into 4DCT volumes. These binning errors could be avoided by sorting the data in sinogram space prior to reconstruction (Pan et al. 2005, Rietzel et al. 2005a), as is done in the helical method, but this approach is not usually adopted, maybe because the binning errors are relatively small when enough Cine CT volumes are reconstructed. The continuous cine method has also been commercially adopted and widely used to study respiratory motion.

When using the stop-and-go method the CT gantry will perform a complete rotation acquiring enough data to reconstruct a volume. There will then be a brief period of 'dead-time', with the x-ray switched off and the gantry still, before the gantry performs another complete rotation (with the x-ray switched on) to acquire another volume. In order to sample the respiratory motion at a high enough temporal frequency data is acquired over a few respiratory cycles. Originally 15 Cine CT volumes were acquired over 11 seconds (Low et al. 2003), but this has been increased to 25 volumes over 18 seconds more recently (Lu et al. 2006). Note, there is a danger that if the length of the respiratory cycle is approximately an integer multiple of the time between Cine CT volumes, the same respiratory states will be sampled from each respiratory cycle, and the effective temporal frequency will be reduced. The stop-and-go cine method is not as efficient for acquiring 4DCT data as the other acquisition methods, as they only need to sample one respiratory cycle at each location. However, the stop-and-go method has still been successfully used in several studies, and having data from multiple respiratory cycles at each location permits the study of some inter-cycle variation using the unsorted Cine CT volumes (Low et al. 2005).

Pan et al. (2004) discuss the difference between the three 4DCT acquisition methods, and perform a more detailed comparison between the helical and continuous cine methods in Pan et al. (2005). He concludes that both methods have their advantages, but more of these are in favour of the continuous cine method. There are some problems common to all methods of acquiring 4DCT volumes. Firstly, they can deliver up to an order of magnitude more radiation than a standard planning CT scan (Li et al. 2005), but as already noted this is small in comparison to the RT dose and it may be possible to substantially reduce this dose using a deformation model (Li et al. 2005). The quality of the 4DCT volumes may be affected by the choice of respiratory surrogate signal and respiratory parameter used to sort the volumes (Lu et al. 2005a, Lu et al. 2006). The different signals and parameters are discussed in Section 2.3. Another drawback with 4DCT is that it assumes that all respiratory cycles are identical and that no inter-cycle variation occurs. Therefore the 4DCT volumes cannot be used to study inter-cycle variation, and can contain artefacts when it occurs.

Recently, a prototype CT scanner has been developed that can acquire 256 x 0.5 mm slices simultaneously, giving a much wider coverage (12.8 cm) at a finer spatial resolution than present scanners (Mori et al. 2006). Although the tumour's motion can be studied by acquiring data at a single couch position (Mori et al. 2007), multiple

couch positions will still be required for full dosimetry calculations, even with the extended coverage of this prototype scanner. However, given current trends it can be envisaged that a scanner capable of acquiring data over the entire lungs simultaneously may be available in the near future. Indeed, Cone-Beam CT systems already acquire data over most of the lungs simultaneously, albeit over multiple respiratory cycles due to a far slower gantry rotation time. Being able to acquire high resolution (both temporal and spatial) 3D data over the whole lungs would be extremely useful for the detailed study of inter-cycle variation.

In-slice motion artefacts can be present in all 4D or Cine CT volumes due to the finite time required to acquire enough projection data to reconstruct a slice (Lu et al. 2005b, Rietzel et al. 2005a). These artefacts can be reduced by using a faster gantry rotation time, employing half-scan reconstruction (Hui et al. 2000, Parker 1982), or possibly from combining data from multiple respiratory cycles in sinogram space (Lu et al. 2005b).

### 2.2.2.4 Cone-Beam CT

Cone-Beam CT (CBCT) uses onboard imaging equipment mounted on the RT treatment machine. This enables the patient to be imaged in position, just prior to treatment, enabling a much more accurate set-up and corresponding reduction in treatment margins (Sonke et al 2008). There are two types of CBCT imaging systems: those that use the LINAC itself to provide the x-rays (MV CBCT), and those that use a separate kV x-ray source usually mounted at right angles to the LINAC (kV CBCT). Both systems operate in a fairly similarly way, although the quality of the images produced and the extra dose required for the imaging can differ. CBCT uses large area detectors in comparison to standard diagnostic CT, with a field of view of 25cm x 25cm or larger. However CBCT machines have a much slower rotation time, of the order of 1 minute, so the projection data from different angles will be acquired from different respiratory states and during different respiratory cycles. If all the CBCT projections are used to reconstruct a 3D volume (as is standard practice for sites not effected by respiratory motion) then the anatomy effected by respiratory motion will appear blurred, similar to its appearance under Slow CT.

It has been demonstrated that CBCT projections can be sorted according to a respiratory signal and then used to reconstruct 4D CBCT volumes at different respiratory states (Sonke et al. 2005). As only some of the projections are used to reconstruct each 4D CBCT volume the quality of the reconstructed volumes will be noticeably degraded. This can be prevented by acquiring more projection data, either by using slower or multiple rotations, or by trying to combine projection data from different respiratory states using non-rigid registration based methods (Li et al. 2006a, Li et al. 2007, see Section 2.6.1) for more details of these methods). Zeng et al. (2007a) have recently proposed a method, also based on non-rigid registration, of deducing the respiratory motion directly from the CBCT projections.

### 2.2.3 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI) is an imaging modality based on the magnetic properties of protons (hydrogen nuclei) and their distribution in the different tissues inside the human body. It is can produce highly detailed 3D volumes of the patient's internal anatomy, and is an extremely versatile imaging modality which can acquire many different types of images. A major advantage of MRI over X-ray and CT imaging is that MRI does not deliver any extra radiation to the patients. This allows volunteers to be studied as well as patients, and means there are no (dose related) limitations on the number or length of scans that a subject can undergo.

Although new and more advanced MRI techniques are constantly being developed there remains an inherent trade off between the spatial resolution, the temporal resolution, the field of view, and the signal to noise ratio in MRI images. Because of this, until very recently, it has only been possible to acquire 2D slices at a high enough spatial and temporal resolution to study respiratory motion in the lung using MRI (Koch et al. 2004, Plathow et al. 2005, Sundaram et al. 2005). Colleagues from our group were able to acquire full 3D volumes over the lungs at a high enough temporal resolution to study respiration, but the images were not of sufficient quality to allow advanced analysis such as non-rigid registration (Blackall et al. 2006). Recent developments in MRI hardware and acquisition schemes have resulted in higher quality MRI images which should permit the detailed study of the non-rigid 3D motion due to respiration, and its variation (Miguel et al. 2007).

Other groups have employed a similar technique to that used for 4DCT, where high quality 2D slices are acquired at different respiratory states and locations over several respiratory cycles, and then sorted according to a respiratory surrogate signal (Remmert et al. 2007). von Siebenthal et al. (2007) have recently reported a more advanced sorting strategy based on acquiring a slice at a fixed location, the navigator slice, between every other slice. Although this paper focuses on imaging the liver the technique can be easily applied to the lungs, and some example lung images are shown in the paper. Up to five easily identifiable structures (veins in the liver) are identified in the navigator slice and are used to assess the similarity between the respiratory states in different data (non-navigator) slices. Using this technique coupled with long acquisition times (approximately 1 hour) they claim to be able to resort the data into coherent volumes that accurately represent the 3D respiratory motion and its inter-cycle variation over the acquisition period.

## 2.3 Respiratory surrogate signals and parameters

Respiratory surrogate signals are used to measure the respiratory parameters which (it is hoped) characterise the respiratory motion of the internal anatomy. These signals have two main uses in lung RT: firstly they are used during data acquisition to sort data into coherent respiratory states, and secondly they are used during the delivery of RT treatment to predict the internal motion when it is not possible to image it directly. There are many different signals and parameters that have been used and reported in the literature.

### 2.3.1 Respiratory surrogate signals

Spirometers are used to measure the volume (or flow) of air being inhaled and exhaled by the patient. Spirometry is a popular choice of respiratory surrogate signal as the signal is physiologically related to the respiratory motion and has historically been used for assessing respiratory performance and patterns (Benchetrit 2000). However, it has been reported that some patients can have difficulty tolerating spirometry for long periods of time (Hoisak et al. 2004), and that spirometry measurements can be subject to time dependent drifts of the end-exhale and end-inhale values due to escaping air and instrumentation errors (Hoisak et al. 2004, Low et al. 2003, Zhang et al. 2003). There have been several methods suggested to correct for this drift, although these either assume that there is no 'real' physiologically drift occurring (and such drifts have been observed in some patients, Seppenwoolde et al. 2002), or use another 'drift-free' surrogate signal to correct for the drifts (Lu et al. 2005a).

Another popular choice of respiratory surrogate signal is to measure the displacement of the patient's chest or abdomen. This is often done using one or more Infra-Red (IR) markers which are tracked optically, and there are a number of commercial system that use this technology: e.g. the Real-Time Position Management (RPM) system (Varian, Palo Alto, California, USA), the Cyberknife (Accuray, Sunnyvale, California, USA), and the Polaris system (NDI, Waterloo, Ontario, Canada). Other methods of tracking the displacement of the chest or abdomen include electromagnetic tracking systems (Hoisak et al. 2004) and laser tracking systems (Tsunashima et al. 2004). Specialised markers that are visible in x-ray images (Ozhasoglu and Murphy 2002), MRI images (Khamene

et al. 2004, Koch et al. 2004, Plathow et al. 2005), and Positron Emission Tomography (PET) images (Nehmeh et al. 2003) have also been used in the literature, but these are either for sorting image data or studying the relationship between the internal and external motion, and cannot be used to predict internal motion during RT treatment (but can easily be replaced by an IR skin marker or other type that can be measured during treatment).

When using any system that measures the displacement of one or more points on the patient's skin, care must be taken to ensure the reproducibility of the location of the point(s) being measured otherwise the signals obtained during different sessions may not be comparable. However, if the set-up is done reproducibly, and if a fixed reference such as the height of the couch is used, it may be possible to detect changes to the base-line of the signal between sessions, which are known to occur in the internal anatomy (Seppenwoolde et al. 2002, Sonke 2008), and cannot be detected by a spirometer. Variations and drifts in the end-exhale and end-inhale values of displacement based signals have also been observed, and these variations have been shown to correspond to variations in the motion of the internal anatomy (Seppenwoolde et al. 2007).

A more advanced option for measuring the displacement of the chest and abdomen is to acquire the full 3D skin surface using stereo imaging techniques (Johnson et al. 2004, Moore and Graham 2000). Recent development at a collaborating commercial partner (Vision RT, London, UK) now permits the processing of the 3D surfaces in 'real-time', allowing them to be used to predict internal respiratory motion during RT treatment, as well as for image sorting, and off-line study of the respiratory motion (Johnson et al. 2004). It is possible to generate a number of different respiratory signals from the surfaces, including tracking the displacement of one or more points on the surface, and calculating the volume underneath the surface, which it is postulated will produce a signal similar to that from a spirometer (Tarte et al. 2006).

Other methods of acquiring respiratory signals that have been reported in the literature include belt based systems that go round the patient's chest and/or abdomen and stretch with respiration (Koch et al. 2004, Kubo and Bruce 1996, Lu et al. 2006), and temperature based systems that record the difference in temperature between inhaled and exhaled air (Kubo and Bruce 1996).

Internal respiratory surrogate signals can also be extracted from images of the internal anatomy. Implanted markers, as discussed under x-ray imaging (Section 2.2.1), can be considered respiratory surrogate signals in a sense, as their motion is used to predict the motion of the tumour (although if they are actually implanted within the tumour this 'prediction' should be 100% accurate). Electromagnetically tracked internal markers have also been developed (Parikh et al. 2005) which avoid the need for x-ray imaging and the associated radiation dose. Another popular choice of internal surrogate signal is too measure the height of the diaphragm, and this has been measured from x-ray imaging (Vedam et al. 2004), 4DCT imaging (Zhang et al. 2007), CBCT projection data (Sonke et al 2005, Zeng et al. 2007a), MRI (Blackall et al. 2006), and Ultra-Sound (US) imaging (Xu and Hamilton 2006). Of these, only x-ray and US imaging could be used to measure the height of the diaphragm during treatment. As mentioned in the MRI section (Section 2.2.3) von Siebenthal et al. (2007) track up to five easily identifiable internal structures (veins in the liver) in MRI 'navigator' slices which are used to sort the MRI data slices. The location of each structure could be considered as a separate surrogate signal (or two, as the position is measured in two dimensions).

### 2.3.2 Respiratory parameters

There are two main respiratory parameters that are used to characterise the respiratory motion: amplitude and phase. The 'amplitude' of the respiratory signal is actually just the value of the signal at the desired time point. Note, this is not the usual mathematical definition of the amplitude of a signal, but 'amplitude' is the name given to the 'value of the respiratory signal' parameter in most of the literature and will be the name used throughout this thesis. The phase of the respiratory cycle indicates how far through the cycle you are, with one point in the respiratory cycle, usually end-exhale, being assigned to 0% (and 100%). Both parameters have their advantages and disadvantages. When using phase the assumption is that the respiratory motion follows the same path during every respiratory cycle. Phase can model hysteresis, which is when the motion follows a different path during inhalation than it does during exhalation. When using amplitude the exact values at end-exhale and end-inhale may differ from one respiratory cycle to the next, allowing some degree of inter-cycle variation to be modelled. However, amplitude on its own cannot model hysteresis. To overcome this problem a distinction is often made between values occurring during exhale and those occurring

during inhale (Lu et al. 2006). Although distinguishing between exhale and inhale values allows hysteresis to be modelled, it effectively constructs two separate models so the overall model will no longer be continuous. Figure 2.1 demonstrates the differences between the motion paths that can be represented using the different parameters.

Phase is sometimes preferred for image sorting (e.g. 4DCT) as the full range of phase values are covered in each respiratory cycle (by definition) so will be acquired at every physical location. If there are large variations between some of the respiratory cycles, there may only be a small range of amplitude values (or even none) that are acquired at all locations, and complete volumes can only be formed for these values. However, if there is a lot of inter-cycle variation then it is highly likely that there will be errors and artefacts in the images reconstructed using phase, that may be reduced in the amplitude based images (Lu et al. 2006, Rietzel et al. 2005a). Another problem with phase is that it can be difficult to calculate accurately in real-time. Algorithms for calculating phase in real-time exist and are used in commercial systems (e.g. the Varian RPM system), but irregular respiration can cause problems for these algorithms. In such cases it has been shown that re-calculating the phase after acquisition using a more accurate (but not real-time) semi-automatic method produces images with less artefacts (Rietzel and Chen 2006).

### 2.3.3 Acquiring multiple respiratory signals and parameters

It should be noted that if both hysteresis and inter-cycle variation occur (which they do to some degree in many patients, Seppenwoolde et al. 2002) then neither phase or amplitude or any other single respiratory parameter will be able to accurately characterise the motion on its own. Distinguishing between inhalation and exhalation can be thought of as introducing another respiratory parameter, but this is a binary parameter, not a continuous parameter like phase or amplitude, and so produces two distinct models and not one continuous model (Figure 2.2). To overcome this problem some groups have suggested acquiring two or more continuous parameters to characterise the respiratory motion. Low et al. (2005) use tidal volume and flow as their respiratory parameters (equivalent to amplitude, and its time derivative, gradient). Zhang et al. (2007) use the amplitude and gradient of the diaphragm motion as their

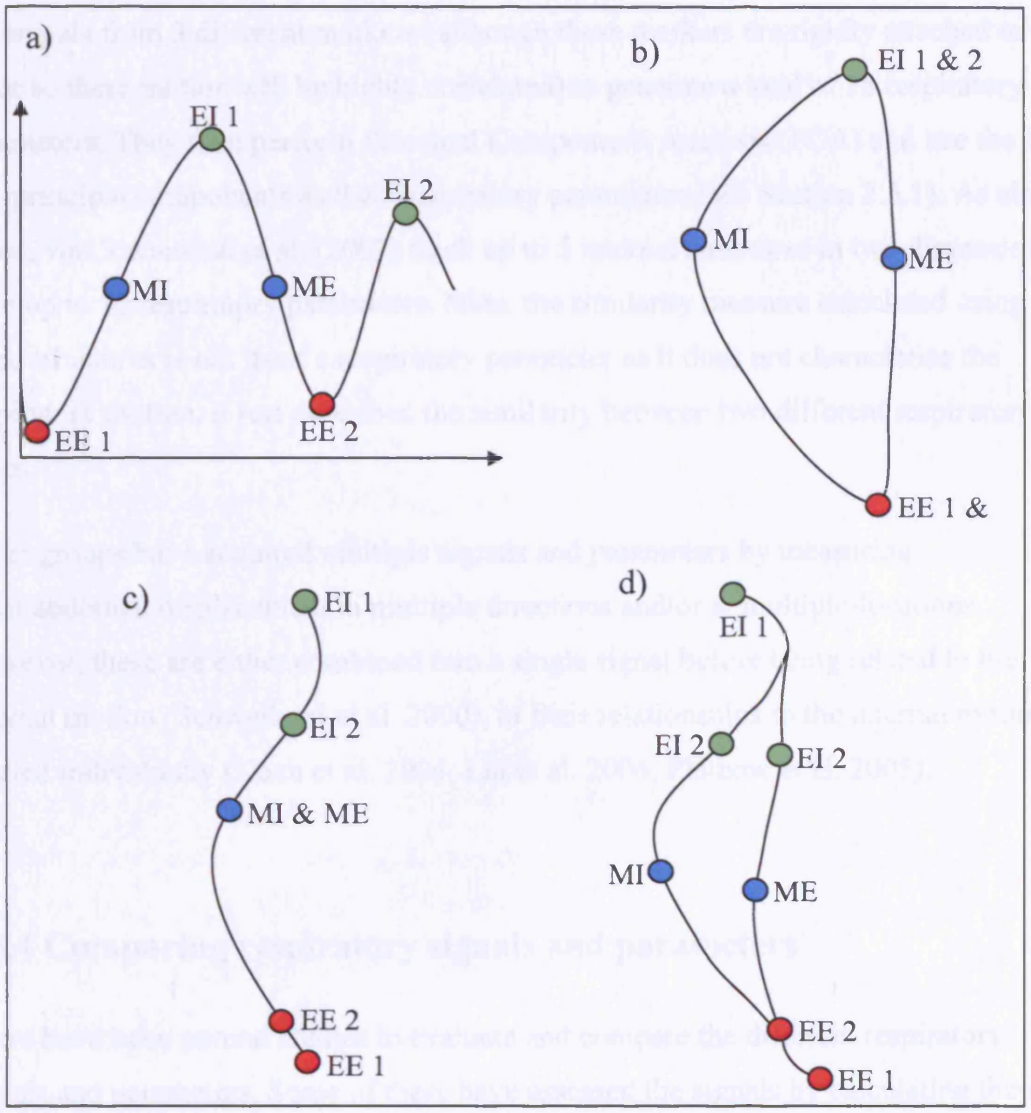


Figure 2.1 - a) An example trace from a respiratory surrogate signal. b-d) Possible internal motion paths that can be modelled using different respiratory parameters: phase (b), amplitude (c), and amplitude + direction (d). b) Phase models the internal motion as a 'loop'. It can model hysteresis so can distinguish between Mid-Inhale (MI) and Mid-Exhale (ME). It cannot model any inter-cycle variation so the End-Exhale (EE) and End-Inhale (EI) points are always at the same location for every respiratory cycle, even if the values of the respiratory signal are different (as for EE1 and EE2, and EI1 and EI2). c) Amplitude models the motion as a line. The motion does not necessarily proceed the same distance along the line during each respiratory cycle, so it can distinguish between different values at EE and EI, and hence can model a limited amount of inter-cycle variation. It cannot model hysteresis so cannot distinguish between MI and ME. d) Amplitude + direction (inhaling or exhaling) effectively creates two separate models, although they may be 'joined-up' at the ends (as in the figure). Using amplitude + direction it is possible to distinguish between different EE and EI values, and between ME and MI. However, if one of the EI (or EE) values occurs in the region where the models are not joined up then the motion will 'jump' from one model to the other and will not be continuous (as at EI 2).

respiratory parameters. Khamene et al. (2004) use the amplitude and gradient of the full 3D signals from 3 different markers (although these markers are rigidly attached to each other so there motion will be highly correlated) to generate a total of 18 respiratory parameters. They then perform Principal Components Analysis (PCA) and use the first two principle components as their respiratory parameters (see Section 2.5.1). As already noted, von Siebenthal et al. (2007) track up to 5 internal structures in two dimensions so have up to 10 respiratory parameters. Note, the similarity measure calculated using these structures is not itself a respiratory parameter as it does not characterise the respiratory motion, it just described the similarity between two different respiratory states.

Other groups have acquired multiple signals and parameters by measuring chest/abdomen displacement in multiple directions and/or at multiple-locations. However, these are either combined into a single signal before being related to the internal motion (Schweikard et al. 2000), or their relationships to the internal motion are studied individually (Koch et al. 2004, Liu et al. 2004, Plathow et al. 2005).

### 2.3.4 Comparing respiratory signals and parameters

There have been several studies to evaluate and compare the different respiratory signals and parameters. Some of these have assessed the signals by calculating their linear correlation with the tumour motion (Hoisak et al. 2004), the motion of another internal structure (Koch et al. 2004, Plathow et al. 2005), or the internal air content calculated from CT images (Lu et al. 2005a, Lu et al. 2006). Although this is an obvious way to assess the surrogate signals it assumes that there is a simple linear relationship between the signal and the tumour motion. This will not be true if there is hysteresis in the internal motion (Seppenwoolde et al. 2007), or 'phase offsets' between the surrogate signals and internal motion (Hoisak et al. 2004). Another way that the respiratory signals and parameters have been assessed is by calculating the residual motion of the tumour or internal structures occurring at a particular value (or range of values) of the respiratory parameter (Berbeco et al. 2005, Liu et al. 2004). This method is often used when the surrogate signal is intended for use in gated RT (see Section 2.1.5).

Spirometry has been compared with skin surface displacements. Spirometry correlated better with the internal air content as measured in Cine CT volumes (Lu et al. 2005a),

but this is not surprising (especially as the spirometry signal is 'drift-corrected' using the internal air content!) and does not necessarily mean that it is better for modelling the internal motion (although this is what they claim). More meaningfully, spirometry also correlated better with Superior-Inferior tumour motion than surface displacement, although for most patients neither had consistent linear relationships with the tumour motion over multiple days (Hoisak et al. 2004).

Different skin marker locations have been compared under different types of breathing (deep, shallow, abdominal , thoracic, etc.) and for predicting the internal motion of different structures and in different directions. The marker location that correlates best with the internal motion depended on the location and direction of the internal motion, and the type of breathing that the patient was performing (Koch et al. 2004, Liu et al. 2004, Plathow et al. 2005).

Three different methods have been used to compare amplitude and phase as respiratory parameters. Amplitude correlates better with internal air content (Lu et al. 2006), but again this is not surprising and not that meaningful. Amplitude based gating usually, but not always, results in less residual tumour motion (Berbeco et al. 2005). 4DCT volumes sorted using amplitude contain less artefacts than those sorted using phase, particularly if there is a lot of inter-cycle variation during the data acquisition (Lu et al. 2006, Rietzel et al. 2005a).

Overall, the ability of respiratory surrogate signals and parameters to characterise the respiratory motion is inconclusive. Most authors agree that there is some relationship between the external surrogate signals and the internal motion. However, this relationship can be more complex than the simple linear relationships assumed by correlation studies, and this relationship may not be stable for long periods of time. The strength of the relationship may depend on where in the lung the motion is being studied as well as which signal/parameter is used. See Section 2.5 for a description of the different models used to relate the internal motion to the respiratory parameters.

## 2.4 Registration

The aim of registration is to find correspondence between different images and/or physical objects. It may be desirable to find correspondence between two different images for many reasons: to compare images of a patient acquired at different times (e.g. to assess progress of disease), to combine information acquired using different imaging modalities, or to compare images from different patients to assess differences in the population. High quality diagnostic images may be registered to physical objects, usually the patient or the treatment room, so that the information in the images can be used to guide interventions (image guided interventions). Physical objects such as scanners or hand held tools may be registered to each other so that the objects (and any images produced by the scanner) share a common co-ordinate system. All the registrations used for constructing the motion models are image to image registrations, although image (or model) to treatment room registrations would also be required if the models were ever used to guide treatment.

Mathematically, registration is the process of establishing a transformation that maps corresponding anatomical points from one co-ordinate space to another. I will adopt the terminology used in our lab and most others for image registration, where the registration result maps from the Target co-ordinate space into the Source co-ordinate space. At first this may seem counter-intuitive, but the transformation is required in this direction in order to form an image in the Target space using the intensity values of the Source image, i.e. to view the transformed Source image aligned with the Target image, which is usually the aim of image registration.

All registration algorithms essentially consist of three components: a transformation model, a similarity measure, and an optimisation method (for more detail on all aspects of medical image registration see Hajnal et al. 2001)

### 2.4.1 Transformation models

The transformation model defines what types of transforms are permitted and how they are parameterised. The number of parameters used to define a transformation is referred to as the number of degrees of freedom, and generally the more degrees of freedom a transformation has the more complex the transformation it can represent. Some popular

transformation models, and their applicability to registering lung images, will now be considered.

### 2.4.1.1 Rigid transformations

A rigid transformation consists of a rotation and a translation, so a 3D rigid transformation will have six degrees of freedom. Rigid transformations maintain distances between points, straight lines, and parallel lines. Rigid transformations are typically used to map between different co-ordinate systems or to align anatomy that is assumed to remain rigid, e.g. bones. Rigid transformations are a subclass of affine transformations.

### 2.4.1.2 Affine transformations

Affine transformations expand on rigid transformations by also including anisotropic scaling, shearing, and reflection, although the reflection is not normally used for registering medical images. 3D affine transformations are often described using fifteen parameters, three each for rotation, translation, and scaling, and six for shearing, although three of the shearing parameters are redundant, and affine transforms only have twelve degrees of freedom. Affine transformations can also be represented in matrix form with twelve parameters. Affine transformations do not preserve distances between points but do preserve straight and parallel lines.

Affine transformations can be useful for registering images that have been subject to scaling and shearing geometric distortions which are sometimes introduced in the imaging process. Although they are generally not adequate to describe the deformation that can occur in soft tissue due to respiration (Coselmon et al. 2004), they can be useful for recovering some of the global motion and deformation, and as a starting estimate to the more complex but less well constrained non-rigid registrations described next.

### 2.4.1.3 Non-rigid transformations

Transformation models that are able to represent more complex deformations, and hence have more degrees of freedom than an affine transform (usually many more, for some as many as three per voxel in the Target image), are usually referred to as non-rigid transformations. There are many different non-rigid transformation models, and many of them have been used for registering lung CT images. The three main types of non-rigid transformation and their application to lung motion modelling will be discussed next.

#### 2.4.1.3.1 Global basis functions transformation models

There are many non-rigid transformation models based on global basis functions, although only thin plate splines have had much use in lung CT registration (Coselmon 2004). Global basis function transformation models are used to interpolate between displaced points. This means that these registrations require a set of approximately corresponding points to be located in each image prior to registration, although this can sometimes be achieved automatically (Berlinger et al. 2006). Also, as the name suggests, global basis functions have a global effect on the image which means they can be computationally expensive to optimise.

#### 2.4.1.3.2 Parametric transformations models

Parametric transformation models use a set of parameters that have a local effect on the image, and hence optimising any one parameter is less computationally expensive than functions with global effect. Although many different functions could be used for parametric transformation models, the ones that have had the most extensive use in image registration are the cubic B-splines. Registrations that use a cubic B-spline transformation model parameterise the transformation using the displacements of regularly spaced Control Points (CPs), which are typically separated by several voxels. Approximating B-splines are used to calculate the deformation field at any point from the displacements of the four control points (in each dimension) that surround that point.

For a 3D cubic b-spline transformation defined by a $n_x \times n_y \times n_z$ grid of CPs $\phi_{i,j,k}$ with uniform spacing of $\delta_x, \delta_y, \delta_z$ in each dimension:

$$T_{B-spline}(x,y,z) = \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)\phi_{i+l,j+m,k+n}$$

[2.1]

where $i = \lfloor x/\delta_x \rfloor - 1$, $j = \lfloor y/\delta_y \rfloor - 1$, $k = \lfloor z/\delta_z \rfloor - 1$, $u = x/\delta_x - \lfloor x/\delta_x \rfloor$, $v = y/\delta_y - \lfloor y/\delta_y \rfloor$, $w = z/\delta_z - \lfloor z/\delta_z \rfloor$, and $B_l$ represents the $l$-th basis function of the B-spline, given by:

$$\begin{aligned} B_0(u) &= (1-u)^3/6 \\ B_1(u) &= (3u^3 - 6u^2 + 4)/6 \\ B_2(u) &= (-3u^3 + 3u^2 + 3u + 1)/6 \\ B_3(u) &= u^3/6 \end{aligned}$$

[2.2]

Registrations using a B-spline transformation model were originally proposed for registration of MRI breast images (Rueckert et al. 1999), but have since been used for registering images from a wide range of anatomical sites and different modalities, including CT images of the lung (Flampouri et al. 2006, Schreibmann et al. 2006, Rietzel et al. 2005b, Zeng et al. 2007a).

### 2.4.1.3.3 Non-parametric transformation models

Non-parametric transformation models are not defined by a set of parameters but are usually defined by a displacement vector for each voxel in the Target image, giving them a very large number of degrees of freedom. To constrain the registrations a regularisation term is added that penalises transformations that deviate from the desired form. Non-parametric transformations may sometimes be expressed and solved using

partial differential equations, and these are often used in fast efficient implementations of these registration algorithms. There are several non-parametric transformation models that have been used for registering lung images, including the linear elastic model (Christensen et al. 2007, Heath et al. 2007, Sarrut et al. 2006), the optical flow model (Guerrero et al. 2006) and the viscous fluid model (Keall et al. 2005, Pevsner et al. 2006, Sundaram et al. 2004).

#### 2.4.1.4 Dimensionality of transformation models

Most of the transformation models in the literature used for registering lung images (and all the ones used in this thesis) are 3D transformation models, that is they map from a 3D space to a 3D space. 2D transformation models have been used to register 2D sagittal MR images of the lung, although this assumes there is negligible through-slice motion (Sundaram et al. 2004). Chandrashekara et al. (2004) proposed a 4D transformation model (used to register cardiac MR images) which is used to map between a static 3D volume (represented as a 4D dataset with no variation along the 4th dimension) and a 4D time series sampled as 3D volumes. Zeng et al. (2007a) also use a 4D transformation model to map between a static 3D volume and a 4D time series, but their 4D time series is sampled using 2D projections from orbiting views. This means that an additional transformation is required from 4D space to 2D space. This transformation represents the imaging process, and its parameters depend on the time and angle that the 2D projection is acquired from. In theory any of the transformation models could be applied in 4D (or 2D), although Chandrashekara et al. (2004) and Zeng et al. (2007a) have both used a B-spline transformation model.

### 2.4.2 Similarity Measures

Similarity measures fall into two main categories, feature based and voxel based. Feature based methods rely on detecting corresponding features in the images to be aligned. These features can be actual corresponding points or corresponding surfaces (where typically the points that define one surface do not directly correspond to the points that define the other surface). Feature based registrations have two main disadvantages: the features need to be detected, and the registrations are only assessed at

the location of the features, so the accuracy of the registration in other regions, particularly those distant from any features, cannot be guaranteed.

Voxel based similarity measures have been developed to overcome the above problems by assessing the similarity between the two images at every voxel. However, voxel based registrations can run very slowly in comparison to feature based registrations, as typically the number of features to be compared (up to several thousand for surface points) is a lot smaller than the number of voxels to compare (can be several million). Many different voxel similarity measures have been developed for many different applications. The choice of similarity measure should take into account the expected differences between the images including whether the images are acquired using different modalities and/or different imaging parameters.

Two voxel similarity measures frequently used for registering lung images are Sum of Squared Differences, SSD, and Cross Correlation, CC. SSD assumes that the images are identical once registered, and so is most applicable when the only differences between the images are due to motion, as is the case (neglecting density change effects) when registering volumes from a 4DCT dataset. CC assumes there is a linear relationship between the intensities in the two volumes being registered, and can be useful when some of the imaging parameters (such as image intensity scaling) have been altered. Mutual Information, MI, and Normalised Mutual Information, NMI, are similarity measures based on information theory which have proved very useful in a variety of applications, particularly for inter-modality registrations (Mattes et al. 2003), and have also been successfully used for registering images of the lung in different respiratory states (Coselmon et al. 2004, Sundaram Cook et al. 2007). It is possible that registration based on voxel similarity measures may not be well constrained enough in regions where there is little variation in intensity to ensure an accurate result is achieved. West et al (2005) have developed a registration method that uses a weighted combination of a point based similarity measure and a voxel similarity measure. This enables the registration to be driven by the intensity values in all of the voxels, but allows a few carefully selected points to be used in regions where the purely voxel based registration does not produce satisfactory results.

### 2.4.3 Optimisation methods

The optimisation method tries to find the parameters of the transformation model that give the best value of the similarity measure. The choice of optimisation method (and implementation) will depend upon the transformation model, and sometimes the similarity measure, being used. Some transformation models will have a standard optimisation scheme, and some can be optimised using a choice of different methods. For some point based registrations it may be possible to find the optimal parameters using a closed form solution (e.g. using least squares) but for all registrations that use a voxel based similarity measure it is necessary to employ iterative optimisation techniques. Different optimisation schemes have different performance characteristics (e.g. speed, robustness) and the development of improved optimisation schemes is an ongoing area of research (Mattes et al. 2003, Schnabel 2001, Shekhar et al. 2007).

Gradient descent optimisation is often used with the B-spline (and other non-rigid) transformation models. This is an iterative technique. At each iteration the gradient of the transformation with respect to the similarity measure needs to be estimated. This can be estimated easily using the finite differences method (Rueckert et al. 1999), although there is evidence that using more complex analytical techniques (Mattes et al. 2003) can greatly reduce the computational time (Klein et al. 2005).

Non-rigid registrations often employ multi-resolution techniques to improve robustness and prevent finding local minima. Multi-resolution techniques can be applied to the images, first registering coarse (down sampled) images, and using the result as input to the registration of the finer resolution images. When using a B-spline transformation model, a multi-resolution technique can also be used with the control point grid, using a progressively smaller control point spacing for each resolution level (Rietzel et al. 2005b).

### 2.4.4 Bio-mechanical models

Bio-mechanical models (often implemented as finite element models) have also been used to perform image registration. This involves first defining regions within the images (such as the different organs) and assigning physical properties to them (e.g. elasticity). Then, corresponding features (e.g. surfaces or landmarks) are detected within

the images, and these are used to calculate the forces that are applied to the bio-mechanical models. The bio-mechanical models are then used to find the displacement of the whole anatomy based on the forces that are applied and the physical properties assigned to the different regions. It has recently been proposed that voxel based image similarity measures can be used to generate the forces that are applied to the bio-mechanical models (Sundaram and Gee 2005). To date there has been relatively little work published on using bio-mechanical models for lung registrations, although they were used by Sundaram and Gee (2005) for 2D lung registrations, and by Brock et al (2005) for a multi-organ registration study which included the lungs.

### 2.4.5 Registration of lung CT images

This section will first describe two problems that can affect the registration of lung CT images: the density change effect, and the sliding problem. It will then review some of the studies that have assessed and compared non-rigid registration algorithms on lung CT images.

#### 2.4.5.1 The density change effect

When the lungs fill with air during respiration they will increase in volume (by up to 65% in some regions, Alasti et al. 2006) and will undergo a corresponding change in density as their mass stays approximately constant (there can be a slight variation due to extra blood mass entering and leaving the lungs, Guerrero et al. 2006). A change in the density of the lungs will produce a corresponding change in the intensity of the lung voxels in CT data.

The density change effect may cause problems when registering CT images from different respiratory states as some corresponding voxels will have different intensities in the different images. However, it can be argued that while the density change effect will certainly affect the value of intensity based similarity measures, it is unlikely to have much effect on the final result of the registrations. This is because the lungs contain a lot of high intensity structures, and it is likely that the alignment of these structures will have more of an effect on the similarity measure than comparatively

small changes to the voxel intensities due to density changes. However, if the transformation model used does not produce smooth transformations, or if there are some regions of the lung with little visible structure, the density change effect may have an adverse effect on the registration results (Sarrut et al. 2006).

The Jacobian of a non-rigid transformation can be used to measure the local change in volume due to the transformation. A Jacobian less than 1 corresponds to a decrease in volume (with a value less than 0 indicating that folding occurs) and a value greater than 1 corresponds to an increase in volume. The Jacobian could in theory be used to correct the intensities of the deformed image at each iteration of the non-rigid registration algorithm, but this would be extremely computationally demanding (Keall et al. 2005). Another approach is to try and estimate and correct for the density change prior to registration. Sarrut et al (2006) describe a simple method for this called A Priori Lung Density Modification (APLDM) which finds corresponding slices in the images and then modifies the lung intensities so that the average intensity is the same in corresponding slices. While they acknowledge that this is a simplification, they claim that the registration results are improved when APLDM is used. Note, this may be because they do not use a very smooth transformation model; even when APLDM is applied there are still some regions of the deformation field that have a negative Jacobian.

If dose calculations are performed on a deformed volume there may be errors in those calculations if density changes have not been accounted for correctly when deforming the volume (Heath et al. 2007), although there is evidence that these errors will be minimal (Webb et al. 2006). The Jacobian of the transformation can again be used to account for the changes in density (although there will be problems if there are any negative Jacobian values), or if the deformed volume is generated by 'interpolating' between the two registered volumes (see Section 2.5.2) the intensity value can be interpolated from the two images (Sarrut et al. 2006).

### 2.4.5.2 The sliding problem

The sliding problem is caused when two parts of the anatomy slide past each other, e.g. the lung and the chest wall during respiration. Almost all transformation models try to produce a smooth and continuous deformation field. For most applications this is a

desired constraint, as the deformation within organs, and even between most, will indeed be continuous. Registration algorithms have even been modified to try and ensure a smooth and continuous transformation for lung CT registrations (Heath et al. 2007, Shekhar et al. 2007). However, when sliding motion occurs the deformation is not continuous, and so any continuous deformation will inevitably include errors.

It has been proposed that the sliding problem can be overcome by segmenting the image into a region that moves with the lungs during respiration and region that does not, and then registering the two regions separately (Flampouri et al. 2006, Rietzel et al 2005b). This technique may produce problems at the boundary between the two regions, as it is likely that this boundary will be transformed to a slightly different location by each registration, creating gaps and areas of overlap between the two regions. This potential problem at the boundary of the two regions has not been addressed in the literature, although it seems likely that the magnitude of the problem and its impact on RT would be minimal in comparison to the sliding problem.

### 2.4.5.3 Assessing lung CT registrations

Most (but not all) studies that have used non-rigid registration have attempted to assess the accuracy of their registrations. This is either done through visual inspection with the aid of colour overlays or difference images, or by using delineated structures (such as the tumour or the lung) and/or manually identified landmarks (such as bifurcations in the lung). In the later case the structures or landmarks are delineated in both volumes, the registration result is used to deform them from one image to the other, and the accuracy can be assessed by the misalignment between the deformed and manually delineated structures and landmarks.

Recently there have been two multi-institution studies conducted to assess and compare the accuracy of a number of non-rigid registration algorithms and implementations (Brock et al. 2007, Kashani et al. 2007b). Note, I participated in both of these studies, as did other members of our lab.

Brock et al. (2007) asked participants to register clinically obtained 4DCT images of the lung at end-inhale and end-exhale (as well as other images from other organs). The images contained noticeable artefacts at the boundaries of different couch positions,

especially in the inhale image, but these were deemed representative of most clinically acquired 4DCT images. They assessed the results using 17 bifurcations in each lung as well as 2 calcifications in the Heart and 2 in the Aorta, manually identified in each image. 18 different implementations of various non-rigid registration algorithms were assessed. All participants had an absolute mean error of less than the voxel size (2.5 mm) in each direction. It was also noted that some of the implementations that had the highest accuracies also had relatively fast runtimes (10 minutes or less), showing that a long runtime is not necessary for accurate registration results.

Kashani et al. (2007b) used a specially designed "breathing" lung phantom (Kashani et al. 2007a). This had 48 small plastic markers embedded in it distributed across all parts of the phantom. These markers could be easily and accurately detected in the images and used to assess the registration results. Once the markers had been located the images were edited to remove the markers before the images were distributed to the participants for registration. The phantom was CT scanned in two different deformation states, corresponding to end-inhale and end-exhale. A total of 8 different registration implementations were assessed. The average errors ranged from 1.5 - 3.9 mm. However, no algorithm was accurate over all regions of the image with maximum errors ranging from 5.1 - 15.4 mm.

The multi-institution studies above have shown that lung CT images can generally be registered accurately using current non-rigid registration algorithms. None of the underlying transformation models consistently performed better than the others. Different implementations of similar registration algorithms from different groups had markedly different results, showing that the implementation and the parameters used for the registrations can affect the result as much as which transformation model is used.

## 2.5 Modelling respiratory motion

A respiratory motion model describes the motion of the tumour (and sometimes the rest of the anatomy) due to respiration. The model usually, but not always, relates the internal motion to one or more respiratory parameters measured from a respiratory surrogate signal.

### 2.5.1 Point based respiratory motion models

Point based respiratory motion models only describe the motion of one internal point of interest, normally a point in the tumour. Point based motion models are often used for gated and tracked treatment (see Section 2.1.5) when internal markers are not used or are only imaged intermittently. Linear correlation is often used to assess different respiratory signals and parameters, and the linear models produced can be considered respiratory motion models, although as noted in Section 2.3.4 such linear models do not always model the internal motion very accurately.

The Synchrony Respiratory Tracking System used with the Cyberknife (Accuray, Sunnyvale, California, USA) extends the simple linear model by relating the internal motion to the external signal with two 2nd order polynomials, one for inhaling and one for exhaling (Seppenwoolde et al. 2007). The system acquires both internal motion data (x-ray imaging of an implanted marker) and an external surrogate signal (skin markers) during a training period prior to treatment. Although 6 skin markers are tracked, their collective centre of mass is found and used to generate a single respiratory signal (Schweikard et al. 2000). This data is used to fit the parameters of the motion model. Once the motion model has been established the external surrogate signal can be used to predict the internal motion. Internal data can be acquired intermittently throughout treatment and used to verify and update the motion model. The model reverts to a simple linear model if the value of the respiratory surrogate exceeds the range of values observed during training so as to prevent large extrapolation errors. It has been shown that the polynomial model can predict the internal motion more accurately than a simple linear model when there is hysteresis or a phase offset between the internal and external motion. As the model is based on a single continuous parameter (and one binary

parameter), it will not be able to accurately characterise inter-cycle variation. However, the results show that the model can predict the internal motion well, even in the presence of irregular breathing. Fig. 4 in Seppenwoolde et al. (2007) gives a good illustration of the situation: the shape of the model does not resemble the actual motion trajectories very closely, but the prediction errors are still relatively small.

Low et al. (2005) propose a 'Five Dimensional' (5D) motion model based on two continuous respiratory parameters, the volume and flow measured by spirometry (equivalent to amplitude and gradient). The location of an internal point will depend on its location in the reference volume (3 dimensions) and the current values of the volume and flow (the other 2 dimensions). The internal motion is related to the volume and flow using a simple linear model. The model assumes that the motion lies on a plane. The points used in this study all had negligible L-R motion, so the plane was always fit in the sagittal orientation, although the model could easily be extended to a plane with any orientation. The motion in one direction on the plane will be caused by changes to the tidal volume, whilst motion in another direction on the plane is caused by the value of the airflow. The airflow component of the model can account for hysteresis, while different combinations of tidal volume and airflow can account for inter-cycle variation.

To construct the models Cine CT data was acquired with simultaneous spirometry for 11 seconds (15 volumes) at each couch position. Points of interest were then automatically tracked over the 15 volumes, giving the location of the points and corresponding volume and flow values over two to three respiratory cycles. The implementation used in the paper was unable to track points across different couch positions, but they claim this should be possible in the future. A model was then fit to each point of interest by minimising the root-mean least-squares average distance between the fits and measurements.

Seventy-six points were tracked in 4 patients. The models showed a very good fit to the data, even when there were large inter-cycle variations. The mean discrepancies between the fit and the actual locations ranged between 0.28mm and 1.71mm for the four patients (average 0.75mm), and the maximal discrepancies ranged between 0.64mm and 3.31mm. It should be noted that these errors are the errors in fitting the model to the data. Leave-one-out experiments were not performed so the ability of the

models to predict unseen data and the possibility of over-fitting the data were not evaluated.

Khamene et al. (2004) also built a model based on two continuous respiratory parameters, but these parameters are found by applying PCA (Principal Components Analysis) to a larger number of respiratory parameters. They track the full 3D motion of three skin markers, although these markers are rigidly attached to each other so their motion will be highly correlated. So that the model can account for hysteresis they also use the gradient of each 3D signal, giving a total of 18 external respiratory parameters. PCA is then performed on these 18 parameters, and the first two principal components used as the respiratory parameters. The 3D location of a point of interest (the portal vein in the liver) and its velocity in each direction (the gradient) are used as internal parameters. PCA is again performed on the internal parameters, and the first two principal components are used. For both the internal and the external data the first two principal components account for 98% of the variation.

Note, in this paper PCA is being used as a dimensionality reduction tool and not to establish a relationship between the internal and external data. This relationship is established by specifying some other (non-linear) functions that relate the values of the external principle components to the internal principle components. They give b-splines as example functions, and say the function are fit to the data using the standard least squares method, but give very little detail on how this part of their method is performed (although the paper is a short conference paper, so space may have been limited). As only two functions are used to map the two external principle components to the two internal principle components, this would imply that mapping from external to internal principle components is one to one, although this is not clearly stated. They use leave one out experiments, where the model is constructed using all but one of the data points (approximately 100 per study), and the model is then used to predict this data point. Using this method they find the mean error in predicting the location of the internal point of interest to be 1.8mm, which they state is about 15% of the average target motion.

### 2.5.2 Deformation field based motion models

Several different motion models have been proposed that model the internal motion using a dense deformation field. This means that the models describe the motion of all of the anatomy covered by the deformation field, not just at some specific points as for the models above. The deformation fields are often obtained as the result of non-rigidly registering images acquired at different respiratory states. Some of the deformation field based models of the lungs reported in the literature are described in more detail below. Deformation field based models of the respiratory motion of other organs, including the liver (Blackall et al. 2005, Rohlfing et al. 2004) and the heart (Ablitt et al. 2004), have also been reported in the literature.

Some authors have proposed non-rigidly registering CT (or MR) volumes acquired at inhale and exhale breath hold, and then linearly interpolating the results to predict the intermediate deformations, and from these the intermediate volumes (Sarrut et al. 2006, Schweikard et al. 2004, Zeng et al. 2007b). Schreibmann et al. (2006) have proposed a similar model, but use volumes acquired while the patient is freely breathing to overcome the potential differences between the anatomy at breath hold and the anatomy during free breathing (Blackall et al. 2006, Rietzel 2005a, and Section 7.3 of this thesis). There are a number of problems with the linear interpolation based models. The predicted volumes will be affected by the density change effect (see Section 2.4.5.1). Sarrut et al. (2006) is the only author that acknowledges and accounts for the density change effect. The models are based on the assumption of linear motion over the respiratory cycle, so cannot model hysteresis (or any other non-linear motion). Schreibmann et al. (2006) suggest including an intermediate volume between inhale and exhale in the model. If two such volumes are used, one from mid-inhale and one from mid-exhale, then hysteresis can potentially be accounted for, although this is not mentioned in the paper (only inhale volumes are used and predicted, so the effects of hysteresis are not assessed). There is a further problem that does not appear to have been identified in any of the papers referenced above: the inverse of the interpolated transformation is required to form a new image. It may be possible to estimate the inverse of the interpolated transform (see Appendix A), but none of the authors mention doing this. If the transformations contain regions where folding occurs, as those used by

Sarrut et al. (2006) do, then the inverse does not exist in these regions and it cannot be estimated.

The intermediate volumes are related to an external respiratory surrogate signal in different ways. Schreibmann et al (2006) find the location of a skin marker in their intermediate volumes. Sarrut et al. (2006) use the air content of the lungs in the intermediate and reference images to calculate the tidal volume as would be read by a spirometer. Schweikard et al. (2004) and Zeng et al. (2007b) both propose a similar idea, although Schweikard et al. (2004) had not implemented the idea, where the intermediate volumes are compared directly to x-ray images either acquired during RT treatment (Schweikard et al. 2004) or as part of a Cone-Beam CT acquisition (Zeng et al. 2007b). The x-ray images can be thought of as the respiratory surrogate signal, and values are assigned to the intermediate volumes by creating simulated x-ray projections (digitally reconstructed radiographs) through these volumes.

Many authors have proposed non-rigidly registering the volumes from a 4DCT dataset for a variety of different applications including RT planning, image acquisition, and assessment of lung physiology (see Section 2.1.5, Section 2.6.1, and Section 2.6.2 respectively). The results of these registrations can be considered a motion (or deformation) model, and they are sometimes referred to as such. As a respiratory parameter is required for sorting the 4DCT data, the internal motion and deformation is implicitly related to this parameter. The deformation is only known for a few values of the respiratory parameter, those which correspond to the 4DCT volumes. Therefore these models are discrete rather than continuous models.

Blackall et al. (2006) propose a continuous model built using the results of registering 4D MR volumes. Due to restrictions with the data only affine registrations were possible. The affine registrations are used to propagate a triangulated lung and tumour surfaces from the reference image (end-exhale) to all of the other images. For each point on the surfaces, a 3rd order polynomial is fit that relates the motion of the point to the height of the diaphragm (the respiratory signal). As a 3rd order polynomial is used non-linear motion can be modelled. Separate models are built for inhalation and exhalation allowing hysteresis to be studied and modelled (but not with one continuous model).

Note, this model is technically lots of point based models rather than a deformation field based model. However, if non-rigid registration had been possible then registration control points could have been modelled instead of the surface points using exactly the same method (as done for the liver in Blackall et al. 2005). In addition, most deformation field based models can be thought of as a multiple-point-based model. Hence this paper has been covered in this section and not the previous section on point based models.

Zhang et al. (2007) propose a deformation field based motion model that uses two continuous respiratory parameters, the current height of the diaphragm and its height 1.5 seconds earlier (which is similar to using the gradient). The deformation fields are generated by non-rigidly registering the volumes from a 4DCT dataset to a reference volume (the end-exhale volume of the 4DCT dataset). The dense deformation fields are then related to the respiratory parameters using Principal Components Analysis (PCA). Note, unlike in Khamene et al. (2004 - discussed in the previous section), here PCA is not just being used to reduce the dimensionality, but also to find the relationship between the internal motion and respiratory parameters. Once PCA has been performed the two respiratory parameters can be used to calculate the value of the first two principal components, and from these the values for the entire deformation field can be calculated. Only two principal components can be used as there are only two respiratory parameters, but the authors show that these account for between 83%-90% of the variation for the four patients studied.

As the model uses two continuous parameters it can potentially model hysteresis and inter-cycle variation. However, only one 4DCT dataset is used as training data to build the model, so the training data does not actually sample any inter-cycle variation. The results show that the model can predict the centre of the tumour to within 2mm in each direction in the training data (i.e. fitting error). Errors at the lung and tumour surfaces are mostly within 5mm, and it is shown that these are comparable to inter-observer error for the tumour delineations. The authors state that the errors are mostly due to the PCA removing "noise" from the deformations. In support of this three examples are shown where the model predictions have greatly reduced artefacts in comparison to the original 4DCT images. This implies that the models are actually more representative of the true respiratory motion than the original data.

The models are also used to predict 4DCT data acquired in a different session, approximately a week after the first (both before RT treatment has commenced). The results show that the models can predict the data from the second session relatively well (mean error < 2mm in all directions for centre of tumour), even when there is a considerable base-line shift in the breathing between the two sessions for one patient. However, the data acquired for the second session is lower quality (5mm slices, 1 second gantry rotation, data sorted into four respiratory states) making it difficult to accurately assess the errors, and the overlay images indicate that the prediction errors are clearly larger than for the training data (as would be expected). The authors also acknowledge that the tumours studied are all located in the vicinity of the diaphragm, and that the models may not perform so well at predicting other sites in the lung.

Zeng et al. (2007a) use a 4D transformation model for their non-rigid registrations (to Cone-Beam projection data). Therefore the registration result is a continuous motion model, although the motion is modelled over the acquisition time and is not related to a respiratory parameter. The height of the diaphragm (detected in the cone-beam projections) is utilised in their method, but only to help regularise the registration which is not well constrained by the data. The deformation of the anatomy can vary from one respiratory cycle to the next, even when the diaphragm is at the same height. While this will be very useful for studying the inter-cycle variation of the internal anatomy, it means that this model cannot be used to predict the deformation in the future from a respiratory surrogate signal. Of course, a further model could be constructed that relates the internal motion modelled by the 4D transformation to one or more respiratory signals (but that was not the aim of this paper).

In a later paper, Zeng et al. (2007b) also propose another method to help regularise the registration. They use a simple linear motion model (based on interpolating the result of non-rigidly registering breath hold images, see above) to provide a good starting estimate to the non-rigid registration. Although the use of the starting estimate increased the accuracy and reduced the time to convergence for the non-rigid registration, when testing the method on phantom data they found that the linear model was actually more accurate than the result of the non-rigid registration (i.e. the non-rigid registration made the result worse!). They said this was due to noise and artefacts present in the Cone-Beam projection images, and possibly because the phantom motion can be well

approximated by the linear model. This reveals that while the non-rigid registration of Cone-Beam projection data is a very promising technique, particularly for studying inter-cycle variation, the robustness and accuracy needs to be improved before the results can be considered reliable.

Sundaram et al. (2004) propose a method of modelling the average deformation and appearance of the lungs from two separate individuals. The paper appears to only be a proof of concept paper, but a follow up paper could not be found. 2D MR coronal slices are used to study the respiratory motion, so there is a high chance of non-negligible through slice motion (in the A-P direction) which will cause errors. The images from each individual are registered to the next image for that individual (using 2D non-rigid registration), forming a motion model for each individual. These models are parameterised according to the normalised lung area, with the lung area at exhale given a value of 0 and the lung area at inhale a value of 1. New images can be generated by linearly interpolating the registration results, effectively modelling the motion as a series of connected line segments. For each patient images are generated corresponding to normalised lung areas of 0, 0.25, 0.5, 0.75, and 1. Only data from exhalation is used and modelled so hysteresis effects can be ignored, but a separate model and images could be created for inhalation as well in an identical manner. The inter-subject average images are then calculated by performing a simultaneous, symmetric registration of the two images from different subjects, and averaging the intensity values. The inter-subject average deformations are calculated by composing the subject specific deformation fields between different lung areas with the deformation fields resulting from the simultaneous registrations described above, and averaging them.

Due to the large variability in respiratory motion between individuals (see Section 2.4) it is unlikely that cross population models will ever be able predict the respiratory motion for a particular individual to the accuracy required for RT planning and treatment. However, cross population models potentially have many other uses including learning about lung physiology, assessing and diagnosing individuals, and to provide good starting points and constraints for patient specific analysis (e.g. non-rigid registration).

## 2.6 Other uses for respiratory motion models

Although this project is primarily focussed on constructing motion models for use in lung cancer RT there are several other potential uses for respiratory models, including image acquisition and assessing local lung function. See Section 2.1.5 for details of how respiratory motion models may be utilised in RT.

### 2.6.1 Image acquisition

Respiratory deformation motion models can be of use when reconstructing any data that may be subject to respiratory motion during acquisition. Respiratory models have been used in 4DCT reconstructions to remove artefacts (Zhang et al. 2007) and to permit a reduction in the imaging dose without degrading the image quality (Li et al. 2005). Respiratory models have been used in 4D Cone-Beam CT reconstruction to improve the image quality without acquiring extra projections, reducing the time and imaging dose required for 4D Cone-Beam CT (Li et al. 2006a, Li et al. 2007). The motion models can be constructed from standard 4DCT data acquired prior to the Cone-Beam acquisition (Li et al. 2006a), or directly from the 4D Cone-Beam (Li et al. 2007, Zeng et al. 2007a). Respiratory motion models constructed from 4DCT can be used in a similar way to reconstruct 4D PET (Positron Emission Tomography) volumes (Li et al. 200b). Respiratory motion models could also be used in MR image reconstruction to remove motion artefacts (Batchelor et al. 2005).

### 2.6.2 Assessing local lung function

Deformation models of the lung can be used to determine the local change in density or volume of the lung tissue, which will be indicative of the local lung function. The change in density can be calculated from the CT intensities in the 4DCT images (the deformation fields are used to find corresponding voxels in the different images, Guerrero et al. 2006) or the change in volume can be calculated directly from the Jacobian of the deformation field (Christensen et al. 2007). The Jacobian based technique has been shown to produce good agreement with Xenon CT based measurements in one animal (Reinhardt et al. 2007).

## 2.7 Conclusions to literature review

This literature review has shown that the study and modelling of respiratory motion has recently been, and remains, a very active area of research. Although many of the 'quick-wins' made possible by imaging and modelling respiratory motion have already been realised, and some have entered clinical practice, there is still much to do in order to fully address the issue of respiratory motion in RT and in other applications.

Developments in imaging technologies over the last five years have permitted the detailed study of respiratory motion in the lungs and other anatomy that was not previously possible. Continued developments are allowing ever more detailed data to be acquired. Initial studies into the effects of long and short term inter-cycle variations have been reported in the last year. Non-rigid registration is regarded by many as a vital tool for the detailed study of the effects of respiratory motion, but most would not yet consider the methods robust or fast enough for routine clinical use. It is widely agreed that being able to predict the internal motion from external surrogate signals would be extremely useful, and there have been many studies on the relationship between internal and external data. However, there is as yet no consensus as to which signals and parameters are best for predicting the internal motion, how they should be related to the motion, how stable the relationships are, or even if it is possible to accurately predict the internal motion from any external surrogate signals.

# Chapter 3

# Data acquisition

This chapter describes the data that has been acquired for the motion models and the methods and protocols that have been used in acquiring this data.

## 3.1 Introduction

In order to construct the motion models it is necessary to acquire data that describes the internal motion of the lungs and surrounding anatomy over the respiratory cycle. It is also necessary to acquire a respiratory surrogate signal to calculate the respiratory parameters, such as the respiratory phase, that correspond to the internal data. Additionally, a reference volume that contains all of the anatomy being modelled is required.

Section 2.2 in the literature review describes and critiques some of the popular methods of acquiring data on the internal respiratory motion, and Section 2.3 covers the different external respiratory surrogate signals. The following sections detail the modalities and protocols adopted for acquiring the internal and surrogate data used throughout the course of this research. The processing required to calculate the respiratory parameters from the external surrogate signals is also described.

## 3.2 Patients studied

Data from a total of 11 patients has been acquired (although not all of the data was used for all experiments, see later chapters for details of which patients have been used for which experiments). Most patients had upper-lobe tumours but one patient had a hilar tumour and one a lower-lobe tumour (Table 3.1). The datasets from different patients were acquired throughout the course of the PhD research, with the data from the initial four patients being acquired prior to the PhD being started, and the different protocols and methods used for acquiring the internal motion data and the respiratory surrogate signals are described in the following sections.

## 3.3 Cine CT data

Cine CT has been used throughout the course of this research to acquire data that describes the internal respiratory motion. Cine CT data is acquired with the CT scanner in Cine mode, which means that the couch is stationary when data is being acquired. Data is acquired for a period of time at one couch position, and then the scanner stops acquiring data and moves to another couch position, where it resumes acquiring data. At each couch position a series of 3D volumes is acquired while the patient is freely breathing. As the data is acquired with the couch stationary, the number of slices and Superior-Inferior (S-I) coverage of the reconstructed volumes is limited by the number and size of the slices that the CT detector can acquire simultaneously. Most CT scanners, including the ones used for this research, acquire relatively few slices (12-16) which only cover a thin slab of the patient's anatomy (approximately 2-3cm). Therefore the couch positions are usually acquired contiguously (or slightly overlapping), so that

| Patient | Tumour Location | Cine CT Protocol | Couch Positions Acquired | Cine CT Volumes per Couch Position | Respiratory Signal |
|---|---|---|---|---|---|
| 1 | LUL | P | 3 | 20 | Cine CT |
| 2 | RUL | P | 3 | 20 | Cine CT |
| 3 | LUL | P | 4 | 20 | Cine CT |
| 4 | LUL | P | 4 | 20 | Cine CT |
| 5 | LUL | P | 4 | 20 | Polaris |
| 6 | RUL | P | 3 | 30 | Polaris |
| 7 | RUL | S | 7 | 40 | Vision RT |
| 8 | RH | S | 7 | 40 | Vision RT |
| 9 | RLL | S | 7 | 40 | Vision RT |
| 10 | LUL | S | 7 | 40 | Vision RT |
| 11 | RUL | S | 7 | 40 | Vision RT |

Table 3.1. Details of the data acquired for each patient, including the tumour location (LUL - left upper lobe, RUL - right upper lobe, RH - right hilum, RLL - right lower lobe), the Cine CT protocol used (P for the protocol developed for the Philips MX8000 CT scanner, and S for the protocol developed for the Siemens Somaton Sensation CT scanner), the number of couch positions acquired, the number of Cine CT volumes acquired per couch position, and the respiratory signal used to derive the respiratory parameters.

data from the different couch positions can be combined to produce predicted volumes over a larger area.

Cine CT was originally acquired, by other groups, for sorting into 4DCT volumes (although 4DCT volumes can also be formed from helical CT data and Cone Beam CT data, see Section 2.2.2.3 and Section 2.2.2.4). Many research groups and clinical institutions use 4DCT data for studying respiratory motion. However, the main drawback with using current 4DCT data is that it assumes that all respiratory cycles are identical and that no inter-cycle variation occurs. Therefore the 4DCT volumes do not contain any useable information on how the respiratory motion can vary, and will contain artefacts when inter-cycle variation does occur. As it is well known that inter-cycle variation does occur (see Section 2.1.4.1) it was decided to use the unsorted Cine CT data (acquired over a few respiratory cycles per couch position) to try and sample some of the inter-cycle variation, so that it could be addressed by the motion models.

A CT scanner will acquire cine CT data at each couch position in one of two different ways, referred to in this thesis as the "stop-and-go" method and the "continuous" method. When using the stop-and-go method the CT gantry will perform a complete rotation (360°) acquiring enough data to reconstruct a volume. There will then be a brief period of 'dead-time', with the x-ray switched off and the gantry still, before the gantry performs another complete rotation (with the x-ray switched on) to acquire another volume (Figure 3.1a). This is repeated until the desired number of volumes have been acquired. When using the continuous method the CT gantry is continuously rotating and acquiring data during the acquisition time for each couch position. The desired number of volumes can then be reconstructed from this data. This number can be greater than the number of gantry rotations by using the 'sliding-window' technique, where the data used to reconstruct one volume overlaps with the data used to reconstruct the previous/next volume (Figure 3.1b).

### 3.3.1 Cine CT acquisition protocol P

The first Cine CT acquisition protocol was developed for the Philips MX8000 Multi-slice CT scanner and was used to acquire data from six patients. This scanner could acquire 16 x 1.5 mm slices simultaneously, giving a S-I coverage of 24 mm per couch position. Each Cine CT volume contained 512 x 512 x 16 voxels and the voxel dimensions were 0.68 mm x 0.68 mm x 1.5 mm. These were reconstructed using the standard in-slice field of view of 350 mm. The standard x-ray tube voltage of 120 kV and tube current of 100 mAs (per Cine CT volume) were used. Either three or four couch positions were acquired per patient, enough to fully cover the extent of the tumour and its respiratory motion. Table 3.1 gives the number of couch position, as well as the other details of the data, acquired for each of the patient.

The Philips MX8000 scanner uses the stop-and-go Cine acquisition method described above. A gantry rotation time of 0.75 seconds was used for patients 1-5. The scanner paused for approximately 0.25 seconds between rotations, and 20 volumes were acquired at each couch position, so data was acquired for approximately 20 seconds at each couch position. For the final patient acquired on the Philips scanner, Patient 6, a

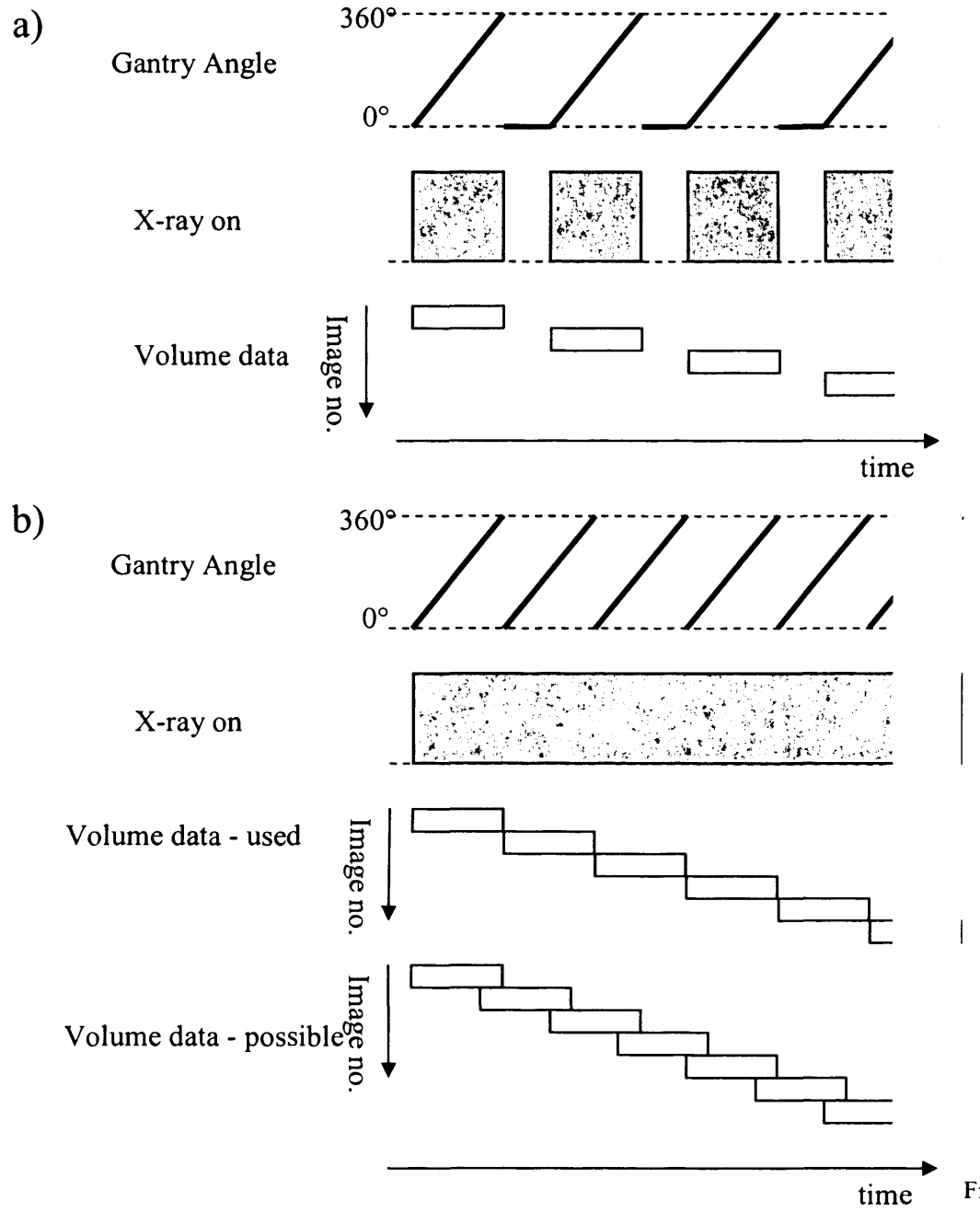


Figure 3.1 - a) Stop-and-go Cine CT acquisition method. The gantry performs a full 360° rotation with the x-ray on, acquiring enough data to reconstruct a single volume. There is then a short period of 'dead-time' while the gantry is stationary and the x-ray is switched off, before the gantry performs another full rotation with the x-ray on, acquiring the data for another volume. This is repeated until the desired number of volumes have been acquired.

b) Continuous Cine CT acquisition method. The gantry is continually rotating and acquiring data (i.e. the x-ray is on) for a predetermined period of time. Volumes can then be reconstructed from this continuous projection data. We reconstructed a separate volume for each 360° of projection data (Volume data - used), although it is possible to achieve a higher imaging frequency by reconstructing temporally adjacent volumes from overlapping projection data using the 'sliding window technique' (Volume data - possible).

faster gantry rotation time of 0.42 seconds was used. The pause between rotations was again approximately 0.25 seconds but now 30 volumes were acquired at each couch position, so the total acquisition time was again approximately 20 seconds per couch position.

### 3.3.2 Cine CT acquisition protocol S

The second Cine CT acquisition protocol was developed for the Siemens Somaton Sensation Multi-slice CT scanner and was used to acquire data from five patients. This scanner could acquire 12 x 2.4 mm slices, giving a S-I coverage of 28.8 mm per couch position. Each Cine CT volume contained 512 x 512 x 12 voxels with voxel dimensions in-slice field of view of 500 mm. The standard x-ray tube voltage of 120 kV was used but the x-ray tube current was lowered from the standard value of 80 mAs to 40 mAs (per Cine CT). Seven couch positions were acquired per patient.

The Siemens Somaton Sensation scanner uses the continuous Cine acquisition method described earlier. A gantry rotation time of 0.5 seconds was used and data was acquired over 20 seconds. The sliding window technique was not employed so 40 volumes were reconstructed per couch position.

### 3.3.3 Discussion of Cine CT issues

Cine CT was chosen as the method of acquiring internal respiratory data for a number of reasons. 3D volumes that were of high enough quality to permit non-rigid registration were required for building the motion models. This meant that either CT or MR data was required. In addition the data needed to be acquired while the patient was freely breathing, as they would be during RT treatment. It has been suggested that it may be possible to use volumes acquired at breath-hold, and to interpolate these to predict the anatomy during free breathing (see Section 2.5.2), although studies by other members of our group on MR data (Blackall et al. 2006), and performed as part of this research on CT data (see Section 7.3) have shown that breath-hold images are not a good predictor of free breathing motion. It was decided that this research would focus on the use of CT data, as CT data is currently required for RT planning, and because at the start of this

research MR images of the lung were not of high enough quality to permit non-rigid registration (Blackall et al. 2006). However, recent developments by other researchers from our group and from collaborating institutions have enabled high quality MR volumes, which cover the entire lungs and permit non-rigid registration, to be acquired at a temporal resolution comparable to cine CT while the patient is freely breathing (Miquel et al. 2007). These developments mean that MR could become a very useful modality for studying respiratory motion, and especially its variation, in the future.

It was decided that acquiring data over approximately 20 seconds (at each couch position), corresponding to 3-6 breath cycles in most patients, was a good balance between sampling some of the inter-cycle variation and not excessively increasing the scan time or the radiation dose delivered to the patient from the scan. In order to fully study and possibly model the inter-cycle variation it is expected that it would be necessary to acquire data for longer periods of time, and over multiple sessions for studying longer term (i.e. inter-fraction) variation. Such studies are very difficult with current CT technology as the radiation they deliver to the patient limits the length and number of CT scans that a patient can undergo. Therefore it was decided that devising methods of acquiring enough data to fully sample the inter-cycle variation was beyond the scope of this PhD project.

A gantry rotation time of 0.75 seconds was initially used as this was the default rotation time on the Philips scanner for Cine acquisition. This scanner was only capable of acquiring Cine data using the stop-and-go method. The pause time between gantry rotation could not be manually set, and was approximately 0.25 seconds. During the course of the research it was discovered that a faster rotation time of 0.42 seconds could be used. This meant that each volume was acquired in less time and so should be less susceptible to motion artefacts (Figure 3.2).

The faster gantry rotation time also meant that the number of volumes that could be acquired over the same 20 second period could be increased from 20 to 30, giving a higher temporal resolution, and hence more detailed information on the motion occurring during this time. However, the image quality depends on the amount of x-ray radiation delivered to the patient while acquiring each volume, the less radiation the more noise is present in the volume. Hence acquiring more volumes either means increasing the noise in the volumes or increasing the dose of radiation delivered to the patient. For the initial increase from 20 to 30 volumes it was decided to maintain the

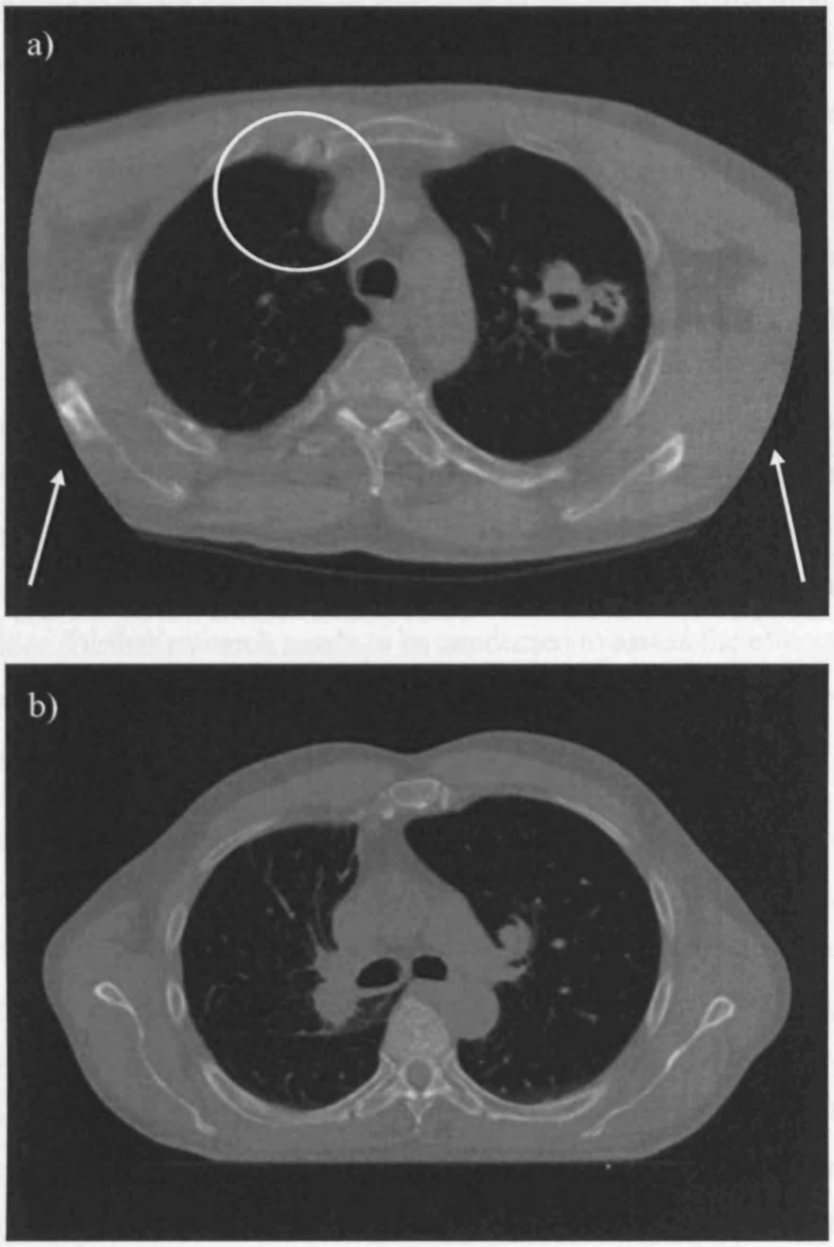


Figure 3.2 - a) a slice from a Cine CT volume acquired from patient 5 at a phase close to mid-inhalation (≈25%) on the Philips MX8000 CT scanner, b) a slice from a Cine CT volume acquired from patient 8 at a phase close to mid-inhalation (≈25%) on the Siemens Somaton Sensation CT scanner. It can be seen that the Cine CT volume from patient 5, which was acquired with a gantry rotation time of 0.75 seconds, contains much more severe motion artefacts (particularly in right lung, circled) than the Cine CT volume from patient 8, which was acquired with a gantry rotation time of 0.5 seconds. A direct comparison of motion artefacts is not possible as the scans are of different patients. Note, not all Cine CT volumes acquired with a 0.75 second gantry rotation time contained severe motion artefacts, the example shown is one of the worst cases observed. The 'false' edges produced by using an insufficient field-of-view can be seen in the Cine CT volume from patient 5 (indicated by arrows).

image quality and increase the radiation delivered to the patient, as the total radiation was still well within the limits for which ethical approval had been obtained (as only three couch positions were acquired for this patient the total number of Cine CT volumes acquired over all couch positions only increased from 80 to 90).

The fastest available gantry rotation time, 0.5 seconds, was used on the Siemens scanner. This scanner could acquire Cine CT data using both the continuous and stop-and-go methods. It was decided to use the continuous method as this gave a higher temporal resolution. The sliding window technique could have been used to reconstruct more volumes from the data and so increase the temporal resolution even further. This was not done as there was already a considerable amount of data to process for each patient, and the temporal resolution was already better than had been possible using the Philips scanner. Further research needs to be conducted to assess the effects of the temporal resolution on the motion model's accuracy and the amount of processing required to construct them.

As more volumes were acquired at each couch position (and more couch positions were acquired) with the Siemens scanner, it was necessary to reduce the x-ray tube current from the default 80 mAs per Cine CT volume to 40 mAs to prevent the x-ray tube from overheating and to reduce the radiation dose delivered to the patient (although the total dose using 80 mAs would still have just been within the limits of our ethics approval). Test volumes of a specially designed phantom acquired using 80 mAs and 40 mAs indicated that this reduction in the x-ray tube current caused little degradation in image quality, and it was not expected to affect the accuracy of the non-rigid registrations performed on the Cine CT volumes. Although further reductions in the x-ray tube current (or x-ray tube voltage) may noticeably degrade the image quality, the motion models presented in this thesis only require the Cine CT images to be of high enough quality to permit accurate non-rigid registration. It is thought that this will still be possible with images that are considerably lower than diagnostic quality. Therefore it is expected that the imaging dose delivered to the patient could be further reduced without affecting the accuracy of the motion models built from the data. More tests would be required to establish how much the imaging dose could be reduced without having an adverse effect, but it is estimated that reductions of one or even two orders of magnitude may be possible.

The minimum slice thickness and maximum number of slices permitted by both scanners were used so as to give the best spatial resolution over the largest area. The Siemens scanner covered a slightly larger area at each couch position, but the Philips scanner acquired more slices and these were considerably thinner, giving a higher spatial resolution than for the Siemens scanner. The higher spatial resolution should permit more accurate non-rigid registrations, and hence more accurate motion models. However, it is not clear whether the higher temporal resolution provided by the Siemens scanner or the higher spatial resolution provided by the Philips scanner will result in a more accurate motion model, and this will need to be further investigated in the future.

For the early acquisitions (on the Philips scanner) it was decided to only acquire enough couch positions to fully cover the tumour and the extent of its motion, as this is the most important area to validate the models on. It is acknowledged that data would be required over more couch positions for accurate dose calculations and clinical use, but it was thought that three or four couch positions per patients provided enough data to validate the motion models while not exposing the patients to excessive radiation dose or scan times. For later acquisitions (on the Siemens scanner) it was decided to increase the number of couch positions acquired to seven, so as to allow dose calculations (being undertaken by collaborating researchers) to be performed on the motion models results. It is again acknowledged that more couch positions may be required for some patients in clinical situations.

The standard (in slice) field-of-view of 350 mm used on the Philips scanner meant that for most Cine CT volumes some of the anatomy lay outside of the field of view and was not reconstructed, producing 'false' edges in the volumes (Figure 3.2). These 'false' edges had to be correctly accounted for when performing registrations on the data (see Section 4.1.2) and meant that the models could not accurately predict the anatomy in these regions. The Philips scanner did allow an extended field-of-view to be used, but the problems caused by the standard field-of-view were not fully realised until after all the data had been acquired on the Philips scanner. The extended field-of-view of 500 mm was used on the Siemens scanner as this was large enough to include all of the anatomy and prevent any 'false' edges from being produced.

## 3.4 Reference CT volume

A reference volume that contains all of the anatomy being modelled is also required for the motion models. As it is desirable to use a high quality, artefact free, volume as the reference volume, it was decided to acquire the reference volume at breath-hold. The breath-holds were performed at tidal exhale. This was initially thought to be a fairly arbitrary decision, and exhale was chosen as it has been shown by others that the location of the tumour is more reproducible at exhale (Seppenwoolde et al. 2002). However, reproducibility is only important if the Cine CT volumes are being concatenating together into a 4DCT volume, and this is being used as the reference volume (as has been proposed by other groups, Flampouri et al. 2006). As the reference volume used here is acquired in a single breath-hold reproducibility is not an issue. However, residual motion and gradual drifts during the breath-hold become a concern, and may be difficult to detect in the reconstructed volumes.

It could be argued that acquiring the breath-hold at inhale may be a better choice, as it is easier to hold your breath at inhale, and so there is less chance of the patient moving due to getting short of breath. However, it could also be argued that the anatomy occupies less space at exhale, enabling a smaller volume to be acquired which will take less time, and hence have less chance of patient motion. None of the breath-hold volumes acquired at exhale (or at inhale) have shown any signs of patient motion, even though some of the breath-hold scans can take up to 25 seconds. One reason that this has not yet caused problems may be that the longer scans are required by patients with larger lungs, who are probably more capable of holding their breath.

There are further advantages to both inhale and exhale breath-holds being used as the reference scans. The anatomy will appear larger in the inhale scan, so will be imaged in more detail, and that the registrations will be contracting the anatomy rather than expanding it. However, there is some evidence that the exhale breath hold scan is more similar to the free breathing anatomy (see section 7.3), so the registrations may have less deformation to account for when using an exhale breath hold. More detailed studies are required comparing the different possible reference volumes before any conclusions can be drawn as to which will produce the best results.

The breath-hold volumes had voxel dimensions of 0.98 mm x 0.98 mm x 1 mm. Each slice had a field of view of 500 mm and contained 512 x 512 voxels. Enough slices were

used to fully cover both lungs and 5-10 mm above and below the lungs (Figure 3.3). The number of slices required ranged between 224 and 355. A 2D projection "scout scan" was acquired prior to the rest of the CT data so that the coverage of the reference volume and the couch positions for the Cine CT scans could be set correctly.

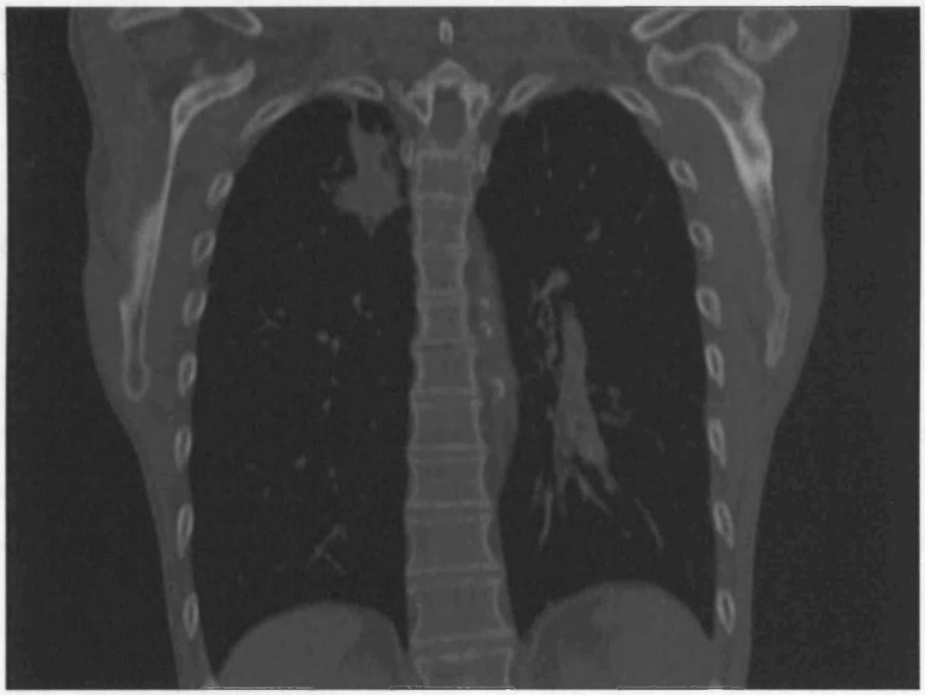

Figure 3.3 - A coronal slice from a reference CT volume covering the entire lungs acquired during an end-exhale breath-hold (from patient 7). It can be seen that the image is high quality and free of motion artefacts.

## 3.5 Respiratory surrogate signal and respiratory parameters

In order to construct the motion models the respiratory parameters, such as phase, that correspond to each of the Cine CT volumes need to be calculated. Although it may be possible to do this directly from the internal data, if the respiration is going to be monitored during treatment (e.g. for gated or tracked treatment deliveries) it would be preferable to use the same (or a similar) respiratory surrogate signal and parameters that will be used to monitor the respiration during treatment.

### 3.5.1 Respiratory surrogate signals

#### 3.5.1.1 Patients 1-4: extracted from the Cine CT volumes

There was no external respiratory surrogate signal acquired for the first four patients as this data had been acquired prior to the start of my PhD. Therefore the respiratory signal had to be derived from the Cine CT volumes. The skin surface was located in each volume. This could be achieved accurately using a simple thresh-holding method as there is a clear distinction between the skin and the surrounding air in the Cine CT volumes. The height of the skin surface was summed over all locations in the volumes so as to reduce the noise in the resulting respiratory signal.

#### 3.5.1.2 Patients 5-6: Polaris IR tracking system

For patients 5 and 6 the respiratory signal was generated by measuring the location of an Infra-Red (IR) reflective marker attached to the patient's abdomen (approximately midway between the Umbilicus and the Xiphisternum) using the Polaris IR tracking system (NDI, Waterloo, Ontario, Canada). This recorded the 3D location of the reflective marker approximately 15 times per second. A 1D signal was generated from the 3D locations by applying Principle Components Analysis (PCA) and using the first principle component as the respiratory signal. A stationary marker was also attached to the CT couch so that couch motion between different couch positions could be corrected for, giving a consistent signal across all couch positions.

Figure 3.4 - Example 3D skin surface acquired by Vision RT system from patient 7. You can see that most of the abdomen is acquired but some of the chest is not acquired due to occlusion (amount of surface acquired depended on patient and couch position). The surfaces from parts of the arms and face can be seen in the top half of the image.

### 3.5.1.3 Patients 7-11: Vision RT 3D surface acquisition system

For patients 7 to 11 the respiratory signal was derived from 3D surfaces acquired using a prototype system developed by Vision RT (London, UK). This system can acquire a 3D surface covering most of the patient's abdomen and part of the chest (Figure 3.4) between 10 and 15 times per second. The system employs stereo vision techniques to generate the surfaces, and consists of two video cameras and a projector which projects a pseudo random speckle pattern that is used to find correspondence in the stereo images.

The respiratory signal was derived from the surfaces using a technique developed by a colleague from our laboratory, Segolene Tarte (Tarte et al. 2006). This technique calculates the volume between the patient's surfaces and the CT couch to produce a respiratory signal. The volume is only calculated within a manually defined bounding box to avoid problems with occlusion.

The Vision RT system can reconstruct the 3D surface (and track multiple points on the surface) at approximately 10 frames per second (with the exact frame-rate depending on the hardware used and the size and resolution of the surface being reconstructed). Currently, the volume calculations are performed 'off-line' after the acquisitions, but it is expected that the calculations could be implemented to run in real-time for patient monitoring during treatment. The Vision RT system was calibrated so as to have the same axis as the CT scanner. This meant that the surfaces from different couch positions could easily be shifted to a common reference frame before calculating the volumes, giving a consistent signal across all couch positions.

### 3.5.2 Respiratory parameters

Three different respiratory parameters have been used for this research: the 'phase' of the respiratory cycle, and the 'amplitude' and the 'gradient' of the respiratory signal. The phase of the respiratory cycle indicates how far through the cycle you are, with one point in the respiratory cycle, usually end-exhale, being assigned to 0% (and 100%). The amplitude of the respiratory signal is just the value of the signal at the desired time point. The gradient of the respiratory signal is the time derivative of the signal at the desired point in time.

#### 3.5.2.1 Patients 1-4: using Cine CT derived signal

As the respiratory signal generated from the Cine CT volumes had a comparatively low temporal resolution it was not possible to accurately discern the end-inhale and end-exhale times. Therefore a sine wave was fitted to the respiratory signal using non-linear least squares, and the corresponding phase from the sine wave was assigned to each Cine CT volume (Figure 3.5).

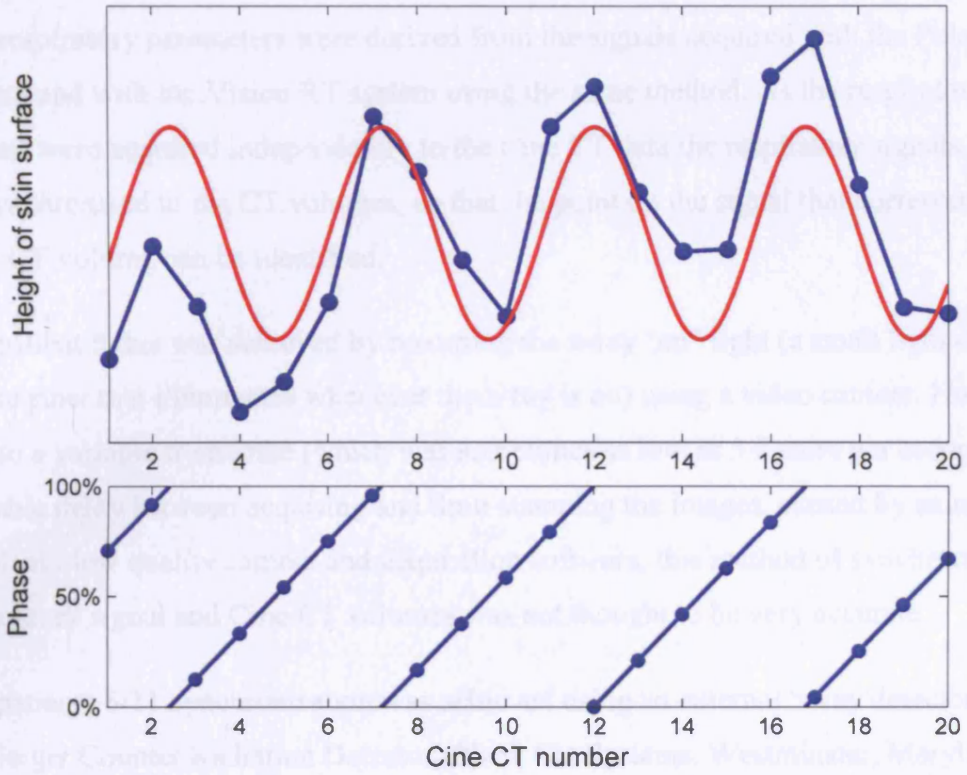


Figure 3.5 - Top - plot of skin height calculated from the Cine CT volumes (blue) for one couch position from Patient 3, with the fitted sine wave shown (red). Bottom - phase of fitted sine wave, markers indicate the respiratory phase assigned to the Cine CT volumes.

Although the amplitude value (the height of the skin) was known accurately for each Cine CT volume, the area of skin surface used to calculate this value was different for each couch position, so the amplitude values were not consistent across different couch positions. Due to the low temporal resolution of the signals it was not possible to accurately calculate the gradient of the signal. It is thought the phase values from the different couch positions should be reasonably consistent as there was only a maximum of four contiguous couch positions acquired. However, it is possible that there could have been small phase offsets between the respiratory signals generated for different couch positions, and having a common respiratory signal for all couch positions would be preferable.

### 3.5.2.2 Patients 5-11: using external respiratory surrogate signal

The respiratory parameters were derived from the signals acquired with the Polaris system and with the Vision RT system using the same method. As the respiratory signals were acquired independently to the Cine CT data the respiratory signals need to be synchronised to the CT volumes, so that the point on the signal that corresponds to each CT volume can be identified.

For patient 5 this was achieved by recording the x-ray "on" light (a small light on the CT scanner that illuminates whenever the x-ray is on) using a video camera. However, due to a variable frame rate (which was sometimes as low as 5 frames per second) and a variable delay between acquiring and time-stamping the images, caused by using a relatively low quality camera and acquisition software, this method of synchronising the respiratory signal and Cine CT volumes was not thought to be very accurate.

For patients 6-11 synchronisation was achieved using an external 'x-ray detector' (GM-10 Geiger Counter Radiation Detector, Black Cat Systems, Westminster, Maryland). This device is connected to the serial port of a PC, and once initialised sends a signal to the computer every time it detects some radiation. The x-ray detector signal was read approximately 15 times per second, and had a value of 0 (or occasionally 1) when the CT was not scanning and a value of 20-50 (depending on how close the detector was positioned to the CT bore) when it was.

As each Cine CT volume is acquired over a period of time the respiratory parameters assigned to the volume are those that correspond to the mid-time of the acquisition period for that volume. As the signals sometimes appeared to contain some noise they were first smoothed by convolving them with a Gaussian function with a standard deviation of 1 or 2 samples. The amplitude value was then calculated at the mid-time for each Cine CT by linearly interpolating the smoothed respiratory signal. The gradient of the signal was calculated by finding the difference and elapsed time between successive points in the signal, and then linearly interpolating these values to obtain the gradient value at the Cine CT mid-times.

To calculate the respiratory phase from the Polaris and Vision RT signals the end-exhale and end-inhale times were established. This was performed semi-automatically by finding the zero-crossings of the gradient of the signal. For a few couch positions from some patients extraneous zero-crossing were found that did not correspond to an

end-inhale or end-exhale time, and these needed to be manually detected and removed (Figure 3.6). The end-exhale times were assigned a phase of 0% (or 100%). The average position of the end-inhale times relative to the end-exhale time was calculated (for each patient), and this value was assigned as the phase of the end-inhale times (typical value between 40% and 60%). The phase at the Cine CT mid-times were then calculated by linearly interpolating between the phases assigned to the end-exhale and end-inhale times (Figure 3.6).

### 3.5.3 Discussion of respiratory signal and parameter issues

Although it was possible to produce a respiratory signal directly from the Cine CT volumes for patients 1-4, this signal had a number of disadvantages compared to the external signals later acquired with the Polaris and Vision RT systems. The signal from the Cine CT volumes had a low temporal resolution (approximately 1 frame per second) making it difficult to accurately calculate the phase and gradient parameters. The signals for different couch positions were derived from different parts of the skin surface, so it would not be possible to reliably combine the results of the models from different couch positions built using the amplitude values. As there appeared to be no phase offsets between the different couch positions it should be possible to combine the results of models built using phase from different couch positions. All of the Cine CT signals were derived from the skin surface in the chest region, which often exhibits less motion than the abdominal skin surface (and the internal anatomy). In addition, it is not possible to calculate the respiratory signal from the Cine CT volumes during treatment, although a signal derived from the same region of skin surface could potentially be acquired with the Vision RT system.

The signal acquired by the Polaris system (after being reduced to one dimension using PCA) is very similar to the signal acquired by the RPM system (made by Varian), which is one of the most commonly used respiratory monitoring systems. However, both the Polaris and the RPM system are susceptible to the problem that the signal can be very dependent on the location of the marker. It was possible to track multiple markers with the Polaris system (e.g. one on the chest and one on the abdomen), but due to difficulties attaching the markers this was not successively achieved for either patient (the RPM system can only track the displacement at one location).

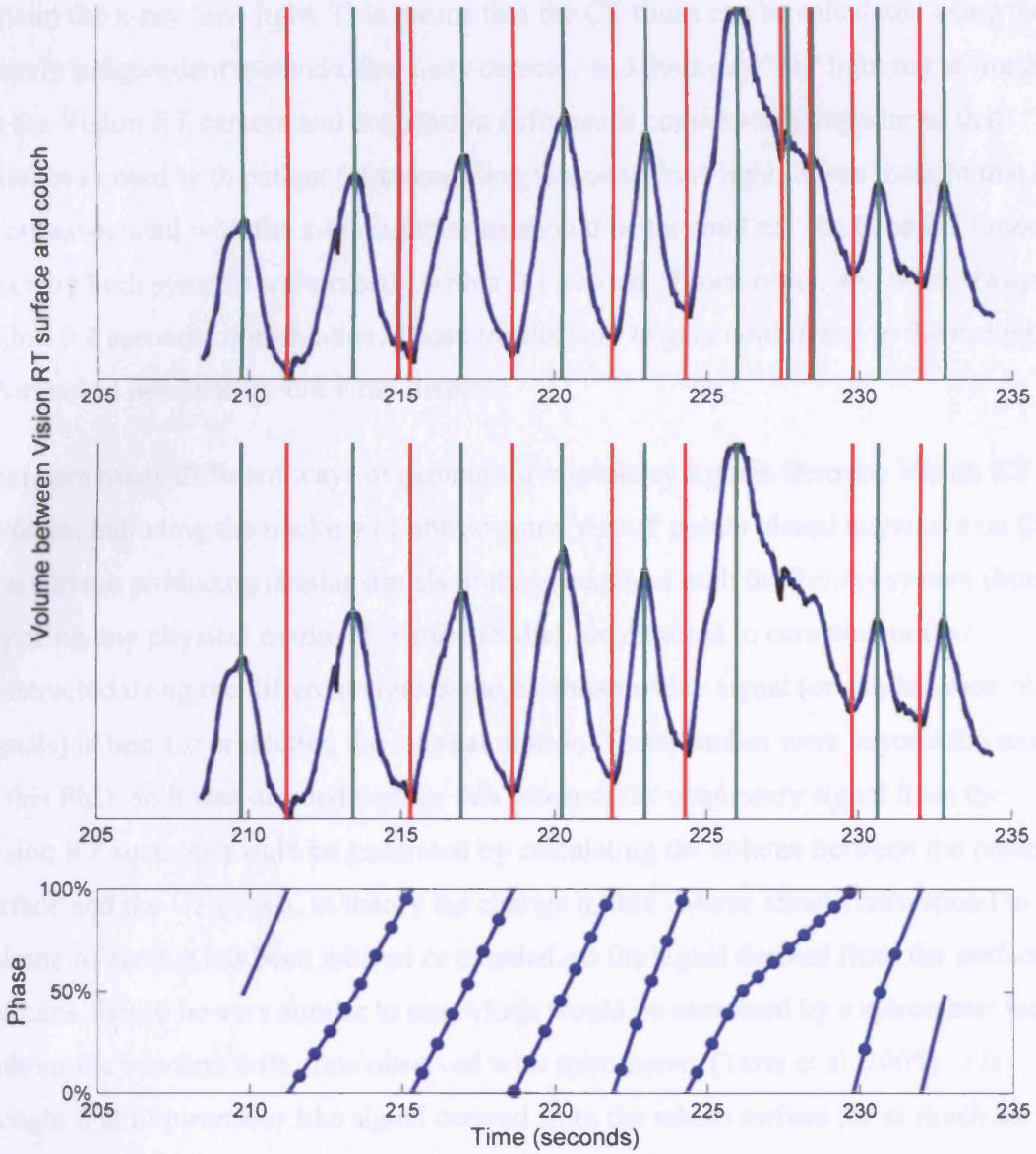


Figure 3.6 - Top and middle - respiratory signal generated from Vision RT surfaces, before (black) and after (blue) smoothing. Top - automatically detected end-inhale (green) and end-exhale (red) times, showing how spurious end-inhale and end-exhale times can sometimes be detected. Middle - the spurious end-inhale and end-exhale times need to be manually removed. Bottom - the respiratory phase calculated from the corrected end-inhale and end-exhale times (the markers indicate the mid-times of the Cine CT volumes).

It has been possible to assess the timing of the x-ray detector and Vision RT system, as the images from the Vision RT camera (from which the surfaces are reconstructed) contain the x-ray "on" light. This means that the CT times can be calculated using two entirely independent methods, the x-ray detector and the x-ray "on" light in the images. As the Vision RT camera and acquisition software is considerably superior to that which was used with patient 5 for recording the x-ray "on" light, it was thought that the errors associated with the x-ray light times should be far smaller. The Cine CT times given by both systems were mostly within 0.1 second of each other, and were always within 0.2 seconds of each other. These results help to give confidence in the timing information provided by the x-ray detector.

There are many different ways of generating respiratory signals from the Vision RT surfaces, including the tracking of one or more 'virtual' points placed anywhere on the skin surface producing similar signals to those acquired with the Polaris system (but not requiring any physical markers). Future studies are required to compare models constructed using the different signals and establish which signal (or combination of signals) is best for predicting the internal motion. These studies were beyond the scope of this PhD, so it was decided that for this research the respiratory signal from the Vision RT surfaces would be generated by calculating the volume between the patient's surface and the CT couch. In theory the change in this volume should correspond to the volume of air that has been inhaled or exhaled, so the signal derived from the surface volumes should be very similar to one which would be measured by a spirometer but without the baseline drift often observed with spirometers (Tarte et al. 2005). It is thought that a spirometer like signal derived from the whole surface (or as much as possible, as the complete surface is not acquired) will be better for modelling the respiratory motion than a signal derived from a single point on the surface (Lu 2005b), although more thorough investigations need to be conducted to confirm (or disprove) this hypothesis.

It is acknowledged that fitting a sine wave to the Cine CT signal is not a very accurate method of calculating the phase, particularly if the respiration is not very regular, and that there will be errors in the assigned phase values that may result in less accurate models. However, due to the low temporal resolution of the signal a more accurate method of assigning the phase could not be devised. The method of calculating the phase from the Polaris and Vision RT signals was much more accurate, and could correctly assign phase values in the presence of irregular breathing. There was only one

couch position (couch position 1 from patient 7) where it was not possible to assign phase values to some of the Cine CT volumes due to extremely irregular respiration. As this couch positions was the most Superior couch position and only contained a few slices of lung, it was expected that there would be very little motion occurring at this couch position anyway, and it was decided to exclude the data from this couch position from the motion modelling (but not registration) experiments.

Neither of the methods for calculating the phase values could be implemented to run in real-time, so neither could be used to monitor the patient's respiration during treatment. Methods of calculating the phase in real-time exist and are available with some commercial products (e.g. the RPM system from Varian), so monitoring the phase during treatment is feasible. It would have been more meaningful to have calculated the phase for the models using a method that could also have been used to calculate the phase during treatment, but devising an accurate method of calculating the phase in real-time is non-trivial and was beyond the scope of this PhD, and none of the commercial real-time phase detection systems were available.

## 3.6 Summary

This chapter has described the data that was acquired for constructing the motion models and the protocols used for acquiring this data, and has discussed many of the issues relevant to the data acquisition. Cine CT data was acquired at several couch positions per patient while they were freely breathing to study the motion of the internal anatomy during respiration. A respiratory signal was simultaneously acquired for calculating the phase, amplitude, and gradient respiratory parameters. This was derived directly from the Cine CT volumes for some patients (1-4), acquired using the Polaris system for some patients (5-6), and acquired using the Vision RT system for the others (7-11). Also, a reference CT volume was acquired for each patient. This was acquired at tidal exhale breath-hold so that the volume was high quality and artefact free.

# Chapter 4

# Non-rigid registrations

This chapter describes the non-rigid registrations that have been performed and evaluates their results. In order to construct the motion models described in chapters 5 and 6 it is necessary to register the reference CT volume to the Cine CT volumes. The registrations will then describe how the anatomy in the reference volume has moved and deformed, so as to appear as it does in the Cine CT volumes.

## 4.1 Protocols used for registering lung Cine CT volumes

All the registrations that have been performed as part of this work have used a B-spline transformation model (see Section 2.4.1.3.2). The main reasons for this were that the B-spline transformation model was well understood and produced good results in the majority of cases. Another advantage of using the B-spline transformation model over a

non-parametric transformation model is that the displacement of each control point can be modelled rather than the displacement of each voxel, resulting in more compact models. Sum of Square Differences (SSD) has been used as the similarity measure for all registration performed as part of this research, as the anatomy will have the same intensities in both images being registered (neglecting density change effects and image noise). Density change effects have not been explicitly addressed when performing the registrations, as it was thought that the structure within the lungs combined with the inherent smoothness of the B-spline transformation model will prevent the density change effect causing errors in the registrations. The gradient descent (estimated via finite differences) optimisation method has been used for these registrations as it is easily understood and there was a useable implementation already available within our laboratory.

The same transformation model and general algorithm have been used for all of the lung Cine CT registrations performed throughout the course of this research but the parameters used and the protocol followed for the registrations have evolved during the research. There are three main registration protocols that have been used at some time during the research. There now follows a brief description of each, the differences between them, and the reasons for changing them.

### 4.1.1 Registration protocol 1

The registrations that were initially used to construct the motion models were performed by Adam Chandler as part of the work for his PhD (Chandler 2005). Prior to performing the registrations the ipsilateral lung was segmented in each of the Cine CT volumes, the Target images, to prevent problems caused by sliding tissue (see Section 2.4.5.2) and to speed up the registrations. Although the segmentation was performed semi-automatically (using the Analyze software package, Mayo Clinic, Rochester, Minnesota) it could still take up to a few hours per patient due to the large amount of data. Only one image resolution, with the original voxel dimensions (0.68 mm x 0.68 mm x 1.5 mm) was used for the registrations. The control point spacing was approximately 20 mm in plane but only 3.75 mm through plane. Normalise Mutual Information (NMI) was used as the similarity measure. The full parameters used for the registrations are given in chapter 7 of Adam Chandler's PhD thesis (Chandler 2005).

Each registration result had an affine and a B-spline component. It was decided to 'incorporate' the affine component into the B-spline transformation, so that only the B-spline control point displacements would need to be modelled. This was done by applying the affine transformation at the location of each of the B-spline control points to find the displacement from the affine transformation, and adding this to the control point displacement. The error in incorporating the affine component into the B-spline transformation in this way was assessed on several registrations and in all cases was found to be in the order of $10^{-14}$ mm, i.e. negligible.

### 4.1.2 Registration protocol 2

So that the motion models could be used to perform full dosimetric calculations it was decided to attempt to perform the registrations on the entire anatomy and not just the segmented ipsilateral lung. This meant that the run time was increased considerably (as many more voxels are involved in the registrations) and the results were susceptible to problems caused by sliding tissue. All voxels outside of the patient were set to the background value using a simple thresh-holding technique. A 'border' of approximately 10 voxels outside the patient's skin was left with its original intensity values so that the skin surface would be used to guide the registrations. Care had to be taken with the images that had been reconstructed using the standard field of view as they contained some 'false' edges due to truncation at the edge of the field of view (see Section 3.3.3 and Figure 3.2). These 'false' edges had to be ignored in the registrations since a corresponding edge was not present in the reference images. This was achieved by only leaving the border of voxels at 'real' edges caused by the skin surface, and not at the false edges caused by the restricted field of view.

Again, each Cine CT was registered independently and an affine registration was initially performed. The affine registrations were performed with the same parameters that had been previously used for the segmented lung registrations. As these produced reasonable results, achieved a relatively short run time in comparison to the non-rigid registrations, and were only being used to obtain a good starting estimate for the B-spline registrations, the affine parameters were not investigated further.

The B-spline registrations used a multi-resolution approach for the images but not for the CP (Control Point) grid. The registration were run with two image resolution levels,

with voxel dimensions of 2.72 mm x 2.72 mm x 6 mm (four times the original dimensions) and 1.36 mm x 1.36 mm x 3 mm (two times the original dimensions) in the target image. Tests were performed that proceeded to use the original images as a third resolution level, but these showed that the extra resolution level increased the run time of the registrations by a factor of 20 with a negligible increase to their accuracy. The B-spline CP spacing was approximately 20 mm in all directions.

### 4.1.3 Registration protocol 3

The final registration protocol employed in this research features a number of improvements over the previous protocols, as well as some modifications so that the results from the models for the different couch positions can be combined into a single transformation (see Section 5.5).

The registrations for patients 1-6 (who had Cine CT volumes with 1.5 mm slices) were again performed using two image resolution levels. However, the images were now re-sampled using 3 mm (x 3mm x 3 mm) isotropic voxels for the first image resolution level, and using 1.5 mm isotropic voxels for the second. Isotropic voxels were used so that the images had a superior through-slice resolution to the previous protocol, but still contained less voxels than the original images (so the registration ran quicker). The registrations for patient 7-11 (who had Cine CT volumes with 2.4 mm slices) were only performed at a single resolution level using 2.4 mm isotropic voxels. Test registrations that started with a coarser image resolution were also performed for these patients, but were found to give no improvement over the results using a single resolution level, and in some cases made the results worse. The test registrations were performed using both 4.8 mm followed by 2.4 mm isotropic voxels, and using 3 mm followed by 1.5 mm isotropic voxels.

The B-spline registrations were now performed sequentially in a chain, using the result from the previous registration as the starting estimate for the next registration. It was thought that the previous B-spline registration may provide a better starting estimate than an affine registration result, particularly if there were large non-rigid deformations occurring. For each couch position the Cine CT volume with a respiratory phase closest to end exhale was registered first, and the result of this registration was then used as input to the registrations of the Cine CT volumes acquired immediately before and after

the first volume. The result of each registration was then used as input to the registration of the next volume until all the volumes from that couch position had been registered. This meant that only the first registration used the affine input (approximated as a B-spline transformation) the others all used the result from the previous registration. When the registrations were performed using two image resolution levels it was the result from the first resolution level that was used as input to the next registration.

Usually the CP grid is constructed so that it just covers the image being registered, with CPs placed at the edge of the image (and one outside these so that the transformation is defined to the edge of the image). However, the registration can be performed using an extended CP grid defining a transformation over a larger area (or a smaller one, but then the registration will not be defined over the whole image). So that the results from the models for the different couch positions can be combined into a single transformation (see Section 5.5), the registrations were now performed with an extended CP grid that defined a transform over the area covered by all of the couch positions, not just the one for the image being registered (Figure 4.1). This meant that some of the CPs had no effect on the image being registered and hence were not optimised as part of that registration. In order to assess the effects of using the extended CP grid the registrations were also performed using a standard CP grid that just defined the transform over the image being registered. A CP spacing of approximately 20 mm was again used.

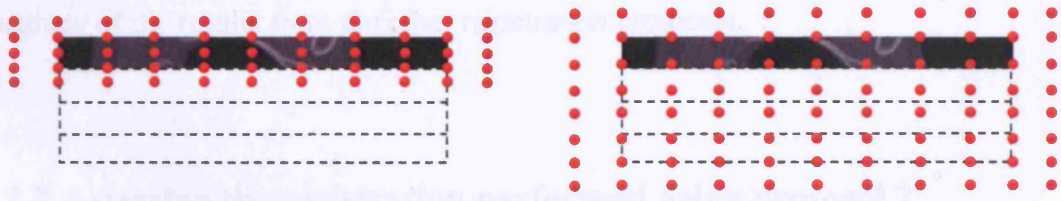

Figure 4.1 - Illustration of Control Point (CP) placement in the standard (left) and extended (right) CP grids. The Cine CT volume being registered (from the most Superior couch position) is shown, the dashed lines indicate the other three couch positions where data has been acquired, and the CP positions are shown in red. It can be seen that the standard CP grid only covers the image being registered (plus one CP outside so that the deformation is defined over the whole image), whereas the extended CP grid covers all of the couch positions where data has been acquired.

## 4.2 Assessing the lung Cine CT registrations

Assessing non-rigid registrations is a very difficult task and several techniques have been employed, described in the following sections.

### 4.2.1 Assessing the registration performed using protocol 1

The segmented lung registrations (performed using protocol 1) were assessed by Adam Chandler as part of the work for his PhD. The tumour and ipsilateral lung were delineated in all of the Cine CT volumes and the reference volume by a clinical oncologist, and these delineations were used to assess the registrations using a number of different measures. For a full explanation of these measures see chapter 7 of Adam Chandler's PhD thesis (Chandler 2005).

After studying the results it was discovered that there were some large inconsistencies in the delineations, even though they were all performed by the same clinical oncologist in a single session. These were due to ambiguities in the images, particularly for patients 1 and 3 where the tumour is located at the edge of the lung, making the tumour and the lung boundaries very hard to identify. These inconsistencies reduced the values of the accuracy measures and meant that the results were less meaningful as it was often found that it was the delineations rather than the registrations that were the cause of bad results. For these reasons it was decided not to use delineations for assessing the accuracy of the results from the other registration protocols.

### 4.2.2 Assessing the registration performed using protocol 2

The registration results achieved using registration protocol 2 were assessed using two techniques.

#### 4.2.2.1 Expert visual assessment

The reference CT was deformed by each of the registration results and was compared to the corresponding Cine CT volume to visually assess the registration results. For each

result three orthogonal slices passing through the centre of the tumour were selected. A combination of overlays and interactive displays were then used to examine the differences between the deformed and the target volume, and, where necessary, direct measurements of the degree of misalignment were taken.

Each of the results were given three scores by a clinical oncologist, firstly for the tumour, secondly for the ipsilateral lung, and thirdly for any other tissue. This is because it is more important to predict the motion of the tumour and (to a lesser degree) the ipsilateral lung correctly, as most of the dose will be distributed in these regions. The results were scored according to the following criteria:

1. No visually detectable misalignment over 1.5 mm (one slice thickness).
2. No visually detectable misalignment over 4.5 mm (three times the slice thickness).
3. Visually detectable misalignment over 4.5 mm.

We specified that at least 90% of the region of interest must meet the criteria above for that region to be given the corresponding score, so that very small areas of misalignment would not affect the score for the whole region. Figure 4.2 shows some example registration results to demonstrate the visual assessment scoring. There is an example of a 'bad' registration, which was given a score of 2 for the tumour and 3 for the ipsilateral lung and the other tissue, and an example of a 'good' registration, which was given a score of 1 for all regions.

In order to assess the intra-user reproducibility of the visual assessment, it was repeated for two patients (using a total of 170 registrations) a few weeks after the initial assessment by the same clinical oncologist.

### 4.2.2.2 Landmark tracking

A clinical oncologist was asked to select an easily identifiable anatomical point, a landmark, in each of the Cine CT volumes acquired for a particular couch position, and then to attempt to identify the same landmark in the reference volume.

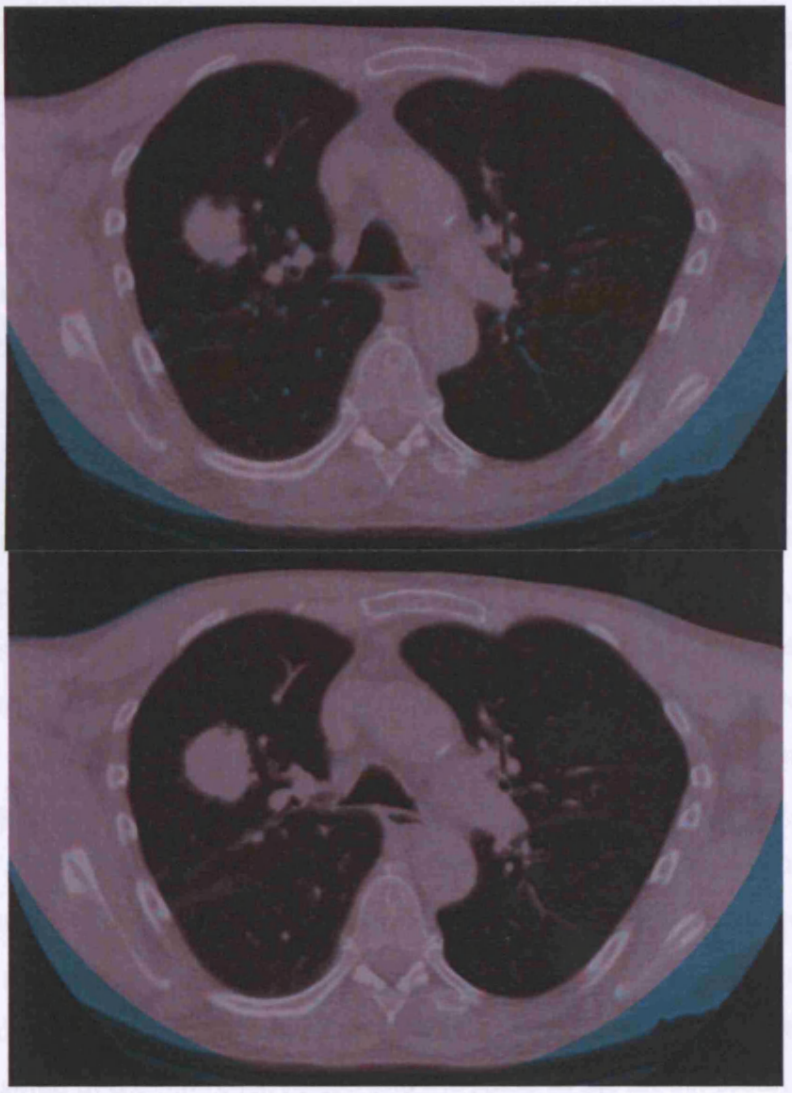

Figure 4.2 – Example registration results (from patient 2) to demonstrate the visual assessment scoring. The results are displayed using a colour overlay, with the Cine CT shown in red and the reference volume deformed by the registration result shown in cyan. This causes voxels which have the same intensity in each volume to appear grey. Top – example of a 'bad' registration, which received a score of 2 for the tumour and 3 for the ipsilateral lung and other tissue when being visually assessed. Bottom – example of a 'good' registration, which received a score of 1 for all regions.

This was repeated using a different landmark for each couch position, for all couch positions from each of the five patients that this registration protocol was assessed on. Although only one landmark was chosen per couch position, a variety of different landmarks were chosen overall, including some which exhibited negligible motion over

the respiratory cycle, some which exhibited similar motion to the tumours over the respiratory cycle (including some landmarks within or on the boundary of the tumours), and landmarks that exhibited different motion to the tumours over the respiratory cycle. For all of the couch positions, landmarks were chosen that remained within the extent of the Cine CT volume for all of the volumes acquired.

The clinical oncologist then attempted to re-identify the same landmark in the Cine CT volumes on a different occasion, in order to estimate the localization error for manually identifying the landmarks. It should be noted that this is only an estimate of the error for identifying the landmark in the same volume, and not the error for identifying the same landmark in different volumes, which may well be larger.

The reference landmark was transformed using an estimate of the inverse of each of the non-rigid registration results (see Appendix A). For each Cine CT volume, the Euclidean distance was calculated between the two locations of the landmark identified on different occasions – an estimate of the manual Target Localisation Error (TLE), and the midpoint of the two landmark locations and the reference point transformed by the (inverse of the) non-rigid registration result – an estimate of the Target Registration Error (TRE). The distance between the Cine CT landmarks and the reference points prior to registration, the Target Error Before Registration (TEBR), was also calculated. To calculate meaningful TEBR values a rigid registration was used to align the reference and Cine CT spaces. The rigid registration was performed once per patient between the spine in the reference image and the spine in the model predicted end exhale volume (over all couch positions from the combined model, see section 5.5).

### 4.2.3 Assessing the registration performed using protocol 3

The results from registration protocol 3 were assessed using similar techniques to those used for assessing protocol 2. Visual assessment was again performed, but this was only done subjectively to see if any large errors had occurred, and was not performed by a clinical expert or scored as before. Landmark tracking was performed in exactly the same way as for the previous protocol, and the same landmarks were used for the patients that were registered using both protocols. The registrations performed with the standard CP grid and the registrations performed with the extended CP grid were both assessed so that the effects of using the extended grid could be evaluated.

## 4.3 Results of assessing the lung Cine CT registrations

### 4.3.1 Results of assessing the registrations performed using protocol 1

Registrations protocol 1 was assessed on data from four patients (1, 2, 4, and 5) using a total of 330 registrations. The mean (standard deviation) percentage overlap was 95.4% (0.01%) for the lung delineations, and 86.62% (0.03%) for the tumour delineations. The mean error for the centre of mass of the tumour was 1.79 mm, and the mean error at the edge of the tumour was 2.30 mm. More detailed results can be found in Chapter 7 of Adam Chandler's PhD thesis (Chandler 2005). As previously mentioned the results were strongly affected by inconsistencies in the segmentations, but still indicated that the registrations had done a reasonably good job of recovering the deformation.

### 4.3.2 Results of assessing the registrations performed using protocol 2

Registration protocol 2 has been assessed on data from five patients (2-6) using a total of 330 registrations. The registrations were also attempted on another patient (1) but the registrations were judged to have failed and so were not assessed further. The reason that the registrations failed for this patient was because of the sliding tissue problem (see Section 2.4.5.2). The tumour was located very near the back edge of the lung in close proximity to the ribs, and the tumour would slide past the ribs during respiration. The sliding tissue problem appears to have only caused small errors, if any, in the results for the other patients.

The exact run time taken for the registrations was not recorded but was between one and three hours (on a 2 GHz CPU) for most of the registrations. These results were first published in McClelland et al. (2006a).

#### 4.3.2.1 Expert visual assessment results

The registrations were scored in three regions, the tumour, the ipsilateral lung, and the rest of the anatomy. For the tumour 64.7% of the registrations were given a score of 1 (no visually detectable misalignment over 1.5 mm), 33.6% a score of 2 (no visually detectable misalignment over 4.5 mm), and only 1.7% a score of 3 (visually detectable

misalignment over 4.5 mm). These results are encouraging with the tumour very rarely exhibiting large errors. For the ipsilateral lung 41.3% of the registrations were given a score of 1, 49.8% a score of 2, and 8.8% a score of 3. Most of the registrations still exhibited reasonable alignment for the ipsilateral, lung but there was more misalignment than for the tumour as it was not always possible to correctly align all of the internal structure of the lung using the 20 mm control point spacing chosen. For the rest of the anatomy only 2.4% of the registrations were given a score of 1, 80.9% a score of 2, and 16.7% a score of 3. There was more misalignment exhibited for the rest of the anatomy but this was expected as all of the remaining anatomy were considered and there were frequently small changes in posture or other changes such as the patient swallowing for which the full motion and deformation could not be recovered using the control point spacing chosen. However, these misalignments were mostly less than 4.5 mm, and restricted to small areas that would generally have a negligible effect on any radiotherapy plans and calculations.

Given the subjective nature of the visual assessment, and that there was no guarantee that exactly the same three orthogonal slices were chosen when the experiment was repeated, the results demonstrated a good degree of reproducibility. None of the repeat scores differed from the original by 2, i.e. none of the results were given a score of 1 on one occasion and 3 on the other. For the tumour 70% of the repeat scores were the same as the original scores, for the ipsilateral lung 71.2% of the repeat scores were the same, and for the "other tissue" 91.2% of the repeat scores were the same.

### 4.3.2.2 Landmark tracking results

The mean, standard deviation, and maximum value for the Target Error Before Registration (TEBR), the Target Localisation Error (TLE) and the Target Registration Error (TRE) for each patient, and over all patients that were assessed, can be found in Table 4.1. It can be seen that the landmarks can be located very accurately, with a mean TLE of 0.85 mm, just over half a slice thickness (0.75 mm), and a maximum TLE of 2.76 mm, under 2 slice thicknesses (3 mm). The registrations have also performed very well with a mean error over all the data of 1.26 mm, less than the slice thickness of 1.5 mm. The standard deviation of the TREs is very similar to that of the TLEs, 0.67 and 0.64 respectively. The maximum TRE of 6.24 mm is reasonable, but when the results

for each patient are examined, it can be seen that only one of the five patients had a maximum TRE this high (patient 2), and the maximum for the other four is 3.32 mm, indicating that the maximum TRE is just over two slice thicknesses for most patients. The TEBR values show that without performing the non-rigid registrations there are relatively large errors, as would be expected due to the fact that (at least some of) the landmark points move with respiration, and because the reference volume is acquired at breath-hold which does not always give a good representation of the anatomy during free breathing (see section 7.3).

| Patient | TEBR (mm) | | | TLE (mm) | | | TRE (mm) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. |
| 2 | 4.95 | 2.19 | 11.25 | 0.72 | 0.48 | 2.08 | 1.65 | 0.92 | 6.24 |
| 3 | 5.48 | 2.70 | 11.65 | 0.76 | 0.58 | 2.18 | 1.14 | 0.53 | 3.28 |
| 4 | 4.49 | 2.49 | 10.58 | 0.47 | 0.55 | 2.23 | 1.07 | 0.64 | 3.32 |
| 5 | 5.08 | 2.44 | 10.51 | 1.35 | 0.54 | 2.76 | 1.39 | 0.68 | 2.99 |
| 6 | 4.58 | 1.88 | 10.00 | 0.92 | 0.67 | 2.73 | 1.17 | 0.45 | 2.17 |
| All | 4.87 | 2.35 | 11.65 | 0.85 | 0.64 | 2.76 | 1.26 | 0.67 | 6.24 |

Table 4.1 - The landmark tracking results for registration protocol 2. The mean, standard deviation ($\sigma$), and maximum Target Error Before Registration (TEBR), Target Localisation Error (TLE), and Target Registration Error (TRE) are given for each patient and over all patients that were assessed (2-6).

### 4.3.3 Results of assessing the registrations performed using protocol 3

Registrations from protocol 3 have been assessed on data from seven patients (2-8) using a total of 930 registrations. As registration protocol 3 also used images of the full anatomy it was not expected to perform any better that protocol 2 on the data from patient 1, so the registrations were not attempted on this patient. However, the registrations were attempted on data from a further three patients (9-11) and visual inspection revealed that the registrations failed on some of the Cine CT volumes (usually from couch positions nearer the diaphragm) from all of these patients. It was thought that the sliding problem (see Section 2.4.5.2) was again the cause of some of the failed registrations. The registrations had been successful for many of the Cine CT volumes from patients 9-11, and it may have been possible to use the successful registration results to construct models for some, if not all, of the couch positions from

these patients. However, it was decided not to use these patients to assess the registrations (or construct the motion models) until the cause of the failed registrations could be further investigated (and unfortunately this proved to be beyond the scope of this PhD).

The results of the landmark tracking results for patients 2-8 are given in table 4.2. The values for the TLE for patients 2-6 are the same as when assessing protocol 2, as the same landmark points were used. The mean TLE values for patients 7 and 8 are less than for patients 2-6. It is thought that this is because the larger slice thickness used for patients 7 and 8 meant that there was less ambiguity about which slice a landmark was located in. However, the larger slice thickness used for these patients means that there is actually likely to be more error associated with the landmark locations in these patients, even though the TLE values were smaller (this is a problem with estimating the TLE in this way).

The Target Registration Error ($TRE_{Standard}$) values when using the standard CP (Control Point) grid were similar to those from assessing protocol 2. The results for the two patients with the larger 2.4 mm slices (patients 7 and 8), were comparable to the results for the other patients with 1.5 mm slices (the mean values were slightly better than most of the other patients but the maximum values were worse). The mean value of the $TRE_{Standard}$ over all patients was 1.12 mm. The exact run time for all of the registrations was also recorded and the mean time was just over one hour with a maximum run time of just under two and a half hours (on a 2 GHz. CPU). The registrations for the patients 7 and 8 were considerably faster than the other registration (mean time of approximately 45 minutes compared to 85 minutes for the other patients) as they only used one image resolution level.

The Target Registration Error ($TRE_{Extended}$) values for the extended CP grid were similar to those for the standard CP grid ($TRE_{Standard}$), indicating that using the extended CP grid had little effect on the accuracy of the registrations. However, the run time of the registrations was increased when using the extended CP grid, with a mean run time of approximately one and a half hours and a maximum run time of just over three and a half hours.

| Patient | TEBR (mm) | | | TLE (mm) | | | $TRE_{Standard}$ (mm) | | | $TRE_{Extended}$ (mm) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. |
| 2 | 4.95 | 2.19 | 11.25 | 0.72 | 0.48 | 2.08 | 1.72 | 0.97 | 6.37 | 1.76 | 0.97 | 6.29 |
| 3 | 5.48 | 2.70 | 11.65 | 0.76 | 0.58 | 2.18 | 1.20 | 0.46 | 1.93 | 1.33 | 0.53 | 2.34 |
| 4 | 4.49 | 2.49 | 10.58 | 0.47 | 0.55 | 2.23 | 1.12 | 0.53 | 3.43 | 1.12 | 0.53 | 3.23 |
| 5 | 5.08 | 2.44 | 10.51 | 1.35 | 0.54 | 2.76 | 0.98 | 0.57 | 2.63 | 1.05 | 0.56 | 2.84 |
| 6 | 4.58 | 1.88 | 10.00 | 0.92 | 0.67 | 2.73 | 1.21 | 0.49 | 2.46 | 1.22 | 0.50 | 2.75 |
| 7 | 16.06 | 3.29 | 28.53 | 0.50 | 0.44 | 2.60 | 1.13 | 0.52 | 3.61 | 1.12 | 0.52 | 3.21 |
| 8 | 2.59 | 1.43 | 7.05 | 0.55 | 0.54 | 2.89 | 0.97 | 0.60 | 3.15 | 0.97 | 0.63 | 3.25 |
| All | 7.56 | 6.18 | 28.53 | 0.66 | 0.58 | 2.89 | 1.12 | 0.61 | 6.37 | 1.13 | 0.62 | 6.29 |

Table 4.2 - The landmark tracking results for registration protocol 3. The mean, standard deviation ($\sigma$), and maximum Target Error Before Registration (TEBR), Target Localisation Error (TLE), and Target Registration Error for the registrations with the standard control point grid ($TRE_{Standard}$) and for the registrations with the extended control point grid ($TRE_{Extended}$) are given for each patient and over all patients that protocol 3 was assessed on (2-8).

## 4.4 Conclusion

The results show that the registrations used to construct the motion models were generally very accurate. Using an extended CP grid does not appear to have any effects on the accuracy of the registrations, but it did cause an increase in their run time (although it is thought that was mainly due to using an inefficient implementation of the registration algorithm). The TRE values would indicate that the Cine CT volumes with 2.4 mm slices can be registered just as accurately as the Cine CT volumes with 1.5 mm slices. However, the registrations failed for three of the five patients with 2.4 mm slices, compared to only one of the six patients with 1.5 mm slices, which indicates that the thicker slices may be contributing to the failed registrations.

Obviously, having some of the registrations fail for such a large proportion of the patients (particularly when 2.4 mm slices were used) is very concerning and needs to be addressed. Unfortunately, the majority of these failed registrations were performed towards the end of this research, and so it was deemed beyond the scope of this PhD to address the failed registration. It was thought that the sliding problem (see Section 2.4.5.2) was the cause of many of the failed registrations, although there are many other factors that could have contributed, including: the density change problem (see section 2.4.5.1), registering relatively thin 'slabs' of data (only 12 or 16 slices), not having good enough starting estimates for the registrations, and possible 'bugs' in the code used to perform the registrations.

All of these issues will need to be addressed in future work. In addition, different protocols, different implementations of the B-spline registration algorithm, and different registration algorithms all need to be investigated to see if they can improve the robustness of the registrations (without degrading the accuracy). The run time of the registrations also needs to be reduced considerably, although it is thought that the use of more efficient optimisation schemes (Klein et al. 2005) and possibly GPUs (Graphics Processing Units, Sharp et al. 2007) should help to achieve this.

In conclusion there remains much work that needs to be done before registrations similar to those performed in this chapter are sufficiently robust and fast enough to be used routinely in a clinical situation. However, the considerable experience gained from performing and evaluating the registration in this chapter has lead me to believe that this

is an achievable goal, and I expect that many of the registration problems currently faced will be solved in the next few years. Additionally, it should be stressed that the main focus of this PhD has been the construction of the motion models and not on performing the registrations. The results in section 4.3 show that the registrations for patients 2-8, that have been used for constructing the motion models in the following chapters, had a high level of accuracy (given the resolution of the data), and should be more than sufficient for constructing and assessing the motion models.

# Chapter 5

# One parameter motion models

This chapter describes how to construct motion models that relate the internal motion to a single respiratory parameter, the phase of the respiratory cycle. It evaluates a number of different potential functions that can be used to describe the relationship between the phase and the internal motion. It then assesses the ability of the motion models to predict the CT data. Finally it explains how the model predictions from different couch positions can be combined into a single continuous prediction over all of the couch positions.

## 5.1 Introduction

The motion models developed for this research are built from the results of the non-rigid registrations described in the previous chapter. The output of the motion model is a non-rigid transformation that can deform the reference CT volume to produce a prediction of the Cine CT volume corresponding to the chosen respiratory parameter. The transformation can also be used to propagate target delineations, perform composite dose calculations, and to provide the deformation maps required for tracked treatment delivery (see Section 7.5). The motion models are continuous in space and over the respiratory parameter, allowing the deformation to be calculated at exactly the desired point in space and value of the respiratory parameter.

As the data is acquired for each couch position separately, each couch position must be modelled separately. The results from each model can simply be concatenated together over all the couch positions. However this will produce discontinuities in the deformation field at the boundaries between adjacent couch positions, which can cause artefacts in the predicted volumes (similar to, but usually smaller than, those seen in 4DCT volumes) and errors in planning calculations. A more advanced technique has been developed for combining the results from the models for different couch positions into a single, continuous, non-rigid transformation and this is described and evaluated in Section 5.5.

The motion models in this chapter relate the internal motion to a single respiratory parameter. This means that the motion models can model the variations in the registration results that correspond to changes in the respiratory parameter, but average out variations in the data that occur at the same value of the respiratory parameter. The phase of the respiratory cycle (calculated as described in Section 3.5.2) was used as the respiratory parameter. Phase was chosen for a number of reasons. It is one of the two most widely used parameters in the literature. It could be calculated for the patients for whom there was no external respiratory signal (but much less accurately than for those which do). The hysteresis that can occur during respiratory motion can be modelled by a fully continuous model when phase is used as the respiratory parameter. Amplitude, the other popular choice for the respiratory parameter can only model hysteresis when a distinction is made between values occurring during inhalation and those occurring during exhalation. This effectively constructs two separate models, one for inhalation

and one for exhalation, and the overall model is no longer continuous (see Section 2.3.2 and Figure 2.1).

## 5.2 Constructing the motion models

The B-spline registration results are defined by a regular grid of Control Points (CPs), each of which has a 3D displacement associated with it. To construct the motion models a separate 1D function is fitted to the displacement of each CP in each of the three dimensions. This can be illustrated using a 2D plot with the CP displacement on the vertical axis and the phase on the horizontal axis (Figure 5.1). The function is then fitted to this data using least squares (implemented in Matlab, MathWorks, Natick, Massachusetts, USA). Once the function has been fitted it can be used to predict the value of the CP displacement for any desired phase. This can be done for each separate function, thus predicting all of the CP displacements, i.e. the full B-spline non-rigid transformation.

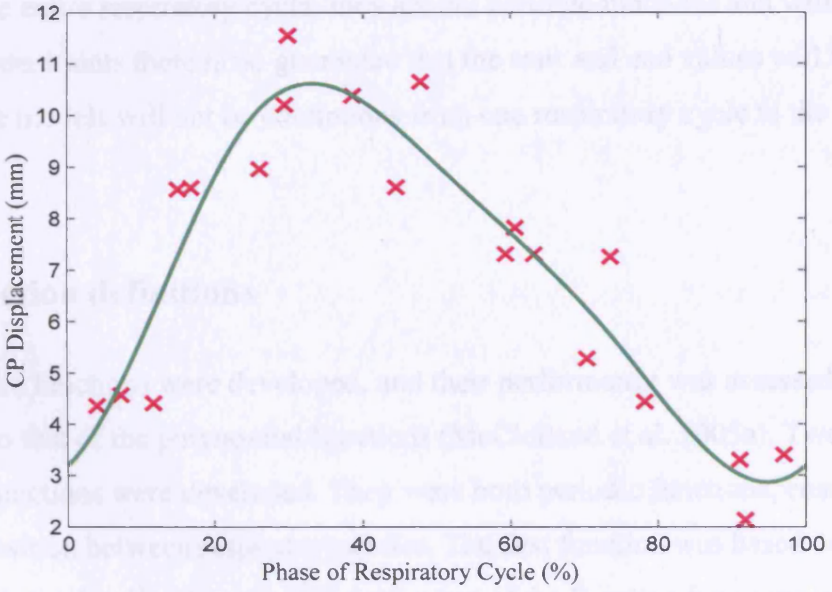


Figure 5.1 - Plot showing a 1D function (green) relating the Control Point (CP) displacement to the phase of the respiratory cycle. The magenta crosses show the values from the registration results for each of the 20 Cine CT volumes from one couch position.

## 5.3 The choice of function for the motion models

In theory any one dimensional function that takes one value as input (the phase) and gives another value as output (the value of the CP displacement) could be used for motion models, but some function will obviously be better suited to modelling the respiratory motion than other functions. An initial feasibility study was performed for the motion models (McClelland et al. 2004). This study visually assessed the results from one patient and concluded that the proposed method was feasible. For this initial study 3$^{rd}$ order polynomial functions were used for constructing the motion models as these functions that had been used in previous work on respiratory motion models in our lab (Blackall et al. 2005). However, in the previous work, due to restrictions with the image acquisition only half of the respiratory cycle, from end-exhalation to end-inhalation (or vice-versa), was modelled, i.e. the model did not allow for hysteresis. A polynomial function was appropriate for the previous models as the position of the anatomy at end-inhalation was different to its position at end-exhalation. The models in this chapter are trying to relate the internal motion to the phase of the respiratory cycle and the position of the anatomy at 100% should be identical to its position at 0%, so a periodic function is required for the models. Although polynomial functions can be used to model the entire respiratory cycle, they are not periodic functions and without additional constraints there is no guarantee that the start and end values will be the same, so the models will not be continuous from one respiratory cycle to the next.

### 5.3.1 Function definitions

Two periodic functions were developed, and their performance was assessed and compared to that of the polynomial functions (McClelland et al. 2005a). Two different candidate functions were developed. They were both periodic functions, ensuring a smooth transition between respiratory cycles. The first function was based on a 1D approximating cubic B-spline (the 1D equivalent of the B-spline functions used for the non-rigid registrations). The function is defined by a number of control points equally spaced over the values of the phase (0%-100%), and an offset defining where the first control point is placed in relation to phase 0%. These control points will be referred to as respiratory control points from here on (not to be confused with the control points (CPs) in the non-rigid registrations). To make the function periodic the respiratory

control points are treated as being cyclic over the phase, i.e. a respiratory control point that is located at 10% can also be thought of as being located at 110% (and -90% if required). This function will be referred to as the cyclic B-spline function from now on.

The value of a cyclic B-spline function, $F_{CBn}(p)$, with $n$ respiratory control points, $\phi_1, \phi_2, ..., \phi_n$, and an offset of $o\%$, at phase $p\%$ is given by:

$$F_{CBn}(p) = \sum_{l=0}^{3} B_l(u)\phi_{i+l(\mathrm{mod}\, n)} \qquad [5.1]$$

where: $\delta = \frac{100\%}{n}$, $i = \left\lfloor \frac{p-o}{\delta} \right\rfloor - 1$, $u = \frac{p-o}{\delta} - \left\lfloor \frac{p-o}{\delta} \right\rfloor$, and $B_l$ represents the $l$-th basis function of the B-spline, given in [2.2]. Figure 5.2 shows an example of the cyclic B-spline function.

The second function that was defined was based on a Fourier series representation of a periodic signal using a limited number of harmonic frequencies. It is defined by the amplitude and phase offset of each frequency component used plus the zero frequency value. As the function is a Fourier series it is by definition periodic. This function defined using a Fourier series will be referred to as the Fourier series function from now on.

The value of the Fourier series function, $F_{FSn}(p)$, defined by $n$ frequencies components having amplitudes $\alpha_1, \alpha_2, ..., \alpha_n$ and phase offsets $\omega_1, \omega_2, ..., \omega_n$, and a zero frequency value of $\alpha_0$ at phase $p\%$ is given by:

$$F_{FSn}(p) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \sin(2\pi i p + \omega_i) \qquad [5.2]$$

Figure 5.3 shows an example of the Fourier series function.

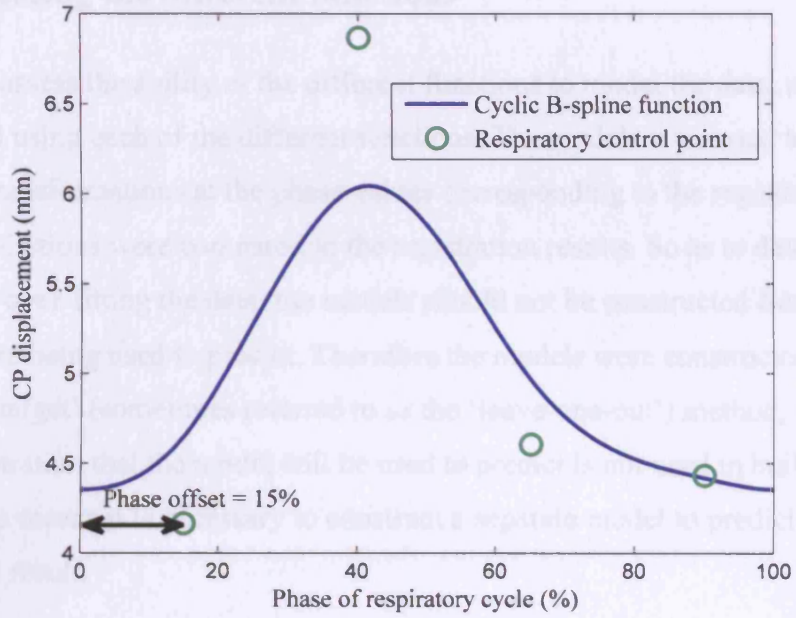


Figure 5.2 - An example of the cyclic B-spline function with four respiratory control points, relating the phase of the respiratory cycle to the registration Control Point (CP) displacement. The respiratory control points and the value of the phase offset are shown.

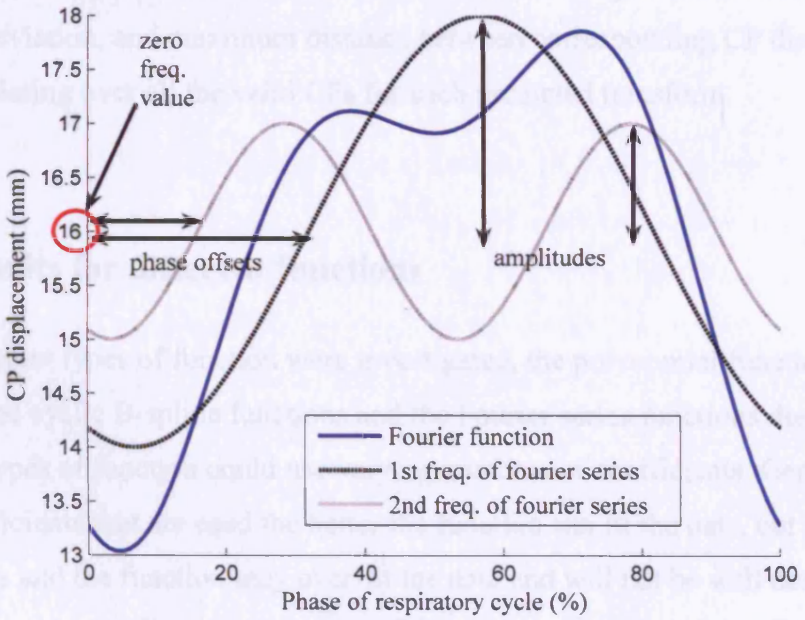


Figure 5.3 - An example of the Fourier function with two frequency components, relating the phase of the respiratory cycle to the registration Control Point (CP) displacement. These individual frequency components are shown as dashed lines, and the zero frequency value and the phase offsets and amplitudes of the two frequency components are also indicated.

### 5.3.2 Assessing the different functions

In order to assess the ability of the different functions to model the data, models were constructed using each of the different functions. The models were used to predict the non-rigid transformations at the phase values corresponding to the registration results, and the predictions were compared to the registration results. So as to detect when the models are over-fitting the data, the models should not be constructed from the same data they are being used to predict. Therefore the models were constructed using the 'leave-out-target' (sometimes referred to as the 'leave-one-out') method, where the target registration that the model will be used to predict is not used in building the model. This means it is necessary to construct a separate model to predict each registration result.

The predicted non-rigid transformations were compared to the original registration results by calculating the 3D Euclidean distance between the CP (Control Point) displacement in the predicted transformations and the corresponding CP displacement in the original registrations. Only CPs that were contained within the segmented lung were used, as the other CPs were not constrained by the registrations. The mean, standard deviation, and maximum distance between corresponding CP displacements were calculating over all the valid CPs for each predicted transform.

### 5.3.3 Results for different functions

Three different types of function were investigated, the polynomial functions originally used and the cyclic B-spline functions and the Fourier series functions described above. All three types of function could use varying numbers of coefficients. Generally, the more coefficients that are used the better the function can fit the data, but too many coefficients and the function may over-fit the data and will not be well defined, causing large errors when predicting unseen data. Therefore several functions of each type were investigated. Table 5.1 lists the different functions that were tested, and gives the number of coefficients that define each function.

The models were built using the results from lung registration protocol 1 (which used the segmented lungs, see section 4.1.1 for details). The models were constructed for all patients who were registered using this protocol (1, 2, 3 and 5).

| Functions tested | Number of coefficients |
|---|---|
| 2nd order polynomial function | 3 |
| 3rd order polynomial function | 4 |
| 4th order polynomial function | 5 |
| 5th order polynomial function | 6 |
| Cyclic B-spline function with 2 respiratory control points | 3 |
| Cyclic B-spline function with 3 respiratory control points | 4 |
| Cyclic B-spline function with 4 respiratory control points | 5 |
| Cyclic B-spline function with 5 respiratory control points | 6 |
| Cyclic B-spline function with 6 respiratory control points | 7 |
| Cyclic B-spline function with 8 respiratory control points | 9 |
| Fourier series function with 1 frequency component | 3 |
| Fourier series function with 2 frequency component | 5 |
| Fourier series function with 3 frequency component | 7 |
| Fourier series function with 4 frequency component | 9 |

Table 5.1 - List of different functions assessed for the single parameter (phase) models, and the number of coefficients that defined each function.

Table 5.2 gives the distance between corresponding CP displacements from the original registration and the transformation predicted by the motion models, for each of the functions assessed. The mean, standard deviation, and maximum value is given for each patient and over all patients. It can be seen that the cyclic B-spline function with 3 respiratory control points performed best overall. However, several of the functions had very similar performance, with the cyclic B-spline functions with 2, 4, and 5 respiratory control points, and the Fourier series functions with 1 and 2 frequency components all having average mean values within 5% of the value for the cyclic B-spline function with 3 respiratory control points.

The two periodic functions that were developed both performed better than the polynomial functions that had been previously used. For each of the function types, the function using the most coefficients performed worse than the other functions, indicating that these function were over-fitting the data. This justifies not investigating functions with more coefficients. Examination of the results for individual registrations shows that there was much more variation between the results for the different

registrations than there was between the results for the majority of the models. Table 5.3 gives the results for two individual registrations (those for the $7^{th}$ and $18^{th}$ cine CT volumes from couch position 2 of patient 1) as an example. This would indicate that sources of 'noise' such as inter-cycle variation, registration errors, and errors in calculating the phase (which could be relatively large for patients 0, 1, 2, and 3 as there was no external respiratory signal recorded for these patients) have more of an impact on the results than the exact function or number of coefficients used.

### 5.3.4 Function chosen for future models

It was decided to adopt the cyclic B-spline function for constructing the motion models. This was because it performed slightly better than the Fourier series function, allows more choice on how many coefficients to use, and is very similar to the B-spline function used for the non-rigid registrations. In addition, if the offset term, $o$, is removed so that the first phase control point is always places at phase 0%, the B-spline function can be calculated very quickly and efficiently for all of the control point displacements in the non-rigid transformation. Some test models were built to predict a small subset of the data used in the previous experiment and it was found that the offset term could be removed with little effect on the accuracy of the models.

Although models using three respiratory control points performed slightly better than those using four respiratory control points it was decided to adopt four control points for future models. This was because it was expected that four respiratory control points should perform better if the some of the sources of noise mentioned above can be reduced. It was hoped that this may be achieved by acquiring more cine CT data at each couch position, by always acquiring an external respiratory surrogate signal for calculating the phase, and by performing more accurate registrations.

| Patient | | Polynomial function (mm) order: | | | | Cyclic B-spline function (mm) number of phase control points: | | | | | | Fourier series function (mm) number of harmonic frequencies: | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2nd | 3rd | 4th | 5th | 2 | 3 | 4 | 5 | 6 | 8 | 1 | 2 | 3 | 4 |
| 1 | Mean | 3.58 | 3.99 | 4.14 | 5.44 | 3.28 | 3.27 | 3.52 | 3.60 | 4.09 | 5.17 | 3.25 | 3.54 | 4.28 | 11.47 |
| | $\sigma$ | 2.19 | 2.45 | 2.64 | 3.43 | 2.03 | 2.08 | 2.22 | 2.26 | 2.49 | 3.37 | 2.04 | 2.23 | 2.63 | 5.85 |
| | Max. | 21.92 | 28.85 | 60.35 | 119.13 | 20.89 | 21.93 | 26.16 | 25.22 | 23.07 | 75.02 | 22.02 | 25.99 | 26.72 | 178.92 |
| 2 | Mean | 3.21 | 2.90 | 2.99 | 2.89 | 2.76 | 2.63 | 2.55 | 2.62 | 2.73 | 2.85 | 2.69 | 2.56 | 2.70 | 2.85 |
| | $\sigma$ | 1.71 | 1.66 | 1.72 | 1.69 | 1.62 | 1.55 | 1.53 | 1.60 | 1.66 | 1.72 | 1.57 | 1.53 | 1.62 | 1.76 |
| | Max. | 18.90 | 21.36 | 20.58 | 22.89 | 21.64 | 21.19 | 17.28 | 16.06 | 18.59 | 17.31 | 21.70 | 17.15 | 19.62 | 18.76 |
| 3 | Mean | 2.82 | 2.78 | 2.80 | 2.96 | 2.68 | 2.65 | 2.75 | 2.81 | 3.03 | 3.28 | 2.65 | 2.76 | 3.05 | 3.45 |
| | $\sigma$ | 1.62 | 1.61 | 1.62 | 1.75 | 1.57 | 1.54 | 1.57 | 1.60 | 1.70 | 1.78 | 1.55 | 1.58 | 1.74 | 1.93 |
| | Max. | 34.07 | 35.14 | 35.23 | 66.51 | 35.06 | 35.05 | 35.10 | 35.09 | 35.11 | 35.46 | 35.02 | 35.14 | 37.42 | 37.66 |
| 5 | Mean | 2.90 | 2.44 | 2.50 | 2.48 | 2.45 | 2.32 | 2.34 | 2.37 | 2.49 | 2.73 | 2.35 | 2.35 | 2.52 | 2.86 |
| | $\sigma$ | 1.61 | 1.54 | 1.60 | 1.59 | 1.56 | 1.51 | 1.49 | 1.50 | 1.56 | 1.74 | 1.51 | 1.50 | 1.60 | 1.82 |
| | Max. | 20.45 | 19.97 | 19.58 | 19.81 | 19.44 | 19.40 | 19.64 | 19.85 | 20.91 | 34.40 | 19.39 | 19.93 | 24.94 | 33.44 |
| All | Mean | 3.06 | 2.92 | 2.99 | 3.23 | 2.74 | 2.66 | 2.72 | 2.78 | 2.99 | 3.34 | 2.69 | 2.73 | 3.03 | 4.47 |
| | $\sigma$ | 1.73 | 1.74 | 1.81 | 1.97 | 1.66 | 1.62 | 1.65 | 1.68 | 1.79 | 2.02 | 1.63 | 1.66 | 1.82 | 2.51 |
| | Max. | 34.07 | 35.14 | 60.35 | 119.13 | 35.06 | 35.05 | 35.10 | 35.09 | 35.11 | 75.02 | 35.02 | 35.14 | 37.42 | 178.92 |

Table 5.2 - The distance between corresponding Control Point (CP) displacements from the original registration and the transformation predicted by the motion models, for each of the functions assessed. The mean, standard deviation, and maximum value is given for each patient and over all patients.

| Cine CT Image: | | Polynomial function (mm) order: | | | | Cyclic B-spline function (mm) number of phase control points: | | | | | | Fourier series function (mm) number of harmonic frequencies: | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $2^{nd}$ | $3^{rd}$ | $4^{th}$ | $5^{th}$ | 2 | 3 | 4 | 5 | 6 | 8 | 1 | 2 | 3 | 4 |
| $7^{th}$ | Mean | 0.98 | 1.01 | 1.53 | 1.45 | 1.40 | 1.16 | 1.13 | 1.15 | 1.36 | 1.73 | 1.27 | 1.19 | 1.28 | 1.75 |
| | $\sigma$ | 0.66 | 0.62 | 0.92 | 0.99 | 0.72 | 0.65 | 0.83 | 0.83 | 0.93 | 1.02 | 0.67 | 0.90 | 0.82 | 0.99 |
| | Max. | 5.15 | 4.42 | 4.79 | 6.53 | 3.78 | 3.86 | 7.68 | 5.94 | 6.87 | 8.26 | 3.82 | 7.68 | 5.95 | 7.14 |
| $18^{th}$ | Mean | 4.11 | 3.95 | 4.25 | 5.65 | 4.15 | 4.29 | 5.24 | 5.10 | 4.94 | 5.09 | 4.14 | 5.24 | 5.21 | 5.50 |
| | $\sigma$ | 2.09 | 2.05 | 2.34 | 3.11 | 2.11 | 2.16 | 2.72 | 2.48 | 2.59 | 2.53 | 2.10 | 2.72 | 2.85 | 3.05 |
| | Max. | 12.67 | 12.92 | 12.93 | 18.25 | 13.67 | 13.44 | 14.00 | 14.09 | 14.12 | 14.50 | 13.25 | 14.00 | 14.15 | 17.18 |

Table 5.3 - The distance between corresponding Control Point (CP) displacements from the original registration and the transformation predicted by the motion models, for each of the functions assessed. The mean, standard deviation, and maximum value are given for the results from two individual Cine CT volumes from patient 1. It can be seen that the differences in the results for the two Cine CT volumes are greater than the differences between the different models.

## 5.4 Evaluating the motion models

Using the CP displacements is the most obvious way to assess the different functions as it is measuring the error in the actual data that the models are directly predicting. However, it should be noted that the magnitude of the error in the CP displacements may not be indicative of the magnitude of the error in the tissue displacements (as the tissue displacements are calculated from several CP displacements). In addition, comparing the model predictions to the registration results does not account for errors in the registrations. Therefore, while assessing the model predictions using the CP displacements is a valid method for comparing the relative performance of different models, it does not assess the ability of the models to predict the location of the anatomy during RT treatment, which is ultimately what the models are trying to achieve.

In order to assess the models ability to predict the location of the actual anatomy the model predictions were evaluated by comparing them to the Cine CT volumes (McClelland et al. 2006a, McClelland et al. 2005b). The models were again constructed using the 'leave-out-target' method, constructing the models from all of the registration results except for the target registration, and then using the model to produce a predicted transformation at the same phase as the target registration. The models were assessed using the expert visual assessment and landmark tracking techniques also used for assessing the registrations used to construct the models.

### 5.4.1 Expert visual assessment

For the expert visual assessment the reference volume was deformed by each of the predicted transformations to produce a predicted CT volume, which was visually compared to the corresponding Cine CT volume. Each prediction was given three scores by a clinical oncologist, one for the tumour, one for the ipsilateral lung, and one for all the other anatomy, according to the following:

1. No visually detectable misalignment over 1.5 mm (one slice thickness).
2. No visually detectable misalignment over 4.5 mm (three times the slice thickness).

3. Visually detectable misalignment over 4.5 mm.

It was specified that at least 90% of the region of interest must meet the criteria above for that region to be given the corresponding score. The intra-user reproducibility of the visual assessment was again estimated by repeating the visual assessment for two patients (using a total of 170 registrations) a few weeks after the initial assessment by the same clinical oncologist.

### 5.4.2 Landmark tracking

A clinical oncologist was asked to locate an easily identifiable anatomical point, a landmark, in each of the Cine CT volumes acquired at a particular couch position, and to identify the same landmark in the reference volume. This was repeated using a different landmark for each couch position, for all couch positions from each patient. The same landmarks were used for assessing the models as were used to assess the registrations (see section 4.2.2.2 for more details on the landmarks used). The clinical oncologist attempted to re-identify the landmark in the Cine CT volumes on a different occasion, in order to estimate the localization error for manually identifying the landmarks. The reference landmark was transformed using an estimate of the inverse (see Appendix A) of the predicted transformations and the registration results. For each Cine CT volume, the Euclidean distance was calculated between:

i. the two locations of the landmark identified on different occasions – an estimate of the manual Target Localisation Error (TLE),

ii. the midpoint of the two landmark locations and the reference point before registration (rigidly aligned to Cine CT space using the spine) – an estimate of the Target Error Before Registration (TEBR),

iii. the midpoint of the two landmark locations and the reference point transformed by the registration result – an estimate of the Target Registration Error (TRE),

iv. the midpoint of the two landmark locations and the reference point transformed by the transformation predicted by the model – an estimate of the Target Model Error (TME),

### 5.4.3 Results of evaluating the motion models

The models were built from the results of registration protocol 2 (which used the complete field of view and not just the segmented lung, see section 4.1.2 for details). The models were built for all couch positions from patients 2-6, giving a total of 18 couch positions modelled. Compared to the time required for the registrations, the motion models could be constructed very quickly, requiring in the order of a 10 seconds per couch position on a 2.8 GHz CPU.

The expert visual assessment results are given in Table 5.4. The results for the registrations are given for comparison. It can be seen that the model prediction results are only slightly worse than the registrations results. The tumour can usually be predicted accurately with over 60% of the predictions scoring 1 (no visually detectable misalignment over 1.5 mm) and only 2.8% scoring 3 (visually detectable misalignment over 4.5 mm). The ipsilateral lung exhibited more misalignment than the tumour. Most of the results still exhibited reasonable alignment, with over 40% of the model predictions scoring a 1, and less than 12% scoring a 3. The results for the other tissue exhibited the most misalignment, but this was usually less than 4.5 mm with only 18.5% of the model predictions receiving a score of 3, and only occurred in small regions that would generally have a negligible effect on any radiotherapy plans and calculations. The repeated results again suggested a high degree of intra-user reproducibility, with 86% of the tumour scores, 68% of the ipsilateral lung scores, and 96.5% of the other tissue scores not differing from the originals and none of the scores differing by 2.

| Score | Registrations (%) | | | Model predictions (%) | | |
|---|---|---|---|---|---|---|
| | T | IL | OT | T | IL | OT |
| 1 | 64.6 | 41.3 | 2.4 | 61.9 | 40.1 | 1.2 |
| 2 | 33.6 | 49.8 | 80.9 | 35.3 | 48.0 | 80.2 |
| 3 | 1.7 | 8.8 | 16.7 | 2.8 | 11.9 | 18.5 |

Table 5.4 - The expert visual assessment results for the registrations and the model predictions. The results are given as the percentages of the registrations/models to be awarded each score: 1, 2, and 3. A score of 1 indicates there was no visually detectable misalignment over 1.5 mm, a score of 2 indicates no misalignment over 4.5 mm, and a score of 3 indicates no misalignment over 4.5 mm. Separate scores were given for the Tumour (T), the Ipsilateral Lung (IL), and all of the Other Tissue (OT).

| Patient | TEBR (mm) | | | TLE (mm) | | | TRE (mm) | | | TME (mm) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. |
| 2 | 4.95 | 2.19 | 11.25 | 0.72 | 0.48 | 2.08 | 1.65 | 0.92 | 6.24 | 2.07 | 1.03 | 5.51 |
| 3 | 5.48 | 2.70 | 11.65 | 0.76 | 0.58 | 2.18 | 1.14 | 0.53 | 3.28 | 1.87 | 0.91 | 4.53 |
| 4 | 4.49 | 2.49 | 10.58 | 0.47 | 0.55 | 2.23 | 1.07 | 0.64 | 3.32 | 1.28 | 0.66 | 3.20 |
| 5 | 5.08 | 2.44 | 10.51 | 1.35 | 0.54 | 2.76 | 1.39 | 0.68 | 2.99 | 1.46 | 0.67 | 3.79 |
| 6 | 4.58 | 1.88 | 10.00 | 0.92 | 0.67 | 2.73 | 1.17 | 0.45 | 2.17 | 1.27 | 0.50 | 3.11 |
| All | 4.87 | 2.35 | 11.65 | 0.85 | 0.64 | 2.76 | 1.26 | 0.67 | 6.24 | 1.56 | 0.81 | 5.51 |

Table 5.5 - The landmark tracking results. The mean, standard deviation ($\sigma$), and maximum of the Target Modelling Errors (TME) are given for each patient and over all patients (2-6) that the models were assessed on. The values for the Target Error Before Registration (TEBR), Target Localisation Error (TLE), and Target Registration Error (TRE) are also given for comparison.

The landmark tracking results are given in Table 5.5. The results for the TEBR, TLE, and TRE were also presented in chapter 4 but have been repeated here for comparison to the model prediction results. The TME results show that the motion models perform very well at predicting the landmark locations. The mean TME for all patients was 1.56 mm, only just over the slice thickness, and the maximum TME was 5.51 mm, less than four slice thicknesses. The mean TME was larger than the mean TRE for every patient as would be expected, but the differences were small, especially for patients 5 and 6 (the patients where the Polaris system had been used to determine the phase).

## 5.5 Combining the motion model predictions from different couch positions

As mentioned in the introduction to this chapter a separate model is built for each couch position. To produce a prediction of the combined volume over all of the couch positions the predictions from each individual model need to be combined.

The simplest way to do this is to use each model to predict the deformation field over the corresponding couch position, and to concatenate the results to produce a deformation field over the all the couch positions. However, this will result in discontinuities at the boundaries between the different couch positions. This can cause artefacts in the concatenated volumes, although these are noticeably smaller than the artefacts in corresponding 4DCT volumes (Figure 5.4).

Although simply concatenating the model predictions reduces the artefacts in comparison to 4DCT, small artefacts can still be present, and more importantly the deformation fields from the different models will not be continuous and may violate a one-to-one mapping. This means that the inverse transformation cannot be estimated in this region (Appendix A), as is required for some RT dose and tracking calculations. To overcome these problems a method has been developed to combine the individual transformations from each model into one continuous transformation which can produce 'artefact free' combined volumes.

Section 5.5.1 assesses the magnitude of the discontinuity in the deformation fields when simply concatenating the predictions from each model. Section 5.5.2 describes how to

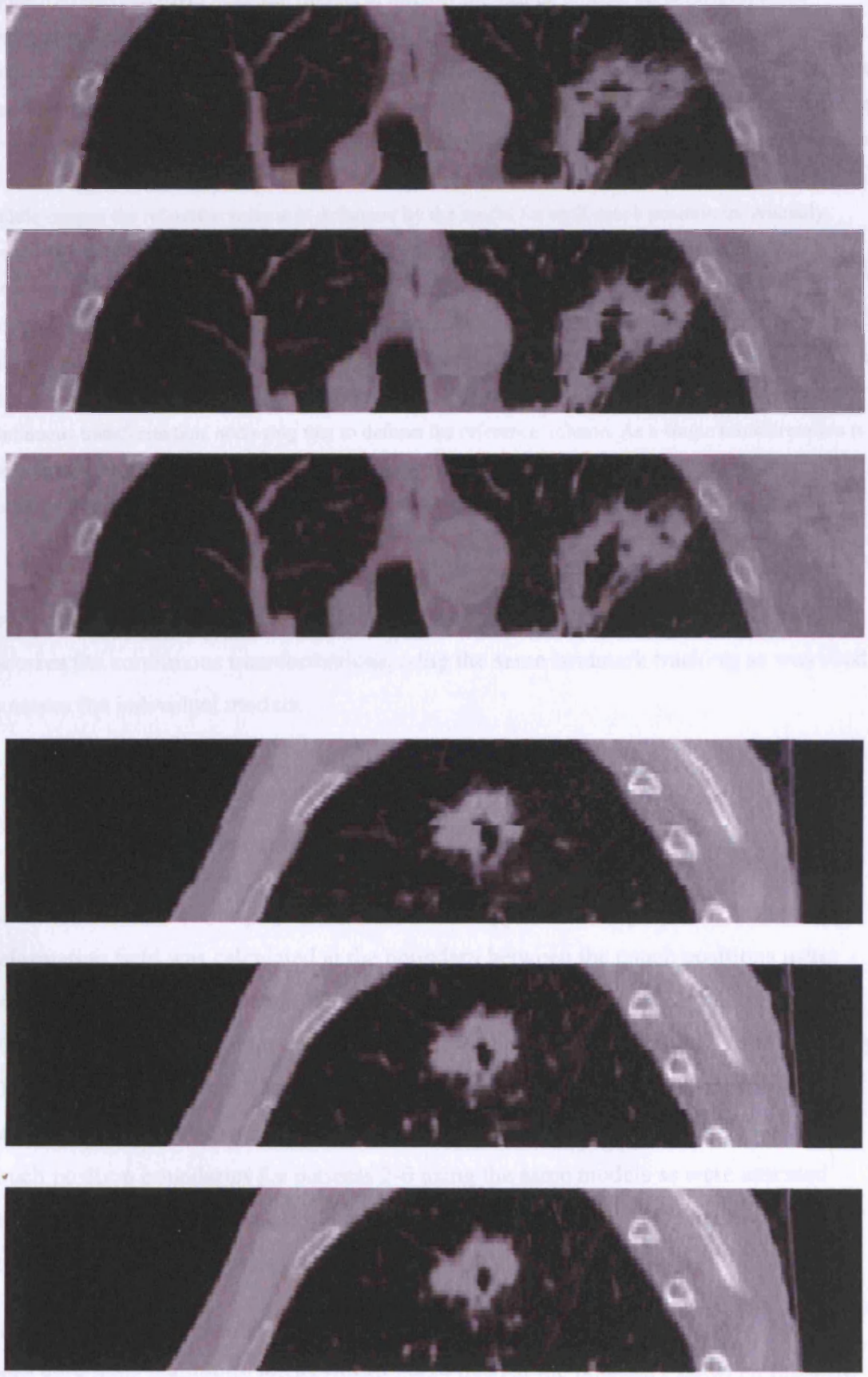

Figure 5.4 (previous page) - Coronal (top set of images) and sagittal (bottom set of images) slices produced by sorting the Cine CT volumes into 4DCT volumes (top image in each set), concatenating the predicted volumes from the individual couch position models (middle image), and combining the motion model results into a single continuous transformation (bottom image). The 4DCT (top) images are formed by concatenating the Cine CT volumes from each couch position that are closest to the desired phase, and so contain many 'discontinuity' artefacts at the boundaries of adjacent couch positions. To form the middle images the reference volume is deformed by the model for each couch position individually, producing a separate predicted volume for each couch position. These are then concatenated together to form a predicted volume over all the couch positions. The artefacts are reduced in these volumes as the models average out some of the inter-cycle variation and do not contain 'binning errors', but the images still contain some noticeable artefacts as a separate transformation is used to predict each couch position . The bottom images are produced by combining the transformations from each model into a single continuous transformation, and using that to deform the reference volume. As a single transformation is used the predicted volumes contain no discontinuity artefacts at the boundaries between adjacent couch positions. These images are from patient 5 for a phase value of 50% (near end inhale).

produce a continuous transformation across all the couch positions, and section 5.5.3 assesses the continuous transformations using the same landmark tracking as was used to assess the individual models.

### 5.5.1 Assessing the discontinuity

To assess the discontinuity between models for adjacent couch positions the deformation field was calculated at the boundary between the couch positions using both models (McClelland et al. 2006a). The deformation field was calculated at 100 x 100 points evenly distributed over the majority of the anatomy visible at each boundary. The continuity error was calculated as the Euclidean distance between the two deformation vectors at each point. The continuity error was assessed at each of the couch position boundaries for patients 2-6 using the same models as were assessed above. For each patient the continuity error was assessed at 10 phase values equally spaced over the respiratory cycle.

The continuity error results are summarized in Table 5.6. It should be noted that there were very large continuity errors (mean 20.24 mm) at the boundary between the third and fourth couch position from patient 3. However, examination of the predicted volumes and the original Cine CT volumes from these couch positions revealed that the

patient had shifted between the acquisitions of these couch positions. This was the cause of the large discontinuities and therefore the results from this boundary were omitted. The mean continuity error over the five patients was 2.15 mm, and the maximum error was 8.76 mm. This shows that even though the artefacts at the couch position boundaries appear greatly reduced in the concatenated volumes when compared to 4DCT volumes, the discontinuities that are present are larger than the modelling errors for the individual couch positions.

| Patient | Continuity Error | |
|---|---|---|
| | Mean | Max. |
| 2 | 2.35 | 8.11 |
| 3 | 2.62 | 8.76 |
| 4 | 1.79 | 7.12 |
| 5 | 2.52 | 6.96 |
| 6 | 1.45 | 6.44 |
| All | 2.15 | 8.76 |

Table 5.6 - The continuity error is a measure of the discontinuity in the deformation field at the boundary between adjacent couch positions when the individual model predictions are simply concatenated together. The mean and maximum continuity errors are given for patients 2-6 and over all five patients.

### 5.5.2 Producing a continuous transformation

The first step towards producing a continuous transform is to perform the registrations with an extended Control Point Grid (CPG), as described in section 4.1.3 and Figure 4.1. When an extended CPG is used the registration result defines a transformation over all of the couch positions, but as data is only present at one couch position the transformation is only constrained in the corresponding region. The contribution that each CP makes to the transformation in this region will depend on the location of the CP relative to the couch position being registered. CPs that are far away will make no contribution to the transformation and will not be optimised as part of the registration. CPs that are outside the couch position but are close (within two CPs) will make some contribution to the transformation and will be optimised during the registration, but may not be very well constrained. CPs that are located within the couch position will make a large contribution to the transformation.

Most CPs will contribute to the registrations at multiple couch positions, but by varying amounts. A measure of the contribution that a CP makes to a particular couch position can be found by calculating the value of the B-spline weight ( $B_n(w)$ in equation [4.1]) for the CP at each row in that couch position, and summing them. The contribution will be the same for all CPs on the same row in the CPG.

If a separate model is built for the each couch position as before, but using the results with the extended CPG, then each model will produce a transformation with the same CPG. The results from the different modes can be combined into a single transformation using the contributions calculated above as weighting factors. For each CP, the combined CP displacement is the sum over all the models of the CP displacement weighted by the contribution that the CP made to the registrations used in that model.

It should be noted that there is no guarantee when combining the transformations in this way that folding or other extreme and unrealistic deformations will not be produced. However, in such cases there would have been extreme discontinuities between adjacent couch positions, indicating that some other problem has occurred during data acquisition (e.g. patient shifts or coughs) or during the data processing (e.g. large errors in the registrations).

### 5.5.3 Assessing the continuous transformations

To assess the results of combining the model predictions into a single continuous transformation the extended registrations from registration protocol 3 were used to construct the motion models. The landmark tracking validation (section 5.4.2) was then repeated twice, once for the results from the individual couch position models ($TME_{individual}$) and once using the results combined into a single transformation as described above ($TME_{combined}$). Note, the individual couch position models will give slightly different results to those in section 5.4.3 and Table 5.5 as the models were constructed from different registration results. As before, the models were constructed leaving out the target Cine CT volume. The models for the couch positions that did not contain the target Cine CT volume (used for the combined transformations) were constructed from all the volumes at those couch positions.

| Patient | $TRE_{extended}$ (mm) | | | $TME_{individual}$ (mm) | | | $TME_{combined}$ (mm) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. | Mean | $\sigma$ | Max. |
| 2 | 1.76 | 0.97 | 6.29 | 2.12 | 1.01 | 5.49 | 2.35 | 1.16 | 6.44 |
| 3 | 1.33 | 0.53 | 2.34 | 1.93 | 0.98 | 4.73 | 2.20 | 1.32 | 7.38 |
| 4 | 1.12 | 0.53 | 3.23 | 1.32 | 0.57 | 3.47 | 1.29 | 0.54 | 2.98 |
| 5 | 1.05 | 0.56 | 2.84 | 1.21 | 0.63 | 2.89 | 1.60 | 0.61 | 3.51 |
| 6 | 1.22 | 0.50 | 2.75 | 1.31 | 0.51 | 3.22 | 1.43 | 0.54 | 3.13 |
| All | 1.27 | 0.66 | 6.29 | 1.52 | 0.81 | 5.49 | 1.71 | 0.93 | 7.38 |

Table 5.7 - The landmark tracking results for the extended CPG registrations ($TRE_{extended}$), the individual models built from the extended registrations ($TME_{individual}$), and the motion models combined into a single transformation ($TME_{combined}$). The mean, standard deviation ($\sigma$), and maximum value are given for patients 2-6 and over all five patients.

Table 5.7 gives the landmark tracking results for the extended registrations ($TRE_{extended}$), the individual models ($TME_{individual}$) and the models combined into a continuous transformation ($TME_{combined}$). As can be seen the extended registration results and the individual model results are very similar to those achieved with the previous registrations (see Table 5.5), indicating that using an extended CPG has no adverse effects on the accuracy. The combined transformation results are marginally worse than the individual model results, but this would be expected given that the average respiratory motion will vary slightly between couch positions. The increase in the landmark tracking error for the combined models is relatively small in comparison to the discontinuity errors (Table 5.6), which by definition are 0 for the combined transformations. Although the combined transformations are slightly worse at predicting the data used to construct the models (as assessed by the landmark tracking), they produce predictions that are more physically plausible than concatenating the individual model predictions (Figure 5.4), and so should actually be better for predicting future data. The CD accompanying this thesis contains animations showing coronal and sagittal slices over a complete respiratory cycle from patients 2-8 for the 4DCT volumes, the individual model predictions concatenated together, and the prediction from the model results combined into a single continuous transformation.

## 5.6 Conclusions

This chapter has introduced and described how to construct single parameter respiratory motion models from B-spline non-rigid registration results. The motion models use one dimensional functions to relate the CP displacements that define the B-spline transformations to the phase of the respiratory cycle.

Two types of function were developed that could be used for the motion models, the cyclic B-spline function and the Fourier series function. Both functions were periodic and could use varying numbers of coefficients. The performance of these two functions was compared to that of polynomial functions, which had previously been used for other motion models, but were not periodic and so not expected to perform so well for the phase based models. To determine how many coefficients should be used, several functions of each different type, with varying number of coefficients were used. In total, four polynomial functions, six cyclic B-spline functions, and four Fourier series functions were assessed. Overall, the cyclic B-spline function with three respiratory control points had the best results, although four respiratory control points were adopted for future models, as their performance was very close to that of the three respiratory control points and they were expected to perform better on future data.

After establishing which function should be used for the motion models, the accuracy of the motion models was evaluated by comparing the model predictions to the original cine CT data. These motion models were constructed from the results of the all anatomy registration protocol 2, and so modelled all of the anatomy, as this was required for Radiotherapy dose calculations. Two different methods were used to evaluate the motion models: expert visual assessment and landmark tracking. The expert visual assessment evaluated the registration over the whole anatomy but only gave a subjective measure of how accurate the models were. The landmark tracking results gave an objective estimate of the accuracy of the models in mm (the Target Modelling Error, TME), but were only evaluated at the landmarks.

The results suggest that the models can predict the Cine CT volumes very accurately, with a mean TME of just over one slice thickness (1.56mm) and a maximum TME less than four slice thicknesses (5.51mm). The TME error was only marginally larger than the Target Registration Error (TRE). It is expected that the TME would be larger than

the TRE as the models average out the inter-cycle variation present in the registration results.

Finally, this chapter presented and assessed a method of combining the model predictions from different couch positions into a single continuous transformation. The combined transformation produced a plausible prediction of the anatomy with no discontinuities between adjacent couch positions, with only a small decrease in the accuracy for predicting the individual Cine CT volumes (mean error 1.71 mm). It should be stressed that even though the combined predictions are less accurate at predicting the CT data acquired for constructing the models, they should in theory be better at predicting future data (i.e. during treatment) as their predictions are physically plausible (unlike the individual model predictions or 4DCT volumes, which can have discontinuities between adjacent couch positions).

# Chapter 6

# Two parameter motion models

This chapter explains how to construct motion models that relate the internal motion to two respiratory parameters. It also describes several example functions that can be used in the two parameter models, and evaluates their performance.

## 6.1 Introduction

The models presented in the previous chapter related the internal respiratory motion to a single respiratory parameter, the phase of the respiratory cycle. This means they cannot model any inter-cycle variation that may be present in the data, and the best they can

achieve is to model the motion for an average respiratory cycle. The models presented in this chapter relate the internal motion to two external respiratory parameters, with the intention that this will allow some of the inter-cycle variation to be included in the models.

## 6.2 Constructing the motion models

The two parameter models are constructed in an identical manner to the single parameter models. A separate 2D function is fitted using least squares for each Control Point (CP) displacement (in each direction) in the registration. This function relates the value of the CP displacement to the values of the two respiratory parameters. This can be illustrated using a 3D plot with the value of the CP displacement on the vertical axis and the values of each surrogate respiratory parameter along the two horizontal axes (Figure 6.1c-f and Animations 6.1c-f on the accompanying CD). The function then defines a surface in this plot (Figures 6.3 - 6.4 and Animations 6.3 - 6.4 on the accompanying CD). As with the single parameter models, once the function has been fitted it can be used to predict the value of the CP displacement for any pairs of values of the respiratory parameters. When this is repeated for all the CP displacements the entire transformation can be predicted. As with the single parameter models, it is still necessary to construct a separate model for each couch position, but the predicted transformations from each model can be combined into a single transformation using the method described in section 5.5.

## 6.3 The choice of function for the motion models

There are a great many different functions that could be used for the 2 parameter models. In theory any function that takes two input values and gives one output value could be used, although there are a number of factors that will influence the appropriateness of the function.

One factor that needs to be considered when choosing the function is the respiratory parameters that are being used. Three different respiratory parameters were acquired that could be used for the two parameters models: the phase of the respiratory cycle

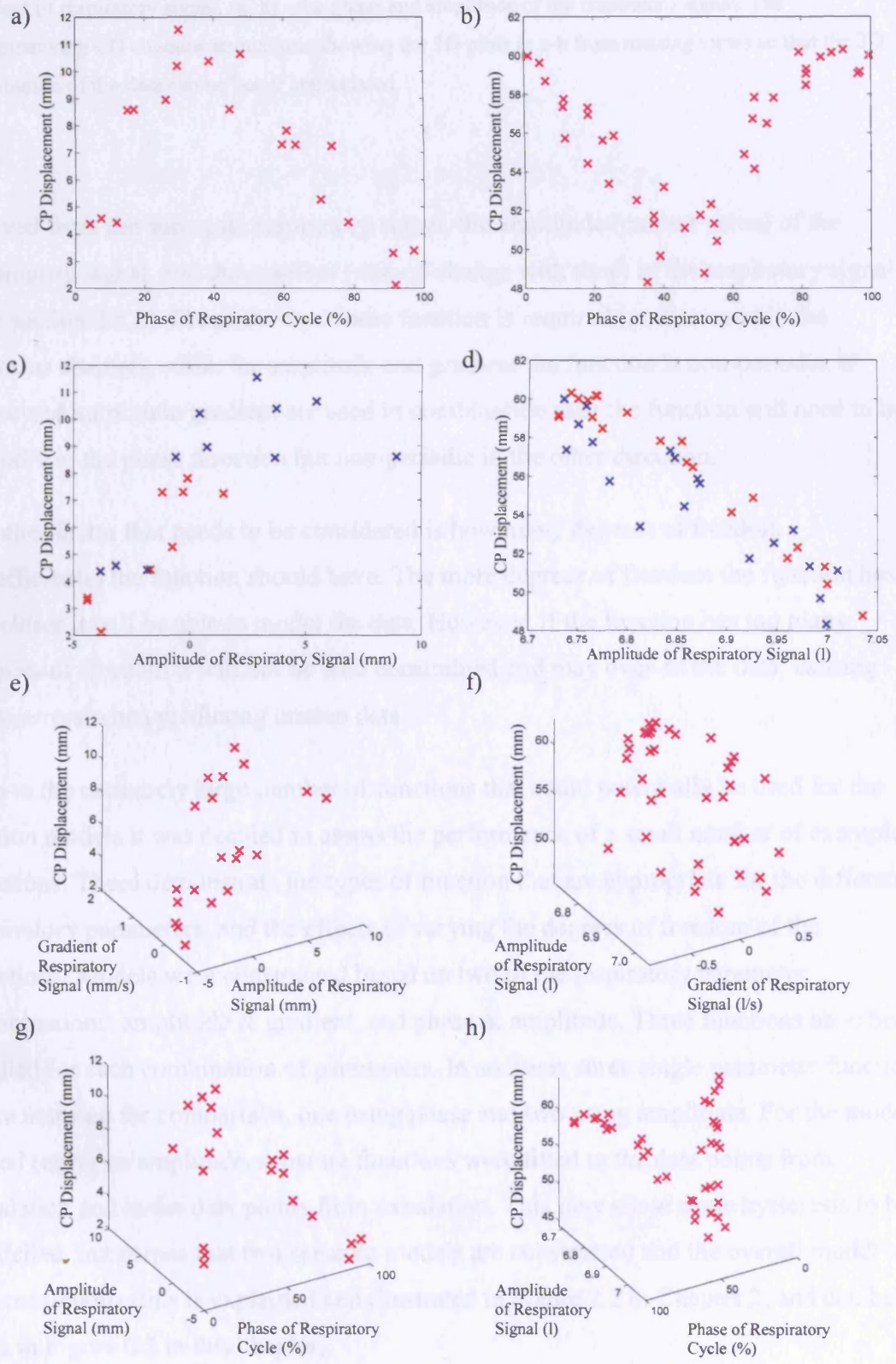


Figure 6.1 - plots for an example Control Point (CP) from the registration results for patients 5 (a, c, e, g) and 8 (b, d, f, h). All plots for each patient show the displacement of the same CP (in one direction) plotted against: (a, b) - the phase of respiratory cycle, (c, d) - amplitude of respiratory signal (CP displacements from inhalation are blue and those from exhalation are red), (e, f) - the amplitude and

gradient of respiratory signal, (g, h) - the phase and amplitude of the respiratory signal. The accompanying CD contains animations showing the 3D plots in e-h from rotating views so that the 3D distribution of the data can be better appreciated.

derived from the surrogate respiratory signal, the amplitude (current value) of the respiratory signal, and the gradient (rate-of-change with time) of the respiratory signal (see section 3.5.2). For phase a periodic function is required (as discussed in the previous chapter), while for amplitude and gradient the function is non-periodic. If phase and amplitude/gradient are used in combination then the function will need to be periodic in the phase direction but non-periodic in the other direction.

Another factor that needs to be considered is how many degrees of freedom (coefficients) the function should have. The more degrees of freedom the function has the closer it will be able to model the data. However, if the function has too many degrees of freedom it will not be well constrained and may over-fit the data, causing large errors when predicting unseen data.

Due to the extremely large number of functions that could potentially be used for the motion models it was decided to assess the performance of a small number of example functions. These demonstrate the types of function that are appropriate for the different respiratory parameters, and the effects of varying the degrees of freedom of the functions. Models were constructed based on two of the respiratory parameter combinations: amplitude & gradient, and phase & amplitude. Three functions have been studied for each combination of parameters. In addition, three single parameter function were assessed for comparison, one using phase and two using amplitude. For the models based (only) on amplitude, separate functions were fitted to the data points from inhalation and to the data points from exhalation. This may allow some hysteresis to be modelled, but means that two separate models are constructed and the overall model is not continuous (this is explained and illustrated in Figure 2.2 in Chapter 2, and can be seen in Figure 6.2 in this chapter).

The purpose of studying these different functions was to establish a framework for constructing and assessing motion models based on one or two respiratory parameters, and to gain an understanding as to how different functions may perform. The next section describes each of the example functions that have been studied.

### 6.3.1 Function definitions

#### 6.3.1.1 Single parameter functions

The function used for the single parameter models based on phase was the cyclic B-spline function with four control points that was adopted and assessed in the previous chapter (Figure 6.2a, b). The equation for the general cyclic B-spline with $n$ control points is given in [5.1]. When $n = 4$ (and $o = 0$) this becomes:

$$F_{CB4}(p) = \sum_{l=0}^{3} B_l(u)\phi_{i+l(\mathrm{mod}\,4)} \qquad [6.1]$$

where $p$ is the phase (between 0-100 %), $\phi_0,...,\phi_3$ are the 4 control points, $i = \left\lfloor \frac{p}{25\%} \right\rfloor - 1$, $u = \frac{p}{25\%} - \left\lfloor \frac{p}{25\%} \right\rfloor$, and $B_l$ represents the $l$-th B-spline basis function, given in [2.2].

The first function used for the amplitude based single parameter models was a 3rd order polynomial (cubic) function (Figure 6.2c, d):

$$F_{P3}(a) = c_3a^3 + c_2a^2 + c_1a + c_0 \qquad [6.2]$$

where $a$ is the amplitude of the respiratory signal, and $c_n$ is the $n^{th}$ coefficient of the polynomial function.

The second function used for the amplitude based single parameter models was a 1D (non-cyclic) B-spline with four control points (Figure 6.2e, f). This is the minimum number of control points required to define a B-spline function and the function is only defined between the central two control points. Therefore, these control points are located at the maximum and minimum values of the surrogate signal observed (over all

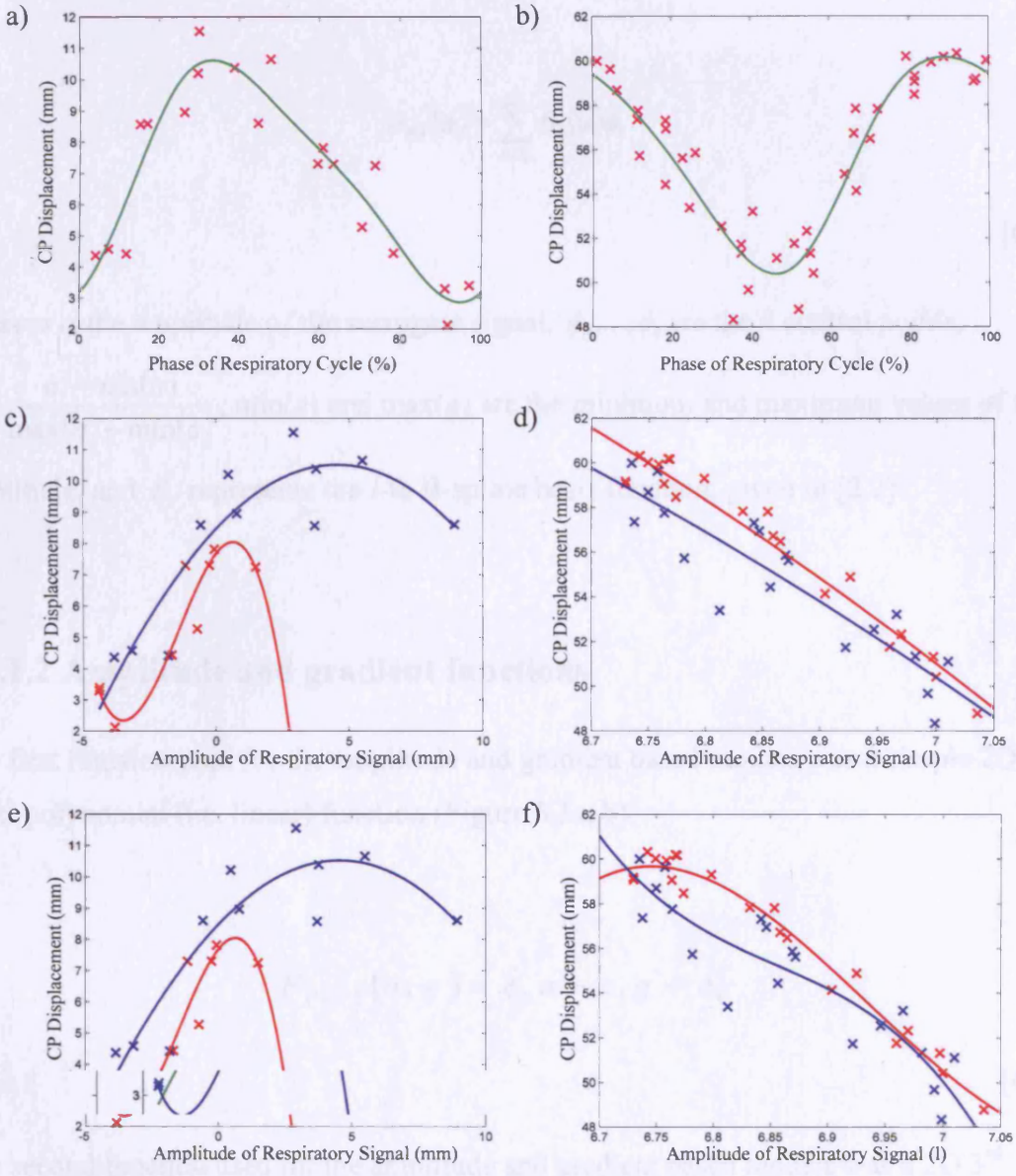


Figure 6.2 - Examples of the one parameter functions for Patient 5 (a, c, e) and Patient 8 (b, d, f). (a, b) - Cyclic B-spline function ($F_{CB4}(p)$) relating the CP displacement to the phase of the respiratory cycle. (c, d) - 3rd order polynomial functions ($F_{P3}(a)$) relating the CP displacement to the amplitude of the respiratory signal. (e, f) - B-spline function ($F_{B4}(a)$) relating the CP displacement to the amplitude of the respiratory signal. A separate function is fit to the values occurring during inhalation (blue) and during exhalation (red).

couch positions), and it is not possible to use the B-spline model to predict a value outside of this range.

$$F_{B4}(a) = \sum_{l=0}^{3} B_l(u)\phi_l \qquad [6.3]$$

where $a$ is the amplitude of the surrogate signal, $\phi_0,...,\phi_3$ are the 4 control points, $u = \frac{a - \min(a)}{\max(a) - \min(a)}$, $\min(a)$ and $\max(a)$ are the minimum and maximum values of the amplitude, and $B_l$ represents the $l$-th B-spline basis function, given in [2.2].

### 6.3.1.2 Amplitude and gradient functions

The first function used for the amplitude and gradient based models was a simple 2D 1st order polynomial (i.e. linear) function (Figure 6.3a, b):

$$F_{P1,P1}(a, g) = c_2 a + c_1 g + c_0 \qquad [6.4]$$

The second function used for the amplitude and gradient based models was a 2D 3rd order polynomial function with no cross terms (Figure 6.3c, d):

$$F_{P3,P3}(a,g) = c_6 a^3 + c_5 a^2 + c_4 a + c_3 g^3 + c_2 g^2 + c_1 g + c_0 \qquad [6.5]$$

where $a$ is the amplitude of the respiratory signal, $g$ is the gradient of the respiratory signal, and $c_n$ is the $n^{th}$ coefficient of the polynomial function in both equations [6.4] and [6.5].

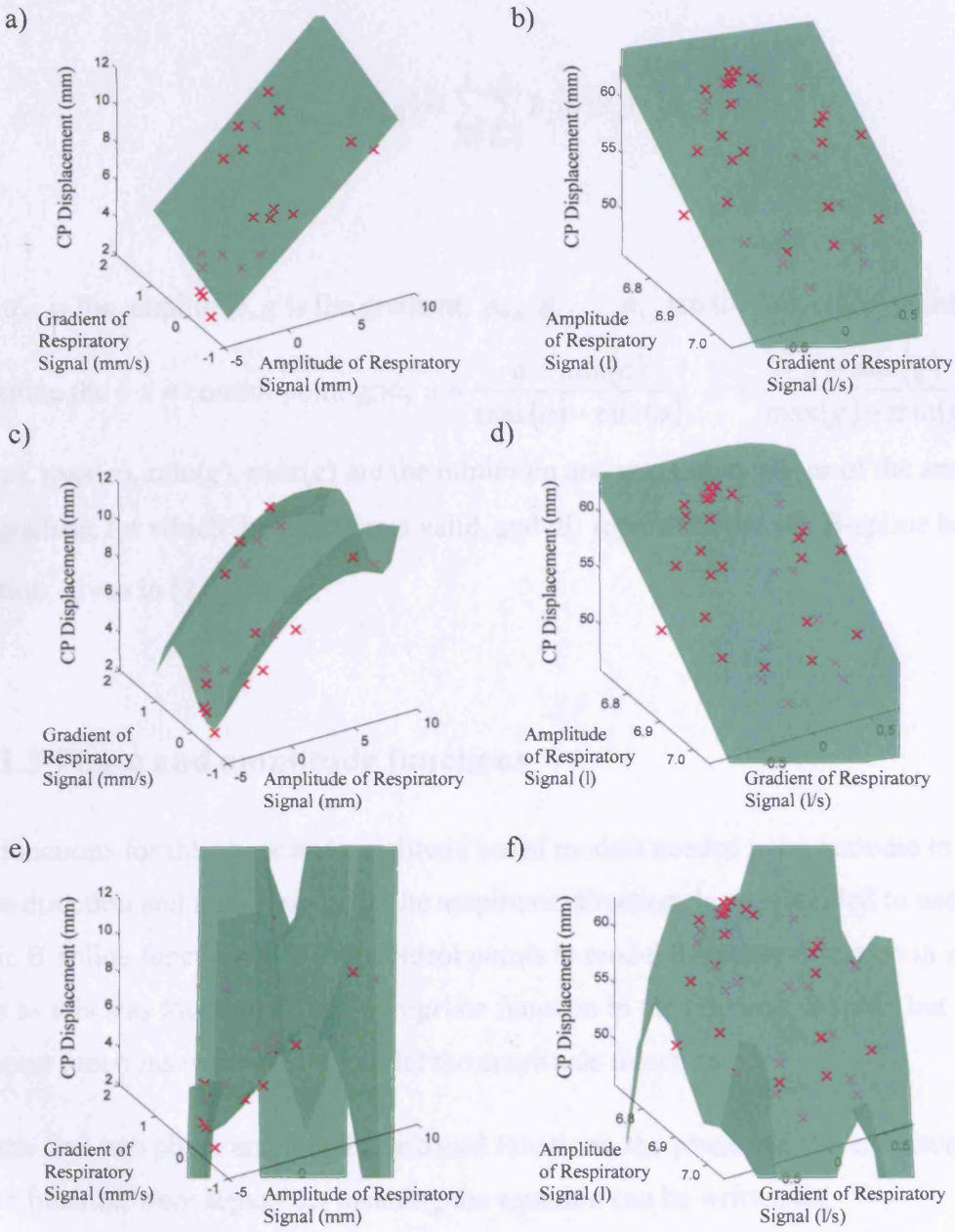


Figure 6.3 - Examples of the two parameter functions relating the CP displacement to the amplitude and gradient of the respiratory signal for Patient 5 (a, c, e) and Patient 8 (b, d, f). (a, b) - 2D 1st order polynomial function ($F_{P1P1}(a,g)$). (c, d) - 2D 3rd order polynomial function ($F_{P3P3}(a,g)$). (e, f) - 2D B-spline function ($F_{B4B4}(a,g)$). The accompanying CD contains animations showing these 3D plots from rotating views so that the 3D distribution of the data and the fitted function can be better appreciated.

The third function used for the amplitude and gradient based models was a 2D B-spline function with four control points in each direction (Figure 6.3e, f):

$$F_{B4,B4}(a,g)=\sum_{l=0}^{3}\sum_{m=0}^{3}B_l(u)B_m(v)\phi_{l,m} \qquad [6.6]$$

where $a$ is the amplitude, $g$ is the gradient, $\phi_{0,0},\phi_{0,1},...,\phi_{3,3}$ are the 16 control points that constitute the 4 x 4 control point grid, $u=\frac{a-\min(a)}{\max(a)-\min(a)}$, $v=\frac{g-\min(g)}{\max(g)-\min(g)}$, min($a$), max($a$), min($g$), max($g$) are the minimum and maximum values of the amplitude and gradient for which the function is valid, and $B_l$ represents the $l$-th B-spline basis function, given in [2.2].

### 6.3.1.3 Phase and amplitude functions

The functions for the phase and amplitude based models needed to be periodic in the phase direction and non-periodic in the amplitude direction. It was decided to use the cyclic B-spline function with four control points to model the phase direction in all cases as this was found to be an appropriate function in the previous chapter, but three different functions were used to model the amplitude direction.

For the first two phase and amplitude based functions, the phase and the amplitude parts of the function were separable, meaning the equation can be written as:

$$F(p,a)=F(p)+F(a) \qquad [6.7]$$

where $F(p)$ is $F_{CB4}(p)$ from equation [6.1]. Note, the first two amplitude and gradient functions were also separable as no cross terms were used.

The first function used to model the amplitude direction for the phase and amplitude based models was a 1st order polynomial function (Figure 6.4a, b):

$$F_{CB4,P1}(p,a) = F_{CB4}(p) + F_{P1}(a)$$

$$F_{P1}(a) = c_1 a + c_0$$

[6.8]

where $p$ is the phase of the respiratory signal, $a$ is the amplitude of the respiratory signal, and $c_n$ is the n$^{th}$ coefficient of the polynomial function.

The second function used to model the amplitude direction was a 3rd order polynomial function (Figure 6.4c, d):

$$F_{CB4,P3}(p,a) = F_{CB4}(p) + F_{P3}(a)$$

[6.9]

where $F_{P3}(a)$ is given in [6.2].

The third function used for the phase and amplitude based models was a 2D 'semi-cyclic' B-spline, which was made cyclic in the phase direction but remained non-cyclic in the amplitude direction (Figure 6.4e, f). Four control points were used in each direction.

$$F_{CB4,B4}(p,a) = \sum_{l=0}^{3}\sum_{m=0}^{3} B_l(u) B_m(v) \phi_{i+l(\mathrm{mod}\,4),m}$$

[6.10]

where $p$ is the phase of the respiratory signal, $a$ is the amplitude of the respiratory signal, $\phi_{0,0}, \phi_{0,1}, \ldots, \phi_{3,3}$ are the 16 control points that constitute the 4 x 4 control point grid, $i = \left\lfloor \frac{p}{25\%} \right\rfloor - 1$, $u = \frac{p}{25\%} - \left\lfloor \frac{p}{25\%} \right\rfloor$, $v = \frac{a - \min(a)}{\max(a) - \min(a)}$, and $B_l$ represents the $l$-th B-spline basis function, given in [2.2].

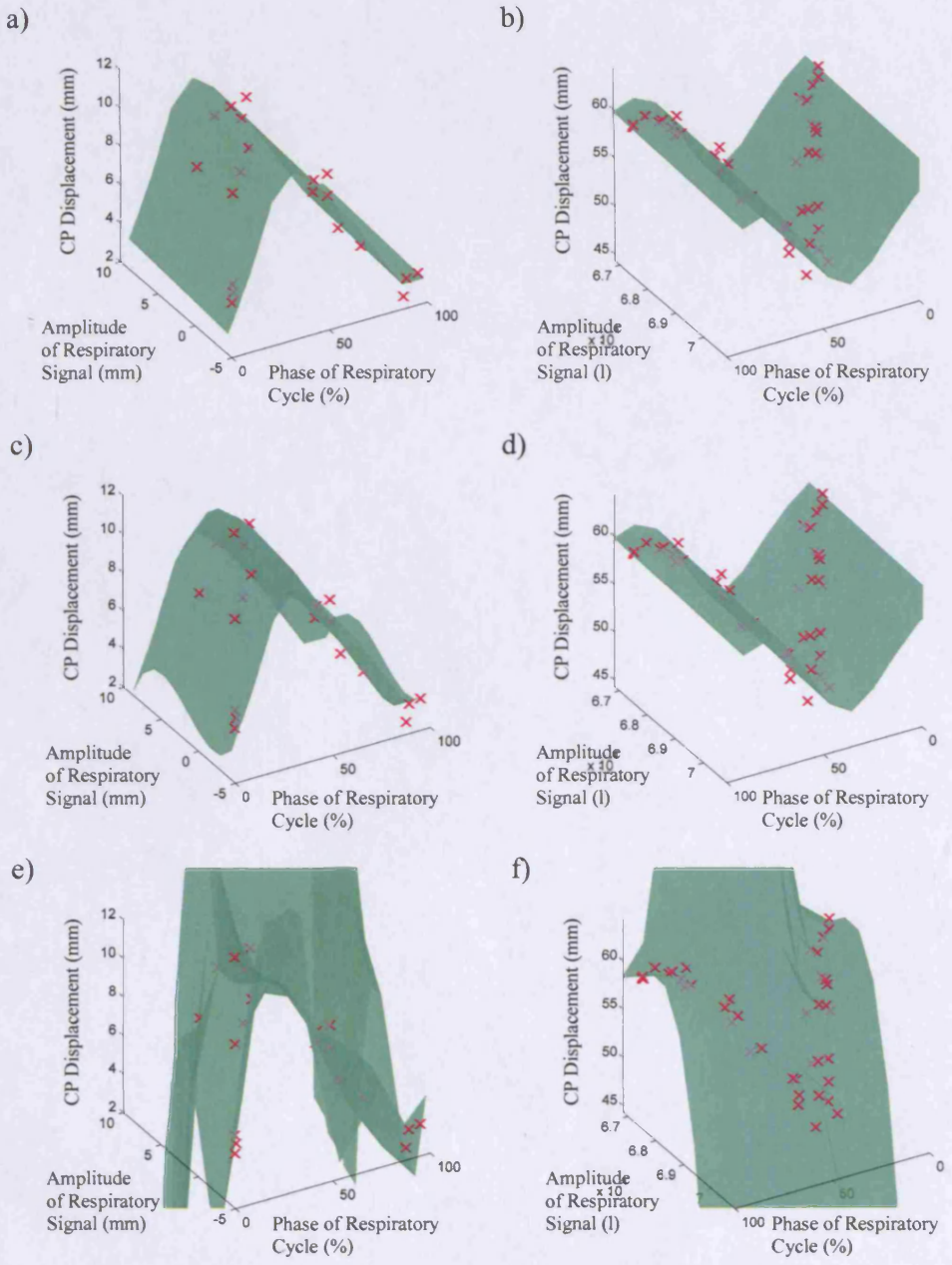


Figure 6.4 - Examples of the two parameter functions relating the CP displacement to the phase and amplitude of the respiratory signal for Patient 5 (a, c, e) and Patient 8 (b, d, f). (a, b) - 2D function ($F_{CB4P1}(p,a)$) with cyclic B-spline ($F_{CB4}(p)$) and 1st order polynomial ($F_{P1}(a)$) components. (c, d) - 2D function ($F_{CB4P3}(p,a)$) with cyclic B-spline ($F_{CB4}(p)$) and 3rd order polynomial ($F_{P3}(a)$) components. (e, f) - 2D 'semi-cyclic' B-spline function ($F_{CB4B4}(p,a)$). The accompanying CD contains animations showing these 3D plots from rotating views so that the 3D distribution of the data and the fitted function can be better appreciated.

## 6.4 Evaluating the different functions

To evaluate the ability of the different parameters and functions to model the data, models were constructed using each of the nine different functions described in the previous section. As before, the models were constructed using the 'leave-out-target' method, to assess the models ability to predict unseen data and to detect over-fitting.

The models were constructed using all but one of the registration results. The models were then used to predict the transformation with the parameter values corresponding to each of the (left out) registration results. The predicted transformations were assessed against the Cine CT data by calculating the Target Modelling Error (TME) using the landmark tracking method described in 5.4.2.

The predicted transforms were also assessed against the original registration results by comparing the deformation fields generated. The deformation fields were sampled at approximately 600 points, spaced 20 mm apart and covering all of the anatomy. Sample points outside of the patient were excluded. At each sample point the displacement vector was calculated from each transformation, and the 3D Euclidean distance between the two displacement vectors, the Deformation Field Error (DFE), was calculated. The DFE assesses the model over the whole region being modelled and not just at specific points (as for the landmark tracking error). However, it is based on the assumption that the registrations are themselves correct, and errors in the registrations cannot be measured or accounted for by the DFE. In the previous chapter the different (single parameter, phase based) models were assessed by comparing the B-spline control point displacements (section 5.3.2). The DFE is more informative than assessing the control point displacements directly, as the value of the DFE is indicative of the actual tissue displacements, which is not necessarily true of the control point displacements.

Both the TME and the DFE where calculated using the predictions from the individual couch position models, $TME_{individual}$ and $DFE_{individual}$, and the combined predictions from all couch positions using the method described in section 5.5, $TME_{combined}$ and $DFE_{combined}$.

## 6.5 Results of evaluating the different functions

The models were constructed using the results of registration protocol 3 (see section 4.1.3 for details), which uses an extended Control Point Grid so that the model results from the individual couch positions can be combined into a single transformation (see section 5.5 for details). The models were only constructed for those patients from whom an external respiratory surrogate signals had been acquired, as an external signal was required to calculate consistent amplitude and gradient parameters across different couch positions (see section 3.5.2.2 for details). Models were constructed for all couch positions from patients 5-8, for a total of 20 couch positions modelled. The models were assessed against a total of 690 Cine CT volumes (20-40 per couch position). As the leave-out-target method was used a separate model was constructed for each Cine CT volume, giving a total of 690 x 9 (number of functions assessed) = 6210 models constructed. The computation time required to fit each model to the registration results was in the order of 10 seconds (implemented in Matlab (Mathworks) on a 2.4 GHz. Intel CPU). This is very fast in comparison to the computation time required for the non-rigid registrations.

The $TME_{individual}$ and $DFE_{individual}$ results are given in tables 6.1 and 6.2. Both sets of results support the same conclusions regarding the comparative performance of the different functions used for the models. Overall, the differences in the mean results for the different models are negligible, except for the two functions with the most degrees of freedom ($F_{CB4,B4}(p,a)$ and $F_{B4,B4}(a,g)$). The differences in the maximum and standard deviation are more noticeable for some functions.

The phase based models using the cyclic B-spline function, $F_{CB4}(p)$, evaluated in the previous chapter, again performed very well for all patients. The amplitude based models using the 3rd order polynomial function, $F_{P3}(a)$, performed very well for patients 7 and 8, slightly outperforming the $F_{CB4}(p)$ models by some measures, and performed reasonably well for patient 6, although the results were noticeably worse than for the $F_{CB4}(p)$ models. The results for the $F_{P3}(a)$ models were not so good for patient 5. Only 20 Cine CT volumes had been acquired per couch position for this patient (30 were acquired for patient 6 and 40 for patients 7 and 8). As separate models were constructed for inhalation and exhalation it was possible that very few volumes were used as training data for some of the models. This would have meant that the functions were not very well constrained by the data, and may have given large errors when predicting the

| Patient | | $TME_{individual}$ (mm) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Single parameter models | | | Amplitude and Gradient based models | | | Phase and Amplitude based models | | |
| | | $F_{CB4}(p)$ | $F_{P3}(a)$ | $F_{B4}(a)$ | $F_{P1,\,P1}$ $(a,g)$ | $F_{P3,\,P3}$ $(a,g)$ | $F_{B4,\,B4}$ $(a,g)$ | $F_{CB4,\,P1}$ $(p,a)$ | $F_{CB4,\,P3}$ $(p,a)$ | $F_{CB4,\,B4}$ $(p,a)$ |
| 5 | Mean | 1.21 | 1.50 | 1.50 | 1.37 | 1.69 | 2.76 | 1.22 | 1.17 | 2.94 |
| | $\sigma$ | 0.63 | 1.10 | 1.10 | 0.72 | 3.20 | 3.44 | 0.65 | 0.63 | 4.87 |
| | Max. | 2.89 | 6.16 | 6.16 | 4.16 | 29.13 | 18.74 | 3.03 | 2.88 | 34.88 |
| 6 | Mean | 1.55 | 1.56 | 1.56 | 1.67 | 1.60 | 2.00 | 1.55 | 1.54 | 1.66 |
| | $\sigma$ | 0.56 | 0.61 | 0.61 | 0.62 | 0.62 | 1.62 | 0.56 | 0.58 | 0.84 |
| | Max. | 3.22 | 3.82 | 3.82 | 3.77 | 3.11 | 13.77 | 3.23 | 2.95 | 6.36 |
| 7 | Mean | 1.23 | 1.17 | 1.21 | 1.17 | 1.20 | 1.49 | 1.25 | 1.25 | 1.26 |
| | $\sigma$ | 0.62 | 0.54 | 0.65 | 0.54 | 0.58 | 1.65 | 0.65 | 0.65 | 0.80 |
| | Max. | 3.44 | 3.95 | 6.37 | 3.76 | 3.66 | 16.35 | 3.68 | 3.68 | 7.83 |
| 8 | Mean | 1.25 | 1.17 | 1.24 | 1.17 | 1.20 | 2.20 | 1.24 | 1.24 | 1.24 |
| | $\sigma$ | 0.77 | 0.72 | 1.04 | 0.75 | 0.80 | 12.01 | 0.78 | 0.78 | 0.88 |
| | Max. | 4.59 | 3.90 | 10.52 | 3.71 | 4.04 | 199.15 | 4.55 | 4.55 | 7.06 |
| All | Mean | 1.28 | 1.26 | 1.30 | 1.26 | 1.31 | 1.99 | 1.28 | 1.27 | 1.50 |
| | $\sigma$ | 0.69 | 0.72 | 0.89 | 0.68 | 1.28 | 7.85 | 0.70 | 0.70 | 1.90 |
| | Max. | 4.59 | 6.16 | 10.52 | 4.16 | 29.13 | 199.15 | 4.55 | 4.55 | 34.88 |

Table 6.1 - The mean, standard deviation, and maximum Target Modelling Error for the individual model predictions, $TME_{individual}$, for each patient and over all patients. The results are given for each of the 9 functions assessed, and are grouped according to the respiratory parameters used.

| Patient | | $DFE_{individual}$ (mm) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Single parameter models | | | Amplitude and Gradient based models | | | Phase and Amplitude based models | | |
| | | $F_{CB4}(p)$ | $F_{P3}(a)$ | $F_{B4}(a)$ | $F_{P1,\,P1}$ $(a,g)$ | $F_{P3,\,P3}$ $(a,g)$ | $F_{B4,\,B4}$ $(a,g)$ | $F_{CB4,\,P1}$ $(p,a)$ | $F_{CB4,\,P3}$ $(p,a)$ | $F_{CB4,\,B4}$ $(p,a)$ |
| 5 | Mean | 0.87 | 1.30 | 1.30 | 0.86 | 1.16 | 15.43 | 0.86 | 0.93 | 4.12 |
| | $\sigma$ | 0.79 | 2.59 | 2.59 | 0.77 | 2.05 | 74.46 | 0.77 | 0.96 | 15.19 |
| | Max. | 9.37 | 73.65 | 73.65 | 10.02 | 43.46 | 1923.95 | 9.96 | 20.96 | 466.77 |
| 6 | Mean | 0.62 | 0.71 | 0.71 | 0.64 | 0.69 | 1.39 | 0.62 | 0.63 | 0.90 |
| | $\sigma$ | 0.55 | 0.63 | 0.63 | 0.56 | 0.61 | 3.23 | 0.54 | 0.55 | 0.88 |
| | Max. | 8.56 | 10.22 | 10.22 | 8.59 | 8.52 | 111.53 | 8.51 | 8.31 | 21.04 |
| 7 | Mean | 0.57 | 0.57 | 0.66 | 0.59 | 0.65 | 1.06 | 0.57 | 0.57 | 0.78 |
| | $\sigma$ | 0.53 | 0.55 | 0.92 | 0.56 | 0.60 | 2.18 | 0.52 | 0.52 | 1.38 |
| | Max. | 11.75 | 11.33 | 35.36 | 12.07 | 11.66 | 87.16 | 11.83 | 11.83 | 65.76 |
| 8 | Mean | 0.68 | 0.64 | 0.73 | 0.64 | 0.67 | 1.35 | 0.67 | 0.67 | 0.80 |
| | $\sigma$ | 0.66 | 0.61 | 0.91 | 0.61 | 0.65 | 6.96 | 0.63 | 0.63 | 1.01 |
| | Max. | 10.86 | 11.41 | 33.73 | 11.96 | 14.09 | 442.76 | 10.87 | 10.86 | 40.74 |
| All | Mean | 0.66 | 0.70 | 0.77 | 0.65 | 0.72 | 2.89 | 0.65 | 0.66 | 1.19 |
| | $\sigma$ | 0.63 | 1.06 | 1.23 | 0.62 | 0.93 | 26.21 | 0.61 | 0.65 | 5.39 |
| | Max. | 11.75 | 73.65 | 73.65 | 12.07 | 43.46 | 1923.95 | 11.83 | 20.96 | 466.77 |

Table 6.2 - The mean, standard deviation, and maximum Deformation Field Error for the individual model predictions, $DFE_{individual}$, for each patient and over all patients. The results are given for each of the 9 functions assessed, and are grouped according to the respiratory parameters used.

transformations for amplitude values not observed in the training data (Figure 6.2c, e), i.e. the function was over-fitting the data. The amplitude based models using the B-spline function, $F_{B4}(a)$, had very similar results to the $F_{P3}(a)$ models for patients 5 and 6, but were noticeably worse than the $F_{P3}(a)$ models for patients 7 and 8. The results for the $F_{B4}(a)$ models were worse than for the $F_{CB4}(p)$ models for all patients. This indicates that the $F_{B4}(a)$ models had started to over-fit the data in some circumstances.

The overall performance of the amplitude and gradient based models using the 2D 1st order polynomial function, $F_{P1,P1}(a,g)$, was very similar to that of the $F_{CB4}(p)$ models.

There were marginal differences in $TME_{individual}$ results for the $F_{P1,P1}(a,g)$ and $F_{CB4}(p)$ models, with the $F_{P1,P1}(a,g)$ performing worse for patients 5 and 6 but slightly better for patients 7 and 8. The differences in the mean and standard deviation of the $DFE_{individual}$ results between the $F_{P1,P1}(a,g)$ models and the $F_{CB4}(p)$ models were negligible, but the maximum $DFE_{individual}$ was larger for the $F_{P1,P1}(a,g)$ models for all patients. The amplitude and gradient based models using the 2D 3rd order polynomial function, $F_{P3,P3}(a,g)$, had similar, but generally slightly worse results to the $F_{P1,P1}(a,g)$ models for patients 6-8. The $F_{P3,P3}(a,g)$ models were considerably worse for patient 5, indicating that for this patient (for whom only 20 Cine CT volumes had been acquired per couch position) the function was over-fitting the data. The amplitude and gradient based models using the 2D B-spline function, $F_{B4,B4}(a,g)$, clearly over-fit the data for all patients, resulting in very large maximum errors (both $TME_{individual}$ and $DFE_{individual}$), and relatively large mean errors in comparison to the other functions used for the models.

The results for the phase and amplitude based models using the cyclic B-spline and 1st order polynomial function, $F_{CB4,P1}(p,a)$, were very similar to the results of the $F_{CB4}(p)$ models. The results for the phase and amplitude based models using the cyclic B-spline and 3rd order polynomial function, $F_{CB4,P3}(p,a)$, were also very similar to the results of the $F_{CB4}(p)$ models for patient 6-8. The $TME_{individual}$ results for the $F_{CB4,P3}(p,a)$ models for patient 5 are better than the $F_{CB4}(p)$ models, but the $DFE_{individual}$ results are worse. These results indicate that extending the phase based cyclic B-spline models with a (1st or 3rd order) polynomial function to model the variation in the amplitude parameter has little effect on the models if sufficient data is acquired. The results for the phase and amplitude based models using the 2D semi-cyclic B-spline, $F_{CB4,B4}(p,a)$, are noticeably

worse than the results for the other models, except for the $F_{B4,B4}(a,g)$ models. This indicates that the $F_{CB4,B4}(p,a)$ models are over-fitting the data, but are better constrained than the $F_{B4,B4}(a,g)$ models. It should be noted that the actual values of the mean errors for the $F_{CB4,B4}(p,a)$ and the $F_{B4,B4}(a,g)$ models are still quite small for patients 6-8, even though they are noticeably larger than the mean errors for the other models.

The results for the combined predictions from all models, $TME_{combined}$ and $DFE_{combined}$, are given in tables 6.3 and 6.4. It can be seen that the combined prediction results are generally worse than the individual model results, as would be expected. Although the mean $DFE_{combined}$ results are noticeably larger than the $DFE_{individual}$ result, there is less of an effect on the maximum DFE values, and for some patients and functions the maximum $DFE_{combined}$ is actually smaller than the maximum $DFE_{individual}$. The differences between the $TME_{combined}$ and the $TME_{individual}$ results are negligible for patients 6-8, with some functions actually having better $TME_{combined}$ results for some patients. It should be noted that the process of combining the predictions improves the results for the functions with the worst performance ($F_{CB4,B4}(p,a)$ and $F_{B4,B4}(a,g)$), although the combined results still contain unacceptably large errors.

The mean $TME_{combined}$ over all patients was less than 1.4 mm for all except the worst two models, and the maximum $TME_{combined}$ was less than 5 mm for the best four models (and was less than 5 mm for three of the four patients for another two models). These results indicate that the models that perform well can predict the Cine CT data with a high degree of accuracy.

The mean $DFE_{combined}$ over all patients was less than 1.1 mm for all except the worst two models, and the maximum $DFE_{combined}$ was less than 15 mm for the best three models (and was less than 15 mm for three of the four patients for another three models). The $DFE_{individual}$ results were even better, with a mean value over all patients of less than 0.8 mm for all but the worst two models. As can be seen from these values, the DFE results have a smaller mean value, but a larger maximum value than the TME results. This implies that the models can reproduce most of the deformation field from the non-rigid registration very accurately, but there are some small regions and/or some particular values of the respiratory parameters where the predictions are less accurate.

Figure 6.5 and 6.6 show error-maps of the $DFE_{individual}$ and the $DFE_{combined}$ for the $F_{CB4}(p)$ model and the $F_{P1,P1}(a,g)$ model, with iso-intensity lines from the Cine CT

| Patient | | $TME_{combined}$ (mm) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Single parameter models | | | Amplitude and Gradient based models | | | Phase and Amplitude based models | | |
| | | $F_{CB4}(p)$ | $F_{P3}(a)$ | $F_{B4}(a)$ | $F_{P1,\,P1}$ $(a,g)$ | $F_{P3,\,P3}$ $(a,g)$ | $F_{B4,\,B4}$ $(a,g)$ | $F_{CB4,\,P1}$ $(p,a)$ | $F_{CB4,\,P3}$ $(p,a)$ | $F_{CB4,\,B4}$ $(p,a)$ |
| 5 | Mean | 1.60 | 1.88 | 1.88 | 1.72 | 1.90 | 3.07 | 1.61 | 1.58 | 2.52 |
| | $\sigma$ | 0.61 | 1.19 | 1.19 | 0.70 | 2.03 | 3.91 | 0.60 | 0.64 | 2.96 |
| | Max. | 3.51 | 7.82 | 7.82 | 4.75 | 18.85 | 24.46 | 3.33 | 3.60 | 21.81 |
| 6 | Mean | 1.60 | 1.62 | 1.62 | 1.73 | 1.66 | 1.91 | 1.60 | 1.60 | 1.69 |
| | $\sigma$ | 0.65 | 0.69 | 0.69 | 0.65 | 0.65 | 1.06 | 0.65 | 0.65 | 0.83 |
| | Max. | 3.12 | 3.59 | 3.59 | 3.58 | 3.04 | 6.39 | 3.14 | 3.07 | 5.86 |
| 7 | Mean | 1.18 | 1.14 | 1.17 | 1.13 | 1.17 | 1.32 | 1.19 | 1.19 | 1.19 |
| | $\sigma$ | 0.62 | 0.62 | 0.65 | 0.61 | 0.63 | 1.05 | 0.60 | 0.60 | 0.82 |
| | Max. | 3.55 | 4.27 | 4.93 | 3.93 | 4.01 | 9.42 | 2.81 | 2.79 | 8.86 |
| 8 | Mean | 1.31 | 1.27 | 1.30 | 1.27 | 1.30 | 1.76 | 1.30 | 1.30 | 1.49 |
| | $\sigma$ | 0.67 | 0.67 | 0.76 | 0.70 | 0.75 | 2.88 | 0.68 | 0.68 | 1.59 |
| | Max. | 4.27 | 4.00 | 6.60 | 4.12 | 4.42 | 27.59 | 4.46 | 4.47 | 16.37 |
| All | Mean | 1.34 | 1.34 | 1.36 | 1.33 | 1.37 | 1.77 | 1.34 | 1.33 | 1.53 |
| | $\sigma$ | 0.66 | 0.77 | 0.81 | 0.70 | 0.98 | 2.42 | 0.66 | 0.66 | 1.58 |
| | Max. | 4.27 | 7.82 | 7.82 | 4.75 | 18.85 | 27.59 | 4.46 | 4.47 | 21.81 |

Table 6.3 - The mean, standard deviation, and maximum Target Modelling Error for the combined model predictions, $TME_{combined}$, for each patient and over all patients. The results are given for each of the 9 functions assessed, and are grouped according to the respiratory parameters used.

| Patient | | $DFE_{combined}$ (mm) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Single parameter models | | | Amplitude and Gradient based models | | | Phase and Amplitude based models | | |
| | | $F_{CB4}(p)$ | $F_{P3}(a)$ | $F_{B4}(a)$ | $F_{P1,\,P1}$ $(a,g)$ | $F_{P3,\,P3}$ $(a,g)$ | $F_{B4,\,B4}$ $(a,g)$ | $F_{CB4,\,P1}$ $(p,a)$ | $F_{CB4,\,P3}$ $(p,a)$ | $F_{CB4,\,B4}$ $(p,a)$ |
| 5 | Mean | 1.42 | 1.67 | 1.67 | 1.43 | 1.57 | 9.21 | 1.41 | 1.44 | 3.47 |
| | $\sigma$ | 0.96 | 1.90 | 1.90 | 0.96 | 1.40 | 39.46 | 0.96 | 1.02 | 8.96 |
| | Max. | 10.19 | 67.48 | 67.48 | 10.30 | 28.58 | 1160.40 | 10.22 | 17.49 | 333.51 |
| 6 | Mean | 0.95 | 0.99 | 0.99 | 0.97 | 0.99 | 1.51 | 0.95 | 0.95 | 1.10 |
| | $\sigma$ | 0.69 | 0.71 | 0.71 | 0.70 | 0.71 | 2.63 | 0.69 | 0.69 | 0.86 |
| | Max. | 8.91 | 9.54 | 9.54 | 8.97 | 8.65 | 93.74 | 8.86 | 8.66 | 19.88 |
| 7 | Mean | 0.88 | 0.89 | 0.94 | 0.90 | 0.93 | 1.27 | 0.88 | 0.88 | 1.10 |
| | $\sigma$ | 0.65 | 0.66 | 0.80 | 0.67 | 0.69 | 2.45 | 0.65 | 0.65 | 1.16 |
| | Max. | 12.00 | 11.68 | 25.31 | 12.28 | 12.18 | 136.95 | 12.05 | 12.05 | 43.72 |
| 8 | Mean | 1.05 | 1.03 | 1.07 | 1.03 | 1.04 | 1.62 | 1.04 | 1.04 | 1.30 |
| | $\sigma$ | 0.75 | 0.72 | 0.86 | 0.73 | 0.74 | 4.75 | 0.74 | 0.74 | 2.19 |
| | Max. | 14.24 | 13.66 | 23.57 | 13.91 | 13.88 | 313.01 | 14.26 | 14.26 | 208.31 |
| All | Mean | 1.03 | 1.05 | 1.09 | 1.02 | 1.06 | 2.37 | 1.02 | 1.02 | 1.46 |
| | $\sigma$ | 0.76 | 0.95 | 1.03 | 0.76 | 0.85 | 14.11 | 0.75 | 0.76 | 3.53 |
| | Max. | 14.24 | 67.48 | 67.48 | 13.91 | 28.58 | 1160.40 | 14.26 | 17.49 | 333.51 |

Table 6.4 - The mean, standard deviation, and maximum Deformation Field Error for the combined model predictions, $DFE_{combined}$, for each patient and over all patients. The results are given for each of the 9 functions assessed, and are grouped according to the respiratory parameters used.

volume overlaid to show the outline of the patient and their lungs (and the tumour Figure 6.5). Figure 6.5 shows the error-maps for a slice at the edge of a Cine CT volume from patient 5. Figure 6.6 shows the error-maps for a slice in the middle of the Cine CT volume from patient 8. For both patients, a Cine CT volume close to the tumour with large DFE values is shown (for patient 8 all the Cine CT volumes that contained the tumour had very small DFE values, < 2mm, so a Cine CT volume from the couch position below the tumour was chosen to illustrate where the larger DFE values are occurring). It can be seen in Figure 6.5 that the results for both models are very similar, even though the models use different respiratory parameters and modelling functions. In Figure 6.6 there is a more noticeable difference in the magnitude of the DFE values for the two models. The $F_{P1,P1}(a,g)$ model clearly has lower DFE values than the $F_{CB4}(p)$ model, although the distribution of the DFE values is similar for the two models. Both figures also show that the largest $DFE_{individual}$ values are occurring outside of the lungs. In both Figures 6.5 and 6.6 the largest $DFE_{individual}$ value inside the lung for the $F_{CB4}(p)$ model is approximately 3.5 mm. In Figure 6.5 the largest $DFE_{individual}$ value inside the lung for the $F_{P1,P1}(a,g)$ model is also approximately 3.5 mm, but is down to approximately 2.75 mm in Figure 6.6

There was a relatively large difference between the $DFE_{individual}$ values and the $DFE_{combined}$ values for slice shown in Figure 6.5. The $DFE_{combined}$ values inside the lungs are larger than the $DFE_{individual}$ (all are still under 5 mm), although the values in the tumour are still very small (< 2mm), and the maximum $DFE_{combined}$ values (which occur outside the lung) are actually smaller than the corresponding $DFE_{individual}$ values. The difference between the $DFE_{individual}$ and the $DFE_{combined}$ values was much smaller for the slice shown in Figure 6.6, with the $DFE_{combined}$ values inside the lungs all being within 0.5 mm of the $DFE_{individual}$ values. The main reason that the difference is so much greater in Figure 6.5 is because that slice was located near the edge of the Cine CT volume, whereas the slice in 6.6 was located in the middle of the Cine CT volumes (although variations in the respiratory motion modelled for each couch position will also have affected the results).

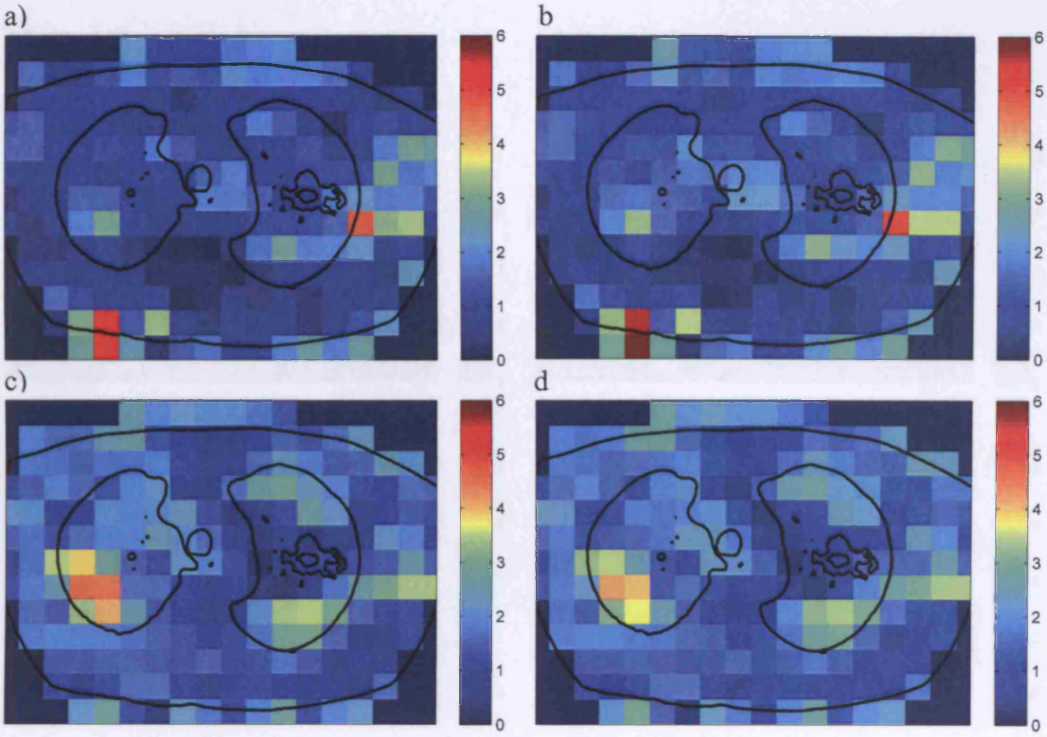


Figure 6.5 - Error-maps showing the Deformation Field Error (DFE) in mm, for a slice from the edge of a Cine CT volume from patient 5. The iso-intensity lines from the Cine CT volume are overlaid to show the outline of the patient, their lungs, and the tumour. a) $DFE_{individual}$ for the $F_{CB4}(p)$ model, b) $DFE_{individual}$ for the $F_{P1,\,P1}(a,g)$ model, c) $DFE_{combined}$ for the $F_{CB4}(p)$ model, and d) $DFE_{combined}$ for the $F_{P1,\,P1}(a,g)$ model. It can be seen that the two different models produce very similar results (a and c compared to b and d), and that combining the model predictions from all the couch positions has a lot of effect on this slice (a and b compared to c and d).

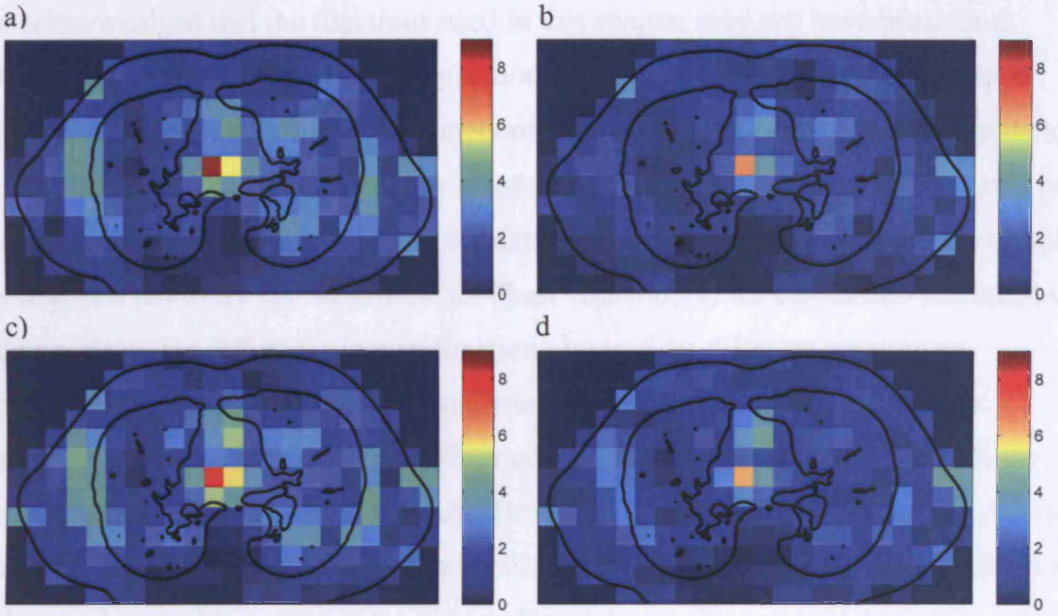


Figure 6.6 - Error-maps showing the Deformation Field Error (DFE) in mm, for a slice from the middle of a Cine CT volume from patient 8. The iso-intensity lines from the Cine CT volume are overlaid to show the outline of the patient, and their lungs. a) $DFE_{individual}$ for the $F_{CB4}(p)$ model, b) $DFE_{individual}$ for the $F_{P1, P1}(a,g)$ model, c) $DFE_{combined}$ for the $F_{CB4}(p)$ model, and d) $DFE_{combined}$ for the $F_{P1, P1}(a,g)$ model. It can be seen that for this slice there is a more noticeable difference between the two different models (a and c compared to b and d), but that combining the model predictions from all the couch positions has little effect on this slice (a and b compared to c and d).

## 6.6 Discussion and conclusions

This chapter has presented a framework for constructing and assessing motion models based on two respiratory parameters. It has also presented a number of example functions that might be used for the motion models, and evaluated their performance on data from four patients.

It is acknowledged that the functions used in this chapter may not have been ideal functions for modelling the data. Simple modifications and/or constraints may have improved the results of some of the functions (e.g. applying smoothness constraints to the B-spline functions), and there may be other functions that were not investigated here that may be more suited to modelling the data. The intention of evaluating the example function was not to try and determine the "best" function(s) for the motion models, but to demonstrate the different types of function required for different respiratory parameters and the effects of using functions with too many degrees of freedom. Despite this, most of the functions performed remarkably well at predicting both the Cine CT data and the registration results. However, it is also acknowledged that to truly evaluate the model's performance they should be assessed using an entirely different set of data to that used to construct the model. This data may be acquired during the same session to test the model's ability to account for short term inter-cycle variation, or during a different session hours or days later to test the model's ability to account for longer term variation. In addition, the models would need to be assessed on many more patients before they could be relied on for clinical use.

The results for the functions with high degrees of freedom, $F_{B4,B4}(a,g)$ and $F_{CB4,B4}(p,a)$, confirm the need for performing the experiments using the leave-out-target method. When the target registration result is also used to construct the models the $F_{B4,B4}(a,g)$ and $F_{CB4,B4}(p,a)$ functions perform better than the other functions. It is only when the target result is left out that it becomes clear that the functions are over-fitting the data.

Although the two parameter models presented in this chapter perform no better than the single parameter models, this does not necessarily mean that the principle of using two respiratory parameters to model the inter-cycle variation is flawed. The single parameter models generally perform very well, and limitations with the data or the pre-modelling processing steps may be preventing the two parameter models from performing any better. There may have been very little inter-cycle variation in the data (although the

respiratory traces indicate that there was considerable variation for some couch positions, see Figure 3.6). There may not be sufficient data from each couch position to fully sample the inter-cycle variation; more samples per respiratory cycle and/or samples from more respiratory cycles may be required. Errors in the registrations and/or the respiratory parameters may be preventing the inter-cycle variation from being measured accurately enough to be modelled. Also, the function used in this chapter for the two parameter models may have been unable to accurately model the relationship between the respiratory parameters and the inter-cycle variations without over-fitting the data.

The two parameter motion models presented here are based on the assumption that there is a strong relationship between the respiratory parameters and the inter-cycle variations in the respiratory motion. It is intuitively thought that the respiratory parameters used in this chapter should have some relationship to some of the variations in the respiratory motion. However, there is no guarantee that this relationship is particularly strong for the parameters/signals used here. Much more work is required to investigate the strength, robustness, and nature of the relationships between the internal respiratory motion (including its variation) and the many different possible respiratory parameters (and the external signals they are derived from), and in particular how stable these relationships are over different time frames. The study and comparison of the relationships to the internal motion for different respiratory signals and parameters was beyond the scope of this thesis, although such studies are essential to the development of accurate models, and are currently underway in our own lab and many others.

The motion modelling method presented in this chapter could be directly applied to respiratory parameters from two different respiratory signals, rather than two parameters from the same signal as used here. They could also easily be extended to three (or more) respiratory parameters, by simply fitting 3D functions that relate the parameters to the B-spline Control Point displacements. However, this would almost certainly require more data to be acquired at each couch position, so that the relationship between the respiratory parameters and the internal motion was sufficiently sampled and to prevent the functions over-fitting the data.

# Chapter 7

# Using the motion models

This chapter describes some of the ways the motion models can be used for lung RT, and details two experiments performed as part of this research that demonstrate the use of the motion models: the comparison of breath-hold and free-breathing motion models, and the delivery of tracked IMRT (Intensity-Modulated Radiation Therapy) treatment.

## 7.1 Introduction

There are many potential uses for the motion models described in the previous two chapters. Section 2.6 describes some of the possible uses for the motion models in applications other than RT treatment of lung cancers, but these were not the focus of this research. Section 2.1.5 describes some of the ways of accounting for respiratory motion when delivering RT to lung tumours and all of these (except for those based on

treating at breath-hold) may benefit from utilising the models. This chapter will give more specific descriptions of how the models presented in this thesis can be used for lung RT. The motion models can be used for 'medical research applications', where the aim is to learn something about the nature of the internal respiratory motion and/or its relationship to external surrogate signals. The idea is that the knowledge gained will be in some way beneficial to lung RT and may well affect clinical practice, but the methodology used is not itself intended for clinical practice. The motion models can also be utilised in methods that are (ultimately) intended to translate to clinical practice, although it is acknowledged that the accuracy and robustness of the motion models (and the rest of the methodology) would first need to be demonstrated to a much higher level before they could be clinically relied on.

## 7.2 Using the motion models for 'medical research applications'

The framework presented in the previous chapters for constructing and evaluating respiratory motion models can be used to try and answer many questions regarding the nature of the internal respiratory motion and its relationship to external surrogate signals. The models can be used to assess the strength and validity of proposed relationships (functions) between external respiratory parameters derived from surrogate signals and the internal respiratory motion and its variation. These relationships can be assessed over the whole region of interest, or at any specific points within the region. Different relationships, respiratory parameters, and surrogate signals can be evaluated and compared to each other, as was demonstrated for a number of example relationships and parameters in chapter 6. If data can be acquired for longer periods of time and/or over multiple sessions then the models can be used for a detailed analysis of the effects of inter-cycle variation on the internal motion and its relationship to the external respiratory parameters. The models can also be used to determine how much data is required to sample the variation over different time frames. The effects of audio and visual feedback and other breath training techniques can be studied with the models. As an example of a 'medical research application', the motion models have been used to compare the motion that occurs between inhalation and exhalation breath-holds and that which occurs during free-breathing (McClelland et al. 2006b).

## 7.3 Comparing models derived from breath-hold and free-breathing data

### 7.3.1 Introduction

Several authors have proposed that motion models can be constructed from CT volumes acquired (exclusively) at inhalation and exhalation breath-holds, and that these models could be used to predict the motion that occurs during free-breathing (Sarrut et al. 2006, Schweikard et al. 2004, Zeng et al. 2007b, see Section 2.5.2). Others in our lab (Blackall et al. 2006) and elsewhere (Rietzel et al. 2005a) have shown that volumes acquired at breath-hold do not always give a good representation of the motion that occurs during free-breathing, and it was decided to see if the motion models supported this finding. As this experiment was performed before the two parameter models had been developed only the phase based cyclic B-spline models were used, although it could easily be modified to use any of the models presented in the previous chapter.

### 7.3.2 Method and materials

Breath-hold volumes were acquired at both tidal inhalation and tidal exhalation for patients 2, 3, 5, and 6. The exhalation scan was non-rigidly registered to the inhalation scan using registration protocol 2 (see Section 4.1.2 for details). The non-rigid registration result was used to determine the motion between the two breath-hold scans. The motion models that were evaluated in Section 5.4 were used to determine the respiratory motion during free-breathing.

A clinician identified the centre of the tumour and each nodal area that was included in the models. A total of 14 points of interest were studied over the four patients (4 tumour points, 10 nodal points). The clinician manually located these points in the exhalation breath-hold scans. The breath-hold registrations were used to identify the location of the points of the interest in the inhalation breath-hold scans. The motion models were used to identify the location of the points of interest during free-breathing at end-exhalation, end-inhalation, and 18 intermediate phases. Note, in order to do this it was necessary to estimate the inverse of the non-rigid transformations predicted by the motion models.

See Appendix A for a description of how to estimate the inverse of a non-rigid transformation.

### 7.3.3 Results

For each point of interest, the motion between the breath-hold scans was compared to the free-breathing motion in a number of ways. The distance between the location of the point in the exhalation breath-hold scan and at end-exhalation during free-breathing was measured (Figure 7.1a). The mean distance over the 14 points was 2.2 mm and the maximum distance was 3.9 mm. The distance between the location of the point in the inhalation breath-hold scan and at end-inhalation during free-breathing was also measured (Figure 7.1a). The mean distance was 4.9 mm and the maximum distance was 7.8 mm. The magnitude of the breath-hold motion, the distance between the locations of the point in the two breath-hold scans, was compared to the magnitude of the free-breathing motion, the distance between the end-exhalation and end-inhalation locations during free breathing (Figure 7.1b). The magnitude of the breath-hold motion was between 0.6 and 3.3 times that of the free-breathing motion, although the mean difference between the free-breathing magnitude and the breath-hold magnitude was only 1.6mm, and the maximum difference was 3.4 mm. The angle between the direction of the breath-hold motion and the primary direction of the free-breathing motion (defined as the direction of a straight line drawn between the end-exhalation and end-inhalation locations) was also measured (Figure 7.1c). This angle was over 45° for half the points and was over 90° in one case. Finally, to assess the extent that the free-breathing motion deviated from a straight line, the distance was measured between the intermediate free-breathing locations and the corresponding points on a straight line drawn between the end-exhalation and end-inhalation locations (Figure 7.1d). For three of the four patients at least one point of interest had a mean distance to the straight line greater than 2 mm, and two patients had at least one point with a maximum distance greater than 5 mm.

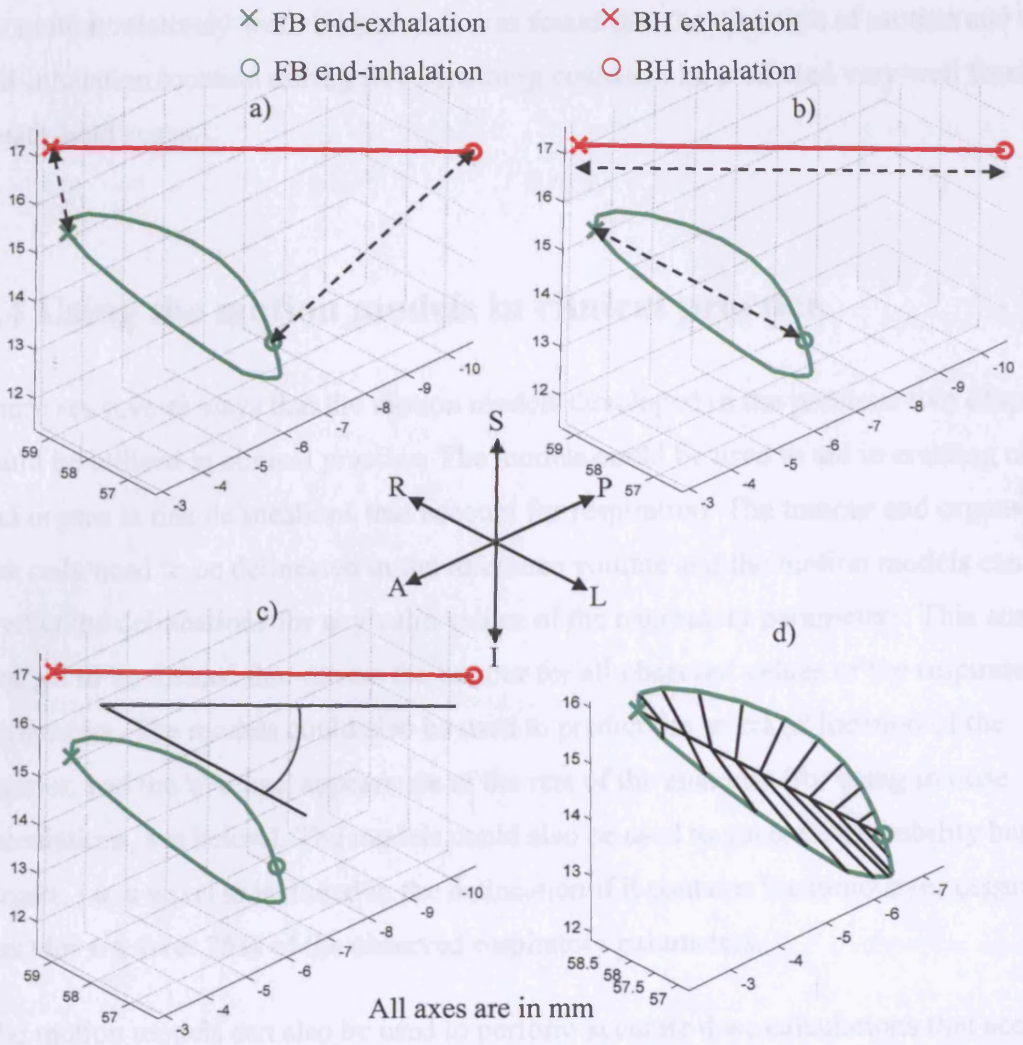


Figure 7.1 - The difference between the Breath-Hold (BH) and Free Breathing (FB) motion was assessed for each point by comparing: a) the location at FB end-exhalation/end-inhalation with the location at BH exhalation inhalation, b) the magnitude of the BH motion with the magnitude of the FB motion, c) the direction of the BH motion with the direction of the FB motion. d) To assess the extent that the free-breathing motion deviated from a straight line, the distance was measured between the intermediate free-breathing locations and the corresponding points on a straight line drawn between the end-exhalation and end-inhalation locations

### 7.3.4 Conclusion

The results above show that the motion between breath-hold volumes and the motion that occurs during free-breathing are not always the same, and therefore breath-hold scans cannot be used to model the motion that occurs during free-breathing. The breath-

hold scans were able to predict the end-exhale locations and the overall magnitude of the motion relatively well. However, it was found that the direction of motion and the end-inhalation location during free-breathing could not be predicted very well from the breath-hold scans.

## 7.4 Using the motion models in clinical practice

There are several ways that the motion models developed in the previous two chapters could be utilised in clinical practice. The models could be used to aid in creating target and organs at risk delineations that account for respiration. The tumour and organs at risk only need to be delineated in the reference volume and the motion models can then predict the delineations for any valid values of the respiratory parameters. This enables a target to be formed that covers the tumour for all observed values of the respiratory parameters. The models could also be used to predict the 'average' location of the tumour, and the 'average' appearance of the rest of the anatomy (for using in dose calculations, see below). The models could also be used to generate probability based targets, i.e. a voxel is included in the delineation if it contains the tumour (or organ at risk) for e.g. over 75% of the observed respiratory parameters.

The motion models can also be used to perform accurate dose calculations that account for the respiratory motion. The dose delivered to the reference anatomy at any values of the respiratory parameters can be calculated. The dose can be accumulated over all the values of the respiratory parameters observed during planning to generate the planned dose distribution. If the respiratory parameters are also measured during treatment the actual dose distribution delivered to the patient can be estimated using the model, and evaluated against the planned dose distribution. The models could also be used to perform probabilistic based dose calculations.

The motion models can be used in the ways described above to aid in the planning of standard 'static' treatments, where the beam is always switched on and motionless during treatment delivery. However, they are potentially even more useful in the planning and delivery of 'active' respiratory treatments, such as gating and tracking (see Section 2.1.5). The motion models relate the internal motion to external respiratory parameters that can be measured during treatment, and hence used to guide the gated or

tracked treatment. For gated treatments, targets and dose distributions can be calculated for various ranges of the respiratory parameters in order to determine the optimum gating parameters. A detailed example of how the motion models can be used to plan tracked IMRT treatment is given in the next section, and the motion models could be used in a similar way for tracked treatment using a robotically mounted LINAC.

Several authors have proposed using the results of non-rigidly registering 4DCT datasets for one or more of the uses of the motion models described above (see Section 2.1.5). However, using the motion model predictions has a number of advantages over using 4DCT registration results. The motion models are potentially better at predicting the motion that occurs during treatment as they are able to average out or model some of the (short term) inter-cycle variation. 4DCT registrations will attempt to recreate any artefacts that occur between couch positions in Cine 4DCT volumes, whereas the motion models (combined using the method in Section 5.5) produce a smooth and continuous deformation that predicts an artefact free volume.

The motion models are continuous over the respiratory parameters and can predict the deformation for any values of the respiratory parameters, whereas 4DCT registrations are discrete and only sample the deformation at particular values of the respiratory parameters. This could be particularly useful for gated and tracked treatments where it may be desirable to estimate the deformation of the anatomy at a higher temporal resolution than can be provided by 4DCT.

Some of the target definition and dose distribution applications described above are currently being investigated by collaborators at the Institute of Cancer Research and Royal Marsden Hospital, using the models developed and evaluated in chapter 5 (the results of these are to be submitted for publications soon). The next section describes how the motion models can be used to plan tracked IMRT treatment. This work was conducted in collaboration with Prof. Steve Webb from the Institute of Cancer Research and Royal Marsden Hospital, and was originally presented as two oral presentations at XVth International Conference on the Use of Computers in Radiation Therapy, ICCR 2007 (McClelland et al. 2007a, Webb et al. 2007), and written up as a paper published in Physics in Medicine and Biology with myself as first author (McClelland et al. 2007b).

## 7.5 Using the models for tracked IMRT treatment

### 7.5.1 Introduction

IMRT, delivered via the Dynamic Multileaf Collimator (DMLC) technique (see Section 2.1.2), can accommodate intra-fraction motion, in particular respiratory motion, if the corresponding motion is applied to the travelling leaves. If this is achieved correctly the target will appear stationary in the "breathing leaves" frame of reference, and a treatment plan made on the static 3D anatomy can be correctly delivered to the breathing anatomy. This was first demonstrated for 1D rigid body translation (Keall et al. 2001), and it has been shown that exact tracking solutions exist for cases of rigid-body translation and deformations in the direction of leaf travel (Papiez et al. 2005), and, with some experimental limitations, for 2D rigid body translations (McQuaid and Webb 2006). However, if non-rigid deformations occur in 2D or 3D, as is generally the case for real patients, then it is no longer possible to produce an exact tracking solution. This is because the anatomical points which are in line with a particular beam element (bixel) at the planning phase of the respiratory cycle will undergo "differential motion" and will no longer lie in a straight line at other phases of the respiratory cycle.

### 7.5.2 The 2D differential motion tracking strategy

A strategy that minimises the geometrical tracking error in a least-squares sense was proposed by Webb and Binnie (2006). Motion along the direction of the beam cannot be accounted for by moving the leaves, and it has been shown that it produces a 2nd order effects in comparison to motion perpendicular to the beam (Webb 2006). Therefore the strategy proposed by Webb and Binnie (2006) only tracks the motion in the beams-eye-view, but samples the motion at many points along the beam, and the motion may be different at each point (differential motion). The original method only tracked motion in the direction of leaf motion, but the method was extended to track full 2D motion in the beams-eye-view in Webb et al. (2007). The 2D differential motion tracking strategy is described by the following equation:

$$[x_L(t_k)]_{q,p} = \frac{\left(\sum_{i_{start}(p)_{CTV}}^{i_{fin}(p)_{CTV}} + \sum_{i_{start}(p)_{OAR}}^{i_{fin}(p)_{OAR}}\right) I_{q,p,i} x_{B_i}(t_k)_{q,p} \delta_1(q) + \sum_{q'(q'\neq q)} \sum_{p'(t_k)} \sum_{i'} I_{q',p',i'} x_{B_{i'}}(t_k)_{q,q',p'} \delta_2(q,q')}{\left(\sum_{i_{start}(p)_{CTV}}^{i_{fin}(p)_{CTV}} + \sum_{i_{start}(p)_{OAR}}^{i_{fin}(p)_{OAR}}\right) I_{q,p,i} \delta_1(q) + \sum_{q'(q'\neq q)} \sum_{p'(t_k)} \sum_{i'} I_{q',p',i'} \delta_2(q,q')}$$

[7.1]

Equation [7.1] was re-written as equation [7.2] so that it is easier to relate to the computational algorithm used to implement the tracking strategy:

$$r_{k,q}(t) = \frac{\sum_{q'} \sum_{j_{start}(q')}^{j_{fin}(q')} \sum_{i_{start}}^{i_{fin}} I(p_{k,q'}(t), j, i) x(p_{k,q'}(t), j, i, t) \delta(q, y(p_{k,q'}(t), j, i, t))}{\sum_{q'} \sum_{j_{start}(q')}^{j_{fin}(q')} \sum_{i_{start}}^{i_{fin}} I(p_{k,q'}(t), j, i) \delta(q, y(p_{k,q'}(t), j, i, t))}$$

[7.2]

Equation [7.1], and the tracking strategy itself, were devised by Steve Webb, and are explained in full detail in Webb et al. (2007) and McClelland et al. (2007b). The differences between equation [7.1] and [7.2] are explained in Appendix B of McClelland et al. (2007b). Note, there are some minor notational differences between equation [7.2] presented here and how the equation was presented in Appendix B of McClelland et al. (2007b). Equation [7.2], the computational implementation of the tracking strategy, and its subsequent evaluation using one of the patient motion models from Section 5.5 of this thesis, were performed by me as part of my PhD research.

There are $Q$ leaf-pairs in the MLC which independently determine the fluence delivered to each channel $q$=1,2,...,$Q$. Each channel, $q$, is sampled at the points $j=j_{start}(q),\ldots,j_{fin}(q)$, equally spaced ($j_{spacing}$ mm apart) along the leaf-end. For each leaf sample point, $j$, the motion is sampled at points $i=i_{start},\ldots,i_{fin}$, equally spaced ($i_{spacing}$ mm apart) along the beam (parallel beams are assumed) such that all of the anatomy to be tracked is sampled.

$I(p,j,i)$ is the Importance factor (weighting value) given to point $p,j,i$ in the planning volume.

$x(p,j,i,t)$ is the position in the x-direction (the direction of leaf-travel) that the point originating at $p,j,i$ in the planning volume had displaced to at time $t$.

$y(p,j,i,t)$ is the position in the y-direction (across leaf direction) that the point originating at $p,j,i$ in the planning volume had displaced to at time $t$.

$\delta(q,y)$ is a function that is 1 if position $y$ lies within channel $q$, and is 0 otherwise.

$p_{k,q}(t)$ is the planned position of the ($k$=leading or $k$=trailing) leaf-end in channel $q$ at time $t$.

$r_{k,q}(t)$ is the respiratory tracking position of the ($k$=leading or $k$=trailing) leaf-end in channel $q$ at time $t$, that is the where the leaf should be positioned so as to optimally track the respiratory motion, according to this strategy. This is the weighted average of the positions of the sample points that lie within channel $q$ at time $t$.

### 7.5.3 Computational implementation of the motion tracking strategy

To apply the motion tracking strategy it is necessary to know where any point in the planning anatomy has moved to at any time, i.e. it must be possible to calculate $x(p,j,i,t)$ and $y(p,j,i,t)$ in equation [7.2]. This information can be provided by the motion models developed in this thesis. The single parameter models developed in chapter 5 were used for this work. This experiment also assumed that the respiratory cycle had a fixed period and that the treatment started at a known phase. While it is acknowledged that these assumptions are clearly not realistic, this was an initial study into tracking differential motion from real patient data and accounting for inter-cycle variation was not the focus of this work. Also, in this initial implementation, tissue moving in and out of the beam was ignored.

The computational algorithm used to implement the differential motion tracking strategy specified in equation [7.1] and [7.2] is:

For every time point:

For each set of leaves (leading and trailing):

- In the beam's-eye view place leaf sample points along each leaf-end spaced $j_{spacing}$ mm apart,

- Project the leaf sample points into 3D by placing motion sample points along the beam from $i_{start}$ to $i_{fin}$ equally spaced $i_{spacing}$ mm apart, such that all of the anatomy being tracked is sampled,
- Calculate the respiratory parameter(s) corresponding to this time point,
- Using the motion model, find the 3D location of each motion sample point for these respiratory parameter(s) – the tracked sample points,
- Project the tracked 3D sample points back into the beam's-eye view,
- For each leaf find the tracked sample points that are aligned with that leaf:
    - Find the mean of those tracked sample points (weighted by their importance factors) – this is the optimal leaf-end position,
    - Find the standard deviation of these points (weighted by their importance factors) – this is a measure of how well the leaves can track the motion, referred to as the tracking error.

This algorithm is illustrated in Figure 7.2.

Note, when using the motion models to find the displacement of the motion sample points from the planning volume to the current respiratory state, the first task is to transform the sample points from the planning volume to the model's reference volume. This can be done using the forward transform used to form the planning volume. In theory the reference volume could be used as the planning volume, but due to the differences between breath-hold and free breathing motion demonstrated in Section 7.3, it was decided to plan on one of the free-breathing volumes predicted by the model. The second task is to find the position of the sample points in the current respiratory state using the transformation predicted by the model for the current respiratory parameter(s). This requires the inverse of the transformation to be estimated using the method described in Appendix A.

Applying this algorithm will calculate the leaf-end trajectories modified so that they track the respiratory motion. However, these modifications will have caused an increase to the leaf speed during some parts of the respiratory cycle. If the original leaf speeds

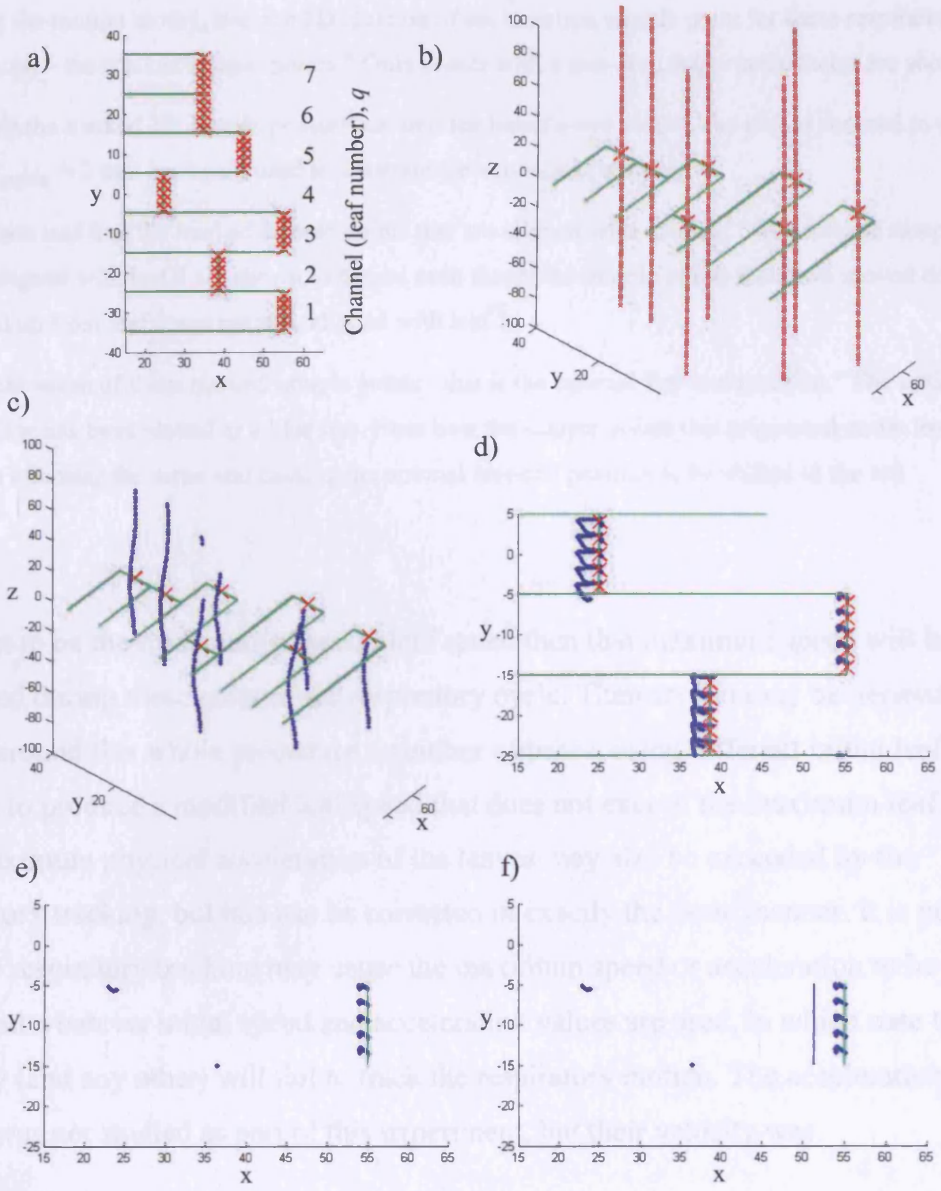


Figure 7.2 - Illustration of the computational algorithm used to implement the differential motion tracking strategy. In all of the above plots the x-axis is the direction of leaf motion, the y-axis is the across-leaf direction, and the z-axis (when shown) is beam direction.

The planned (trailing) leaf positions are shown in green, the sample points are shown in red, and the tracked sample points are shown in blue.

a) "In the beam's-eye view place leaf sample points along each leaf-end spaced $j_{spacing}$ mm apart." $j_{spacing}$ = 2 mm in example shown.

b) "Project the leaf sample points into 3D by placing motion sample points along the beam from $i_{start}$ to $i_{fin}$ equally spaced $i_{spacing}$ mm apart, such that all of the anatomy being tracked is sampled." $i_{start}$ = +100 mm, $i_{fin}$ = -100 mm, $i_{spacing}$ = 2 mm in example shown. Only 1 sample per leaf-end ($j_{spacing}$ = 10 mm) is shown so that the plot is not over-crowded.

c) "Using the motion model, find the 3D location of each motion sample point for these respiratory parameter(s) – the tracked sample points." Only points with a non-zero importance factor are shown.

d) "Project the tracked 3D sample points back into the beam's-eye view." The plot is zoomed in on leaves 2-4 and $j_{spacing}$ = 2 mm has again used to illustrate the across leaf motion.

e) "For each leaf find the tracked sample points that are aligned with that leaf." The tracked sample points that are aligned with leaf 3 are shown. It can be seen that some sample points that have moved down from leaf 4 and up from leaf 2 and are now aligned with leaf 3.

f) "Find the mean of those tracked sample points - this is the optimal leaf-end position." The optimal leaf-end position has been plotted as a blue line. Note how the sample points that originated under leaves 2 and 4 are affecting the mean and causing the optimal leaf-end position to be shifted to the left

were set to be the maximum physical leaf speed then that maximum speed will be exceeded during these parts of the respiratory cycle. Therefore, it may be necessary to iterate around this whole procedure a number of times, using different initial leaf speeds, to produce a modified leaf speed that does not exceed the maximum leaf speed. The maximum physical acceleration of the leaves may also be exceeded by the respiratory tracking, but this can be corrected in exactly the same manner. It is possible that the respiratory tracking may cause the maximum speed or acceleration to be exceeded whatever initial speed and acceleration values are used, in which case this strategy (and any other) will fail to track the respiratory motion. The acceleration of the leaves was not studied as part of this experiment, but their velocity was.

### 7.5.4 Applying the motion tracking strategy to patient data

The combined results from the phase based cyclic B-spline models (see Section 5.5) were used for this experiment. As this experiment was intended to investigate the feasibility of the tracking strategy it only used data from one patient, patient 5. A 3-field IMRT plan was produced (by collaborators at the Institute of Cancer Research and Royal Marsden Hospital) with beams at 0° (A-P), 145°, and 210° (Figure 7.3a). The plan was made on the end-exhale volume predicted by the model. The 210° beam was used for this experiment, as the respiratory motion in the beams-eye-view was greatest for this beam.

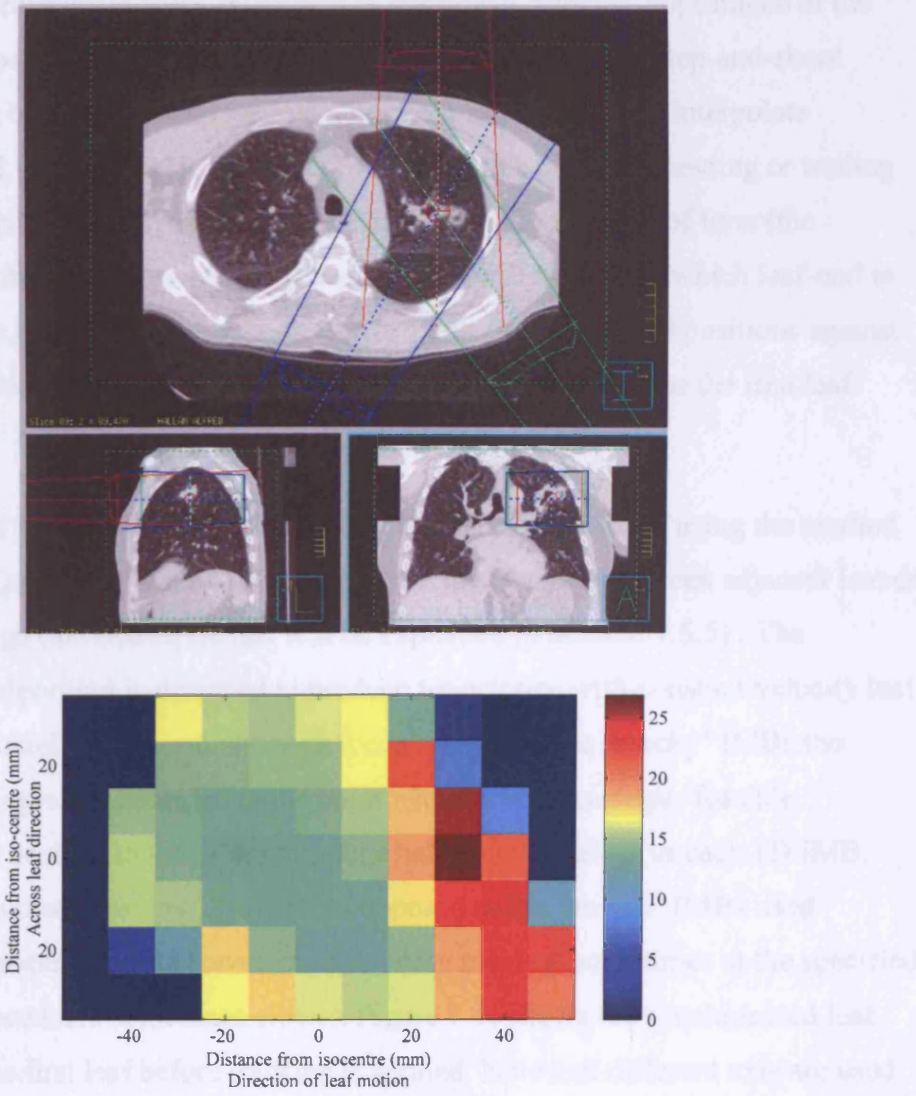


Figure 7.3 - a) Three field IMRT plan - the 210° beam (shown in blue) was used for this study, b) The 2D Intensity Modulated Beam (IMB) for the 210° beam - the scale shows the number of seconds each bixel should be irradiated for.

The 2D Intensity Modulated Beam (IMB) produced by the plan covered 70 mm in the SI direction (Figure 7.3b). The Multi-Leaf Collimators (MLCs) were set to move horizontally (most of the respiratory motion is in the horizontal direction for the patient used) and the IMB was formed using 7 x 10 mm wide leaves. The trajectories of the leading and trailing ends needed to form the IMB were calculated for each of the seven sets of leaves using the equations in Stein et al. (1994). A constant delivery rate for the LINAC of 400 MU/min and a maximum leaf speed of 25 mm/s were assumed. The leaf-edge positions were calculated every 0.1 s over the total treatment time.

Initially the leaf trajectories were calculated to reproduce "blocky" appearance of the IMB that would be produced if the same plan was delivered via the step-and-shoot method. This can be seen as using nearest neighbour interpolation to interpolate between the bixel values in each 1D IMB. To achieve this either the leading or trailing leaf-end will remain stationary at the edge of each bixel for a period of time (the difference in the fluence values of the adjacent bixels will determine which leaf-end is stationary and for how long). Figure 7.4a shows a plot of the leaf-end positions against time that will deliver the "blocky" 1D IMB to the planning volume for the first leaf (most inferior).

It was later found that the leaf trajectories needed to be synchronised using the method described in McQuaid and Webb (2006) to prevent the distance between adjacent leaves becoming too large (the reason for this will be explained in Section 7.5.5) . The synchronisation algorithm is designed to produce trajectories with constant velocity leaf motion between bixel centres and although it can reproduce the 'blocky' IMBs the delivery time was greatly extended to the point where it was no longer feasible. Therefore, it was decided to linearly interpolate between the values in each 1D IMB, producing continuously varying 1D IMBs as opposed to the 'blocky' IMBs used previously. This meant that the leaves are constantly moving, sometimes at the specified maximum leaf speed and sometimes slower. Figure 7.4c shows the synchronised leaf trajectories for the first leaf before tracking is applied. Note that different axis are used to the previous plot as the synchronisation causes the leaves to start further to the left and the delivery to take longer.

The respiratory period was set to 3 seconds (approximately the average respiratory period observed during data acquisition for this patient) and the starting phase was set to end-exhale. For this initial experiment, only tissue within the lungs (including the tumour) was tracked, so all voxels within the lungs in the planning volume (end-exhale) were given an importance factor of 1, and all voxels outside the lungs were given an importance factor of 0. Nearest neighbour interpolation was used to calculate the importance factor at arbitrary points in the planning volume.

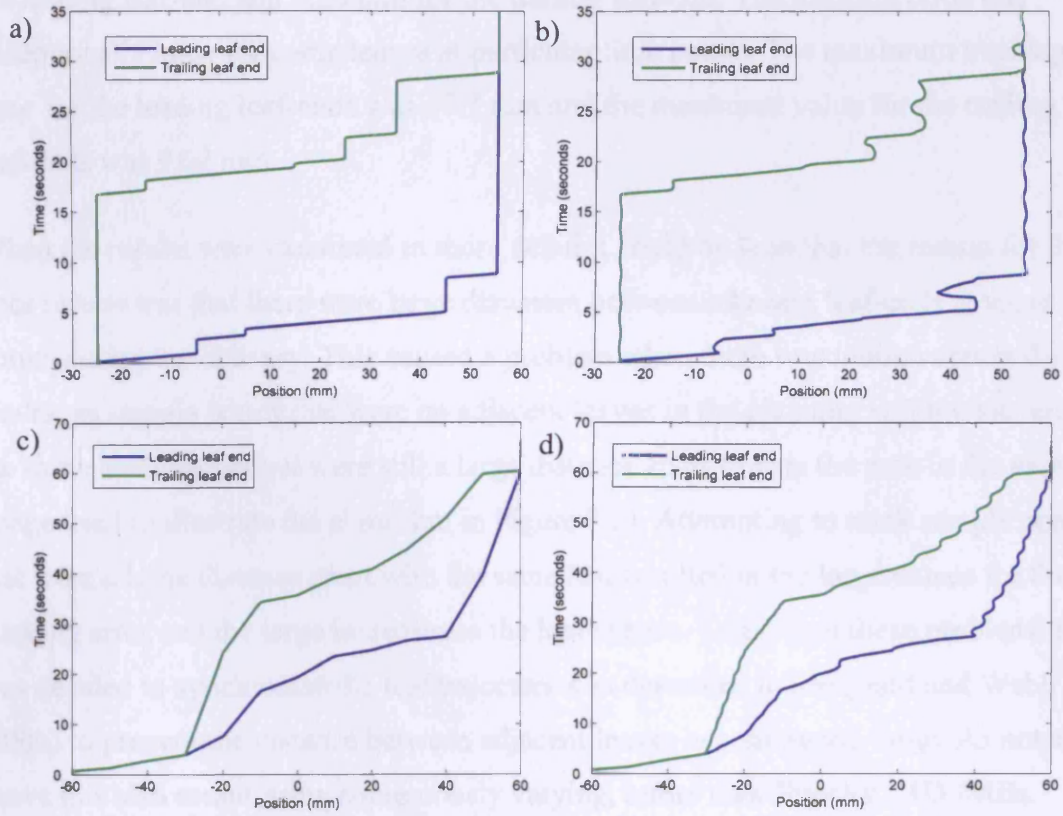


Figure 7.4 - Plots showing the leading and trailing leaf-end trajectories for leaf 1. The position of the leaf-end (relative to the iso-centre) is plotted on the horizontal axis and the time (relative to the starting the delivery of treatment) is plotted on the vertical axis. a) The planned leaf-end trajectories required to deliver the "blocky" 1D IMB to the static planning anatomy. b) The modified leaf-end trajectories that attempt to deliver the "blocky" 1D IMB in the presence of differential motion. c) The planned synchronised leaf-end trajectories required to deliver the "continuous" 1D IMB to the static planning anatomy. b) The modified synchronised leaf-end trajectories that attempt to deliver the "continuous" 1D IMB in the presence of differential motion.

### 7.5.5 Results of applying the motion tracking strategy to patient data

The algorithm presented in Section 7.5.3 was first applied to the leaf trajectories for the "blocky" IMBs with $j_{spacing}$ = 2 mm and $i_{spacing}$ = 2 mm. Figure 7.4b shows the leaf trajectories from Figure 7.4a modified to track the respiratory motion. An initial leaf speed of 25 mm/s (the maximum leaf speed) resulted in a leaf speed of 52 mm/s when tracking the respiratory motion. Initial leaf speeds as low as 5 mm/s still resulted in leaf speeds greater than 25 mm/s when tracking the respiratory motion, meaning that the modified trajectories could not be physically delivered. In addition the mean tracking errors over all leaves and time points were relatively large, with a value of 1.71 mm for

the leading leaf-end and 1.20 mm for the trailing leaf-end. The tracking error was exceptionally large for some leaves at particular time points. The maximum tracking error for the leading leaf-ends was 17.7 mm and the maximum value for the trailing leaf-ends was 9.02 mm.

When the results were examined in more detail it could be seen that the reason for the poor results was that there were large distances between adjacent leaf-ends at some points during the delivery. This caused a problem when there was motion across the leaves, as sample points that were on adjacent leaves in the planning volume moved into the same leaf channel but were still a large distance apart (this is the case in the example image used to illustrate the algorithm in Figure 7.2). Attempting to track sample points that were a large distance apart with the same leaf resulted in the large values for the tracking error and the large increases to the leaf speeds. To prevent these problems it was decided to synchronise the leaf trajectories as described in McQuaid and Webb (2006) to prevent the distance between adjacent leaves becoming too large. As noted above this also meant using continuously varying, rather than "blocky", 1D IMBs.

The motion tracking algorithm was then applied to the synchronised leaf trajectories, again with $j_{spacing}$ = 2 mm and $i_{spacing}$ = 2 mm. By iterating over the algorithm twice it was found that ensuring the initial leaf speed did not exceed 17 mm/s resulted in a leaf speed that was always under 25 mm/s when tracking the respiratory motion, i.e. the modified trajectories were now physically deliverable. The resulting trajectories for the first leaf are shown in Figure 7.4d, and a movie comparing the planned and motion tracking trajectories of all the leaves (animated over a projection through the lung) is included in the CD accompanying this thesis. The tracking error was also considerably improved when using the synchronised trajectories. The mean tracking error for the leading leaf-ends was 0.73 mm and the maximum was 3.87 mm. The mean tracking error for the trailing leaf-ends was 0.73 mm and the maximum was 4.12 mm.

In order to further assess the effects of the respiratory motion and the ability of the proposed strategy to correct for those effects, the unattenuated fluence delivered to a plane of tissue perpendicular to the beam direction can be calculated. This can be done using the planned leaf trajectories and the (static) planning volume to produce the planned fluence, the 2D IMB, or using the planned leaf trajectories and the motion model to produce the fluence delivered in the presence of motion. Also, the modified

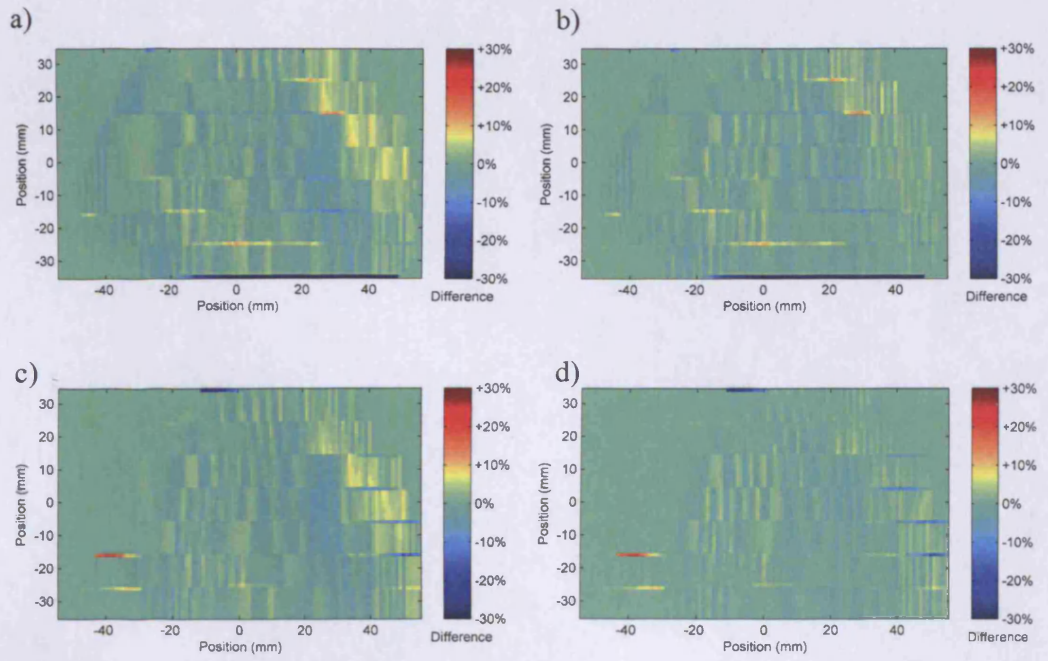


Figure 7.5 - Difference maps showing the difference between the planned unattenuated fluence and that delivered in the presence of differential motion: a) using the planned leaf trajectories for a plane passing through the iso-centre, b) using the modified leaf trajectories for a plane passing through the iso-centre, c) using the planned leaf trajectories for a plane 20 mm in front of the iso-centre, and d) using the modified leaf trajectories for a plane 20 mm in front of the iso-centre. In all maps the difference is given as a percentage of the maximum planned unattenuated fluence.

leaf-end trajectories can be used instead of the planned leaf-end trajectories to calculate the unattenuated fluence delivered when motion compensation is applied. Difference maps have been produced between the planned fluence and the fluence delivered in the presence of motion, both with and without motion compensation applied. These results were generated using the synchronized leaf trajectories. Figure 7.5a shows the difference map for a plane passing through the isocentre without motion compensation applied and Figure 7.5b shows the same plane with motion compensation applied. Figure 7.5c shows the difference map for a plane 20 mm in front of the isocentre without motion compensation applied and Figure 7.5d shows the same plane with motion compensation applied. In all the maps the differences are expressed as a percentage of the maximum unattenuated fluence. The maps are from the beam's-eye-view and the location (0,0) is the isocentre. The large differences present at the bottom of Figures 7.5a and 7.5b are caused by the tissue moving outside of the area covered by the initial 2D IMB. When this happens, the tissue is ignored in the current

implementation, and this highlights the need to address this problem in the future. Figures 7.5a and 7.5c show that although the differences to the fluence caused by motion are similar for the two planes, they are not identical, confirming that differential motion was occurring. Figures 7.5b and 7.5d show that using motion compensation reduces the differences between the planned fluence and the fluence delivered in the presence of motion, but does not remove them entirely, and in a few small areas even appears to increase them. This result is exactly what would be expected when trying to compensate for differential motion.

### 7.5.6 Conclusion and discussion

This work has demonstrated how the motion models can be used to implement a strategy for tracking the 2D differential motion in the beam's eye view when delivering IMRT via the DMLC technique. This work only applied the strategy to one example case, and the strategy would need to be assessed on far more data and far more rigorously to truly evaluate how well the strategy can account for respiratory motion. However, this test case still provided a useful insight into how the strategy and method can be further developed.

This work has demonstrated the need for using synchronised leaf trajectories when there is motion across the leaves, which there will almost always be in clinical practice. The difference maps in Figure 7.5 show that differential motion does occur as the effects of the respiratory motion are different on the two planes. Tracking the differential motion helps to compensate for the effects of the motion on both planes. It is postulated that if the motion was only tracked on one plane (e.g. the one passing through the isocentre), then the motion compensation would be better for this plane but would be worse for the other.

Of course, it may be desirable to track the motion at the isocentre more accurately than at other points, as there will typically be more dose delivered near the isocentre than at other points. This is why the Importance factor, $I(p,j,i)$, has been included in equations [7.1] and [7.2]. In this work the Importance factor was simply set to 1 for all voxels within the lungs (including the tumour) and to 0 for all voxels outside, meaning all voxels within the lungs were tracked equally and voxels outside the lungs were ignored. The Importance factor could be used much more intelligently so that the tracking

focuses on regions of high dose or high dose gradients, or critical structures where the dose must be delivered accurately.

There are many parts of the methodology that need further investigation or development. A few different values of $j_{spacing}$ were investigated as part of this work, and it was found that using only one sample per leaf-end ($j_{spacing}$ = 10 mm) was insufficient to capture the across leaf motion. However, the effects of sampling the motion at different resolutions (changing $i_{spacing}$ and the temporal resolution as well as $j_{spacing}$) needs to be further investigated. The method can be directly applied to 2D IMBs of any size and resolution (providing they are not larger than the model) and for any size MLC leaf, and the bixel and leaf sizes do not necessarily need to be equal as they were here. The method can be easily extended to use diverging beams rather than parallel beams. The method needs to be extended so that it can account for tissue moving into and out of the leaves used to form the original 2D IMB. This should be relatively straight forward by including extra leaves above and below those currently used.

This implementation of the strategy assumed completely reproducible breathing (i.e. no variation in period or deformation), which is clearly unrealistic. The strategy itself does not require that the breathing is reproducible, just that the full internal motion and deformation is known at all time points. Two parameter models can potentially model some of the inter-cycle variation, and could be used to provide information on the internal deformation at any time point, accounting for variations in the motion. If it is possible to measure the respiratory parameters, obtain the model prediction, and then apply all the tracking calculations in "real-time", then accounting for variations in the respiratory motion is straight forward. However, if it is not possible to perform all these calculations in "real-time" (which is likely with current hardware) then results must be pre-calculated. This can be done by creating a look-up table giving the leaf-end positions for all permitted values of the respiratory parameters at all time points throughout delivery (this will require a huge number of calculations, especially for the two parameter models). This can be refined, particularly if the respiration can be made somewhat reproducible (e.g. by using visual feedback), by restricting the permitted respiratory parameters at different time points during delivery and only calculating the leaf-end positions for these respiratory values (the implementation used here is an extreme example of this, with the leaf-end positions only calculated for one value of the respiratory parameter). If the values of the respiratory parameters observed at a particular time point during treatment are within the permitted range, the leaf-end

positions can be found in the look-up table, if they are not within the permitted range then the treatment will need to be paused until the parameters return to the permitted range (this is similar to the idea presented by Neicu et al. 2003, where the tracked delivery is paused until synchronisation between the leaf positions and the respiratory parameter is resumed.).

The tracking strategy presented here was based on minimising the geometrical tracking error, but it may be more meaningful to base the strategy on minimising the error in the fluence or dose delivered. However, this is still based on the idea of planning on a single 3D volume and then trying to reproduce the 3D plan or moving anatomy. Superior results may well be achieved by trying to incorporate the motion information earlier and actually planning on the moving anatomy.

In conclusion, there is still a long way to go and a lot of work that needs to be done before IMRT tracking will be ready for clinical use, although it is, in my opinion, a potentially promising method, and is an active area of research in several labs.

As previously noted, the tracking strategy itself was developed by Steve Webb, and while I personally think there is some potential for the strategy, the aim of this work was not to develop the tracking strategy or assess its performance, but to demonstrate how the motion models could be utilised in a way that may one day translate to clinical practice.

# Chapter 8

# Conclusions and future work

This final chapter discusses the findings and conclusions from each chapter. It also suggests some of the future work that may follow on from this research.

## 8.1 Conclusions and discussion

This thesis has presented a methodology for constructing respiratory motion models that relate the complex internal motion and deformation, described by non-rigid registrations, to one or two respiratory parameters derived from an external respiratory surrogate signal. A framework for constructing and evaluating the models has been

developed, and a number of different possible models have been evaluated. The results show that some of these models can predict the respiratory motion very accurately. One example application of the models in lung RT, planning tracked IMRT treatment, has been demonstrated in some detail.

It is acknowledged that there still remains a lot of work that must be done before respiratory motion models can be used to fulfil their full potential. However, the motion models presented in this thesis potentially offer substantial improvements over current methods of imaging and analysing the respiratory motion. Many institutions are now acquiring 4DCT to obtain information on the respiratory motion. Several authors have highlighted the benefits of non-rigidly registering the 4DCT datasets, such as propagating targets delineations and performing dose calculations that account for respiration. As the models in this thesis are also based on non-rigid registrations they also provide all of these benefits. In addition the motion models have two further advantages over registered 4DCT datasets.

Firstly, the motion models are fully continuous over the respiratory parameter(s) used. This can be very useful for tracking and gating applications were it may be desirable to estimate the motion at a finer temporal resolution than provided by 4DCT. This is demonstrated by the IMRT tracking application in Section 7.5, where the internal motion is sampled every 0.1 seconds, corresponding to 30 times per respiratory cycle (a typical 4DCT volume only sample the motion 8-10 times per respiratory cycle).

Secondly, the models are constructed by fitting a function to 'unsorted' data. This means that the models can account for variation in the respiratory motion which is not possible when using 'sorted' 4DCT, as the sorting process assumes there is no variation. As the models fit a function to the data any variation not explained by the respiratory parameter(s) is averaged out. Additionally, the models can use two respiratory parameters (and could be extended to more, see future work below) which may permit some of the variation to be modelled. It also means that the strength of the relationship between the internal motion and the respiratory parameters can be assessed prior to combining the data from different couch positions. If the relationship is not sufficiently strong then different parameters and functions can be assessed to see if a better relationship can be established.

There now follows a summary of the findings and conclusions from each chapter in this thesis (other than chapter 2 which contained the background and literature review).

### 8.1.1 Chapter 3 summary and conclusions

Chapter 3 described the data that was acquired for constructing the motion models and the protocols used for acquiring the data. Cine CT data was acquired at several couch positions to sample the internal respiratory motion. A respiratory signal was simultaneously acquired for calculating the phase, amplitude, and gradient respiratory parameters. This was derived directly from the Cine CT volumes for some patients (1-4), by tracking a marker attached to the abdomen for some patients (5-6), and by acquiring 3D surface data and using it to calculate internal volume changes for the others (7-11). Also, a reference CT volume was acquired for each patient. This was acquired at breath-hold so that the volume could be high quality and artefact free.

In conclusion, the data acquired was sufficient for developing the motion modelling methodology and for constructing the models in the following chapters. However, there are, potentially, numerous ways of improving the data acquisition methods to generate superior data for constructing the models (these are discussed below in Section 8.2 Future work).

### 8.1.2 Chapter 4 summary and conclusions

Chapter 4 described the non-rigid registrations that were performed to recover the internal respiratory motion. It details the different protocols used to perform the registrations and evaluates their results. The results show that the registrations used to construct the motion models in the following chapters were generally very accurate (mean errors of 1.26 mm and 1.13 mm for the two different registration protocols assessed as part of this research). However, the registrations failed for a relatively high proportion of the patients (one in six when 1.5 mm slices used, but three in five when 2.4 mm slices were used). Not all registrations failed for these patients, although it was decided to exclude all of their results from further study so as not to bias later results. The registrations that failed were usually for volumes acquired near end-inhale (the

reference volume was acquired at exhale breath-hold). It was thought that many of the failed registrations were caused by lung tissue sliding past the chest wall (the sliding problem - see Section 2.4.5.2). However, there are many other factors that could have contributed, including: the density change effect (see section 2.4.5.1), registering relatively thin 'slabs' of data (only 12 or 16 slices), not having good enough starting estimates for the registrations, and possible 'bugs' in the code used to perform the registrations. In addition, the current run-time for the registrations needs to be dramatically reduced for their use to be clinically viable, although there is evidence that this should be achievable using more advanced implementations of the registration algorithm and dedicated hardware such as Graphics Processing Unit (GPU).

In conclusion, very good registration results were achieved for some patients, and these provided adequate data for constructing the models in the following chapters. However, the robustness and run-time of the registrations both need addressing.

### 8.1.3 Chapter 5 summary and conclusions

Chapter 5 described how to construct respiratory motion models that related the internal motion to a single respiratory parameter. The motion models use one dimensional functions to relate the registration results to the phase of the respiratory cycle. As the models use a single respiratory parameter they attempt to average out any inter-cycle variation in the data used to construct the models (20 seconds of data, corresponding to 2-8 respiratory cycles, were used for each model). Two functions were developed that could be used for the motion models, the cyclic B-spline function and the Fourier series function, and their performance was compared to that of polynomial functions. Each of the functions was evaluated with varying degrees of freedom. The ability of the different model to predict the registration results was assessed. Many of the functions performed well, and the cyclic B-spline function with four respiratory control points was adopted for further evaluation. All experiments in this chapter were performed using the leave-out-target technique so that over-fitting could be detected.

The cyclic B-spline models were evaluated in more detail by assessing their ability to predict the Cine CT volumes and manually identified landmark points located within them. The results showed that the motion models could predict the Cine CT data with a high degree of accuracy (mean error 1.56 mm). A method of combining the model

predictions from different couch positions into a single continuous transformation was also described and assessed in this chapter. The combined transformation produced a plausible prediction of the anatomy with no discontinuities between adjacent couch positions, with only a small decrease in the accuracy for predicting the individual Cine CT volumes (mean error 1.71 mm). It should be stressed that even though the combined predictions are less accurate at predicting the CT data acquired for constructing the models, they should in theory be better at predicting future data (i.e. during treatment) as their predictions are physically plausible (unlike the individual model predictions or 4DCT, which can have discontinuities between adjacent couch positions).

In conclusion, a method has been developed for constructing and assessing respiratory motion models related to the phase of the respiratory cycle. The results have shown that a cyclic B-spline function can be used to construct motion models that are able to predict the Cine CT volumes very accurately.

### 8.1.4 Chapter 6 summary and conclusions

Chapter 6 extended the motion models developed in chapter 5 so that the internal motion can be related to two respiratory parameters, with the idea that this may have permitted some of the inter-cycle variation to be modelled, thus producing more accurate models. Two parameter models were constructed based on the amplitude and gradient parameters and on the phase and amplitude parameters, and were compared to the phase based models developed in the previous chapter and single parameter models based on the amplitude. A number of example function were presented that could be used for the one or two parameters models, and their performance was evaluated both using the registration results and manually located landmarks in the Cine CT data. Both methods of assessing the models found similar results. Most of the models produced very good results, although the models with the most degrees of freedom were clearly over-fitting the data.

The overall results showed that the phase based models developed in the previous chapter were just as accurate as any of the two parameter models. This does not necessarily mean the idea of using two respiratory parameters to model the inter-cycle variation is flawed; there may have been too little variation in the data (only a few respiratory cycles were acquired at each couch position), the variation may not have

been sampled well enough, the respiratory parameters (and the surrogate signals they are derived from) chosen may not have a strong relationship with the inter-cycle variation, or the functions used for the two parameter models may not have been appropriate. The inter-cycle variation must be studied in much more detail, a far wider range of respiratory parameters and surrogate signals need to investigated, and more appropriate functions for modelling the relationships between the parameters and the variation need to be developed, before firm conclusions about the ability to model the inter-cycle variation can be drawn. The aim of evaluating the different models was to demonstrate the different types of function required for different respiratory parameters and the effects of using functions with too many degrees of freedom, and this was achieved.

In conclusion, a framework has been presented for constructing and evaluating respiratory motion models based on one or two respiratory parameters, and this has been demonstrated on a number of example models. The results show that the models can predict the respiratory motion very accurately, although their ability to predict variations in the respiratory motion is inconclusive.

### 8.1.5 Chapter 7 summary and conclusions

Chapter 7 described a number of possible uses for the motion models related to lung RT. These included 'medical research applications', where the aim is to learn something about the nature of the internal respiratory motion and/or its relationship to external surrogate signals, but the methodology used is not itself intended for clinical practice. As an example of a simple 'medical research application' the motion models were used to compare the motion that occurs during free breathing to the motion that occurs between two breath hold volumes.

Also, a number of uses of the motion models that are (ultimately) intended to translate to clinical practice were described. These included using the motion models to generate more accurate target delineations, using the models to perform accurate dose calculations that account for the respiratory motion, and using the models to plan and deliver 'active' treatment that responds to the respiration, such as gating and tracking. As an example of one of the 'clinical' uses of the motion models, they were used to implement a strategy for tracking the 2D differential motion during the delivery of

IMRT developed by Prof. Steve Webb, a collaborator from the Institute of Cancer Research and Royal Marsden Hospital. This work demonstrated how the motion models could be utilised to plan IMRT treatment, and that there is some promise for tracking respiratory motion during the delivery of IMRT, although there is still a lot to do before tracked IMRT will be clinically viable.

In conclusion, there are a wide variety of potential uses for the motion models in lung RT. They can be a valuable tool for analysing the internal respiratory motion and its relationship to the respiratory parameters, and can also be utilised in methods that are intended to translate to clinical practice. Two example applications demonstrated the use of the motion models in further detail: examining the difference between breath-hold and free breathing motion models, and planning IMRT tracking.

## 8.2 Future work

The models need to be assessed on much more data to truly evaluate their performance. Not only do they need to be evaluated on data from more patients, but the models need to be evaluated on entirely different data to that used to construct the motion models. Data needs to be acquired for longer periods of time, and over several different sessions that are days and weeks apart, in order to study the short and long term variation in the motion. The nature of the variation needs to be fully investigated. It needs to be established how much variation can occur and over what time frames. The effects that visual feedback and other respiratory training techniques can have on the reproducibility of the respiratory motion need to be studied. The amount of data required to adequately sample the variation needs to be determined.

It is thought that it will be possible to relate some of the variations, particularly the shorter term variations, to external surrogate signals. This may entail the development of more appropriate modelling functions for relating the motion and its variation to the respiratory parameters. This may include extending the models to use three or more parameters (although it is thought that substantially more data will be required to prevent models of three or more parameters from over-fitting) or using statistical models (these are already being investigated by colleagues in our laboratory). However, it may not be possible to relate some of the longer term variation to any external

surrogate signals. In this case methods will need to be developed for updating the models to account for the longer term variation. The Vision RT system may prove very useful when it comes to studying long term variations in the respiratory motion as it may be able to detect base-line shifts between different sessions, which are difficult to detect with other surrogate signals.

Acquiring enough data to study and model the inter-cycle variation may prove very difficult with (current) CT technology due to dose considerations. Therefore methods of reducing the dose for CT and other imaging modalities that do not deliver radiation (such as MR) need to be investigated. Having data that covered a larger region of the anatomy, ideally the entire region of interest, would also be beneficial.

Different respiratory surrogate signals need to be acquired from the same patients so that their ability to predict the internal motion and its variation can be directly compared. The models could be used to relate the motion to two (or more) parameters from separate surrogate signals, which may perform better than relating the motion to two parameters derived from the single signal. Statistical parameters that incorporate information from many different surrogate signals could also be used. The Vision RT system may prove very useful for this as it is possible to track the motion at several different points or regions on the skin surface simultaneously. If phase continues to prove to be a useful parameter, as it did in this research, then robust methods of estimating the phase in real time need to be developed.

There are a number of issues with the registrations that need addressing, the most important being to establish why some of the registrations failed and how they can be made to succeed. The sliding problem (i.e. the failure of the registration algorithm when two tissues slide against each other, see Section 2.4.5.2) needs to be formally addressed and transformation models that can account for the sliding motion need to be developed. The effect that density changes within the lung (and hence its intensity value in CT, see Section 2.4.5.1) has on the registration results needs to be examined. If the density change causes errors in the registrations then a method of accounting for the density change needs to be developed. Better starting estimates for the non-rigid registrations may help improve their robustness, and protocols for obtaining appropriate starting estimates need to be developed.

Different implementations of the B-spline registration algorithm and different non-rigid registration algorithms may have superior performance, in terms if accuracy, speed, or run-time, over the implementation of the B-spline algorithm used in this research. Implementing the non-rigid registration algorithm on a Graphics Processing Unit (GPU) may facilitate considerable reductions in the run time of the registrations. Bio-mechanical constraints and models can be utilised in the registrations to help improve the robustness and prevent unfeasible deformations.

The motion models can be considered 4D or 5D transformation models, as they describe the 3D deformation over one or two respiratory dimensions (parameters). Therefore the models could be used to perform 4D, or maybe even 5D, registrations, effectively performing the registration and modelling steps used in this thesis all one process. This approach was not pursued during this research for two reasons: the predicted run-time and problems with accounting for variation. Using the 3D registration run-time as a guide it was expected that a 4D registration would have a very long run-time (and even more so for a 5D registration). Although this may actually be shorter than the sum of the run-times for all the separate 3D registrations, the different 3D registrations can easily be run in parallel on the large computer cluster available in our laboratory, whereas substantial work would be required to parallelise a 4D registration algorithm to efficiently use more than one processor. By fitting a function to the data, the models in this thesis average out variation not explained by the respiratory parameters used. However, it is unclear how 4D registrations would behave when there is variation in the data although it is unlikely they would correctly average it out without further development.

Despite these two potential problems, using the models developed here for 4D or 5D registrations certainly warrants study in the future. Another possibility that should be investigated is performing 4D registrations where time is the 4th dimension. This relates the motion to time rather than the respiratory parameters, so the registration does not need to do any averaging. The 4D transformation could then be used to sample the 3D deformation at several time-points, and the motion modelling methods from this thesis could then be used to establish the relationship between the deformation and the respiratory parameters.

Unsorted Cine CT data was used in this thesis as it provides 3D data, albeit with a limited coverage in the Superior-Inferior direction, that can be considered a 'snap-shot'

of the motion. However, this data is actually acquired over 0.5 seconds (or 0.75 seconds at the start of the research), so can be subject to motion artefacts itself. Recent work by Zeng et al. (2007a) has proposed performing the registrations directly to Cone-Beam CT projection data. Each projection is acquired extremely quickly so motion effects will be truly negligible. This method can potentially provide 3D motion information over a large field of view without needing to sort the data according to a respiratory parameter. The method proposed by Zeng et al. (2007a) needs further development, but if this method can be made to work it could provide an invaluable source of data on the variation in the respiratory motion, and for using to construct the motion models presented in this thesis. It may also be possible to apply similar ideas to MR data, using individual slices, or possibly even k-space data, for the registrations.

In addition to the work that can be done on improving the motion modelling methodology and registrations used to construct them, described above, the many potential uses for the motion models, both in lung RT and in other applications need further development. Chapter 7 discussed many of the ways that the motion models can be utilised in lung RT, including defining targets and performing dose calculation that account for respiration, and for planning gated or tracked treatment. More work is required both to develop the RT methodologies themselves (especially the more adventurous ones, such as IMRT tracking), and to incorporate the models developed in this thesis into those methodologies.

As discussed in Section 2.6 there are a number of uses for the motion models outside of lung RT. These include other image directed therapies or image guided interventions, e.g. RT for other sites in the thorax and abdomen, High Intensity Focused Ultrasound for the treatment of liver and breast cancer, Radio-Frequency Ablation of liver tumours, and Cardiac EP interventions. The models could also be used to help reconstruct motion free images when the image data is corrupted by motion, e.g. high resolution MR, Cone-Beam CT, or PET. Finally, the motion models can be used to learn about the physiology of respiration. They could be used to study the effects of different diseases, such as Chronic Obstructive Pulmonary Disease (COPD), on local lung function, and for assessing the response to potential treatments to these diseases.

# Bibliography

Ablitt N A, Gao J, Keegan J, Stegger L, Firmin D N, and Yang G-Z, "Predictive cardiac motion modeling and correction with partial least squares regression," *IEEE Trans. Med. Imaging* 23, 1315-1324 (2004).

Alasti H, Cho Y B, Vandermeer A D, Abbas A, Norrlinger B, Shubbar S, and Bezjak A, "A novel four-dimensional radiotherapy method for lung cancer: imaging, treatment planning and delivery," *Phys. Med. Biol.* 51, 3251-3267 (2006).

Allen A M, Siracuse K M, Hayman J A, and Balter J M, "Evaluation of the influence of breathing on the movement and modeling of lung tumors," *Int. J. Radiat. Oncol., Biol., Phys.* 58, 1251-1257 (2004).

Balter J M, Ten Haken R K, Lawrence T S, Lam K L, and Robertson J M, "Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing," *Int. J. Radiat. Oncol., Biol., Phys.* 36, 167–174 (1996).

Batchelor P G, Atkinson D, Irarrazaval P, Hill D L G, Hajnal J, and Larkman D, "Matrix description of general motion correction applied to multishot images," *Magnetic Resonace in Medicine* 54, 1273-1280 (2005).

Benchetrit G, "Breathing pattern in humans: diversity and individuality," *Respiration Physiology* 122, 123-129 (2000).

Berbeco R I, Nishioka S, Shirato H, Chen G T Y, and Jiang S B, "Residual motion of lung tumours in gated radiotherapy with external respiratory surrogates," *Phys. Med. Biol.* 50, 3655-3667 (2005)

Berlinger K, Roth M, Sauer O, Vences L, and Schweikard A, "Fully automatic detection of corresponding anatomical landmarks in volume scans of different respiratory state," *Med. Phys.* 33, 1569-1572 (2006).

Blackall J M, Penney G P, King A P, and Hawkes D J, "Alignment of sparse freehand 3-D ultrasound with preoperative images of the liver using models of respiratory motion and deformation," *Trans. Med. Imaging* 24, 1405-1416 (2005).

Blackall J M, Ahmad S, Miquel M E, McClelland J R, Landau D B, Hawkes D J, "MRI Based Measurement of Respiratory Motion Variability and Assessment of Imaging Strategies for Radiotherapy Planning," *Phys. Med. Biol.* 51, 4147-4169 ( 2006).

Bortfeld T, Jiang S B, and Rietzel E, "Effects of motion on the total dose distribution," *Semin. Radiat. Oncol.* 14, 41-51 (2004).

Britton K R, Starkschall G, Tucker S L, Pan T, Nelson C, Chang J Y, Cox J D, Mohan R, and Komaki R, "Assessment of gross tumour volume regression and motion changes during radiotherapy for non-small-cell lung cancer as measured by four-dimensional computed tomography," *Int. J. Radiat. Oncol., Biol., Phys.* 68, 1036-1046 (2007).

Brock K K, Sharpe M B, Dawson L A, Kim S M, and Jaffray D A, "Accuracy of finite element model-based multi-organ deformable image registration," *Med. Phys.* 32, 1647-1659 (2005).

Brock K K, on behalf of the Deformable Registration Accuracy Consortium, "A multi-institution deformable registration accuracy study," *Int. J. Radiat. Oncol., Biol., Phys.* 69, S44 (2007).

Chandler A G, "Slice-to-volume registration for tomographic images," PhD thesis, University of London (2005).

Chandrashekara R, Mohiaddin R H, and Rueckert D, "Cardiac motion tracking in tagged MR images using a 4D B-spline motion model and nonrigid image registration," *IEEE International Symposium on Biomedical Imaging: Macro to Nano*, 468-471 (2004).

Chen Q, Weinhous M S, Deibel F C, Ciezki J P, and Macklis R M, "Fluoroscopic study of tumour motion due to breathing: Facilitating precise radiation therapy for lung cancer patients," *Med. Phys.* 28, 1850-1856 (2001).

Christensen G E, Song J H, Lu W, Naqa I E, and Low D A, "Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry," *Med. Phys.* 34, 2155-2163 (2007).

Coselmon M M, Balter J M, McShan D L, and Kessler M L, "Mutual information based CT registration of the lung at exhale and inhale breathing states using thin-plate splines," *Med Phys* 31, 2942-2948 (2004).

Dawson L A, Brock K K, Kazanjian S, Fitch D, McGinn C J, Lawrence T S, Haken R K T, and Balter J M, "The reproducibility of organ position using active breathing control (ABC) during liver radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 51, 1410-1421 (2001).

de Mey J, Van de Steene J, Vandenbroucke F, Trappeniers L, Meysman M, Everaert H, Noppen M, Storme G, and Bossuyt A, "Percutaneous Placement of Marking Coils before Stereotactic Radiation Therapy of Malignant Lung Lesions." *J. Vasc. Inter. Radiol.* 16, 51-56 (2005).

Flampouri S, Jiang S B, Sharp G C, Wolfgang J, Patel A A, and Choi N C, "Estimation of the delivered patient dose in lung IMRT treatment based on deformable registration of 4D-CT data and Monte Carlo simulations," *Phys. Med. Biol.* 51, 2763-2779 (2006).

Ford E C, Mageras G S, Yorke E, and Ling C C, "Respiration-correlated spiral CT: A method of measuring respiratory induced anatomic motion for radiation treatment planning," *Med. Phys.* 30, 88-97 (2002).

George R, Chung T D, Vedam S S, Ramakrishnan V, Mohan R, Weiss E, and Keall P J, "Audio-visual biofeedback for respiratory-gated radiotherapy: impact of audio instruction and audio-visual biofeedback on respiratory-gated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 65, 924-933 (2006).

Giraud P, De Rycke Y, Dubray B, Helfre S, Voican D, Guo L, Rosenwald J C, Keraudy K, Housset M, Touboul E, and Cosset J M, "Conformal radiotherapy (CRT) planning for lung cancer: analysis of intrathoracic organ motion during extreme phases of breathing," *Int. J. Radiat. Oncol., Biol., Phys.* 51, 1081-1092 (2001).

Graham M V, Purdy J A, Emami B, Harms W, Bosch W, Lockett M A, and Perez C A, "Clinical dose–volume histogram analysis for pneumonitis after 3d treatment for non-small cell lung cancer (NSCLC)," *Int. J. Radiat. Oncol., Biol., Phys.* 45, 323-329 (1999).

Guckenberger M. Wilbert J, Meyer J, Baier K, Richter A, and Flentje M, "Is a single respiratory correlated 4D-CT study sufficient for evaluation of breathing motion?" *Int. J. Radiat. Oncol., Biol., Phys.* 67, 1352-1359 (2007).

Guerrero T, Sanders K, Castillo E, Zhang Y, Bidaut L, Pan T, and Komaki R, "Dynamic ventilation imaging from four-dimensional computed tomography," *Phys. Med. Biol.* 51, 777–791 (2006).

Hajnal J V, Hawkes D J, and Hill D L G, *Medical Image Registration*, CRC Press LLC, (2001).

Hanley J, Debois M M, Mah D, Mageras G S, Raben A, Rosenzweig K, Mychalczak B, Schwartz L H, Gloeggler P J, Lutz W, Ling C C, Leibel S A, Fuks Z, and Kutcher G J, "Deep inspiration breath-hold technique for lung tumors: the potential value of target immobilization and reduced lung density in dose escalation," *Int. J. Radiat. Oncol., Biol., Phys.* 45, 603-611 (1999).

Heath E and Seuntjens J, "A direct voxel tracking method for four-dimensional Monte Carlo dose calculations in deforming anatomy," *Med. Phys.* 33, 434-445 (2006).

Heath E, Collins D L, Keall P J, Dong L, and Seuntjens J, "Quantification of accuracy of the automated nonlinear image matching and anatomical labeling (ANIMAL) nonlinear registration algorithm for 4D CT images of lung," *Med. Phys.* 34, 4409-4421 (2007).

Hoisak J D, Sixel K E, Tirona R, Cheung P C, and Pignol J P, "Correlation of lung tumour motion with external surrogate indicators of respiration," *Int. J. Radiat. Oncol., Biol., Phys.* 60, 1298-1306 (2004).

Hui H, Pan T, and Shen Y, "Multislice helical CT: Image temporal resolution," *IEEE Trans. Med. Imaging* 19, 384–390 (2000).

International Commission on Radiation Units and Measurements. *ICRU report 50: Prescribing, recording and reporting photon beam therapy.* Bethesda: ICRU (1993).

International Commission on Radiation Units and Measurements. *ICRU report 62: Prescribing, recording and reporting photon beam therapy (supplement to ICRU report 50).* Bethesda: ICRU (1999).

Johnson U, Landau D, Lindgren-Turner J, Smith N, Meir I, Howe R, Rodgers H, Davit S, and Deehan C, "Real Time 3D Surface Imaging for the Analysis of Respiratory Motion during Radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 60, S603-S604 (2004).

Kashani R, Lam K, Litzenberg D, and Balter J, "Technical note: A deformable phantom for dynamic modeling in radiation therapy," *Med. Phys.* 34, 199-201 (2007a).

Kashani R, Balter J, Kessler M, Hub M, Dong L, Zhang L, Xing L, Xie Y, Hawkes D, Schnabel J, McClelland J, and Joshi S, "Objective assessment of deformable image registration in radiotherapy - a multi-institution study," *Med. Phys.* 34, 2545 (2007b).

Keall P J, Kini V R, Vedam S S, and Mohan R, "Motion adaptive x-ray therapy: a feasibility study," *Phys. Med. Biol.* 46, 1-10 (2001).

Keall P J, Starkschall G, Shukla H, Forster K M, Ortiz V, Stevens C W, Vedam S S, George R, Guerrero T, and Mohan R, "Acquiring 4D thoracic CT scans using a multislice helical method", *Phys. Med. Biol.* 49, 2053-2067 (2004).

Keall P J, Joshi S, Vedam S S, Siebers J V, Kini V R, and Mohan R, "Four-dimensional radiotherapy planning for DMLC-based respiratory motion tracking," *Med. Phys.* 32, 942-951 (2005).

Keall P J, Mageras G S, Balter J M, Emery R S, Forster K M, Jiang S B, Kapatoes J M, Low D A, Murphy M J, Murray B R, Ramsey C R, Van Herk M B, Vedam S S, Wong J W, and Yorke E, "The management of respiratory motion in radiation oncology report of AAPM Task Group 76," *Med. Phys.* 33, 3874-3900 (2006).

Khamene A, Warzelhan J K, Vogt S, Elgort D, Chefd'Hotel C, Duerk J L, Lewin J, Wacker F K, and Sauer F, "Characterization of internal organ motion using skin marker positions," MICCAI 2004, 526-533 (2004).

Kini V R, Vedam S S, Keall P J, Patil S, Chen C, and Mohan R, "Patient training in respiratory-gated radiotherapy," Medical Dosimetry 28, 7-11 (2003).

Klein S, Staring M, and Pluim J P, "Comparison of gradient approximation techniques for optimization of mutual information in nonrigid registration," *Proceedings of SPIE Medical Imaging 2005: Image Processing*, 192-203 (2005).

Koch N, Liu H H, Starkschall G, Jacobson M, Forster K, Liao Z, Komaki R, and Stevens C W, "Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part I – correlating internal lung motion with skin fiducial motion," *Int. J. Radiat. Oncol., Biol., Phys.* 60, 1459-1472, (2004).

Kong F M, Ten Haken R K, Schipper M J, Sullivan M A, Chen M, Lopez C, Kalemkerian G P, and Hayman J A," High-dose radiation improved local tumor control and overall survival in patients with inoperable/unresectable non-small-cell lung cancer: Long-term results of a radiation dose escalation study," *Int. J. Radiat. Oncol., Biol., Phys.* 63, 324-333 (2005).

Kubo H D and Bruce C H, "Respiration gated radiotherapy treatment: a technical study," *Phys. Med. Biol.* 41, 83-91 (1996).

Lagerwaard F J, Van Sornsen De Koste J R, Nijssen-Visser M R J, Schuchhard-Schipper R H, Oei S S, Munne A, and Senan S, "Multiple "slow" CT scans for incorporating lung tumour mobility in radiotherapy planning," *Int. J. Radiat. Oncol., Biol., Phys.* 51, 932-937 (2001).

Li T, Schreibmann E, Thorndyke B, Tillman G, Boyer A, Koong A, Goodman K, and Xing L, "Radiation dose reduction in four-dimensional computed tomography," *Med. Phys.* 32, 3650-3660 (2005).

Li T, Schreibmann E, Yang Y, and Xing L, "Motion correction for improved target localization with on-board cone-beam computed tomography," *Phys. Med. Biol.* 51, 253-267 (2006a).

Li T, Thorndyke B, Schreibmann E, Yang Y, and Xing L, "Model-based image reconstruction for four-dimensional PET," *Med. Phys.* 33, 1288-1298 (2006b).

Li T, Koong A, and Xing L, "Enhanced 4D cone-beam CT with inter-phase motion model," *Med. Phys.* 34, 3688-3697 (2007).

Li X A, and Keall P J, "Point/Counterpoint: Respiratory gating for radiation therapy is not ready for prime time," *Med. Phys.* 34, 867-870 (2007).

Liu H H, Koch N, Starkschall G, Jacobson M, Forster K, Liao Z, Komaki R, and Stevens C W, "Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part II – margin reduction of internal target volume," *Int. J. Radiat. Oncol., Biol., Phys.* 60, 1473-1483 (2004).

Low D A, Nystrom M, Kalinin E, Parikh P, Dempsey J F, Bradley J D, Mutic S, Wahab S H, Islam T, Christensen G, Politte D G and Whiting B R, "A method for the

reconstruction of four-dimensional synchronized CT scans acquired during free breathing," *Med. Phys.* 30, 1254-1263 (2003).

Low D A, Parikh P, Lu W, Dempsey J, Wahab S, Hubenschmidt J, Nystrom M, Handoko M, and Bradley J, "Novel breathing motion model for radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 63, 921–929 (2005).

Lu W, Low D A, Parikh P J, Nystrom M M, El Naqa I M, Wahab S H, Handoko M, Fooshee D, and Bradley J D, "Comparison of spirometry and abdominal height as four-dimensional computed tomography metrics in lung," *Med. Phys.* 32, 2351-2357 (2005a).

Lu W, Parikh P J, Hubenschmidt J P, Politte D G, Whiting B R, Bradley J D, Mutic S, and Low D A, "Reduction of motion blurring artefacts using respiratory gated CT in sinogram space: A quantitative evaluation," *Med. Phys.* 32, 3295-3304 (2005b).

Lu W, Parikh P J, Hubenschmidt J P, Bradley J D, and Low D A, "A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT," *Med. Phys.* 33, 2964-2974 (2006).

Mageras G S, Pevsner A, Yorke E D, Rosenzweig K E, Ford E C, Hertanto A, Larson A M, Lovelock D M, Erdi Y E, Nehmeh S A, Humm J L, and Ling C C, "Measurement of lung tumour motion using respiration-correlated CT," *Int. J. Radiat. Oncol., Biol., Phys.* 60, 933-941 (2004).

Mah D, Hanley J, Rosenzweig K E, Yorke E, Braban L, Ling C C, Leibel S A, and Mageras G S, "Technical aspects of the deep inspiration breath-hold technique in the treatment of thoracic cancer," *Int. J. Radiat. Oncol., Biol., Phys.* 48, 1175-1185 (2000).

Mattes D, Haynor D R, Vesselle H, Lewellen T K, and Eubank W, "PET-CT Image registration in the chest using free-form deformations," *IEEE Trans. Med. Imaging* 22, 120-128 (2003).

McClelland J R, Chandler A G, Blackall J M, Ahmad S, Landau S, Hawkes D J, "Feasibility study of a novel technique for constructing respiratory motion models, for use in 4D lung cancer radiotherapy planning." *Proceedings of Medical Image Understanding and Analysis (MIUA) 2004*, 220-223 (2004).

McClelland J R, Chandler A G, Blackall J M, Ahmad S, Landau D, Hawkes D J, "4D motion models over the respiratory cycle for use in lung cancer radiotherapy planning." *Proceedings of SPIE Medical Imaging 2005: Visualization, Image-Guided Procedures, and Display*, 5744, 173-183 (2005a).

McClelland J R, Blackall J M, Hughes S, Ahmad S, Landau D, Hawkes D J, "4D models of respiratory motion for use in RT planning and their advantages over 4DCT techniques," *Radiother. Oncol.* 76 S2, S93 (2005b).

McClelland J R, Chandler A G, Blackall J M, Tarte S, Hughes S, Ahmad S, Landau D, Hawkes D J, "A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy," *Med. Phys.* 33, 3348-3358 (2006a).

McClelland J R, Hughes S, Ahmad S, Chandler A G, Blackall J M, Tarte S, Landau D, Hawkes D J, "Comparing breath-hold and free breathing CT derived motion models for predicting respiratory motion of lung cancer and lymph nodes," *Radiother. Oncol.* 81 S1, S213 (2006b).

McClelland J R, Webb S, Hawkes D J, "Tracking 2D 'differential organ motion' using the DMLC technique and 4DCT motion models for the delivery of IMRT to sites affected by respiratory motion," *Proceedings of the 15$^{th}$ International Conference on the use of Computers in Radiation therapy (ICCR) 2007*, 123-127 (2007a).

McClelland J R, Webb S, McQuaid D, Binnie D M, and Hawkes D J, "Tracking 'differential organ motion' with a 'breathing' multileaf collimator: magnitude of problem assessed using 4D CT data and a motion-compensation strategy," *Phys. Med. Biol.* 52, 4805-4826 (2007b).

McClelland J R, Blackall J M, Tarte S, Hughes S, Hawkes D J, "Non-rigid registration based respiratory motion models of the lung using two parameters," *Med. Phys.* 34, 2516 (2007c).

McQuaid D and Webb S, "IMRT delivery to a moving target by dynamic MLC tracking - delivery for targets moving in two dimensions in the beam's eye view," *Phys. Med. Biol.* 51, 4819-4839 (2006).

Miquel M, Blackall J M, Uribe S, Leussler C, Schaeffter T, and Hawkes D J, "Dynamic 3D lung MRI using a 32 channel coil array for the construction of respiratory motion models", *Proceedings of ISMRM 2007*, Berlin, p. 586 (2007).

Moore C J and Graham P A, "3D dynamic body surface sensing and CT-body matching: a tool for patient set-up and monitoring in radiotherapy," *Computer Aided Surgery* 5, 234-245 (2000).

Mori S, Kanematsu N, Mizuno H, Sunaoka M, and Endo M, "Physical evaluation of CT scan methods for radiation therapy planning: comparison of fast, slow and gating scan using the 256-detector row CT scanner," *Phys. Med. Biol.* 51, 587-600 (2006).

Mori S, Endo M, Komatsu S, Yashiro T, Kandatsu S, Baba M, "Four-dimensional measurement of lung tumour displacement using 256-multi-slice CT-scanner," *Lung Cancer* 56, 59-67 (2007).

Murphy M J, "Tracking moving organs in real time," *Semin. Radiat. Oncol.* 14, 91-100 (2004).

Nehmeh S A, Erdi Y E, Rosenzweig K E, Schoder H, Larson S M, Squire O D, and Humm J L, "Reduction of respiratory motion artefacts in PET imaging of lung cancer by respiratory correlated dynamic PET: methodology and comparison with respiratory gated PET," *The Journal of Nuclear Medicine* 44, 1644-1648 (2003).

Neicu T, Shirato H, Seppenwoolde Y, and Jiang S B, "Synchronised moving aperture radiation therapy (SMART): average tumour trajectory for lung patients," *Phys. Med. Biol.* 48, 587-598 (2003).

Neicu T, Berbeco R, Wolfgang J, and Jiang S B, "Synchronized moving aperture radiation therapy (SMART): improvement of breathing pattern reproducibility using respiratory coaching," *Phys. Med. Biol.* 51, 617-636 (2006).

Ozhasoglu C and Murphy M j, "Issues in respiratory motion compensation during external-beam radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 52, 1389-1399 (2002).

Paganetti H, Jiang H, Adams J A, Chen G T, and Rietzel E, "Monte Carlo simulations with time-dependent geometries to investigate effects of organ motion with high temporal resolution," *Int. J. Radiat. Oncol., Biol., Phys.* 60, 942-950 (2004).

Pan T, Lee T Y, Rietzel E, and Chen G T Y, "4D-CT imaging of a volume influenced by respiratory motion on a multi-slice CT" *Med. Phys.* 31, 333-340 (2004).

Pan T, "Comparison of helical and cine acquisitions for 4D-CT imaging with multislice CT," *Med. Phys.* 32, 627-634 (2005).

Papiez L, Rangaraj D, and Keall P J, "Real-time DMLC IMRT delivery for mobile and deforming targets," *Med. Phys.* 32, 3037-3048 (2005).

Parker D L, "Optimal short scan convolution reconstruction for fanbeam CT," *Med. Phys.* 9, 254–257 (1982).

Parkin D M, Bray F, Ferlay J, and Pisani P, "Global Cancer Statistics, 2002," *Cancer Journal Clinicians* 55, 74-108 (2005).

Pevsner A, Davis B, Joshi S, Hertanto A, Mechalakos J, Yorke E, Rosenzweig K, Nehmeh S, Erdi Y E, Humm J L, Larson S, Ling C C, and Mageras G S, "Evaluation of an automated deformable image matching method for quantifying lung motion in respiration-correlated CT images," *Med. Phys.* 33, 369-376 (2006).

Plathow C, Zimmermann H, Fink C, Umathum R, Schöbinger M, Huber P, Zuna I, Debus J, Schlegel W, Meinzer H P, Semmler W, Kauczor H, and Bock M, "Influence of different breathing maneuvers on internal and external organ motion: use of fiducial markers in dynamic MRI," *Int. J. Radiat. Oncol., Biol., Phys.* 62, 238-245 (2005).

Reinhardt J M, Christensen G E, Hoffman E A, Ding K, and Cao K, "Registration-derived estimates of local lung expansion as surrogates for regional ventilation," *Proceedings of IPMI 2007*, 763-774 (2007).

Remmert G, Biederer J, Lohberger F, Fabel M, and Hartmann G H, "Four-dimensional magnetic resonance imaging for the determination of tumour movement and its evaluation using dynamic porcine lung phantom," *Phys. Med. Biol.* 52, N401-N415 (2007).

Rholfing T, Maurer C R Jr, O'Dell W G, and Zhong J, "Modeling liver motion and deformation during the respiratory cycle using intensity-based free-form registration of gated MR images," *Med. Phys.* 31, 427-432 (2004).

Rietzel E, Pan T, and Chen G T Y, "Four-dimensional computed tomography: Image formation and clinical protocol," *Med. Phys.* 32, 874-889 (2005a).

Rietzel E, Chen G T Y, Choi N C, and Willet C G, "Four-dimensional image-based treatment planning: Target volume segmentation and dose calculation in the presence of respiratory motion," *Int. J. Radiat. Oncol., Biol, Phys.* 61, 1535-1550 (2005b).

Rietzel E and Chen G T Y, "Improving retrospective sorting of 4D computed tomography data," *Med. Phys.* 33, 377-379 (2006).

Rosu M, Chetty I J, Balter J M, Kessler M L, McShan D L, and Ten Haken R K, "Dose reconstruction in deforming lung anatomy: Dose grid size effects and clinical implications," *Med. Phys.* 32, 2487-2495 (2005).

Rueckert D, Sonoda L I, Hayes C, Hill D L, Leach M O, and Hawkes D J, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Trans. Med. Imaging* 18, 712–721 (1999).

Sarrut D, Boldea V, Miguet S, and Ginestet C, "Simulation of four-dimensional CT images from deformable registration between inhale and exhale breath-hold CT scans," *Med. Phys.* 33, 605-617 (2006).

Schnabel J A, Rueckert D, Quist M, Blackall J M, Castellano-Smith A D, Hartkens T, Penney G P, Hall W A, Liu H, Truwit C L, Gerritsen F A, Hill D L, and Hawkes D J, "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," *Proceeding of MICCAI 2001*, 573-581 (2001).

Schreibmann E, Chen G T Y, and Xing L, "Image Interpolation in 4D CT using a bspline deformable registration model," *Int. J. Radiat. Oncol., Biol., Phys.* 64, 1537-1550 (2006).

Schweikard A, Glosser G, Bodduluri M, Murphy M, and Adler J R, "Robotic Motion Compensation for Respiratory Movement during Radiosurgery," *Comput. Aided Surg.* 5, 263–277 (2000).

Schweikard A, Shiomi H, Fisseler J, Dötter M, Berlinger K, Gehl H B, and Adler J R, "Fiducial-Less Respiration Tracking in Radiosurgery," *Proceedings of MICCAI 2004*, 992-999 (2004).

Seppenwoolde Y, Shirato H, Kitamura K, Shimizu S, van Herk M, Lebesque J V, and Miyasaka K, "Precise and real-time measurement of 3Dtumor motion in lung due to breathing and heartbeat, measured during radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 53, 822–834 (2002).

Seppenwoolde Y, Berbeco R I, Nishioka S, Shirato H, and Heijmen B, "Accuracy of tumour motion compensation algorithm from a robotic respiratory tracking system: A simulation study," *Med. Phys.* 34, 2774-2784 (2007).

Sharp G C, Jiang S B, Shimizu S, and Shirato H, "Prediction of respiratory tumour motion for real-time image-guided radiotherapy," *Phys. Med. Biol.* 49, 425-440 (2004).

Sharp G C, Kandasamy N, Singh H, and Folkert M, "GPU-based streaming architectures for fast cone-beam CT image reconstruction and demons deformable registration," *Phys. Med. Biol.* 52, 5771-5783 (2007).

Shekhar R, Lei P, Castro-Pareja C R, Plishker W L, and D'Souza W D, "Automatic segmentation of phase-correlated CT scans through nonrigid image registration using geometrically regularized free-form deformation," *Med. Phys.* 34, 3054-3066 (2007).

Shimizu S, Shirato H, Kagei K, Nishioka T, Bo X, Dosaka-Akita H, Hashimoto S, Aoyama H, Tsuchiya K, and Miyasaka K, "Impact of respiratory movement on the computed tomographic images of small lung tumors in three-dimensional (3D) radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 46, 1127–1133 (2000).

Shirato H, Shimizu S, Kunieda T, Kitamura K, van Herk M, Kagei K, Nishioka T, Hashimoto S, Fujita K, Aoyama H, Tsuchiya K, Kudo K, and Miyasaka K, "Physical aspects of a real-time tumor-tracking system for gated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 48, 1187–1195 (2000).

Shirato H, Harada T, Harabayashi T, Hida K, Endo H, Kitamura K, Onimaru R, Yamazaki K, Kurauchi N, Shimizu T, Shinohara N, Matsushita M, Dosaka-Akita H, and Miyasaka K, "Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumour tracking in radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 56, 240-247 (2003).

Sonke J J, Zijp L, Remeijer P, and van Herk M, "Respiratory correlated cone beam CT," *Med. Phys.* 32, 1176 –1186 (2005).

Sonke J J, Lebesque J, and van Herk M, "Variability of four-dimensional computed tomography patient models," *Int. J. Radiat. Oncol., Biol., Phys.* 70, 590-598 (2008).

Stein J, Bortfeld T, D¨orschel B and Schlegel W, "Dynamic x-ray compensation for conformal radiotherapy by means of multileaf collimation," *Radiother. Oncol.* 32, 163-173 (1994).

Stevens C W, Munden R F, Forster K M, Kelly J F, Liao Z, Starkschall G, Tucker S, and Komaki R, "Respiratory-driven lung tumour motion is independent of tumour size, tumour location, and pulmonary function," *Int. J. Radiat. Oncol., Biol., Phys.* 51, 62-68 (2001).

Sundaram T A, Avants B B, and Gee J C, "A Dynamic Model of Average Lung Deformation Using Capacity-Based Reparameterization and Shape Averaging of Lung MR Images," *Proceedings of MICCAI 2004*, 1000-1007 (2004).

Sundaram T A, and Gee J C, "Towards a model of lung biomechanics: pulmonary kinematics via registration of serial lung images," *Med. Image Anal.* 9, 524-537 (2005).

Sundaram Cook T, Tustison N, Biederer J, Tetzlaff R, and Gee J, "How do registration parameters affect quantitation of lung kinematics?" *Proceedings of MICCAI 2007*, 817-824 (2007).

Tarte S, McClelland J R, Hughes S, Blackall J M, Landau D, Hawkes D J, "A non-contact method for the acquisition of breathing signals that enable distinction between abdominal and thoracic breathing," *Radiother. Oncol.* 81 S1, S209 (2006).

Tsunashima Y, Sakae T, Shioyama Y, Kagel K, Terunuma T, Nohtomi A, and Akine Y, "Correlation between the respiratory waveform measured using a respiratory sensor and 3d tumour motion in gated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* 60, 951-958 (2004).

Underberg R W M, Lagerwaard F J, Cuijpers J P, Slotman B J, van Sörnsen de Koste J R, and Senan S, "Four-dimensional CT scans for treatment planning in stereotactic radiotherapy for stage I lung cancer," *Int. J. Radiat. Oncol., Biol., Phys.* 60, 1283-1290 (2004).

Webb S, *Intensity modulated radiation therapy*, Bristol: IOP Publishing (2000).

Webb S, "Does elastic tissue intrafraction motion with density changes forbid motion-compensated radiotherapy?" *Phys. Med. Biol.* 51, 1449-1462 (2006).

Webb S and Binnie D M, "A strategy to minimise errors from differential intrafraction organ motion using a single configuration for a "breathing" multileaf collimator," *Phys. Med. Biol.* 51, 4517-4531 (2006).

Webb S, Binnie D, McClelland J R, "General formulation of the problem of 'differential organ motion', magnitude of problem assessed using 4D CT data and motion-compensation strategy suitable for a 'breathing' multileaf collimator," *Proceedings of the 15th International Conference on the use of Computers in Radiation therapy (ICCR) 2007*, 118-122 (2007).

Weiss E, Wijesooriya K, Ramakrishnan V, and Keall P J, "Comparison of intensity-modulated radiotherapy planning based on manual and automatically generated contours using deformable image registration in four-dimensional computed tomography of lung cancer patients," *Int. J. Radiat. Oncol., Biol., Phys.*70, 572-581 (2008).

Werner-Wasik M, Scott C, Curran W J, and Byhardt R, "Correlation between acute esophagitis and late pneumonitis in patients (pts) with locally advanced non-small cell lung cancer (LA-NSCLC) receiving concurrent thoracic radiotherapy (RT) and chemotherapy: a multivariate analysis of the Radiation Therapy Oncology Group (RTOG) database," *Pro. Am. Soc. Clin. Oncol.* 21, abstr. 1192(2002).

West J B, Maurer C R, and Dooley J R, "Hybrid point-and-intensity-based deformable registration for abdominal CT images," *Proceedings of SPIE Medical Imaging 2005: Image Processing*, 204-211 (2005).

van Herk M, "Different styles of image-guided radiotherapy," *Semin. Radiat. Oncol.* 17, 258-267 (2007).

van Sörnsen de Koste J R, Lagerwaard F J, Nijssen-Visser M R J, Graveland W J, and Senan S, "Tumour location cannot predict the mobility of lung tumours: a 3d analysis of data generated from multiple ct scans," *Int. J. Radiat. Oncol., Biol., Phys.* 56, 348-354 (2003).

von Siebenthal M, Székely G, Gamper U, Boesiger P, Lomax A, and Cattin Ph, "4D MR imaging of respiratory organ motion and its variability," *Phys. Med. Biol.* 52, 1547-1564 (2007).

Vedam S S, Keall P J, Kini V R, and Mohan R, "Determining parameters for respiration-gated radiotherapy," *Med. Phys.* 28, 2139-2146 (2001).

Vedam S S, Keall P J, Kini V R, Mostafavi H, Shukla H P, and Mohan R, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal," *Phys. Med. Biol.* 48, 45–62 (2003).

Vedam S S, Keall P J, Docef A, Todor D A, Kini V R, and Mohan R "Predicting respiratory motion for four-dimensional radiotherapy," *Med. Phys.* 31, 2274-2283 (2004).

Xu Q, Hamilton R J, "A novel respiratory detection method based on automated analysis of ultrasound diaphragm video," *Med. Phys.* 33, 916-921 (2006).

Zeng R, Fessler J A, and Balter J M, "Estimating 3-D respiratory motion from orbiting views by tomographic image registration," *IEEE Trans. Med. Imaging* 26, 153-163 (2007a).

Zeng R, Fessler J A, and Balter J M, "A simplified motion model for estimating respiratory motion from orbiting views," ," *Proceedings of SPIE Medical Imaging 2007: Image Processing*, 651240 (2007).

Zhang T, Keller H, O'Brien M J, Mackie T R, and Paliwal B, "Application of the spirometer in respiratory gated radiotherapy," *Med. Phys.* 30, 3165-3171 (2003).

Zhang Q, Pevsner A, Hertanto A, Hu Y-C, Rosenzweig K E, Ling C C, and Mageras G S, "A patient-specific respiratory model of anatomical motion for radiation treatment planning," *Med. Phys.*34, 4772-4781 (2007).

# Appendix A

# Estimating the inverse of a B-spline transformation

The non-rigid registrations have been performed in the direction that allows the motion models to predict the Cine CT volumes. This means the transformations actually map from the predicted volumes to the reference volume, as is required for image formation. The transformations effectively tell us where any point in the predicted volume 'comes from' in the reference volume, they do not tell us where any point in the reference volume 'goes to' in the predicted volume. However, there are a number of times when it is desirable to know where a specific point defined in the reference volume has moved to in the predicted volume(s), e.g. when assessing the registrations and models, or planning tracked treatment. As the exact inverse of a B-spline transformation (and most other non-rigid transformations) cannot be calculated exactly it is necessary to estimate the inverse of the predicted transformation.

The following algorithm was used to estimate the inverse of a non-rigid transformation at a particular point. This algorithm has only been applied to the B-spline non-rigid transformation in this work, but it could be used for any smooth and continuous non-

rigid transformation. The algorithm is iterative and will continue until the estimated inverse has been found to the desired accuracy or the maximum number of iterations has been reached. The algorithm needs to be provided with a starting estimate of the inverse. This can be the centre of the region the transformation is defined over if a better starting estimate is not available.

Inputs to algorithm: $\mathbf{r}$ - the point in the reference volume where the inverse is required, $\mathbf{T}$ - the non-rigid transformation, $\mathbf{p}_{est}$ - the initial estimate of point $\mathbf{r}$ transformed by the inverse of $\mathbf{T}$, i.e. $\mathbf{p}_{est}$ is an estimate of $\mathbf{p} = \mathbf{T}^{-1}(\mathbf{r})$

Algorithm parameters: *accuracy* - the desired level of accuracy (set to 0.01 mm), *max_iter* - the maximum number of iterations permitted (set to 100).

- $iter = 0$
- $\mathbf{r}_{est} = \mathbf{T}(\mathbf{p}_{est})$
  (transform the initial estimate by the forward transform),
- while $\mathbf{r} - \mathbf{r}_{est} > accuracy$ AND $iter < max\_iter$
  (while the estimate is not within the desired accuracy and the maximum number of iterations has not been reached),
  - $\mathbf{p}_{new} = \mathbf{p}_{est} + (\mathbf{J}(\mathbf{p}_{est}))^{-1}(\mathbf{r} - \mathbf{r}_{est})$
    (find the values of the Jacobian matrix, $\mathbf{J}$, of the transformation $\mathbf{T}$, at point $\mathbf{p}_{est}$, and use this to find the new estimate of $\mathbf{p}$, $\mathbf{p}_{new}$)
  - $\mathbf{p}_{est} = \mathbf{p}_{new}$, $\mathbf{r}_{est} = \mathbf{T}(\mathbf{p}_{est})$, $iter = iter + 1$
    (update estimates and iteration counter)

It was found that generally very few iterations (<10) were required, even when a high accuracy (0.01 mm) was specified. If the transformation is not continuous (e.g. because it contains folding) then the inverse may not be defined for some points, and the maximum number of iterations will be reached. Also, as the B-spline transformations are only defined over a finite region, it is possible that the inverse for the desired point does not lie within this region. Care had to be taken when implementing the algorithm that points outside of the transformation, and points inside but near the edge of the transformation, were handled correctly.

# Contents of accompanying CD

Movie_comparing_volumes_coronal_Patient_2( to 8 ).avi,
Movie_comparing_volumes_sagittal_Patient_2( to 8 ).avi

Animations of coronal and sagittal slices over a complete respiratory cycle from patients 2-8, comparing the 4DCT volumes, the individual model predictions concatenated together, and the prediction from the model results combined into a single continuous transformation - see Chapter 5, page 123

fig6_1a( to f ).avi, fig6_3a( to f ).avi, fig6_4a( to f ).avi

Animations of 3D plots from rotating views so that 3D data distribution and surfaces can be appreciated - see Chapter 6, page 127.

imrt_anim.avi

Animation of implemented IMRT tracking strategy. The dashed lines show the planned leaf positions, the solid lines show the leaves modified to track the differential motion - see Chapter 7, page 166.