POLITECNICO DI MILANO

Scuola di Ingegneria Industriale e dell'Informazione Corso di Laurea Magistrale in Ingegneria Biomedica

EVALUATION OF TOOLS USED FOR 4D CBCT MOTION MODELLING

Relatore: Prof. Guido Baroni
Correlatore: Prof. Chiara Paganelli

Tesi di Laurea Magistrale di:
Omar Martínez Monsiváis matr. 862464

Anno accademico 2017/2018

Index

Summary

One of the most effective strategies to control and eradicate tumoral tissues within the human body is radiation therapy. This treatment has the aim to deliver the maximum dose to the target while spearing the surrounding healthy tissues. A way to overcome this issue is the procedure known as fractionation, in which the dose is delivered in different days to allow recovery of healthy tissues. Safety margins are also clinically defined in order to accurately hit the target while spearing the healthy anatomy.

A frontier in radiotherapy has been introduced by its combination with image modalities (Image Guided Radiotherapy, IGRT), whose aim is to guide the treatment by reducing treatment margins, allowing to perform safe dose escalation and improving patient treatment. The IGRT workflow consists of two phases:

- Planning: A Computed Tomography (CT) image of the area of interest is acquired which allows to define safety margins and to optimize the dose to deliver.
- Irradiation: Just prior treatment, the patient setup is performed, and imaging data are typically acquired to evaluate potential deviations from the planning CT (i.e. Cone-beam CT acquisition (CBCT)); then the plan is delivered.

One issue to deal with is organ motion, which can be divided into two classes: inter-fraction, i.e., between radiation delivery sessions; and intra-fraction movement, which occurs during the ongoing session.
In order to accurately treat the target and spearing surround tissues, organ motion needs to be carefully quantified and compensated for.
As a consequence, several motion management techniques have been developed both in treatment planning and treatment delivery. For what concerns treatment planning, breath-hold or gating techniques (i.e. acquiring the patient images during a breath-hold or during a specific respiratory phase) have been considered. Currently, the clinical standard relies on 4DCT images which allows to acquire data in free-breathing and to retrospectively sort them in order to describe a representative respiratory cycle. This can be also extended on CBCT data, in which 4D CBCT reconstruction can be performed to account for motion just prior treatment. As regards treatment, breath-hold, gating or tumor tracking schemes can be

adopted as to replicate the planning scenario. Tumor tracking (in which the patient breath freely and the tumor is tracked) represents the clinical goal, since it is expected to improve treatment performance and efficiency. Some of the clinically implemented solutions for tumor tracking, rely on the construction of a local internal-external correlation motion model, which allows to estimate the position of the target by means of a correlation with an external surrogate.

Motion modeling in IGRT typically relies on the use image registration, which consists in the integration of different images. Image registration it is divided into rigid registration which can perform three translation and three rotation operations; and Deformable Image Registration (DIR) which can compensate for tissue deformation due to organ motion, and provides as outputs a registered volume and a deformation field. The main problem with DIR is validation; there are no standard approaches to assess its results. Nonetheless, DIR algorithms are fundamental in IGRT.

Thanks to the advantages provided by DIR in accounting for organ motion, studies in the literature proposed the development of global motion models. Differently from the clinically implemented local motion models, they allow to estimate the whole anatomy, rather than the tumor position only. They typically rely on establishing a correlation between the deformation field derived by DIR between different respiratory phases of a 4DCT and an external surrogate. However, the difficulties in the validation of these approaches (due to dependency on DIR, capability to account for irregular motion, etc) limit their introduction in the clinical practice.

The aim of this project consists in applying DIR and global motion models to 4DCBCT data, in order to extend the use of these strategies to imaging data acquired just before treatment, conversely to 4DCT which are acquired in the planning phase. The validation of these strategies is also provided by means of a local and global assessment, relying on markers implanted in the tumor and on an analysis on the whole image.

The first step to perform these assessments was to execute a DIR between the 4D CBCT images: all the respiratory phases were registered on the reference exhale phase.

The error measurement for the DIR was divided into local and global errors.

- The local DIR error was measured as the Euclidean distance between the markers on the exhale and the DIR-output volumes, and it was divided into three sections:
    - Isolation of tumoral volumes, done by performing a multiplication between the volumes and the tumor contour.
    - Marker Segmentation: performed to isolate the markers implanted in the tumor. It consisted on setting a threshold value, which meant that everything above this threshold was extracted from the tumoral volume.
    - Local Error Measurement: the distance between the segmented markers was computed. This is done by computing the centroids of the marker regions, then the distance between these centroids was measured, which was interpreted as the target registration error.
- Regarding the Global Registration Error, the measurement of this metric was divided in two measurements, both were performed by subtracting all the voxel values of the involved volumes.
    - Global difference between volumes before registering them
    - Global difference between volumes after registering them

The next error measurement corresponded to the motion model error: the model used throughout this work was developed by Vandemeulebroucke et al. in 2009. This model was directly applied to the 4DCBCT dataset and was used to estimate one of the respiratory volumes in a leave-on-out approach. The inputs of this model are described as follows:

- Reference: A volume used as the reference for motion estimation.
- Phase: A text file that describes the value of the phase of the cycle.
- Amplitude: A text file is also used to describes the value of the amplitude of the phase of the cycle.
- DIR VFs: The deformable vector fields obtained as outputs from the DIR.

The error measurements for the motion model were also divided into local and global measures:

- Local Estimation Error: the distance between markers was measured between the estimated and the ground truth respiratory phase. The process was the same as for the DIR local error measurement.
- Global Estimation Error: identical process like the one outlined for the DIR global error. This global measurement was also divided into two steps:
    - Measuring the global difference between the ground truth volumes and the DIR volumes.
    - Measuring the global difference between estimated volumes and ground truth volumes

The observations regarding the obtained results are as follows:

- Locally, lower errors were obtained for the phases that were closer to the reference phase. It was proven that the volume estimation is highly influenced by DIR error. Also, it is worth highlighting that the model was able capable to compensate for motion, with a median error of 1.29 mm. It should be noted however that higher errors could also be attributed to low performance in marker segmentation.
- Globally, the results pointed out that the estimation differences (with a median of 1.81%) were lower than the DIR differences (median 1.87%), which in turn were also lower than the before DIR measurements (median of 2.04%). Also, the global error was again higher for the phases that were further from the reference phase. Additionally, it was noted that the DIR error was influencing the estimation error. So for this metric, the image registration was also critical for obtaining accurate outputs from the motion model.

Future developments would include more extensive research on DIR algorithms and an optimized marker segmentation strategy. Both of which seemed to have an impact on the accuracy of the obtained results. Additionally, the application of the global motion model to data of different treatment sessions will be performed, as well as a comparison with a conventional application on 4DCT data.

Sommario

Una delle strategie più efficaci per gestire ed eliminare i tessuti tumorali all'interno del corpo umano è la radioterapia. Questa strategia ha l'obiettivo di somministrare la dose massima al tessuto bersaglio risparmiando però il più possibile i tessuti sani circostanti. Un approccio per superare questo problema è la procedura nota come frazionamento della dose, in cui la dose viene somministrata in diversi giorni con l'obiettivo di consentire il recupero dei tessuti sani. In aggiunta, margini di sicurezza sono definiti clinicamente con lo scopo di colpire il bersaglio con l'accuratezza richiesta, mentre allo stesso tempo si cerca di risparmiare l'anatomia sana.

Una frontiera in radioterapia è stata introdotta dalla sua combinazione con le diverse modalità di immagini (Image Guided Radiotherapy, IGRT), in modo da diminuire i margini della terapia, consentendo così di aumentare la dose somministrata e contemporaneamente migliorare la cura della malattia. Il flusso di lavoro IGRT è costituito da due fasi:

- Pianificazione: Si acquisisce una Tomografia Computerizzata (CT) dell'area di interesse che consente di definire i margini di sicurezza e ottimizzare la dose da erogare.
- Irradiazione: Precedente alla terapia si esegue la correzione della posizione del paziente e vengono acquisite immagini (ad esempio, acquisizione cone-beam CT, CBCT) per valutare potenziali deviazioni dalla CT di pianificazione; dopodiché la terapia viene erogata in conformità alla pianificazione.

Una problema da affrontare è il movimento degli organi, che può essere diviso in due classi: inter-frazione, cioè tra sessioni di erogazione di radiazioni; e movimento intra-frazione, che si verifica durante la sessione in corso. Al fine di trattare con accuratezza il bersaglio e prendere cura dei tessuti circostanti, il movimento degli organi deve essere accuratamente quantificato e compensato. Di conseguenza, sono state sviluppate diverse tecniche di gestione del movimento sia nella pianificazione della terapia che nella sua erogazione. Per quanto riguarda la pianificazione, sono state prese in considerazione le tecniche di breath-hold o gating (ovvero l'acquisizione delle immagini del paziente durante una pausa nel respiro o durante una specifica fase respiratoria). Attualmente, lo standard clinico si basa su

immagini 4DCT che consentono di acquisire dati durante respirazione libera e di ordinarli retrospettivamente per descrivere un ciclo respiratorio rappresentativo. Questo può essere esteso anche sui dati CBCT, in cui è possibile eseguire la ricostruzione 4D CBCT per tenere conto del movimento appena prima del trattamento. Per quanto riguarda il trattamento, possono essere adottati schemi di breath-hold, gating o inseguimento (tracking) del tumore per replicare lo scenario di pianificazione. Il tracking del tumore (in cui il paziente respira liberamente e il tumore è inseguito) rappresenta l'obiettivo clinico, poiché si prevede che questo approccio migliorerà le prestazioni e l'efficienza della terapia. Alcune delle soluzioni clinicamente implementate per il tracking del tumore, si basano sulla costruzione di un modello di movimento sulla base di correlazione interna ed esterna locale, che consente di stimare la posizione del bersaglio tramite una correlazione con un surrogato esterno.

La modellazione del movimento nell' IGRT si basa sulla registrazione dell'immagine, che consiste nell'integrazione di immagini diverse. La registrazione delle immagini è suddivisa in registrazione rigida, che può eseguire tre operazioni di traslazione e tre di rotazione; e registrazione deformabili delle immagini (DIR), che riesce a compensare la deformazione del tessuto dovuta all'organ motion e fornisce come uscite un volume registrato e un campo di deformazione. Il problema principale con DIR è la sua validazione; non ci sono approcci standard per valutare i suoi risultati. Nonostante ciò, gli algoritmi DIR sono fondamentali nell'IGRT.

Grazie ai vantaggi forniti dalla DIR intorno al movimento degli organi, gli studi in letteratura hanno proposto lo sviluppo di modelli di movimento globali. A differenza dei modelli di movimento locale implementati clinicamente, quelli globali consentono di stimare l'intera anatomia del paziente, piuttosto che la sola posizione del tumore. Di solito si basano sulla correlazione tra il campo di deformazione derivato dalla DIR tra diverse fasi respiratorie di una 4DCT, e un surrogato esterno. Tuttavia, le difficoltà nella validazione di questi approcci (a causa della dipendenza alla DIR, la capacità di spiegare il movimento irregolare, ecc.) limitano la loro introduzione nella pratica clinica.

Lo scopo di questo lavoro consiste nell'applicare DIR e modelli di movimento globale a dati 4DCBCT, al fine di estendere l'uso di queste strategie immagini acquisite appena prima del trattamento, al contrario dei dati 4DCT che vengono acquisiti nella fase di

pianificazione. La validazione di queste strategie è anche fornita mediante una valutazione locale e globale, basata su marcatori impiantati nel tumore e su un'analisi dell'intera immagine.

Il primo passo per eseguire queste valutazioni consiste nell'eseguire una DIR tra immagini 4D CBCT: tutte le fasi respiratorie sono state registrate sulla fase di espirazione di riferimento.

La misura dell'errore per la DIR è stata suddivisa in errore locale e globale.

- L'errore DIR locale è stato misurato come la distanza euclidea tra i marcatori sui volumi di espirazione e quelli ottenuti dalla DIR, il percorso per ottenere la tale valutazione è consistito in:

    - Isolamento dei volumi tumorali, ottenuto eseguendo una moltiplicazione tra i volumi e il contorno tumorale.
    - Segmentazione dei Marker: eseguita per isolare i marker impiantati nel tumore. Si è impostato un valore di soglia che ha consentito l'identificazione dei marker.
    - Misura dell' Errore Locale: la distanza tra i marker segmentati è stata calcolata. I centroidi dei marker sono stati calcolati e la distanza tra questi centroidi misurata, che è stata interpretata come errore di registrazione nel target.
- Per quanto riguarda l'errore di registrazione globale, la quantificazione è stata suddivisa in due parti, entrambe eseguite sottraendo tutti i valori dei voxel dei volumi coinvolti, e in particolare si cono calcolate:
    - Differenza globale tra i volumi prima di essere registrati.
    - Differenza globale tra i volumi dopo la loro registrazione

Successivamente è stato quantificato l'errore del modello di movimento globale: il modello utilizzato in questo lavoro è stato sviluppato da Vandemeulebroucke et al. nel 2009. Questo modello è stato applicato direttamente ai dati 4DCBCT ed è stato utilizzato per stimare dei volumi respiratori. Gli input di questo modello sono:

- Riferimento: volume utilizzato come riferimento per la stima del movimento.
- Fase: un file di testo che descrive il valore della fase del ciclo da stimare.
- Ampiezza: un file di testo viene anche utilizzato per descrivere il valore dell'ampiezza della fase del ciclo da stimare.
- VF DIR: i campi di deformatizione ottenuti come output dalla DIR.

Le misura dell'errore per il modello di movimento è stata divisa anche in misure locali e globali:

- Errore di stima locale: la distanza tra i marker è stata misurata tra la stima e quella della specifica fase respiratoria (ground truth). Il processo è stato lo stesso rispetto alla misura degli errori DIR locali.
- Errore di stima globale: processo identico a quello delineato per l'errore DIR globale. Questa misura globale è stata anche divisa in due fasi:
    - Misura della differenza globale tra i volumi ground truth e quelli ottenuti dalla DIR.
    - Misura della differenza globale tra volumi stimati e volumi ground truth.

Le osservazioni riguardanti i risultati ottenuti sono le seguenti:

- A livello locale, sono stati ottenuti errori più bassi per le fasi più vicine alla fase di riferimento. Si è inoltre notato che la stima del volume è fortemente influenzata dall'errore della DIR. Inoltre, è stato osservato che il modello era in grado di compensare i movimenti, fornendo un errore con una mediana di 1.29 mm. Errori più elevati possono essere attribuiti alla bassa performance nella segmentazione dei marker.
- Globalmente, i risultati hanno evidenziato che gli errori di stima del modello (con una mediana di 1.81%) sono inferiori alle quelli della DIR (mediana pari a 1.87%), che a loro volta sono anche inferiori agli errori ottenuti prima della DIR (mediana di 2.04%)). Inoltre, l'errore globale si è mostrato di nuovo più alto per le fasi che sono più lontane dalla fase di riferimento. Ulteriormente, è stato notato che l'errore DIR influenza fortemente l'errore di stima del modello. Pertanto, si può dire che la

qualità della DIR è fondamentale per ottenere risultati accurati dal modello di movimento.

I futuri sviluppi includeranno una ricerca più ampia degli algoritmi DIR e una strategia di segmentazione dei marker ottimizzata, i quali hanno mostrato un impatto sull'accuratezza dei risultati ottenuti. Inoltre, verrà eseguita l'applicazione del modello di movimento globale ai dati di diverse sessioni di trattamento, nonché un confronto con un'applicazione convenzionale su dati 4DCT.

Chapter 1. Introduction

Radiation therapy is a treatment strategy aimed to control and, in the best-case scenario, eradicate malignant tissue structures. It works by causing damage to aberrant cells by ionizing radiation. The main problem with this is that high amounts of energy are delivered to the tissue. Since millimetric precision is required during every stage of this procedure, any small deviation of the anatomical geometry with respect to the one that was used a priori to design the treatment, will for sure have some impact over the healthy tissue that is adjacent to the cancerous formations.

1.1) Radiotherapy

Radiotherapy uses high energy x-rays for the treatment of tumors. This high energy dose leads to the production of free radicals, which in turn, damage the genetic code of the tumoral cells, thus making them incapable of reproducing. The outcome of this will be a gradual cease in tumor growth.

Whereas healthy tissues can repair DNA damage reasonably quickly, cancer cells have less efficient recovery mechanisms. This discrepancy between the two types of cell capability to recover is the reason why it is suitable to supply repeated radiation doses at intermediate time intervals between the recovery time of the tumor tissue and the healthy tissue, this is known as dose fractioning. Fractioning enables to have a more effective treatment because while receiving the same absorbed energy, the damage to the tumor tissue is higher than that of healthy tissue (Barnett, 2009).

Although it is a widely used technique, traditional radiotherapy has contraindications linked to the use of ionizing radiation, the most worrisome is the possibility of damaging healthy tissues and creating secondary lesions (Schulz-Ertner, 2006). High radiation doses ensure better local control of the pathology. Nevertheless, it is essential to acknowledge the tolerability of the surrounding healthy tissues (Orecchia, 1998); it is critical to plan the treatment with an extremely accurate execution, guaranteeing the delivery of the dose on the

volume of the tumor and sparing the adjacent healthy tissues. This can be accomplished by the definition of margins of the tumoral volume and the surrounding organs at risk (OAR).

Specifically, three volumes are defined:

- Gross Tumor Volume (GTV): represents the macroscopic tumor volume.
- Clinical Target Volume (CTV): which adds to the GTV any microscopic extensions that are not visible in the diagnostic images. This quantity is called subclinical involvement, and it is a value established by medical experience and clinical studies.
- Planning Target Volume (PTV) obtained by adding two additional safety margins to the CTV: The Internal Margin and the Setup Margin, which take into account the patient's anatomopathological variations, the uncertainties between the image acquisition methods and the position of the patient, and geometric and dosimetric inaccuracies related to the equipment and the irradiation technique (Van Herk, 2004).

Additionally, the use of multiple beams in a different direction, allows a higher geometric composition of the dose distribution over the tumor, releasing high-density energy at the malignant mass (Goldsmith, 2012). It should be noticed that misalignments between the particle beam and the target may cause collateral damage and inadequate control of the disease (Kang, 2015).

1.2) Organ Motion

Organ motion stands for either physiological or pathologic mechanisms that produce ambiguities in patient anatomy when delivering radiotherapy. There are two classes of organ motion: inter-fraction, i.e., between radiation delivery sessions; and intra-fraction movement, which occurs during the ongoing session. Inter-fraction changes involve changes in tumor size, variations in the weight of the patient, and geometric errors and uncertainties; these changes occur in a day to day basis. Intra-fraction displacements happen in minutes or seconds and are usually a consequence of the physiological activity of the patient, which includes gastrointestinal, cardiac, muscular and respiratory responses. The latter is particularly relevant for lung cancer therapies, given that normal breathing motion has a

significant impact on the affected anatomical region (Keall P. , 2004), respiratory displacement is regarded as a non-rigid deformation, and it has been described as: natural, cyclic, irregular and complex mechanism (Korreman, 2012), which translates to several issues during image acquisition (Chen, 2004) and the whole IGRT process (Keall P. J., 2006). That is why this motion is the reason for having large safety margins. Therefore, a critical challenge for lung cancer radiation treatment is organ motion (Goitein, 2004).

Motion might lead to a poor dose delivery of the Gross Tumor Volume (GTV) and an undesired and relevant dose delivery of healthy tissues; these effects can be avoided by reducing ambiguities in the tumor's position: fewer uncertainties lead to a better treatment result (Ling, 2004). Motion management techniques (Langen, 2001) involve several image modalities such as CT, 4DCT, CBCT, among others. It consists of the use of several techniques, as motion limitation (like breath holding, coached breathing), residual motion quantification (intra and inter-fraction), residual motion management, and treatment delivery control. Overcoming motion from an imaged anatomy at several breathing states is similar to how movement from day to day images is approached. Two inherit characteristics of this movement are exploited to address the uncertainties that it creates. First, breathing motion is cyclic, and second, changes in lung tissue density, caused by inspired air, are constant.

Under those considerations, a successful model is expected to take into account this motion and adequately address it.

1.3) Image Guided Radiotherapy, IGRT

The goal of radiotherapy is to achieve excellent local control of a tumoral lesion, sparing as much as possible the surrounding healthy tissues and organs at risk, this means providing a high dose of radiation to the tumor while minimizing the dose to which surrounding tissues are exposed. To accomplish this goal, image modalities are employed to accurately localize the target and the healthy structure during the whole treatment. This is known as image-guided radiotherapy (IGRT). In the case of organ motion, IGRT allows to acquiring images to monitor changes in the shape and position of organs due to respiration during treatment.

The goal of using IGRT is to reduce treatment margins, allowing to perform safe dose escalation and improve patient treatment (Sarrut, 2006).

1.4) Imaging Modalities

1.4.1) Computerized Tomography, CT

CT is an imaging modality that uses X-ray equipment to produce cross-sectional images, i.e., "slices" of the patient being imaged. A movable table carries the patient through a round opening in the system. Then, an X-ray source and a detector within the system rotate around the patient. At this point, the X-ray source generates a thin, fan-like beam of X-rays which goes through a particular section of the patient's body. Detectors in rows placed opposite to the X-ray source register the X-rays that are going through the patient. Several acquisitions (at numerous angles through the patient) are collected during one complete rotation. This data is sent to a dedicated computer to reconstruct from it one or multiple cross-sectional images of the internal anatomy.
CT images present higher clarity and details than traditional X-ray images (U.S. Food and Drug Administration, 2018).

The evolution of CT scanners has enabled the 3D reconstruction of a patient's soft tissue structures with an always increasing resolution for each scan slice. Visualization of digital images on computers allows fusing reconstructed images from CT scans with the rendering of the patient's treatment beams; this enables radiotherapists to adjust the beam to the patient's anatomy with millimetric precision (Nguyen, 2015).

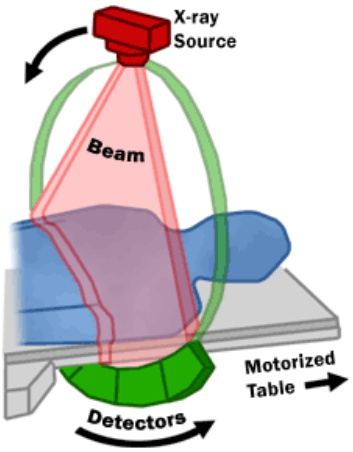

Figure 1. Drawing of CT fan beam and patient in a CT imaging system. U.S. Food and Drug Administration

1.4.2) Cone-Beam Computerized Tomography, CBCT

This modality was developed in the late 90s. A 2D detector system allows the creation of a complete volumetric data set from a unique gantry rotation as opposed to conventional CT which utilizes a one-dimensional detector or series of detectors (Jaffray, 2000).

Consequently, CBCTs are 3D images reconstructed from 2D projections acquired at different angles around a patient using a rotating system composed of an X-ray source and a detector panel (Chan, 2011). Usually, around 650 projections are acquired during a complete 360 gantry rotation; the rotation speed determines the number of images acquired per revolution. The multiple projections acquired at different angles are processed using a Filtered Back-Projection algorithm (Feldkamp, 1984) which provides a reconstructed 3D volume (D'Acunto, 2016).

The use of CBCT improves radiation therapy (Xing, 2006) because these scans are usually performed just before the treatment, to verify and overcome any possible error in patient positioning. To achieve this CBCT scans are rigidly registered with planning CT images. Based on this, the corresponding adjustments are calculated, this allows to have a treatment position of the patient that is aligned with the one from the setup phase (Bombardieri, 2016). CBCT is therefore often present directly in the treatment room, to acquire data immediately before irradiation.

1.4.3) 4D Computerized Tomography, 4DCT

4DCT imaging (Keall P. , 2004) is the acquisition of a series of 3DCT image sets over sequential segments of a breathing period. This type of acquisition is restricted by the velocity of the detectors and gantry, and by the quantity of delivered dose. 4D reconstruction derives from a synchronized respiratory signal, this sync operation relies on an external signal, like surface motion tracking, and a lung volume signal, like spirometer. The resulting 4D dataset is constituted by ten temporal and consecutive 3D reconstructed images.

Patient setup is similar as in a standard 3DCT exam: the subject is immobilized on the scanner bed and aligned using room and scanner lasers. Sagittal and coronal images are used to verify positioning, and the setup is adjusted if necessary.

The incorporation of several devices assesses the patient's position: an infrared source, a camera, and a reflecting block. The latter is fastened to the patient's abdomen; the camera then captures the motion of the block. Next, the movement is analyzed in real-time. The breathing pattern is recorded throughout the scan and is called "respiratory trace." Once the scan is over, the software computes the phase at each point of the respiratory trace by determining the location of the peaks at end-inspiration and assigning percentages to interpeak positions based on the peak-to-peak distance. According to this, the end-inspiration occurs at 0%, whereas end-expiration arises around the 50 – 60%.

Irregularities in a patient's respiratory pattern can be reduced by helping the patient to breathe calmly and consistently. Once a sufficiently regular breathing has been established, the CT data is acquired in "cine" mode. The gantry rotates several times at each bed position to collect the data over the full respiratory cycle. This data is then divided into bins according to a selected time interval, later CT slices are reconstructed from these bins.

The respiratory and scanned data are combined using the respiratory trace to classify the CT slices according to their respiratory phase. Missing phases for any position are replaced with their nearest neighbor (Xing, 2006).

1.4.4) 4D Cone-Beam Computerized Tomography, 4DCBCT

This modality refers to CBCT images resolved in time. 4D CBCT volumes are reconstructed by synchronizing a surrogate respiratory signal with a CBCT system; similarly, as for 4DCT. Projections are therefore retrospectively sorted using the external surrogate which allows the description of inter-fraction variations with respect to the planning CT as well as intra-fraction variations due to the description of the respiratory motion. To improve the 4DCBCT reconstruction, motion compensated techniques exploiting the 4D motion derived by a planning 4DCT can be used (Bombardieri, 2016). By using the Motion Compensated Reconstruction (MCR) method (Rit S., 2009), ten 4D CBCT volumes can be reconstructed, each one associated to a different respiratory phase. The advantage of the MCR technique is the possibility to use all the CBCT projections from the available dataset to reconstruct a single respiratory phase, integrating the information of the movement of the planning 4D CT, without the need of acquiring further projections.

1.5) Radiotherapy Workflow

Radiotherapy treatment consists of two phases: planning and irradiation, in both of them image acquisition plays a relevant role.

1.5.1) Planning

It involves the acquisition of CT images from the anatomical area of interest. The tumor is identified on these images, manually contouring the target volume to be irradiated. In the case of organ motion, 4DCT represents the standard image modality for treatment planning. Planning could also be complemented with Magnetic Resonance Imaging (MRI) or functional imaging such as Positron Emission Tomography (PET).

1.5.2) Irradiation

The irradiation phase consists of patient setup and the dose delivery.

1.5.2.1) Patient setup

This phase is performed at the beginning of each treatment session. It consists in positioning the patient according to the geometric configuration of the planning phase.

First, the initial alignment is carried out using laser systems whose meeting point fixes the isocenter of the treatment room. The lasers are aligned with the fiducial points defined in the planning phase. The positioning of the target can be verified and adjusted through the acquisition of images of the patient's anatomy. Some systems use double radiographic projections, acquired through two pairs of symmetrical source-detectors, to identify the 3D position of the internal target (Kilby, 2010). The location of the internal lesion is then compared to the area defined in the planning phase by comparing the images acquired in the treatment room with the DRR images (Digitally Reconstructed Radiographs), generated by projecting the planning CT volume to a given angle. From this comparison, we obtain the roto-translation parameters to be adopted to correct the positioning with respect to the planning phase. The most common 3D imaging system is that of the Cone-Beam CT. A rigid registration is initially performed between the planning CT image and the CBCT image acquired during setup with the purpose of verifying and correcting the patient's position. Following this, the planning position and the current position are lined up.

Patient alignment can also be achieved using 4D CBCT, as mentioned before, these volumes are produced using datasets acquired during the setup phase, they compensate for the inter-fraction variability between planning and treatment phases and prevent delivering an additional dose to the patient given that there is no further need to acquire more x-ray images (Bombardieri, 2016).

1.5.2.2) Delivery phase

The radiating beams defined in the treatment plan are emitted, the beam shape is matched to the target volume geometry to get a steep dose gradient between the tumor and the surrounding organs at risk. These are high energy photon beams; generally, megavoltage X-rays generated by linear accelerators mounted on rotating gantries.

Higher accuracy and performance can be achieved if charged particle beams deliver the dose, these particles are accelerated by circular machines such as cyclotrons or synchrotrons (Amaldi, 2005).

1.6) Motion Management Strategies

Respiration, even though is a natural mechanism, can produce undesired consequences during radiotherapy, among those a blurring effect is worth highlighting given that it affects the dose distribution. This side effect influences the adherence to the treatment plan regarding the amount of delivered dose. In consequence, it is highly relevant to counteract movement happening during dose delivery. A successful compensation could translate into a more definitive therapy. Two methods could be mainly exploited in free-breathing condition: beam gating and beam tracking (Bert, 2011) (Rietzel, 2010).

1.6.1) Beam Gating

This technique allows the irradiation solely when the moving object is inside a predefined limiting window. Otherwise, the beam will not be supplied.

A movement tracking device estimates a signal that determines if the tumor is inside the limits (Minohara, 2000). The lesion will only be irradiated for a small amount of time because the window is usually defined at the maximum exhale phase, which is the steadiest point throughout the respiration cycle and lasts less than a third of it (Bert 2012).

1.6.2) Beam tracking

This strategy enables on-the-go regulation of beam direction and energy, through it a moving tumor can be irradiated without requiring limiting windows or other compensation methods. Consequently, it is said that this is the most efficient approach because it supports simultaneous motion compensation and maximum duty cycle (Keall P. a., 2001).

This method is exploited in photon radiation therapy; in particular, two commercially available systems make use of this technique: Synchrony RTS for CyberKnife (Accuray, Sunnyvale, California, USA) and VERO system (Vero GmbH, Germany). Further detail of these devices is presented in the following sections.

Regarding particle therapy, researches have also proposed the implementation of this technique for this type of therapy (Grözinger et al. 2004, Bert 2007, Saito 2009), but are not applied in the clinical practice yet, with beam gating the most used technique for the treatment of moving sites.

1.7) Photon Therapy State of the Art

Two commercial systems are mainly used for beam tracking: The CyberKnife System and the Vero System, both of them relying on the implementation of local motion models. Here we overview the two systems, and a detailed explanation of local motion model will be provided in the following chapters.

1.7.1) CyberKnife System

It is a radiosurgery device with the purpose of directing with high accuracy a beam in three planes through the coordinates given by images, to irradiate a particular section of the patient's body (Merriam-Webster).

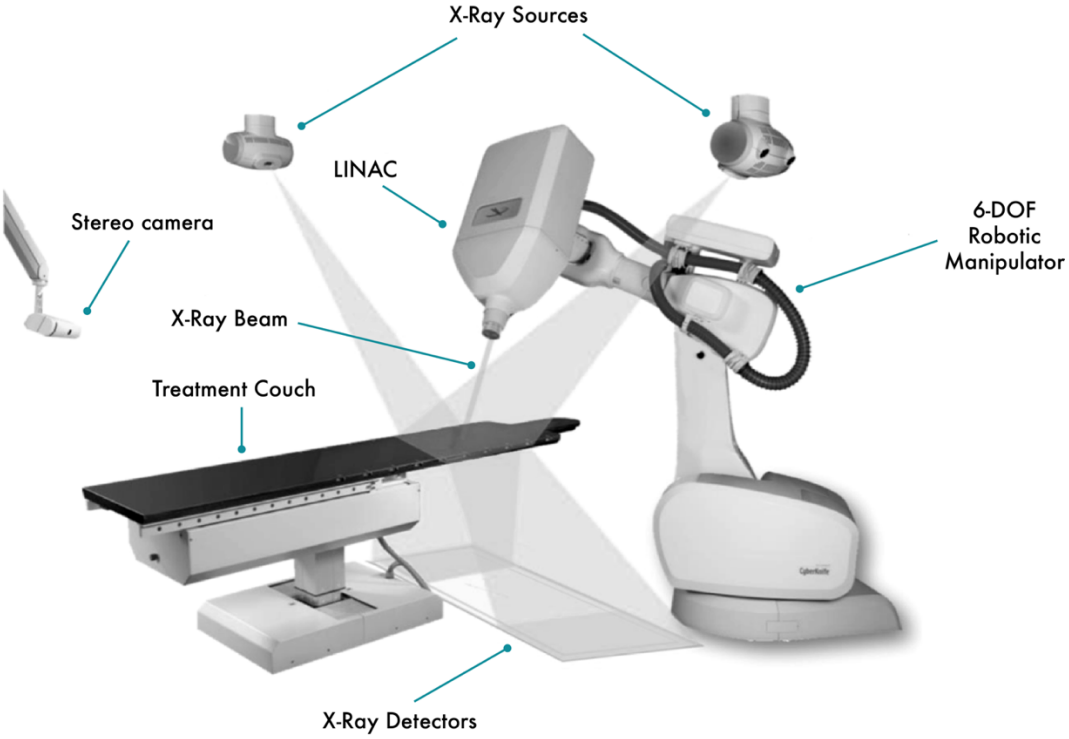

Figure 2. CyberKnife components.

The device has 6 freedom of movement, also thanks to its non-isocentric geometrical arrangement, it can supply non-coplanar treatment beams, i.e., each beam can be directed to a distinct location and can be produced from any angle (Schweikard 2000, Kilby 2010).

The elements of the device are as follows:

- LINAC, linear accelerator, this component provides an x-ray beam of 6 MV with a dose delivery of 1000 cGy/min, and this accelerator is attached to a robotic manipulator that has 6 degrees of freedom.
- Treatment couch, which can have up to 6 degrees of freedom.
- Two diagnostic X-rays attached to the operating room roof.
- A stereo camera that maps the location of the optical markers placed over the patient.

The previously mentioned components, in combination with the Synchrony Respiratory Tracking System, allow the machine to perform a live tracking of moving tumors.

The tracking system works in the next way: the tumor location is delimited before the actual irradiation; this is done for several points in time throughout separate respiratory phases using two X-ray images.

Surface markers are scanned using the system's stereo camera, then, a polynomial model associates the location of the lesion to the makers which allow this correlation model to predict the tumor's location throughout therapy. The system formulates a correlation model separately for each marker that is being used; the individual results are then averaged to get a single prediction. As the procedure goes on, further x-ray acquisitions are performed at a reduced frequency in time, approximately one acquisition per minute, which allows verifying and refreshing the model on demand.

1.7.2) Vero System

This is an image-guided stereotactic radiotherapy system; it is produced by BrainLAB AG, Feldkirchen, Germany; and Mitsubishi Heavy Industries, Tokyo, Japan [Kamino 2006].
It features a LINAC fixed to a gantry; the latter can rotate ± 60° to deliver non-coplanar radiation. Also, the LINAC can move in a pan-and-tilt fashion. The combination of both of the previously mentioned movements allows track tumor motion on the go.

Additionally, the system's architecture integrates two imaging acquisition pieces of equipment, which are fixed to the gantry. Moreover, for marker positioning, the device relies on infrared automatic optical localization.

In this system, X-ray acquisition is synched with the breathing signal which is obtained through infrared markers arranged over the patient's abdomen. This synched acquisition comprises four breathing cycles.

The motion model implemented in this system works with the breathing signal and the location of the implanted markers as inputs. This data is fitted to the correspondence pattern. Thus, markers and interior movement directions are separately adjusted according to the next polynomial:

$$F(s) = a^s + bs + c + d\dot{s}^2 + e\dot{s}$$

Having s as the surrogate signal, $\dot{s}$ corresponds to the surrogate's signal first derivative and $F(s)$ as the predicted target location over an individual path.

The system computes a single prediction by taking the mean of the estimations for individual markers. The surrogate's first derivative is employed to represent the hysteretic paths [Depuydt 2013]. The accuracy of this system is comparable with the one of the Synchrony RTS.

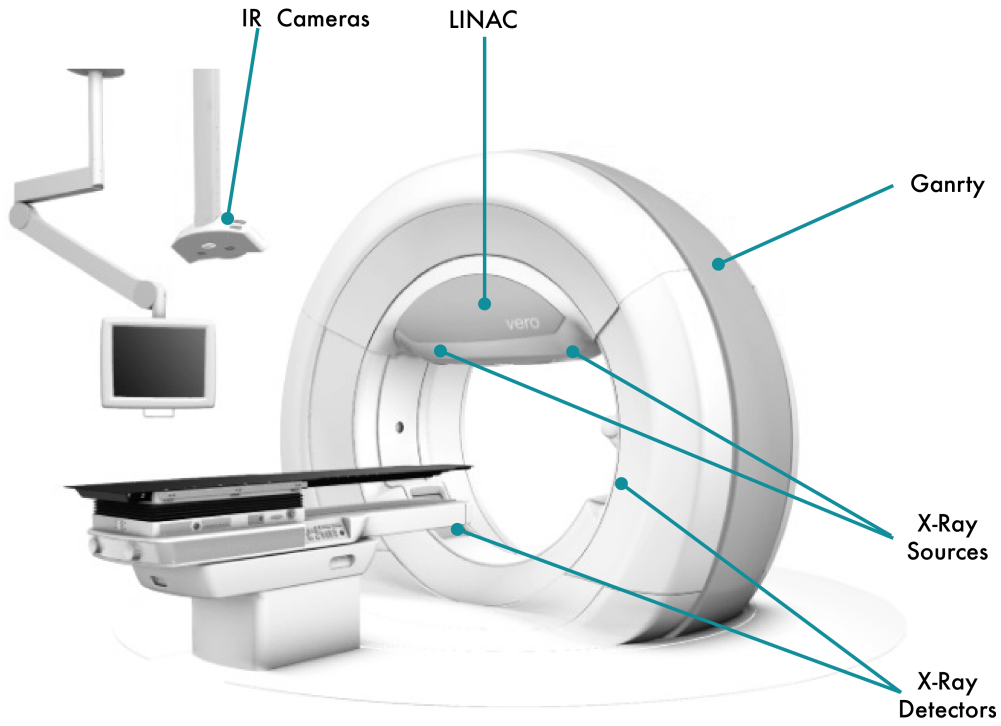

Figure 3. Vero system components.

1.8) Medical Image Registration

The goal of this operation is to match, confront and blend data coming from distinct images (Maintz et al. 1998a, 1998b). This procedure can be performed to integrate different images. These could be multi-modal, like CT and MRI, or mono-modal, as images that were generated by the same system (usually with different acquisition times). Image registration is particularly valuable in radiotherapy because it allows measuring the movement of the

organs of a patient, and also it allows to address the planning and treatments setups, which, for example, is helpful to align the position of a patient. Image registration is an optimization problem. Conceptually the performed operation is a spatial transform which associates a given position in a reference image with the reciprocal position in other(s) moving image(s).

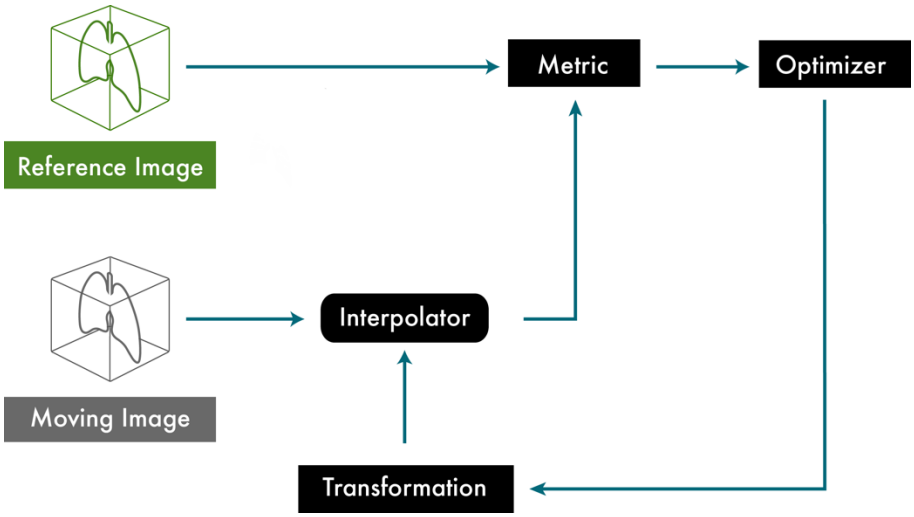

Figure 4. Image registration workflow, adapted from PPT Paganelli.

There are four components in the image registration workflow. In the first place, there is the interpolator block, which receives the coordinates of the moving image and converts them into the space of the reference image. Next, the metric segment determines the similarity metric among both images. Afterward, the optimizer gets the optimal specifications of the transformation which will allow getting a maximum/minimum similarity/difference. Finally, the transformation block, which provides the transformation to overlap the moving image on the reference one.

Two types of registration methods can be performed: Rigid and Deformable registration, both have a tradeoff between accuracy and speed, and as usual, the best option depends on the current application (Sarrut, 2006).

1.8.1) Rigid Registration

Rigid registration is only able to perform three translation and three rotation operations. Lines are kept straight, and distances are not altered. It is heavily used to compute patient

positioning corrections throughout the treatment phase; hence, it is essential for treatment design (Albertini S., 2016). The tissue deformation is neglected. This assumption presents a trade-off: it simplifies the operation, but at the same time it limits the applicability of the method because, in practice, organs present a high grade of deformability, as it is seen in the cardiac and respiratory cycles (Sarrut, 2006).

1.8.2) Deformable Image Registration, DIR

On the other hand, DIR manages to compensate for tissue deformation. The output of the DR is a deformation field (DF) of vectors; it displays the similarity between individual voxels in the first image and the corresponding voxel in the second one. The length of the vectors corresponds to the magnitude of the 2D displacement.

Similarly, warping consists of applying a deformation to one image in order to simulate another one, in this way all corresponding voxels have the same coordinates in both images. This deformed is also used to carry out data fusion by constructing a composite image representing information from both images using transparency colors.

DFs can also be represented as vectors on image slices, this way only a planar displacement (two of three dimensions) is depicted on each slice.

Biomechanical models may be an alternative to simulate organ deformation through the physical properties of the involved material and constraints based on the initial and final states of the organs. Finite-Element Method (FEM) and elastic model equations allow simulating individual organ deformation. Triangular meshes are used for surface models and tetrahedral meshes for volume models. The particular properties of the organ are described with parameters such as Young's modulus and Poisson's ratio (Sarrut, 2006).

1.8.2.1) B-Splines

Deformable Vector Fields can be calculated using different approaches, in this work the method referred to as free-form deformation based on B-splines was applied. This approach

consists of superimposing a grid of B-splines over the image's grid of voxels. The output of the interpolation of the B-spline and the voxel grids is, accordingly, an interpolated DVF (Shackleford 2010).

DVFs are computed through the next formula:

$$v(x,y,z) = \sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{k=1}^{4} P(i,j,k)B_i(x)B_j(y)B_k(z)$$

- $B_i(x)B_j(y)B_k(z)$: Previously computed B-spline basis functions.
- $P(i,j,k)$: B-splines coefficients, which are going to be optimized.

1.8.2.2) DIR Validation

The main problem in DR is validation; there are no standard approaches to assess its results. Some validation methodologies are:

- Simulated data: it involves the creation of an artificial image deformation with a numerical transformation. Its benefit is that the gold standard is known. Nonetheless, this technique is not linked to reality. It must be used with additional approaches.
- Phantom data: it is more attached to reality, but it is still somehow ambiguous. The gold standard is assessed through known deformations and implanted landmarks.
- Patient data: the gold standard is not known. It is possible to specify it by the manual selection of landmarks by specialists. Implanted landmarks can also be used; these small objects are readily detected in the acquired images (Seppenwoolde, 2002). Other metrics could also rely on manually delineated contours, such as Dice Coefficient (i.e., contours overlap) or Center of Mass Distance between the contours.

These approaches have proved their convenience in several applications, an example is when the propagation of organs/tumor contours requires to be analyzed: the distance among deformed contours and the ones defined by experts must be between the variability values

determined by specialists. Despite its drawbacks, deformable image registration algorithms are fundamental in IGRT (Sarrut, 2006).

1.9) Motion Model

We can distinguish two types of motion modeling:

- Local motion model, If we are planning to track just local points in the tumor
- Global motion model, if we are planning to track the whole anatomy

1.9.1) Local motion model

These approaches aim to track an external marker or (breathing) signal, i.e., a surrogate, that has a clear correlation in time and space with the motion of the tumor (Vedam et al. 2003b). The model will be updated continuously using the tumor's latest available position, which is determined using an imaging system (Murphy et al. 2006). Conventionally, x-rays are used to quantify the patient's internal movements; this method can be implemented locally using implanted markers, concentrating only in some points of the anatomy, or globally, by tracking an entire structure. This technique needs at least a double projection (Schweikard 2000).

These tools try to detect the association that exists between motion and appropriate surrogate data. This strategy is helpful in cases when the motion of interest cannot be measured directly, it might be too invasive, or the acquired data might not have the required temporal resolution. Moreover, as previously stated, models may be preferred when the amount of dose delivered to the patient needs to be restricted.

Since there are several options to choose from among motion sources and magnitudes to be measured as surrogates, some properties have been suggested to select a surrogate source data correctly: it should be readily determined, have a clear correlation with the motion, and a sharp resolution in time.

In the case of IGRT for lung tumor treatment, it is convenient to highlight the possibility of using optical tracking methods. This kind of technology tracks the movement of a point or a set of points on the surface of the patient's chest, usually through the use of lasers.

Researchers have brought to light that a customized respiratory motion supervision strategy seems to be the proper approach given that motion properties vary between patients and treatment sessions. Recently, commercially available systems are making use of this kind of approaches, like Cyberknife System by (Accuracy, Sunnyvale, California, USA) (Kilby et al. 2010) and Vero System by BrainLAB AG, Feldkirchen, Germany and Mitsubishi Heavy Industries, Tokyo, Japan (Kamino 2006).

Roughly, models are constructed and implemented as follows: the internal motion and the surrogate must be acquired at the same time. Next, during the training phase, the acquired data is used to fit a correspondence model. When the model has been successfully trained, it is ready to be used to estimate the internal motion. It will continue to use as an input a surrogate signal.

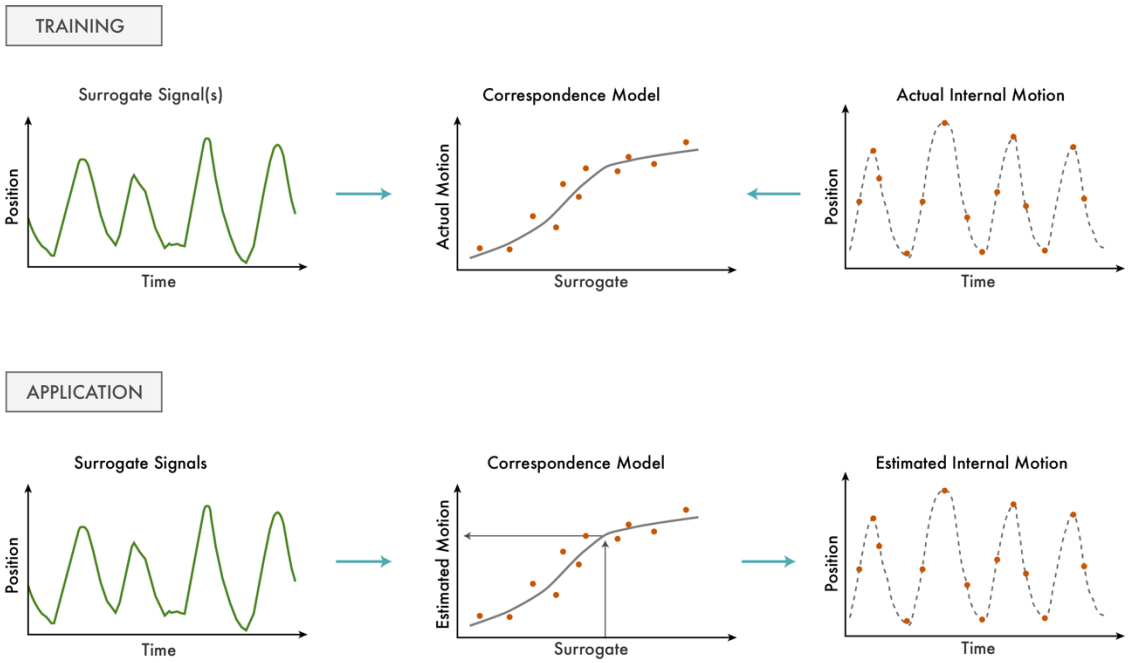

Figure 5. Motion model descriptive operation. Adapted from Bombardieri.

The model relies on the input signals, the characteristics of the movement that is modeled, and the properties of the correspondence model that will depict the surrogate and motion

association. To reach the model's highest potential more than one surrogate signal should be used; this will raise the complexity of the model though.

The general expression to define the internal-external correlation function is a polynomial formulation:

$$M(s): \sum_{i=0}^{n} a_i s^i$$

The components of this function are as follows: s stands for the surrogate, M(s) is the modeling output, a refers to the polynomial coefficients.

Polynomial models present a deficiency when the surrogate signal is outside of the training data extent, in this scenario they will render large extrapolation errors.

1.9.2) Global motion model

Researchers have introduced global respiratory motion models based on images, in this case, the motion is predicted using image registration and allows to estimate the position of the tumor and the whole anatomy (McClelland et al. 2013). In this way, they are providing relevant details like position, shape, and volume (Seregni, 2013). These models essentially define the correlation between the internal anatomy, which is derived through deformable image registration, with an external surrogate. Specifically, deformable registration is performed between different respiratory phases of a planning 4DCT, and the motion field is related to the acquired external surrogate. During treatment, the model is applied and, using the sole external surrogate; it is possible to derive an estimated current CT. The correlation between the internal motion field and the external surrogate can be derived by both principal component analysis (Fayad et al. 2012) or interpolation of the motion field (Vandemeulebroucke et al. 2009). Motion models are typically constructed on a 4DCT and can be updated through a CBCT acquisition to compensate for inter-fraction variations (Fassi et al. 2015).

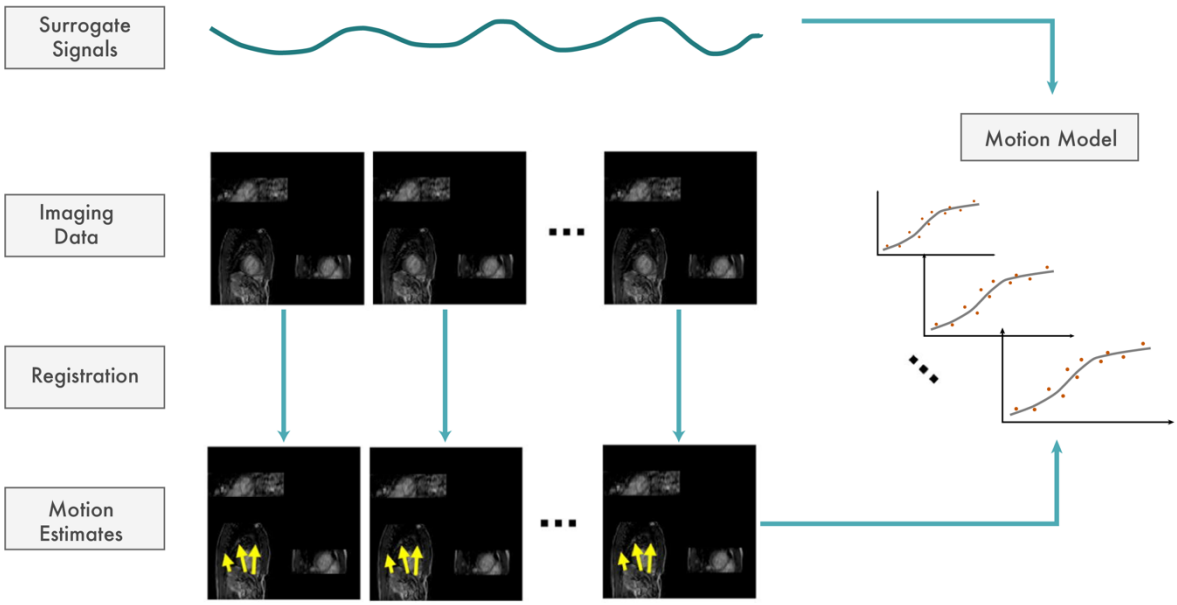

Figure 6. Motion model descriptive operation. Adapted from McLeland.

These models are not used in the clinic yet, but they are widely explored in literature, and also in the next Methods section of this work.

1.9.2.1) Validation of global motion models

One of the main limitations of global motion models is that they are sensitive to new input: if the model is trained on a planning 4DCT, variations from this scenario could affect the motion model performance.

Validation of the performance of this approach is therefore mandatory to evaluate if the model can estimate a respiratory phase in case of both regular and irregular breathing patterns. In particular, when multiple acquisitions are performed during the treatment workflow, the model has to be tested to evaluate its capability to compensate for motion accurately.

Two possible strategies can be exploited to evaluate the motion model:

1. Leave one out: Build the motion model on a 4DCT and leave one phase of the 4DCT out of the training, thus subsequently testing the performance of the motion model on the remaining respiratory phase. This was the strategy selected in this work.
2. Multiple acquisitions: train and test the motion model on multiple acquisitions during the treatment workflow. This could have been implemented in this work with the use of a series of repeated 4D CBCT images.

1.10) Aim of the project

Considering the previously mentioned developments in image-guided radiotherapy, it is relevant and of common interest to develop and apply validation strategies to account for the accuracy of deformable image registration and global motion models.

Hence, this work is meant to validate the image registration process together with the performance appraisal of a global motion model, created by Vandemeulebroucke et al. (2009) for 4DCBCT applications.

This will exploit the availability of a dataset composed of different 4DCBCT with implanted markers which can be used to estimate the target position accurately.

Chapter 2. Materials and Methods

This thesis has two purposes, first to quantify the quality of the image registrations of 4DCBCT data, and additionally, to evaluate the quality of the performance of the global motion model developed by Vandemeulebroucke et al. in 2009.

Roughly this will be made by measuring the distance between the implanted markers detected in a reference image and in the image generated as an output of the given processing method, which is either the DIR image or the image estimated by the motion model.

2.1) Image Datasets

The present work was developed using as data sources the image acquisitions obtained from three patients who endured lung cancer and had implanted markers in their respective tumoral lesions. Also, it is worth mentioning that The University of Sydney in Australia provided the dataset employed in this research.

A multi-modality approach was carried out, justified by the requirements of the motion model and also by the workflow of the study. The involved modalities consist of a) ten-bin 4D CBCTs for each acquisition session, representing the whole respiratory cycle, from end-exhale to end-inhale, b) the tumor contour delineated on the exhale phase of a 4DCT. The number of image acquisition sessions is as follows: three sessions for patient A, two sessions for patient B, and three sessions for patient C, all of them with pre-treatment data.

The treatment dataset corresponds to the 4D CBCT projections; these volumes were reconstructed from a respiratory signal and by following the motion-compensated approach (Rit S., 2009).

The tumor contour files were elaborated manually by a medical doctor who was in charge of highlighted this structure.

Figure 7. Tumor contour.

As mentioned in the previous section, the usefulness of these images relies upon the fact that the imaged patients have implanted markers in their tumoral lesions, which allows performing more accurate assessments, due to the low contrast of the CBCT data.

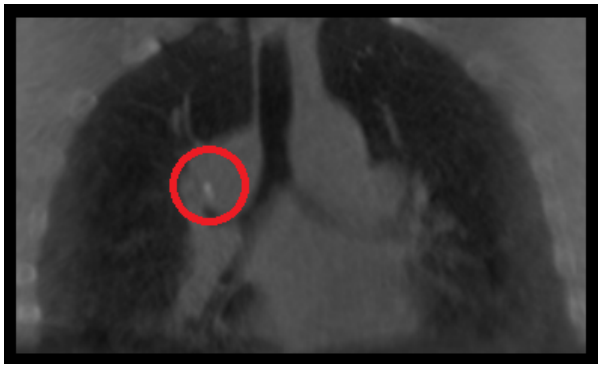

Figure 8. 4DCT image is displaying the implanted marker encircled in red.

2.2) Data Processing

2.2.1) Pre-Processing

The dimensions of the images were modified with the purpose of image processing optimization and artifact reduction. The following table highlights the before and after image properties.

Average CTs	Original Dimension in Pixels	Resolution in mm	Modified Dimension in Pixels
Patient A	512 x 142 x 512	0.9765 x 3 x 0.9765	362 x 88 x 280
Patient B	512 x 149 x 512	0.9765 x 3 x 0.9765	269 x 86 x 194
Patient C	512 x 131 x 512	0.9765 x 3 x 0.9765	284 x 89 x 222
4D CBCTS	**Original Dimension in Pixels**	**Resolution in mm**	**Modified Dimension in Pixels**
Patient A	450 x 220 x 450	1 x 1 x 1	286 x 168 x 226
Patient B	450 x 220 x 450	1 x 1 x 1	244 x 171 x 179
Patient C	450 x 220 x 450	1 x 1 x 1	269 x 176 x 169
Tumor CTs	**Original Dimension in Pixels**	**Resolution in mm**	**Modified Dimension in Pixels**
Patient A	512 x 142 x 512	0.9765 x 3 x 0.9765	Not modified
Patient B	512 x 149 x 512	0.9765 x 3 x 0.9765	Not modified
Patient C	512 x 131 x 512	0.9765 x 3 x 0.9765	Not modified

Table 1. Original and modified properties of the images (defined in the right-let, superior-inferior and anterior-posterior directions).

The Tumor CT volume did not require any cropping because of the nature of the image, i.e., the image is composed of a tumor contour with a null-data background.

Then, a Gaussian filter with a standard deviation of 0.6 was applied to these images to reduce their noise components.

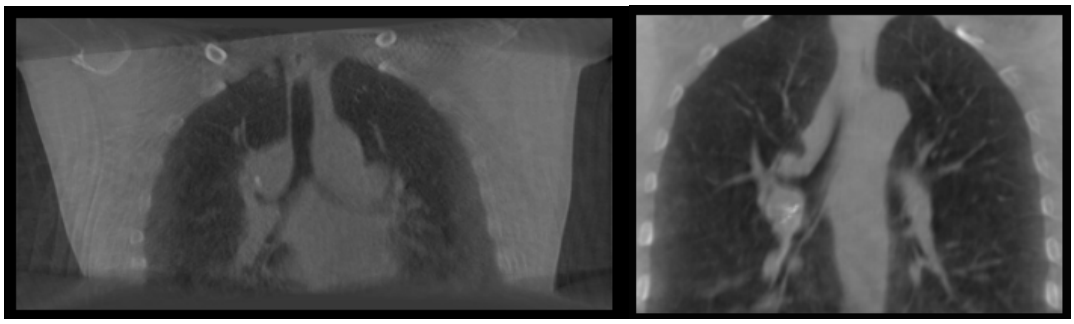

Figure 9. Left: One of the original images, a wider field of view is seen. Right: Cropped and filtered image, reduced FOV and noise are recognized.

2.2.2) Registrations

2.2.2.1) *First step: Deformable Image Registration*

The first step in the proposed procedure was to perform a DIR between the 4D CBCT images that describe the full respiratory cycle, specifically among:

a) The fixed volume which corresponds to the cycle's exhale phase (the most stable phase of the respiratory cycle).
b) The rest of the breathing cycle's volumes, which are referred to as the moving volumes.

The outputs of this operation are a vector field and the DIR volume itself.

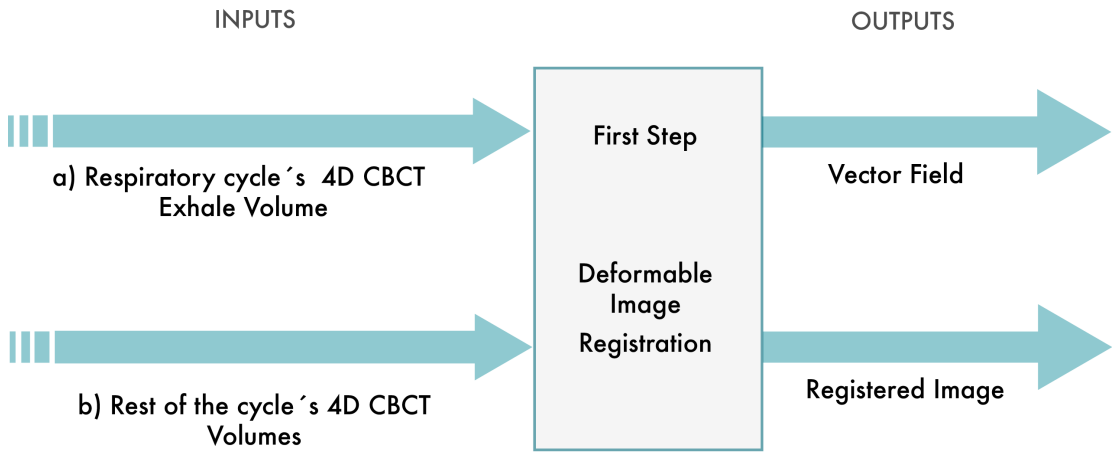

Figure 10. Diagram showing the inputs and outputs involved in the Deformable Image Registration.

Each DIR had four different stages, all of them were performed using the B-Spline method (Shackleford, 2010). Usually, more than one stage is needed to perform a proper registration; this is because registrations are an optimization problem, after each stage the accuracy of the procedure is increased.

Mainly, two registration parameters, when properly tuned, allow enhancing the DIR quality, these are RES and GRID_SPACE, both of them become smaller and smaller after each stage, this, in turn, increases the definition of details, yielding an improved registration after each stage.

This couple of parameters are defined as follows:

- RES: It characterizes the dimension of the image regarding the pixel/voxel ratio. Having a small value for this parameter translates into a better representation of details.

- GRID_SPACING: This determines the size of the grid that is superimposed to the image over which the registration is performed.

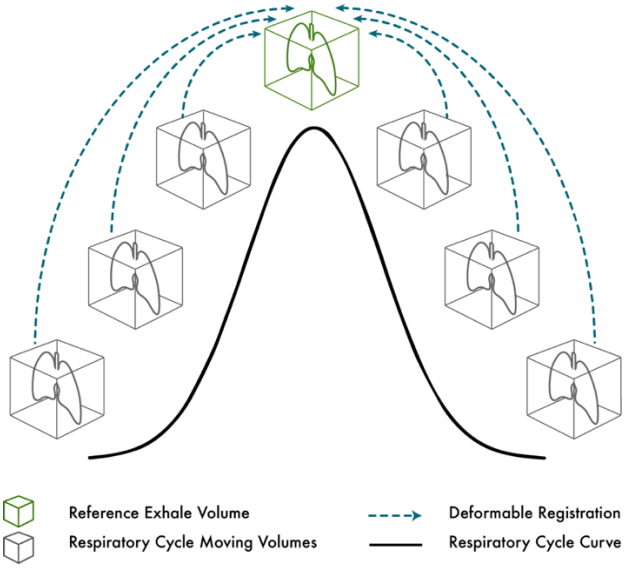

Figure 11. Conceptual diagram of the Deformable Image Registration process.

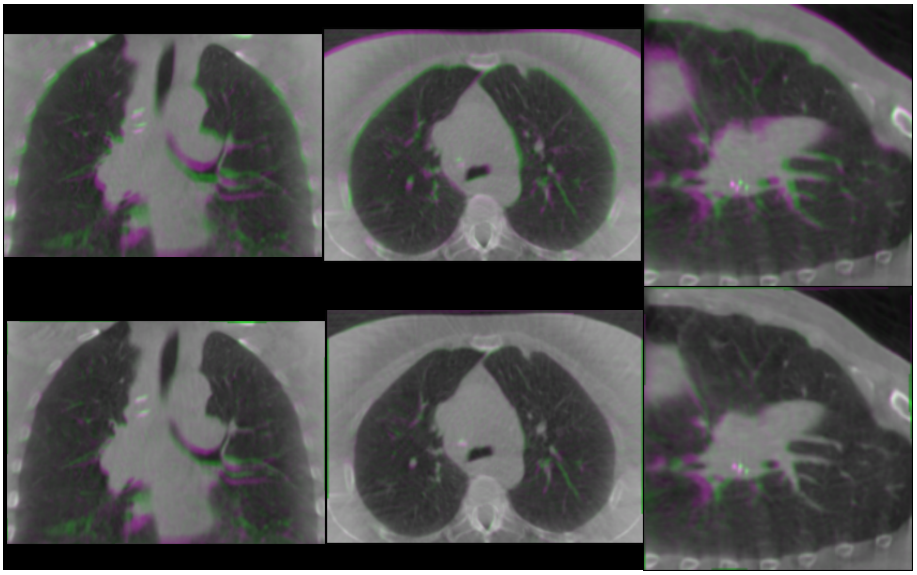

Figure 12. Visual evaluation of the DIR output. Upper row: Different views of an overlay between the unregistered reference image (green) and the moving image (violet). Lower row: Different views of an overlay between the unregistered reference image (green) and the registered moving image (violet).

2.2.2.2) Second Step: Rigid + Deformable Registration

Then, for each session, a Rigid and Deformable Image Registration was performed between:

a) The ten CBCT volumes that describe the whole breathing cycle
b) The CT inhale volume

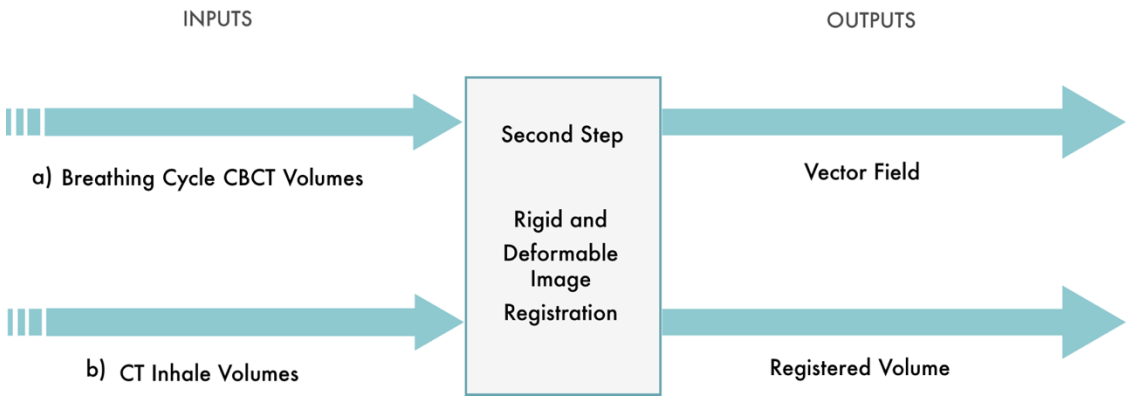

Figure 13. Diagram showing the inputs and outputs involved in the Rigid & Deformable Image Registration.

Several tests involving the registration parameters were made with the aim of finding the best possible registrations. Different image dimensions were first explored in these tests to find the optimal registration while respecting the relevant field of view. This was followed by the search for the optimal registration parameters for each involved stage. It was noted that the best registration output was rendered with a succession of eight stages of rigid registration combined with three stages of deformable registration, which made use of the B-spline method. The registration parameters were tuned throughout the eight stages as follows: the GRID_SPACE component was gradually reduced from 20 to 5 in all directions, and the value of the RES element from 8 to 2 units, thus providing a registration with increased detail and alignment.

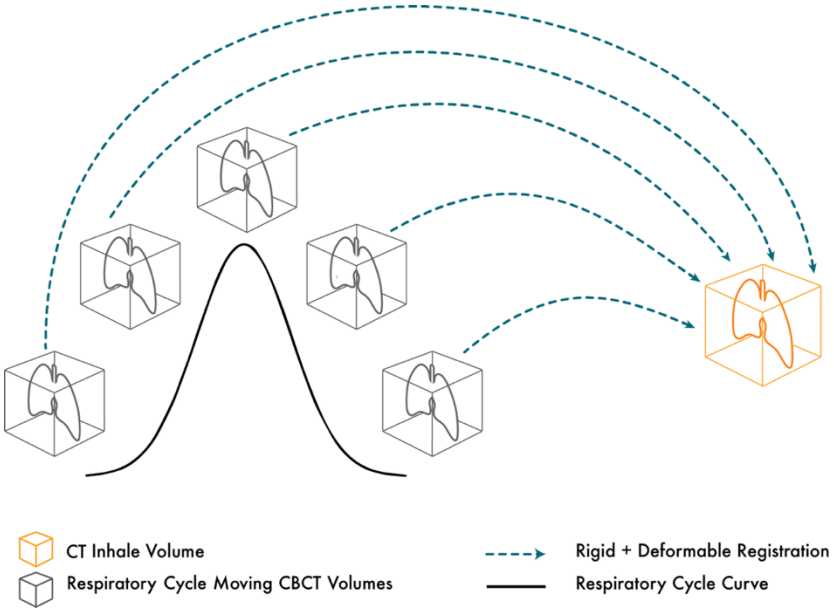

Figure 14 Conceptual diagram of the Rigid & Deformable Image Registration.

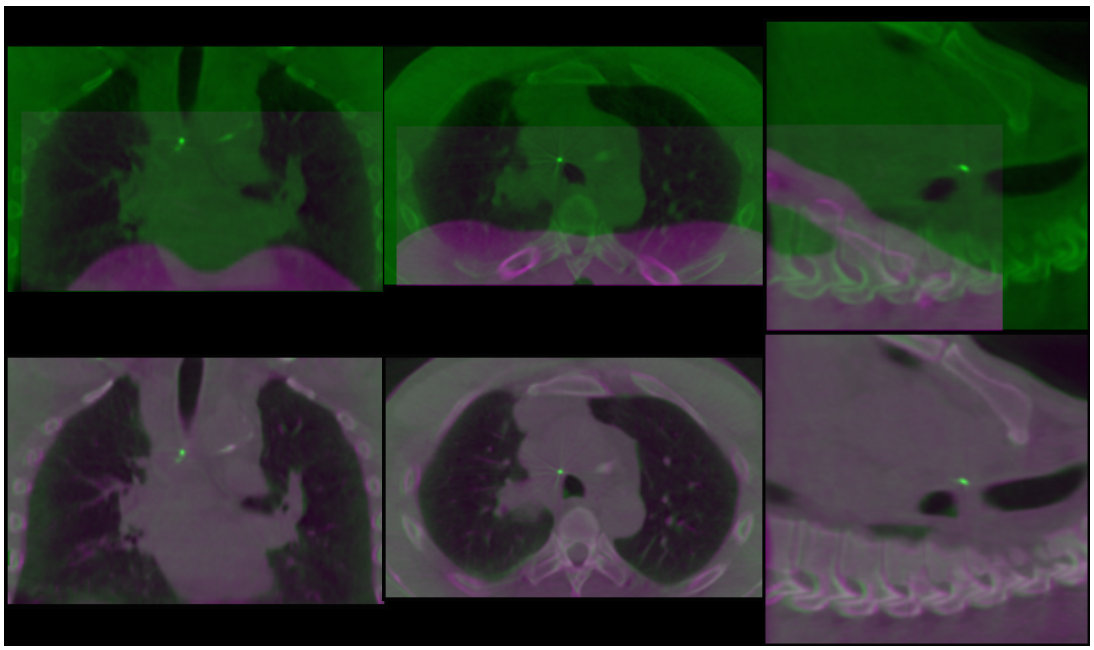

Figure 15. Visual evaluation of the R+DIR output. Upper row: Different views of a simple overlay between the unregistered images of a moving CBCT volume (green), and a CT inhale volume (violet) . Lower row: Different views of an overlay between the minimum exhale phase, and the output of the R+DIR between the minimum exhale phase and the 4D CT inhale volume.

2.2.3) Warps

The purpose of this procedure is to fit the tumor contour into the CBCT-inhale space and subsequently into the CBCT-exhale space. This is done by taking the tumor contour, and by performing translations and deformations to it. In this case, the Rigid + Deformable registration step between the CT and the CBCT at the inhale phase was used to warp the tumor volume delineated on the CT to the CBCT inhale phase (Warp1), whereas the step of DIR between the 4DCBCT respiratory phases was used to warp the tumor volume from the CBCT inhale to the CBCT exhale (Warp2), as depicted in Figure 16.

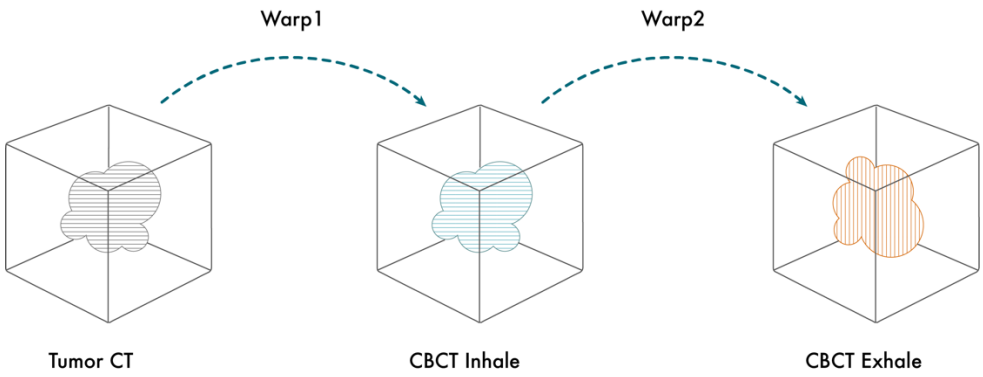

Figure 16. The final output of the two warps is a tumor contour which fits in the CBCT exhale space.

2.2.3.1) Third Step: Warp 1

In this first warp, the vector fields obtained from the Rigid + Deformable registration (c) were applied to the tumor volume, as shown in the following diagram.

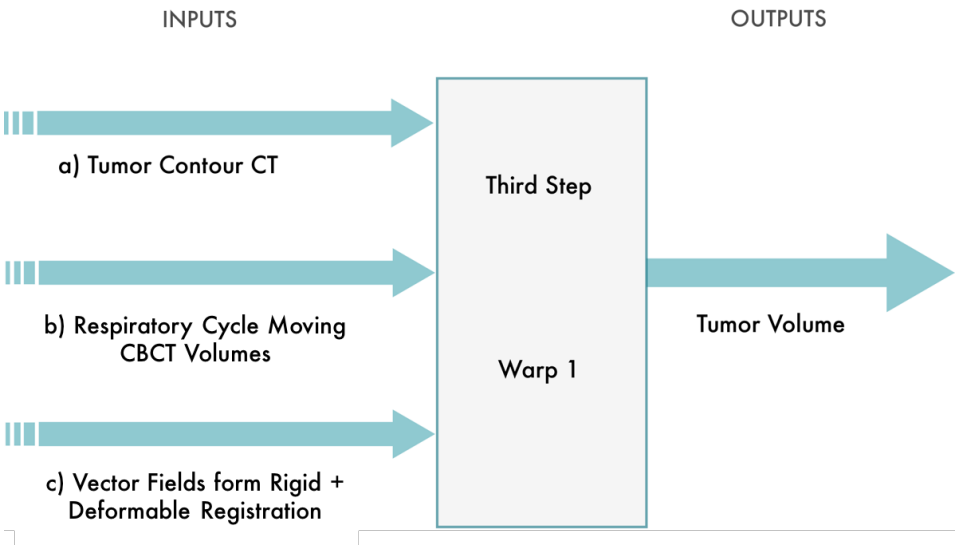

Figure 17. Diagram showing the inputs and outputs involved in the first warp.

Figure 18. Illustrative output of one of the warping operations.

2.2.3.2) Fourth Step: Warp 2

This second warp had as inputs a) the output from the previous warp, i.e., the tumor fitted to the inhale volume, and b) the vector field that was obtained as output from the first DIR, i.e., the registration that uses as a reference the exhale volume.

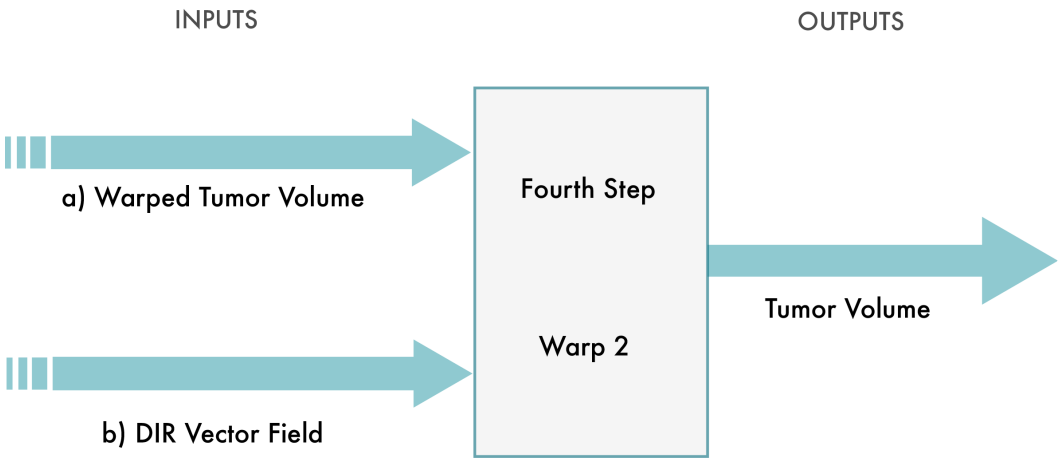

Figure 19. Diagram showing the inputs and outputs involved in the second warp.

The final output is a tumor contour that matches the space of the respiratory cycle moving phases.

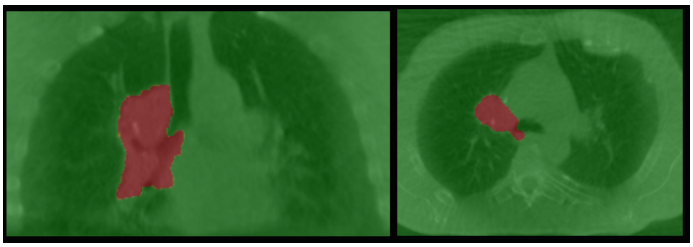

Figure 20. Fusion image between a given phase of the respiratory cycle and the artificially generated tumor contour (in red).

2.3) Measuring the Local Registration Error

The computation of the Euclidean distance between the markers on the exhale and the DIR-output volumes is used to assess the magnitude of the registration error. This distance is proportional to the registration error, i.e., the smaller the distance, the smaller the registration error (higher registration quality), and likewise, the bigger the distance, the bigger the error (worse registration quality). This procedure is depicted in the following image.

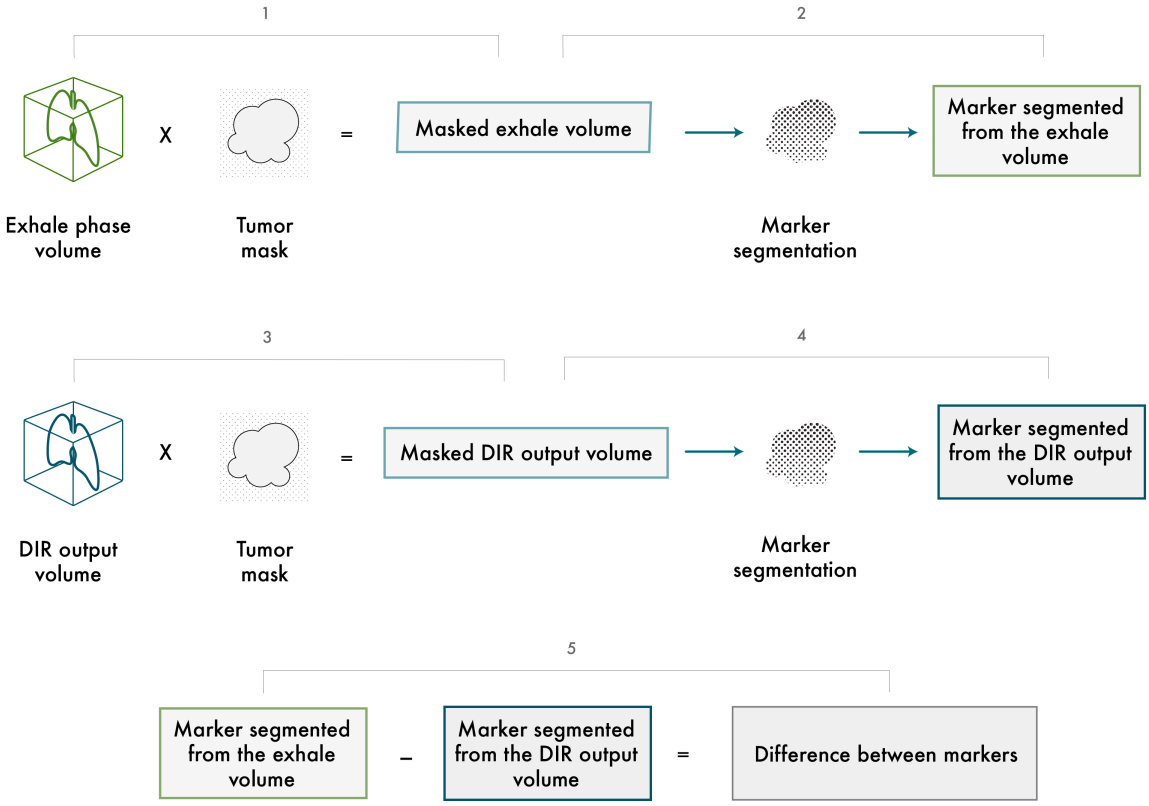

Figure 21. Computation of the Euclidean distance between the markers.

The process outlined in Figure 21 is carried out in order to measure the distance between markers, i.e., the distance between the implanted marker located on the tumor of the exhale volume, and the implanted marker located on the tumor of the registered volume, as depicted in step 5.

This method is divided into three major sections:

- First: Isolation of tumoral volumes on the exhale phase volume, as in step 1, and also on all of the volumes obtained as outputs of the deformable image registration as in step 2.
- Second: the implanted markers must be segmented from the already isolated tumoral volume of the exhale phase (as in 2) and of the DIR outputs (as in 4).
- Third: Distance (error) measurement, as in step 5.

More details are discussed in the following subsections.

2.3.1) Isolating Tumoral Volumes (Masking the Volumes)

The goal of this first step is to isolate the tumor volume from the exhale and DIR-output volumes; in other words, the aim is to get rid of all of the non-tumoral data within the exhale and DIR-output volumes. This is done by taking the volumes mentioned above and performing a multiplication between them and a special tumor contour volume, known as the tumor mask, thus discarding the non-relevant data.

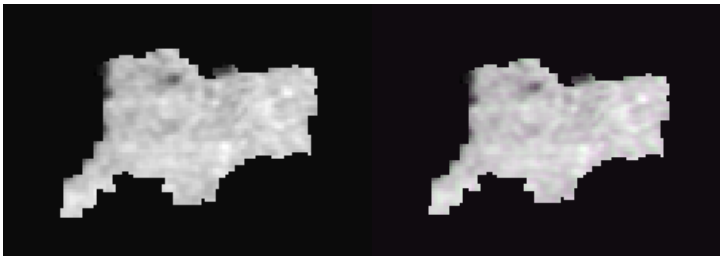

Figure 22. Masked volumes visualization. Left: Tumoral volume extracted from the exhale phase. Right: Overlay between the tumoral volumes extracted from the exhale phase and a given DIR-output.

2.3.2) Marker Segmentation

In the previous step, the tumor was extracted from the exhale and DIR-output volumes; now, this current stage aims to isolate the markers from these extracted tumor volumes. This process of marker isolating is referred to as marker segmentation.

The first step to extract the markers is to set a threshold value, then everything with a value higher than this threshold will be extracted from the tumoral volume, following that, a value of one is assigned to the extracted data. The outcome of this is considered as the segmented marker; it is an image that is composed of a null (zero) background and a marker contour which is made up of only values equal to one. The result is represented in the following image.

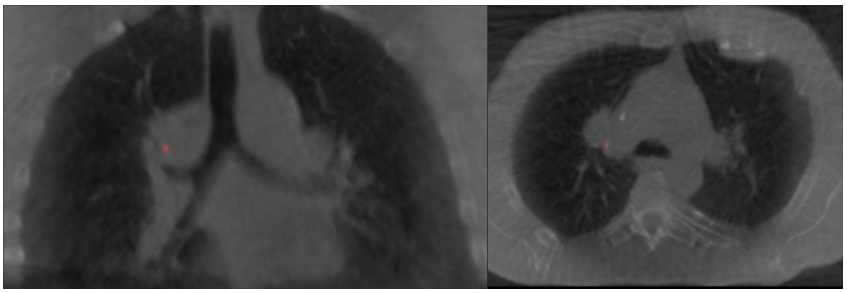

Figure 23. Visual representation of the segmented marker. This image was created by fusing the exhale phase and the segmented marker (in red).

This step was repeated sever times using different thresholds; this was done to determine which was the optimal threshold level.

2.3.3) Error Measurement

The final step is to compute the distance between the segmented markers; this is done by taking the markers extracted in the previous step and computing the centroids of the marker regions. Then the distance between the centroids is measured, this is interpreted as the registration error. Additionally, the root-mean-square error is computed for this measurement. The results of this procedure are explored in the following chapter.

2.4) Measuring the Global Registration Error

Fundamentally, it can be said that the structures that require the highest registration accuracies are the implanted marker and the volume comprising the tumoral lesion.

Nevertheless, since in this study the image registrations were carried out using a global approach, and given that the study aims to evaluate the capabilities of the employed image registration methods, then it is also required to evaluate the global outcome, which is done by computing the global registration error.

This global measurement was divided in two:

- Measuring the global difference between volumes before registering them

- Measuring the global difference between volumes after registering them

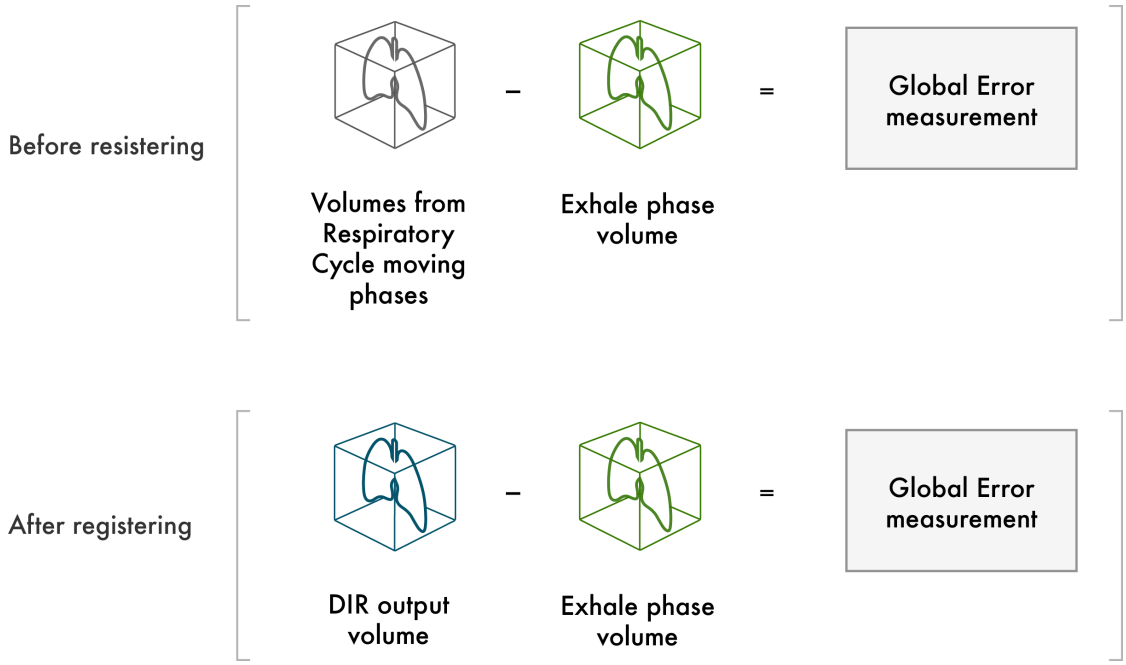

Figure 24. Global registration error measurement.

2.5) Motion Model

The second part of this work consisted of using a global motion model (GMM) to estimate the different phases of a given respiratory cycle. Furthermore, the estimations were compared to their corresponding ground truth volumes by measuring the distance between the predicted implanted marker position and the ground truth marker position (an almost identical process as in the DIR-output error measurement).

Hence, this section can be divided into:

1) Model Description.
2) Model Implementation.
3) Measuring the model's error.

2.5.1) Model Description

The motion model applied throughout this work was developed by Vandemeulebroucke et al. in 2009.The following diagram displays the workflow that describes how was this model implemented to generate the volume estimations for the different respiratory phases.

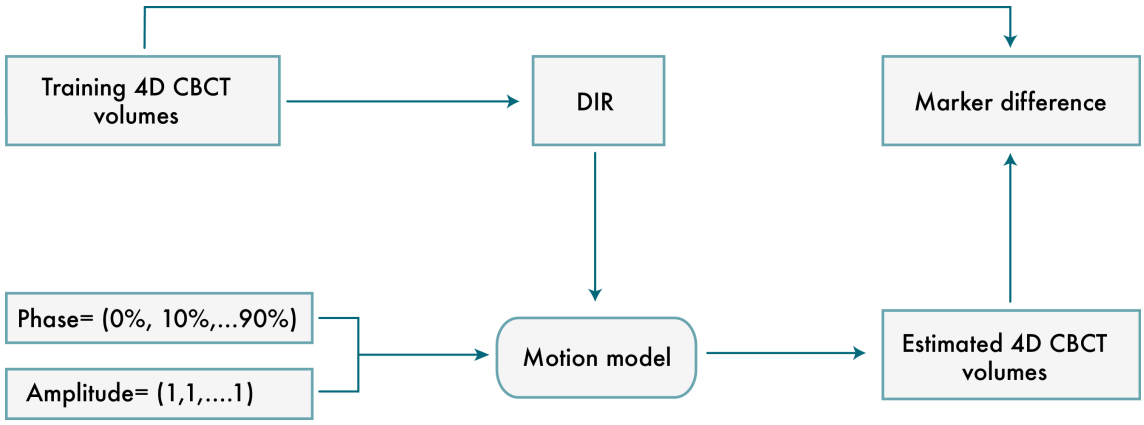

Figure 25. Workflow describing the implementation of Vandemeulebroucke's motion model and further error assessment. Adapted from Fassi.

2.5.1.1) Model's inputs

The GMM requires various inputs; Figure 26 is displaying the different data requirements of the model.

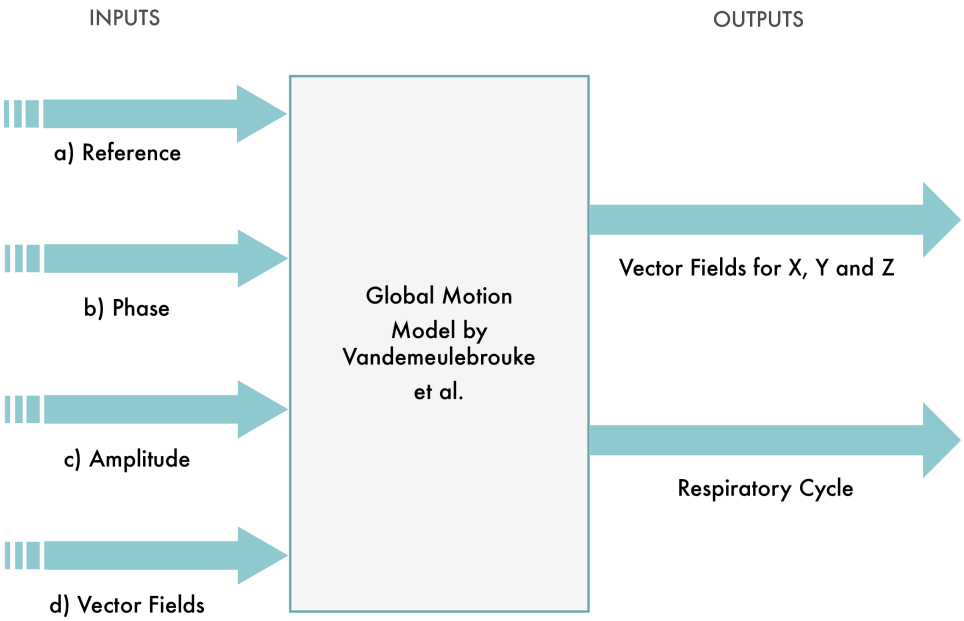

Figure 26. Diagram showing the inputs and outputs involved in the GMM implementation.

The inputs are described as follows:

a) Reference: A volume used as reference for motion estimation.
b) Phase: A text file that describes the value of the phase of the cycle.
c) Amplitude: A text file is also used to describes the value of the amplitude of the phase of the cycle.
d) DIR VFs: The deformable vector fields obtained as outputs from a DIR between each volume and the reference volume of the 4DCBCT dataset.

2.5.1.2) Model's Mathematical Function

The model works by implementing the following function:

$$S = s + \Delta s + \alpha_t * D_{\phi_t}$$

The individual terms are defined as follows:

S: the value of a particular motion state

s: surrogate baseline, which was neglected in this study

Δs: baseline shift, as for the previous term, this current value was not implemented in this study because the scope of this work was only to assess the GMM prediction capabilities.

α_t: instantaneous amplitude

ϕ_t: instantaneous phase

D_{ϕ_t}: interpolation results

The instantaneous phase and amplitude are accountable for the generation of uniquely shaped paths and also for diversified respiratory level values (Vandemeulebroucke et al. in 2009).

The deformable vector fields (which are fed as inputs to the model) are interpolated according to the phase ϕ_t, which ranges from 0 to 1. This operation is performed by using the b-splines method, the output of which is assigned to D_{ϕ_t}.

Then, a new deformable vector field is generated, it describes the displacement of the moving acquisitions for the current phase and reference volume. Additionally, this vector field is also adjusted in proportion to the instantaneous amplitude fed to the model.

To conclude, the reference volume is transformed (warped) with respect to the newly created deformable vector field; the outcome of this final step is the estimated volume.

2.5.2) *Model implementation*

The model was directly applied on the 4DCBCT dataset and used to estimate the one of the respiratory volumes in a leave-one-out approach. The method was implemented for the three patients that took part in this study, this includes all of the sessions underwent by each patient, i.e., three sessions for patient A, two sessions for patient B, and three sessions for patient C. The quantitative assessment of these estimated 4D CBCT volumes is discussed in the following chapters.

2.6) Measuring the Local Estimation Error

In a similar fashion of the DIR evaluation, also in this case, the distance between markers is measured between the estimated and the ground truth. The results of this operation are exposed in the next chapter.

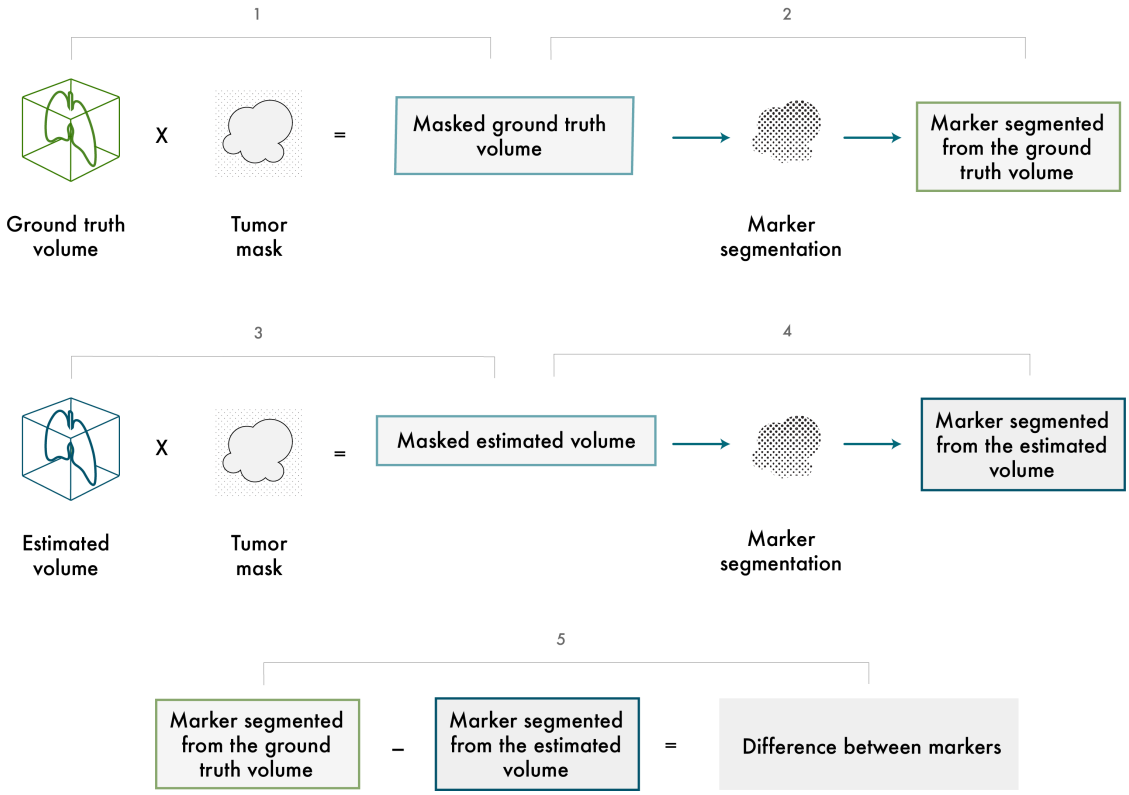

Figure 27. Computation of the Euclidean distance between the markers.

2.7) Measuring the Global Estimation Error

It is highly relevant to measure the global modeling capabilities of Vandemeulebroucke's model, not limiting thus, the quantification of the proficiency of the motion model. Therefore, as in the case of the assessment of the image registration output, after weighing the local outputs of the model, at this point it is required to measure the global estimation error of the model.

This global measurement is also divided into two steps:

- Measuring the global difference between the ground truth volumes
- Measuring the global difference between estimated volumes and ground truth volumes

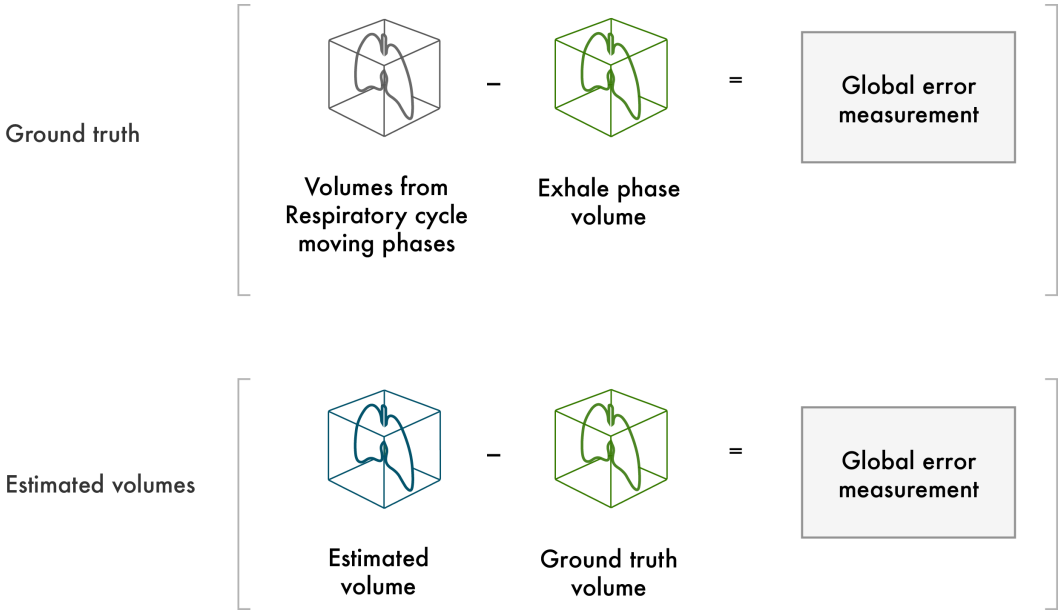

Figure 28. Global model error measurement.

The absolute difference between the estimated volume and the ground truth volume is computed for each respiratory phase and for every session of each patient.

For example, the estimated volume corresponding to the first respiratory cycle phase is compared against the ground truth volume that corresponds to the first respiratory cycle phase, the same is done for the second phase and the third, and so on. This computation is done for every session of each patient.

The results of these operations are exposed in the next chapter.

Chapter 3. Results and Discussion

The rest of the chapter will present an evaluation of the accuracy of the registration process and the motion model. Having an accurate DIR is crucial given that the accuracy of the model depends, in part, on the accuracy of the DIR.

The results are presented under two different approaches: first, a synthesized qualitative representation is exposed, later, the quantitative output of the procedures carried out throughout the work are shown. The results will be divided into two subsections: local error and global error.

3.1) Local Error

The local error measurements represent the Euclidean distances between the implanted markers, defined in millimeters. This was computed for each session of each patient, as well as a grouped analysis was performed.

3.1.1) Qualitative Threshold Selection

In order to compute the distance between pre-registration and post-registration markers, i.e., the registration error, one of the involved steps was the segmentation of the markers, also known as the process of isolating the marker from the tumor volume. A fundamental step in this segmentation process is the selection of an optimal threshold; this consisted in exploring different threshold values to determine which was the optimal threshold level.

This section will describe how this process was performed in order to justify the final threshold selection. The image below corresponds to the threshold comparison performed for the first session of patient A. Two different rows of images can be identified in this figure:

- The upper row entitled as Exhale Segmented Markers is displaying the same exhale reference volume in grayscale with different threshold values (labeled with Greek letters) applied to the marker segmentation process in order to isolate the implanted

marker. The result of the marker segmentation is highlighted in red; this area is intended to be the marker.

- Similarly, the lower row, which is entitled DIR output Segmented Markers presents the same image registration output volume in grayscale with the corresponding highlighted marker also in red.

This same outline will be used for the rest of the threshold identification processes of patient A, as well as for patients B and C and their corresponding sessions.

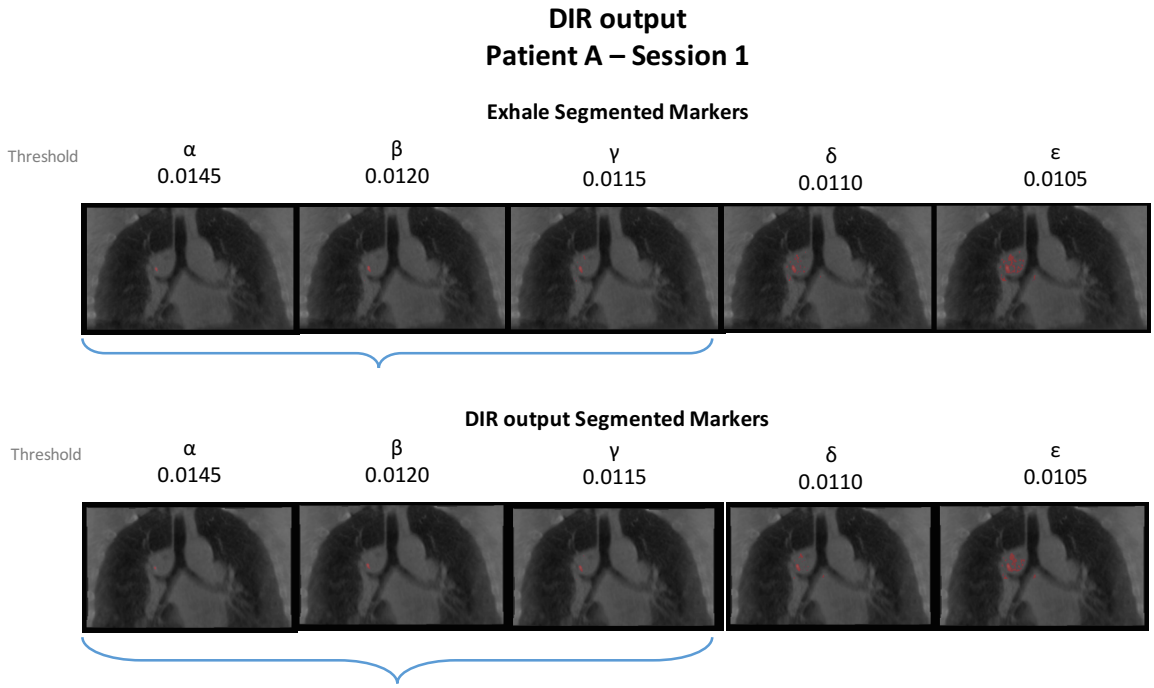

Figure 29. Comparison between different thresholds for patient A, session1. Volumes related to the DIR output.

In Figure 29, it can be seen that the relation between the threshold and the quality of the segmentation is proportional, i.e., for a higher threshold level, the segmentation output is finer: it has a higher quality; on the other hand, for a lower threshold the quality of the segmentation is lower as well.

After analyzing figure 29, it can be qualitatively said that the optimal threshold is the one labeled as α, which is the higher. Threshold β and γ also represent good options, but not the optimal.

The next set of figures depict the rest qualitative evaluations performed for the rest of the patient's A.

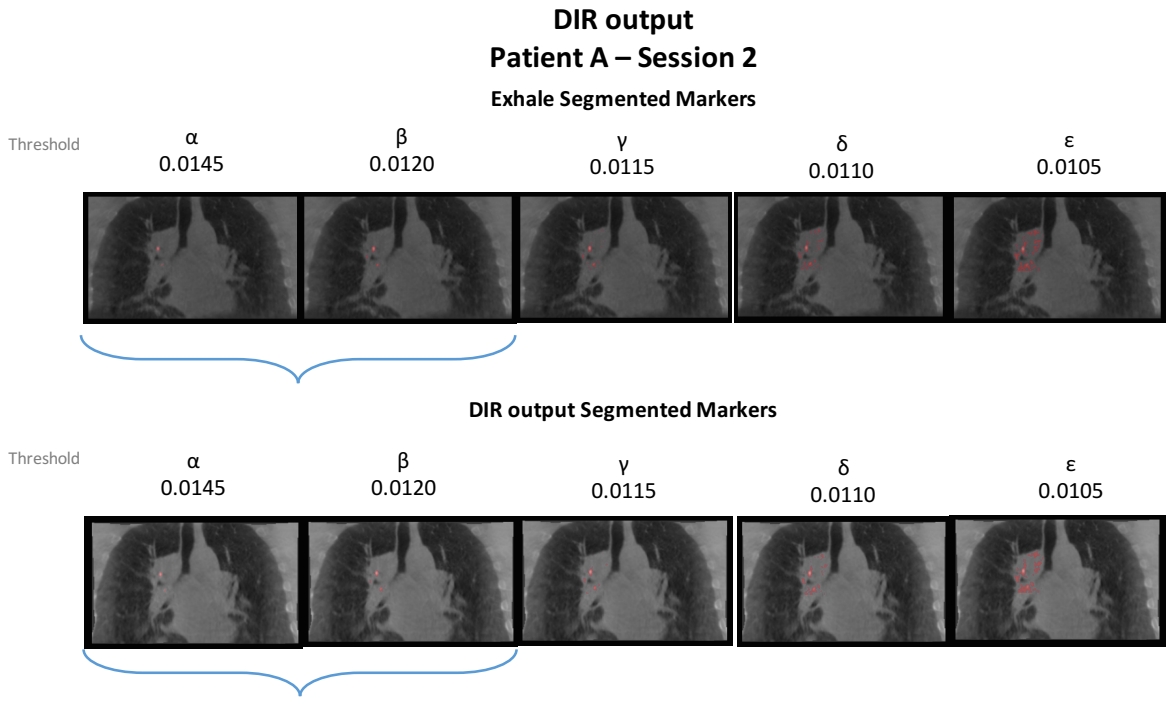

Figure 30. Comparison between different thresholds for patient A, session2. Volumes related to the DIR output.

In the figure above, it is seen that possible thresholds choices could be α and β. However, the best choice is α. This is given by the fact that a higher threshold represents a more adequately isolated and highlighted marker.

The next figure shows the same procedure for the third and last session of patient A. The comparison displays that the best choice is the threshold labeled as α.

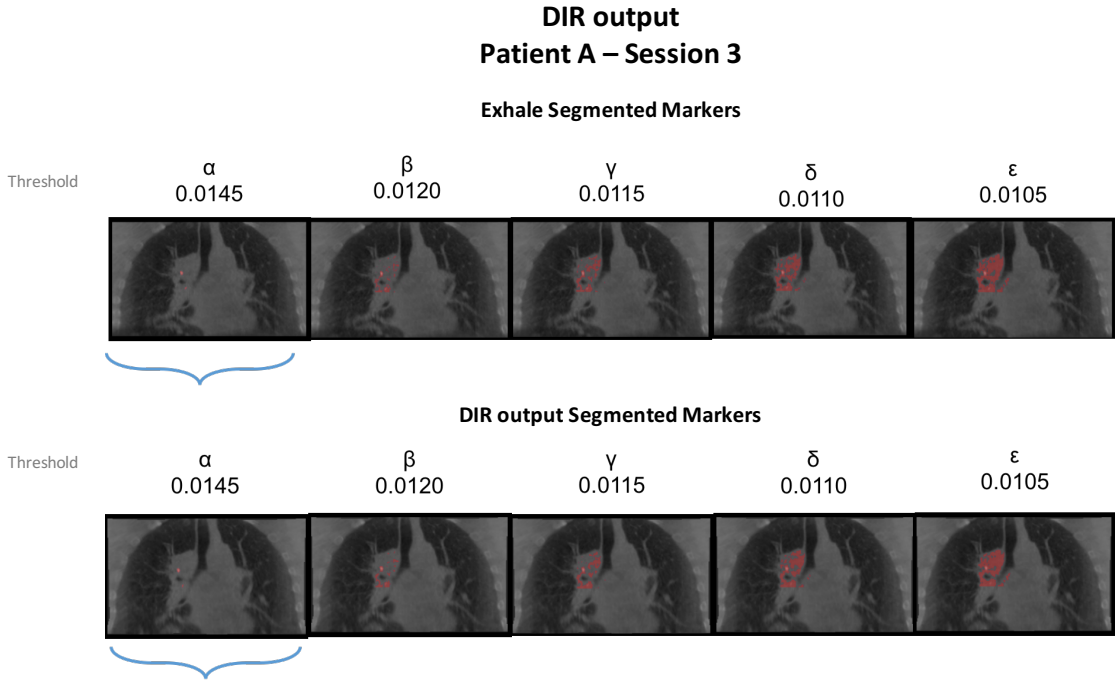

Figure 31. Comparison between different thresholds for patient A, session 3. Volumes related to the DIR output.

The threshold comparison that has been done so far corresponds to the markers found within the different volumes that are going to be used for measuring the local error for the image registration process. Given that it is also required to measure the motion model error, it is then needed to perform the same threshold comparison for the marker segmentation process carried out for the volumes involved in the motion model, which are the ground truth and the estimated volumes.

Figure 32 shows this previously mentioned threshold comparison for the volumes involved in the motion estimation for the first session of patient A.

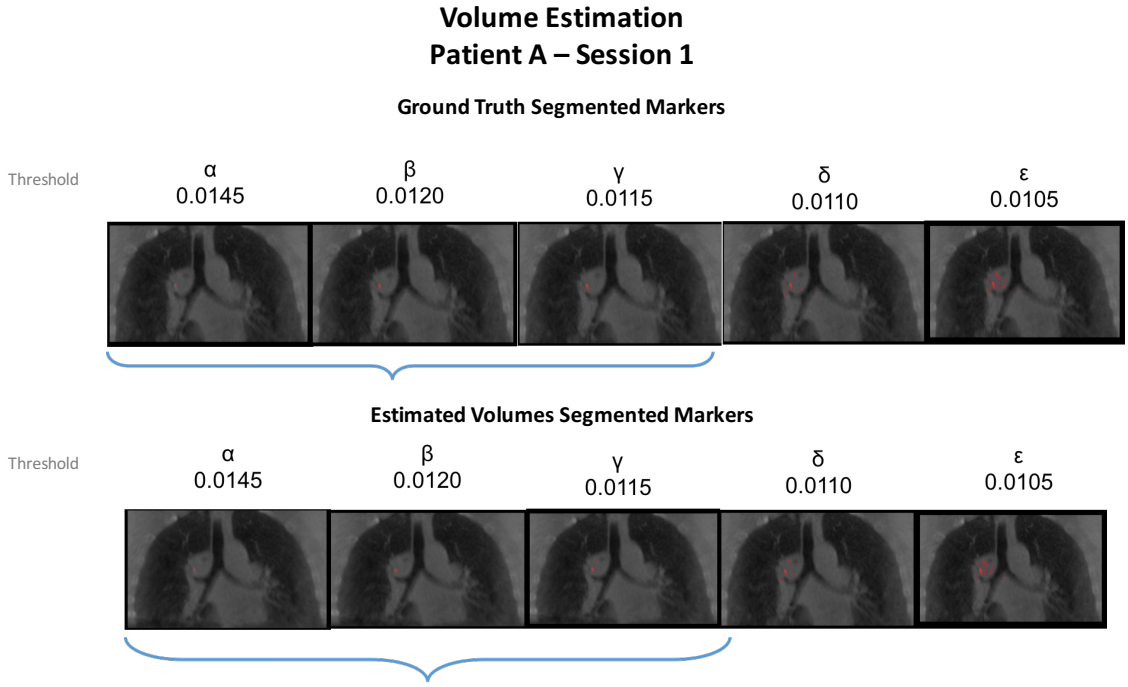

Figure 32. Comparison between different thresholds for patient A, session1. Images related to the volume estimation process

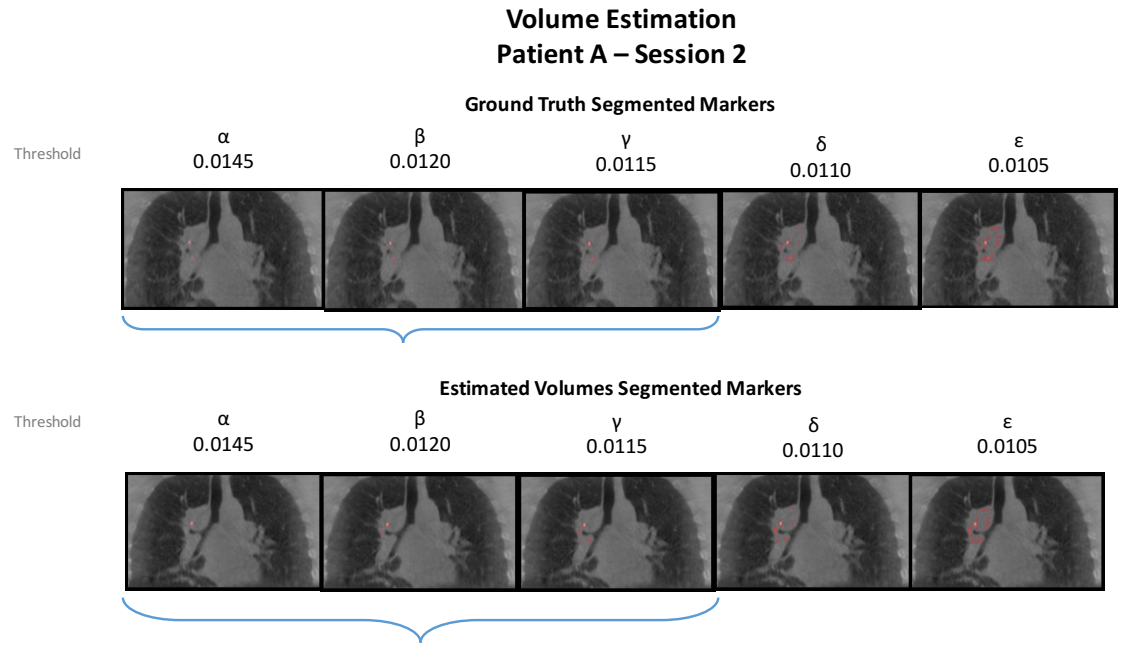

Figure 33. Comparison between different thresholds for patient A, session 2. Images related to the volume estimation process.

The previous two figures show the results for the volume estimation for the first and second sessions of patient A. It can be seen that even though the thresholds labeled as α, β, and γ,

are performing all right, the optimal output is obtained when selecting the threshold α. This is given by the fact that the higher threshold will always be the one that isolates the marker with the highest precision (as seen in the previous subsection), then, the preferred threshold will always be α. Another consideration for this choice is that, after reviewing all the results for the different thresholds, sessions, and patients, it was required, in order to hold consistency, to choose the same threshold value for all the sessions and patients. The previous and the following figures will support the fact that, the best threshold for all of the data is α. More details will be given in the following subsections.

The next image shows how, for the third session of the current patient, threshold α is the one providing the best segmentation output.

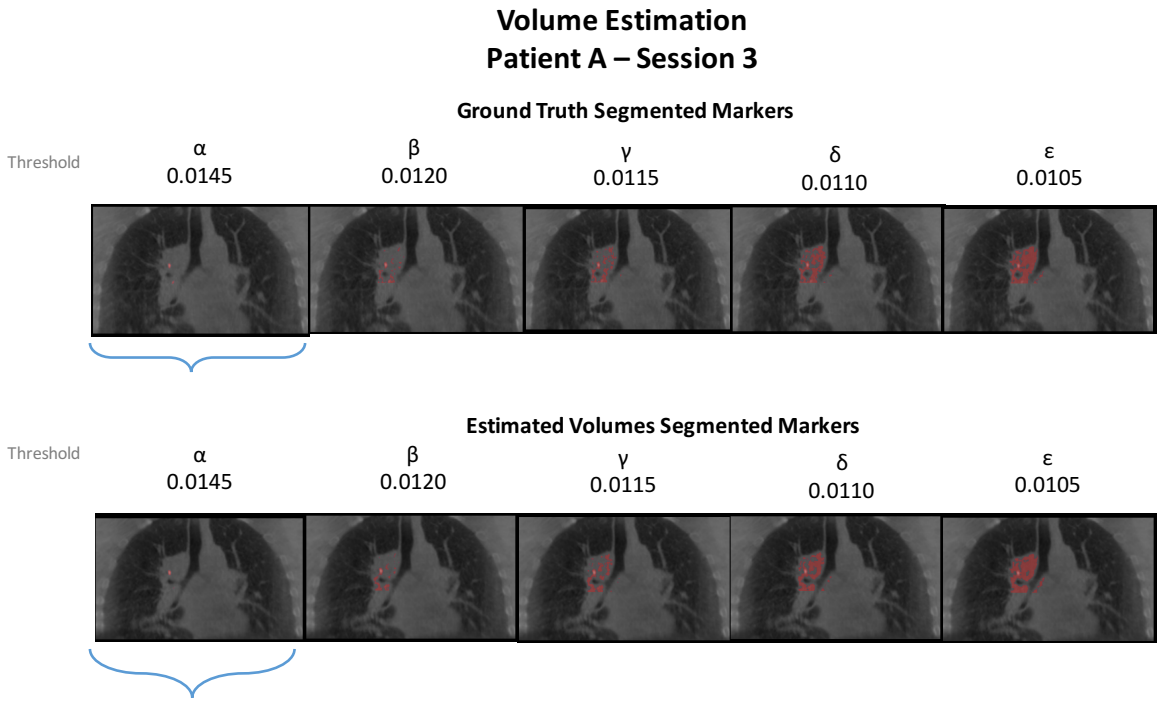

Figure 34. Comparison between different thresholds for patient A, session 3. Images related to the volume estimation process.

After performing the analysis for patient A, for both DIR output and volume estimation output, the results can be summarized as follows: with respect to the marker segmentation for the registered images, it can be seen that sessions 1 and 2, display a good output when selecting thresholds α, β, and γ. However, for session 3, threshold α is the only one that is

displaying an acceptable output. Likewise, concerning the marker segmentation for the volume estimation process, the thresholds are providing the same results as for the DIR output, i.e., sessions 1 and 2 have good outputs for more than one threshold but the third session has only one good result, and it arises only with threshold α. Consequently, since threshold α is the one for which all of the segmenting processes are performing correctly, it is then said that the best threshold for patient A is threshold α.

A similar analysis was also performed for patient B, and C. For patient B the selected threshold was α for both session 1 and 2. Nevertheless, the situation was different for patient C because it was not possible to use threshold α for session 2: this threshold did not return any data; consequently, it was decided to segment the marker of session 2 using a lower threshold β. The rest of the sessions made use of threshold α.

3.1.2) Quantitative Results

To perform quantitative evaluations, the general workflow resulted as shown in Figure 35:

- Before DIR: The first row of images displays the results of an overlay between the exhale phase, in green, and a moving phase, in violet, before performing the corresponding DIR.
- After DIR: The second row shows an overlay between the reference exhale volume (green) and the output of the DIR (violet). The accuracy of the DIR method can be visually assessed by comparing this row with the previous one (Before DIR).
- Warp: The third pair of images displays the output of the warping operation. The tumor is highlighted in red over the CBCT volume in green thanks to the image fusing process.
- Marker: In the fourth row, the implanted marker can be easily identified in red thanks to the marker segmentation technique that was performed throughout this work. This image is a fusion between the segmented marker and the CBCT volume.
- Prediction: This last row is an overlay between the ground truth, in green, and the estimated volume, in violet, which allows assessing the accuracy of the motion model visually.

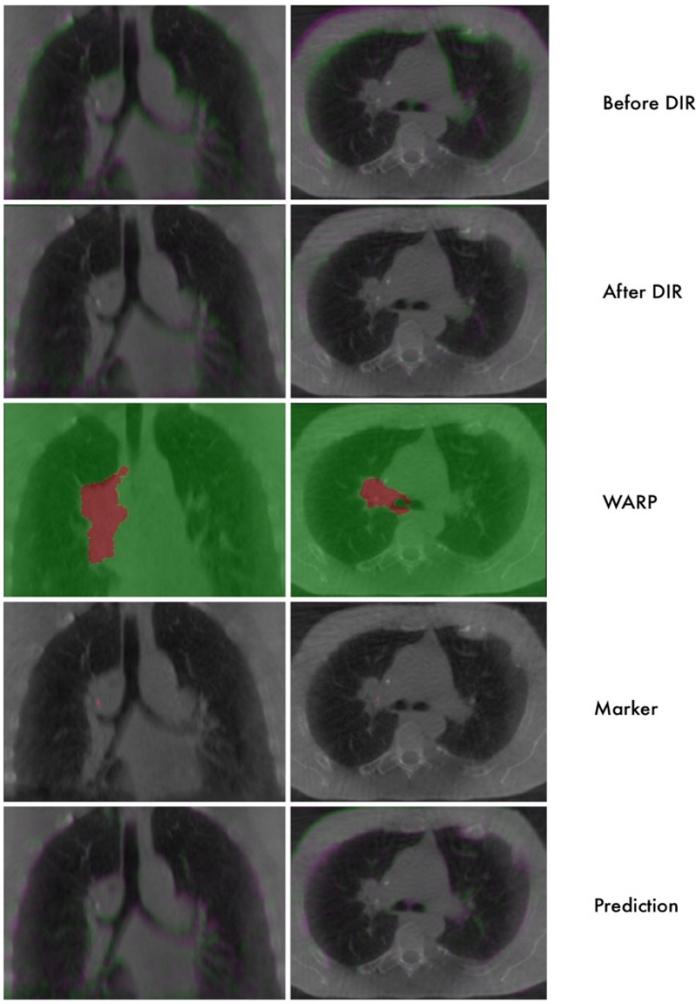

Figure 35. Summary of results for Patient A, Session 1.

To quantitative evaluate the perfromace of the registration and of the global motion model, the distance between the segmented markers was computed. This results will be organized in tables. The general layout of these tables is as follows:

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average	RMSE Std. Deviation
x											
y											
z											
RMSE											

Table 2. General layout used for displaying the local error.

The row displaying the text *Phase XX* refers to the phase of the respiratory cycle. As explained in previous chapters, 4D CBCT were represented by 10 respiratory phases, labeled

as 01, 02, ... 10. The phase number 06, which is missing from the table, symbolizes the reference peak exhale phase, which is the most stable one. All of the other phases are also called the moving phases they represent a displacement with respect to the reference exhale phase.

Then, the first rows of the first column represent the 3D direction over which the distance between the markers was measured. The last row in that same column, which reads, *RMSE*, stands for the root-mean-square error computed for the three dimensions x, y, and z.

An ideal result would be in the order of the voxel resolution, which is 1x1x1mm for the 4DCBCT data.

Patient A – Session 1: Threshold α: 0.0145

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	0.36	0.39	0.04	0.00	0.23	0.79	0.53	0.09	0.38		
y	0.91	0.27	0.48	0.08	0.32	0.37	0.47	0.30	1.02		
z	0.50	0.30	0.50	0.62	0.39	0.41	0.36	0.48	0.24		
RMSE [mm]	1.10	0.56	0.70	0.62	0.56	0.97	0.80	0.57	1.11	0.78	0.23

Table 3. Local registration error for patient A, session 1.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	0.68	0.73	2.10	0.02	0.58	3.25	2.98	1.84	0.50		
y	2.21	0.21	1.48	0.43	0.38	1.88	2.08	1.50	1.35		
z	4.97	1.78	0.52	1.78	0.51	3.82	3.10	2.25	1.20		
RMSE [mm]	5.49	1.94	2.62	1.83	0.86	5.36	4.78	3.28	1.88	3.11	1.71

Table 4. Local error in estimated volume for patient A, session 1.

For the first session of patient A, it can be seen that the image registration process has a good performance. It is worth noting that it presents a couple of small peaks in the y-direction for phases 01 and 10, both are the farthest phases from the reference exhale phase; consequently, it is expected that they present a more significant error given that they are demanding more computation resources from registration algorithms and procedures.

On the other hand, the estimation error is greater than the registration error. Table 4 shows that the lowest estimation errors are measured in the phases which are close to the phase

number 06 (which is the exhale reference phase), after which the error suddenly increases. Given that the DIR error has a small magnitude throughout the respiratory cycle, it could be said that the different error magnitudes obtained for the estimated volumes could mean that the model is not able to totally compensate for motion in this session.

As a final remark, the above tables are also showing that both the average and standard deviation of the RMSE are higher for the estimation error than for the registration error.

Patient A – Session 2: Threshold α: 0.0145

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	0.37	1.25	0.33	1.41	0.66	0.69	2.82	0.75	2.37		
y	0.30	0.09	0.39	0.29	0.47	0.11	2.39	0.45	0.11		
z	0.13	0.16	0.19	0.05	2.01	0.31	3.80	0.15	0.03		
RMSE [mm]	0.50	1.26	0.54	1.44	2.16	0.76	5.30	0.89	2.38	1.69	1.51

Table 5. Local registration error for patient A, session 2.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	1.75	1.68	1.83	0.62	2.38	5.22	2.53	4.53	0.66		
y	0.64	0.38	0.51	0.76	0.22	0.85	1.77	1.90	1.19		
z	1.58	0.66	0.97	2.59	0.08	1.38	2.75	2.98	0.80		
RMSE [mm]	2.45	1.84	2.14	2.77	2.39	5.46	4.14	5.75	1.58	3.17	1.56

Table 6. Local error in estimated volume for patient A, session 2.

The scenario is slightly different for this second session of patient A. The registration error is maintained low for most of the respiratory cycle except for phase 8 in which a combination of errors in all directions produce an RMSE peak of 5.30, this might be a consequence of sudden lack of stability in the breathing pattern of the patient. It is worth noting that the DIR error is close to zero for some of the phases of this cycle, especially in the z-direction.

Concerning the estimation error, the first remark is that the error is higher in the x-direction, which is consistent with the DIR error; consequently, this is suggesting a correlation between DIR error and volume estimation error. Another observation is that after the exhale reference phase (number 06), phases 07, 08 and 09 are displaying an increased error, the biggest one showing up for phase 07, so it could be said that the remaining peak errors for phases 08 and 09 might reflect the model's attempt to adapt and correct to the previous peak seen for phase

07. The fact that the error decreases considerably for phase 10 might confirm that the model successfully corrected the instability that arose for phase 07.

The last note regarding this session is that the RMSE average is higher for the estimation error than for the DIR error, as in the previous case; nevertheless, the standard deviation of the RMSE is almost the same for both DIR and estimation errors. Consequently, it could be said that the model is working better for this session than for the previous one.

Following this, the computed error magnitudes acquired for the third session of the current patient will be exposed below.

Patient A – Session 3: Threshold α: 0.0145

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	0.04	0.17	0.60	0.16	0.16	4.20	0.97	0.02	0.19		
y	0.07	0.22	0.62	0.04	0.18	0.78	0.91	0.19	0.47		
z	0.12	0.08	0.01	0.08	0.97	2.67	1.49	0.53	0.82		
RMSE [mm]	0.14	0.29	0.86	0.18	0.99	5.04	2.00	0.57	0.96	1.23	1.54

Table 7. Local registration error for patient A, session 3.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	0.64	3.68	1.89	1.31	0.40	2.81	1.45	0.71	0.05		
y	1.33	1.39	1.12	0.55	0.25	1.42	1.55	0.77	1.38		
z	4.00	1.41	0.30	2.33	0.56	2.53	2.93	0.48	1.36		
RMSE [mm]	4.27	4.18	2.22	2.73	0.73	4.04	3.62	1.15	1.94	2.76	1.34

Table 8. Local error in estimated volume for patient A, session 3.

There are a couple of interesting observations concerning the final session of patient A:

First, the DIR error is again kept low for most of the phases of the cycle. Nonetheless, an error peak is showing up again, this time for phase 07. This phase is the very next cycle phase after the reference one (phase 06), which could be interpreted as a potential struggle that the registration algorithm is undergoing after processing the registration for the relatively less complicated reference exhale image. This peak error is showing up in the x-direction, which so far has been the most sensitive to the DIR process for this current patient.

Afterwards, it can be seen how following phase 07, the error in the registration decreases which reflects a good performance form the DIR algorithm.

The overall performance of the registration and the model were good, with errors of about 2 voxels.

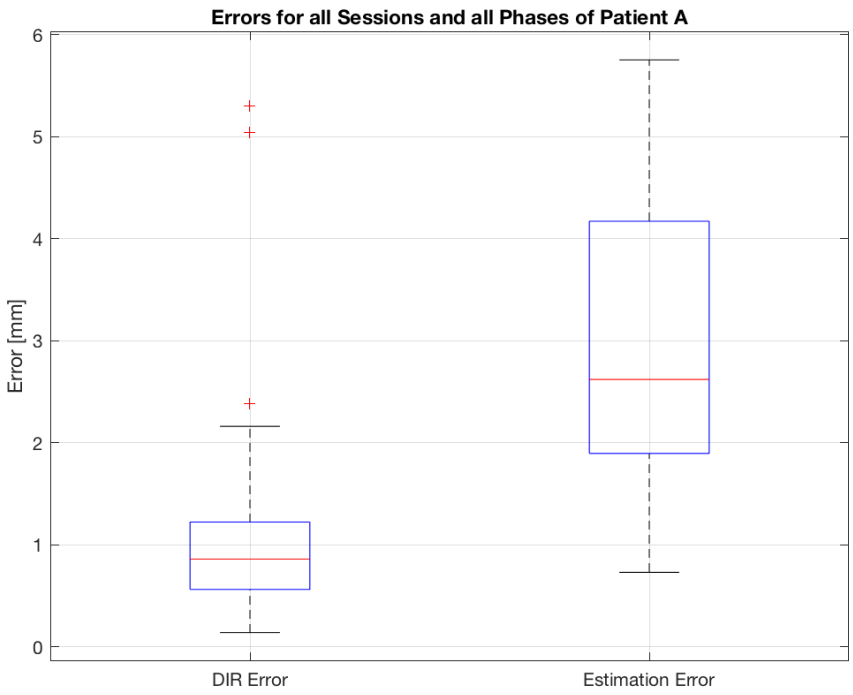

Figure 36. Overall performance for patient A.

Figure 36 is showing the boxplot for DIR output and volume estimation local errors for all sessions and all phases of patient A. Each one accordingly labeled on the x-axis. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol. The median and the interquantile range resulted higher for the global motion model performance than the DIR error. Overall the accuracy of the DIR is better than that of the motion model, nevertheless, taking into account the inherent DIR error, the model is performing reasonably good.

Next, the results for patient B will be displayed in the following tables and figures. It is worth noting that this patient only underwent two sessions as opposed to the previous patient for which more data was available.

Patient B – Session 1: Threshold α: 0.0145

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	4.63	3.66	3.31	1.41	0.40	0.31	1.00	3.80	3.64		
y	0.98	0.82	0.82	0.39	0.13	1.25	0.99	2.27	1.30		
z	0.14	0.15	0.00	0.19	0.33	1.51	0.93	1.56	0.19		
RMSE [mm]	4.73	3.76	3.41	1.47	0.54	1.98	1.69	4.69	3.87	2.91	1.52

Table 9. Local registration error for patient B, session 1.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	5.68	4.87	4.19	2.15	0.35	0.08	0.15	2.30	3.45		
y	1.30	1.16	0.96	0.44	0.09	0.96	0.92	2.31	1.60		
z	0.73	0.69	0.23	0.42	0.34	0.77	0.57	1.34	1.08		
RMSE [mm]	5.87	5.06	4.31	2.23	0.50	1.23	1.09	3.52	3.96	3.09	1.90

Table 10. Local error in estimated volume for patient B, session 1.

Regarding the DIR error, this session shows a steeper error in the x-direction, and also, the error is higher for the moving phases. The previous situation in which the error is higher for the phases which are farther from the reference phase (i.e. phase 06) is repeating for this session. As a consequence, the DIR error is lower for the middle phases, which correspond to the ones that are closer to the reference phase.

As to the estimation error, it is depicting a very similar trend that the one shown by the DIR error: A similar pattern and also the same tendency of miscalculating the x-direction. Hence, this estimation error seems to be highly influenced by the DIR output.

As the last comment for this session, it can be said that the RMSE's average and standard deviation are comparable for both DIR and estimation errors, which can be interpreted as a good performance of the model.

The following paragraphs expose the results for the second, and last, session for patient B.

Patient B – Session 2: Threshold α: 0.0145

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	4.03	2.79	3.20	1.18	0.36	0.43	1.24	3.02	2.67		
y	1.01	0.68	0.88	0.28	0.07	1.30	1.23	1.70	0.92		
z	0.26	0.21	0.10	0.16	0.18	1.33	0.74	0.57	0.00		
RMSE [mm]	4.17	2.88	3.32	1.22	0.41	1.91	1.90	3.51	2.83	2.46	1.19

Table 11. Local registration error for patient B, session 2.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	5.25	4.12	4.08	2.00	0.30	0.78	0.88	1.91	2.76		
y	1.62	1.04	1.13	0.36	0.06	0.57	0.55	2.23	1.70		
z	0.63	0.55	0.21	0.44	0.23	0.60	0.40	1.30	0.86		
RMSE [mm]	5.53	4.29	4.24	2.08	0.39	1.13	1.11	3.21	3.36	2.81	1.74

Table 12. Local error in estimated volume for patient B, session 2.

The results for this second session are very similar to those measured for session 1. An error in the x-direction is dominating the overall measures, and also a strong influence by the DIR is present in the volume estimation, which supports the previously made hypothesis regarding the impact that the DIR has in the motion modelling process.

Figure 37 reports the boxplot of the DIR and the volume estimation for patient B.

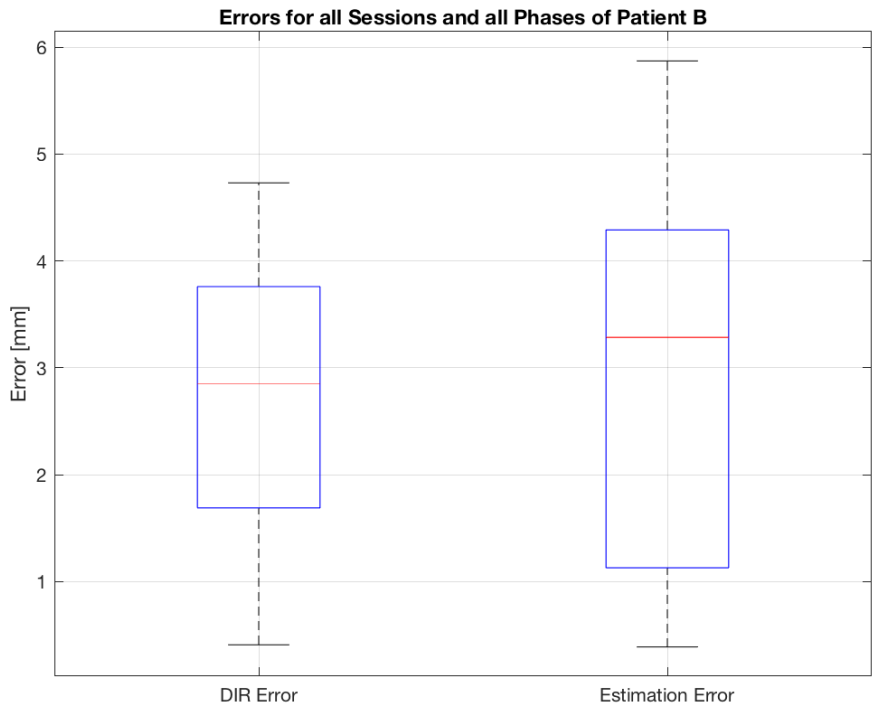

Figure 37. Overall performance for patient B.

The above plot is illustrating a dynamic relatively different for patient B than the one shown by patient A. Regarding the DIR error it has a median value of 2.85mm, which is also relatively close to the median value for the estimation error which is 3.28mm. A larger variability of the motion model performance was observed with respect to the accuracy of the registration. It can also be seen that the range for the 25th and 75th percentile is wider for the volume estimation error than for the DIR error, which is also expected.

Finally, the error metrics for the last patient, labeled as C are going to be displayed in the next subsection. Following that, a general evaluation regarding the results acquired for all of the patients will be presented.

Patient C – Session 1: Threshold α: 0.0145

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	1.28	4.22	4.15	2.27	0.46	5.82	3.07	0.25	4.85		
y	2.68	0.02	0.17	0.23	0.32	2.51	0.89	1.13	0.19		
z	0.73	1.99	1.74	0.80	0.24	1.23	0.91	0.25	2.15		
RMSE [mm]	3.06	4.67	4.51	2.42	0.61	6.46	3.33	1.18	5.31	3.50	1.92

Table 13. Local registration error for patient C, session 1.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	1.84	5.88	4.40	2.27	0.46	8.52	2.19	4.49	4.11		
y	1.99	0.27	0.15	0.23	0.32	3.73	0.69	2.62	0.37		
z	1.46	2.44	1.64	0.80	0.24	2.02	0.34	1.71	1.39		
RMSE [mm]	3.08	6.37	4.70	2.42	0.61	9.52	2.32	5.47	4.35	4.32	2.64

Table 14. Local error in estimated volume for patient C, session 1.

Given that the next session is making use of threshold β (it is the only session that is not using threshold α), the following considered an example of the impact on the accuracy of the results after employing a different threshold for measuring the distance between markers. Higher errors could be also attributed to a worst performance of the marker segmentation.

Patient C – Session 2: Threshold β: 0.0120

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	4.81	1.40	3.88	0.38	0.09	0.63	1.83	4.34	2.94		
y	0.02	0.01	0.09	0.24	0.16	0.08	0.41	0.56	0.38		
z	1.68	1.02	1.00	0.11	0.05	1.08	0.13	0.08	1.01		
RMSE [mm]	5.09	1.74	4.00	0.46	0.19	1.25	1.88	4.38	3.13	2.46	1.76

Table 15. Local registration error for patient C, session 2.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	4.83	2.01	3.86	0.34	0.16	0.83	1.24	3.52	3.24		
y	0.01	0.08	0.08	0.25	0.18	0.08	0.34	0.63	0.22		
z	1.48	0.88	1.00	0.12	0.06	1.08	0.30	0.22	0.66		
RMSE [mm]	5.05	2.20	3.99	0.44	0.24	1.36	1.32	3.58	3.32	2.39	1.68

Table 16. Local error in estimated volume for patient C, session 2.

The results for this session, show a relationship between the data similar to what has been seen so far:

- The model is replicating the errors present in the DIR.
- There is a high dependence of the model on the quality of the DIR, this is expected since the registered volumes are the primary inputs given to the motion model.

Patient C – Session 3: Threshold α: 0.0145

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	1.04	0.78	0.35	0.47	0.47	0.44	0.75	1.04	0.66		
y	1.12	0.86	1.00	0.53	0.31	0.41	0.28	0.35	1.11		
z	1.05	0.36	0.75	0.21	0.26	0.99	0.03	0.13	0.03		
RMSE [mm]	1.86	1.22	1.30	0.74	0.62	1.16	0.80	1.10	1.29	1.12	0.37

Table 17. Local registration error for patient C, session 3.

Direction	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10	RMSE Average [mm]	RMSE Std. Deviation [mm]
x	1.29	0.58	0.32	0.45	0.47	0.37	0.59	0.95	0.42		
y	1.18	0.91	1.00	0.52	0.31	0.44	0.12	0.18	1.09		
z	1.16	0.38	0.75	0.21	0.26	0.98	0.05	0.05	0.01		
RMSE [mm]	2.10	1.15	1.29	0.72	0.62	1.13	0.61	0.96	1.17	1.08	0.46

Table 18. Local error in estimated volume for patient C, session 3.

These last results are showing a tendency which is similar to what has been seen before: Higher values on the phases that are more distant from the reference phase, and a peak showing after the reference phase (which the model fails to compensate). Nevertheless, better performance in the DIR was achieved with respect to session 2, thus leading to better results for the volume estimation. This is easily confirmed by the decrease in the RMSE average and the standard deviation.

The figure below is illustrating the general behavior of the DIR and estimation errors for patient C.

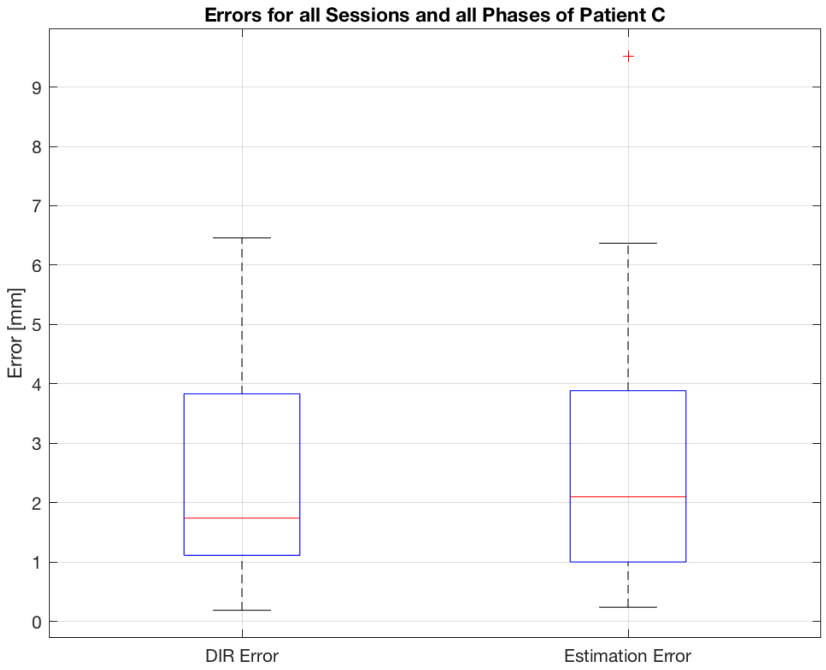

Figure 38. Overall performance for patient C.

The errors appraisal for patient C shows an outcome that is very similar to the results obtained for patients A and B, as reported by Figure 38. It should be highlighted that even though the DIR error reaches its maximum levels for this patient, the values for the estimation error are kept stable, which points out the robustness of the model.

3.1.3) Grouped Evaluation

A grouped assessment of the retrieved results is carried out in this final subsection. Below, Figure 39 is depicting the DIR and estimation errors for all phases of every session and every patient. The overall median error was 1.29 mm for DIR and 2.40 mm for the estimation of the motion model, attesting the error of both the approaches was below 3 voxels. This is comparable to the literature findings, where the error for global motion models is typically quantify around 2/3mm (Fassi et al. 2015, McClelland et al. 2013). The higher error for the motion model attested that model is replicating the errors present in the DIR. Moreover, the minimum values are near 0mm for both types of errors, with the estimation error slightly higher. The maximum values are also very similar for both DIR and estimation, and it is

worth noting that the highest value is equal to 6.46 mm and it corresponds to the DIR error category.

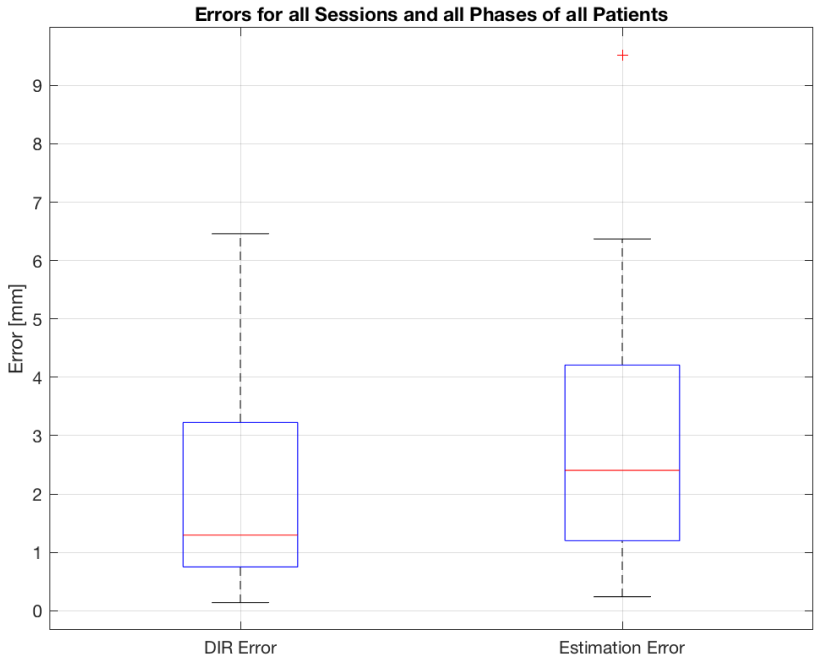

Figure 39. Errors for all phases of every session and every patient.

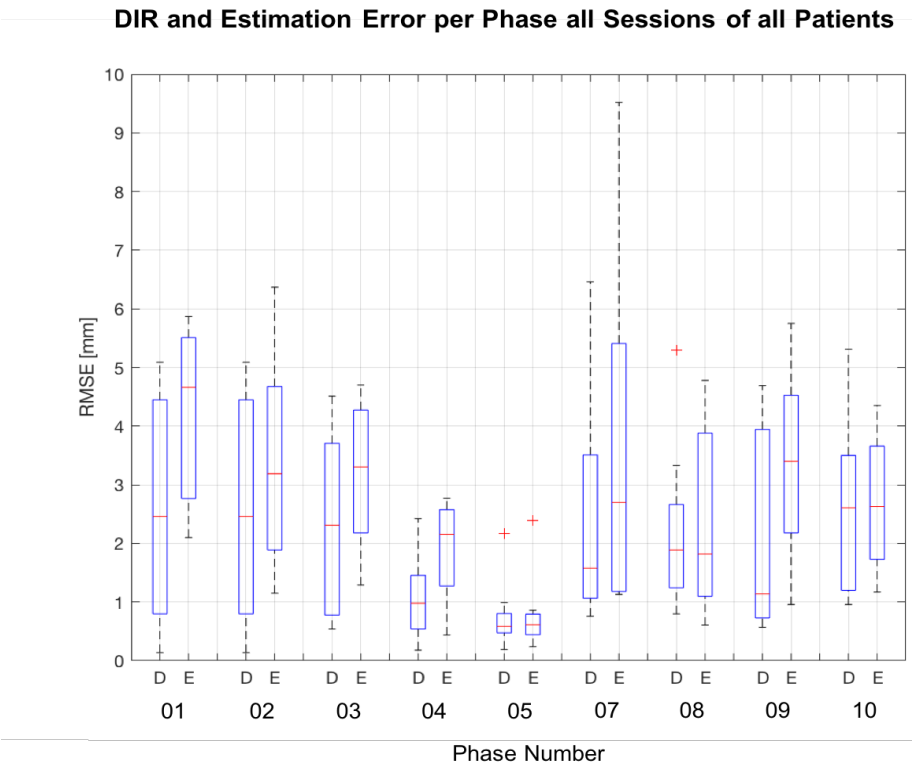

Figure 40. Grouped results for all Sessions.

Similarly, Figure 40 is showing the data for all of the patients and their sessions for each scenario is presented above. A clear influence from the DIR results is seen in in the volume estimation. Most of the cases will show that if an error is present in the DIR, it is most likely that the model is going to upscale this error. Therefore, it is fundamental to have a flawless DIR in order to have good results when predicting volumes. For this work, a high number of experiments were performed in order to find the best DIR parameters. The presented results were by far the best ones that could be found by the author after several tests. Nevertheless, the numbers and figures show that there is still room for improvement.

The final step regarding the assessment of the local error estimation is the implementation of a statistical test, particularly the Wilcoxon rank sum test. The output of this analysis establishes if two populations of data are significantly diverse between them or not. In this case, one population is going to be all of the data measured for the DIR output error (the data from all sessions and all patients); and the second population will correspond to all of the data measured for the estimation error. After performing the tests, the results point out that the two populations (DIR error and estimation error) are significantly different (Wilcoxon, α =5%).

3.2) Global Error

The arrangement for this category is the same as the one used for the local error section. The difference with respect to the previous section is that in here the measurement is made on a global scale, and also that there was no need to use any thresholds. Additionally, the median, range and the maximum were computed for all of the involved phases; these measurements were selected in order to provide a more intuitive view of the measured differences.

The ideal outcome would be a decrease in the global metric after the registration process; this would mean that the registration process is able to reduce the global differences between the images.

$$\text{Global Difference Before Registration} > \text{Global Difference After Registration}$$

The same applies to the global measure applied to the motion model, i.e., the metric before registration should decrease for the estimated volumes.

Global Difference Before Registration > Global Difference for Estimated Volumes

3.2.1) Patient A

3.2.1.1) Patient A – Session 1

	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10
Median	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Range	0.13%	0.12%	0.09%	0.07%	0.04%	0.06%	0.09%	0.11%	0.12%
Max	4.62%	4.50%	3.13%	1.52%	3.15%	3.98%	1.53%	4.59%	4.58%

Table 19. Global difference before image registration for patient A, session 1.

	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10
Median	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Range	0.07%	0.07%	0.06%	0.05%	0.04%	0.05%	0.07%	0.07%	0.07%
Max	3.48%	4.03%	1.86%	1.46%	3.15%	3.94%	1.36%	4.00%	3.87%

Table 20. Global difference after image registration for patient A, session 1.

	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10
Median	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Range	0.11%	0.11%	0.09%	0.07%	0.04%	0.07%	0.08%	0.09%	0.09%
Max	1.77%	1.78%	1.89%	1.90%	1.83%	1.65%	1.58%	1.53%	1.46%

Table 21. Global difference between the estimated volumes and the corresponding ground truth volume for patient A, session 1.

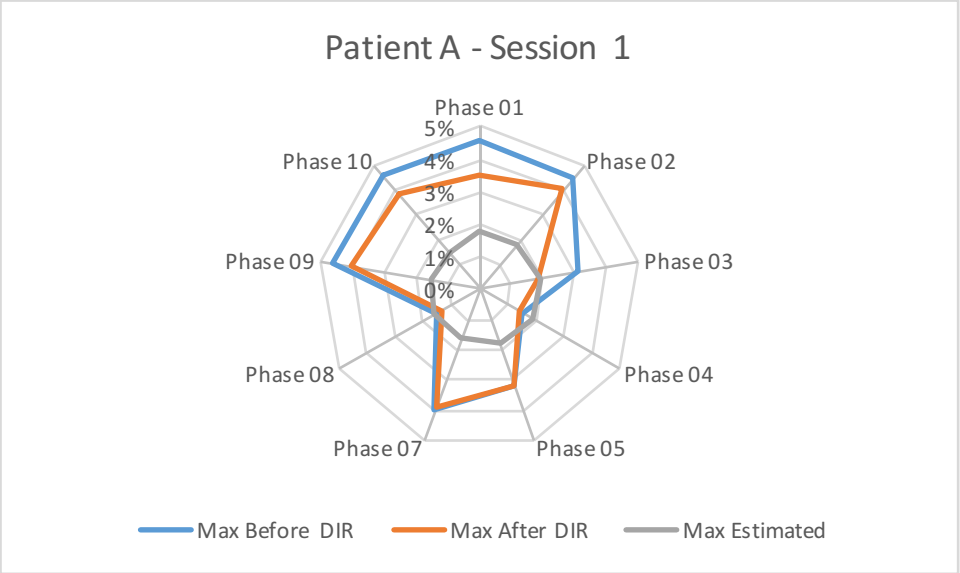

Figure 41. Comparision between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient A, session 1.

Regarding the relation between the maximum before DIR and the maximum measured for the volume prediction, it can be seen that for seven out of nine comparisons, the results behave in the same way as expected.

Similarly, when comparing the same measurements before and after DIR, it can be seen that the results perform as anticipated for nine out of nine measurements. Therefore, it can be said that the performance is good.

Also, the maximum differences computed for the predicted volumes are considerably lower than those obtained before registering the volumes. Regarding the before and after DIR errors, it is seen that the maximums for both quantities are similar. It is worth noting that the error present after DIR slightly decreases when compared to the error before the DIR.

Finally, the plot allows to see that the error is consistently higher for the moving phases of the respiratory cycle, i.e., those that are not occurring near the exhale phase.

3.2.1.2) Patient A – Session 2

	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10
Max Before DIR	2.52%	2.11%	1.84%	1.41%	1.93%	4.66%	2.07%	2.16%	2.15%
Max After DIR	1.79%	1.70%	1.71%	1.30%	1.93%	4.02%	1.47%	2.30%	1.60%
Max Estimated	1.84%	1.84%	1.81%	2.39%	1.87%	1.48%	1.61%	1.59%	1.56%

Table 22. Maximum global differences for patient A, session 2.

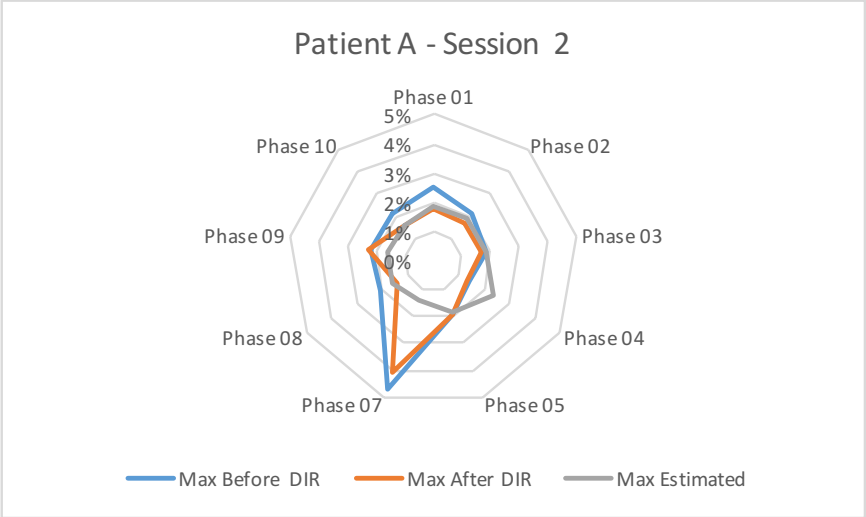

Figure 42. Comparison between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient A, session 2.

For this second session, it can be seen that regarding the global maximum error for the estimated volumes, 8 out of 9 phase comparisons are performing as expected, i.e., the error is lower for the estimated volumes when compared to the error before DIR. When making the same comparison for the errors before and after DIR, it is seen that 8 out of 9 results are showing the anticipated behavior.

3.2.1.3) Patient A – Session 3

	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10
Max Before DIR	3.61%	3.42%	1.99%	1.49%	3.02%	3.96%	2.02%	3.36%	3.98%
Max After DIR	3.05%	3.15%	1.69%	1.89%	3.01%	3.88%	1.63%	2.99%	3.61%
Max Estimated	2.19%	2.39%	2.37%	2.18%	2.94%	2.03%	1.54%	2.21%	1.73%

Table 23. Maximum global differences for patient A, session 3.

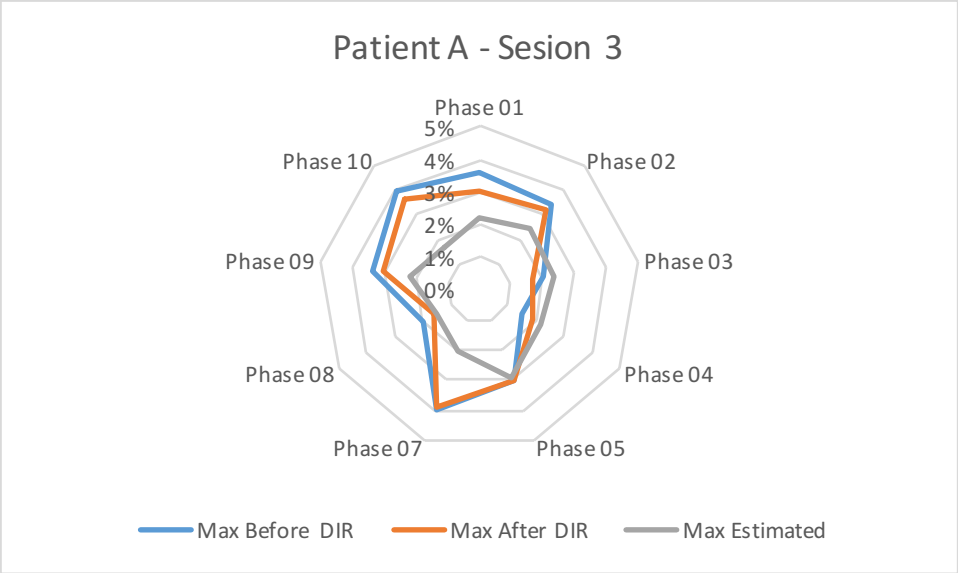

Figure 43. Comparision between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient A, session 3.

In this third session, the global maximum error for the estimated volumes is performing as anticipated for 7 out of 9 phases. The before and after comparison it is behaving as expected for 8 of 9 phases, and again phase 07 shows a higher maximum error for the measurements regarding the before and after DIR.

3.2.1.4) Comparison between the three sessions of patient A

In this subsection, an analysis of the three sessions of patient A is presented for the involved metrics.

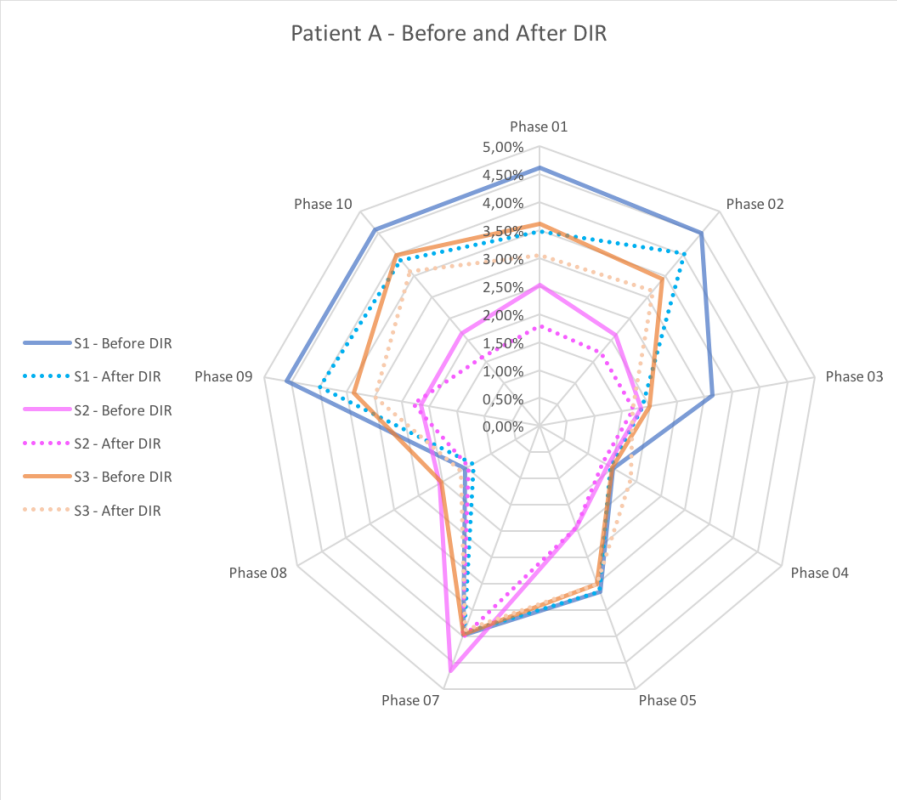

Figure 44. Comparison between before and after DIR global errors for the three sessions of patient A.

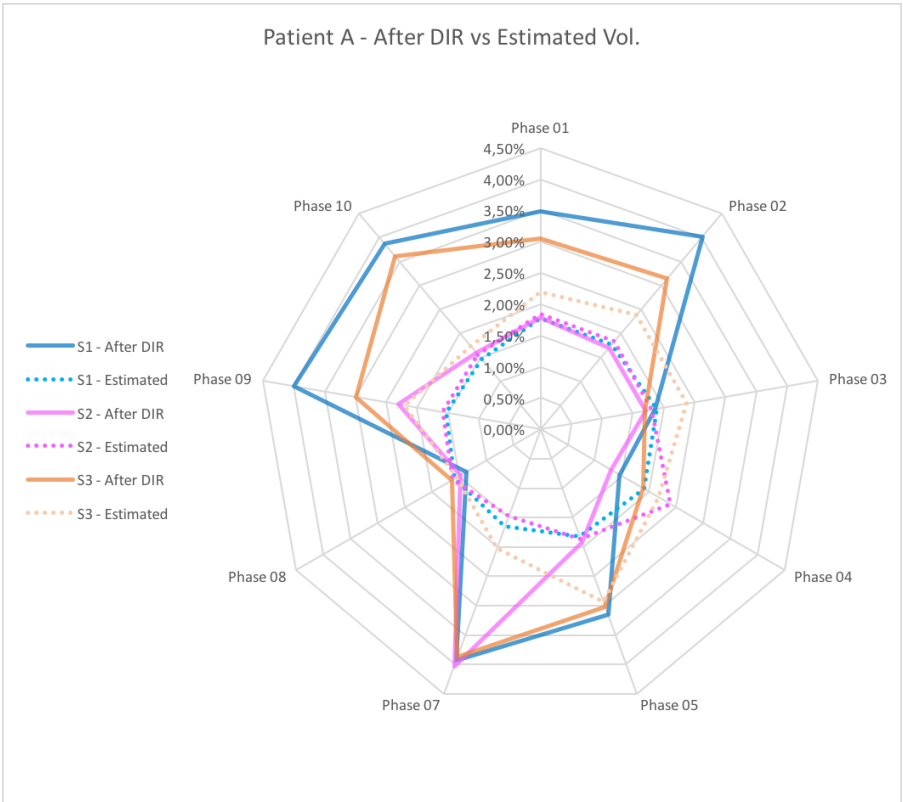

Figure 45. Comparison between after DIR and estimation global errors for the three sessions of patient A.

Regarding the before and after DIR comparisons, shown in Figure 44, it can be seen that the maximum global error decreased after performing the DIR, as expected. Also, session number 2 is the one with the highest performance, i.e., the lowest global error. All of the sessions show a peak error at phase 07 which is present for both before and after DIR.

With respect to the volume estimation, illustrated in Figure 45, sessions 1 and 2 show a comparable behavior, with a maximum global error of 2%. Session 3 displays an error trend that is limited, reaching a maximum of 3% for phase 05, just before the exhale motion.

3.2.2) Patient B

3.2.2.1) Patient B – Session 1

	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10
Max Before DIR	1.84%	1.57%	1.63%	1.61%	1.28%	1.48%	1.41%	1.77%	1.74%
Max After DIR	1.45%	1.41%	1.48%	1.07%	1.27%	1.47%	1.82%	1.34%	1.30%
Max Estimated	1.70%	1.72%	1.89%	1.63%	1.14%	1.77%	1.53%	1.76%	1.54%

Table 24. Maximum global differences for patient B, session 1.

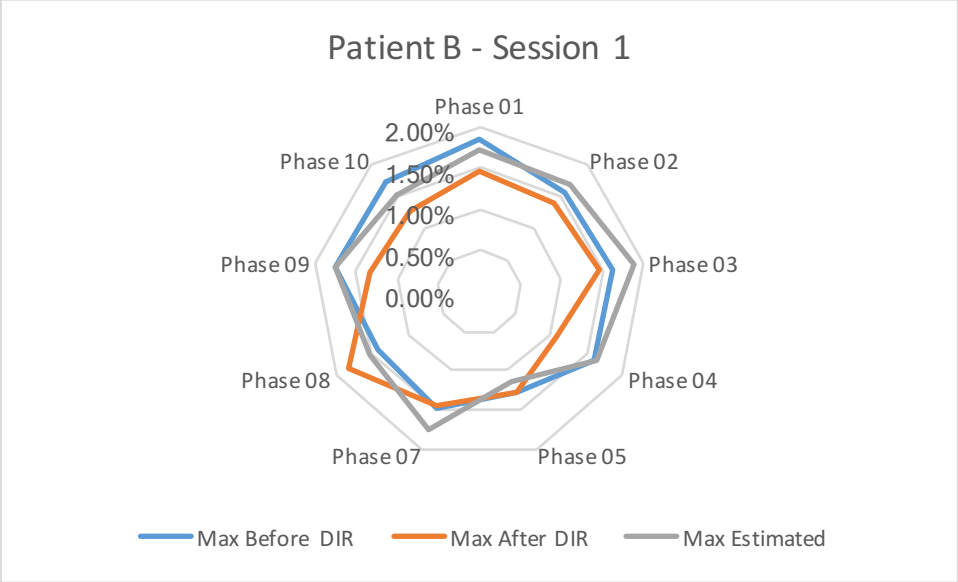

Figure 46. Comparision between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient B, session X.

With respect to this first session of patient 2, the comparison between the maximum error after DIR with the maximum error before DIR shows up that 8 of 9 phases are performing as expected, this means that there are lower values for the measures performed after the DIR. Overall the results are good for the global DIR metric, and on the other hand, the global estimation displays a restrained error, always below 2%.

Regarding the estimation error, these values reflect that the trend for 4 of 9 phases are behaving as expected, in those 4 phases the estimated error is lower than the before-DIR error. For 2 of 9 phases, these comparisons present very similar values, i.e., the maximum between the two metrics are very close. Additionally, for the remaining 3 of 9 phases, the estimated error maximum is higher than the values shown before performing the DIR.

3.2.2.2) Patient B – Session 2

	Phase 01	Phase 02	Phase 03	Phase 04	Phase 05	Phase 07	Phase 08	Phase 09	Phase 10
Max Before DIR	1.75%	1.60%	1.59%	1.16%	1.45%	1.25%	1.24%	1.47%	1.59%
Max After DIR	1.55%	1.20%	1.24%	0.89%	1.45%	1.22%	1.61%	1.06%	1.21%
Max Estimated	1.49%	1.70%	1.42%	1.24%	1.08%	1.71%	1.70%	1.68%	1.32%

Table 25. Maximum global differences for patient B, session 2.

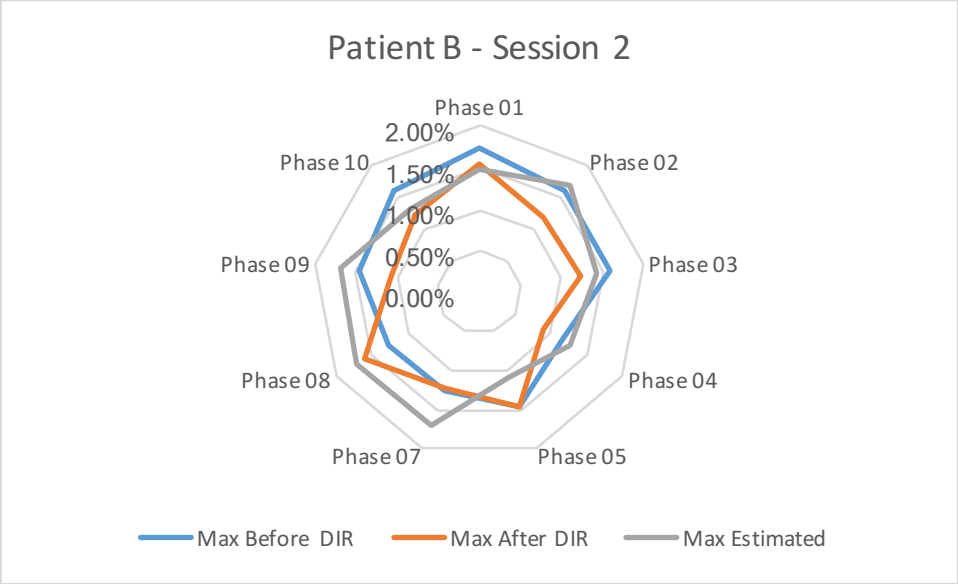

Figure 47. Comparision between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient B, session X.

The results for this second session are painting a similar picture as before, for the volume estimation 4 out of 9 phases are providing good results, 2 of 9 phases show a similar error before DIR and after performing the volume estimation, and 3 of 9 phases have a greater estimation error than its counterparts measured before registering the images. Most of the errors can be seen in the phases that are more distant from the reference phases which has also been seen when performing local error mesures. For this particular case, it is seen that a peak in the DIR error for just one phase is heavily influencing the performance of up to three phases of the volume estimation. Even though the estimation errors that are showing up for this phase, it is clear that image registration is critical for obtaining accurate outputs from the motion model.

Regarding the DIR metric, outstanding performance is seen for 8 out of 9 phases present the expected results (as in the previous session of this patient). An error peak is arising for phase 08, which unfortunately exerts a relevant influence on the estimation results as previously highlighted. The overall error for both of the metrics is kept below 2% which represents a fair output.

3.2.2.3) Comparison between the two sessions of patient B

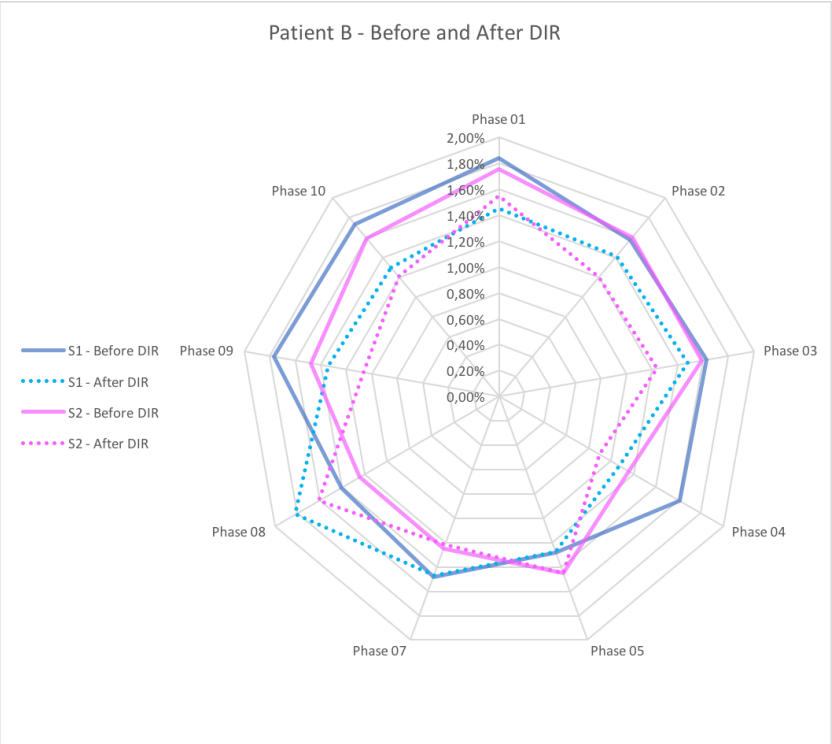

Figure 48. Comparison between before and after DIR global errors for the two sessions of patient B.

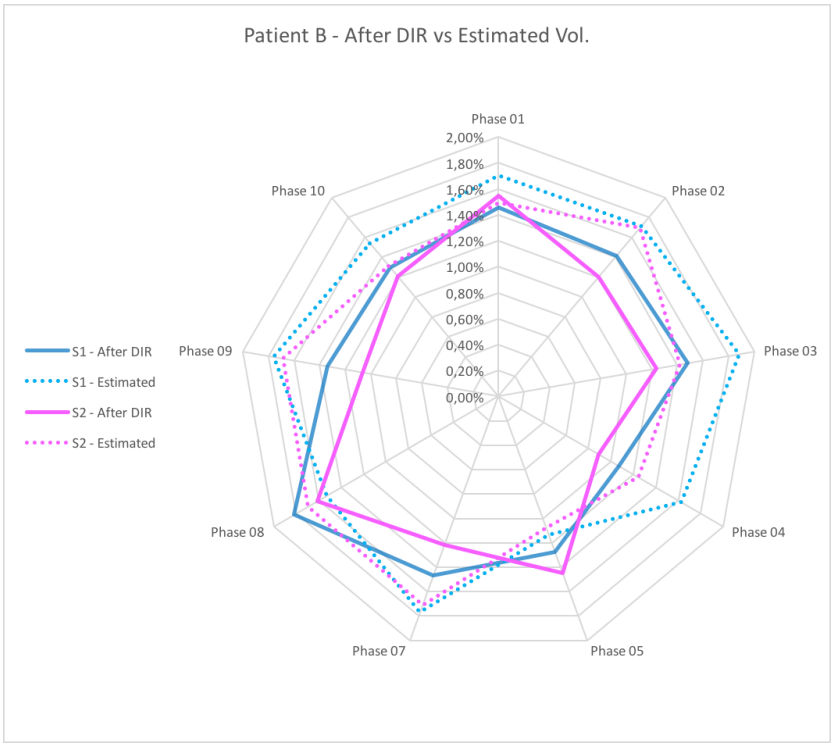

Figure 49 Comparison between after DIR and estimation global errors for the two sessions of patient B.

When comparing the behavior of the metrics for all of the sessions, it is within sight that session two is displaying a slightly better performance for all of the measured magnitudes. Also, the error is continuously kept under 2% for all of the measured conditions.

3.2.3) Patient C

3.2.3.1) Patient C – Session 1

The obtained results for the three sessions of patient C are displayed below. After session 3 a general assessment will be present for this patient.

	01	02	03	04	05	07	08	09	10
Max Before DIR	3.00%	3.06%	2.50%	2.42%	1.58%	3.48%	1.96%	3.09%	3.47%
Max After DIR	2.28%	2.85%	2.44%	2.41%	1.58%	3.47%	1.95%	2.85%	3.23%
Max Estimated	3.23%	3.37%	2.46%	2.37%	1.83%	2.59%	1.72%	2.82%	3.19%

Table 26. Maximum global differences for patient C, session 1.

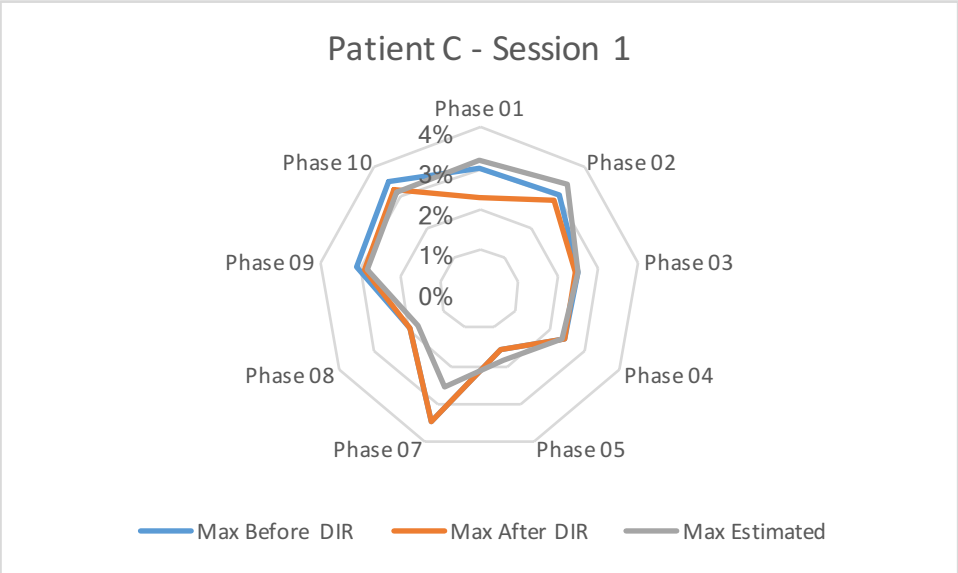

Figure 49. Comparision between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient C, session 1.

3.2.3.2) Patient C – Session 2

	01	02	03	04	05	07	08	09	10
Max Before DIR	2.16%	2.60%	1.89%	1.57%	1.83%	3.10%	1.24%	2.54%	2.96%
Max After DIR	2.01%	2.50%	1.88%	1.57%	1.82%	3.10%	1.24%	2.50%	2.84%
Max Estimated	3.00%	2.98%	2.53%	2.05%	1.57%	2.56%	1.76%	3.15%	3.14%

Table 27. Maximum global differences for patient C, session 2.

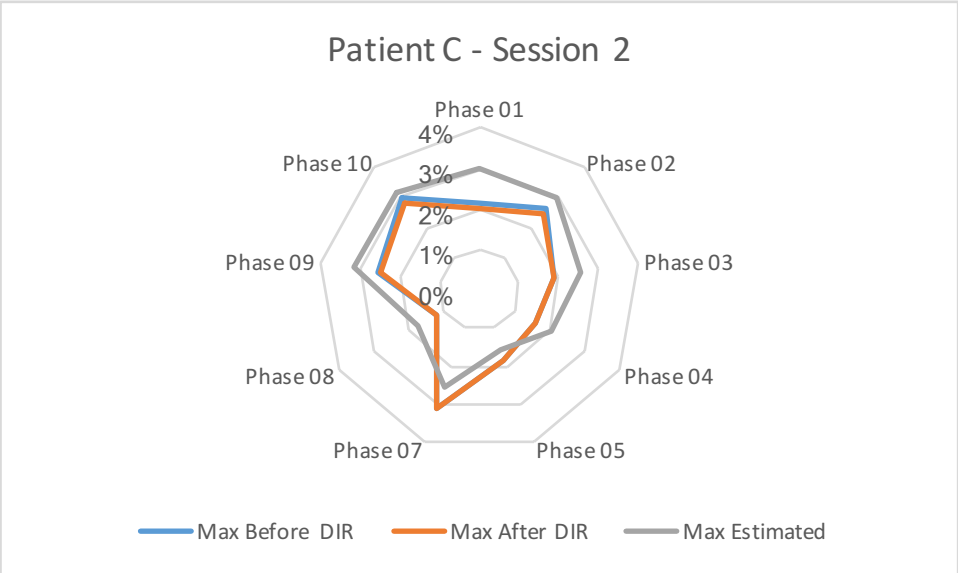

Figure 50. Comparision between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient C, session X.

3.2.3.3) Patient C – Session 3

	01	02	03	04	05	07	08	09	10
Max Before DIR	2.37%	2.52%	1.98%	1.85%	1.64%	3.60%	1.45%	2.21%	2.84%
Max After DIR	2.16%	2.42%	1.97%	1.85%	1.64%	3.59%	1.45%	2.16%	2.65%
Max Estimated	2.09%	2.55%	1.73%	1.66%	1.81%	3.16%	1.56%	3.03%	2.68%

Table 28. Maximum global differences for patient C, session 3.

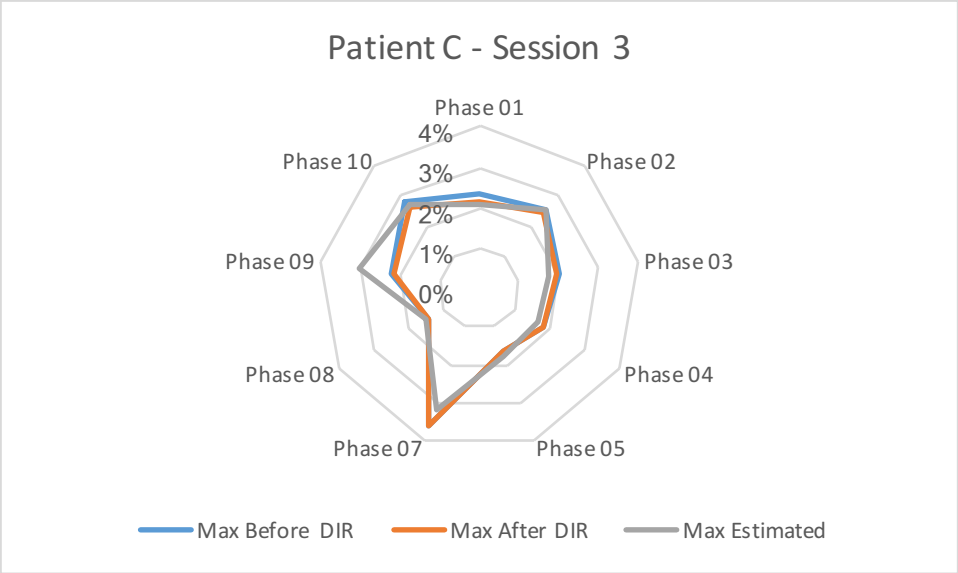

Figure 51. Comparision between the maximum values for the estimated volumes (gray), after DIR (orange), and before DIR (blue), for patient C, session 3.

For the first session of patient C, the estimation error is performing correctly for 6 of 9 phases. The after DIR error is lower when compared to the before DIR error for 7 phases, which is also beneficial. Again, a peak in the estimation error follows a peak in the global DIR error in phase 07, confirming that errors in image registration will show up in the volume estimation.

Regarding session 2, it is discernible that the before and after DIR errors have a very similar trend. Even though the DIR error is slightly lower, this similarity is profoundly influencing the performance of the motion model.

Concerning session 3, it is shown that the estimation error is performing correctly for 5 of 9 phases. Also, its behavior is not the preferred one for 3 of 9 phases. On the other hand, the DIR error is found to be slightly lower for 4 phases and very similar for the remaining cycle phases. So, for this case, it is visible that there is not that much variation between the difference computed before and after registering the images. The relevance of this is that this behavior translates into a decreased performance in the prediction capabilities of the model: it is seen how the trend of the three metrics is almost identical, with a slight peak on phase 09 for the global estimation error, which is just following a smaller peak that is found for both before and after DIR.

3.2.3.4) Comparison between the three sessions of patient C

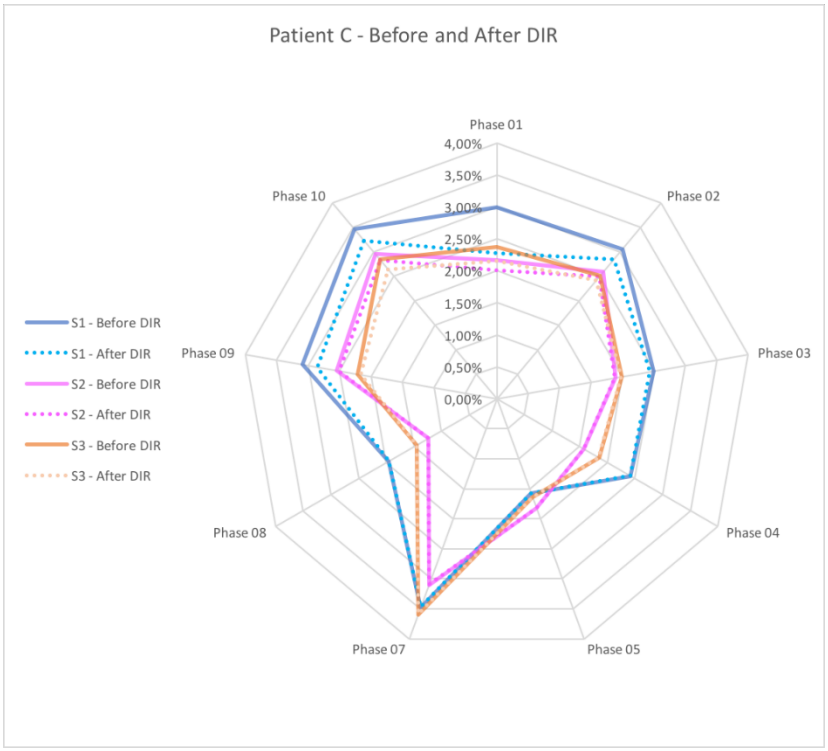

Figure 52. Comparision between before and after DIR global errors for the three sessions of patient C.

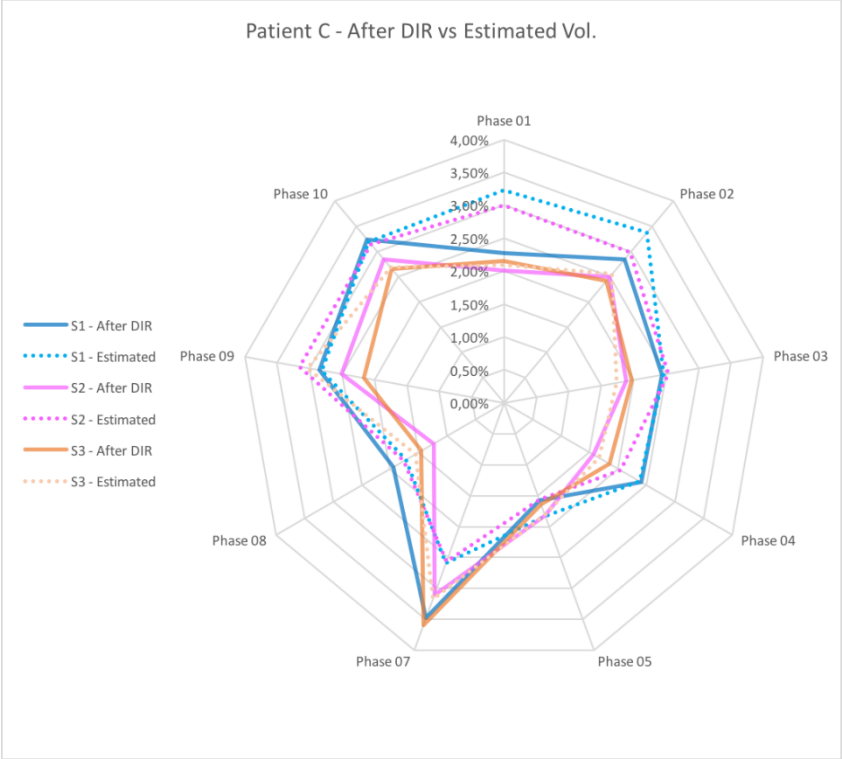

Figure 53. Comparison between after DIR and estimation global errors for the three sessions of patient C.

When evaluating the general performance of the metrics for all of the sessions of this patient it can be seen that session 1 is the one with the weakest execution.

Then, the analysis shows that sessions 2 and 3 present a trend that is very similar between them, over which they depict a decreased error when compared to the first session.

It is worth noting that the volume estimation is showing a decreased error for session 3 even though the error for this last session was regarded as a not ideal performance. However, when comparing it with the previous sessions, it can be seen that it is the one with the lowest magnitude, so as a consequence, its trend outlines a good result when compared to the global estimation errors found for the rest of the sessions.

3.2.4) Grouped Evaluation

This final section will present an evaluation of all of the results obtained for the measurement of the global difference between all of the involved volumes. For this, a boxplot condensing all of the generated data is shown below.

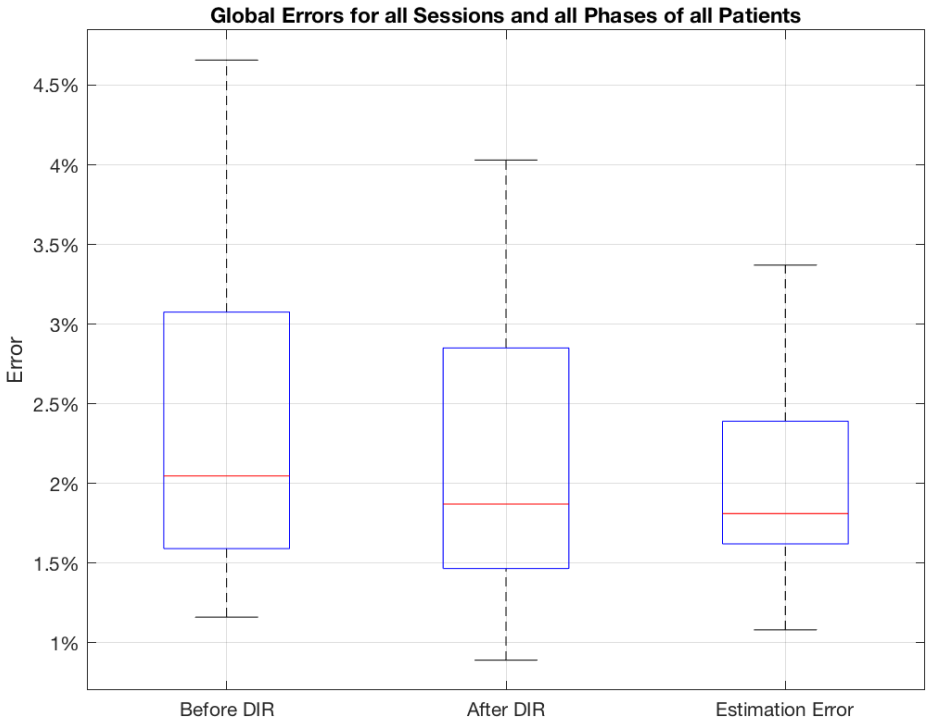

Figure 54. Maximum error percentage for the three patients, their sessions, and all the respiratory phases. The x-axis is describing the following data per each phase: before image registration, after image registration, and estimated volume.

The boxplots in Figure 54 are showing the measured global errors obtained before and after registering the images, and also the global error in the estimated volumes. This plot is condensing the data of all of the respiratory cycles. It is noticeable that the global DIR error is lower than the before DIR error, which is the expected result.

Similarly, it is visible that the performance of the motion model is reasonably good: its median is similar than the median for the after DIR data, and its varaibility is kept close to values similar to the DIR error. The model performance shows that the estimated volume is able to compensate for motion, with errors within the DIR performance values.

The last assessment for the global error measurement is the implementation of a statistical test, which is the same as the one performed for the local error measurement, i.e., the Wilcoxon rank sum test. Recalling, the output of this analysis establishes if two populations of data are significantly diverse between them or not. This test was performed three times:

- In the initial test, the data involved was the before and after DIR global error. So one of the populations was the error measured before implementing the DIR, and the second population was the after DIR error data. The outcome of this test was that the two populations (before and after DIR errors) are not significantly different (Wilcoxon, α =5%). However, a decreased median error was observed for after DIR.
- The second run of the test was carried out using the after DIR error data as one population, and the volume estimation error data as the other population. The analysis pointed out that the two populations (after DIR and estimated errors) are not significantly different (Wilcoxon, α =5%).
- For the third and last time, the before DIR error was taken as one population, and the estimation error was taken as population two. The output of this test resulted showed that the two populations (before DIR and estimation errors) are also not significantly different (Wilcoxon, α =5%).

Chapter 4. Conclusions

Currently, radiation therapy is one of the preferred strategies to manage and control tumoral lesions. This method delivers a high amount of energy to the patient, which could potentially cause secondary lesions. Therefore, high precision is required if this procedure is intended to be successful. In order to fulfill this requirement, radiotherapy is often combined with other image modalities in an attempt of performing a safe dose escalation and improve the patient treatment, this methodology is known as Image-Guided Radiotherapy, IGRT.

IGRT is usually influenced by organ motion, which decreases the accuracy and thus the efficacy of a given treatment. Motion models allow to overcome this obstacle when implemented in IGRT procedures; these models are classified into local and global motion models. They work through the use of image registration techniques which allow quantifying the movement of the patient's organs, and they are also helpful in treatment design and setup. Even though image registration is fundamental for IGRT, an opportunity area in image registration methodologies is their validation, given that they are no standard approaches to assess the outcome of the registration procedure.

This work aimed to validate the tools that are used to perform image registration and motion modeling in 4DCBCT data. First, a deformable image registration was implemented between CBCT respiratory phases to derive the motion field (the exhale phase was kept as reference). To assess the accuracy of this operation, the registered output images were compared against the unregistered images and the difference between both was measured locally and globally. Following that, a global motion model was applied based on 4DCBCT motion fields data in order to estimate one of the respiratory phases in a leave-one-out approach. The estimated output of the motion model was compared with the ground truth data, also in a local and global approach.

Concerning the local error, the following observations can be concluded:

- A higher error is arising from the moving phases, which are the ones that are farther from the reference exhale phase, labeled as 06. Likewise, lower errors were obtained for the phases that are closer to the reference phase.

- Moreover, it was seen that errors that were showing up in a specific respiratory phase or anatomical direction of the DIR, were also present in the volume estimation. Therefore, it can be said that the volume estimation is highly influenced by DIR error, which was expected given that the DIR is the primary input to the model.
- Besides, it was observed that the model was sometimes capable of compensating for the DIR errors: it yielded, for some phases, a lower estimation error than the DIR error, like in the case of patient C. This highlights the robustness of the model.
- Furthermore, the error for both DIR and estimated volumes was below three voxels (i.e. <3mm), the minimum values were near 0 for both metrics, and the maximum value was 6.46 which corresponded to the DIR error. Consequently, it could be concluded that a good DIR will produce a good volume estimation. Overall, similar results were observed for the three patients and confirmed by the literature findings (i.e. median errors around 2/3mm).
- However, it should be noticed that marker segmentation was performed by finding qualitatively the best threshold to mask the data, and this could affect the derived local errors.

Regarding the global error, the following relevant points are presented:

- Globally, the estimation differences are lower than the DIR differences, which in turn are lower than the before DIR measurements; this confirm that the DIR is able to globally account for motion.
- Again, as for the local error, the global error is higher for the phases that are further from the reference phase, and also, the DIR error is influencing the estimation error also for this assessment. Consequently, it can be concluded that the image registration is critical for obtaining accurate outputs from the motion model.
- All over, the performance of the motion model is accurate: its median is similar than the median for the after DIR data, and globally the model is able estimate the specific respiratory phase.

Taking into account the intrinsic noise present within the CBCT data, it should be emphasized that the errors obtained are similar to the voxel resolution, even though this

amount of noise limited the accuracy of the overall image processing performed throughout this work.

The future developments regarding this work are enlisted as:

- Regarding the image registration process, an area of opportunity for future developments would be to perform even more extensive and inclusive research on algorithms for image registration and their parameters.
- The marker segmentation needs a proper validation and markers should be segmented by a clinical expert. During this work, four different thresholds were tested, and the selected threshold was chosen upon qualitative criteria. Given that the local error is influenced by the accuracy of the marker segmentation, a more comprehensive test would include a higher number of thresholds to analyze, which could potentially lead to a better-optimized segmentation.
- In this work, only predefined phase and amplitude values were used for the motion model. Future developments could extend the use of the model in order to apply it also to the training and testing data, which would mean to include the surrogate data into the model computations.

Finally, a comparative analysis between 4DCT and 4DCBCT global motion models will be performed in order to evaluate the capability of 4DCBCT data to create an accurate motion model just prior treatment delivery, without relying on planning 4DCT data which does not account for motion occurring between different fractions.

References

Amaldi, U. a. (2005). Radiotherapy with beams of carbon ions. Reports on progress in physics, 68(8), 1861.

Barnett, G. C. (2009). Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nature Reviews Cancer, 9(2), 134.

Cerutti, S. a. (2011). Advanced methods of biomedical signal processing (Vol. 27). John Wiley \& Sons.

Chan, M. a. (2011). Evaluation of imaging performance of major image guidance systems. Biomedical imaging and intervention journal, 7(2).

Chen, G. T. (2004). Artifacts in computed tomography scanning of moving objects. In Seminars in Radiation Oncology (Vol. 14, pp. 19--26). Elsevier.

D'Acunto, M. a. (2016). 3D image reconstruction using Radon transform. Signal, Image and Video Processing, 10(1), 1--8.

Driver, D. a. (2004). Improvements in radiotherapy practice: the impact of new imaging technologies. Cancer Imaging, 4(2), 142.

Durante, M. a. (2010). Charged particles in radiation oncology. Nature reviews Clinical oncology, 7(1), 37.

Feldkamp, L. a. (1984). Practical cone-beam algorithm. JOSA A, 1(6), 612--619.

Goitein, M. (2004). Organ and tumor motion: an overview. In Seminars in radiation oncology (Vol. 14, pp. 2--9). Elsevier.

Goldsmith, C. a. (2012). Stereotactic ablative body radiotherapy (SABR) for primary and secondary lung tumours. Cancer Imaging, 12(2), 351.

Hajnal, J. V. (1995). A registration and interpolation procedure for subvoxel matching of serially acquired MR images. Journal of computer assisted tomography, 19(2), 289--296.

Hill, D. L. (2001). Medical image registration. Physics in medicine & biology, 46(3), R1.

Jaffray, D. a. (2000). Cone-beam computed tomography with a flat-panel imager: Initial performance characterization. Medical physics, 27(6), 1311--1323.

Kang, K. H. (2015). Complications from stereotactic body radiotherapy for lung cancer. Cancers, 7(2), 981--1004.

Keall, P. (2004). 4-dimensional computed tomography imaging and treatment planning. In Seminars in radiation oncology (Vol. 14, pp. 81--90). Elsevier.

Keall, P. J. (2006). The management of respiratory motion in radiation oncology report of AAPM Task Group 76 a. Medical physics, 33(10), 3874--3900.

Kilby, W. a. (2010). The CyberKnife robotic radiosurgery system in 2010. Technology in cancer research & treatment, 9(5), 433--452.

Langen, K. a. (2001). Organ motion and its management. International Journal of Radiation Oncology* Biology* Physics, 50(1), 265--278.

Lichter, A. S. (1995). Recent advances in radiation oncology. New England Journal of Medicine, 332(6), 371--379.

Ling, C. C. (2004). High-tech will improve radiotherapy of NSCLC: a hypothesis waiting to be validated. International Journal of Radiation Oncology• Biology• Physics, 60(1), 3--7.

Little, J. A. (1997). Deformations incorporating rigid structures. Computer Vision and Image Understanding, 66(2), 223--232.

McClelland, J. R. (2006). A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. Medical Physics, 33(9), 3348--3358.

Nguyen, N. P. (2015). Image-Guided Radiotherapy for Effective Radiotherapy Delivery. Frontiers in oncology, 5, 253.

Orecchia, R. a. (1998). Particle beam therapy (hadrontherapy): basis for interest and clinical experience. European Journal of Cancer, 34(4), 459--468.

Radiological Society of North America, Inc. (2016, 01 26). Fiducial Marker Placement. Retrieved 07 02, 2018, from RadiologyInfo: https://www.radiologyinfo.org/en/info.cfm?pg=fiducial-marker

Richmond, N. a. (2012). Positioning accuracy for lung stereotactic body radiotherapy patients determined by on-treatment cone-beam CT imaging. The British journal of radiology, 85(85), 819--823.

Salsa, V. Z. (2013). A primer on PDEs, Models, Methods, Simulations. Milan: Springer.

Sarrut, D. (2006). Deformable registration for image-guided radiation therapy. Zeitschrift für medizinische Physik, 16(4), 285--297.

Schulz-Ertner, D. a. (2006). Radiation therapy with charged particles. In Seminars in radiation oncology (Vol. 16, pp. 249--259). Elsevier.

Seregni. (2013). Tumor tracking in particle therapy: development of dedicated methods and experimental testing. Doctoral Dissertation. Milano.

Shackleford, J. A., Kandasamy, N., & Sharp, G. C. (2010). On developing B-spline registration algorithms for multi-core processors. Physics in Medicine & Biology, 55(21), 6329.

U.S. Food and Drug Administration. (2018, 03 07). Computed Tomography (CT). Retrieved 06 30, 2018, from U.S. Food and Drug Administration: https://www.fda.gov/Radiation-EmittingProducts/RadiationEmittingProductsandProcedures/MedicalImaging/MedicalX-Rays/ucm115317.htm

Van Herk, M. (2004). Errors and margins in radiotherapy. In Seminars in radiation oncology (Vol. 14, pp. 52--64). Elsevier.

Xing, L. a.-F.-Y. (2006). Overview of image-guided radiation therapy. Medical Dosimetry, 31(2), 91--112.