POLITECNICO DI MILANO
Department of Bioengineering
PhD Program in Bioengineering

Development and testing of extra-cranial tumour tracking methods for intra-fraction organ motion compensation in particle radiation therapy

Supervisor: Prof. Guido BARONI
Tutor: Prof. Antonio PEDOTTI
Coordinator: Prof. Maria Gabriella SIGNORINI

Doctoral Dissertation by:
Aurora FASSI
754121

XXV CYCLE
Academic years 2010-2012

ACKNOWLEDGMENTS

I'm sincerely grateful to my supervisor, Prof. Guido Baroni, for granting me the opportunity to pursue this interesting research project and for guiding me with useful advices throughout this extremely formative experience.

I would like also to thank Dr. David Sarrut for his precious collaboration and valuable ideas, giving me the possibility to spend a really stimulating training period at the Centre Léon Bérard in Lyon.

I'm particularly thankful to the following research colleagues who contributed with high competence to the realization of this work, helping me to achieve these important results: Dr. Joël Schaerer, Dr. Mathieu Fernandes, Dr. Simon Rit, Dr. Jef Vandemeulebroucke and Dr. Rômulo Pinho.

Thanks also to Dr. Marco Riboldi and Dr. Pietro Cerveri for their valid support and suggestions in writing scientific papers.

Finally, a special thanks to the colleagues and friends of the CARTLab / TBMLab and to the whole Algeco team who worked and collaborated with me in Lyon.

CONTENTS

SUMMARY

Hadrontherapy, or charged particle beam therapy, is an emerging technique in the field of radiotherapy for the treatment of cancer pathologies. The use of charged particles, such as protons or carbon ions, offers several physical and radiobiological advantages with respect to conventional radiotherapy with photons. The favourable depth-dose profile of particle beams, which release most of their energy in a confined volume at a specific penetration depth, provides a higher spatial selectivity to the tumour, as well as a greater sparing of the healthy surrounding tissues. Dose conformity to the target volume can be further improved with the advanced techniques of dynamic dose delivery, based on the active scanning of charged particle beams through steering magnets. However, due to the steeper dose gradients of hadrontherapy treatments, tumour-beam misalignments may result in more severe damages to the irradiated structures, causing unacceptable under-dosage of the target volume and over-dosage of the nearby organs at risk. In addition, active beam scanning techniques feature a high sensitivity to intra-fraction tumour motion, caused by physiological organ movements in the thoraco-abdominal sites mainly due to breathing. For these reasons, only tumours with a negligible intra-fraction motion, like head and neck cancers, have been treated with active scanning particle therapy up to now. The opportunity to extend these advantageous techniques to the treatment of extra-cranial moving tumours necessarily requires the application of adequate strategies for dynamic target localization and for organ motion compensation.

Methods for localizing thoracic and abdominal lesions affected by respiratory motion can be generally categorized as direct or indirect tumour tracking. Direct tracking techniques rely on the imaging of the lesion or of implanted clips by means of in-room X-ray imaging systems, such as digital radiography, fluoroscopy or rotational Cone-Beam Computed Tomography (CBCT). The invasivity of this approach, related to the use of ionizing radiation and to the possible clinical complications induced by clip implantation, restricts its applicability to a limited temporal window of the treatment session. Indirect tracking methods infer target motion from external respiratory surrogates, such as the thoraco-abdominal surface displacement. Non-invasive optical localization devices can be applied to acquire breathing surface motion, by reconstructing the trajectory of

passive markers placed on the patient's skin or through markerless scanning of the entire surface. The non-invasivity of this approach allows a continuous monitoring of intra-fraction organ motion during the whole treatment course. However, the use of respiratory surrogates entails the definition of appropriate external-internal correlation models to estimate target position from surface displacement, requiring frequent verification and on-line adaptation of the correlation parameters to encompass intra-fraction breathing irregularities.

Different methods have been proposed to compensate for the geometric uncertainties that result from respiratory motion, based on the information derived from direct or indirect tumour localization techniques. A motion-compensated irradiation approach is breath-hold immobilization, consisting in the delivery of the treatment dose while the patient holds his breath at specific inspiratory volumes. Appropriate methodologies need to be applied to verify the stability and reproducibility of breath-hold maneuvers repeated intra and inter-fractionally. In the respiratory gating approach, the radiation beam is activated only during a limited phase interval of the breathing cycle, involving long treatment duration and a residual tumour motion within the gating window. The most efficient motion compensation strategy for accurate dose delivery with continuous irradiation in free-breathing conditions is real-time tumour targeting, based on the dynamic repositioning of the radiation beam according to the instantaneous target position. Motion-correlated dose delivery approaches make use of time-resolved imaging for treatment planning, such as 4D Computed Tomography (4D CT), providing information on tumour motion patterns during each phase of the breathing cycle.

The aim of this PhD project concerns the development and testing of surface-based tumour tracking methods for the compensation of intra-fraction organ motion, with application for real-time tumour targeting in active scanning particle treatments of extra-cranial lesions. The proposed approach is based on the non-invasive estimation of target position from the displacement of the thoraco-abdominal surface, by using patient-specific external-internal correlation models. The method relies on the integration of different imaging modalities, such as planning 4D CT, in-room X-ray imaging systems and three-dimensional (3D) optical devices for dynamic surface scanning. The correlation between external and internal breathing motion is derived directly from 4D CT images acquired during treatment planning. The anatomical information in the CT volumes associated to each respiratory phase provides a patient-specific model describing the 3D lesion movement during the breathing cycle. The 4D CT motion model is parameterized as a function of three main parameters, represented by the respiratory baseline, amplitude and phase. During each treatment session, the model's parameters are retrieved from the motion of the patient's thoraco-abdominal surface, dynamically acquired using non-invasive optical tracking systems. A surrogate

respiratory signal is extracted from the external surface displacement and modelled to derive the instantaneous values of amplitude and phase variables. To take into account possible inter-fraction anatomical variations that may occur between planning and treatment time, the tumour baseline in the 4D CT model is updated according to the daily information on target localization derived from in-room cone-beam CT imaging. The combination of the adapted 4D CT motion model with the breathing parameters extracted from the external surface motion provides the non-invasive estimation of the internal tumour trajectory along the three spatial directions during the whole treatment course.

The extraction of a surrogate breathing signal from the external thoraco-abdominal surface acquired with an optical imaging system requires different implementation steps. Markerless optical surface detection suffers from the lack of surface point correspondence between consecutive meshes. We proposed to apply deformable mesh registration to extract the multi-dimensional breathing motion of correspondent surface points, taking into account the deformation of the thoraco-abdominal surface induced by respiration. The developed deformable mesh registration approach is a non-rigid extension of the Iterative Closest Point (ICP) algorithm, based on a locally affine regularization. The general cost function is composed by a distance term, minimizing the distance between the deformed source and the target surface, and a regularization term, penalizing the difference among the transformations of neighbouring vertices. The proposed mesh registration algorithm was assessed in a group of 5 male healthy volunteers, by pair-wise registering the thoraco-abdominal surfaces optically acquired at three different respiratory phases. The method accuracy was quantified in terms of residual surface distance between each point of the deformed source mesh and the closest point on the corresponding target surface. We also estimated the registration error based on the known position of multiple surface control landmarks, represented by ten star-shaped markers placed on the thorax and abdomen of the subjects. The 95^{th} percentile of the residual surface distance after deformable registration did not exceed 1.2 mm, whereas the median errors on marker localization ranged between 1.0 and 2.3 mm. The motion tracking accuracy proved to be maximal in the abdominal region, where breathing motion mostly occurs, with median errors of 1.1 mm over all subjects.

We investigated and compared different approaches for summarizing the breathing trajectories of all surface points in a single respiratory surrogate signal, including Principal Component Analysis (PCA), k-means clustering and Self-Organizing Map (SOM). The three synthesis techniques were tested on 5 healthy subjects, acquiring the displacement of passive markers placed on the thoraco-abdominal surface with an opto-electronic system, simultaneously with the non-invasive imaging of the diaphragm motion obtained with a markerized ultrasound

probe. The spatial variability of breathing surface movement was investigated by applying k-means algorithms, which allow to group marker trajectories according to their motion pattern. Two different cluster validity indices were evaluated to select the optimal number of clusters for surface markers. According to both Silhouette and Dunn indices, the external surface can be mainly partitioned in two distinct regions represented by the thorax and abdomen, usually featuring breathing motion with different amplitude and phase shift. For the PCA approach, the overall respiratory signal summarizing breathing information from all surface control points was obtained by combining the signals associated to the first two principal components as a function of their explained data variance. For k-means and SOM techniques, we averaged the signals associated to the first two clusters and to the first two neurons. The breathing signals obtained with the three methods of synthesis were compared to the ultrasound diaphragm displacement, representing an accurate and reproducible surrogate for lung tumour motion. The respiratory signals obtained with k-means clustering proved to be more correlated with the reference diaphragmatic variable, obtaining a median value of the Pearson coefficient of 0.93 and a root-mean-square error of 0.11 between the normalized signals.

The estimation of the breathing amplitude and phase parameters to give in input to the 4D CT motion model requires the comparison of the external surface displacement acquired during planning and treatment phases, to take into account possible differences in respiratory motion patterns. The triangulated meshes corresponding to the thoraco-abdominal surfaces are extracted from each 4D CT phase image. The intersection between the surface portion scanned in the planning CT and the optical surfaces acquired during treatment is taken as the reference region of interest (ROI). For both 4D CT and optical surfaces, the 3D breathing trajectories for each surface point belonging to the selected ROI are estimated using the developed deformable registration algorithm. A mono-dimensional signal is derived for all ROI points, by computing the time series of the distance from the lowest position assumed by each point during the corresponding 4D CT scan or optical acquisition. K-means clustering is applied to summarize in a single respiratory surrogate signal the distance variables obtained for all ROI points. The breathing phase parameter associated to each frame of the optical surface acquisition is obtained by applying the Hilbert transform to the surrogate signal derived from the optical surfaces. The respiratory amplitude is instead parameterized according to a scaling factor computed as the difference between the amplitude of the optical surrogate signal and the amplitude of the 4D CT surface signal at the corresponding breathing phase.

The proposed patient-specific breathing motion model is built directly from the planning 4D CT volumes, by applying deformable image registration. The model relies on the computation of

the mid-position (MidP) CT image, representing a time-weighted average of all 4D CT phase volumes. The deformation vector fields between the MidP and each 4D CT phase image are estimated through B-spline deformable registration algorithms. Specific adaptation strategies are performed at each treatment fraction, in order to adapt the model parameters to the inter-fraction variations of the anatomo-pathological configuration and motion patterns. The daily information on tumour localization derived from the 3D CBCT image acquired before each treatment session is used to update the target baseline, featuring consistent differences from planning to treatment phases. Deformable registration is applied between the MidP CT and the CBCT volumes to generate an adapted baseline image. This image is mapped according to the deformation vector fields of the a priori 4D CT breathing motion model, thus obtaining the updated tumour positions at each phase of the breathing cycle. The 3D tumour coordinates associated to each frame of the optical surface acquisition are obtained by linearly interpolating the updated tumour positions of the 4D CT motion model at the corresponding instantaneous value of the phase parameter. The vector difference between the interpolated lesion coordinates and the adapted baseline image is multiplied by the amplitude scaling factor, to amplify or reduce tumour breathing motion with respect to the planning CT scan. The described approach allows to obtain the target trajectory in the three spatial dimensions during the whole course of the treatment session.

The proposed tumour tracking method was experimentally tested on a clinical database, including the synchronized information on internal and external breathing motion acquired from 7 lung cancer patients. Data collected for each patient include a planning 4D CT image set acquired with the Philips Brilliance CT scanner, providing ten phase-sorted volumes and the Maximum Intensity Projection (MIP) image. A cone-beam CT scan, consisting in about 650 X-ray projections of the internal anatomy acquired over 360° of gantry rotation, was collected for each patient using the Elekta Synergy system. The XVI software was used to reconstruct the 3D CBCT volumes from the two-dimensional (2D) cone-beam projections. The patient's thoraco-abdominal surface motion was acquired during the entire CBCT scan with the VisionRT optical imaging system. The VisionRT device employs non-invasive stereo-photogrammetric techniques to reconstruct the 3D topological model of the monitored surface. A speckle light-pattern is projected on the patient's skin and acquired by means of Charge-Coupled Device (CCD) cameras used for stereo-vision. The experimental testing of the developed tumour tracking technique is based on the comparison between the real lesion trajectories identified on cone-beam images and the target trajectories estimated from the external surface motion combined with the 4D CT breathing motion model.

An image-based approach was developed to synchronize CBCT and VisionRT data acquisitions, in order to retrieve the temporal correspondence between cone-beam projections and

optical surfaces. The proposed synchronization method is based on the tracking of the rotating CBCT flat-panel detector in the 2D images captured by the VisionRT stereo-cameras. The beginning of cone-beam CT scan is detected in the stereo-images by identifying the starting motion of the cone-beam CT panel, thus obtaining the temporal difference between CBCT and VisionRT acquisitions. A linear model was built to relate the position of the CBCT detector identified in the VisionRT photos with the projection angles associated to cone-beam images, which is correlated to the time elapsed from the beginning of cone-beam CT scan. Interpolation procedures were introduced to increase the temporal resolution of the developed synchronization technique, which allows to compute the relative starting time of CBCT with respect to VisionRT acquisition with millisecond resolution. The accuracy of the proposed image-based approach was evaluated on a moving phantom, placing a radiopaque marker on its flat moving surface. Eight simultaneous CBCT/VisionRT acquisitions of the phantom were performed, varying the frequency of its sinusoidal motion. The marker trajectory obtained from cone-beam projections was compared to the motion of the phantom surface derived from the VisionRT system. The median temporal difference between the zero-crossings of the two signals ranged between -8.1 and 19.5 msec, with a maximum inter-quartile range of about 70 msec.

The experimental testing of the proposed tumour tracking method required the identification of lung target position on cone-beam CT images. However, lung lesions often lack sufficient contrast in X-ray projection images and may be difficult to distinguish at particular rotation angles due to the superimposition of the surrounding high-intensity structures, such as the spine or the rib cage. We developed a novel automatic method for contrast enhancement and robust markerless tracking of lung tumours in X-ray images, exploiting the anatomical information derived from 4D CT planning volumes to increase target region visibility and reduce the overlapping effect of the nearby anatomy. The enhancement of lung lesion contrast is based on the subtraction between the original projection and the Digitally Reconstructed Radiograph (DRR) generated at the corresponding angle from the CT volume, in which lung regions including the tumour are masked. The tracking of lesion position on the contrast-enhanced images is obtained through template matching algorithms, using as reference template the corresponding projection of the tumour segmented in the planning CT image. The developed contrast enhancement technique was tested on the 650 cone-beam projections acquired for the 7 lung cancer patients of the database. The proposed method provided a 150% increase of the tumour contrast-to-noise ratio and a 24% increase of the rate of target identification with respect to conventional tracking methods without contrast enhancement. The errors in 3D target localization proved to be lower than 0.5 mm in all spatial directions.

The accuracy of the developed surface-based tumour tracking method was evaluated on the synchronized CBCT/VisionRT scans of the 7 lung cancer patients included in the database. From about 20 to 35 seconds of acquisition were analyzed per patient, corresponding on average to 7 breathing cycles. The real tumour trajectories identified on CBCT images were compared to the trajectories obtained by projecting the 3D lesion coordinates estimated from the synchronized optical surface acquisition with the 4D CT breathing motion model. The median differences between the real and estimated tumour trajectories ranged between 0.7 and 2.4 mm, with median values of 1.5 mm along both the horizontal and vertical image dimensions. The differences between the real and estimated respiratory parameters associated to the analyzed breathing cycles were computed for each patient. Median errors of the respiratory baseline and amplitude did not exceed 2.2 and 2.6 mm, respectively. The median values of the measured phase shifts were lower than 7% of the cycle length, corresponding to a maximum of 170 msec of delay.

A novel surface-based approach for the real-time tracking of extra-cranial moving tumours was developed and investigated. The proposed method relies on the non-invasive optical acquisition of the external surface displacement and on patient-specific breathing motion models, which are estimated directly from 4D CT planning images and updated at each treatment fraction to compensate for inter-fraction variations. The developed technique was assessed on 7 lung cancer patients, obtaining tumour tracking errors lower than the safety margins normally applied for the treatment of moving lesions with conventional X-ray radiotherapy. Although the validation of the proposed approach requires the experimental testing on a wider patient population, the presented work is put forward as a feasibility study for the clinical applicability of the developed surface-based tumour tracking method, providing a robust estimate of the achievable accuracy and geometrical uncertainties in dynamic target localization. The investigated technique is expected to represent an effective tool for the non-invasive compensation of intra-fraction organ motion, potentially allowing the extension of the advanced active scanning particle therapy techniques to the treatment of extra-cranial tumours.

PREFACE

The subject of this PhD thesis concerns the development and testing of real-time tumour tracking methods for the compensation of intra-fraction organ motion in hadrontherapy treatments of extra-cranial lesions. The work is part of the European project ULICE (Union of Light Ion Centres in Europe), under the work-package WP4 named "Ion therapy for intra-fraction moving targets". The WP activity aims at extending the advanced techniques of active particle beam scanning to the treatment of extra-cranial moving tumours, for a potential future application at the National Hadrontherapy Cancer Centre (CNAO, Pavia, Italy), the first Italian carbon-ion therapy facility. The research project was performed in collaboration with the Centre Léon Bérard (CLB, Lyon, France), where I spent the first six months of my PhD program. The stage period at CLB was focused on the acquisition of the patient database for the experimental testing of the proposed tumour tracking method.

The dissertation is divided into 4 chapters. The Introduction chapter describes the basic concept of particle therapy, focusing on different dose delivery systems. The problematic of organ motion for the treatment of extra-cranial lesions is discussed, mostly referring to lung cancers affected by breathing-related movements. A state-of-the-art review of the current strategies for the compensation of intra-fraction organ motion is reported, analyzing in particular the most effective approach of real-time tumour targeting based on direct or indirect target localization. Methods chapter presents an overview of the developed algorithms for the non-invasive prediction of the lesion motion from the displacement of the external thoraco-abdominal surface. The implementative steps for the extraction of the breathing parameters from the dynamic optical surface acquisition are described in details. The estimation of the patient-specific motion model from 4D CT planning images is also illustrated, along with the adopted strategy for the adaptation of model parameters to anatomical inter-fraction variations. In addition, Methods chapter describes the proposed approach for the evaluation of the developed tumour tracking method, based on the comparison between the real tumour trajectories obtained from cone-beam projections and the target trajectories estimated from the external surface combined with the 4D CT motion model. We present the collected lung cancer patient database, including synchronized acquisitions of CBCT images and surface motion captured with the VisionRT optical device. The proposed strategy for the synchronization of CBCT

and VisionRT systems is reported, along with the developed contrast enhancement technique for the identification of lung lesions on cone-beam projections. The third chapter is dedicated to the illustration of the results obtained from the experimental evaluation of the accuracy of the developed methods and algorithms. We focus in particular on the assessment of the proposed techniques for the extraction of the surrogate breathing signal from the dynamic optical surfaces and for the identification of tumour position on CBCT images. Target tracking errors related to the estimation of the lesion trajectories from the external surrogate signal are reported for all patients of the collected database. Chapter 4 finally reports general discussion and conclusions related to the presented results.

1. INTRODUCTION

1.1 Particle radiation therapy

1.1.1 Physical and radiobiological aspects

Radiotherapy is one of the most common and effective therapies for curative and palliative management of patients with cancer. It consists of locally exposing tumoral cells to ionizing radiations, with the aim of causing irreparable damages to their DNA. Approximately 50% of all patients with localized malignant cancers are treated with radiotherapy [Durante 2010], which can be used as a single treatment strategy or combined with other types of therapies, such as surgery or chemotherapy. The main goal of radiation treatments consists in achieving local tumour control with minimal side-effects to surrounding healthy tissues. The success rate of radiotherapy depends on the total radiation dose delivered to the tumour, whilst preserving adjacent structures. Normal tissue sparing is accomplished by treatment fractionation, consisting in delivering the radiation dose in multiple subsequent fractions. While healthy tissues show a greater ability to repair DNA damages induced by ionizing radiations, tumoral tissues feature less efficient recovery mechanisms. Multiple dose delivery fractions repeated at time intervals intermediate between the recovery time required by tumour and healthy tissues cause more lethal consequences on cancer cells than on normal cells. The preservation of healthy tissues can also be improved by geometrically conforming the radiation dose to the target volume, releasing a higher energy density to the tumour whilst maintaining an acceptable dose level to adjacent structures.

Conventional external beam radiation therapy is based on high energy photons, such as X-rays or gamma-rays, whose potential for treating cancer was recognized soon after their discovery in 1895. Gamma-rays are generated from high-activity radioactive sources, like cobalt, whereas X-rays are produced by electron linear accelerators (linac). Particle radiation therapy, clinically called hadrontherapy, is a more recent radiotherapy technique employing non-elementary charged particles, in particular protons and carbon ions, which are accelerated by cyclotrons or synchrotrons.

The clinical use of particle beams for the treatment of localized cancer was proposed in 1946 by Robert Wilson [Wilson 1946], based on the physical and radiobiological advantages of accelerated particles compared with photons.

1.1.1.1 Depth-dose profile of photons and particles

Figure 1.1 depicts the depth-dose distribution profiles for different radiation beams employed in radiotherapy. X-rays or gamma-rays show an exponential decrease in the radiation dose with increasing penetration depth, since they are progressively absorbed by tissues. The higher amount of energy is released at the body's surface, whereas dose profile is attenuated in depth. For photon beams with an a energy of 8 MeV, the maximum dose deposition is reached at a depth of 2–3 cm, while at 25 cm of penetration depth the dose is about one-third [Amaldi 2005]. The shift of the peak dose by a few centimetres from the patient's surface is due to a build-up effect, mainly caused by forward scattered Compton electrons [Schardt 2010]. Differently from photons, charged particles penetrate tissues with a limited absorption and lateral scattering, releasing most energy in a distinct narrow Bragg peak at the end of their range, with a sharp fall-off at the distal edge (Figure 1.1). For a 400 MeV/u carbon ion beam, the fall-off from 80% to 20% of the Bragg peak is about 10 mm [Amaldi 2007]. The position of the Bragg peak can be precisely adjusted to the desired penetration depth by changing the kinetic energy of the incident particle beams, knowing the attenuation profiles of the crossed tissues.

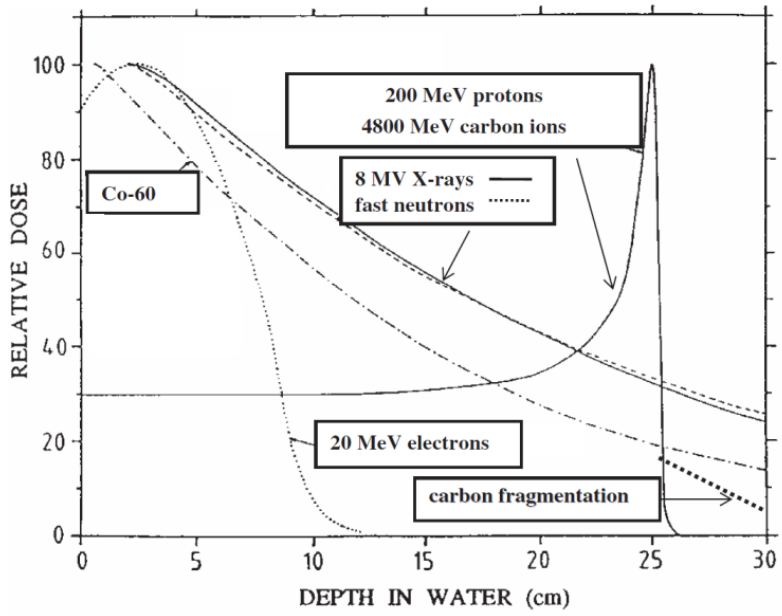


Figure 1.1. Comparison of depth-dose distribution for different radiation beams. While X-rays or gamma-rays show an exponential attenuation, with the maximum dose released at few centimetres of depth, particles deposit most energy at a higher penetration distance near the end of their range, focusing a higher dose in a narrower peak [Amaldi 2005].

The favourable depth-dose distribution of charged particles offer a number of theoretical advantages over conventional radiotherapy [Orecchia 1998]. The superior spatial selectivity provides an improved conformity of the particle beam to the target volume. The release of higher energy density at a precise and confined depth allows to focus the dose on the tumoral area and to minimize the irradiation of the healthy surrounding tissues, thus improving both local tumour control and normal tissue sparing. Particle radiation therapy is mainly indicated to treat inoperable deep-seated lesions or cancers located in proximity to organs at risks, such as eye and spine cord tumours. As shown in Figure 1.1, the energies for reaching 25 cm of penetration depth in water are of the order of 200 MeV for protons and 4800 MeV (corresponding to about 400 MeV/u per nucleon) for carbon ions. This means that in every cell a carbon ion leaves on average 24 times more energy than a proton having the same depth range [Amaldi 2007]. Energetic protons can be obtained either with cyclotrons (normal or superconducting) or with synchrotrons having a diameter of about 7 meters. Till now, only synchrotrons are used to produce carbon ions of about 400 MeV per nucleon. Their magnetic rigidity is about three times larger than 200 MeV protons, so that synchrotrons with more than 20 m of diameter are needed to generate carbon beams with a similar range in depth [Amaldi 2005].

The electric charge of particle beams can be exploited to control with extreme precision their trajectory by means of variable magnetic fields. Beside the configuration at fixed horizontal beam lines, isocentric gantries can be used to generate multiple radiation fields with different orientation focused on the tumour region [Pavlovic 1997]. Gantries consist in rotating mechanical structures, which support bending magnets and quadrupoles to deflect beam's direction. Due to the high magnetic rigidity of particles, very large and heavy gantries are required, with a typical radius of 4-5 m for 200 MeV proton beams [Amaldi 2007]. The superimposition of multiple particle beams magnetically guided and modulated in intensity, by delivering non-uniform dose pattern from each field direction according to tumour shape, is called Intensity-Modulated Particle Therapy (IMPT). This advanced technique allows to obtain steeper dose gradients and sharper penumbra, further improving the conformity of the treatment beam to the target volume and reducing the irradiation of surrounding healthy tissues by a factor ranging from 2 to 3 [Amaldi 2005]. As depicted in Figure 1.2, IMPT plans provide excellent dose distributions and a significant sparing of normal tissues and critical structures when compared with the most advanced conventional X-ray treatments, such as Intensity-Modulated Radiation Therapy (IMRT). IMRT techniques usually make use of 6-10 non-coplanar X-ray beams, generated by a linear accelerator mounted on a rotating gantry and focused towards the geometrical centre of the target. Beam intensity is varied across the irradiation field by means of dynamic multi-leaf collimators. Although IMRT treatments lead to tumour dose

conformity comparable to particle therapy, a larger volume of surrounding healthy tissues is exposed to irradiation (Figure 1.2).

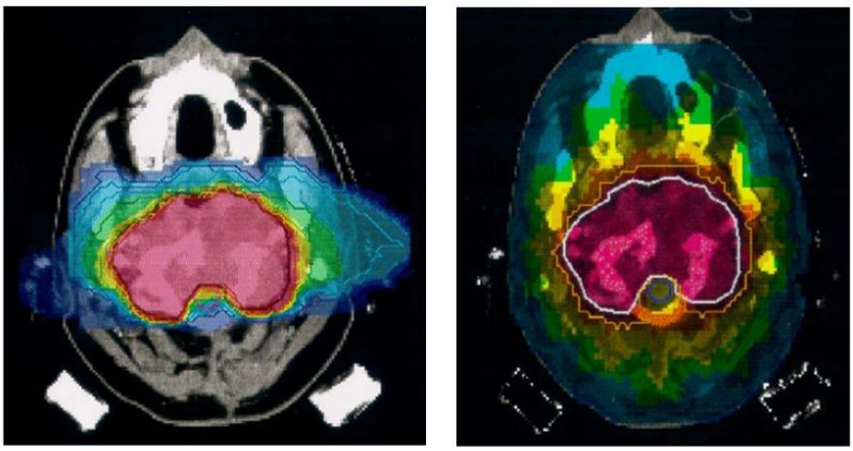


Figure 1.2. Comparison of treatment plans with 2 fields of carbon ion IMPT (Panel a) and with 9 fields of X-rays IMRT (Panel b). In both cases, a good conformity to the target volume is achieved, but in the IMPT plan the dose to the normal tissues is significantly lower [Amaldi 2005].

1.1.1.2 Biological effects of particles

A further advantage of charged particles with respect to conventional radiotherapy concerns the radiobiological properties, which are measured as Relative Biological Effectiveness (RBE). RBE is defined as the ratio between the dose of photons and the dose of the investigated radiation needed to produce the same biological effect, expressed in terms of ionization of the DNA molecules of the irradiated cells. While proton beams have almost the same biological effectiveness as X-rays, featuring a constant RBE value of 1.1 as a function of penetration depth, carbon beams show a favourable RBE depth profile. The biological effectiveness of carbon ions significantly increases in the Bragg peak region to values between 2 and 5 (depending on the type of tissue) with respect to the entrance channel, where mostly repairable damages are produced [Amaldi 2005]. Due to the higher electron-density, carbon ions show a greater Linear Energy Transfer (LET), equivalent to the energy loss per unit track length, especially in the last few millimetres of their range [Schardt 2010]. A greater ionization density is produced at the Bragg peak, causing double strand breaks and multiple damage sites to the DNA of the irradiated tumoral cells. The increased biological effectiveness in carbon ions can be restricted mainly to the target volume [Kraft 1990], whereas for heavier ions the range of elevated RBE also extends to normal tissues located before and after the tumour. As a result of this favourable radiobiological property, carbon ion beams are particularly indicated for the treatment of targets that are radio-resistant both to X-rays and to protons.

1.1.2 Dose delivery systems

As depicted in Figure 1.1, mono-energetic particle beams feature a very narrow Bragg peak, with a width of few millimetres. To irradiate larger tumours, beams of different energies have to be superimposed to produce an enlarged Spread-Out Bragg Peak (SOBP), as shown in Figure 1.3. By varying the particle beam energy during irradiation, the peak of the Bragg curve can be shifted at different depths in order to cover the whole tumour extent. The physical dose profile is corrected according to the RBE of the specific particle to provide a flat SOBP in the tumour region [Krämer 2000]. The spread-out Bragg peak allows to deposit a homogenous biologically effective dose within the entire tumour volume, but increases the energy density released to normal tissues in the entrance channel. Due to the favourable RBE profile, carbon ions provide SOBP with steeper dose gradients, improving the ratio between the dose deposited in the tumour region and in the entrance channel (Figure 1.3).

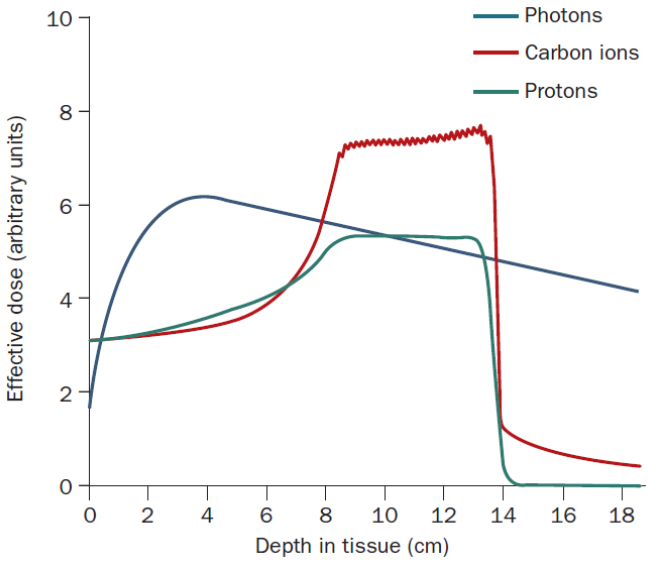


Figure 1.3. Spread-out Bragg peaks obtained by modulating the energy in overlapping beams of protons and carbon ions, providing a constant biologically effective dose within the tumour [Durante 2010].

The modulation of the particle beam energy to produce the SOBP can be obtained in two possible ways [Durante 2010]. A fixed energy beam can be attenuated by interposing absorbing materials and range shifters of variable thickness in the beam path (passive modulation). Alternatively, the energy of the particle beam can be modulated by the accelerator during the irradiation, without the need of passive attenuators (active modulation). In synchrotrons, the energy of the extracted pulsed beams can be varied from one cycle to the next one in steps of few MeV by tuning the accelerator parameters, providing essential advantages for conformal treatments. Active

modulation is instead not feasible in cyclotrons, requiring energy absorbers to adapt the particle range to the depth of the irradiated target, which may involve beam scattering and degradation of the energy sharpness [Schulz-Ertner 2006]. According to Bert and Durante [Bert 2011], mainly two treatment delivery options exist to conform the charged particle dose distribution to the clinical target: passive beam shaping [Chu 1993], which has been used in all hadrontherapy centres till 1997, and active beam scanning, introduced at the GSI (Gesellschaft für SchwerIonenforschung, Wixhausen, Germany) for carbon ions [Haberer 1993] and at the PSI (Paul Scherrer Institute, Villingen, Switzerland) for protons [Pedroni 1995]. Both treatment delivery approaches can be used either with active or passive energy modulation.

1.1.2.1 Passive beam shaping

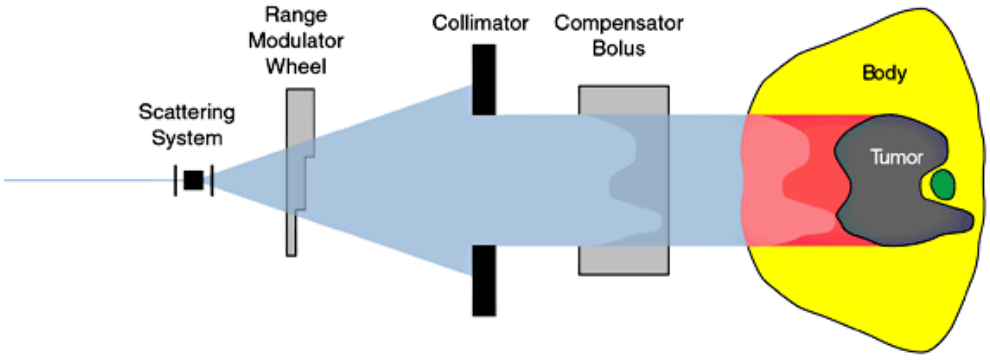


Figure 1.4. Schematic drawing of a passive beam shaping system. The initially narrow beam is broadened by a scattering device and adapted to the target volume by various passive field-shaping devices.

In passive beam shaping systems, the whole tumour volume is irradiated almost simultaneously, conforming the radiation field to the target contours by means of passive field-shaping elements, like scatterers, compensators and collimators (Figure 1.4). A scattering foil is interposed at the beginning of the beam path to laterally widen the pencil beam to the desired field size (typically 20 cm). The energy of the beam is degraded by means of range modulators and shifters, consisting in rotating wheels with varying thickness, in order to broaden the Bragg peak and to conform its width to the target longitudinal extent. Radiation field is adapted to the distal contour of the tumour by using appropriate compensating absorbers. The transverse form of the irradiation field is defined by means of custom-made collimators. Passive beam shaping systems are relatively simple to implement but feature several disadvantages, such as a low efficiency in the use of beam energy and the need of patient-specific compensators and collimators, defined by the size and shape of the irradiated target. In addition, this technique provides a good dose conformity only

for tumour with simple geometrical features and does not allow to tailor dose distribution to the proximal target shape, inevitably generating unwanted normal-tissue doses in the proximal end [Schardt 2010].

1.1.2.2 Active beam scanning

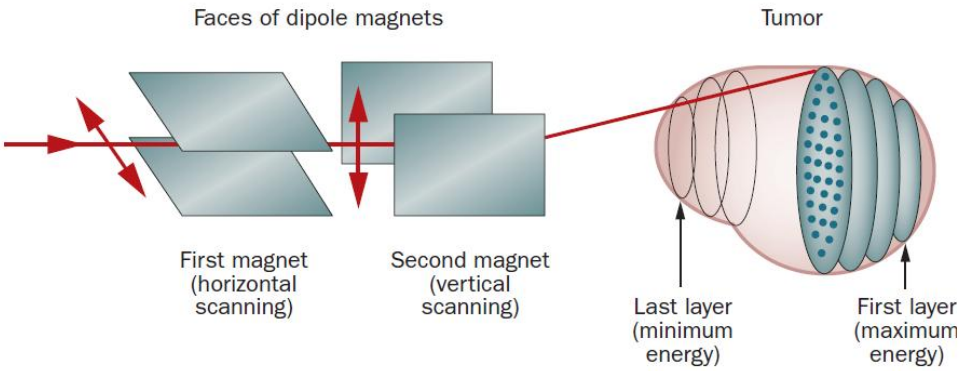


Figure 1.5. Schematic drawing of an active beam scanning system. The target volume is sequentially irradiated by bending a pencil-like ion beam in the horizontal and vertical directions with two couples of fast orthogonal magnets over thin layers at constant energy [Durante 2010].

Active beam scanning represents the most widespread technique for dose delivery in the newly designed and upcoming particle therapy facilities. In active scanning systems, the target volume is irradiated sequentially with a dynamic dose delivery approach based on magnetically guided particle beams. The electric charge of particles is exploited to deflect and focus the beam direction through orthogonal computer-controlled magnets. As shown in Figure 1.5, the target volume is divided in layers of equal beam energy (iso-energy slice), each one composed by a grid of finite volume elements (voxel). Each tumour voxel is sequentially irradiated by magnetically steering an individual pencil-like beam. Active beam scanning has several advantages compared to scattered beams, mostly due to the improved target conformity and to the possibility of beam shaping in the proximal tumour area. Since the dose is distributed voxel by voxel, any irregular-shaped lesions can be treated with optimal precision, giving minimal doses to the surrounding healthy tissues. Moreover, neither field-specific nor patient-specific hardware is required, eliminating any material in the beam path that can generate beam losses and production of secondary particles, like neutrons. On the other hand, active beam scanning implies strict demands on the beam control and safety systems and places strong requirements on the accelerator performance, such as stability and reproducibility of the beam position [Schardt 2010].

Two different beam scanning techniques have been proposed, namely spot-scanning [Pedroni 1995, Pedroni 2004] and raster-scanning [Haberer 1993]. The spot-scanning system uses

circular-shaped pencil beams which are positioned by a fast sweeper magnet in single spots of the target volume, with discrete steps of typically 5 mm. When the required dose has been released at one spot, the beam is switched off by a fast kicker magnet, the next spot is selected by the sweeper magnet and the beam is switched on again. This approach has been tested both with horizontal beam lines [Pedroni 1995] and isocentric gantries [Pedroni 2004], equipped with a double magnetic scanning system, which significantly speed up the beam delivery. In raster-scanning technique, a continuous path is selected in each iso-energy slice, made of subsequent raster-points with a spacing of typically 2 mm. When the desired dose is delivered in one raster-point, the pencil beam is moved to the next one without turning it off. After completion of one slice, the synchrotron beam extraction is interrupted and the beam energy for the next slice is selected. In both scanning techniques, a beam monitoring system located close to the patient measures the instantaneous beam position and controls the dose delivered to each tumour voxel.

1.1.3 Current status of particle therapy

According to the most recent survey of the Particle Therapy Cooperative Group [PTCOG 2011], 33 and 6 centres in the world are nowadays treating patients with protons and carbon ions, respectively, whereas 25 new particle facilities are being planned or under construction. In Europe, only two facilities are now performing carbon-ion treatments: HIT (Heidelberg Ion Therapy, Heidelberg, Germany) and CNAO (National Centre for Oncological Hadrontherapy, Pavia, Italy), which started treating patients in 2010 and 2011, respectively. CNAO is in particular the first combined proton and carbon-ion facility in Italy equipped with a 25 m diameter synchrotron and active scanning delivery techniques [Orecchia 2009]. To date, about 68000 patients have been treated with protons worldwide and about 9000 with carbon ions. Including also the centres now out of operation, the total number of patients that underwent a charged particle treatment is about 97000 [PTCOG 2011]. Uveal melanomas and lesions at the skull base were among the first diseases treated with particle beams, followed by tumours localized in head and neck, liver, brain, upper abdomen, pelvis, lung and spinal cord [Greco 2007]. The current results of clinical phase I-II trials, although involving a limited sample of patients, support that the favourable dose distribution and the radiobiological properties of charged particles lead to improved patient's outcomes, especially for tumours in proximity to critical organs or resistant to conventional treatments [Durante 2010]. Even if clinical data are still not sufficient to establish firm conclusions on the cost-effectiveness of this treatment modality, the increasing number of excellent patient's outcomes provides evidence that charged particle therapy might be beneficial in several tumour sites [Schardt 2010].

1.2 Radiation treatment process

1.2.1 Anatomical and functional imaging

The radiation treatment process, both for conventional and particle therapy, consists of two main steps: treatment planning and treatment delivery. Medical imaging plays an essential role in all state-of-the-art radiotherapy techniques, since it enables to plan the desired treatment and to verify that the treatment is delivered as planned [Evans 2008]. Input data for treatment planning process are obtained from the acquisition of anatomical and functional images of the region of interest, used for the identification of the lesion and surrounding Organs At Risks (OAR). The main imaging modalities applied for treatment planning include Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET). CT and MRI are high resolution anatomical imaging techniques, which measure the physical characteristics of the tissue, whereas PET provides complementary functional imaging of various aspects of tumour biology, such as tissue metabolism [Nestle 2009]. Three-dimensional (3D) CT is the most common volumetric imaging modality applied for radiation treatment planning, since it allows the accurate delineation of bony structures and the analysis of the dose distribution decay. The gray levels of the CT image, coded in terms of Hounsfield Units, represent a map of the radiation attenuation coefficients, which can be converted in tissue density with appropriate scaling factors [Kalender 2006]. MRI scans give instead a proton density map of the tissue, obtained non-invasively by using static or gradient magnetic fields and radiofrequency excitation. This technique have gained particular importance for the treatments of tumours localized in the central nervous system, abdomen and pelvis, thanks to its superior ability to discriminate soft tissues.

1.2.1.1 Image registration

Since information obtained from different imaging modalities is usually of a complementary nature, separate images (such as CT and MRI or CT and PET) are often integrated and combined to improve the interpretation of clinical data and to facilitate the delineation of the structures of interest [Veninga 2004]. The integration process requires a registration step, for the spatial alignment of the different set of images, and a fusion step, for the combined display of the data. Registration is defined as the process of determining the optimal transformation that provides the maximum similarity between the reference source and the target image. All registration algorithms essentially consist of three components: a similarity metric, a transformation model and an optimization method [Sarrut 2006a]. The similarity metric represents the cost function to be

optimized, which provides a fitness value of the match between the source and the transformed target image. Feature-based approach relies on corresponding features detected in the two images, such as anatomical landmarks, organ contours or segmented surfaces. Registration is assessed only at feature locations, without guaranteeing the accuracy in other regions, particularly those distant from any feature. In the intensity-based approach, the grey-levels distribution of the two images are matched, thus assessing the image similarity at each voxel. For mono-modality image registration, the most frequently used intensity-based similarity metrics are the Mean Square Error (MSE) and the Cross Correlation (CC), which are based on the assumption that the registered images have an equal or linear intensity relationship, respectively. For multi-modal cases, the Normalized Mutual Information (NMI) metric is usually applied, which assumes a probabilistic relationship between image intensities and is based on the computation of the entropies of the intensity distributions [Maintz 1998].

The transformation model defines how the target image coordinates are mapped to the reference source coordinates to improve the similarity [Zitova 2003]. The number of parameters used to define the transformation represents the number of degrees of freedom. Rigid transformations, implying that the relative distances between different points of the two images remain constant, have six degrees of freedom, since they are parameterized with three rotations and three translations. Rigid transformations are typically used to align anatomy that is assumed to remain rigid, such as bones. They are often applied during treatment setup to calculate the couch parameters for the correction of patient positioning errors. Rigid transformations are a sub-class of affine transformations, which include also anisotropic scaling, shearing and reflection, without preserving distances between points. The more complex non-rigid transformations account for elastic deformation of the matched volumes. The output of deformable image registration is a Deformation Vector Field (DVF), composed by a vector map describing the 3D displacement of each voxel from the source to the target image, as shown in Figure 1.6.

Non-rigid transformations are classified in parametric and non-parametric [Kybic 2003]. In the parametric approach, based on a limited number of parameters, the deformation is constrained by a set of corresponding landmarks defined in the source and target images. A linear combination of radial basis functions is used to interpolate the transformation between landmarks. Thin Plate Splines (TPS) are the most popular global basis functions [Bookstein 1989], in which each landmark pair globally impacts the transformation over the entire image. Cubic B-splines are an example of basis functions with a local effect, since transformation is affected only in a spatially limited neighbourhood of the landmarks, thus reducing the computational costs [Shackleford 2010]. B-splines are piecewise polynomials, often described by a freeform deformation model, based on a

regular grid of control points with uniform spacing. Approximating B-splines are used to calculate the deformation field at any image voxel from the displacements of the surrounding control points. In the non-parametric methods, the transformation function relies on physical properties and is basically unconstrained, thus resulting in a very large number of degrees of freedom. A regularization term is added to penalize deformation fields that deviate from the desired form. The most commonly applied regularizing energy functions are the linear elastic model, the optical flow model and the viscous fluid model [Cachier 2004]. In a multi-institutional study on deformable registration accuracy, the results of deformable image registration for multiple organs performed by 21 participating institutions were compared to human segmentation based on natural landmarks, such as bifurcation points of blood vessels [Brock 2010]. The study showed that the majority of registration algorithms perform at an accuracy equivalent to the voxel size. For 4D CT registration of the lung, the average absolute errors were 0.6–1.2 mm, 0.5–1.8 mm and 0.7–2.0 mm in the medio-lateral (ML), anterior-posterior (AP) and superior-inferior (SI) directions, respectively.

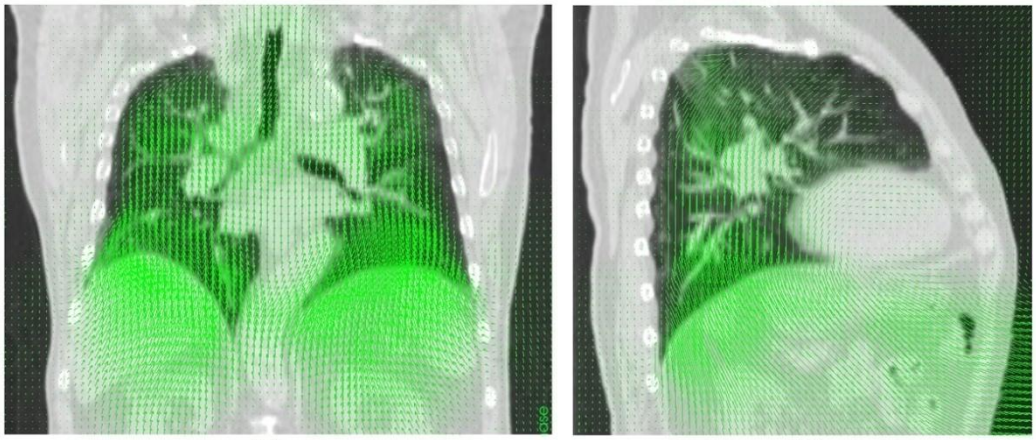


Figure 1.6. Anterior (Panel a) and lateral (Panel b) views of the deformation vector field computed by non-rigidly registering two 4D CT phase images of a lung cancer patient at different breathing states (end-exhale and end-inhale). The DVF is visualized by using the open-source image viewer VV [Rit 2011]. The length of the vectors corresponds to the entity of the 3D voxel displacement, projected on the corresponding plane.

The optimization method tries to find the parameters of the transformation model that give the best value of the similarity metric [Zitova 2003]. Closed form solutions, such as least squares methods, can be used for some feature-based registrations, whereas all intensity-based registrations require iterative optimization techniques. Various optimization schemes have been proposed, with different performance in terms of speed and robustness. Steepest gradient descent technique, consisting in the iterative minimization of the gradient of the transformation with respect to the similarity measure, is often used with B-splines and other non-rigid transformation models.

Levenberg-Marquardt and Gauss-Newton algorithms, such as the Limited-memory Broyden Fletcher Goldfarb Shanno (L-BFGS), are other examples of optimization methods. As demonstrated by Fischer and Modersitzki [Fischer 2008], the optimization problem for image registration is ill-posed, since small changes in the input images or in the initialization can lead to completely different registration results. Its solution may not be unique, since several equally optimal solutions can be found to match the input datasets. Multi-resolution techniques are often employed in non-rigid registrations to improve robustness and to prevent finding local minima.

1.2.1.2 Time-resolved imaging

All imaging modalities applied in thoracic and abdominal sites are affected by artefacts due to physiological movements, such as breathing and peristalsis [Keall 2006]. Motion artefacts can lead to errors in the delineation of target and normal tissues and adversely affect the accuracy of dose calculation. New protocols for image acquisition have been proposed in order to account and compensate for breathing movements, such as respiratory-gated scanning [Keall 2002]. The gating approach consists in acquiring all CT data with the patient holding his breath at a specific respiratory phase, such as deep-inspiration or deep-exhalation. An example of the difference between a non-gated and a respiratory-gated CT scan is shown in Figure 1.7. Whilst respiratory-gating greatly reduces motion artefacts in the acquired images, it does not represent an ideal solution for treatment planning [Cheung 2003]. Radiotherapy treatments generally take several minutes, during which a single breath-hold is not possible. Several breath-holds could be used but questions exist about the reproducibility of repeated breath-holds, as further described in Section 1.4.2.

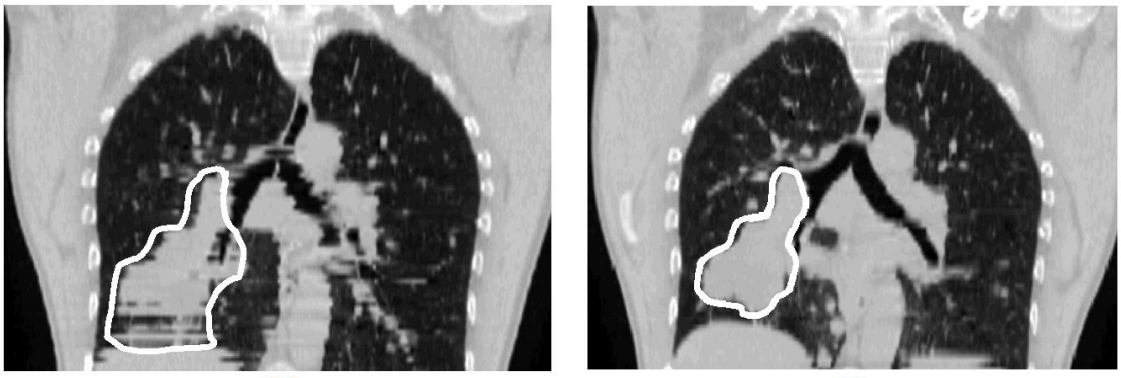


Figure 1.7. Coronal views of a CT volume of the same patient taken during free-breathing (Panel a) and with respiratory-gated scanning at end-exhalation (Panel b). The contoured tumour volume is outlined in each image [Keall 2002].

Time-resolved volumetric imaging, such as four-dimensional MRI or CT (4D CT), have been developed to extend treatment planning capabilities for tumour sites influenced by respiratory motion [Von Siebenthal 2007, Rietzel 2005]. 4D CT technique consists in the simultaneous acquisition of CT data and of a respiratory surrogate signal, which is used to sort the acquired images into bins according to the breathing state. The three-dimensional reconstruction of each bin yields a set of respiratory-correlated CT volumes, representing different respiratory phase states during an averaged breathing cycle. The Maximum Intensity Projection (MIP) image, corresponding to the highest CT number found for each voxel in the set of 4D CT volumes, can be used to compute the motion-encompassing tumour volume including target positions in all respiratory phases. A CT scan can be performed both in cine-mode, in which CT data are acquired at discrete stationary positions of the treatment couch, and in spiral or helical mode, involving the slow couch motion during the continuous CT data acquisition. An illustration of 4D CT scanning and image reconstruction is depicted in Figure 1.8.

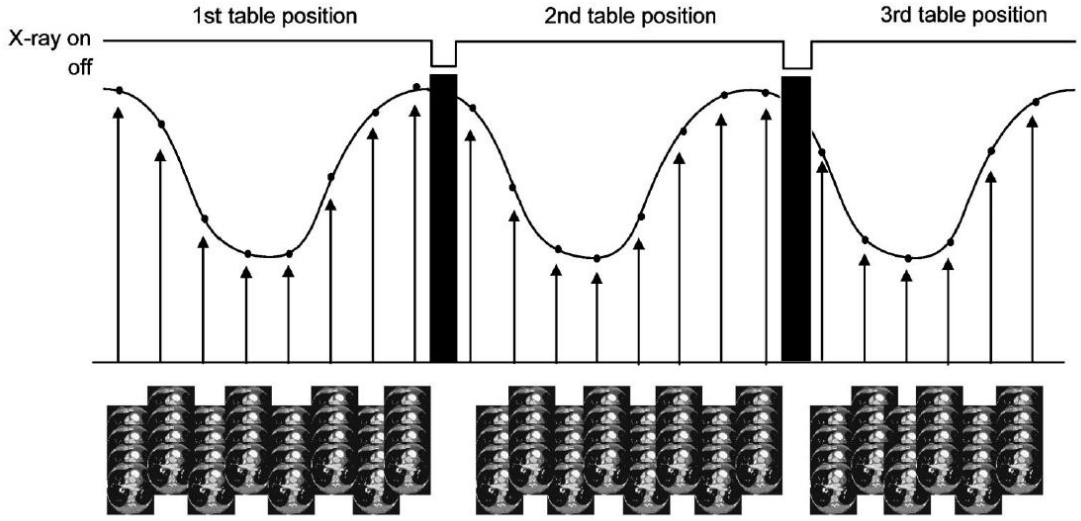


Figure 1.8. 4D CT scan acquisition process using a multi-slice CT scanner operating in cine-mode. The breathing cycle is divided into distinct bins and the CT slices are sorted into these bins according to the respiratory state associated to their acquisition [Pan 2004].

The most common respiratory surrogates applied for 4D CT imaging include spirometric volume [Low 2003], abdominal height [Vedam 2003] and internal anatomical motion [Pan 2004]. Different approaches have been proposed for 4D CT data sorting: amplitude gating, phase gating or a combination of both [Wink 2006, Abdelnour 2007]. In amplitude gating, the CT acquisition is triggered when the breathing trace is at a certain position, whereas in phase gating the triggering is activated when the respiration is at a given phase. Current 4D CT techniques are based on the

assumption that the breathing cycle is periodical and that the respiratory surrogate can be used to uniquely determine the breathing state of the acquired CT data. However, respiratory irregularities or incorrect breathing state identification may lead to blurred images and motion-induced artefacts, which were observed for about the 80% of 4D CT acquisitions [Yamamoto 2008]. 4D CT volumes sorted using amplitude gating generally result in less image artefacts than those sorted using phase gating, particularly in presence of high inter-cycle variations during data acquisition [Lu 2006]. The combination of multiple respiratory-related signals coming from different portions of the thoraco-abdominal surface was proved to reduce artefacts in 4D CT imaging [Gianoli 2011].

1.2.2 Treatment planning

Radiation treatment planning can be divided in two different phases: contouring and plan optimization. During the contouring phase, the tumour and critical structures are manually delineated on the acquired volumetric images. According to the International Commission on Radiation Units and Measurements Report 62 [ICRU 1999], the visible target outlined on the planning images is called Gross Tumour Volume (GTV), as shown in Figure 1.9a. Safety margins are applied to take into account different sources of geometrical uncertainties, such as variations between treatment and planning scan (systematic errors) or day-to-day variations during a single treatment fraction (random errors) [Van Herk 2004]. A margin is added to the GTV contour to include microscopic tumour extensions, thus obtaining the Clinical Target Volume (CTV). To take into account changes in size, shape and position of the CTV, which may be caused by patient's weight loss, tissue response to radiation or intra-fraction motion due to breathing, the CTV is extended to the Internal Target Volume (ITV) by applying Internal Margins (IM). Setup Margins (SM) are finally added to include possible inter-fraction geometric variations, such as setup errors in patient positioning or beam misalignments, thus obtaining the Planning Target Volume (PTV) used for dose prescription. In the plan optimization phase, the appropriate geometric and radiological aspects of the treatment, such as the number, energy, shape and direction of the radiation beams, are determined by using dedicated Treatment Planning Systems (TPS) to simulate the dose deposition (Figure 1.9b). Treatment plan is optimized in order to achieve the prescribed uniform dose to the PTV, while minimizing the irradiation of organs at risk and healthy tissues, evaluated by means of dose-volume histograms.

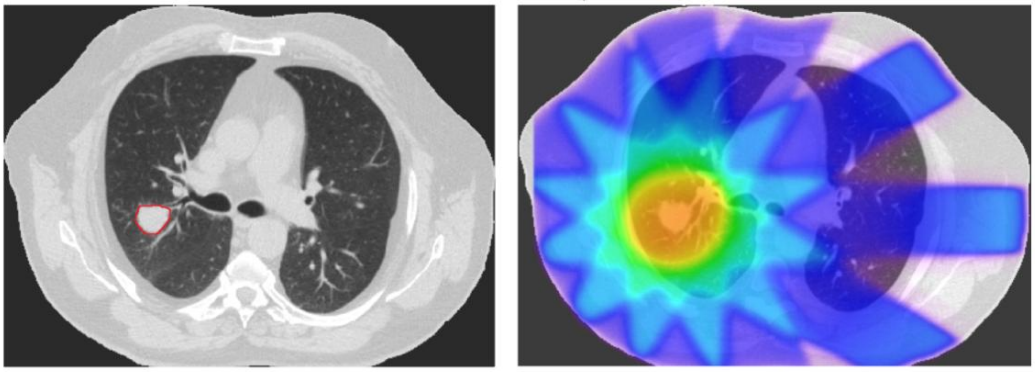


Figure 1.9. GTV contours (red line) outlined in the planning CT image for a lung cancer patient (Panel a) and corresponding treatment plan obtained by simulating dose deposition with TPS systems (Panel b).

1.2.3 Treatment delivery

The second step of a radiation treatment is the delivery of the planned dose, fractionated in multiple sessions. The amount of delivered radiation is measured in Gray (Gy) and varies depending on the type and stage of the tumour being treated. A typical conventional radiation treatment for a localized solid cancer consists of 60–70 Gy delivered to the tumour in 30–35 daily fractions of 2 Gy over 6–7 weeks [Durante 2010]. Hypo-fractionation schemes, involving fewer treatment sessions with higher doses, have recently received considerable attention, especially for lung cancer therapy, for which remarkably high tumour control rates have been reported [Lagerwaard 2008]. Each treatment fraction is preceded by a setup phase in which the patient is re-positioned in the planning reference configuration, by aligning skin markers with the treatment room lasers. In order to minimize setup uncertainties, immobilization devices are generally used, including patient-specific thermoplastic masks or customized stereotactic body frames. In the most advanced Image Guided Radiation Therapy (IGRT), inter-fraction positioning errors are monitored and corrected prior to irradiation by using in-room imaging systems, to improve the geometric accuracy of the delivered treatment [Jaffray 2007]. The most common in-room imaging modalities are represented by two-dimensional (2D) X-ray projections or volumetric CT images. The agreement with the planning CT is usually checked based on bony anatomy or soft tissues, such as OAR and tumour volume, updating couch position according to the computed rigid corrections. The discrepancies between planned and daily acquired images can be also used to modify and adapt the treatment plan, in a process known as Adaptive Radiation Therapy (ART) [Webb 2008].

The most widespread in-room approach for patient setup verification is represented by portal images, consisting in 2D projections acquired before each treatment fraction and compared to Digitally Reconstructed Radiographs (DRR) obtained by projecting the planning CT volume at the

corresponding angle [Sherouse 1990]. Portal imaging provides information on internal anatomy, especially bony structure. Soft-tissue positioning is usually determined with the aid of implanted fiducial markers as tumour surrogates [Chen 2007]. Megavoltage (MV) electronic portal images are obtained by using the actually X-ray treatment beam, hence involving no extra dose [Herman 2005]. However, the higher energy of the employed beam limits the quality of the acquired images, due to the lower quantum efficiency of the detectors and to the lower intrinsic contrast between various tissues at MV energies. By incorporating a diagnostic-energy X-ray imaging system into the treatment room, it is possible to acquire kilovoltage (kV) portal images with better quality and contrast [Jaffray 1999]. A favoured approach consists in integrating at least two kV systems to enable stereoscopic 3D imaging based on triangulation techniques, thus providing information on tumour localization in three dimensions [Berbeco 2004]. Proton radiography is currently gaining a growing interest, since it allows a significant reduction of the dose delivered to the patient. Despite the lower spatial resolution compared to X-ray imaging, proton radiography is more sensitive to tissue density variations, offering a better soft and bone tissue contrast. This leads to sharper edge detection and better target localization for specific tumour cases, such as lung cancer [Depauw 2011].

3D anatomical information including soft tissue can be obtained through kV Cone-Beam CT (CBCT), consisting in an on-board imaging equipment with a source-detector system integrated in the linear accelerator [Jaffray 2002]. As shown in Figure 1.10, the X-ray tube is typically mounted on a retractable arm perpendicular to the treatment gantry, opposite to the X-ray flat-panel detector. CBCT imaging involves the acquisition of multiple kV radiographs as the gantry rotates around the patient positioned on the treatment couch. A filtered back-projection algorithm is employed to reconstruct a volumetric image from the set of 2D projections [Feldkamp 1984]. Geometric non-idealities in the rotation of the gantry system are measured and corrected during the reconstruction. To reduce motion artefacts and geometrical uncertainties due to breathing, a respiratory-correlated CBCT (4D CBCT) has been proposed [Sonke 2005]. The approach consists of retrospective sorting the acquired projections according to the breathing phase, subsequently reconstructing a volumetric CBCT image for each phase. The breathing signal used for respiratory correlation is directly extracted from the 2D projection data, removing the need for an additional respiratory monitor system. Despite the valuable information provided by CBCT technique for IGRT application, concerns are addressed to the additional dose released to the patient's skin and normal tissues, which may entail the risk of cancer induction [Evans 2008].

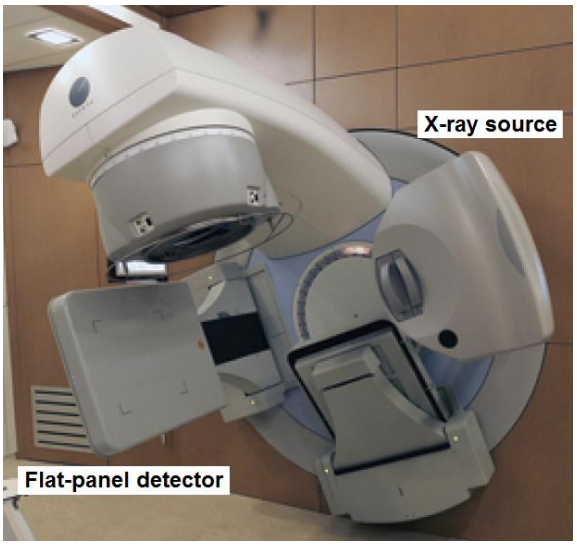


Figure 1.10. Kilovoltage cone-beam CT scanner mounted on a linear accelerator, consisting of an X-ray source opposite to a flat-panel detector.

1.3 Organ motion in radiotherapy

The issue of organ motion is becoming increasingly important in the contest of radiotherapy, since it introduces intra-fraction geometrical uncertainties that influence the accuracy of imaging, treatment planning and radiation delivery [Keall 2006]. Organ motion includes all physiological movements associated to the respiratory, skeletal muscular, cardiac, and gastrointestinal systems, which induce not negligible displacement of the tissues of interest, and consequently of the included tumoral mass. The most relevant source of intra-fraction organ motion is respiration, which affects all tumour sites in the thorax, abdomen and pelvis, such as lung, liver or pancreas cancers [Bert 2011]. According to the US National Cancer Institute, these pathologies currently account for about 20% of new cancer cases, with 5-year overall survival rates as low as 6–16% [Howlader 2011]. Since the respiratory frequency of a healthy person at rest is about 12–15 cycles per minute [Shirato 2004], breathing motion acts on a time-scale of seconds to minutes. A continuous real-time monitoring of organ motion is therefore required during the delivery of treatment fractions. Breathing motion patterns vary markedly between patients, indicating that a patient-specific approach for respiration management is advised. No correlation between the tumour size or location and the motion occurrence or magnitude was found, suggesting that tumour movement should be assessed individually [Stevens 2001]. Intra-patient variations in amplitude, period and regularity of the breathing pattern, as well as systematic changes in the respiratory baseline, were also observed during imaging and treatment sessions [Seppenwoolde 2002].

The treatment of moving organs is one of the major challenges in particle therapy, representing a very active field of research. The clinical impact of intra-fraction motion consists in blurring the dose gradients from target volume to normal tissues and in changing the distribution of the deposited dose, since the motion of organs within the beam path alters the radiological path-length [Bortfeld 2004]. The higher energy density and the steeper dose gradients of particle treatments may lead to more severe consequences in case of target-beam misalignment due to organ motion. Geometric errors in the presence of intra-fraction tumour motion may seriously affect the effectiveness of particle radiation, causing unacceptable under-dosage of the target volume and over-dosage of the surrounding healthy tissues. Treatment of intra-fractionally moving organs with a passive shaped beam is routinely performed in several clinical centres, since dedicated treatment planning procedures that incorporate the radiological changes induced by organ motion are sufficient to guarantee an adequate CTV coverage [Bert 2011].

An even more challenging issue is represented by the capability to treat moving tumours with the new advanced dose delivery technique based on active beam scanning, described in Section 1.1.2.2. Scanning beam systems are in fact more sensitive to intra-fraction organ motion, due to the interference effects between the moving target and the scanned beam [Grözinger 2008]. This results in a deteriorated and non-homogeneous dose distribution within the irradiated volume, which hinders an adequate CTV coverage despite the use of safety geometrical margins [Lambert 2005]. Bert and colleagues [Bert 2008] quantified the interplay effects of uncompensated motion on dose uniformity in carbon-ion therapy with scanned beams. For a spherical water phantom with an amplitude of 15 mm, only 47% of the target volume received 80% of the planned dose. To date, only tumours that are not affected by respiratory motion have been clinically treated with scanned particle beams [Bert 2011]. To make the most of the excellent target dose conformity and geometrical accuracy of active scanning systems, dedicated measures are needed to compensate for the dosimetric influence of intra-fraction target motion. High precision techniques for tumour localization and real-time motion mitigation are necessarily required to extend the application of active scanning treatments to intra-fractionally moving tumours, with the potential benefit of improving local tumour control and reduce normal tissue complications.

1.3.1 Lung cancer

The most prevalence and relevant disease affected by breathing-induced organ motion is lung tumour, representing the most common cause of cancer-related death for men and the second for women, behind breast cancer. According to the last GLOBOCAN statistics on cancer incidence

and mortality, lung tumour accounts for an estimated 1092 thousand new cases and 949 thousand deaths each year among men, and 516 thousand new cases and 428 thousand deaths among women [Ferlay 2008]. The appropriate therapy depends on the stage and type of lung cancer, which can be classified as small cell (14%) or non-small cell (85%). For localized Non-Small Cell Lung Cancer (NSCLC) patients, surgery is usually the treatment of choice, often combined with radiotherapy or chemotherapy to improve survival rate. According to the American Cancer Society, the 5-year survival rate for all lung cancer stages is only 16% [ACS 2012]. The 1-year relative survival increased from 37% in 1975–1979 to 43% in 2003–2006, largely due to the improvements in surgical techniques and combined therapies. There is a clinical evidence of local control and survival advantages for higher dose levels delivered to lung tumours [Wulf 2005], but lung complications have been shown to correlate with the mean lung dose [Kwa 1998]. Due to the improved capability of delivering a highly conformal radiation dose distribution, charge particle therapy can provide substantial benefit for lung cancer treatments. Both protons and carbon ions have been extensively used for the treatment of inoperable NSCLC cancers, obtaining improved clinical results in terms of local tumour control and survival rate [Shioyama 2003, Koto 2004]. However, the advanced active scanning techniques have never been applied to date for lung cancer treatments, due to the high sensitivity to intra-fraction respiratory motion.

As shown in Figure 1.11, lung tumours generally show a great variation in the trajectory of motion induced by breathing. Several studies confirmed that the main motion component is in the SI direction, with lower lobe tumours exhibiting larger motion amplitude than upper lobe tumours. The path followed by lung tumours during inhalation is different from exhalation, leading to hysteresis loops of even several millimetres [Seppenwoolde 2002]. A variable speed is usually observed during the breathing cycle, spending more time near exhalation than inhalation. Tumour position at the exhale phase is found to be more stable than the inhale phase [Seppenwoolde 2002]. Sonke and colleagues [Sonke 2008] quantified the inter-fraction variability of lung tumour trajectories and of respiratory baseline positions in 56 patients with lung tumours, by using 4D cone-beam CT data. On average, a peak-to-peak tumour motion amplitude of 9 mm (range 2.1–25.0 mm) was found. The tumour trajectory shape proved to be stable inter-fractionally, with mean variability not exceeding 1 mm of standard deviation in each direction for the inhale and exhale phases. Large inter-fractional variations in tumour baseline were observed, with 3.9, 2.8 and 1.6 mm as systematic errors and 2.4, 2.2 and 1.2 mm as random components in SI, AP and ML directions, respectively. Liu and colleagues [Liu 2007] assessed intra-fraction variations of respiration-induced motion amplitude for 152 lung cancer patients by analyzing 4D CT data acquired during normal breathing. For 95% of the tumours, the magnitude of target motion was less

than 13.4, 4.0 and 5.9 mm in SI, ML and AP directions. Tumour motion larger than 5 mm along the SI, ML and AP axes was found in 39.2, 1.8 and 5.4% of all patients, respectively.

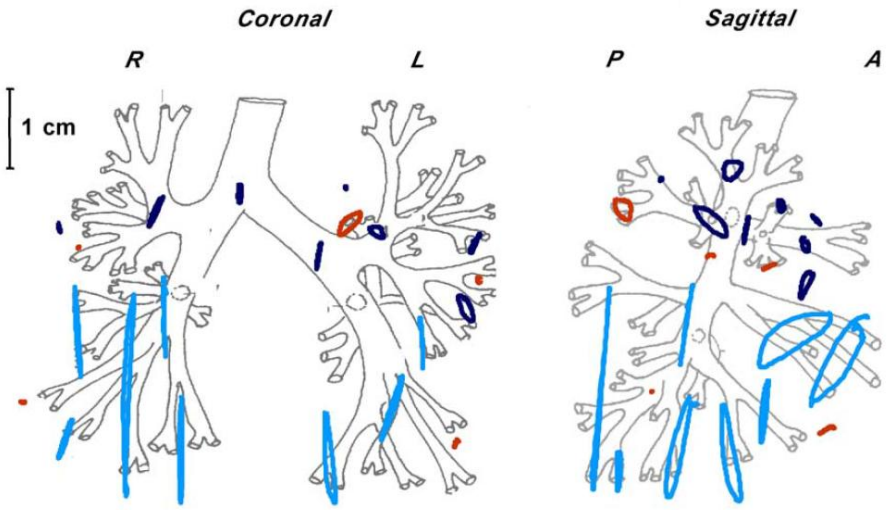


Figure 1.11. Orthogonal projections of the 3D trajectories of 21 lung tumours on the coronal (Right-Left) and the sagittal (Posterior-Anterior) planes, measured using implanted markers and real-time stereoscopic fluoroscopy. Tumours that were attached to bony structures are coloured in red, while upper-lobe and lower-lobe tumours are coloured in dark blue and light blue, respectively [Shirato 2004].

1.3.2 Respiratory motion modelling

A respiratory motion model, representing a spatio-temporal description of the motion of an anatomy of interest due to breathing, is required to explicitly include organ motion in radiation treatment planning and delivery processes, thus reducing the associated geometrical uncertainties. Application of motion modelling include the reduction of motion-induced artefacts in planning or in-room images [Zhang 2007, Rit 2009], simulation of the absorbed dose distribution for treatment planning [Lujan 1999] and prediction of tumour motion during treatment delivery [Sharp 2004]. Various respiratory motion models have been proposed in literature, including signal models, biomechanical models and registration-based models. Signal models represent a mathematical description of the motion trajectory of single points or structures due to breathing. For instance, Lujan et al. [Lujan 1999] modelled respiratory motion using the following modified cosine function:

$$x(t) = x_0 - b\cos^{2n}(\pi t/\tau - \phi), \tag{1.1}$$

where x_0 is the position at exhale, b is the motion amplitude, τ is the fixed period of the breathing cycle and ϕ is the starting respiratory phase. The parameter n controls the shape (steepness and flatness) of the model, allowing for an asymmetric breathing trajectory with more time spent at

exhale than inhale phase, as observed in real patient data [Seppenwoolde 2002]. Another example of signal model is reported by Ruan et al. [Ruan 2009], which propose a method for the real-time profiling of respiratory motion, by decomposing it into baseline, phase and fundamental pattern (amplitude and shape) variations. Baseline and phase parameters are computed by applying ellipse fitting techniques in an augmented state space, whereas fundamental motion pattern are obtained by unwarping the motion signal with the estimated respiratory phase.

Differently from signal models, biomechanical and registration-based approaches describe the internal motion of the entire thoraco-abdominal region. Biomechanical modelling of the lung aims at simulating breathing processes based on physiological and anatomical observations derived from CT and MRI images, usually using finite element methods [Villard 2005]. To this group belongs the NURBS-based Cardiac-Torso Phantom (NCAT), a geometric model of the thorax with a realistic respiratory mechanic and cardiac motion, widely employed as a simulation tool for dynamic imaging and dose deposition [Segars 2001]. Registration-based models rely on intensity-based deformable registration algorithms usually applied to time-resolved CT or MRI images. The respiratory motion of inner lung structures is expressed in terms of deformation vector fields, obtained by applying non-rigid registration to images at different respiratory states. Some authors have proposed to non-rigidly register CT volumes acquired at inhale and exhale phases during breath-hold or free-breathing, and then linearly interpolating the resulting vector fields to predict the intermediate deformations and volumes [Sarrut 2006b, Schreibmann 2006]. This interpolation models are based on the assumption of a linear motion over the respiratory cycle, without taking into account the hysteresis of the breathing process. To overcome this limitation, many authors have proposed to non-rigidly register the volumes from a 4D CT dataset to a reference breathing state image, usually corresponding to the end-exhale phase [McClelland 2006, Zhang 2007, Yang 2008]. Since registration-based models are built from a single average breathing cycle obtained from the 4D CT scan, they have a limited potential to model intra- and inter-cycle motion variability. The accuracy of this modelling approach strongly depends on the accuracy of the applied deformable registration algorithm. Standard registration methods do not take into account the sliding motion of the lung with respect to the chest wall, which may locally reduce matching accuracy [Wu 2008]. On the contrary, biomechanical modelling allows to explicitly model the physical properties of physiological processes, such as sliding motion. Despite the increasingly detailed models, current biomechanical approaches does not include most of the motion information available from anatomical images, resulting in a less accurate registration of the inner lung structures compared to registration-based methods [Sarrut 2007].

1.4 Dynamic tumour localization

In order to allow a proper management of intra-fraction organ motion during a radiotherapy treatment, the continuous and precise information on tumour position is necessarily required. While the proposed strategies for organ motion compensation will be described in details in the next section, here we will focus on the developed techniques for the real-time monitoring of target motion during dose delivery. The time-resolved localization of the moving tumour can be achieved through two main approaches: direct imaging of the lesion or of implanted fiducial markers and indirect tumour tracking by means of surrogate breathing signals [Keall 2006, Riboldi 2012].

1.4.1 Direct tumour imaging

The most widespread image-based approach for the tracking of intra-fraction tumour motion is dynamic in-room X-ray imaging, including MV portal images used in cine-mode, multiple kV radiographs episodically acquired during treatment and kV fluoroscopy, consisting in the continuous acquisition of 2D X-ray projections from a single or dual angular directions with a high temporal resolution [Murphy 2007]. Fluoroscopic imaging is provided both by ceiling/floor-mounted systems, in which the X-ray tubes and the flat-panel detectors are mounted permanently either to the ceiling or in the floor of the treatment room, or by gantry-mounted system, in which the imaging system is mounted on the treatment gantry, usually orthogonal to the central axis of the treatment beam. In the dual orthogonal fluoroscopy, realized through a pair of source-detector systems, the image data from the two views are combined with stereoscopic methods to obtain the 3D tumour trajectories. Fluoroscopic imaging has been applied for the assessment of intra-fraction tumour motion both before and during a radiation treatment for periods ranging from 30 sec to several minutes, with an acquisition rate up to 30 frames per second [Shirato 2000].

Image-based tumour localization methods can be divided in markerless and marker-based [Riboldi 2012]. The markerless approach is based on the direct identification of the moving tumour in the radiographic/fluoroscopic images acquired during treatment. However, soft tissue tumours often present low contrast and undefined contours on X-ray projection images, mainly due to the superimposition of the surrounding high-intensity anatomical structures [Keall 2006]. Although several manual and automatic tracking approaches have been proposed [Lin 2009], markerless tumour detection in X-ray images is usually difficult to achieve, particularly at specific projection angles. This problem can be overcome by using radio-opaque fiducial markers implanted inside or close to the tumour, which are easily visualized in X-ray projections, even with low quality megavoltage imaging [Keall 2004]. Figure 1.12 shows a radiographic image of a lung tumour

containing four metal markers, which are the most common implanted fiducials. Radioactive markers, such as positron emission markers, have also been proposed for tumour tracking applications [Chamberland 2011].

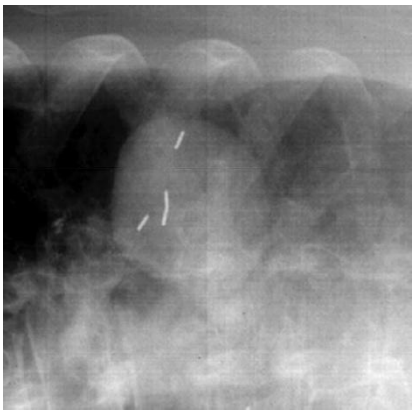


Figure 1.12. A radiographic projection of a lung tumour with four implanted gold fiducial markers, obtained with a flat-panel amorphous silicon detector [Murphy 2004].

Tumour tracking based on internal fiducial markers proved to be highly accurate, with reported localization uncertainty within 2 mm [Shirato 2012]. However, drawbacks associated to marker-based approaches concern the invasiveness of the implantation procedure, which can involve the risk of pneumothorax in case of clip insertion in the lung parenchyma, and the possibility of marker migration from the intended location. Variations in the distance between markers within 2 mm are reported, along with significant changes in the position of the markers relative to the tumour over a few weeks [Imura 2005]. In particle therapy applications, implanted markers can also affect dose distribution, resulting in critical dose deviations up to 82% [Newhauser 2007]. The major disadvantage of X-ray imaging is the extra dose of ionizing radiation delivered to the patient, which is associated to an additional risk of radiation-induced cancer pathology. Murphy and colleagues [Murphy 2007] reported that the imaging dose delivered during a 5 min fluoroscopic study can exceed 100 mGy. Dose constraints limit the applicability of image-based techniques for intra-fraction motion management, due to the impossibility of continuously monitoring tumour position during treatment.

Non-radiographic image-based approaches have also been proposed for intra-fraction tumour motion tracking in radiotherapy treatments. The main non-invasive technique applied for real-time image guidance is ultrasound, which allows soft tissue visualization without additional dose [Hsu 2005]. Ultrasound method usually provides low quality images, whose interpretation is operator-dependent and can be affected by tissue distortion induced by the probe pressure. A remote

approach for ultrasound image acquisition, based on a stable holder for the probe that applies a steady contact force to the patient, has been developed and tested, obtaining motion tracking accuracy of about 2 mm [Schlosser 2010]. In-room MRI scanners integrated with X-ray radiotherapy units have recently been developed for image guidance in radiotherapy. Prototype systems demonstrated that good quality MRI images can be acquired, with minimal interference from treatment irradiation [Fallone 2009]. The potential of non-invasive in-room MRI to capture soft tissue images in real-time during treatment could offer unique opportunities for tumour tracking, although the coupling of MRI scanners with high-energy particle beams poses specific challenges for an application in particle therapy [Raaymakers 2008]. A non-radiographic marker-based alternative for real-time tumour localization relies on implanted transponders, which are detected continuously in three dimensions by means of external electromagnetic receivers with sub-millimetre accuracy [Balter 2005]. To date, this technology have been applied only in prostate cancer radiotherapy, but future clinical applications for the treatment of moving targets are expected, since the US Food and Drug Administration has approved the use of transponders anywhere in the body.

1.4.2 Indirect measurement of breathing surrogates

Indirect tumour localization during radiotherapy treatments can be achieved by inferring its position from breathing surrogates expected to have a close relationship with the target. Since respiratory surrogates are acquired through non-invasive techniques, this approach can be applied to continuously monitor target motion during the whole treatment duration. The use of surrogate breathing signals to predict tumour position requires the modelling of the relationship between the target displacement and the respiratory surrogate. Since a non-zero phase-offset is usually observed between the surrogate and the internal tumour signals, models that compensate for phase shifts between the two motions should be taken into account to obtain a reliable target prediction [Hoisak 2004]. Extensive analysis of tumour trajectories also suggests that non-linear correlation and hysteresis need to be modelled to accurately depict target displacement [Pepin 2010]. Several correlation models with different complexity have been proposed, ranging from piece-wise linear-polynomials to advanced machine learning methods, such as fuzzy systems and artificial neural networks [Torshabi 2010, Seregni 2011]. All studies show that the correlation is patient-specific and time-dependent, due to the non-stationary characteristics and transient changes of breathing motion. This requires frequent verification of the model estimation and on-line adaptation of the correlation parameters to encompass intra-fraction breathing irregularities during treatment

[Ozhasoglu 2002]. The accuracy and robustness of surrogate-based approaches for dynamic tumour localization can be improved by periodically updating the correlation model based on the acquisition of X-ray images of the tumour position synchronously with the breathing surrogate [Kanoulas 2007]. Examples of surrogate signals applied for radiotherapy treatments include the respiratory volume, the displacement of the external thoraco-abdominal surface and the motion of internal structures, such as the diaphragm. The performance of the different respiratory surrogates has been evaluated and compared by analyzing the correlation with the tumour displacement or by calculating the target localization errors after prediction.

The volume of inhaled and exhaled air is a popular choice of respiratory surrogate signal, since it is physiologically related to breathing motion. During normal respiration, the lung volume typically changes by 20% (from 3.3 to 4.1 Litres on average), whereas at deep inspiration, the increase in lung volume is approximately 3 to 4 times that of normal breathing [Keall 2006]. Spirometric devices are usually applied to non-invasively measure the flow of air entering and exiting the lungs by means of differential pressure pneumotachographs [Mah 2000]. As depicted in Figure 1.13a, the spirometer is composed by a flexible arm ending with the pneumotachographic sensor, an antibacterial filter and a mouthpiece. Video-glasses showing the acquired breathing traces are often provided to the patient for visual guidance (Figure 1.13b). A daily calibration procedure, performed with a syringe containing a known volume of air (Figure 1.13a), is required to convert the flow measurements in respiratory volumes. The accuracy of volume estimation is typically ±2% for flow rates ranging approximately between 0.5 and 3.0 L/s. Examples of commercial spirometry clinically employed during radiotherapy treatments are the VMAX Spectra 20C (VIASYS Healthcare Inc, Yorba Linda, CA) and the SpiroDyn'RX (Dyn'R, Muret, France) [Haack 2009]. A drawback of the volume-based approach for dynamic tumour localization is that patients can have difficulty in tolerating spirometry for long periods of time. Moreover, spirometric measurements can be subject to time-dependent drifts of the end-exhale and end-inhale values due to escaping air and instrumentation errors. The use of drift-free surrogate signal, such as the internal air content measured in CT images, has been proposed to correct for spirometric drifts [Lu 2005]. In a 11-patients dataset, the correlation coefficients between tumour displacement and respiratory volume ranged from very high (0.99) to very low (0.51), finding an inconsistent surrogate-target relationships over multiple days for most patients [Hoisak 2004].

Different anatomical structures inside the thorax, such as diaphragm, carina or pulmonary veins, have also been used as surrogate breathing signals for the estimation of tumour position [Van der Weide 2008]. The motion of the internal structures can be tracked through image-based methods, including X-ray fluoroscopy [Cerviño 2009a], time-resolved MRI [Von Siebenthal 2007]

and non-invasive ultra-sound imaging [Xu 2006]. The most common internal surrogate principally applied for lower lobe lung cancers, liver and pancreas tumours is the diaphragm, which represents the main inspiratory muscle. During inhalation, the contracted diaphragm descends and the abdomen is forced inferiorly and anteriorly, increasing the SI dimension of the chest cavity. The diaphragm position in the superior-inferior direction is easily detectable in X-ray projections, as well in ultrasound images, as shown in Figure 1.14. The correlation analysis between tumour and diaphragm motion in fluoroscopic image sequences from 10 lung cancer patients shows average correlation factor of 0.98 when phase delays are taken into account [Cerviño 2009a]. The average localization error for tumour prediction proved to be 0.8 mm, with an error at the 95% confidence level of 2.1 mm. However, for selected patients a much weaker correlation was found, suggesting that tumour-diaphragm relationship should be assessed on a patient-specific basis.

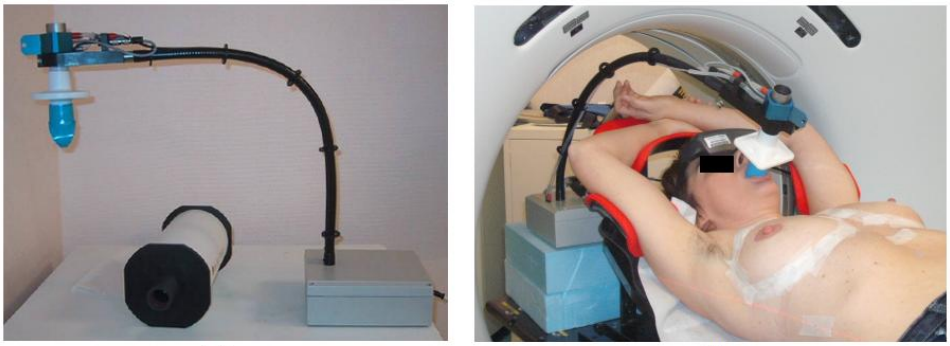


Figure 1.13. SpiroDyn'RX spirometer clinically employed during radiotherapy treatments, with the 3 L syringe used for daily system calibration (Panel a). The patient breathes through a mouthpiece connected via flexible tubing to the spirometric sensor (Panel b).

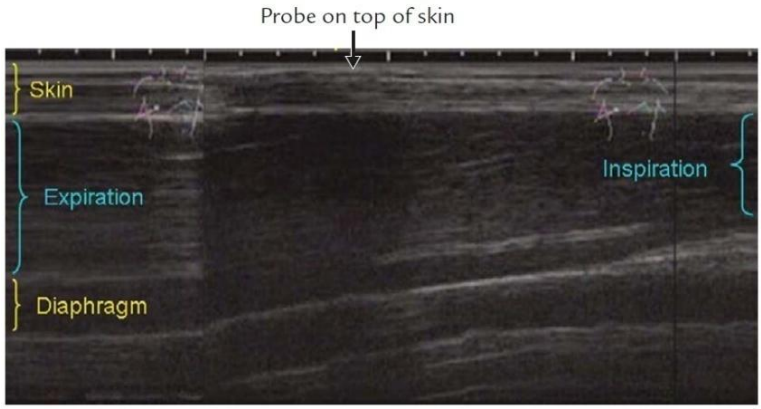


Figure 1.14. Ultrasound image obtained from the right-side abdominal rib cage, showing the diaphragmatic motion during inspiration and expiration phases [Wang 2009].

The most commonly used surrogate breathing signal for tumour localization in radiotherapy treatments is the displacement of the patient's thoraco-abdominal surface. The intercostal muscles connecting adjacent ribs contract actively during inhalation, pulling the ribs superiorly and anteriorly, thus increasing both the ML and AP diameters of the thorax and abdomen. During exhalation, the lung and chest wall return passively to their positions at rest, due to elastic properties. The displacement of the external surface induced by breathing can be acquired non-invasively by means of optical localization systems with non-ionizing energies, thus allowing the continuous motion monitoring during the whole treatment duration. Optical tracking consists in localizing an object by measuring the light either emitted or reflected by the object, usually operating in the infrared range to avoid the interference of the ambient light [Meeks 2005]. Surface motion can be tracked directly by using markerless surface imaging systems, which capture the entire thoraco-abdominal surface of the patients through video-cameras [Smith 2003, Bert 2005]. Indirect surface tracking can also be performed, by recording the position of fiducial markers placed on the patient's skin through infrared opto-electronic systems [Spadea 2006, Baroni 2007]. Marker-based and markerless optical surface tracking will be described in details in the next sections.

Several respiratory surrogates can be derived from both marker-based and markerless optical surface tracking, including the trajectory of one or more surface points, the diameter in AP direction, the cross-sectional area or the volume enclosed by the surface [Romei 2010]. Opto-Electronic Plethysmography (OEP) techniques allows for example the accurate measurement of the volumetric variations of the chest wall and its compartments (rib cage and abdomen), starting from the 3D coordinates of fiducial markers positioned on the subject’s thoraco-abdominal surface and acquired by an opto-electronic system [Cala 1996]. After defining a geometrical model that connects the reconstructed surface points to form triangles, the enclosed volume and its changes due to breathing are calculated by surface integration using the Gauss theorem. Volumetric surrogate signals derived from both marker-based and markerless surface tracking proved to be highly correlated with spirometric volume measurements [Aliverti 2001, Tarte 2006]. Unlike spirometry, surface-derived volume signals are not subject to time drifts and provide a compartmental analysis of surface motion, allowing to differentiate between diverse modes of respiration, such as abdominal or thoracic breathing patterns [Hughes 2009].

The use of external surface surrogates entails the definition of appropriate external-internal correlation models to infer target localization. Composite breathing signals obtained from multiple surface control points were proposed to improve the robustness of the correlation between surface and tumour motion [Yan 2006, Yan 2008]. Different locations of surface markers have been compared for the prediction of the internal motion of lung structures, including tumours or veins,

under different types of breathing (deep, shallow, abdominal, thoracic, etc.). The marker location showing better correlation depended on the position and direction of the internal structure and on the patient's respiratory pattern [Koch 2004, Yan 2008]. The overall accuracy of target position estimation from external surface motion was quantified in retrospective clinical studies for 44 lung cancer patients [Hoogeman 2009]. The mean errors of the external-internal correlation model were less than 0.3 mm, with maximum standard deviations, describing intra-fraction variations around the whole-fraction mean error, of 2.5 mm for AP direction and 1.9 mm for ML and SI directions. The use of correlation models featuring a higher level of complexity than linear-polynomial correlation was reported to increase the accuracy in target position estimation, particularly in presence of breathing irregularities [Torshabi 2010].

1.4.2.1 Marker-based optical tracking

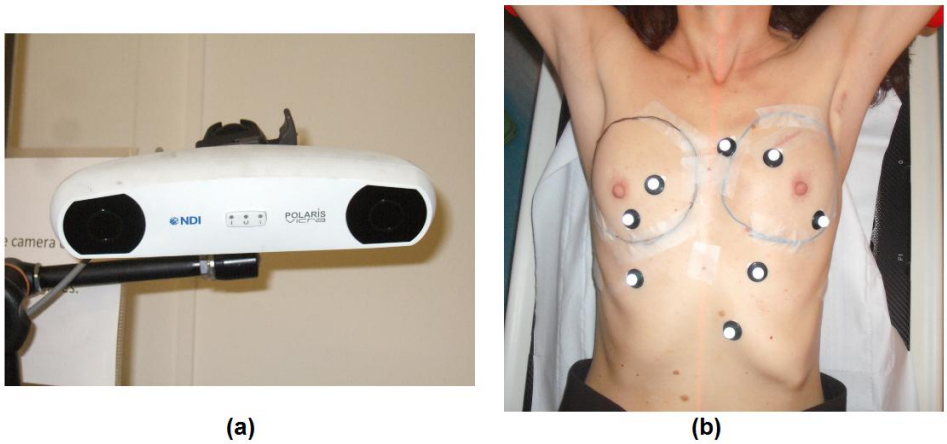
Opto-electronic localizers, such as the ELITE or SMART systems (BTS Engineering, Milan, Italy), measure in real-time the three-dimensional coordinates of active light-emitter or passive light-reflecting markers with a sub-millimetre accuracy at an acquisition rate up to 100 frame per second [Ferrigno 1994]. Passive markers are generally 6 mm diameter spheres coated with a highly reflective surface that reflects the infrared light emitted from an external illuminator. The detectors of the emitted or reflected marker light consist in at least two solid-state Charge-Coupled Device (CCD) cameras with an enhanced sensitivity in the infrared spectrum. A dedicated processor recognizes and calculates the 2D positions of the markers recorded by the camera sensors, using pattern recognition algorithms [Ferrigno 1990]. The reconstruction of the 3D marker coordinates is based on stereoscopic techniques, requiring the calibration of the physical and geometrical parameters of the CCD cameras. This is performed by using a calibration apparatus composed by passive markers at known distances and placed at known coordinates relative to the isocentric reference frame of the radiotherapy machine. The three-dimensional marker location is found through triangulation as the intersection of the 3D rays determined by each camera, passing through the detected 2D marker positions and the calibrated camera optical centres. The calibration procedure provides also the transformation matrix to convert camera coordinates into room coordinates, thus allowing to determine the 3D marker positions with respect to the treatment isocenter.

One of the first commercially available optical tracking system for radiotherapy is the Real-time Position Management (RPM) system (Varian, Palo Alto, California, USA) [Ford 2002]. As shown in Figure 1.15a, a plastic block with infrared reflective markers is placed on the patient’s abdominal surface, typically midway between the xyphoid process and the umbilicus to maximize

the respiratory-induced motion. The AP trajectory of the block is captured by a single CCD camera, usually mounted on the ceiling of the treatment room (Figure 1.15b). The Polaris system (Northern Digital Inc, Waterloo, Ontario, Canada) is another commercial example based on optical tracking technology [Bova 1997]. The sensor unit (Figure 1.16a), consists in an array of two CCD cameras that track the 3D position of a set of fiducial markers either attached to the patient's skin or to a holder. Figure 1.16b depicts a configuration of multiple passive markers placed on the thoraco-abdominal surface of a patient.

(a) (b)

Figure 1.15. RPM optical system for the tracking of external surface motion. The device is composed by a plastic block with passive reflective markers placed on the patient's abdomen (Panel a). The AP trajectory of the block is recorded by a ceiling-mounted CCD camera, equipped with infrared light-emitting diodes for marker illumination (Panel b).



(a) (b)

Figure 1.16. Polaris sensor unit, composed by an array of two CCD cameras for optical marker tracking (Panel a). Configuration of multiple passive markers placed on different anatomical landmarks of the patient's thoraco-abdominal surface (Panel b).

Marker placement is usually performed manually by selecting appropriate anatomical features, such as nevi or scars. Care must be taken to ensure the reproducibility of marker location on the patient's skin, to allow the comparison of the motion signals obtained during different

treatment sessions. Marker-based optical systems allow the fast real-time 3D localization of corresponding control points solid to the patient surface, thus yielding motion tracking at specific anatomical landmarks. However, they do not provide a complete description of the anatomical area involved in the irradiation, since the motion monitoring is under-sampled only at specific surface landmarks [Spadea 2006]. Moreover, they typically involve long patient preparation and may not be repeatable due to inaccuracy in marker placement [Wang 2001].

1.4.2.2 Markerless surface imaging

An alternative option for measuring the displacement of the chest and abdomen is based on markerless optical surface imaging techniques that capture the whole thoraco-abdominal surface without the use of passive markers. The clinical application of body surface sensing in radiotherapy was proposed by Moore and Graham [Moore 2003], who described a method based on laser interferometry, consisting in the projection of interferometer fringes on the patient's skin to generate a height map of the external surface. Commercial laser-based surface scanning devices for radiotherapy application include the Sentinel system (C-RAD AB, Uppsala, Sweden) [Brahme 2008] and the Galaxy system (LAP Laser, Lüneburg, Germany) [Moser 2011]. These devices allow the real-time reconstruction of a 3D surface model by scanning the patient with a laser beam, while a camera records its reflections. The geometrical resolution in surface reconstruction is about 0.1 mm [Brahme 2008], whereas the accuracy in surface motion estimation tested on healthy volunteers proved to be lower than 1.7±1.5 mm [Moser 2011]. Further methodological investigations were focused on the evaluation of video-based systems for the instantaneous 3D surface imaging through stereo-photogrammetric techniques [Smith 2003]. The projection of a light-pattern on the patient's skin provide a high density 3D topological model of the monitored surface, with few millimetres spacing between surface points. The reconstructed surface model is defined by vertices connected into triangular faces. Each triangular face is parameterized by three 3D vertices for the surface topology and by three direction cosines of the triangle normal for surface orientation.

Markerless optical surface imaging allows to overcome the problems of marker-based optical techniques, related to the reproducibility of marker replacement and to the under-sampling of surface motion. Surface imaging provides a complete description of the patient's thoraco-abdominal region, with redundant information accounting for inter-session surface deformations [Spadea 2006]. Potential drawbacks of surface sensing methods are the computational bulkiness, which limits the real-time performance, and the lack of spatial correspondence between points of subsequently acquired surfaces. Appropriate methods based on surface registration are required to derive a breathing surrogate signal from markerless optical surface acquisitions. In case of low

spatial frequency of the controlled anatomical surface, such as the thorax and abdomen, registration algorithms may feature a relatively low robustness due to local minima in the iterative surface fitting process [Baroni 2003]. Registration robustness can be improved by using constrained surface fitting, which represents a trade-off between the overall surface congruence and the matching of selected landmarks [Riboldi 2009].

The most popular commercially available video-based surface imaging system applied in radiotherapy is the VisionRT (Vision RT, London, UK), whose technical details are extensively described by Bert et al. [Bert 2005]. As depicted in Figure 1.17a, the system is composed by two imaging pods suspended from the treatment room ceiling. Each pod is equipped with two CCD cameras for stereo-vision, a texture camera and a speckle projector. 3D surface reconstruction is performed by projecting a pseudo-random speckle pattern on the patient's skin (Figure 1.17b), which facilitates the identification of corresponding surface points on the images acquired by the two cameras of the imaging pod. The depth values for all 2D points covered in both stereo-images are triangulated using the baseline parallax between the camera sensors.

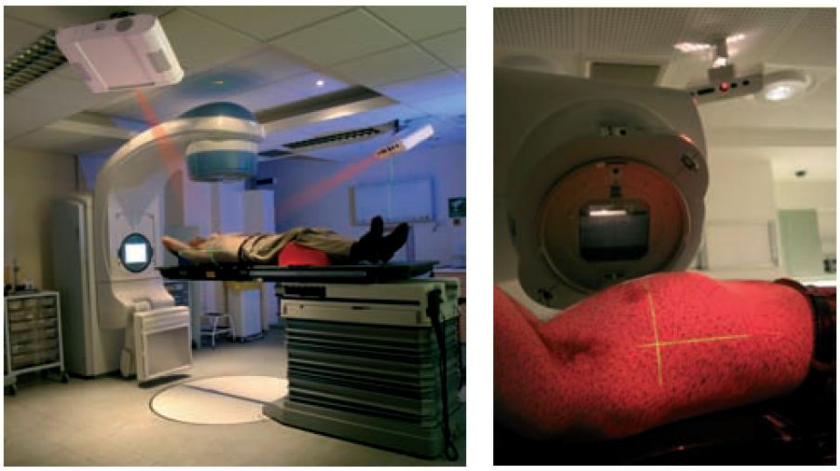


Figure 1.17. VisionRT surface imaging system, composed by two imaging pods fixed to the room ceiling at symmetrical positions with respect to the treatment couch (Panel a). Each pod projects a red and black random pattern on the patient's skin (Panel b), used for the reconstruction of the 3D surface model through stereo-photogrammetric techniques.

A calibration procedure of the VisionRT system is required to compute the optical properties as well as the position and orientation of each camera with respect to the isocenter, which is performed by means of a planar grid of black spots. During system calibration, the grid is positioned on the treatment couch aligned with the room lasers, and grid images are acquired for multiple couch heights. Each imaging pod acquires 3D surface data over approximately 120° in the

axial plane. In order to obtain a field of view of 180°, data from the two pods are merged in a single integrated 3D surface image, with less than 1 mm of root-mean-square discontinuity error in the overlap region near midline. As shown in Figure 1.18a, the surface model acquired with the VisionRT system is defined by triangular tiles with adjustable vertex separation (typically 6 mm). The additional texture camera capture a gray-scale image of skin features, that can be superimposed to the surface topology to generate a 3D textured model (Figure 1.18b).

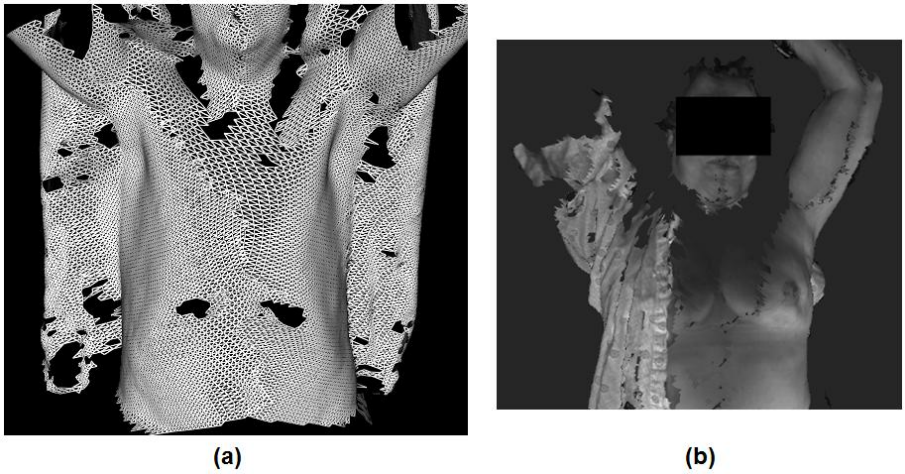


Figure 1.18. Representative examples of 3D surface model acquired with the VisionRT system, defined by vertices connected into triangular faces (Panel a), and of 3D textured surface, in which each triangular face is coloured with a gray-level video image of the patient's skin captured by the texture camera (Panel b). The merging line connecting the surface models acquired by the two imaging pods is visible in both figures.

The VisionRT system can operate in two distinct principal modes. The AlignRT program is mainly used for setup verification to correct patient position before treatments. The daily acquired surface is compared to the reference model through rigid surface registration algorithms, computing the translational end rotational couch corrections that minimize the distance between the two surfaces. AlignRT program allows the real-time reconstruction of the whole 3D surface or of a region of interest selected by the user. Usually, a single static frame is acquired with texture information, merging data captured by both imaging pods. This operation mode also provides the possibility to gate the surface acquisition at a specific phase of the respiratory cycle, tracking the trajectory of a point chosen on the patient surface. The captured volume around isocenter is about 650×1000×350 mm (ML×SI×AP directions), with a typical number of 3D points per surface model of about 10000-20000. According to validation studies performed on rigid phantoms, the 3D surface reconstruction with the AlignRT system features a sub-millimetric accuracy, obtaining a root-mean-square error of 0.65 mm if compared with CT-derived surface topology [Bert 2005]. In the second operation mode of the VisionRT system, surface data are captured continuously from a

single imaging pod at an acquisition rate of about 7.5 frames per second, depending on the size of the monitored surface. In this modality, only the photos captured by the two stereo-cameras are stored in real-time, while the 3D surface reconstruction is performed off-line from the stored photos. Texture mapping and selection of regions of interest (ROI) are not available in the continuous acquisition mode. Phantom studies on a moving surface with a sinusoidal trajectory at different combination of amplitudes (1–5 mm) and periods (2–6 sec) indicate that the system accuracy in the measurement of motion amplitude is 0.11±0.15 mm, with maximum absolute deviations of 0.83 mm [Bert 2005]. The dynamic capture mode include the GateCT software, mainly dedicated to track the external surface motion during planning image acquisition for retrospective 4D CT reconstruction, and the GateRT software, that can be used for real-time monitoring of respiratory motion and patient movements during treatment delivery [Johnson 2004].

1.5 Intra-fraction motion management

This section discusses in detail the different methods for the management of intra-fraction breathing motion during radiotherapy treatment delivery. Several immobilization techniques are applied for the stabilization of the moving tumour and for the reduction of organ motion. Abdominal compression is a reported option since the pressure applied to the abdomen reduces diaphragmatic excursions, while still permitting limited normal respiration [Negoro 2001]. A dedicated abdominal pressing plate can be integrated in the stereotactic body frame used for patient immobilization, as shown in Figure 1.19. The effectiveness of the compression device in the management of respiratory movements has been assessed in 27 lung cancer patients using 4D CT and dosimetric studies [Bouilhol 2012]. The most significant impact of abdominal compression was obtained in patients with lower lobe tumours, where motion amplitude was decreased by 3.5 mm on average. Minor or negative effects of compression were reported for upper/middle lobe locations, obtaining a mean tumour motion reduction of 0.8 mm. Compression increased tumour motion in 5 cases, whereas the dosimetric gain for lung sparing was not clinically relevant. As described in Section 1.2.2, intra-fraction respiratory motion can be accounted by adding treatment margins that encompass the tumour motion during the entire breathing cycle. However, especially for tumours with a large motion amplitude, the use of margins is not ideal since it increases the radiation field size and consequently the volume of healthy tissues exposed to high doses. Alternative strategies for the effective management of intra-fraction tumour motion have been proposed, which can be mainly divided in the following categories: breath-hold methods, respiratory gating techniques and real-time tumour targeting.

Figure 1.19. Stereotactic body frame and abdominal compression plate used for patient immobilization and suppression of internal breathing motion.

1.5.1 Breath-hold method

The reduction of tumour motion amplitude during treatment delivery can be achieved through specific breathing maneuvers that modify the respiratory pattern of the patient, such as breath-hold. In this approach, the radiation dose is delivered while the patient holds his breath at a specific phase of the respiratory cycle, thus providing tumour immobilization and compensation for breathing motion. Deep-inspiration breath-hold (DIBH) is usually performed due to the potential dosimetric advantages of increasing the lung volume. Lung density is in fact reduced during inhalation, thus ensuring a greater sparing of healthy pulmonary tissues [Hanley 1999]. Breath-hold methods have been predominantly applied to lung cancer and left-breast radiotherapy. Although intra-fraction breathing motion can be considered negligible for tumours in the breast region [Smith 2005], the inspiration breath-hold maneuver allows to reduce both cardiac and lung toxicity with respect to free-breathing irradiation [Pedersen 2004]. The distance between the heart and the breast is in fact increased during inhalation, thus minimizing the cardiac volume included in the irradiation field.

Multiple repeated DIBH maneuvers are typically required both during the acquisition of planning CT images and during dose delivery at each treatment fraction. The inter- and intra-session repeatability of breath-holds therefore represents a crucial issue, which entails the application of an appropriate technology for DIBH monitoring. Different methods have been proposed to acquire patient's breathing signal and to assess the reproducibility and stability of breath-hold maneuvers. The respiratory signal allows to enable and disable dose delivery, automatically stopping the radiation beam if the inspiratory level deviates from the intended one. The most common monitoring technique is based on the measurement of the volume of air inhaled

and exhaled by the patient through spirometric devices [Hanley 1999, Mah 2000], described in Section 1.4.2. External surface surrogates have also been used for respiratory monitoring in breath-hold irradiation. Surface motion during DIBH maneuvers has been acquired both with marker-based optical tracking approaches, such as the RPM system [Vikström 2011] or configuration of multiple control points [Stock 2006], and with markerless surface imaging systems, such as the VisionRT [Cerviño 2009b]. The reproducibility and stability of DIBH maneuvers proved to significantly increase with audio-visual guidance provided to the patient [Stock 2006, Cerviño 2009b].

Two different breath-hold strategies can be applied in radiotherapy treatments: self-held breath-hold and active-breathing control. In self-breath holding methods [Kim 2001] the patient voluntarily holds his breath at a specific phase of the respiratory cycle. A training session is usually performed before treatment to instruct patients to reproduce repeated breath-holds by inspiring the same amount of air for about 10–20 seconds, according to patient's ability. A reference inspiratory level and a tolerance interval are selected, generally as a specific percentage of the maximum capacity reached during training. In treatment planning and delivery, the patient is verbally coached to achieve the predefined breath-hold level, often with the aid of video goggles displaying the breathing traces acquired by the monitoring device and the selected tolerance band (Figure 1.20). To improve breath-hold reproducibility, the Active-Breathing Control (ABC) system was developed [Wong 1999]. The ABC apparatus consists of a spirometer that can control patient's inhaled and exhaled air flow via occlusion valves. At preselected respiration levels, the valves are automatically closed to stop patient's breathing at the desired inspired volume. A moderate breath-hold level, set for example at 75% of deep inspiratory capacity, was demonstrated to achieve good reproducibility of internal organ displacement while maintaining patient comfort [Remouchamps 2003].

Although treating at deep-inspiration breath-hold has been shown to have considerable dosimetric advantages, significant concerns are addressed to the low reproducibility of internal anatomy and tumour position in repeated maneuvers performed at constant inspired volume. Residual motion can in fact occur during breath-holds, due to muscle relaxation. Moreover, the complexity of breathing processes, especially during forced inspiration, can involve the composed movements of the abdomen and the thoracic cage [Aliverti 2001]. This means that the same respiratory level can be reached with a different combination of thoracic elevation in CC direction and abdominal filling in AP direction. The reproducibility of moderated DIBH using an ABC-based monitoring was evaluated in lung cancer patients, obtaining a non-negligible inter-maneuver variability in tumour position, with maximum shifts of about 3 mm in all spatial directions [Cheung 2003, Muralidhar 2008]. The applicability of breath-hold methods is also limited by patient's compliance. Approximately 60% of lung cancer patients cannot perform adequate reproducible

maneuvers or cannot achieve sufficient lung inflation, required to have a dosimetric advantage with respect to free-breathing irradiation [Keall 2006]. A further drawback of DIBH treatments consists in the time requirements, since each treatment session usually takes 5–10 minutes longer than a similar beam arrangement performed in free-breathing conditions.

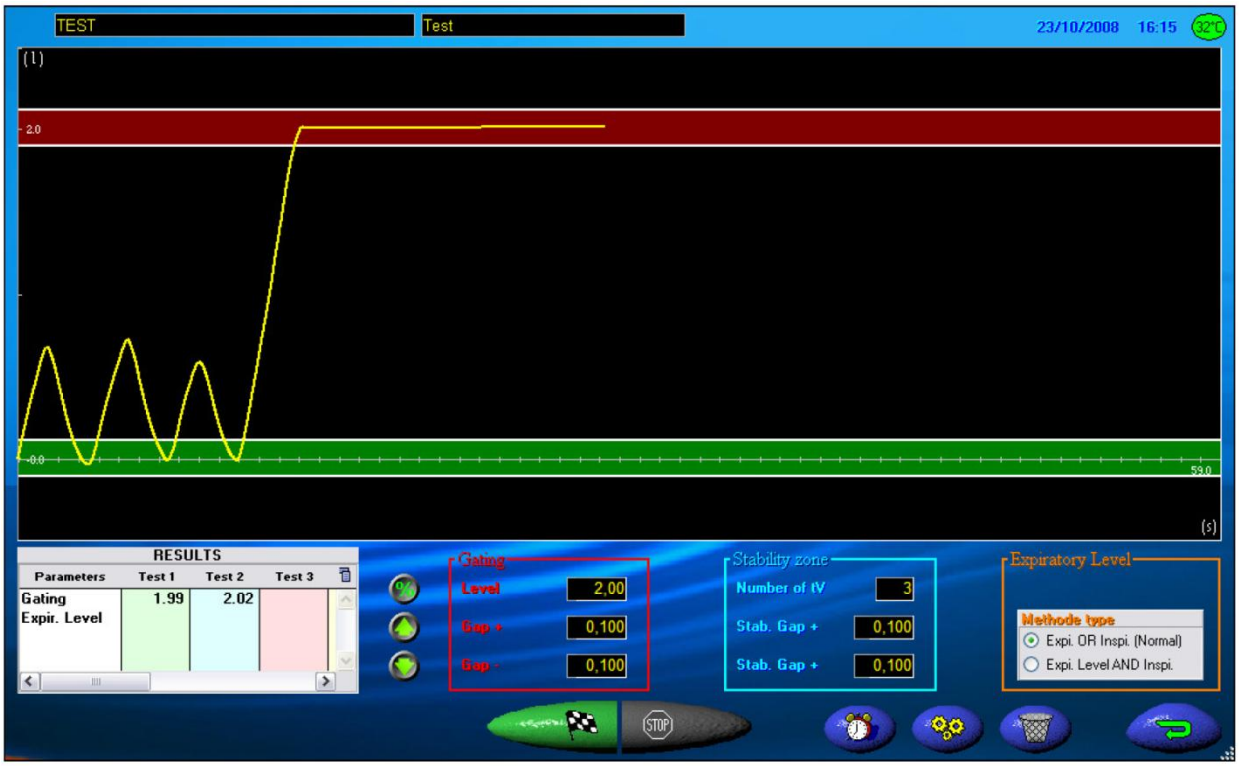


Figure 1.20. Graphical user interface of the SpiroDyn'RX spirometer, showing the patient's breathing curve during a DIBH maneuver. The breath-hold reference level is displayed in red, with a tolerance interval of ±0.1 L.

1.5.2 Respiratory gating

Respiratory gating involves the administration of radiation (during both imaging and treatment delivery) within a specific portion of the patient's free-breathing cycle, commonly referred to as the gate [Ohara 1989]. The radiation beam is automatically switched on and off during certain time intervals, synchronously with the patient's respiration. The end-exhale phase offers preferential conditions, since exhalation is the most reproducible respiratory state and features longer duration [Seppenwoolde 2002]. Respiratory gating requires the monitoring of the patient's breathing motion during treatment delivery, which is usually obtained through external surrogate signals derived from optical surface tracking systems. Due to the non-invasive nature and to the minimal effort required by the subject, gating combined with an external motion signal can be applied to almost all (>90%) patients [Keall 2006]. Currently, the most widespread respiratory

gating system based on a surface-based surrogate signal is the RPM system [Ford 2002, Yorke 2005]. In a gating or breath-hold approach for motion compensation, the tumour-surrogate correlation does not need to be known explicitly, since the surrogate signal is used to monitor the breathing motion, but not to predict the absolute tumour position [Keall 2006]. However, the intra- and inter-fraction correspondence between internal anatomy and external signal used for gating should be accurately verified by acquiring the tumour position through in-room X-ray imaging [Berbeco 2005a].

As shown in Figure 1.21, respiratory gating technique can be based on two different breathing variables (amplitude or phase) derived from the surrogate signal provided by the motion-monitoring device [Vedam 2001]. In amplitude-based gating, dose delivery is activated whenever the respiration signal is within a predefined window between the inhale and exhale extremes of the breathing motion. In phase-based gating, the radiation beam is triggered when the respiratory phase, extracted from the breathing signal through specific algorithms, is within a preset angular phase window. As the beam is not continuously delivered, gated procedures are longer than non-gated procedures. The ratio of the beam-on time within the gate to the overall treatment time is referred to as the duty cycle and represents a measure of the efficiency of the method. Typical duty cycles of 30–50% lead to 4 to 15-fold increase in the delivery time with respect to free-breathing treatments [Keall 2006]. Since the patient is freely breathing in gated radiotherapy, a tumour residual motion still occurs within the gate. The residual tumour motion in gated radiotherapy with external respiratory surrogates has been assessed on lung cancer patients, obtaining a 95^{th} percentile range of residual motion of 0.7–5.8 mm, 0.8–6.0 mm, and 0.9–6.2 mm for 20%, 30% and 40% duty cycle, respectively [Berbeco 2005b]. Amplitude-based gating usually results in less residual tumour motion. The choice of the gate width is a trade-off between the amount of residual motion and the duty cycle. The less residual motion the greater the dosimetric benefit of gated treatment, but the smaller the duty cycle the greater the increase of the treatment time [Vedam 2001].

Respiratory gating technique is currently applied by several clinical centres to account for respiratory motion during conventional radiotherapy of thoracic and abdominal tumours [Mageras 2004, Giraud 2006]. Kubo and Wang [Kubo 2000] demonstrated the feasibility of gating the linear accelerator during a dynamic multi-leaf collimator delivery, thus allowing gated IMRT treatments. Gating has been recently introduced also in particle therapy for several sites influenced by respiratory motion, with an emphasized use for lung cancer and hepatocellular lesions [Minohara 2000, Lu 2007]. Multiple beam on/off cycles during treatment delivery have to be supported by the accelerator and by the energy-selection system through dedicated extraction techniques and intensity modulation [Bert 2011]. Gating is not yet used in particle centres with active scanning

techniques since interplay effects of residual motion within the gate window and the scanned beam can lead to unacceptable tumour under-dosage [Bert 2009]. To overcome the dosimetric influence of the interference effects, gating need to be combined with other dedicated motion compensation methods, such as by increasing the overlap of neighbouring pencil beams in the lateral and longitudinal directions [Bert 2009] or by applying rescanning techniques [Furukawa 2010].

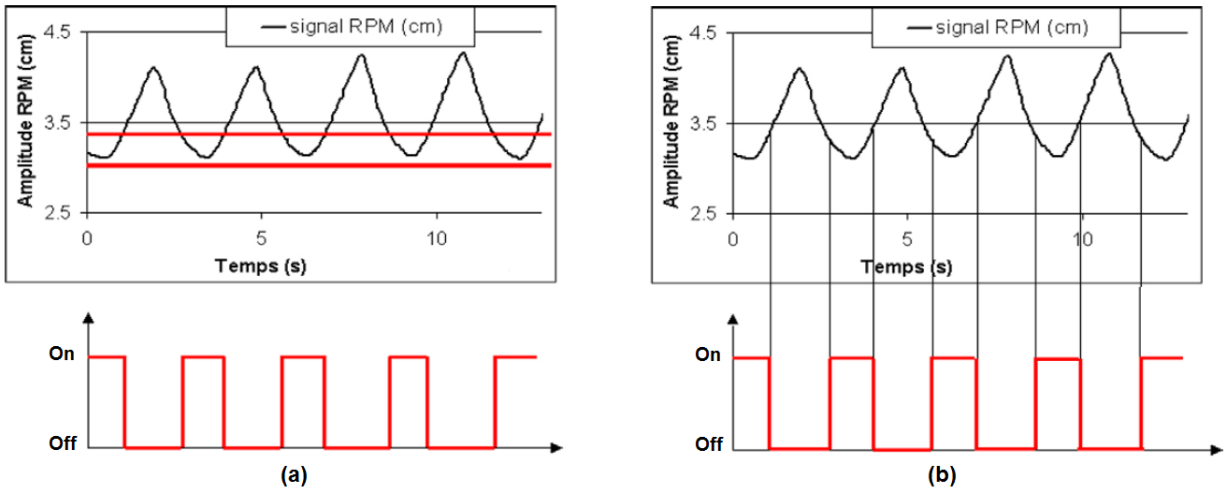


Figure 1.21. Illustration of two types of respiratory gating performed during patient's expiration using the surface-based motion signal derived from the optical RPM system. Synchronization can be based on amplitude with an horizontal gate (Panel a) or on phase with a vertical gate (Panel b). The beam is delivered only when the RPM signal is within this gate (Beam On/Off) [Giraud 2006].

Rescanning (or repainting) is a motion-compensated delivery strategy for scanned particle beams, based on a statistical dose averaging effect by multiple irradiations of the expanded PTV per treatment fraction with a proportionally reduced dose [Phillips 1992]. Assuming that target and beam motion are uncorrelated, the variance of the average dose decreases with a factor of about $1/\sqrt{N}$ (where N is the number of rescans) compared to a single irradiation. If a sufficient number of rescans are applied, the averaging effect will lead to homogeneous dose coverage of the CTV and to a blurred dose distribution in the region of the internal margins. Several rescan options have been proposed in the last few years, including rescanning of the single iso-energy slices and rescanning of the entire volume [Furukawa 2007, Zenklusen 2010]. Specific technological developments to avoid temporal synchronism between beam scanning and target motion are required to guarantee adequate averaging. For instance, Rietzel and Bert [Rietzel 2010] propose random pauses throughout the irradiation process and/or random modulations in the beam extraction, leading to changes in the scanning speed. A more effective solution is to incorporate the breathing pattern in the rescanning process by using the data from a motion-monitoring system, with the so-called

phase-controlled rescanning technique [Seco 2009, Furukawa 2010]. By spreading the rescans throughout the respiration period, irradiation is ensured to be executed in each motion phase of the respiratory cycle. As shown in Figure 1.22, simulation and experimental results clearly demonstrated that phase-controlled rescanning leads to more robust treatment delivery close to the expected one, obtaining high dose homogeneity level within the CTV with a minimal increase in the treatment time [Seco 2009]. The comparison between static and moving measurements showed dose errors less than 2% for pinpoint chambers in the target volume [Furukawa 2010].

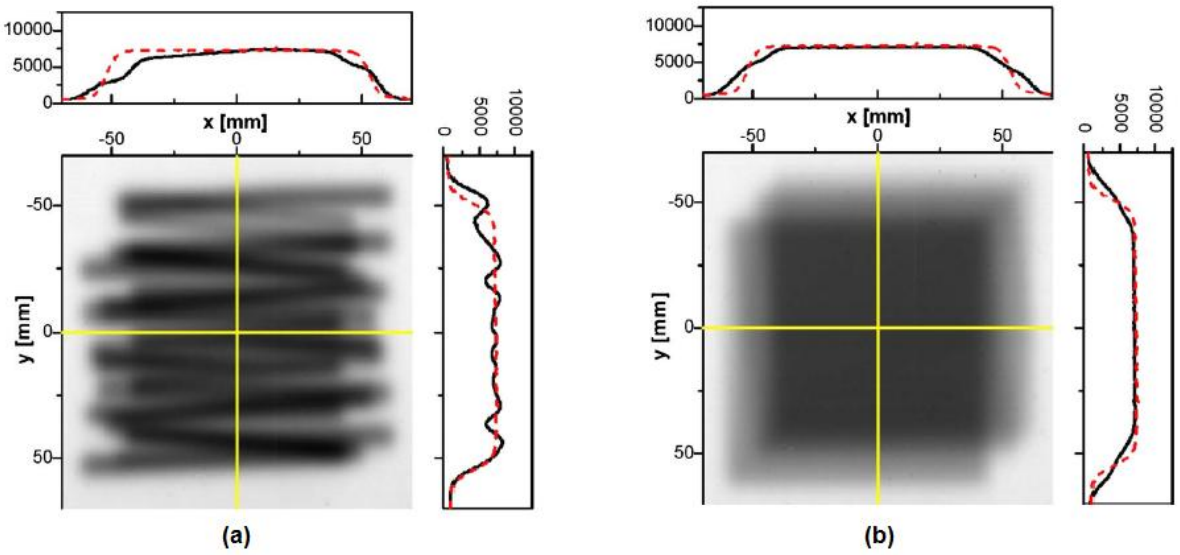


Figure 1.22. Experimental results of deterioration of the lateral dose distribution in uniform scanning of an iso-energy slice without rescanning (Panel a) and with ten fast rescans during two motion cycles, matching the irradiation time with the motion phases (Panel b). The dashed line shows the results without the motion [Furukawa 2010].

1.5.3 Tumour targeting

The most efficient and potentially most precise motion compensation technique in free-breathing conditions is real-time tumour targeting, consisting in continuously following the tumour's changing position through the dynamic adjustment of the radiation beam [Murphy 2004]. Under ideal conditions, the targeting approach can eliminate the need for internal tumour motion margins in the dose distribution, while maintaining a 100% duty cycle for dose delivery, thus with no associated increase in the treatment time [Keall 2006]. Tumour targeting requires the identification of the target position in real-time and the fast repositioning of the radiation beam accordingly, anticipating tumour motion to compensate for system latency and time delays associated to target localization and beam adjustment. The real-time detection of the target position during treatment is the most important and challenging task in tumour targeting. As described in

Section 1.4.1 and 1.4.2, the time-resolved localization of the moving tumour can be achieved through direct imaging of the lesion or of implanted fiducial markers and through indirect measurements of surrogate breathing signals. Differently from gating and breath-hold approaches, targeting techniques require the knowledge of the tumour-surrogate correlation, since the surrogate signal is used to estimate the absolute target position. Adaptive filters and prediction algorithms are usually required to anticipate tumour position, so that the beam adjustment can be synchronized with the exact tumour location [Sharp 2004]. This step is required to compensate for the time delays associated to target detection, such as image acquisition and processing, and to the latency of the beam-positioning system in response to a change in tumour position. For system delays less than 0.4 seconds, the respiratory motion can be predicted with an average error of less than 2 mm [Vedam 2004].

Regardless of the specific technique applied for target localization, a further issue in tumour tracking is represented by the need to dynamically adjust the radiation beam to compensate for the detected motion. In this framework, specific solutions for X-ray radiotherapy and particle therapy have been proposed. Different approaches for delivering targeted treatments in conventional X-ray radiotherapy have been implemented. The first technique is based on dynamic multi-leaf collimators, by adapting the setting of the collimator leaves to the tumour's changing position in real time during irradiation [Keall 2001]. A few preliminary studies report the use of treatment couch movements to compensate for intra-fraction breathing motion. Robotic couch corrections provided 1 mm tracking accuracy in phantom studies [Buzurovic 2011], though patient discomfort due to the continuous applied motion was also reported [Sweeney 2009]. Robotic manipulators for the adjustment of the X-ray source position have also been investigated [Schweikard 2000]. Several therapy machines based on robotic motion compensation for respiratory tumour tracking in X-ray radiotherapy are commercially available, such as the CyberKnife system (Accuray, Sunnyvale, California, USA) and the Vero system (Vero GmbH, Germany).

The CyberKnife Robotic Radiosurgery system, shown in Figure 1.23, consists of a lightweight compact 6 MV linear accelerator mounted on a robotic arm, which can be steered with high precision to compensate for the detected intra-fraction tumour movements. The robotic linac is able to deliver many independently targeted and non-coplanar treatment beams aligned with the moving target. The integrated Synchrony Respiratory Tracking System allows to estimate internal tumour motion from an external surrogate breathing signal, derived from a configuration of three infrared light-emitters attached to a vest worn by the patient (Figure 1.23). The 3D displacement of the surface markers acquired by means of optical tracking stereo-cameras at a frequency of approximately 30 Hz is used as surrogate signal. The external-internal correlation model is based on

polynomial functions, different for inhalation and exhalation, thus allowing to capture hysteresis or phase-offset between tumour and surface motion [Seppenwoolde 2007]. During a training period prior to treatment, the correlation model is initialized by simultaneously acquiring the external surrogate signal and the internal motion data, by means of stereoscopic X-ray images of the tumour or of implanted marker. Internal data are periodically acquired throughout the whole treatment duration typically every 30–60 seconds [Kilby 2010], to continuously verify and update the motion model, thus accounting for inter- and intra-fraction changes in the external-internal correlation. During treatment the system can revert to a simple linear model if the value of the respiratory surrogate exceeds the range of values observed during training, thus preventing large extrapolation errors. An adaptive predictor is used to compensate for system inertia and latency using the history of the target movement. Current tolerance levels stated in the vendor's equipment specifications include a robotic manipulator precision of 0.12 mm and an overall targeting accuracy of less than 0.95 mm for static target and 1.5 mm for target undergoing respiratory motion [Kilby 2010].

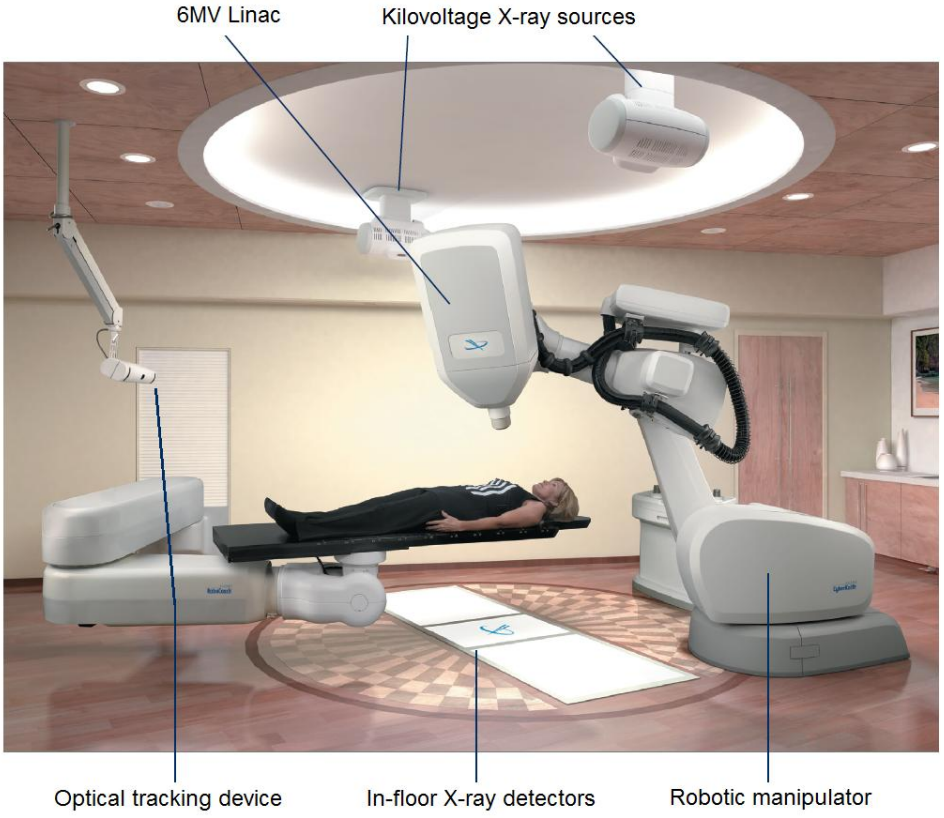


Figure 1.23. The CyberKnife system, composed by a robotic arm carrying the linear accelerator, two ceiling-mounted X-ray sources, two flat-panel detectors on the room's floor and an optical tracking device. The patient wears the Synchrony vest, equipped with three infrared markers.

The Vero system incorporates a novel design in which the X-ray source can pan and tilt on a gimbaled mechanism to provide about 4 cm motion correction around the nominal isocenter [Kamino 2006]. Similarly to the CyberKnife system, motion correction is accomplished by combining infrared tracking of surface surrogates and KV imaging through external-internal correlation models. The system lag was quantified within 50 ms and overall tracking errors of 0.82 mm at 90% confidence level were reported in phantom studies, when latency was compensated by forward motion prediction [Depuydt 2011].

Particle therapy with active beam scanning offers excellent conditions for a real-time motion compensation by tumour targeting techniques [Schardt 2010]. Individual pencil beams have to be adapted not only laterally but also in depth according to the detected tumour motion and to the changes in the radiological path-length. On the plane perpendicular to the beam direction, the target is tracked by directly correcting the lateral beam position through the steering magnets. Movements along the longitudinal beam direction are compensated by using dedicated range shifters [Saito 2009] or by switching the beam energy at frequencies compatible with breathing [Zenklusen 2010]. Real-time tumour targeting for beam scanning systems has been implemented in prototypes and promising results from phantom studies are available. One example is the organ motion compensator integrated into the carbon-ion therapy system at the GSI [Bert 2007]. The system provides real-time motion data feedback to the scanning magnets for lateral motion compensation and to a motor-driven wedge degrader to adjust beam energy for variation of the depth range. System's speed and accuracy in lateral beam corrections were found less than 1 ms and 0.16 mm, respectively [Saito 2009]. A water-equivalence shift of 5 mm in the longitudinal beam direction was obtained in 16±2 ms of delay, with tracking errors lower than 1.08 mm [Saito 2009]. Dosimetric accuracy assessed with ionization chambers deviates by 1±2% from comparable measurements without target motion [Bert 2010]. Another example of a system designed for tumour targeting in active scanning proton therapy is the Gantry 2 at the PSI, featuring planar scanning and energy switching optimized for the efficient implementation of intra-fraction motion compensation strategies [Zenklusen 2010].

Few recent studies report the clinical outcomes of real-time tumour targeting for X-ray radiotherapy treatments delivered with the CyberKnife system [Nuyttens 2012]. The1-year and 2-year overall survival rates of the reported clinical trials proved to increase with respect to SEER (Surveillance, Epidemiology, and End Results) statistics by 41.9% and 25.7% for lung cancer, 30.8% and 19.9% for liver cancer, 22.6% and 10.3% for pancreas cancer, respectively [Riboldi 2012]. The overall survival at 4.5 years is more than 80% in peripheral NSCLC treated with real-time tumour tracking [Brown 2009] and is around 40% in central lung metastases [Nuyttens 2012].

The preliminary clinical outcomes in X-ray photon therapy and the reported feasibility of following moving targets by means of scanned particle beams are boosting the implementation of new designs for particle therapy machines equipped for real-time tumour targeting, with promising perspectives for clinical application in the near future. Besides the already available precise control of the particle beam, the accuracy and robustness of the present tumour localization methods during irradiation need to be further improved to allow for application with active scanning systems [Rietzel 2010, Riboldi 2012]. In X-ray radiation therapy, external-internal correlation modelling, supported by in-room image guidance and optical localization of marker-based surface surrogates, has already been experimentally validated and implemented in new generation treatment machines designed for clinical practice. Research is still needed to investigate the benefit of increasing the complexity of motion correlation models and the redundancy of surface data, for example by using surface imaging techniques to monitor the whole patient's surface, as a way to improve the accuracy of tumour localization [Riboldi 2012].

1.6 Project rationale

The aims of the PhD project concern the development and experimental evaluation of non-invasive methods for real-time tumour tracking, allowing the dynamic localization of target position during radiotherapy treatments for the continuous monitoring and compensation of intra-fraction organ motion due to breathing. The developed strategy exploits the integration of different imaging modalities, such as time-resolved planning imaging, in-room X-ray verification systems and 3D optical surface tracking devices. The proposed tumour tracking approach is based on external respiratory surrogates, inferring tumour motion from the displacement of the patient's thoraco-abdominal surface, acquired by means of non-invasive optical surface imaging systems. The technique relies on patient-specific breathing motion models, derived directly from 4D CT planning images and updated at each treatment fraction according to the anatomical variations obtained from in-room imaging and optical surface devices. The developed method is mainly designed for a possible application with real-time targeting techniques in active beam scanning systems, whose clinical implementation is currently limited by the inadequate accuracy and robustness of the actual target localization techniques [Rietzel 2010, Riboldi 2012].

The developed tumour tracking method features several potential advantages compared to the current techniques clinically applied for dynamic tumour localization during radiotherapy treatments. Since the surface displacement is acquired by means of non-invasive optical tracking systems, the proposed approach allows the continuous monitoring of intra-fraction tumour motion

during the whole treatment duration. This differs from direct target tracking techniques based on radiographic imaging, which can be applied only for a limited temporal window due to the invasiveness of ionizing radiations. In the proposed method, the external surrogate signal is derived from the motion of the entire thoraco-abdominal surface of the patient, dynamically acquired by using 3D surface imaging systems. This approach allows to capture the complex and detailed information of breathing motion coming from an extended surface portion including both thorax and abdomen. The surface-based approaches currently employed in the clinical practice for indirect tumour tracking are instead based on the acquisition of a discrete number of fiducial markers placed on the patient's skin, thus providing an under-sampled description of the surface motion. Finally, the developed technique relies on patient-specific correlation models that are directly derived from planning 4D CT images and updated at each treatment fraction according to patient's daily data. The actual tumour tracking methods do not take into account motion information derived from time-resolved imaging acquired during planning, that can instead provide a valuable description of the target motion pattern during the breathing cycle.

The methodological improvements introduced in the developed tumour tracking approach, associated to patient-specific breathing motion models and to more detailed surface information, are expected to increase the precision of target localization during radiotherapy treatments. The enhanced performances are put forward to guarantee the adequate tumour tracking accuracy and robustness required for the application of targeting techniques in active particle beam scanning. The expected results consist in the clinical extension to extra-cranial moving tumours of active scanning particle therapy methods, presently applied only for the treatment of lesions affected by negligible intra-fraction organ motion [Bert 2011]. This will allow to potentially exploit for a wider range of cancer types the excellent geometrical selectivity and dose conformity of scanned particle beams to improve local tumour control and to minimize complications of healthy surrounding tissues.

2. METHODS

2.1 Overview of tumour tracking technique

The thesis work consisted in developing and testing a surface-based tumour tracking method for the dynamic localization of extra-cranial targets during radiotherapy treatments, providing the real-time continuous monitoring of breathing-induced intra-fraction organ motion. The proposed approach is based on external surface surrogates estimated from non-invasive optical devices, and on patient-specific adaptive motion models derived from time-resolved planning images and in-room X-ray imaging systems. A schematic representation of the developed tumour tracking method is presented in Figure 2.1.

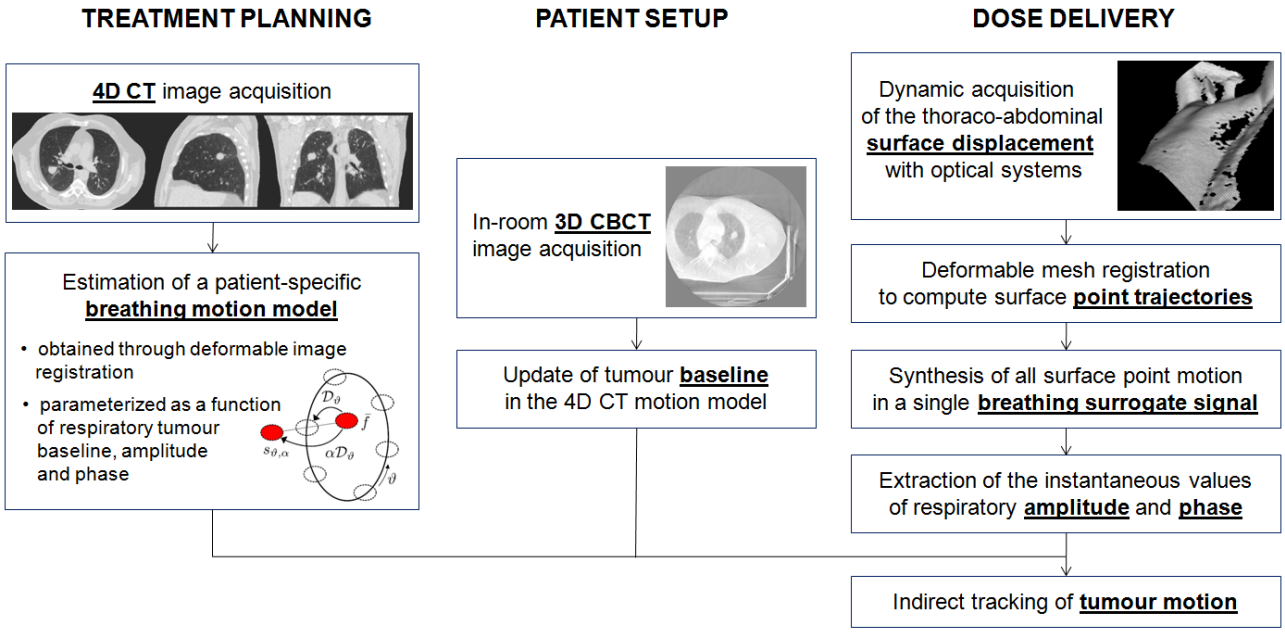


Figure 2.1. Flowchart of the proposed tumour tracking method, based on the integration of different imaging modalities, such as time-resolved computed tomography acquired during planning, in-room CBCT imaging for patient setup verification and 3D optical tracking devices for dynamic surface acquisition during dose delivery.

The 4D CT images acquired during treatment planning are used to estimate a patient-specific respiratory motion model, describing the tumour movement during all phases of the breathing cycle. The model is obtained through deformable image registration, by computing the deformation vector fields between each 4D CT phase image and a reference volume. The mid-position CT image, consisting in a time-weighted average of all 4D CT phases, was chosen as reference. The 4D CT motion model is parameterized as a function of the respiratory tumour baseline, amplitude and phase. The model parameters are adapted at each treatment session according to the inter- and intra-fraction variations in the patient's data obtained by means of in-room X-ray imaging and optical surface devices. The daily information on tumour localization derived from the 3D CBCT image acquired during patient setup is used to update the tumour baseline in the 4D CT respiratory motion model, to take into account possible baseline drifts between planning and treatment time. The breathing amplitude and phase parameters are instead retrieved from the motion of the patient's thoraco-abdominal surface.

During dose delivery, the displacement of the patient's external surface is continuously acquired by means of non-invasive optical tracking systems, providing a 3D topological model of the entire thoraco-abdominal surface. The extraction of a parameterized surrogate breathing signal from the external surface motion required different implementation steps (Figure 2.1). Deformable mesh registration algorithms are applied to the markerless optical surfaces to derive the spatial correspondence between consecutive meshes, thus obtaining the breathing trajectories of all surface points in the three dimensions. The motion information coming from all surface points is summarized in a single surrogate respiratory signal, taking into account both the thoracic and the abdominal breathing patterns. The instantaneous values of the respiratory amplitude and phase parameters associated to each surface frame are extracted from the surrogate signal and adapted according to the difference in surface breathing motion between planning and treatment time. The updated respiratory baseline, amplitude and phase parameters are given in input to the 4D CT motion model to estimate indirectly the three-dimensional tumour trajectory during the whole treatment course.

Figure 2.2 shows the methodological approach developed for the experimental evaluation of the proposed tumour tracking technique. The strategy relies on the synchronized acquisition of internal and external breathing motion information from lung cancer patients. The displacement of the patient's thoraco-abdominal surface is captured by non-invasive optical tracking systems synchronously with a rotational cone-beam CT scan, providing a sequence of 2D projections of the internal anatomy. According to the described tracking approach, the lesion trajectory is indirectly estimated from the external surface motion using the breathing model derived from the 4D CT

planning images of the patient. The tracking accuracy of the proposed method is evaluated by comparing the estimated tumour motion with the real lesion trajectory directly identified on CBCT projections.

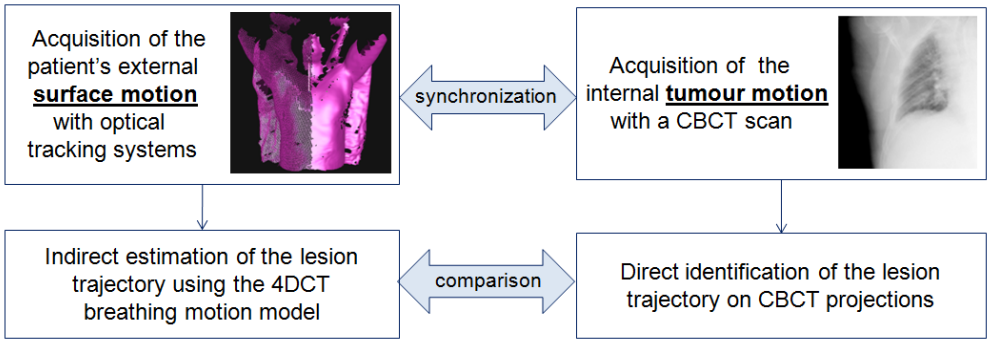


Figure 2.2. Methodological approach for the experimental testing of the developed tumour tracking technique, based on the comparison between the lesion trajectory estimated from the patient's external surface motion and the real lesion trajectory identified on the synchronously acquired CBCT projections.

2.2 Patient-specific breathing motion model

As depicted in Figure 2.1, the developed tumour tracking approach exploits a breathing motion model to predict the internal tumour position from the displacement of the external thoraco-abdominal surface. We propose to make use of the patient-specific motion information contained in the time-resolved imaging data acquired during treatment planning to derive a respiratory model that describes the tumour motion during the breathing cycle. As described in Section 1.2.1.2, the planning of radiotherapy treatments for lesions affected by intra-fraction respiratory motion usually involves the acquisition of a 4D computed tomography of the patient's thoraco-abdominal district, providing high-resolution volumetric images of the internal anatomy for a finite number of breathing phases (typically ten). 4D CT imaging data include valuable anatomical information about the respiration-induced motion of all thoracic and abdominal structures. We propose to exploit 4D CT planning data to estimate a patient-specific tumour motion model to be applied for real-time target tracking and motion compensation during radiotherapy treatments.

2.2.1 Model estimation from 4D CT data

The proposed breathing motion model is estimated from planning 4D CT images through registration-based methods, described in Section 1.3.2. In the selected approach, the tumour motion during the breathing cycle is described by the deformation vector fields computed through

deformable registration applied to the 4D CT image dataset. Due to the elastic nature of lung tissues, deformable image registration is needed to register a pair of pulmonary CT images. B-spline deformable registration algorithms were used for the present application, exploiting the efficient native implementation provided by the open-source software package Plastimatch (www.plastimatch.org) [Sharp 2009]. As introduced in Section 1.2.1.1, B-spline deformable registration uses B-spline basis functions to parameterized the deformation vector field that maps the voxels in the target moving image to those in the reference fixed image. The displacement vector of each individual voxel is obtained by interpolating the coefficients of piecewise cubic B-splines, defined on a set of uniformly-spaced control points. The registration process is an optimization problem where B-spline coefficients are refined iteratively until the warped moving image closely matches the reference image. The optimization process requires the computation of the gradient of a cost function, that quantifies the similarity between the fixed and the moving images, with respect to the coefficient values at each individual control point. B-spline registration algorithms are popular due to their flexibility and robustness both for uni-modal and multi-modal images, but they can result computationally intensive [Rohde 2003]. Plastimatch uses a grid alignment scheme that can significantly accelerate the time-consuming B-Spline interpolation and gradient computation stages, thus speeding up the registration process [Shackleford 2010, Sharp 2010]. As shown in Figure 2.3, the B-spline control point grid is aligned with the voxel grid, partitioning the image volume into many equally-sized tiles. In the 3D case, the vector field at a specific voxel within a tile is influenced by the 64 control points in the tile's immediate vicinity and by the value of the B-spline basis function product, which is dependant only on the voxel's local coordinates (offset) within the tile (Figure 2.3). This property allows to pre-compute all relevant B-spline basis function products once, instead of recomputing them for each individual tile.

The use of the Plastimatch implementation of the B-spline registration algorithm for the estimation of the 4D CT breathing motion model required the selection of several parameters and options, specified in the registration command file. The mean square error was chosen as similarity metric to define the cost function of the optimization process, since the registration of uni-modal CT images is involved. The cost function is optimized using the L-BFGS-B algorithm, an efficient quasi-Newton limited-memory method for solving large non-linear optimization problems, either constrained or unconstrained [Zhu 1997]. A regularization term, based on the second derivative of the vector field, is added in the cost function to increase the smoothness of the registration. The λ parameter that controls the regularization term was set to 0.005, whereas the maximum number of iterations in the optimization process was set to 30. In order to increase the accuracy of registration and to reduce the risk of convergence to local minima, a multi-stage down-sampling approach was

adopted, by progressively refining the B-spline registration parameters in four consecutive stages. The resolution used to subsample both input volumes before running the registration was decreased from 4 voxels along each image dimension to the original voxel size, whereas the spacing between the control points of the B-spline grid was reduced from 80 to 5 voxels. The deformation vector field approximated at a coarser stage with down-sampled images is used as input for finer stages, thus improving the convergence probability of the registration process towards the correct solution.

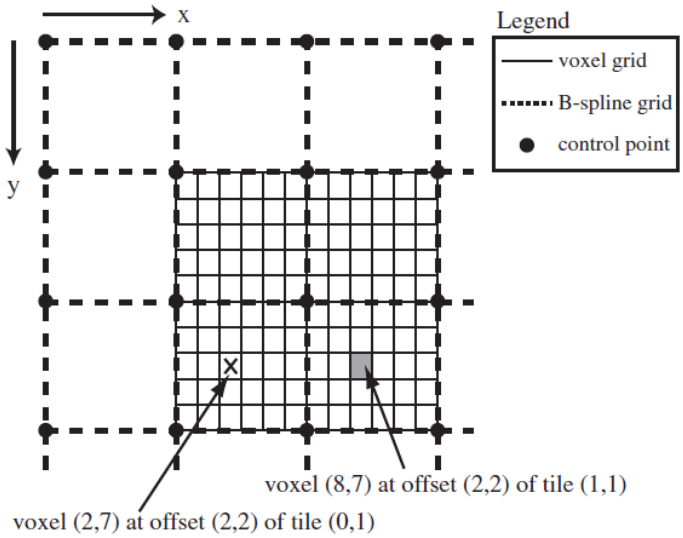


Figure 2.3. Example of a B-spline control point grid aligned with the image voxel grid. Since both the marked voxel and the gray voxels are located at the same relative offset within their respective tiles, the B-spline basis function products evaluated at these two voxels yield identical values [Shackleford 2010].

The patient-specific breathing motion model is estimated from the 4D CT planning dataset by non-rigidly registering each 4D CT phase image to a reference volume. The presence of artefacts in the reference image may have larger impact on the accuracy and robustness of the computed motion model. To reduce the bias caused by the choice of the reference, several approaches have been proposed, including the use of multiple references [Boldea 2008] or the combination of all 4D CT phase images to generate the reference volume. We chose this last solution, by considering as reference the mid-position (MidP) CT image, which represents a time-weighted average of all 4D CT phase volumes [Wolthaus 2008]. The overall procedure for the estimation of the 3D MidP CT image from the 4D CT scan is depicted in Figure 2.4. For each frame of the 4D CT dataset, we compute the deformation vector field with respect to the end-exhale 4D CT frame, which was chosen due to the higher reproducibility of tumour position compared to the other frames. The mean motion of all image voxels during the breathing cycle is calculated by averaging the estimated DVFs. A set of mean-corrected deformation vector fields is derived by subtracting the mean motion

from each DVF. The mean-corrected DVFs are then applied to the original 4D CT scan, transforming each frame into the time-weighted mean position. The resulting set of motion-compensated 4D CT images is finally averaged to obtain the mid-position CT volume, where all the internal structures, including the tumour, assume the exact time-weighted mean position within the 4D CT respiratory cycle. The median average is chosen instead of the standard arithmetic mean for the reconstruction of the MidP CT image since it was proved to be more robust against mis-registration artefacts [Wolthaus 2008]. The MidP scan was found to contain about one-third of the noise and scanning artefacts of the individual 4D CT frames. The tumour shape compared to a breath-hold CT scan was also better represented by the MidP reconstruction than by any other 4D CT image [Wolthaus 2008].

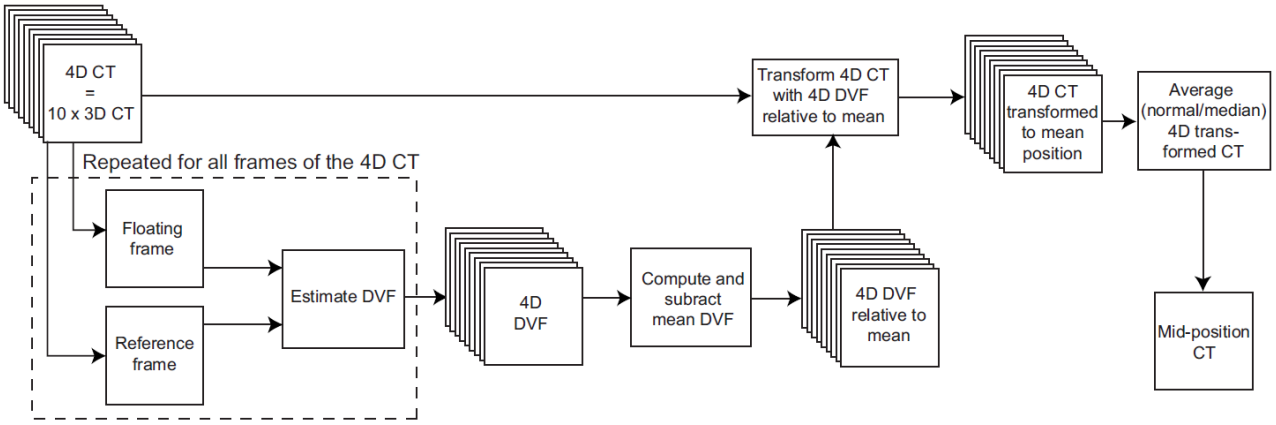


Figure 2.4. Overview of the procedure for the reconstruction of the mid-position CT image. The physical motion in the 4D CT scan is estimated from each frame and subsequently compensated to the time-weighted mean position. The 3D MidP CT scan is obtained by averaging the motion-compensated 4D CT images [Wolthaus 2008].

To improve the local registration accuracy and to reduce the computational cost, the 4D CT breathing motion model is estimated only for the tumour volume of interest (VOI). VOI size are defined by adding a uniform margin of 2 cm in each spatial direction to the tumour extent derived from the clinical GTV contours. Tumour VOI is extracted from the reconstructed mid-position CT and from all 4D CT phase images. B-spline deformable registration is then applied between the MidP and each 4D CT frame VOI, using the described parameters. The resulting deformation vector fields defined the patient-specific breathing model, describing tumour motion in each phase of the respiratory cycle with respect to the time-weighted mean position. The parameterization of the breathing motion plays an important role because it is directly related to the number of degrees of freedom considered for the characterization of the model. In the proposed modelling approach, the respiratory tumour motion is decomposed into three main parameters, represented by baseline,

amplitude and phase (Figure 2.5). The baseline consists in the time-weighted mean position assumed by the tumour during the whole breathing cycle. The angular position within the respiratory cycle is defined by the phase parameter, while the amplitude represents the distance between the tumour position at a specific phase and the breathing baseline. As shown in Figure 2.5, tumour location at a defined breathing phase is obtained by interpolating target positions assumed in the 4D CT images and multiplying the corresponding amplitude by a scaling factor, which allows to model respiratory states that are outside the tumour trajectory followed during the 4D CT scan. The breathing motion model can be represented by the following equation:

$$s_{\vartheta,\alpha} = \bar{s} + \alpha D_{\vartheta} \ , \tag{2.1}$$

where $s_{\vartheta,\alpha}$ is the target position at specific values of the respiratory phase ϑ and amplitude α parameters, $\bar{s}$ is the tumour baseline and D_{ϑ} is the distance between $\bar{s}$ and the position assumed by the tumour at that specific phase during the 4D CT scan.

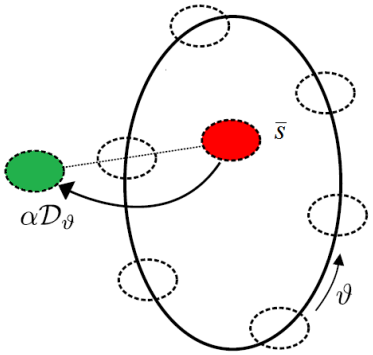


Figure 2.5. Schematic representation of the main breathing parameters that characterize the proposed 4D CT tumour motion model. The time-weighted mean position assumed by the tumour during the respiratory cycle is represented by the red circle ($\bar{s}$). Tumour location at a specific breathing phase ϑ (green circle) is obtained by interpolating target positions estimated in the 4D CT frames (white circles) and multiplying by the scaling amplitude parameter α the distance D_{ϑ} between the tumour position at that phase and the breathing baseline.

2.2.2 Daily adaptation of baseline parameter

Since several weeks may elapse between treatment planning and the end of treatment delivery, breathing variability in short as well as longer terms need to be take into account in motion modelling. Both intra- and inter-fraction variations in the patient's internal anatomy and in the respiratory pattern may occur. The adaptation of the breathing parameters (baseline, amplitude and phase) of the motion model estimated from 4D CT planning images according to the patient's daily situation at each treatment session is required to guarantee the adequate accuracy and robustness of the proposed tumour tracking method. Figure 2.6 shows a schematic representation of a breathing

trace, influenced by different types of motion variability with distinct time-spans, ranging from seconds to weeks. Intra-cycle variability, due for example to changes in the respiratory frequency or in the duration of inhale or exhale phases during a single breathing cycle, is mainly accounted by variations of the phase parameter. The respiratory level can also change irregularly from one cycle to another (inter-cycle variability), causing differences in the end-inhale or end-exhale peak positions and in the shape of the trajectory followed during inspiration or expiration. This type of variability is accounted in the breathing motion model by variations of the amplitude parameters. Systematic drifts or progressive time-trends can finally occur in the respiratory mean position, which are taken into account through variations of the baseline parameter. Systematic shifts are mainly caused by patient positioning errors and by changes in muscle tone or stomach filling, while time-trends are instead attributed to the gravity action on compliant lung tissues or to patient relaxation throughout the treatment [Shirato 2004].

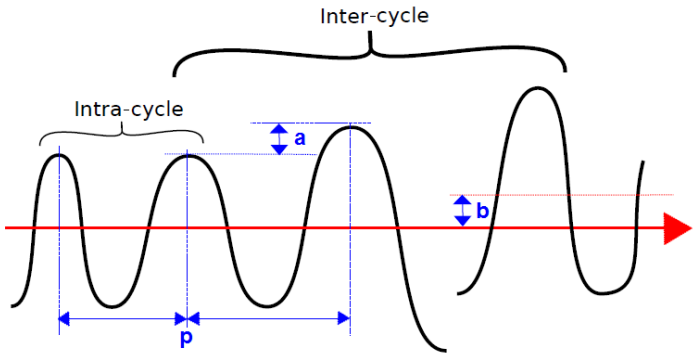


Figure 2.6. Illustration of possible variations of the breathing motion parameters (p = phase, a = amplitude, b = baseline) that can occur within the same respiratory cycle (intra-cycle variability) or among different cycles (inter-cycle variability).

In the proposed tumour tracking application, baseline time-trends and phase or amplitude variations with respect to the 4D CT breathing motion model are retrieved from the external surface motion acquired during treatment delivery by means of optical systems, as will be described in detail in Section 2.3.3. Systematic baseline shifts are instead estimated from in-room volumetric X-ray imaging, acquired at the beginning of each treatment session for patient setup verification. Substantial drifts in the mean target position have been reported by several studies in all three dimensions, potentially involving large dosimetric impact if not properly accounted for [Seppenwoolde 2002, Sonke 2008]. An example of tumour baseline shifts between planning and treatment phases is depicted in Figure 2.7a.

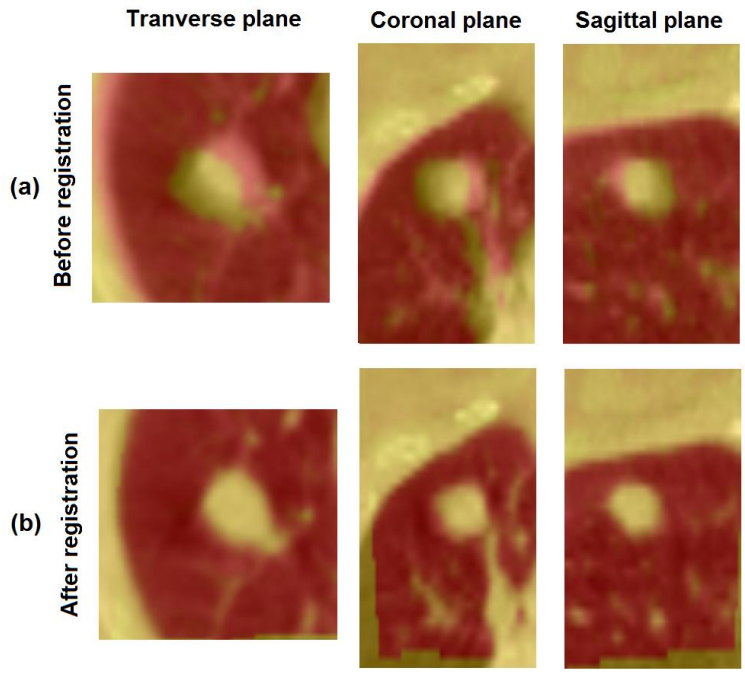


Figure 2.7. Overlap of the tumour VOIs extracted from CBCT volume (in purple) and from MidP CT image (in green) before (Panel a) and after (Panel b) applying B-spline deformable registration. The tumour baseline shifts between planning and treatment phases are properly recovered through the developed registration-based approach.

The adaptation of the tumour baseline in the 4D CT breathing motion model required the daily assessment of the patient's internal anatomy. Different types of in-room X-ray imaging techniques can be applied to retrieve the anatomical information on the daily configuration of the thoraco-abdominal structures, including the lesion. For example, X-ray stereo-projections, clinically available in most radiotherapy and particle therapy centres, provide the required 3D information on tumour and OAR localization during treatment. In the present study, we developed a specific baseline adaptation strategy based on the CBCT volume reconstructed from the 2D projections acquired before each treatment fraction for patient setut. The 3D CBCT image can be considered analogous to the mid-position CT image, since it represents a time-weighted average of the breathing cycles performed during the cone-beam CT scan. At each treatment session, the shifts of the tumour baseline with respect to the planning 4D CT scan are computed as the deformation vector field between the reconstructed CBCT volume and the mid-position CT image. B-spline registration algorithms are applied to the tumour VOIs extracted from the two images. A multi-stage registration process is used, similarly to the previously described approach. The only difference concerns the similarity measure that defines the cost function to be optimized. Since in the actual case multi-modal images with different intensity ranges are involved, we select as metric the normalized mutual information, using the Plastimatch implementation of the Mattes mutual information [Mattes 2001]. Figure 2.7b shows the resulting overlap between the CBCT volume and

the deformed mid-position CT after applying B-spline registration. The deformed MidP represents the updated tumour baseline that is given in input to the 4D CT breathing motion model, thus correcting for possible inter-fractional variations in the patient's anatomo-pathological configuration with respect to the planning phase.

2.3 External surface motion tracking

As shown in Figure 2.1, the breathing amplitude and phase parameters required in input to the 4D CT motion model for the estimation of the internal tumour position are retrieved during treatment from the patient's external surface, dynamically acquired by means of optical surface imaging. External motion tracking is a key aspect in radiotherapy treatments, especially when applied to extra-cranial sites where breathing motion is relevant [Keall 2006]. Real-time optical tracking systems offer a non-invasive way to monitor the intra-fraction motion of the patient's surface, allowing to capture the complex and detailed breathing information coming from an extended portion of the thoraco-abdominal surface. Laser-based or video-based surface scanning devices, described in Section 1.4.2.2, provide the reconstruction of a 3D surface model as a function of time, without the use of passive markers. However, markerless optical acquisitions suffer from the lack of spatial correspondence between points of consecutive surfaces, which hinders the direct measurement of multi-dimensional breathing motion at specific surface landmarks. This is due to the fact that surface detection is performed by projecting laser beams or structured-light patterns over a moving surface from a fixed point of view. This implies changes in the geometrical representation of the surface over time, meaning that vertices and edges of the meshes acquired at different timestamps vary.

Motion tracking from markerless optical surface acquisitions can be derived by applying mesh registration algorithms to establish the spatial correspondence between surface points from a time-series of dynamic meshes. When the thoraco-abdominal surface is considered, registration procedures need to account for a deformation model to adequately describe the motion due to breathing. This means that the transformation that warps each vertex of the source surface onto its corresponding point on the target surface is not rigid. The existing optical tracking systems currently implement rigid surface fitting procedures which provide a global surface motion, without taking deformation effects explicitly into account [Schöffel 2007]. This approach does not allow to capture local surface transformations and complex breathing motion patterns, which generally vary for different regions of the thoraco-abdominal surface. One of the most well-known algorithms for surface registration is the Iterative Closest Point (ICP), which is based on the iterative minimization

of the distance between two sets of 3D points [Besl 1992]. The procedural steps of the ICP algorithm consist in associating points using the nearest-neighbour criteria, estimating the transformation parameters based on a root-mean-square cost function, transforming the points using the estimated parameters and iterating the procedure till convergence is achieved. The ICP technique is limited to global rigid (translation and rotation) or affine (scaling, shearing and reflection) transformations.

Feldmar and Ayache [Feldmar 1996] were among the first to propose a method that allows for deformable surface registration. They initially perform regular ICP to get a rough alignment between the source and the target meshes. They then relax the rigidity constraint by computing affine transformations for spherical subsets of the source surface. Regularization is performed by geometrically smoothing the resulting affine transforms. Allen et al. [Allen 2003] proposed a regularization term based on connectivity rather than spatial proximity. The optimization problem is formulated by assigning an affine transform to each surface vertex, with a global constraint penalizing large differences between transformations of connected points. Amberg et al. [Amberg 2007] showed that this problem can be solved directly using a least-squares approach. The resulting algorithm resembles ICP since it optimizes surface point correspondences sequentially, with the difference that it searches for one affine transformation per vertex of the source mesh instead of a global transformation. Similarly to the presented approaches, we implemented a deformable mesh registration algorithm for the estimation of the spatial correspondence between surface points from subsequent meshes obtained with a markerless optical device [Schaerer 2012]. The algorithm was applied to extract the multi-dimensional breathing motion of the thoraco-abdominal surface, taking into account the deformation induced by respiration.

2.3.1 Deformable surface registration algorithm

The implemented deformable surface registration approach is an extension of the optimal step non-rigid ICP method proposed by Amberg et al. [Amberg 2007], based on a locally affine regularization. The standard ICP algorithm consists in the iterative optimization of a global transformation that minimizes the distances between two sets of 3D points. At each iteration step, each vertex v_i of the source surface S is matched to the closest vertex u_i of the target surface T. The corresponding displacements are used to estimate a global rigid or affine transformation, using a linear least-squares approach. The extension of this process to cope with deformation implies the estimation of one affine transformation X_i for each vertex in S. The general cost function is composed by a distance term, minimizing the distance between the deformed source $X_i v_i$ and the

target surface, and a regularization term, penalizing the difference among the transformations of neighbouring vertices. The distance term E_d is expressed by the following equation:

$$E_d(X) = \sum_{v_i \in V} w_i \, dist^2(T, X_i v_i), \tag{2.2}$$

where V is the set of source vertices and w_i are the weights attached to each vertex. The weights w_i are set to 1 if a minimal distance correspondence is found and 0 otherwise. In order to ensure the convergence of the algorithm towards the correct solution and to avoid false registration, specific constraints are introduced, choosing reasonable thresholds for the proposed application. Unrealistic surface point correspondences are excluded by verifying the following criteria:

- the estimated displacement should not be too tangent to the surface, with a maximum deviation angle from the surface normal of 30°;
- the displacement norm should not exceed a fixed limit, equal to 10 cm;
- the angular difference between the source and the target surface normals should be lower than 30°.

The issue of assigning an affine transformation to each vertex of the source mesh makes the cost function under-constrained and the optimization problem ill-posed. In order to overcome this drawback, an additional regularization term is considered in the overall cost function, by adding constraints to the free variables in the X_i transform set. The constraints represent the transition smoothness across adjacent transforms, which can be modelled as:

$$E_s(X) = \alpha \sum_{\{i,j\} \in \varepsilon} \left\| \left(X_i - X_j \right) \right\|_F^2 , \tag{2.3}$$

where ε is the set of edges of S and α is the stiffness weight that modulates the capability of the surface to deform. This term is used to penalize the difference between the transformations of neighbouring vertices under the Frobenius norm $\| \|_F$, thus regularizing the surface deformation and correctly constraining the equation system. The general cost function is derived by summing up equations (2.2) and (2.3), that can be rearranged to obtain the linear equation system:

$$E(X) = \left\| AX - B \right\|_F^2 , \tag{2.4}$$

which is solved directly using a least-squares approach. In order to compute an efficient solution for the resulting linear system, the matrix A is factorized using the Cholesky decomposition, consisting in decomposing a symmetric positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose.

The implemented non-rigid ICP algorithm is composed by two iterative loops. In the outer loop, the stiffness factor α is gradually decreased with uniform steps, starting from higher values,

which enable to recover an initial rigid global alignment, to lower values, allowing for more localized deformations. For a given value of α, the problem is solved iteratively in the inner loop. At each step, the closest point on the target mesh is computed for each point in the source surface and the optimal set of affine transformations for this correspondence is estimated. Figure 2.8 shows a detail of an intermediate inner step of the registration. The inner iterative process stops when the norm of the difference between two consecutive transforms is lower than a threshold δ. The convergence threshold is adapted for each outer loop step, setting δ equal to a constant fraction of the norm of the transform difference computed in the first inner iteration. The implementation of the deformable registration algorithm was realized through own software developed in C++, whereas Eigen libraries were used to solve the linear system for the estimation of surface point transformations.

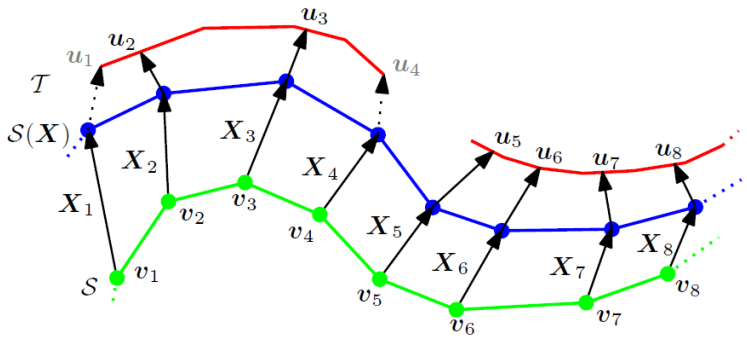


Figure 2.8. Iterative process of the deformable registration algorithm, in which the source surface *S* (green) is deformed by locally affine transformations X_i onto the target surface *T* (red). The closest points u_i for each displaced source vertex $X_i v_i$ is determined and the optimal deformation is found for the stiffness used in this iteration. The process is repeated with lower stiffness values until a stable state is found.

The implemented deformable surface registration algorithm is applied to the time-series of the patient's thoraco-abdominal surfaces optically acquired during radiotherapy treatments for external breathing motion tracking. The first surface captured by the optical imaging system at the beginning of each treatment fraction is taken as the reference source mesh for deformable registration. As illustrated in Figure 2.9, a region of interest is selected in the reference mesh, by considering the intersection between the external surface portions scanned in the planning CT and optically acquired during treatment. The resulting reference ROI includes the patient's thorax and the upper part of the abdomen, excluding the regions not involved in the respiratory motion, such as the arms and the head.

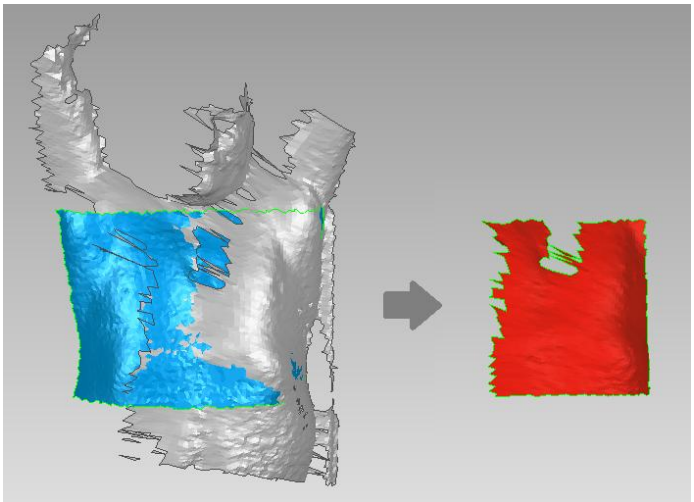


Figure 2.9. Thoraco-abdominal region of interest (red mesh) selected as the intersection between the patient's surface acquired with the optical system during treatment (grey mesh) and the surface extracted from the CT planning image (blue mesh).

Since the accuracy and robustness of the registration algorithm is strongly influenced by the presence of noise and artefacts (such as scanner errors) in the reference source mesh, a pre-processing stage is applied to the selected thoraco-abdominal ROI. Surface pre-processing is performed using the Geomagic Studio software and includes the following steps:

- noise reduction, by moving surface points to statistically correct locations;
- hole filling, by filling the mesh gaps caused by sparse point data;
- spike removal, by detecting and flattening single-point spikes and sharp edges;
- surface smoothing, by minimizing the angle between adjacent triangles and by conforming the triangle size;
- mesh repair, by reversing back-facing triangle normals.

An example of original and pre-processed reference meshes is reported in Figure 2.10. The patient's optical surfaces of the time-series acquired during treatment are registered to the reference source ROI using the implemented deformable registration algorithm. The number of outer iterative loops was set to 10, decreasing the stiffness parameter from a starting values of 150 to an end value of 1. The proposed method allows to retrieve the point-by-point correspondence between consecutive optical surfaces, thus obtaining the 3D trajectory of the generated correspondent surface points. The multi-dimensional breathing motion of the thoraco-abdominal surface is therefore derived from markerless optical imaging, taking into account thorax and abdomen deformations induced by respiration.

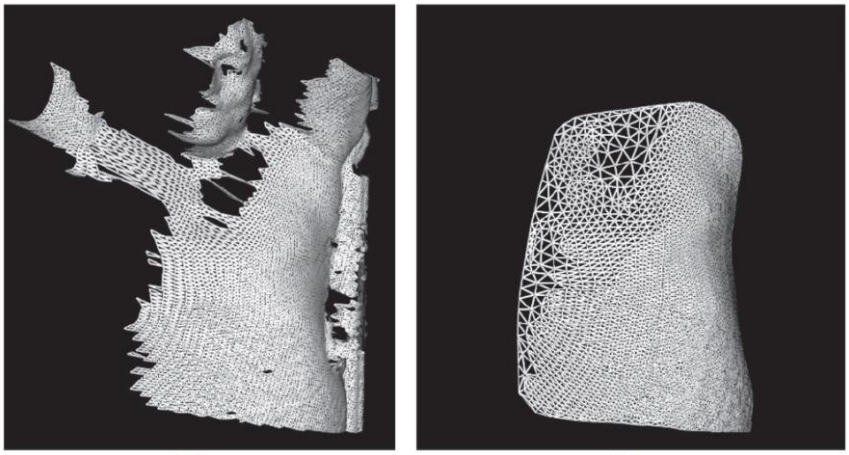


Figure 2.10. Patient's external surface acquired with the VisionRT optical surface imaging system (Panel a) and corresponding thoraco-abdominal ROI after the pre-processing step, which allows to smoothly fill the surface holes (Panel b).

2.3.2 Synthesis of surrogate breathing signal

The motion trajectories of the points of the patient's external surface derived with the implemented deformable registration algorithm are summarized in a single overall signal, in order to retrieve a robust and reliable respiratory surrogate for real-time tumour tracking applications. For each surface point belonging to the selected thoraco-abdominal ROI, a mono-dimensional signal is derived by computing the time-series of the distance from the lowest position assumed by the single points, taking into consideration all the three spatial directions of surface motion. The distance signals obtained from all ROI points are then synthesized in a unique breathing motion surrogate. Different approaches for summarizing the respiratory information coming from all surface points were investigated and compared, including Principal Component Analysis (PCA), k-means clustering and Self-Organizing Map (SOM).

2.3.2.1 Principal component analysis

Principal component analysis is a widespread statistical technique for finding similarity or difference patterns in a dataset consisting of a large number of inter-related variables, thus reducing the dimensionality with minimal loss of information [Jolliffe 2002]. It is based on an orthogonal linear transformation that converts a set of dependent observations into a set of linearly uncorrelated variables, called principal components. This transformation is defined so that the first principal component has the largest possible variation, namely it accounts for the most of the variability in the input data. The original dataset is transformed into a new coordinate system such that the

greatest variance for any projection of the data lies on the first coordinate, the second greatest variance on the second coordinate, and so on. Each principal component is a linear combination of the original variables and all the principal components are orthogonal to each other, so there is no redundant information. PCA results are expressed in terms of component scores and loadings. Scores are the transformed variables corresponding to the original data represented in the principal component space. Loadings are the principal component coefficients, namely the weight by which each original variable should be multiplied to get the component score [Abdi 2010].

PCA is obtained through eigenvalue decomposition of the data covariance matrix, after mean subtraction and normalization of the input data. Given the *m*-by-*n* input matrix *X*, where each of the *n* columns (variables) represents a different repetition of the experiment and each of the *m* rows (observations) gives the results obtained for each repetition, the covariance matrix of *X* is computed as:

$$C = XX^T . \tag{2.5}$$

The reduced-dimensionality representation of the input data *X* in the principal component space is given by:

$$Y = A^T X , \tag{2.6}$$

where columns of the orthogonal matrix *A* are the eigenvectors of the covariance matrix *C*. The eigenvalues of the covariance matrix of *X* represents the variance of the original data explained by each principal component. Given a set of 3D points, as shown in Figure 2.11, the first principal component corresponds to a line that passes through the mean value and minimizes the sum of squares of the distances of the points from the line. The second principal component corresponds to the same concept, after all correlation with the first principal component has been subtracted from the points.

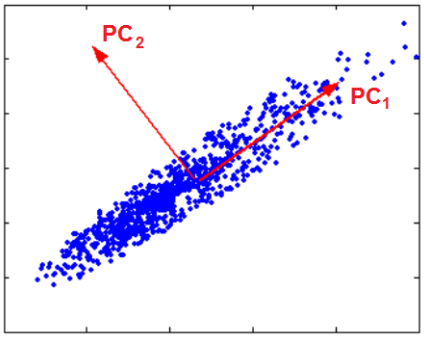


Figure 2.11. Plot of a 3D point cloud and overlay of the respective first two principal components (PC_1 and PC_2). In the new transformed coordinate system, the greatest variance for any projection of the input data lies on the first coordinate (PC_1), while the second greatest variance lies on the second coordinate (PC_2).

2.3.2.2 K-means clustering

K-means clustering is one of the most commonly applied methods for data partitioning, which refers to the process of grouping similar data into the same cluster or splitting dissimilar data into different clusters according to some predefined criteria [MacQueen 1967]. Given a set of m data points $X = \{x_1, x_2, ..., x_m\}$, where each data point is a n-dimensional vector, the k-means clustering algorithm aims to partition the m data points into k clusters $C = \{c_1, c_2, ..., c_k\}$, as illustrated in Figure 2.12, so as to minimize an objective function of dissimilarity [Hartigan 1979]. The objective function is represented by the within-cluster sum of squared errors, using as metric the Euclidean distance between a data point vector x_i in the cluster j and the corresponding cluster centroid v_j. The set of cluster centroids $V = \{v_1, v_2, ..., v_k\}$ is computed as the mean value of the data points belonging to each cluster. The objective function to be minimized is represented by:

$$J(X,V) = \sum_{j=1}^{k}\left(\sum_{i=1}^{m} u_{ij} \cdot \left\|x_i - v_j\right\|^2\right), \tag{2.7}$$

where u_{ij} defines the membership of the data point x_i to the cluster c_j:

$$u_{ij} = \begin{cases} 1, & \text{if } x_i \in c_j \\ 0, & \text{otherwise} \end{cases}. \tag{2.8}$$

The centroid of all the data point vectors in cluster c_j can be computed as:

$$v_j = \frac{1}{\left|c_j\right|}\sum_{i=1}^{m} u_{ij} \cdot x_i\ , \tag{2.9}$$

where $\left|c_j\right| = \sum_{i=1}^{m} u_{ij}$ is the size of the cluster c_j.

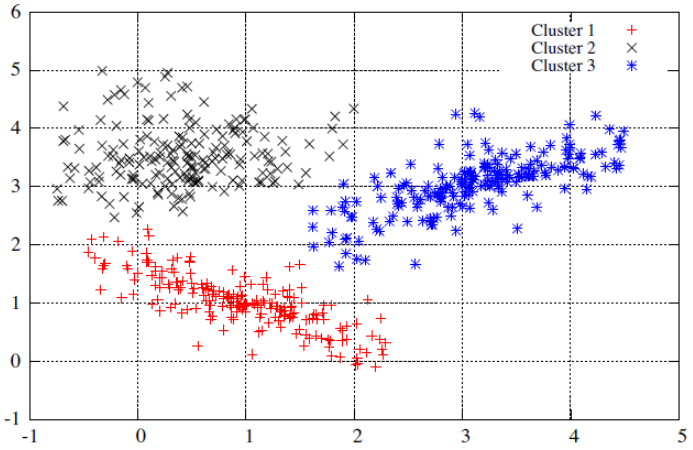


Figure 2.12. Result of partitioning a 3D point dataset into three clusters with the k-means clustering algorithm.

Given an initial set of *k* cluster centroids, obtained by randomly selecting *k* data points from *X*, the k-means algorithm proceeds by alternating between the following two steps:

- assignment step, that assigns each data point to the cluster with the closest centroid, thus determining the membership matrix *U*;
- update step, in which the cluster centroid are updated as the mean of all the data points assigned for that cluster.

These two steps are iteratively repeated until the cluster centroids no longer vary or until there is no significant change in the *J* values of two consecutive iterations. Since the k-means clustering output may be influenced by the initial selection of the cluster centroids, the algorithm is usually run multiple times with different starting conditions, selecting as optimal solution the one associated to the lowest value of the within-cluster sum of point-to-centroid distances. In the present application we chose to replicate the clustering algorithm 10 times, varying the initial set of cluster centroid positions to increase the method robustness.

2.3.2.3 Self-organizing map

Self-organizing maps are a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional discretized representation (map) of the input space of the training samples [Kohonen 1982]. SOM algorithm uses a neighbourhood function to preserve the topological properties of the input space, providing a spatially organized representation of various features in the input data. A self-organizing map consists of components, called nodes or neurons, that are associated to a specific position in the map space and to a weight vector of the same dimension as the input data vectors. The SOM describes a transformation from a higher dimensional input space to a lower dimensional map space. The usual arrangement of map neurons is a two-dimensional regular spacing in a hexagonal or rectangular grid. Similarly to most artificial neural networks, SOMs operate in two modes:

- training, that is the process to build the map using the input examples;
- mapping, which automatically classifies a new input vector, finding the node with the closest weight vector, namely the weight vector with the smallest distance metric to the data space vector.

The goal of SOM training is to associate output nodes with groups in the input dataset, causing different parts of the network to respond similarly to certain input patterns. The weights of the neurons are either initialized to small random values or sampled uniformly from the input data space. Training is based on unsupervised competitive learning, by feeding iteratively the network

with a large number of example vectors, representing as close as possible the kind of vectors expected during mapping [Van Hulle 1997]. For each training example, the Euclidean distance between the input vector and the weight vectors of all neurons is computed. The neuron whose weight vector is most similar to the input vector is called the best matching unit (BMU). The weights *W* of the BMU and of the neighbouring nodes in the map are adjusted towards the input vector *I*, thus conforming the winning neurons to the given training data. At iteration *t*, the new weights of the selected nodes are updated as follows:

$$W(t+1)=W(t)+\Theta(t)\cdot L(t)\cdot[I(t)-W(t)], \qquad (2.10)$$

where $I(t)-W(t)$ is the difference between the node weight and the input vector, $L(t)$ is the current learning rate of the SOM, which decreases exponentially with time, and $\Theta(t)$ is the neighbourhood function. As depicted in Figure 2.13, the alteration of the node weights is inversely proportional to the distance *d* between the neuron and the BMU according to the following neighbourhood function:

$$\theta(t)=e^{-\frac{d^2}{2\sigma^2(t)}} . \qquad (2.11)$$

The radius σ of the BMU's neighbourhood is progressively decreased with time by an exponential decay:

$$\sigma(t)=\sigma_0\cdot e^{-t/\lambda} , \qquad (2.12)$$

where σ_0 is the radius of the map and λ is a time constant, computed as the ratio between the total number of iterations and the map size.

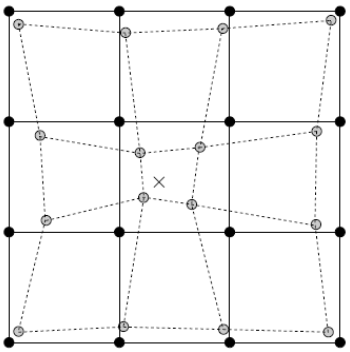


Figure 2.13. Effect of the neighbourhood function in the SOM algorithm. Starting from a perfect node arrangement in a squared map (full lines), the weights of the neurons nearest to the current input (indicated with the cross) receive the largest alteration, resulting in the updated map (dashed lines).

The three presented approaches were tested and compared for the synthesis of a single respiratory surrogate signal from the trajectory of all surface points in the thoraco-abdominal

region. The input matrix was organized by considering each point as a different variable, while the observations were represented by the time-series of the distance of each point from its lowest position during optical surface acquisition. The respiratory surrogate was derived from the algorithm outputs, represented by the coefficients of the principal components for the PCA approach and by the positions of the cluster centroids or map neurons for the k-means or SOM techniques. All the three investigated methods were implemented using specific toolboxes available in Matlab (The Matworks, Natik, MA, USA).

K-means clustering was also applied to investigate the spatial variability of breathing surface motion, since it allows to group surface point trajectories according to their motion pattern. As will be further explained in the Experimental testing chapter, the cluster analysis revealed that the external surface can be mainly divided in two regions with different breathing patterns. As shown in Figure 2.14, the two clusters are mainly localized in the thoracic and abdominal areas, confirming the anatomical partition of the chest wall in two separate compartments (abdomen and rib cage), usually featuring different respiratory kinematics in terms of motion amplitude and phase shift [Aliverti 2001]. According to the obtained results, for the k-means and SOM approaches, the surrogate breathing signal was computed as the linear combination of the motion trajectories of the first two clusters and of the first two neurons, respectively. In case of PCA analysis, the coefficients of the first two principal components were instead combined according to the respective explained variance of the input dataset. Based on the results of the comparative analysis reported in the Experimental testing chapter, k-means clustering was selected as the optimal algorithm for the synthesis of the surrogate breathing signal from the external optical surface acquisitions for tumour tracking applications.

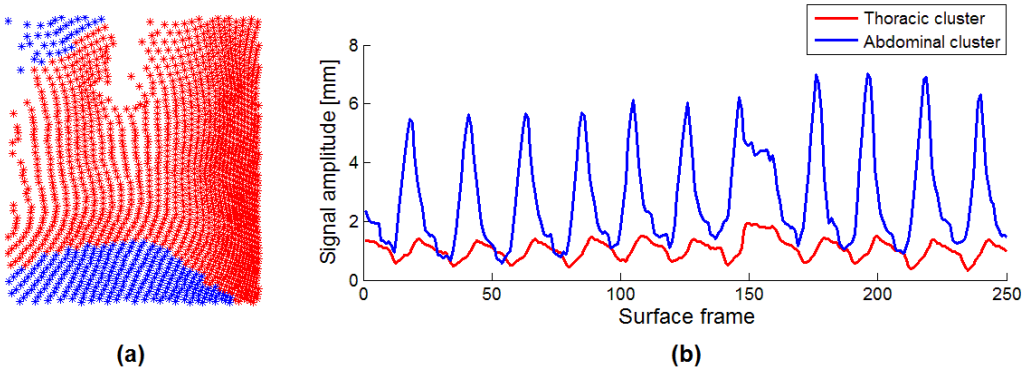


Figure 2.14. Panel a shows the spatial distribution of the first two clusters obtained by applying k-means algorithm to the trajectories of the surface points in the thoraco-abdominal region. Panel b depicts the breathing motion signals associated to each centroid of the two clusters.

2.3.3 Extraction of respiratory parameters

In the proposed tumour tracking application, the respiratory amplitude and phase parameters required as input to the 4D CT motion model for the indirect estimation of tumour position during treatment are retrieved from the surrogate breathing signal derived from the patient's external surface displacement. To take into account possible differences in the respiratory motion patterns between planning and treatment phases, the surface motion acquired with optical tracking systems during each treatment fraction is compared to the displacement of the external surface derived from the 4D CT scan. This process requires the extraction of the patient's thoraco-abdominal surface from each of the 4D CT image of the planning dataset. Threshold-based segmentation algorithms are applied to binarize the volume portion in the 4D CT images corresponding to the patient's body. The open-source Matlab-based "iso2mesh" toolbox [Fang 2009] was used for the extraction of the 3D triangular meshes from the binarized CT volumes. The applied meshing algorithm is based on Delaunay refinement [Shewchuk 2002], a technique for generating unstructured meshes of triangles using the Delaunay triangulation as geometric construction criteria, which is refined by the insertion of additional vertices. The placement of these vertices is chosen to enforce boundary conformity and to improve the mesh quality.

2.3.3.1 4D CT mesh generation

In two dimensions, a triangulation of a set V of vertices is a set T of triangles whose edges do not intersect each other and whose union completely fills the convex hull of V. The Delaunay triangulation D of a vertex set maximizes the minimum internal angle among all possible triangulations. Given two vertices u and v of V, the edge uv is said Delaunay if there exists an empty circle that passes through u and v that does not contain any vertex of V. The triangle composed by Delaunay edges has the property to have an empty circumcircle, namely no vertex of the triangulation is enclosed in the circle that passes through all its three vertices. In three dimensions, the Delaunay condition corresponds to the requirement that the circumscribed spheres of all triangles have empty interiors. For our application, the maximum radius of the Delaunay sphere used for mesh generation was set to 5 mm. Delaunay refinement algorithms are based on the insertion of new vertices in the Delaunay triangulation, until the mesh meets constraints on element quality and size [Chew 1993, Ruppert 1995]. As shown in Figure 2.15, vertices are inserted at the circumcenter of a triangle of poor quality, which is therefore eliminated by the refinement process because its circumcircle is no longer empty. Poor quality is associated to triangles whose circumradius-to-shortest edge ratio, equivalent to the radius of the circumsphere divided by the

length of the shortest edge of the triangle, is greater than a specific bound *B*. Each of the new triangle edges has length at least *B* times that of the shortest edge, thus obtaining a uniform mesh that is size-optimal and topologically well-posed.

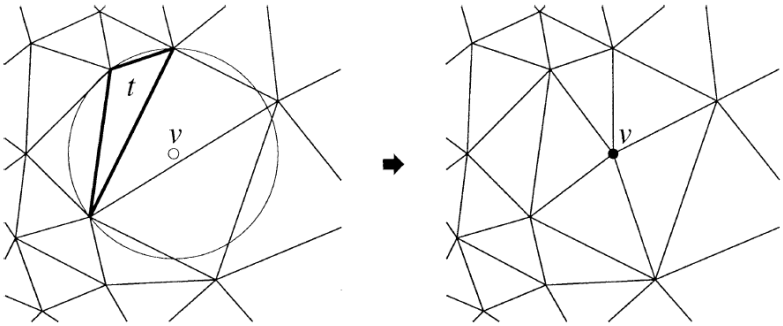


Figure 2.15. Delaunay refinement algorithm provides that any triangle whose circumradius-to-shortest edge ratio is larger than a specific threshold is split by inserting a vertex at its circumcenter [Shewchuk 2002].

The applied mesh generation algorithm provides as output the meshes corresponding to the patient's thoraco-abdominal surface at each breathing phase of the 4D CT scan (Figure 2.9). The mesh are composed by 3D vertices connected to form Delaunay triangles. For each triangle, the surface normal is calculated as the vector cross product of two edges of the triangle. The estimation of the respiratory amplitude and phase parameters from the external surface displacement requires the comparison between the breathing surrogate signal extracted from the optical surfaces acquired during treatment and the surface motion signal derived from the planning 4D CT scan. To allow the comparison, the same procedure is applied for the computation of the respiratory surface signals during planning and treatment phases. The triangulated meshes extracted from each of the 4D CT phase image are registered to the thoraco-abdominal region of interest derived from the first optical surface acquired in the considered treatment session. The breathing trajectories in the three spatial directions for each surface point belonging to the selected ROI are computed using the implemented deformable registration algorithm both for 4D CT and optical surfaces. A mono-dimensional signal is derived for all ROI points, by computing the time-series of the distance from the lowest position assumed by each point during the corresponding 4D CT scan or optical acquisition. K-means clustering is applied to summarize the distance variables derived from all ROI points, thus obtaining a single respiratory surrogate signal for both planning and treatment phases. The cluster assignment obtained with the treatment optical surfaces, namely the index specifying to which cluster each ROI point belongs to, was maintained when applying k-means clustering to the 4D CT surfaces, thus allowing a robust comparison.

2.3.3.2 Phase and amplitude estimation

Different approaches were tested and compared for the estimation of the instantaneous breathing phase from the surface surrogate signals, including linear interpolating phase, Hilbert phase [Gabor 1946] and a modified version of the ellipse-based method proposed by Ruan et al. [Ruan 2008, Ruan 2009]. The linear phase technique relies on the identification of the location of the peaks at end-inspiration and end-expiration in the respiratory trace, assigning percentages to inter-peak positions based on a linear interpolation of the peak-to-peak distance. The linear interpolating phase is given by the following equation:

$$\phi(t) = 2\pi \frac{t - t_n}{t_{n+1} - t_n} + 2\pi\, n \qquad (t_n \le t \le t_{n+1}), \tag{2.13}$$

where t_n represents the timestamp for the n^{th} identified respiratory peak. In addition to its simplicity, an advantage of the linear interpolating phase is that it is locally defined, so perturbations and errors do not propagate over time. The presented approach does not allow however to capture variability in the inhale and exhale durations and hysteresis in the breathing trajectory. The identification of peak values is also very sensitive to noise, especially at extreme tidal positions, such as deep inhale or exhale. Low-pass filtering of the respiratory trace allows noise suppression [Neicu 2003], but introduces uncertainties in the detected peak locations due to the diminished values of local gradients.

Ruan et al. [Ruan 2008, Ruan 2009] proposed to apply elliptic shape tracking in an augmented state space to automatically and reliably detect iso-phase points in the breathing trajectory. A state vector is created from the values of the respiration signal at the current instant and at a previous time, considering a fixed temporal delay of *K* samples. An augmented state space is derived by plotting *N* state vectors. An ellipse model is then fit to the plotted data, as shown in Figure 2.16a. The iso-phase events can be identified as the intersection points between the ellipse model fitted in the augmented state space and an arbitrary Poincaré section, corresponding to a line in two-dimensions. The Poincaré line is constructed starting from the estimated ellipse centre and with a fixed direction relative to the orientation of the ellipse axes, thus guaranteeing that it intersects each breathing period exactly once at the same phase value (Figure 2.16a). The phase estimation for the intermediate positions in the breathing cycle is performed by Ruan et al. by linearly interpolating the detected iso-phase points (Figure 2.16b). Thanks to the elliptic model fitting, this approach has the advantage of being robust to local noise, outliers and partially missing data. We extended the ellipse-based method proposed by Ruan et al. for the instantaneous estimation of the respiratory phase, without relying on linear interpolation. The phase value at each

time instant can be derived by the angle between the orientation of the ellipse's major axis and the line connecting the corresponding state vector in the augmented state space and the centre of the fitted ellipse. This approach allows to capture hysteresis and irregular dynamics in the respiration motion, but proved to be highly sensitive to the choice of the two model parameters, represented by the temporal delay *K* used to construct the state space and the number *N* of previous state vectors used for ellipse fitting.

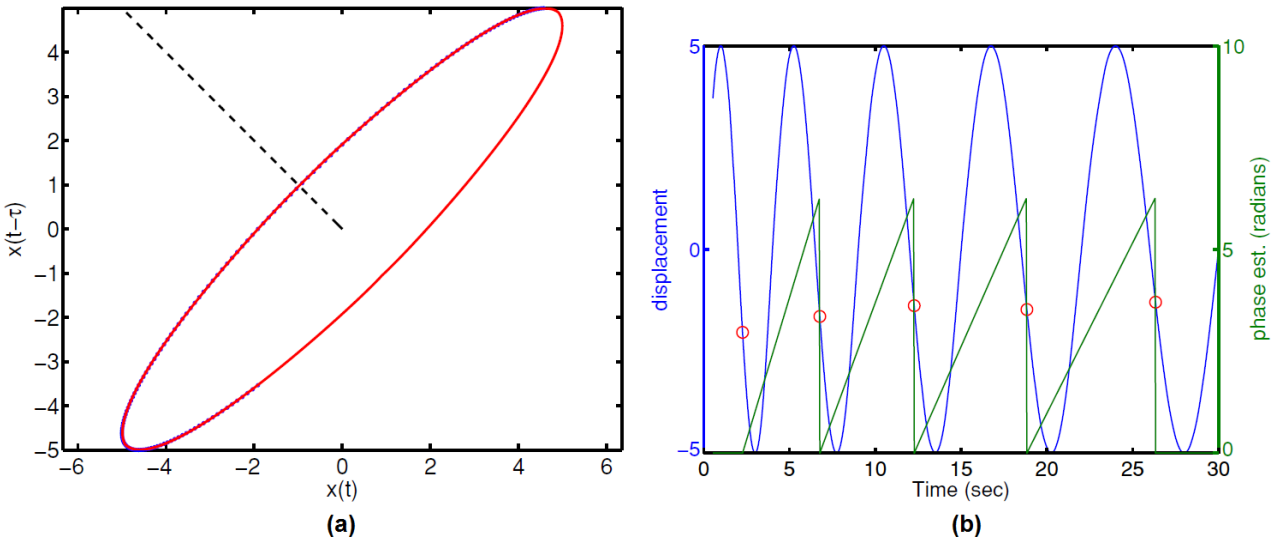


Figure 2.16. Application of the ellipse-based method for phase estimation in a simulated sinusoidal waveform with variation in frequency. Panel a shows the discrete observations (blue dots), the fitted ellipse (red line) in the augmented state space and the Poincaré section (black dashed line) for identifying the iso-phase points. Panel b depicts the simulated trajectory (blue line), the iso-phase markers (red circle) and the estimated linear phase (green line) [Ruan 2009].

The Hilbert transform [Huang 1998] is a spectral analysis method, specifically designed for analyzing data from non-linear and non-stationary processes. It represents the response of a signal to a linear time-invariant filter having impulse response *1/t*. According to the following equation, the Hilbert transform *y(t)* of a an arbitrary time-series *x(t)* is defined as the convolution of *x(t)* with the signal *1/t*:

$$y(t) = H[x(t)] = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{x(u)}{t-u}\, du \ , \qquad (2.14)$$

where *P* indicates the Cauchy principal value of the integral. The Hilbert transform can be used to derive an analytic signal of the real signal *x*(*t*), consisting in the best local fit of an amplitude and phase varying trigonometric function to the time-series *x*(*t*). The analytic signal of *x*(*t*) is represented by a rotating vector with instantaneous phase *φ(t)* and amplitude *A(t)* in the time domain. According to the following polar coordinate expression, the real part of the analytic signal

z(t) is represented by the original signal, while the imaginary part coincides with the Hilbert transform:

$$z(t) = x(t) + iy(t) = A(t) \cdot e^{i\varphi(t)} \ , \tag{2.15}$$

whit $A(t) = \sqrt{x^2(t) + y^2(t)}$ and $\varphi(t) = \arctan\left(\frac{y(t)}{x(t)}\right)$. The phase value at each time instance can be computed as the angle of the Hilbert transformed series, which has the same frequency content as the original data. The Hilbert approach was selected for the estimation of the breathing phase parameter in the present application, since it accounts for hysteresis and irregular variability in respiratory motion, without depending on any specific parameter.

To reduce the acquisition noise, the surrogate breathing signals extracted from the patient's external surface during planning and treatment sessions are smoothed by applying a low-pass moving average filter with a window size of 3 samples. The mean values are subtracted from the corresponding signals to allow the comparison. The presence of time-drifts in the signal baseline can reduce the accuracy and robustness in the estimation of the breathing phase through the Hilbert transform. In order to obtain meaningful instantaneous phase values, the signal must be locally symmetric with respect to the zero mean level [Huang 1998]. Baseline time-trends in the surrogate breathing signal extracted from treatment optical surfaces are computed by smoothing the signal with a moving average filter, whose window size approximately spreads over two respiratory cycles. The number of samples *s* included in the filter window is computed as:

$$s = 2 * \text{round}\,(T * f) + 1, \tag{2.16}$$

where *f* is the frame rate of optical surface acquisition and *T* is the average cycle length, considered equal to 3.5 sec. The no-trend respiratory signal used for the computation of the Hilbert transform is obtained by subtracting the computed baseline from the original signal. Baseline subtraction is not performed for the 4D CT surrogate signal since it represents a single average respiratory cycle, thus not featuring any baseline time-trends. The instantaneous values of the breathing phase associated to each 4D CT mesh and optical surface are estimated from the angle of the Hilbert transform. A representative example of breathing motion signals and corresponding phase values extracted from the patient's surface during planning and treatment phases is depicted in Figure 2.17. Phase values are expressed in radians and ranged between -π and +π. The end-inhale peaks correspond to the zero-crossing values of the Hilbert phase signals, while the end-exhale peaks are determined by phase signal transitions from the maximum to the minimum values.

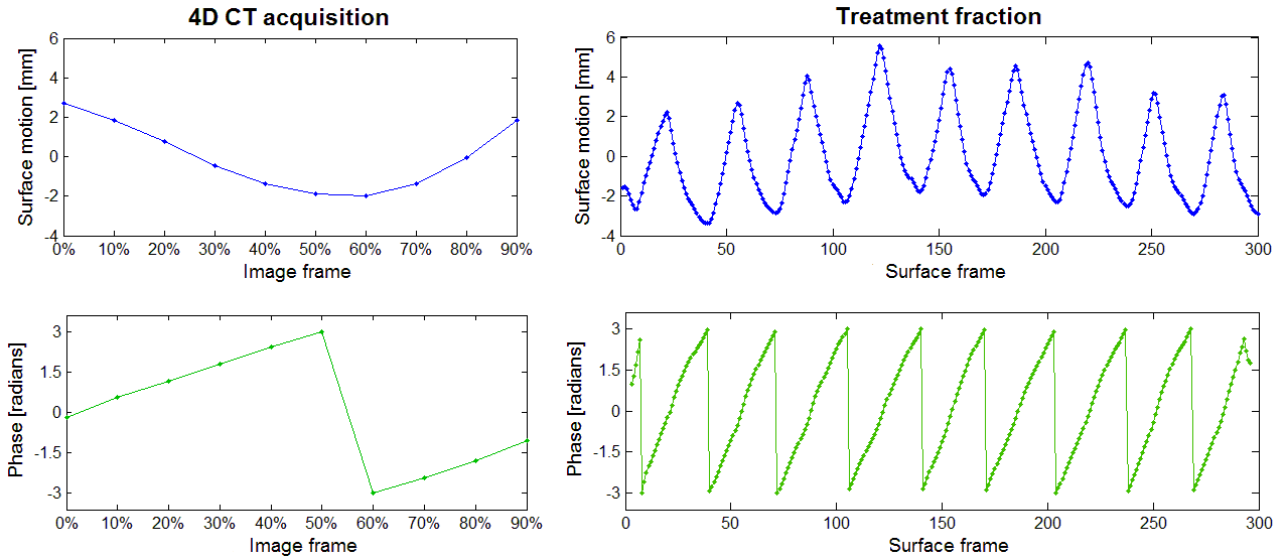


Figure 2.17. Breathing motion signals (blue line) of the patient's thoraco-abdominal surface acquired during planning and treatment phases and corresponding phase values (green line) computed with the Hilbert transform. The planning 4D CT scan is divided in 10 phase volumes, while the optical surface acquisition include 300 surface frames.

The phase and amplitude parameters required by the 4D CT breathing motion model for the estimation of tumour position are obtained by combining the surface motion and phase signals acquired during planning and treatment phases. An example of the phase and amplitude model parameters extracted from the surrogate respiratory signal acquired from the patient's external surface during a treatment fraction is shown in Figure 2.18. For each frame of the optical surface acquisition, the phase parameter specifies the interpolated 4D CT image associated to the same value of the respiratory phase. This parameter is obtained by linearly interpolating the surrogate breathing signal extracted from the 4D CT meshes, thus computing the percentage position between 4D CT images that corresponds to the same respiratory phase of the surface frame (Figure 2.18b). The computed phase parameter ranges between 0 and 1. For instance, in case of a 4D CT scan composed by 10 phase volumes, a phase parameter equal to 0.2 corresponds exactly to the 20% phase CT image. The proposed approach allows to compensate for possible phase shifts between the external surface displacement and the internal tumour motion.

The amplitude of the breathing motion model is instead parameterized according to a scaling factor, obtained by comparing the amplitude of the surface motion measured during planning and treatment phases. For each surface frame of the treatment optical acquisition, we linearly interpolate the amplitude of the 4D CT surface signal at the correspondent breathing phase (Figure 2.18a). The amplitude parameter is computed as the difference between the amplitude of the optical surface signal and the correspondent 4D CT amplitude, normalized by the span (maximum to minimum value) of the 4D CT surface signal. Since the considered breathing signals have zero mean, thus

yielding negative values for the exhale phase, the absolute values of the signal amplitude are considered before the subtraction. A low-pass moving average filter with a window size of 10 samples is applied to obtain a smooth transition of the amplitude values between consecutive time instants. As shown in Figure 2.18c, the resulting amplitude parameter is cantered around 0, with positive or negative values if the amplitude of optical surface motion during treatment is greater or smaller than surface motion amplitude measured in the 4D CT scan. The amplitude parameter accounts both for baseline time-trends and amplitude variations in the breathing motion that can occur between planning and treatment phases. The phase and amplitude parameters are given in input to the 4D CT motion model to estimate the internal tumour position from the external surface displacement, as explained in the following section.

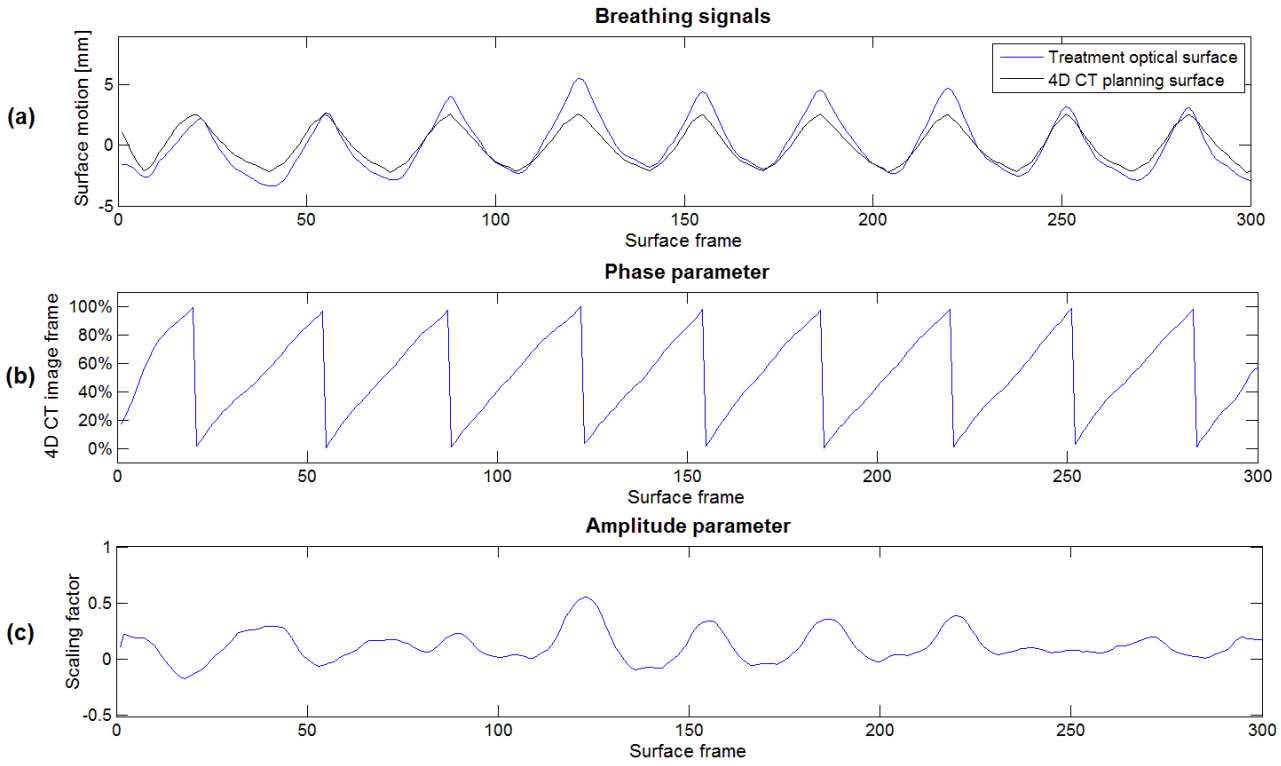


Figure 2.18. Comparison between the surrogate breathing signals extracted from surface motion during treatment (blue line) and derived from the planning 4D CT images at corresponding phases (black line) (Panel a). Panel b and c depicts the phase and amplitude parameters, respectively, representing the input to the 4D CT motion model. These parameters are obtained by integrating the planning and treatment surface motion signals depicted in Figure 2.17.

2.4 Tumour trajectory reconstruction

The proposed tumour tracking method is based on the integration of the external surface displacement optically acquired during treatment with a patient-specific model of tumour breathing motion estimated from planning 4D CT images and updated at each session according to daily in-

room imaging. As explained in Section 2.2, the respiratory model is defined by the mid-position CT image, representing the time-weighted average of tumour position during the 4D CT breathing cycle, and by a set of deformation vector fields computed between the MidP and each 4D CT phase image. The model adaptation to the patient's daily situation is expressed in terms of an additional vector field, describing the baseline systematic changes between planning and treatment phases. From the input vector fields of the breathing motion model we derive the fraction-specific tumour positions at each respiratory phase of the breathing cycle. As performed by Orban de Xivry et al. [Orban de Xivry 2007] the deformation maps obtained through image registration are used to propagate the tumour contours manually delineated in a single reference volume to all the phases of the 4D CT dataset. They did not find a statistically significant difference between tumour segmentation based on deformed contour propagation and manual delineation in lung cancer patients, obtaining a method accuracy comparable to inter-observer variability.

The deformation vector field derived from daily in-room imaging for baseline correction is used to propagate the clinical GTV contours defined on the mid-position CT image. The tumour binary volume is extracted from the propagated contours, thus computing the updated average position of tumour baseline during the considered treatment fraction. The set of vector fields for each 4D CT phase of the breathing motion model is applied to the updated baseline image. An equal number of binary volumes is obtained, representing the tumour positions during all phases of the breathing cycle. The tumour centroid is extracted from each binary volume by computing the average coordinates along each spatial direction. This approach allows to compute the 3D positions of the tumour centroid at each phase of the average breathing cycle of the 4D CT planning scan, including baseline shift correction specific for each treatment fraction. As depicted in Figure 2.19, the tumour trajectories during a respiratory cycle define an ellipse in each orthogonal plane.

The reconstruction of the 3D tumour trajectory during the whole course of the treatment fraction was performed by combining the phase and amplitude parameters extracted from the optical surface acquisition with the information on tumour position at each breathing phase derived from the 4D CT motion model. As described in Section 2.3.3, the phase parameter defines the interpolated 4D CT phase corresponding to each acquired surface frame. At each time instant, the phase parameter is used to select the two 4D CT image frames closer to the considered respiratory phase value. The tumour centroid positions associated to the selected image frames are interpolated according to the percentage value defined by the phase parameter, thus obtaining the 3D coordinates of the tumour centroid at that specific breathing phase. The amplitude parameter specifies instead the difference in the amplitude of breathing motion between planning and

treatment phases. This value is used to scale the amplitude of tumour motion according to the following equation:

$$\bar{p}(t) = \bar{c}(t) + a(t) \cdot \left[\bar{c}(t) - \bar{b}\right], \tag{2.17}$$

where *a(t)* is the amplitude parameter at time instant *t*, $\bar{p}(t)$ is the coordinate vector defining tumour centroid position corrected for amplitude variations, $\bar{c}(t)$ is the non-corrected position derived from the interpolation of the model's tumour centroids and $\bar{b}$ is the updated position of tumour baseline. As stated in the equation (2.17), the amplitude correction factor is obtained by multiplying the amplitude parameter for the vector difference between the non-corrected centroid position and the tumour baseline. An example of tumour motion signals estimated with and without amplitude correction is shown in Figure 2.20. This approach allows a non-uniform amplitude scaling in the three spatial directions, since it depends on the entity of tumour motion along that direction at each specific breathing phase. The output of the proposed tumour tracking procedure corresponds to the 3D trajectory of the tumour centroid for each surface frame acquired during the treatment fraction. A low-pass moving average filter with a window size of 3 samples is applied to smooth and reduce noise from the resulting tumour motion signals.

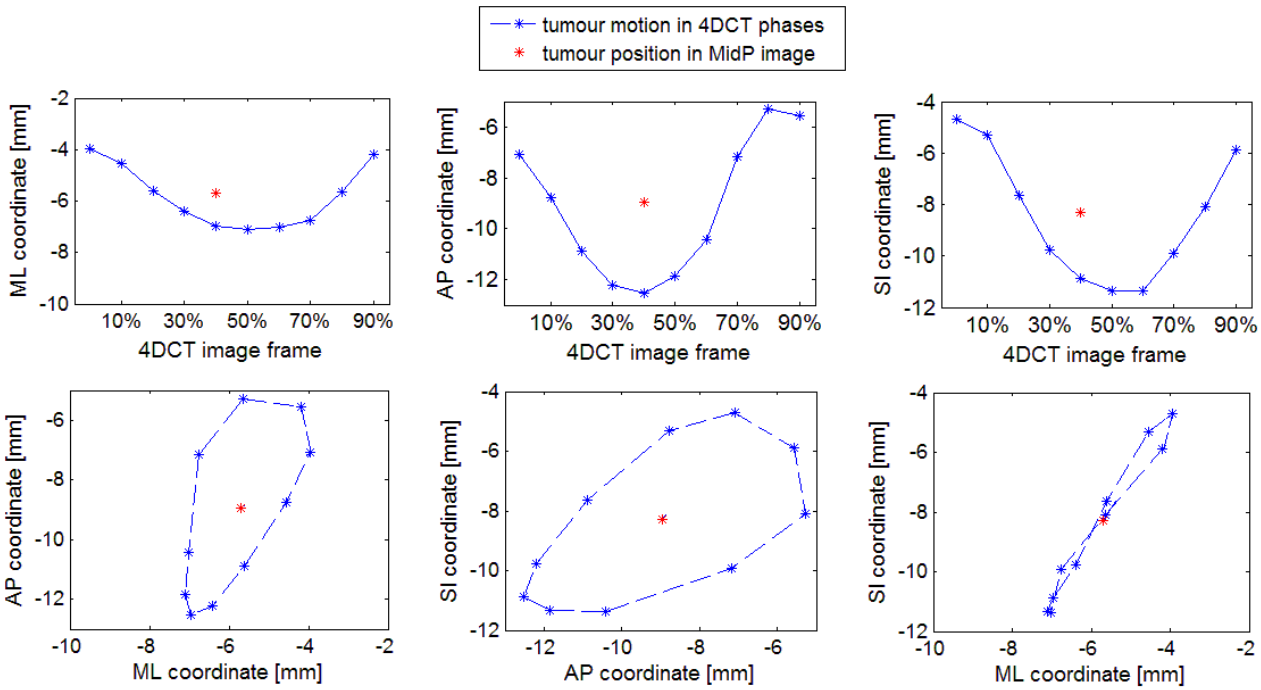


Figure 2.19. Tumour centroid positions in the MidP CT image (red marker) and in each phase volume of the 4D CT scan (blue markers) along the three spatial directions, obtained by applying the deformation vector fields of the 4D CT breathing motion model.

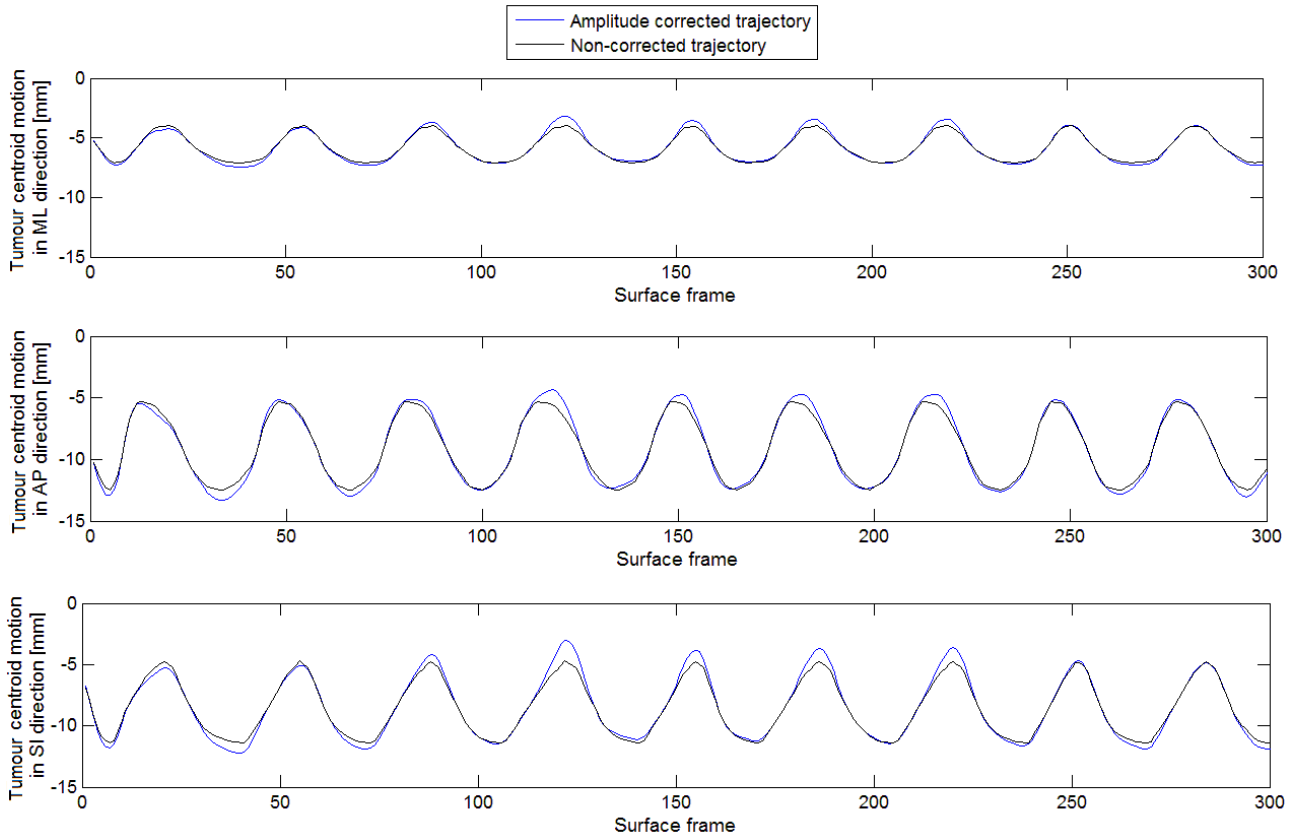


Figure 2.20. Comparison of tumour centroid trajectories along the three spatial directions estimated with and without amplitude correction. The depicted motion signals are obtained by combining the phase and amplitude parameters reported in Figure 2.18 with the tumour centroid positions of the 4D CT motion model represented in Figure 2.19.

2.5 Tumour identification on CBCT projections

As presented in Section 2.1.2, the evaluation of the proposed tumour tracking method is based on the comparison between the target trajectories estimated from the external surface motion combined with the 4D CT breathing model and the real tumour trajectory extracted from the cone-beam CT projections acquired simultaneously. As further explained in the Experimental testing chapter, the assessment of the tracking accuracy is performed on lung cancer patients, featuring non-negligible intra-fraction tumour motion due to the respiration process. The proposed approach for the evaluation of the developed tumour tracking method therefore requires the identification of lung lesions in CBCT projections. Since the considered patients do not present any implanted fiducial marker for lung tumour localization, target detection in X-ray images relies on markerless tracking techniques.

The algorithms proposed for markerless tracking of lung tumours in 2D projection images acquired during radiotherapy treatments can be divided into two groups: classification algorithms and direct tracking algorithms [Rottmann 2010]. In classification algorithms, a sequence of X-ray

projections acquired prior to treatment is used to build a training dataset for identifying different tumour location states. The actual tracking process consists in assigning each new image to one of the previously defined tumour locations. Different machine learning algorithms, such as Support Vector Machines (SVM) and Artificial Neural Networks (ANN), have been reported for tumour position classification in fluoroscopic images [Cui 2008, Lin 2009]. ANN technique allows to achieve a better performance than SVM in terms of classification accuracy and recall rate, although the target coverage is similar for both methods. The classification approach lacks however the ability to adapt to changes in the breathing pattern that are not described by motion observations during the training session. In direct tracking algorithms, a set of features is extracted and localized on each image frame, thus robustly capturing irregular breathing patterns. Single or multi-region template matching methods have been proposed for target tracking in electronic portal images. Tumour templates are obtained from manual or automatic segmentation in a reference projection, usually the first acquired image of the time-series [Arimura 2009, Rottmann 2010]. The multi-region approach based on the selection of a set of landmarks with a small surrounding region proved to be more robust with respect to target deformations and rotations, due to the inherent flexibility of multiple templates. State-of-the-art methods for markerless lung tumour tracking in rotational CBCT projections are instead based on the cross-correlation with templates derived from digitally reconstructed radiographs, obtained by projecting the planning CT volume at corresponding angles [Hugo 2010, Lewis 2010].

The reported direct techniques for markerless tumour tracking are limited by the availability of traceable features in the analyzed images [Rottmann 2010]. Soft tissues, such as lung lesions, usually feature low contrast and poor visibility in X-ray projections. Moreover, lung tumours may be difficult to distinguish at particular rotational angles due to the overlap of the surrounding high-contrast anatomical structures, such as the spine or the rib cage. These difficulties result in a reduced tumour tracking accuracy and in a limited range of projection angles at which lung targets can be tracked [Lewis 2010]. We developed and investigated an automatic method for contrast enhancement and direct markerless tracking of lung lesions in projection images, using planning CT information to increase tumour region visibility and reduce the overlying effect of the nearby anatomy [Fassi 2011]. As will be described in detail in the Experimental testing chapter, we assessed the effect of tumour contrast enhancement on the accuracy of target localization in CBCT projections and we compared tumour detection capabilities with those obtained with state-of-the-art tracking techniques.

2.5.1 Contrast-enhanced tracking technique

Figure 2.21 shows the overall procedure of the proposed markerless lung tumour tracking method. The developed approach consists of two main steps: contrast enhancement of lung tumour region and markerless tracking of target position in projection images. Both these processes exploit the anatomical information derived from 4D CT planning images, in order to reduce the overlying effect of the surrounding high-intensity structures, thus increasing lung tumour visibility. The proposed markerless tracking technique can be generally applied to all types of two-dimensional projection images, such as X-ray digital radiography, fluoroscopy and cone-beam CT. As described in the Experimental testing chapter, rotational CBCT acquisitions were selected for the evaluation of the developed contrast enhancement technique, to allow the assessment of tumour detection capability and accuracy at different projection angles. The two steps of the tracking procedure specifically applied to cone-beam CT projections are described in detail in the following paragraphs.

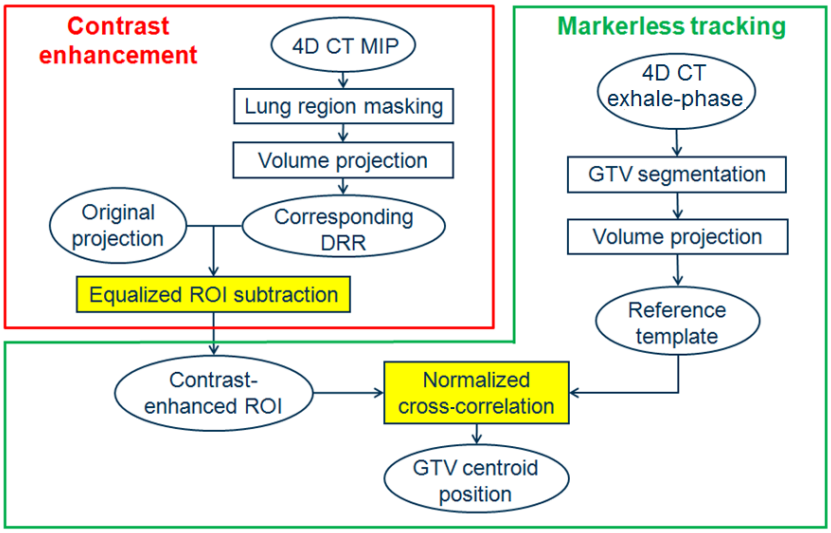


Figure 2.21. Flow-chart of the implemented procedure for lung tumour contrast enhancement and markerless tracking in projection images. All operations except the two highlighted processes can be performed a priori, thus reducing the computational cost.

2.5.1.1 Enhancement of tumour contrast

The enhancement of tumour contrast in projection images is based on masking and subtraction operations (Figure 2.21), which aim at removing from the original cone-beam CT images the contribution of the anatomical structures that obscure the lung lesion, thus enhancing tumour visibility. The developed procedure involves the computation of the digitally reconstructed

radiographs, consisting in two-dimensional simulated projections generated from the planning CT volume at the corresponding angle of the original CBCT images. The maximum intensity projection volume, representing the highest CT number found for each voxel in the set of 4D CT frames, is used for DRR generation, since it contains information about all respiratory phases. As depicted in Figure 2.22, lung regions, including the tumour, are segmented and masked in the MIP image before computing the DRRs, by applying automatic threshold-based segmentation algorithms and morphological operations [Hu 2001]. The clinical contours of the right and left lungs are used to extract a binary volume including lung regions. Since most lung tissues feature an intensity range from -910 HU (Hounsfield units) to -500 HU in CT image data [Wu 1994], a more relaxing threshold of -250 HU was selected for lung tumour segmentation. The masked MIP image was obtained by setting to the minimum intensity value (-1024 HU) all the voxels included in the lung binary volume having a grey-level intensity lower than the selected threshold value. The lung tumour was also masked in the MIP image, using the clinical contours of the planned target volume. The extended PTV contours were considered instead of the gross tumour volume to include possible errors in manual tumour delineation. The final masked MIP image features minimum intensity in all voxels of index *i*, *j*, *k* satisfying the following condition:

$$PTV(i,j,k)=1 \text{ or } \left[L(i,j,k)=1 \text{ and } MIP(i,j,k)<-250\right],$$
$$\text{for } i = 1...N,\ j = 1...M \text{ and } k = 1...Q \qquad (2.18)$$

where *MIP* is the MIP CT image of size *N*x*M*x*Q*, *L* is the binary lung volume and *PTV* is the binary PTV volume. A morphological opening operation, represented by an erosion followed by a dilatation, is performed using a disk-shaped structuring element, in order to refine lung contours and remove small high-intensity structures included in lung regions.

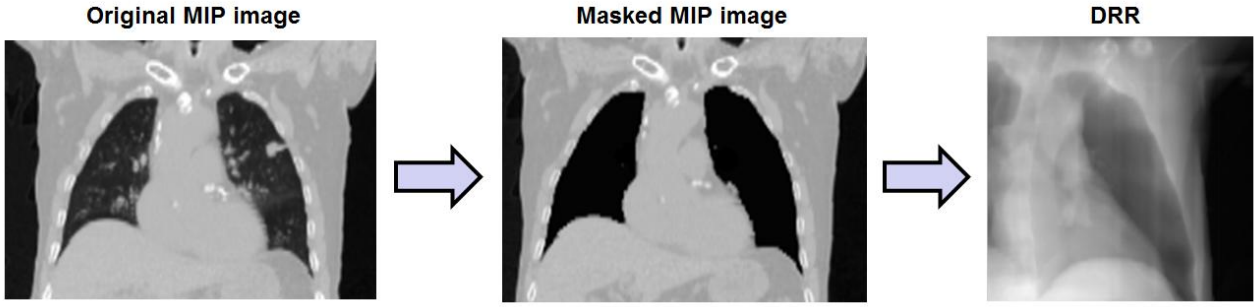


Figure 2.22. Example of digital reconstructed radiograph obtained by projecting the masked MIP image, which is computed by segmenting the lung and tumour regions in the original MIP volume.

Before DRR generation, we apply B-spline deformable registration algorithms [Sharp 2010] to align the masked MIP image with the 3D CBCT volume reconstructed at each treatment fraction,

to take into account non-rigid mismatches of the anatomical structures between planning and treatment phases. A multi-stage registration process is performed, using the normalized mutual information as similarity metric since multi-modal images with different intensity ranges are involved. Digitally reconstructed radiographs are generated by projecting the aligned MIP volume at each CBCT projection angle by using ray-tracing algorithms. DRR are obtained by summing the attenuation of each voxel along known ray paths through the data volume. The number of rays for DRR generation is determined by the image resolution, usually chosen to match that of the flat-panel X-ray detector. In the present application, DRR are obtained by using the Plastimatch open-source software [Sharp 2009], which implements the Siddon's ray-tracing method [Siddon 1985]. Siddon's algorithm is based on the "exact path-length method", which consists in the calculation of the intersection of the radiological path with the voxels of the three-dimensional CT array. Instead of treating the CT image as individuals voxels, Siddon's method considers the CT data as composed by the intersection volumes of three orthogonal sets of equally spaced, parallel planes. The intersection of the ray with the voxels are obtained as a subset of the intersections of the ray with the planes. For each voxel intersection length, the corresponding voxel indices are computed and the products of the intersected length and particular voxel density are summed over all intersections to yield the radiological path. For a 3D CT array of N^3 voxels, the Siddon's algorithm scales with $3N$ number of planes, rather than N^3 number of voxels, thus resulting more efficient and less time-consuming. Plastimatch implementation of the Siddon's ray-tracing method also includes different voxel interpolation methods, which can be used to increase the apparent resolution of DRR construction. An example of DRR generated with Plastimatch software is depicted in Figure 2.22.

The use of Plastimatch DRR program required the specification of the following parameters, describing the imaging geometry of the cone-beam CT system:

- SAD (source-to-axis distance), representing the distance between the X-ray source and the isocenter of the imaging system [mm];
- SID (source-to-imager distance), corresponding to the distance between the X-ray source and the panel detector of the imaging system [mm];
- spatial resolution of the panel detector [pixels];
- physical size of the panel detector [mm].

If used in the rotational mode, Plastimatch program provides the specified number of DRRs at equally spaced angles. However, during a real cone-beam CT scan the difference between subsequent angles is not uniform, since the speed of gantry rotation may vary due to mechanical inertial constraints. We therefore modified the Plastimatch code to accept in input the list of the

rotational positions of the gantry during a specific CBCT acquisition, thus generating the DRRs exactly at the desired angle values.

An example of the implemented contrast enhancement and template matching technique applied to a cone-beam projection is shown in Figure 2.23. To reduce the computational cost, tumour search was limited within a restricted region of interest. Tumour ROI was extracted from both the original cone-beam image and the DRR simulated at the corresponding projection angle. The ROI centre was placed at the imager centre, corresponding to the 2D coordinates of the isocenter of the cone-beam CT imaging system projected on the panel detector. Flexmap corrections, describing the discrepancy between the radiation isocenter and the mechanical isocenter of the CBCT [Bissonnette 2008], were considered for the computation of the ROI centre for the simulated DRR. The ROI size was selected as a function of the tumour extents derived from the clinical ITV contours derived from the MIP CT image, adding a safety margin to take into account possible variations in target baseline and motion amplitude. As depicted in Figure 2.23, the contrast-enhanced tumour ROI is obtained by subtracting the corresponding DRR from the original CBCT projection. Since DRR and CBCT image intensities are not linearly correlated [Yang 2012], non-linear histogram equalization is applied to each ROI in order to match different grey level values before subtraction. The equalization procedure transforms the image intensity values into N discrete gray levels, such that a roughly equal number of pixels is mapped to each of the N levels, thus producing an approximately flat histogram. For the present application, the number of grey levels in the equalized images was set to 64.

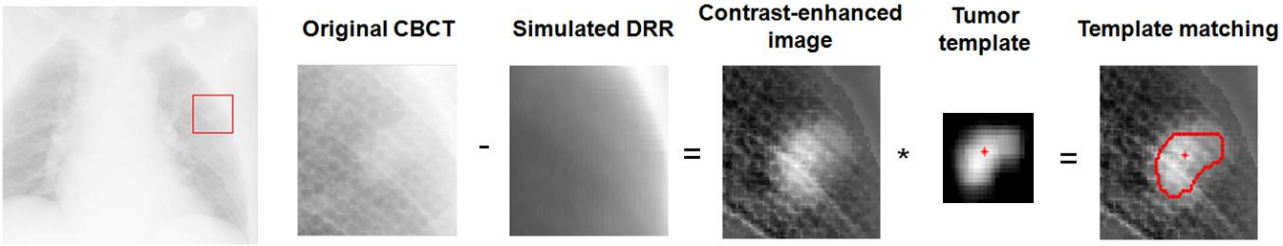


Figure 2.23. Application of the developed contrast enhancement and template matching technique to an exemplificative cone-beam projection. The circular structures visible in the original CBCT image come from the stereotactic body frame applied for patient's immobilization.

2.5.1.2 Tumour template matching

The tracking of tumour position on the contrast-enhanced images is realized through template matching techniques based on fast normalized cross-correlation. The reference tumour template for each CBCT rotational position is obtained by projecting at the corresponding angle the

gross tumour volume extracted from a single 4D CT phase image. The exhale phase was chosen as reference image for template generation due to the higher stability of tumour position at the end of expiration [Seppenwoolde 2002]. Since in our patient dataset the target contours are clinically delineated only in the planning MIP volume, we apply deformable image registration to propagate tumour contours in the 4D CT exhale phase. The deformation vector field generated by registering the MIP and exhale phase images with B-spline deformable registration algorithms is used for contour propagation to extract tumour volume in the selected 4D CT phase. Reference target templates for each cone-beam image are obtained by applying ray-tracing algorithms to the GTV volume extracted from the propagated contours. For each projection angle, the size of the tumour template is selected according to the size of the projected GTV, by adding a 3 pixel uniform margin in each dimension. The GTV centroid is also projected on the two-dimensional plane of each reference template. Tumour centroid is identified on each CBCT image by performing normalized cross-correlation between the contrast-enhanced tumour ROI and the corresponding reference template.

The cross-correlation function between the image *f* and the template *t* located at image coordinates *(u,v)* is based on a pixel-wise comparison according to the following equation:

$$c(u,v) = \sum_{x,y} f(x,y) \cdot t(x-u, y-v) \,. \tag{2.19}$$

The use of cross-correlation for template matching can be negatively influenced by several factors, such as the local intensity value of the image or the size of the template [Briechle 2001]. To overcome these difficulties, the image and template vectors are normalized to unit length, thus obtaining the normalized cross-correlation *n(u,v)*:

$$n(u,v) = \frac{\sum_{x,y} \left[f(x,y) - \bar{f}_{u,v} \right] \cdot \left[t(x-u, y-v) - \bar{t} \right]}{\sqrt{\sum_{x,y} \left[f(x,y) - \bar{f}_{u,v} \right]^2 \cdot \sum_{x,y} \left[t(x-u, y-v) - \bar{t} \right]^2}} \,, \tag{2.20}$$

where $\bar{t}$ is the mean intensity value of the template and $\bar{f}_{u,v}$ is the mean value of the image within the area of the template. Due to the normalization factor, the use of *n(u,v)* for template matching is more robust than other similarity measures, since it's independent to changes in brightness or contrast of *f*. The image coordinates featuring the highest value of the normalized cross-correlation coefficient correspond to the position of the identified GTV centroid. In order to reduce the computational cost, the image cross-correlation function, corresponding to the nominator of the equation (2.20), is computed in the frequency domain using the Fast Fourier Transform. The denominator of (2.20) is instead efficiently calculated by pre-computing running-sum tables containing the integral of the squared image function $f^2(x,y)$. These tables are used to evaluate the

expression $\left[f(x,y)-\bar{f}_{u,v}\right]^2$ at each point *(u,v)* where the normalized cross-correlation coefficient is calculated [Lewis 1995]. The low computational cost of the proposed tumour tracking method is also associated to the possibility of performing in advance the most time-consuming operations, such as deformable image registration and DRR generation. Calculation time only depends on the few processes highlighted in Figure 2.21, namely image equalization, subtraction and cross-correlation, requiring together less than 10 ms on average. The application of tumour search within a restricted region of interest also contributes to reducing the computational cost to levels compatible with real-time performance.

A direct tumour tracking approach similar to the described contrast enhancement method was independently proposed by Yang et al. [Yang 2012], although important differences and methodological improvements are evident. The cited technique involves the generation of two sets of DRRs, the first one including only the tumour and the second one with the complete anatomy without the tumour. The "anatomy-without-tumour" dataset is obtained in Yang et al. by masking the tumour in the CT volume with the average lung tissue value. We masked not only the lesion but the whole lung region, as a way to reduce tracking sensitivity to tumour position in the CT scan. In addition, the rigid registration proposed in [Yang 2012] between the CBCT projections and the "anatomy-without-tumour" DRR may not take into account non-rigid mismatches of the anatomical structures. To overcome this limitation, we applied deformable registration between the "anatomy-without-tumour" CT dataset and the reconstructed 3D CBCT volume before generating the DRR. Finally, Yang et al. used the average CT scan to derive the "tumour-only-templates" for cross-correlation. Instead, we considered a single 4D CT phase image, which may better represent the size and shape of the lesion in CBCT projections.

As described in the Experimental testing section, we developed an improved and more extended methodological approach for the accuracy evaluation of the proposed contrast enhancement method with respect to the experimental testing performed by Yang et al. The tumour tracking accuracy was evaluated in [Yang 2012] only for the SI tumour coordinate, obtaining errors less than 2.2 mm. We instead investigated the tracking performance of the proposed contrast-enhancement technique along all three spatial directions, to provide a more comprehensive and complete assessment of the method accuracy. In addition to the testing procedure described by Yang et al., we performed a comparative analysis of the tumour tracking performance with respect to the state-of-the-art techniques for direct markerless target detection in CBCT projections [Lewis 2010, Hugo 2010], to allow the objective evaluation of the effective gain of the proposed contrast enhancement technique.

2.6 Synchronization between CBCT and surface optical acquisitions

As previously introduced, the accuracy of the proposed tumour tracking method is quantified by comparing the real target motion derived from cone-beam CT acquisitions with the tumour trajectory indirectly estimated from the displacement of the patient's external surfaces simultaneously acquired with optical systems. The evaluation technique therefore required the development of a specific strategy for the synchronization between CBCT and surface optical scans, in order to retrieve the temporal correspondence between each cone-beam image and the dynamic time-series of external surfaces. The first approach implemented for synchronization purpose was based on the use of the BlackCat X-ray detector, which records the number of radiation particles detected in the treatment room, with a temporal resolution of approximately 10 msec. The BlackCat detector was connected directly to a serial port of the optical system, thus yielding the absolute timestamps of the particle measurements expressed in the time frame of the optical system. Figure 2.24 shows the number of radiation particles detected by the BlackCat system during a CBCT scan. The provided information can be used to derive the starting time of the cone-beam acquisition in the temporal frame of the optical system. Since CBCT scan consists in a sequence of multiple X-ray projections acquired at different rotational positions, the measured BlackCat signal has a pulsed time-trend (Figure 2.24). The CBCT starting time can therefore be computed as the absolute timestamp of the first detected radiation pulse.

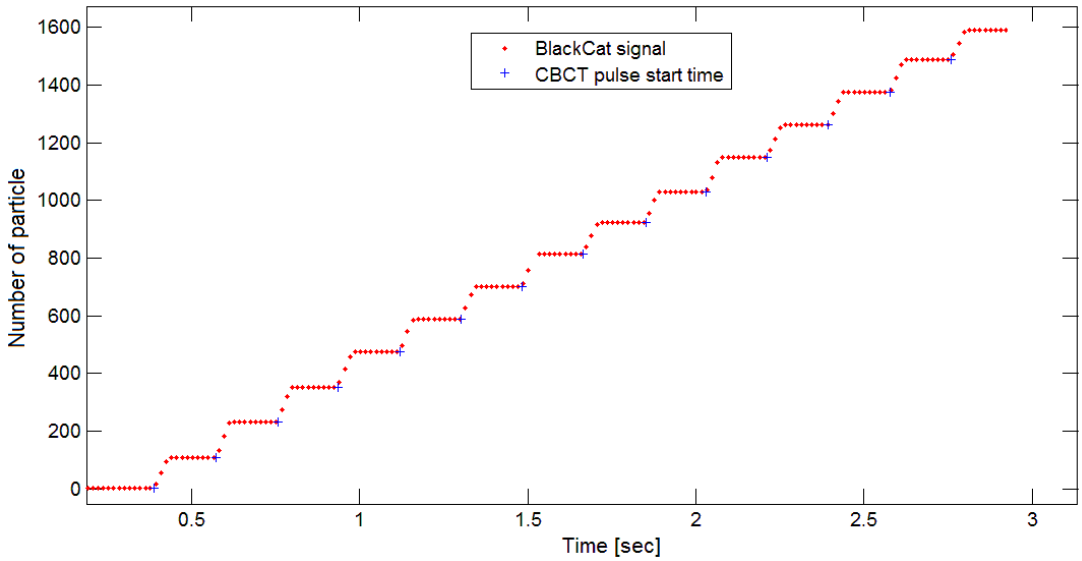


Figure 2.24. Number of radiation particles measured by the BlackCat X-ray detector during the acquisition of cone-beam images. The timestamps associated to each CBCT pulse is also depicted.

Unfortunately, the BlackCat approach could not be applied in our case for the synchronization of cone-beam CT and surface acquisitions. In fact, the video-based optical system employed for surface scanning did not provide any information on the absolute timestamps associated to surface acquisition, hindering the possibility to relate the starting time of the CBCT scan with the surface timestamps. The only information available from the surface imaging system concerns the temporal difference between the acquisition of each surface frame, thus providing the relative timestamps with respect to the starting time of the surface scan. To overcome this limitation of the employed optical device, we developed an image-based approach for CBCT and optical surface synchronization, which does not require the knowledge of the absolute acquisition timestamps of surface data.

2.6.1 Image-based synchronization method

The developed method for the synchronization of CBCT and optical surface acquisitions exploits the two-dimensional photos acquired by the stereo-cameras of the video-based optical system employed for surface scanning. An exemplificative photo captured by the optical device at the beginning of a cone-beam CT acquisition is shown in Figure 2.25a. The three photos acquired simultaneously by the two stereo-cameras and the texture camera of the surface imaging system are superimposed in a single image as different RGB (Red, Green, and Blue) components. In particular, the two photos acquired by the optical stereo-cameras are used for the 3D reconstruction of the patient's thoraco-abdominal surface, while the photo captured by the texture camera is used for grey-level coloration. As depicted in the figure 2.25a, a portion of the flat-panel detector of the CBCT device is visible in the photos captured by the optical system at the beginning of cone-beam CT acquisition, when the panel starts to rotate around the patient. We propose to synchronize CBCT and optical surface acquisitions by tracking the motion of the CBCT detector in the 2D images captured by the optical cameras. The tracking is performed on the blue image component extracted from the captured photos, since it contains a larger portion of the visible panel (Figure 2.25b).

The beginning of cone-beam CT scan can be detected in the stereo-images captured by the optical system by identifying the starting motion of the CBCT rotating panel, thus obtaining the relative starting time of cone-beam CT with respect to the optical surface acquisition. The elliptic-shaped feature shown in Figure 2.25b was selected to track panel motion, since it was clearly visible on the photos captured by the optical system. Threshold-based segmentation algorithms and ellipse fitting operations were applied to estimate the position of the selected feature on the 2D images. A region of interest including the panel's feature is extracted from the captured photos, as shown in

Figure 2.26a. A grey-level threshold is applied to the selected ROI, thus segmenting the elliptic feature of interest (Figure 2.26b). The convex hull, which is the smallest convex polygon enclosing a set of points, is computed from the thresholded image. A least-squares fit of an ellipse model is obtained from the 2D vertices of the extracted convex hull (Figure 2.26a). The estimated trajectory of the panel's elliptic feature is used to derive a mean estimate of the temporal difference between CBCT and surface acquisitions, as described in detail in the next paragraphs. Interpolation procedures are introduced to increase the temporal resolution of the proposed synchronization method, overcoming the limit imposed by the acquisition frame rate of the optical imaging system.

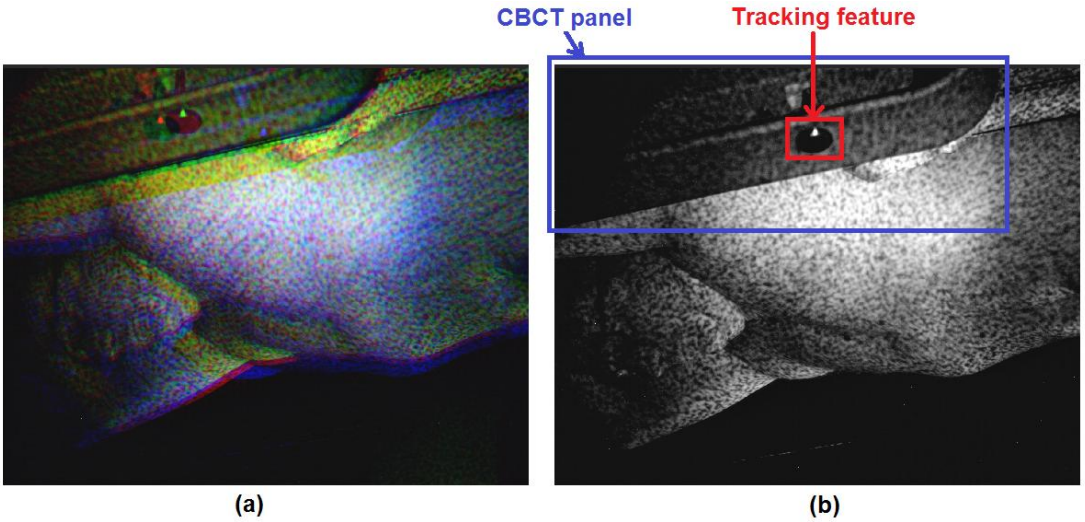


Figure 2.25. Photos captured by the three cameras of the optical surface imaging system superimposed in a single RGB image (Panel a). The blue image component, including the larger portion of the CBCT panel detector, is extracted from the captured photos (Panel b). The elliptic feature used for the tracking of the CBCT panel motion is shown in the figures.

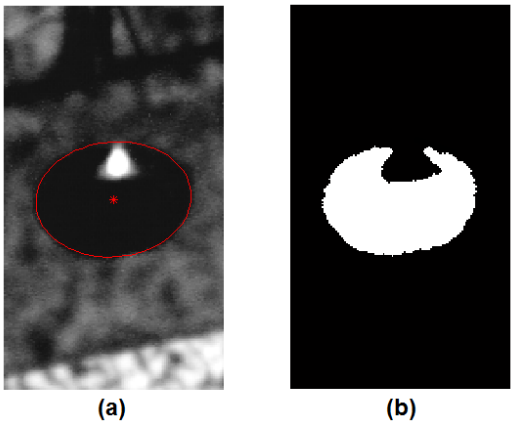


Figure 2.26. ROI of the optical system's photo including the elliptic feature used for the tracking of CBCT panel motion (Panel a). The red star represents the centre of the ellipse fitted to the vertices of the convex hull extracted from the thresholded feature of interest (Panel b).

A linear angular model was built to relate the position of the CBCT panel detector identified in the optical system's photos with the projection angles associated to the corresponding cone-beam images. As depicted in Figure 2.27a, 30 rotational positions of the CBCT device ranging from $-176°$ to $-180°$, which corresponds to the starting cone-beam angle, were considered for the model construction. For each of the 30 selected positions, the gantry was rotated at the corresponding angle and 50 photos were captured with the optical system. The 2D coordinates of the centre of the panel's elliptic feature were identified on the acquired photos using the previously described method. The distance of the panel feature's centre from the top-left corner of the selected ROI was computed and averaged over the 50 captured photos. This process was repeated for each of the CBCT projection angles included in the model, storing the estimated mean distances of the correspondent panel's feature. The stored data are fitted with a first order polynomial (Figure 2.27a), approximating with a linear model the motion of the rotating CBCT panel in the optical system's photos. The fitted polynomial of the angular model is then used for the synchronization of CBCT and surface acquisitions at each treatment fraction according to the following procedure.

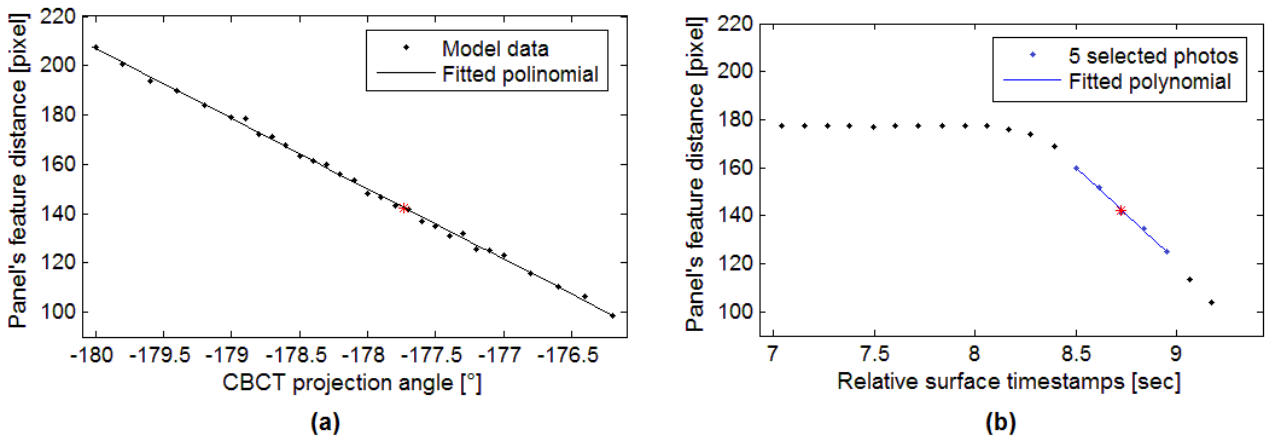


Figure 2.27. Linear angular model relating the CBCT projection angles with the distances of the panel's feature from the top-left corner of the ROI extracted from the photos captured with the optical system (Panel a). The feature's distances computed during a cone-beam acquisition are shown in Panel b. The blue dots represents the 5 selected photos acquired soon after the beginning of the CBCT scan. In both panels, the red stars correspond to the feature's distance (Panel b) and to the corresponding projection angle (Panel a) associated to the mean value of the relative acquisition timestamps of the five selected photos, obtained by interpolating with the respective fitted polynomials.

The first 5 photos captured by the optical surface imaging system after the beginning of the CBCT rotational acquisition are identified based on the estimated distance of the panel's elliptic feature from the top-left corner of the selected ROI. As shown in Figure 2.27b, the feature's distances associated to the selected five photos are linearly interpolated to derive the feature's distance at the mean value τ of the relative acquisition timestamps of the considered photos. The

stored model polynomial is used to estimate the CBCT projection angle correspondent to the mean feature's distance (Figure 2.27a). The retrieved projection angle is then related to the elapsed time since the beginning of the CBCT scan. This relationship is derived from the acquisition log file associated to the specific cone-beam CT scan, which includes the information about the projection angles and the relative timestamps of each acquired CBCT image. The data stored in the log file are linearly interpolated to derive the elapsed time ε since the beginning of the cone-beam CT scan associated to τ. The temporal difference between the CBCT and optical surface acquisitions is then computed as $\tau - \varepsilon$. By subtracting the estimated difference from the relative timestamps associated to each optical surface, we obtain the corresponding timestamps expressed in the temporal frame of the cone-beam CT system, thus allowing the synchronization between CBCT and surface scans. In particular, only surface frames with positive timestamps are acquired during the cone-beam CT scan. The associated timestamp values indicate the elapsed times since the start of CBCT acquisition.

3. EXPERIMENTAL TESTING

3.1 Lung cancer patient database

As described in the Methods chapter, we developed a tumour tracking method for the non-invasive indirect estimation of the internal tumour position from the external surface displacement, for the compensation of intra-fraction target motion mainly due to respiration. A clinical patient database was specifically acquired for the experimental testing of the proposed tracking technique, including the synchronized information on internal and external breathing motion. Lung cancer patients were selected for the evaluation of the method accuracy, since they feature non-negligible intra-fraction tumour motion due to the respiratory process. All patient data were collected at the Centre Léon Bérard (CLB) in Lyon, one of the main clinical facilities in France dedicated to cancer treatment and research. The acquired database includes early-stage non-small cell lung cancer patients, treated at CLB with stereotactic body radiation therapy (SBRT). For this type of malignancies, SBRT provides a substantial improvement of treatment outcomes in terms of local control and survival rate with respect to conventional radiation therapy [Timmerman 2010].

3.1.1 Stereotactic body radiation therapy

SBRT consists in the delivery of biologically equivalent doses that can exceed 150 Gy in hypofractionated schedules with 5 or less fractions [Hiraoka 2007]. Medical indications for this treatment modality include NSCLC peripheral tumours with a diameter smaller than 5 cm and a distance to the proximal bronchial tree greater than 2 cm [Timmerman 2010]. The delivery of such high doses per fraction is achieved with highly conformal beam arrangements, requiring very high setup accuracy and robust motion management techniques. Patient immobilization is usually enforced with specific devices, such as stereotactic body frames, vacuum systems or abdominal compression, as shown in Figure 1.19 [Wulf 2000]. According to the clinical protocol adopted at the Centre Léon Bérard, patients undergoing SBRT treatments are immobilized with the stereotactic

body frame from Elekta (SBF, Elekta, Stockholm, Sweden). A fluoroscopic image acquisition is performed prior to treatment to evaluate the extent of tumour motion. If motion amplitude exceeds 1 cm in the superior-inferior direction, abdominal compression is applied to reduce tumour extent [Gassa 2006].

Treatment planning for NSCLC patients at CLB is routinely performed by acquiring a 4D CT scan. Ten breathing phase-sorted images and the maximum intensity projection are reconstructed from the 4D CT dataset. The internal target volume, encompassing the total extent of tumour motion for all phases of the breathing cycle, is manually delineated by clinicians on the MIP image. The irradiated planned target volume is obtained by adding to the ITV a treatment margin of 5 mm in the transversal plane and 8 mm in the longitudinal plane [Gassa 2006]. The SBRT treatments usually commence within ten days from the 4D CT planning acquisition. For NSCLC patients, a total dose of 48 Gy is usually delivered in four fractions of 12 Gy, performed twice a week for a total of two weeks. This hypofractionated schedule corresponds to a biologically equivalent dose of 105 Gy [Gassa 2006]. The number of treatment sessions and the prescribed dose per fraction can be adapted as a function of the tumour size. Treatment delivery is performed using the Elekta SynergyS linear accelerator (Elekta, Stockholm, Sweden) equipped with the Beam Modulator system, a beam-shaping multi-leaf collimator composed by 40 fully inter-digitating leaves of 4 mm. The Elekta SynergyS machine incorporates different imaging systems, including kV planar radiography, 2D fluoroscopy for real-time motion tracking and 3D cone-beam CT for soft tissue visualization. A photo of the Elekta SynergyS system with CBCT imaging functionality installed at CLB is depicted in Figure 3.1.

According to the clinical protocol adopted at the Centre Léon Bérard for NSCLC treatments with SBRT, at the beginning of each fraction the patient is repositioned in the stereotactic body frame through laser alignment. A cone-beam CT scan is routinely performed to allow a high precision patient setup, using the 3D CBCT volume reconstructed with the XVI software from the acquired two-dimensional cone-beam projections. Rigid registration is performed between the daily CBCT and the planning CT volume, obtaining the roto-translational corrections for target alignment with the treatment isocenter. If the measured translations or rotations are higher than 3 mm or 1°, respectively, patient position is corrected by moving the treatment couch [Gassa 2006]. The Elekta SynergyS system installed at CLB is equipped with the HexaPOD computer-controlled robotic table, which allows a sub-millimetre and sub-angular accuracy in patient positioning with six degrees of freedom, compensating for both translational and rotational geometric errors (roll, pitch and yaw) [Meyer 2007]. The real-time continuous feedback on the actual position of the treatment couch is provided by means of an infrared tracking camera (Polaris, NDI, Waterloo, Ontario,

Canada) mounted to the room ceiling, which records the 3D coordinates of a set of multiple passive reflectors rigidly connected to the treatment couch (Figure 3.1). A photo of the horizontal bar with the passive markers used for the optical infrared tracking of the HexaPOD table motion is shown in Figure 3.2. After treatment couch correction, a second cone-beam CT scan is acquired for setup verification. If the measured roto-translational parameters are below the pre-defined thresholds, the radiation treatment is started. Each SBRT session is performed through 8 to 10 radiation fields, for a total duration of about 30 minutes per fraction [Gassa 2006].

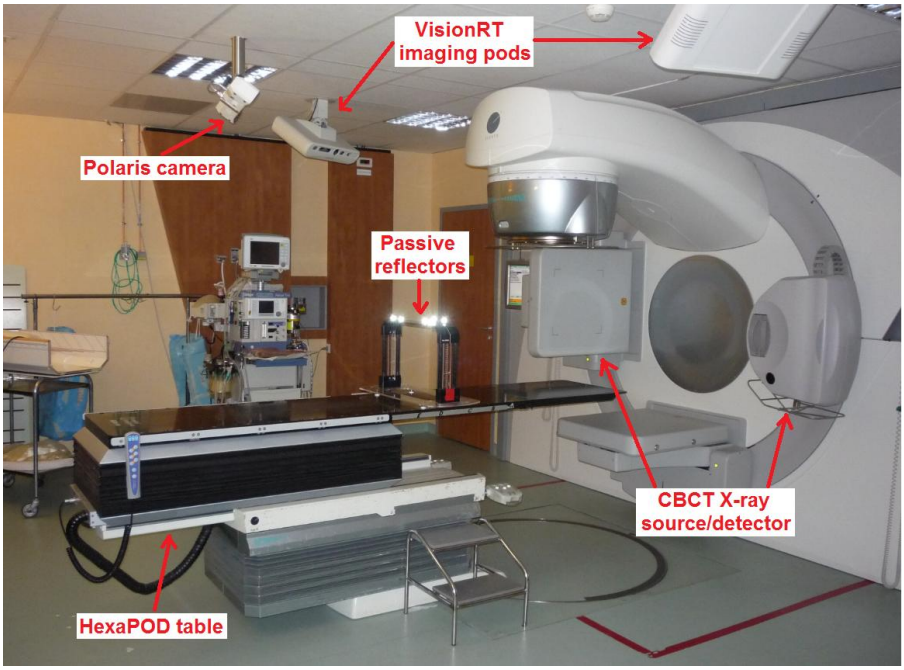


Figure 3.1. Photo of the treatment room at the Centre Léon Bérard where the Elekta SynergyS system is installed. The couple of X-ray source and panel detector for cone-beam CT imaging and the VisionRT pods for surface scanning are depicted in the figure. The HexaPOD table is also highlighted, including the Polaris camera and the passive reflective markers attached to the top of the treatment couch.

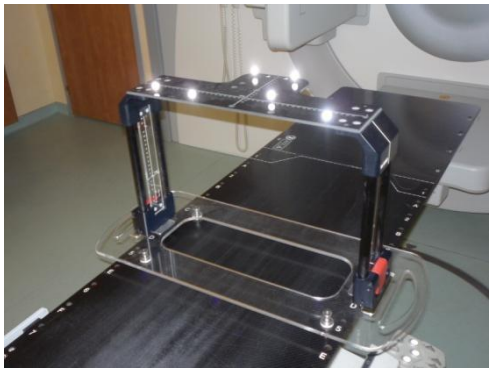


Figure 3.2. Photo of the HexaPOD horizontal bar with multiple passive markers connected to the treatment table, used for the optical infrared tracking of the couch position.

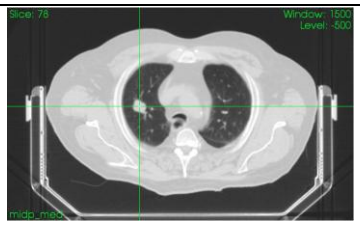
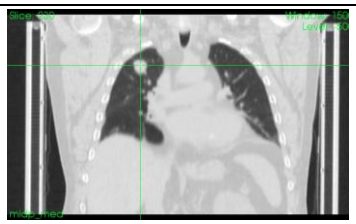
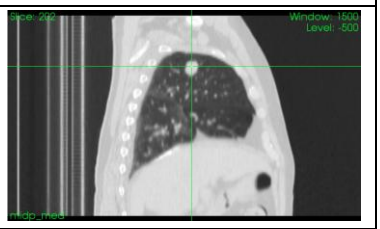
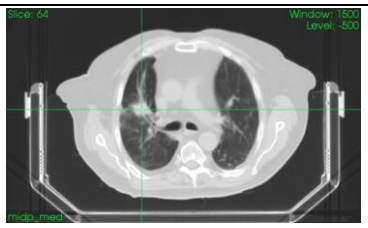
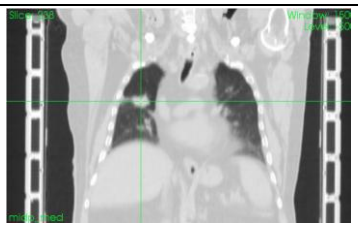
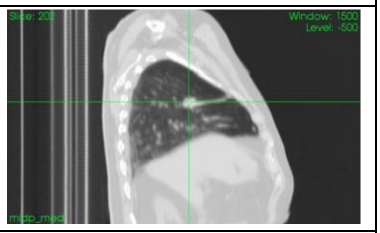
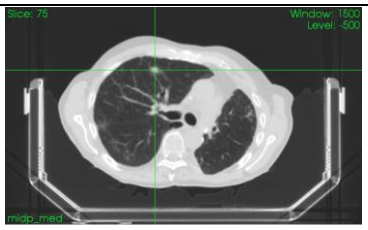
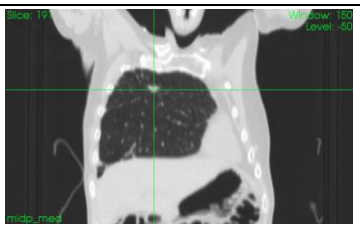
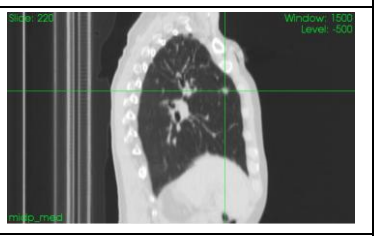
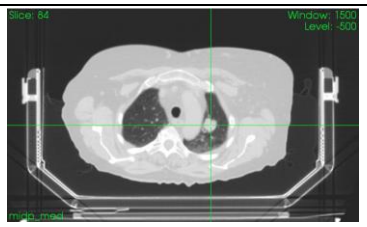
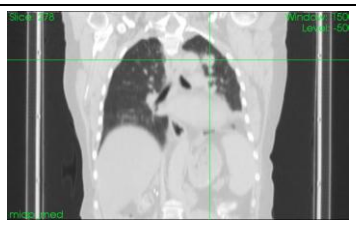
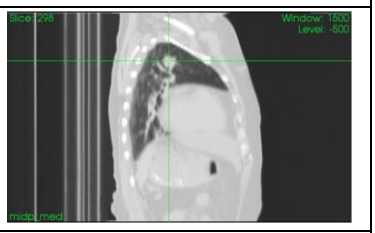
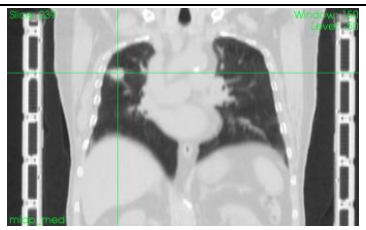
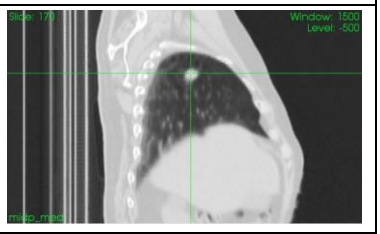
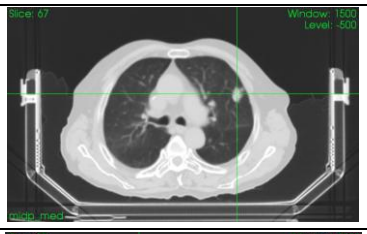
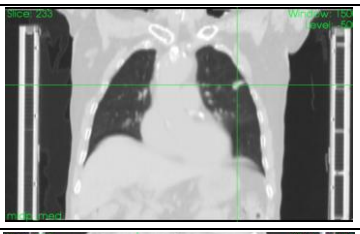
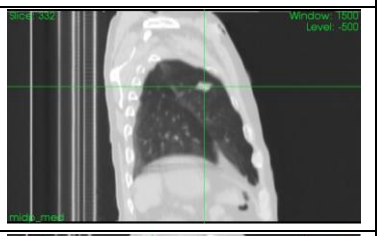
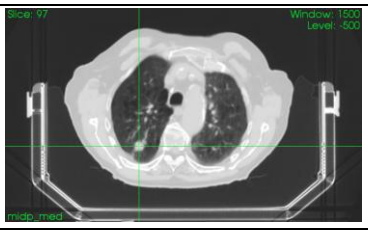
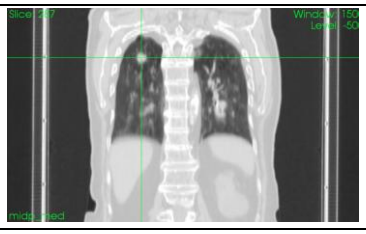
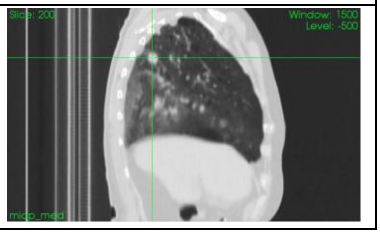
3.1.2 Patient characteristics

The clinical database acquired for the experimental testing of the developed surface-based tumour tracking method includes 7 early-stage non-small cell lung cancer patients undergoing stereotactic body radiation therapy at the Centre Léon Bérard. The acquisition protocol was focused only on patients treated without the abdominal compression, since the horizontal plate pressing on the abdomen covered a large portion of the patient's external surface, thus hindering the possibility to optically capture the intra-fraction surface motion. As described in the previous section, the compression device is prescribed in SBRT treatments at CLB for patients showing an amplitude of tumour motion higher than 1 cm. The requirement of no abdominal compression in the present study implied that only patients with a reduced tumour motion extent were collected in the database. As shown in Table 3.1, the acquired database includes six males and one female patients, with an age ranging between 65 and 85 years. For all the considered patients, the tumoral lesion is located in the upper lobes of the lung. The patient-specific target shape and position inside the lung are depicted in Table 3.2, showing for each patient the reconstructed mid-position CT volume along three orthogonal planes intersecting at the tumour centroid. For 5 out of 7 patients, the tumour is located in the right-side lung. The distance of the target centroid from the corresponding lung apex along the SI direction ranges between 3.5 and 9.0 cm (Table 3.1). The mean size of the GTV segmented in the mid-position CT images is equal to 2.1, 2.1 and 1.8 cm along the ML, AP and SI directions, respectively. The lung tumour volumes, measured as the number of voxel included in the GTV binary image extracted from the mid-position CT multiplied by the unit volume of a single voxel, vary from 0.5 to 5.1 cm^3, with a mean value of 2.7 cm^3.

Table 3.1. Characteristics of the lung cancer patients included in the clinical database specifically acquired for the experimental evaluation of the developed tumour tracking method.

Patient	Gender	Age	Tumour side	SI tumour distance from lung apex (cm)	Tumour size ML×AP×SI (cm)	Tumour volume (cm^3)
P1	Male	79	Right	3.5	2.2×2.2×2.3	3.7
P2	Male	85	Right	7.5	3.1×3.3×1.8	5.1
P3	Male	72	Right	9.0	1.4×0.8×0.8	0.5
P4	Female	71	Left	4.1	1.8×1.9×1.9	3.0
P5	Male	77	Right	5.1	2.4×2.0×1.8	3.0
P6	Male	65	Left	7.2	2.0×2.2×1.7	1.7
P7	Male	84	Right	7.1	1.9×2.0×2.1	2.0

Table 3.2. Mid-position CT image reconstructed from the 4D CT scans for the 7 lung cancer patients of the database. The MidP volumes are visualized along three orthogonal planes intersecting at the tumour centroid, indicated by a green cross. Image visualization is realized though the open-source image viewer VV [Rit 2011], using a lung window cantered in -500 HU.

Patient	Transverse plane	Coronal plane	Sagittal plane
P1			
P2			
P3			
P4			
P5			
P6			
P7			

3.1.3 Internal and external motion data

The patient database specifically acquired for the experimental testing of the developed surface-based tumour tracking method includes clinical data provided by several multi-modality imaging systems installed at the Centre Léon Bérard. The collected data yields synchronized internal and external information on respiratory tumour motion during radiotherapy planning and treatment phases. In the following paragraphs we describe in detail the imaging data collected for each of the 7 NSCLC patients of the database, which include:

- a 4D CT scan acquired for treatment planning;
- a cone-beam CT scan performed at the beginning of a treatment fraction for patient setup;
- an acquisition of the patient's external surface motion using a non-invasive optical tracking device synchronized with the cone-beam CT scan.

As described in the Methods chapter, the real tumour trajectory identified on the CBCT projections is compared to the tumour trajectory estimated from the synchronous optical surface displacement, combined with an a priori patient-specific model of the breathing target motion derived from the planning 4D CT dataset.

3.1.3.1 4D CT planning scan

The 4D CT planning scan is acquired by using the Philips Brilliance CT Big Bore simulation system (Philips Medical Solution, Cleveland, OH), shown in Figure 3.3. The scanner operates in a helical mode, by acquiring 16 slices simultaneously with a rotation period of 0.5 sec and a pitch of 0.1. The pitch is defined as the ratio of the table translating distance per gantry rotation to the thickness of the individual X-ray beam. Breathing synchronization is achieved with the Pneumo Chest Bellows (Lafayette Instrument, Lafayette, IN), consisting in a pressure belt placed around the patient's chest to provide the surrogate breathing signal used for respiration-correlated 4D CT data binning. The acquired CT dataset is sorted into ten breathing phase images, with a slice thickness of 3 mm and a voxel size of $1.17 \times 1.17 \times 3$ mm^3. The spatial resolution of the output image in ML and AP directions is 512×512 pixels, while the number of slices acquired in the SI direction varies according to the vertical lung extent of the specific patient. Typically, about 100-170 slices are reconstructed, resulting in a physical acquired volume of $60 \times 60 \times 40$ cm^3.

The contours of the tumours and of the structures of interest, such as the lungs and the patient body, are manually delineated by a clinician on the maximum intensity projection image derived from the 4D CT dataset. The delineated tumour contours represent the ITV, corresponding to the volume covered by the moving GTV during each phase of the 4D CT breathing cycle. For

each patient included in the database, the 4D CT planning dataset is used to derive a patient-specific model of tumour breathing motion, as described in Section 2.2.1. Deformable image registration is applied to each patient dataset to estimate the mid-position CT image and the ten deformation vector fields between the MidP and every 4D CT phase volume. Deformable image registration is also used to propagate the ITV contours clinically outlined on the MIP image to the GTV contours in the MidP and in each of the ten breathing phase images.

Figure 3.3. The Philips Brilliance CT Big Bore scanner used for the acquisition of the 4D CT planning dataset.

3.1.3.2 Cone-beam CT acquisition

For each patient of the database, a cone-beam CT scan was acquired at the beginning of a treatment fraction using the Elekta SynergyS machine, equipped with the XVI on-board kV imaging system. As shown in Figure 3.1, the CBCT X-ray source and the corresponding flat-panel detector are mounted perpendicularly with respect to the treatment gantry, rotating simultaneously with the gantry. The cone-beam CT scan performed at CLB consists in acquiring approximately 650 2D projections at different angular positions around the patient, covering an entire 360° gantry rotation. The temporal frequency of CBCT image acquisition is 5.46 frames per second, thus requiring about 2 minutes for a full cone-beam CT scan. Motion-induced artefacts in the acquired projections are limited by reducing the acquisition time of each CBCT image to approximately 20–40 msec. The cone-beam projection sequence is acquired clockwise, starting with the X-ray source at the left-side of the treatment couch (gantry at -180°). The cone-beam CT scans for the lung cancer patients of the database were acquired with a medium field of view (FOV) and a M20 collimator. A schematic representation of the Elekta SynergyS system for the considered CBCT geometrical configuration is shown in Figure 3.4.

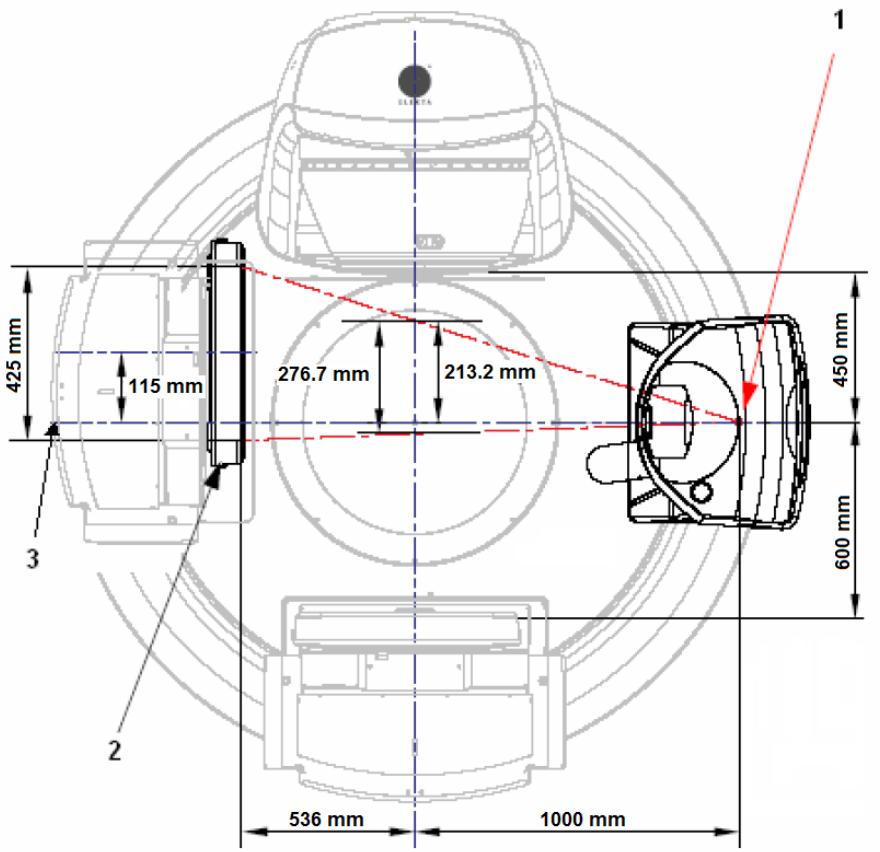


Figure 3.4. Geometrical arrangement of the Elekta SynergyS system for the medium imaging FOV. The corresponding positions of the kV X-ray beam focal spot (1), of the kV panel detector (2) and of the kV imaging reference axis (3) are depicted in the figure.

As reported on the Elekta Clinical User Manual, the source-to-axis distance (SAD) between the X-ray source and the isocenter of the XVI imaging system is 1000 mm, while the total source-to-imager distance (SID) between the X-ray source and the panel detector is 1536 mm. The physical size of the squared CBCT detector is 410×410 mm^2, while the spatial resolution is 512×512 pixels, resulting in an effective pixel size of 0.8 mm/pixel. The kV photons are directed in a circular cone-shaped beam from a focal spot of the X-ray source tube. The uncollimated cone incident perpendicularly on the flat-panel detector has a diameter of 425 mm. As depicted in Figure 3.4, in the medium FOV configuration with gantry at 0°, the inclination of the X-ray source beam is deviated of about 4.15° around the vertical direction (perpendicular to the detector plane) with respect to the imaging reference axis (perpendicular to the detector plane). This deviation results in an offset of 213.2 mm at the imaging isocenter in the horizontal direction (longitudinal to the treatment couch). The detector panel is shifted of 115 mm in the same direction to match the beam offset, thus allowing to scan a wider patient volume (Figure 3.4). The width of the beam in the horizontal direction is defined by inserting a collimator in front of the X-ray source beam. The M20

collimator used for lung cancer patients provides a nominal irradiated length at isocenter of 276.7 mm. The deviation of the kV imaging axis in the medium FOV configuration yields a shift in the horizontal direction of the imager centre, which represents the projection of the isocenter on the detector plane. The horizontal and vertical coordinates of the imager centre in the considered system arrangement are equal to 400 and 256 pixel, respectively.

A calibration procedure of the Elekta SynergyS cone-beam CT system is required to compensate for the flexing that occurs in the rotating arms of the imaging system at various gantry angles, due to the weight of the mechanical components, such as the accelerator gantry, the kV X-ray tube and the flat-panel detector [Sharpe 2006]. Ideally, the X-ray beams generated by the kV source during gantry rotation intersect exactly in a single point, represented by the radiation isocenter. In practice, flexions and torsions of the rotating mechanical systems cause perturbations in the circular scanning trajectory, that can produces a significant distortion in the reconstructed image. The actual isocenter of the imaging system defines a spherical volume that contains the beam intersection points for all gantry orientations. The flexmap represents the discrepancy between the radiation isocenter and the mechanical isocenter of the CBCT imaging system as a function of the projection angle. When integrating the XVI cone-beam imaging system with the treatment unit, flexmaps need to be calibrated to allow the alignment of the CBCT reconstruction centre with the radiation isocenter. These maps are initially computed after the system installation and periodically checked during maintenance operations [Lehmann 2007]. A ball-bearing (BB) phantom placed at the radiation isocenter is used for the geometrical calibration of the cone-beam CT system, computing and storing for each gantry angle the shifts of the projected BB in the two image dimensions. Long-term reproducibility and stability of the flexmaps have been reported [Bissonnette 2007]. An example of flexmaps measured for the Elekta SynergyS system installed at CLB is shown in Figure 3.5.

All the parameters associated to a specific cone-beam CT scan are saved in an acquisition log file by the XVI software. For each acquired CBCT image, the log file specifies the corresponding projection angle, the relative acquisition timestamp with respect to the beginning of the cone-beam CT scan and the flexmap corrections along the horizontal and vertical image dimensions. The temporal and angular information are used for the synchronization of the CBCT and optical surface acquisitions, described in Section 2.6.1. The horizontal and vertical flexmaps are instead considered when comparing the real and estimated tumour trajectories, as will be further explained in the next paragraphs.

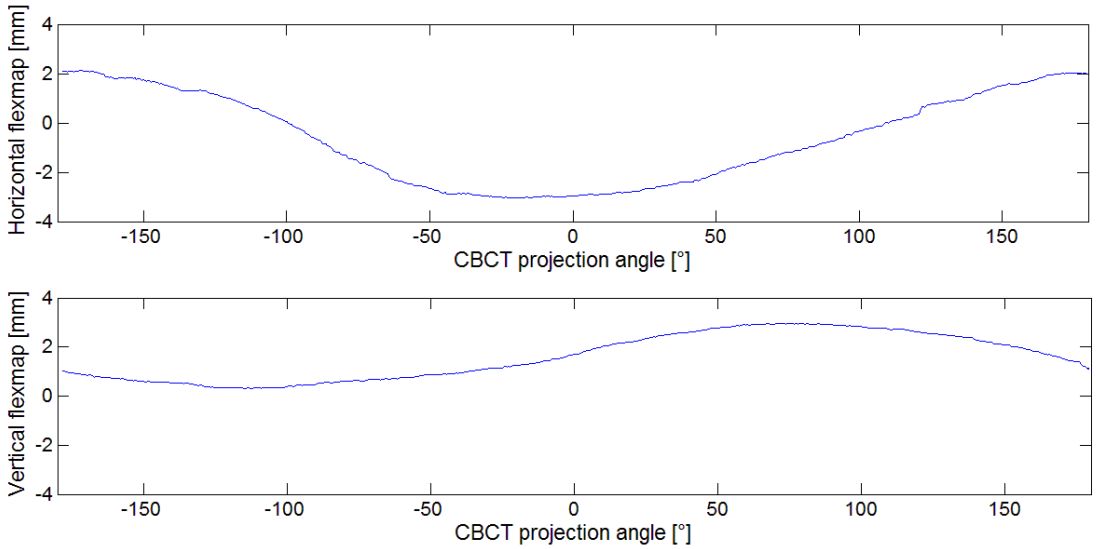


Figure 3.5. Example of flexmaps along the horizontal and vertical image directions associated to each projection angle of a cone-beam CT acquisition performed with the Elekta SynergyS system.

The volumetric CBCT dataset is reconstructed from the acquired sequence of planar projections using the XVI software, which implements the filtered back-projection based on the Feldkamp-Davis-Kress (FDK) algorithm [Feldkamp 1984]. The FDK method can be divided into three different steps: pre-weighting of the projection data, filtering and 3D back-projection. The raw images are initially corrected for a weighting function to compensate for the increased attenuation of photons at higher cone angles, due to the longer path-length travelled in the tissues. The pre-weighting factor is dependent on the cone angle. The weighted projection data are filtered raw-by-raw by multiplying a ramp-filter in the Fourier domain, in order to remove the statistical noise in the reconstructed volume. The filtered images are finally back-projected in 3D, thus determining the summed intensities at each voxel of the CBCT volume. Geometric non-idealities due to flexions and torsions in gantry rotation are corrected by the XVI software in the reconstruction process by using the calibration flexmaps [Sharpe 2006]. The CBCT projections are shifted according to the flexmap corrections estimated at the corresponding angle prior to 3D image reconstruction. As described in Section 2.2.2, the 3D volume reconstructed from the acquired cone-beam projections is used to update the tumour baseline in the 4D CT motion model according to the daily variations of patient's anatomy. The mid-position CT image is non-rigidly registered to the CBCT volume, and the obtained vector field is applied to the planning GTV contours to compute the tumour baseline adapted at each specific treatment fraction. The baseline corrections computed for the lung cancer patients included in the database are depicted in Table 3.3, showing the difference along the three spatial directions between the tumour centroid in the planning mid-position CT image and in the fraction-specific updated baseline volume. Substantial shifts in tumour baseline were measured in

most patients, with maximum absolute values of 6.4, 8.8 and 9.8 mm along the ML, AP and SI directions, respectively. As reported in Table 3.3, the 3D distances of the tumour centroids acquired during planning and treatment phases range between 1.4 and 12.1 mm, with a mean value of 5.6 mm. Since the CBCT scan is acquired before setup corrections, the estimated baseline shifts can be mainly associated to errors in patient positioning, but inter-fraction changes in muscle tone or stomach filling can also be involved [Shirato 2004].

Table 3.3. Shifts between tumour centroids extracted from the mid-position CT planning image and from the fraction-specific updated baseline along ML, AP and SI directions. The 3D tumour distance between planning and treatment phases is reported in the last column of the table for each lung cancer patients included in the database.

Patient	ML baseline shift (mm)	AP baseline shift (mm)	SI baseline shift (mm)	3D baseline shift (mm)
P1	-6.4	-4.5	-1.4	7.9
P2	0.2	-6.6	1.1	6.7
P3	0.2	0.9	1.1	1.4
P4	-2.0	1.3	-0.9	2.6
P5	-3.6	-8.8	-7.5	12.1
P6	0.2	3.1	-1.0	3.3
P7	-4.2	-3.7	-9.8	11.3

3.1.3.3 VisionRT surface imaging

For each patient of the database, the external thoraco-abdominal surface was continuously acquired during the corresponding cone-beam CT scan by using the VisionRT optical system (Vision RT, London, UK). As described in Section 1.4.2.2, the VisionRT is a markerless video-based surface imaging device based on non-invasive stereo-photogrammetric techniques for the reconstruction of a dynamic 3D surface model. The VisionRT installed at the Centre Léon Bérard and used for the acquisition of the surface data included in the database is depicted in Figure 3.1. The two treatment pods are suspended from the room ceiling at a symmetrical position with respect to the treatment couch. The stereo-cameras are calibrated with respect to the isocentric system of the treatment unit using a specific calibration plate with a printed grid of black spots (Figure 3.6). The calibration plate is positioned on the treatment couch relative to the light-field cross wire of the room lasers at three specific heights of the treatment couch, corresponding to -20, 0 and +20 cm from the radiation isocenter. At each couch position, a photo of the calibration grid is acquired by each stereo-camera and the grid points are automatically identified by the VisionRT calibration

software. This procedure allows to compute the transformation matrix that correlates the camera coordinate frame of the VisionRT system to the isocentric coordinate frame of the accelerator. System calibration is performed monthly and verified on a daily basis according to vendor recommendations. The daily process of calibration verification is similar to the full monthly calibration, but only a single plate position corresponding to a zero couch height is required.

Figure 3.6. Photo of the planar grid of black spots used for the calibration of the VisionRT surface imaging system.

According to the protocol defined for surface data acquisition, the GateCT program was selected as the more appropriate operation mode for the present application, since it allows the continuous dynamic scanning of the whole patient surface. Dynamic surface models are provided also by the AlignRT mode at a very low acquisition frame rate, since this program performs in real-time the 3D surface reconstruction by merging data from both imaging pods. Experimental tests performed with the AlignRT system showed frequencies lower than 1 frame per second (fps) when the entire body surface was acquired, whereas for small region of interest with approximately 10 cm side the frame rate proved to be about 5 fps. The GateCT program allows to achieve higher scanning frequencies without selecting any ROI, since surface data are acquired from a single imaging pod and the 3D reconstruction is performed off-line, as explained in Section 1.4.2.2. Table 3.4 lists the mean frame rates of GateCT surface acquisitions obtained for the patients of the collected database. The number of surface points included in the thoraco-abdominal region of interest selected for the extraction of the breathing surrogate signal is also reported in the Table. The number of surface points for the collected patients varied from 871 to 1892, with a mean value of 1340. The surface frame rates ranged between 7.9 and 9.1 fps, depending on the size of the

surface portion visible from the VisionRT stereo-cameras. These surface scanning frequencies were considered adequate for the present application, since they are about twice the frame rate of CBCT projection acquisition.

Table 3.4. Mean value of surface acquisition frequencies obtained with the GateCT program for the patients in the database, and number of points included in the thoraco-abdominal ROI selected for the extraction of the breathing surrogate signal.

Patient	Mean frame rate of surface acquisition (fps)	Number of points in the thoraco-abdominal ROI
P1	7.9	1892
P2	8.2	1173
P3	7.9	1191
P4	8.5	1174
P5	8.1	1572
P6	9.1	871
P7	8.2	1507

The application of the GateCT program for surface motion acquisition involved several disadvantages due to the use of a single imaging pod. An example of patient surface and corresponding thoraco-abdominal ROI acquired with the GateCT operation mode is shown in Figure 2.10. Almost the entire body surface can be captured with the single-pod GateCT scan, excluding about one third of the surface located on the side opposite to the imaging pod. The use of one pod implied also problems related to patient's surface occlusion due to gantry rotation during the CBCT scan. When the rotating structures of the treatment unit (such as the gantry, the CBCT X-ray source and the flat-panel detector) pass in front of the VisionRT imaging pod, the visibility of the patient's external surface is completely hindered for the optical stereo-cameras. The imaging pod located on the right-side of the treatment couch was selected for surface data acquisition since it minimizes the period of occlusion time, given the clock-wise rotation of the gantry and CBCT unit. This allows about 35 seconds of surface scan without occlusions, corresponding to the acquisition of about 300 surface frames and 200 cone-beam images. The number of breathing cycles that can be evaluated in the non-occluded temporal window was considered sufficient for the present tumour tracking study.

Partial occlusions of the patient's external surface during a GateCT scan can also be caused by the inappropriate positioning of the horizontal bar of the HexaPOD system, as shown in Figure

3.7a. If the markerized bar is not attached low enough to the treatment couch, it can cover the abdominal portion of the patient's skin. Surface occlusions due to the lateral handles of the stereotactic body frame (Figure 3.7b) and to the trousers worn by the patients (Figure 3.7c) can also occur. A too or less intense surface illumination may hinder the visibility of the speckle pattern projected on the patient surface, resulting in holes in the reconstructed meshes. The setting of the VisionRT system parameters was optimized according to the specific acquisition conditions. In particular, the value of the "skin tone" parameter (fair, mid or dark), which controls the exposure of the speckle projector, was set on a patient-specific basis, to minimize the presence of surface holes due to inadequate surface illumination. At the end of each GateCT surface scan, the "Review and export" functionality of the GateCT program was used to export the associated acquisition file, containing for each captured frame the information about the relative timestamp, namely the time elapsed since the start of the surface scan. As explained in the Section 2.6.1, this file was used for the synchronization between the cone-beam CT scan and the optical surface acquisition.

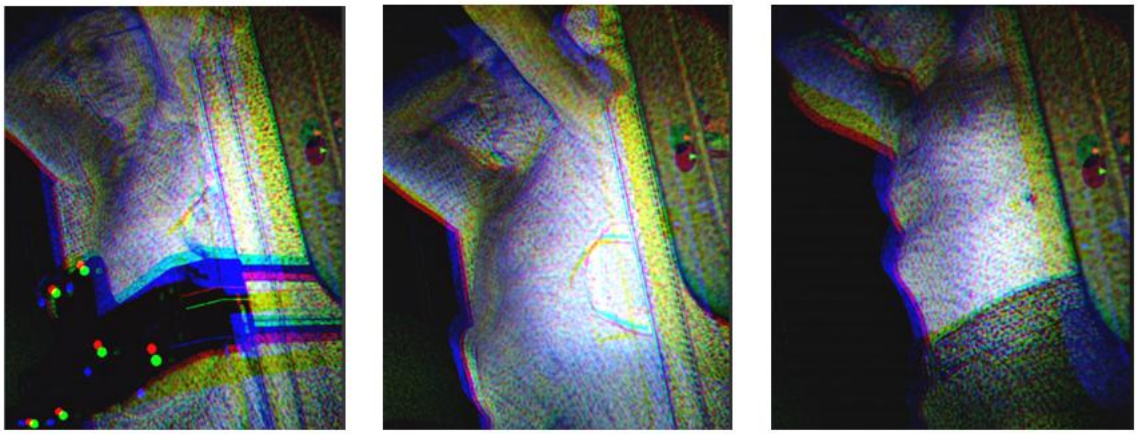


Figure 3.7. Photo acquired by the optical cameras of the VisionRT surface imaging system, in which partial surface occlusion is visible due to the HexaPOD horizontal bar (Panel a), to the lateral handle of the stereotactic body frame (Panel b) and to the trousers worn by the patient (Panel c).

3.2 CBCT/VisionRT synchronization

3.2.1 Temporal accuracy

The temporal accuracy of the image-based method for the synchronization of the cone-beam CT scan and the VisionRT optical surface acquisition, described in Section 2.6.1, was evaluated on a moving phantom. As shown in Figure 3.8, the test phantom is composed by a planar horizontal surface moved by a microprocessor-controlled stepping-motor along a sinusoidal trajectory, thus

simulating the breathing motion of the patient's thoraco-abdominal surface. The employed phantom provides the possibility to adjust the amplitude and frequency parameters of its sinusoidal motion. For the experimental testing of the developed synchronization method, 8 simultaneous CBCT/VisionRT acquisitions of the moving phantom were performed. The amplitude of the sinusoidal motion was set to 15 mm, while the frequency was varied from 0.75 to 1 Hz, as reported in Table 3.5. For each synchronization test, 50 CBCT images were acquired and used for the accuracy evaluation, corresponding to about 9 seconds of phantom motion acquisition.

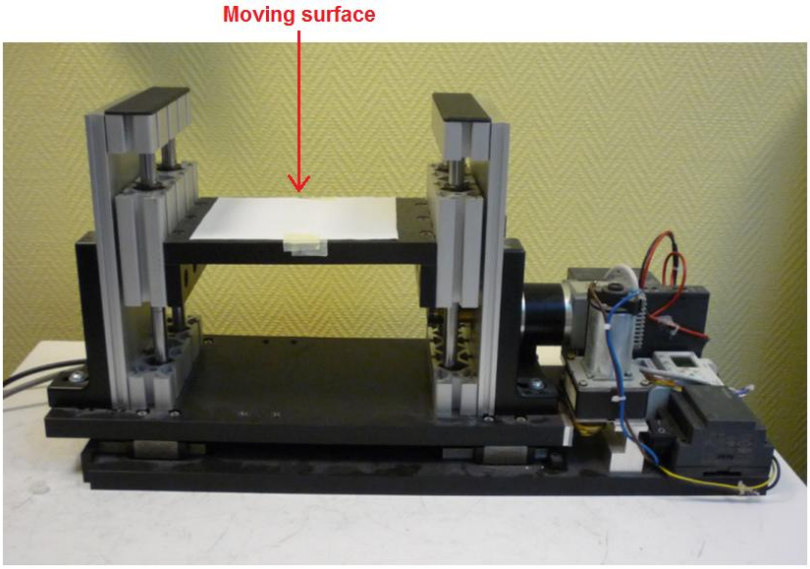


Figure 3.8. Photo of the phantom used for the experimental evaluation of the temporal accuracy of the developed image-based synchronization method. The phantom's moving surface is depicted in the figure.

The phantom motion signals extracted from the CBCT images and from the VisionRT optical surfaces were compared after applying the developed synchronization technique. A radiopaque marker was placed approximately at the centre of the phantom's horizontal surface, thus allowing the tracking of the surface motion in cone-beam CT projections (Figure 3.9). Threshold-based segmentation algorithms were applied to identified the position of the radiopaque marker in the CBCT images, obtaining the vertical trajectory of the phantom's surface. The VisionRT system in the single-pod GateCT operation mode was used to optically acquired the phantom motion. The GateCT modality provides the trajectory associated to a surface point manually selected by the user at the beginning of the optical surface scan. In particular, the program computes the mean vertical motion of the surface points reconstructed in a circular region of interest with 5 cm radius, cantered in the selected point. For the present application, the tracking point was chosen in correspondence of the radiopaque marker. For all the 8 synchronization tests, the motion signal extracted by the

GateCT program from the captured surface frames was stored, along with the relative acquisition timestamps of each surface frame.

Table 3.5. Experimental results of the phantom tests performed to evaluate the accuracy of the developed CBCT/VisionRT synchronization method. The temporal difference between the zero-crossings and the sinusoidal peaks of the phantom motion signals extracted from cone-beam images and optical surfaces are expressed as median value [25^{th} – 75^{th} percentile].

Test	Frequency of phantom motion (Hz)	CBCT starting angle difference (°)	Zero-crossings difference (msec)	Sinusoidal peaks difference (msec)
T1	0.75	0.05	-2.6 [-11.2 – 26.1]	4.4 [-2.8 – 11.6]
T2	0.85	0.07	19.5 [7.6 – 36.3]	12.9 [11.5 – 14.4]
T3	1.0	0.08	-8.1 [-12.3 – 15.8]	1.1 [-0.7 – 2.9]
T4	1.0	0.07	-2.4 [-16.1 – 10.7]	1.2 [1.1 – 1.3]
T5	1.0	0.09	-5.4 [-15.3 – 16.6]	0.8 [-0.2 – 1.8]
T6	1.0	0.09	4.7 [-21.1 – 18.8]	2.8 [0.3 – 5.4]
T7	1.0	0.04	-7.4 [-20.2 – 4.6]	-3.1 [-5.8 – -0.3]
T8	1.0	0.08	6.2 [-24.3 – 45.0]	7.8 [4.1 – 11.4]

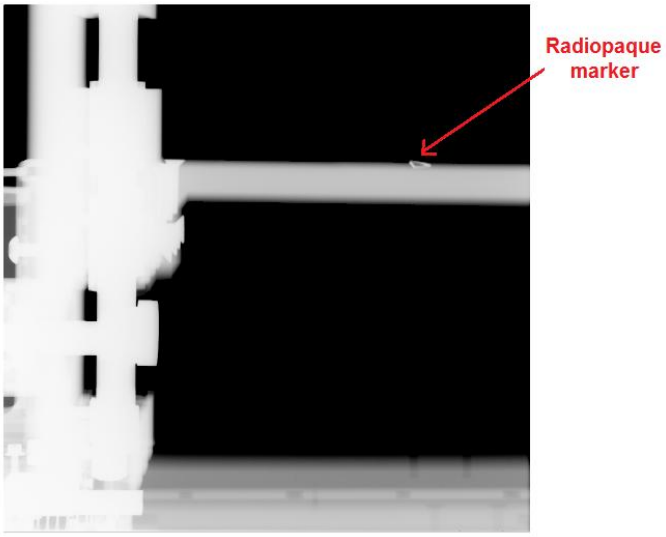


Figure 3.9. Cone-beam CT projection of the phantom employed for the synchronization tests. The radiopaque marker used for the CBCT tracking of phantom's moving surface is shown in the figure.

Two different approaches were developed to compare the phantom's surface trajectories obtained from the synchronized cone-beam CT and VisionRT optical acquisitions. The first evaluation strategy was based on the computation of the zero-crossings of the two motion signals

normalized between -1 and +1. The timestamps associated to the zero-crossings were computed as the intersection points between the zero-value horizontal axis and the motion signals, obtained by linearly interpolating the phantom's surface positions measured in the cone-beam projections or in the VisionRT frames. The estimated temporal differences between the zero-crossing timestamps of the motion signals derived from the synchronized CBCT/VisionRT acquisitions are reported in Table 3.5. The median value of the temporal differences obtained for each phantom test ranged between -8.1 and 19.5 msec. The maximum value of the inter-quartile range, calculated as the difference between the 75th and 25th percentiles of the timestamp differences, proved to be about 70 msec. The linearly fitted motion signals estimated for a phantom test are depicted in Figure 3.10a.

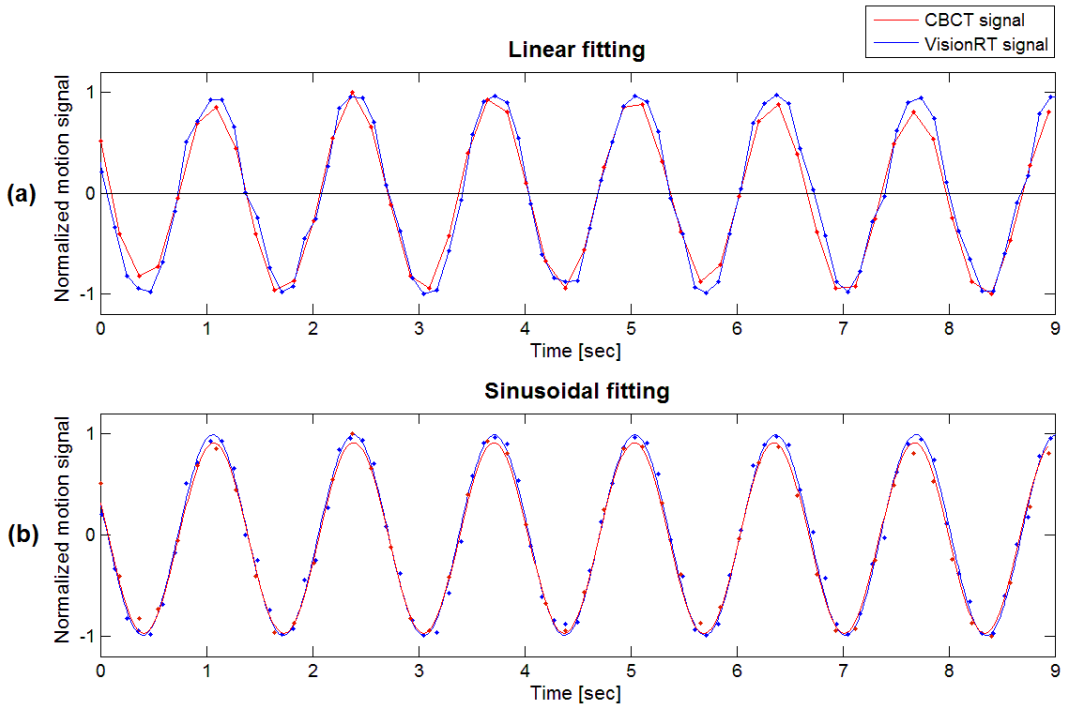


Figure 3.10. Linear (Panel a) and sinusoidal (Panel b) fitting of the phantom's motion signals estimated from the cone-beam projections and from the VisionRT surfaces for the first synchronization test (T1).

To overcome the limitation of the described evaluation strategy related to the linear approximation of the motion signals, a second approach was developed based on a sinusoidal fitting. The signals extracted from the synchronized CBCT/VisionRT scans were fitted with the following sinusoidal model using a non-linear least-squares approach:

$$y(t) = A\sin(2\pi f \cdot x(t) + \varphi), \tag{3.1}$$

where A, f and φ are the amplitude, frequency and initial phase parameters of the phantom's motion signals. The sinusoidal fitting of the motion signals obtained for a phantom test is shown in Figure 3.10b. The timestamps associated to the maximum and minimum peaks of the fitted sinusoids were

computed for both the CBCT and VisionRT scans. Table 3.5 reports the temporal differences between the estimated peaks of the cone-beam and optical signals for each synchronization test. The median value of the measured differences ranged between -3.1 and 12.9 msec, with a maximum inter-quartile range of about 15 msec.

3.2.2 Angular model accuracy

As described in Section 2.6.1, the developed image-based synchronization method used a linear angular model to correlate the cone-beam projection angle with the position of the CBCT panel identified in the captured VisionRT photos. The accuracy of the built linear model, represented in Figure 2.27a, was evaluated for all the 8 phantom tests by considering the starting angle of the cone-beam CT scan. The real value of the initial angular position of the CBCT system was derived from the acquisition file associated to the specific CBCT scan. The obtained value was compared with the angle estimated from the VisionRT photos with the linear angular model. In particular, the 10 photos captured before the starting of cone-beam CT acquisition were considered. The distances of the elliptic feature on the CBCT panel from the top-left corner of the extracted ROI was computed for all the 10 photos according to the procedure explained in Section 2.6.1. The mean value of the feature's distances was used to derive the CBCT starting angle from the model's linear polynomial. The differences between the real and estimated starting angles of the cone-beam CT scan for each of the phantom tests are reported in Table 3.5. All the measured angular differences proved to be lower than 0.1°, with a mean value of 0.07°.

The accuracy in the estimation of the cone-beam CT starting angle by using the linear angular model built for CBCT/VisionRT synchronization was assessed also for the cone-beam CT acquisitions of the lung cancer patients included in the database. This procedure allows to verify the long-term validity of the angular model, estimated several weeks before the patient's treatment fractions. Table 3.6 reports the real angular positions of the CBCT system for the 7 cone-beam CT scans of the collected database, and the respective angles estimated from the 10 VisionRT photos captured before the beginning of the cone-beam CT acquisition. The nominal CBCT starting angle is -180°, since the initial rotational position of the treatment gantry is at -180°, namely the gantry is located below the treatment couch. As reported in Table 3.6, the real initial angle of the CBCT system is lower than the nominal one, due to the mechanical limitations in the treatment gantry rotation. The angular differences measured between the real and the estimated CBCT starting angles varied from -0.1 to 0.2°, with a mean absolute value of 0.08°. The results obtained for patient data

proved to be comparable with the phantom tests, thus confirming the long-term validity of the linear angular model built for CBCT/VisionRT synchronization.

Table 3.6. Comparison between the real rotational positions of the CBCT system at the beginning of the cone-beam CT scans included in the database and the angles estimated from the linear model built for CBCT/VisionRT synchronization.

Patient	Real CBCT starting angle (°)	Estimated CBCT starting angle (°)	Angular difference (°)
P1	-178.07	-178.10	0.03
P2	-178.53	-178.43	-0.10
P3	-178.50	-178.59	0.09
P4	-178.60	-178.65	0.05
P5	-178.27	-178.47	0.20
P6	-178.38	-178.42	0.04
P7	-178.61	-178.67	0.06

3.3 Contrast-enhanced tumour detection

The accuracy of the developed contrast enhancement technique for markerless tracking of lung tumours in 2D projection images, described in Section 2.5.1, was experimentally evaluated on the cone-beam CT acquisitions collected for the 7 lung cancer patients of the database. Each CBCT dataset includes about 650 two-dimensional images, acquired over a full 360° of gantry rotation. This allowed to test the performance of the developed tracking method as a function of the projection angle. Exemplificative results for three selected patients are reported in Figure 3.11, depicting the difference between the original cone-beam projections simply adjusted for intensity and the corresponding contrast-enhanced images, with the overlying contours of the matched tumour templates. These results qualitatively highlight the ability of the proposed method in improving tumour region visibility at different projection angles and reducing the overlying effect of the nearby anatomical structures, such as the spine in the first example and the rib cage in the second one.

The tumour tracking performance of the developed technique was quantitatively evaluated in terms of target identification rate, tumour contrast-to-noise ratio (CNR) and 2D/3D tracking accuracy. The effective gain of contrast enhancement on the above listed indices was assessed by comparing the obtained results with the outcomes of the state-of-the-art techniques based on template cross-correlation with no contrast enhancement. The state-of-the-art methods were

reproduced according to [Hugo 2010, Lewis 2010], implementing the cross-correlation between the original cone-beam images and the digital reconstructed radiographs generated from the MIP CT volume at the corresponding angles. Differently from the proposed contrast-enhanced approach, in the state-of-the-art technique the reference templates include not only the tumour but also the whole surrounding anatomy. The tumour region of interests are extracted from both the CBCT projections and the simulated DRR templates, using the same ROI selection approach developed for the contrast enhancement technique. The grey-level values of the extracted ROIs are scaled between 0 and 1 using a linear mapping, thus adjusting and conforming image intensities before performing the cross-correlation.

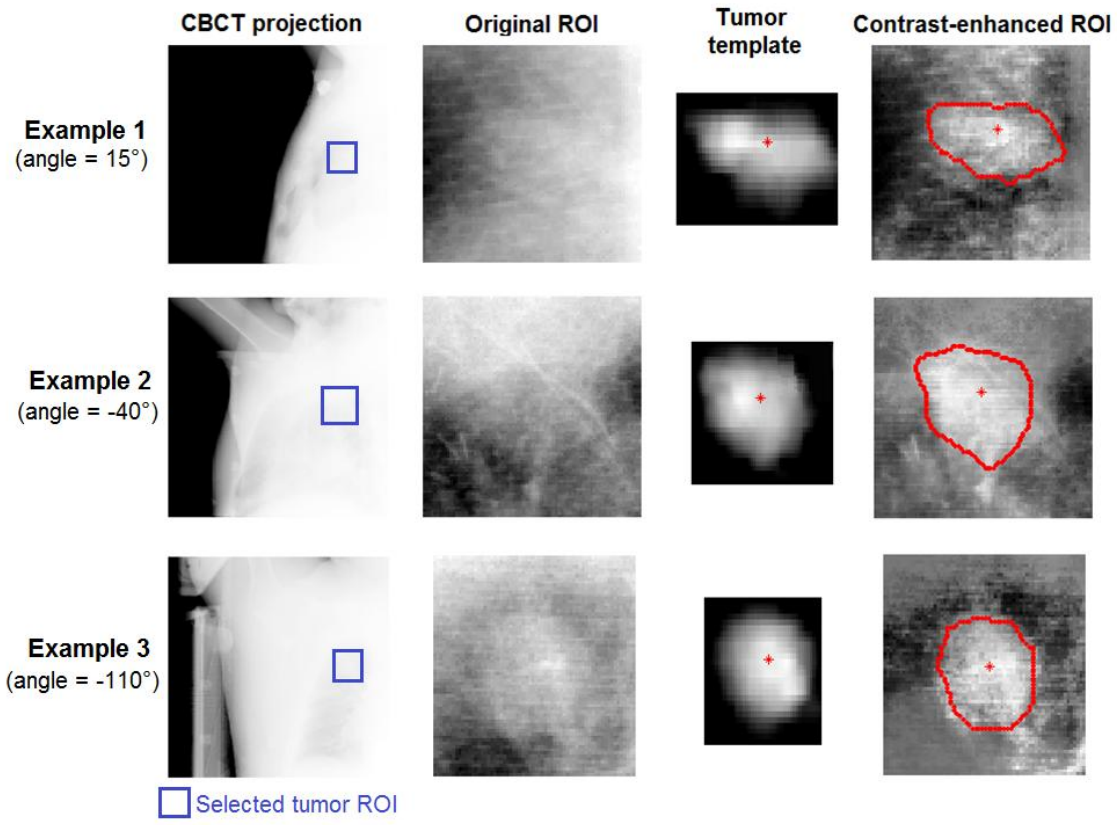


Figure 3.11. Representative examples of the original intensity-adjusted cone-beam projections and corresponding contrast-enhanced images with the overlying contours of the matched tumour templates for three selected lung cancer patients of the database at different CBCT rotational angles.

3.3.1 Target identification rate

For both novel and conventional tracking techniques, target identification rates were estimated as the percentage of CBCT projections over the entire image dataset in which tumour was successfully detected. The distance from the expected two-dimensional coordinates of the 3D

CBCT tumour centroid, as well as the distance from the tumour positions identified in the close projections of the CBCT scan sequence, were considered to assess correct target recognition. The expected tumour centroid positions on each cone-beam image were obtained by projecting at the corresponding angle the tumour centroid segmented in the reconstructed CBCT volume, as further explained below. Visual inspection was required for the estimation of the target identification rates, to exclude wrong detections biased by the surrounding anatomical structures. The visual verification of correct target identification was performed using a specifically implemented graphical user interface (GUI). The GUI shows the tumour ROIs extracted from both the original and the contrast-enhanced CBCT projections, overlaying the tumour contours in the position identified with the state-of-the-art and the proposed novel tracking methods (Figure 3.12). For each rotational position, tumour contours are extracted by segmenting the GTV volume projected at the corresponding angle.

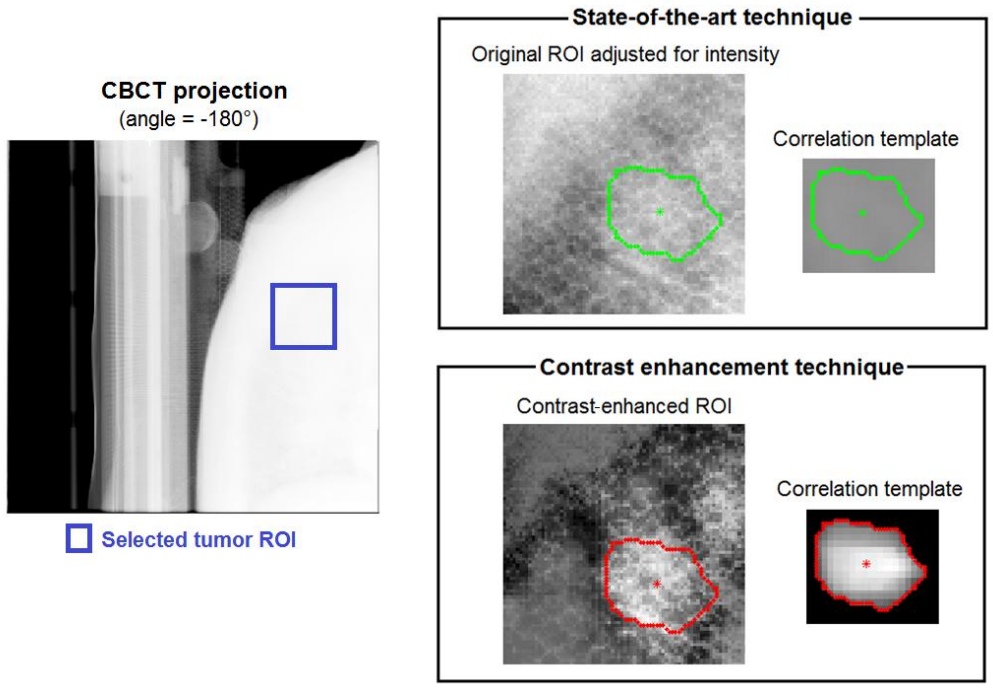


Figure 3.12. Tumour detection performed with the state-of-the-art tracking technique and with the proposed contrast enhancement approach in a cone-beam CT lateral projection of a lung cancer patient included in the database.

Figure 3.13 shows the comparative results related to the target identification rates obtained by applying the proposed contrast-enhanced method and the state-of-the-art tracking techniques for the 7 lung cancer patients included in the database. The percentage of tumour detection was on average 63±5% in the contrast-enhanced CBCT images and 52±7% when conventional detection was applied on original projections with intensity adjustment and no contrast enhancement. The

gains in tumour detection obtained with the proposed tracking method range between 6% and 37%, with a mean value of 24%. This corresponds on average to 76 projections per cone-beam CT scan in which tumour could be detected only by applying the contrast enhancement. The relationship between the target identification rate and the CBCT projection angle was also investigated. Figure 3.14 shows the rate of target recognition successfully achieved with the developed contrast-enhanced method over the entire patient database as a function of the angular position of the cone-beam CT system. The projection angles were divided in two different groups, represented by the lateral and frontal angles, according to the following conditions:

$$angle_{lat} = [-180°,-135°] \cup [-45°,+45°] \cup [+135°,+180°] , \tag{3.2}$$

$$angle_{front} = [-135°,-45°] \cup [+45°,135°]. \tag{3.3}$$

As depicted in Figure 3.14, CBCT images acquired at lateral projection angles featured a tumour detection rate of 46±18%, while frontal angles showed higher rates of about 80±16%. According to our experience, there was no patient case in which the lesion could be detected in the totality of CBCT images. Especially on lateral projections, the overlap of the surrounding anatomical structures completely hindered tumour visibility, thus excluding any detection possibility both with conventional and contrast-enhancement methods. The limited target identification rates could also be related to the small size of the patient's lung lesions included in the selected database, showing a tumour volume down to 0.5 cm^3 (Table 3.1). Lung lesions with reduced sizes are more difficult to visualize on CBCT projections, since they feature lower image contrast. However, patients with higher tumour sizes were not available for this study, since the clinical protocol applied at CLB sets a specific limit for lung lesion's diameter in case of stereotactic body radiotherapy treatments.

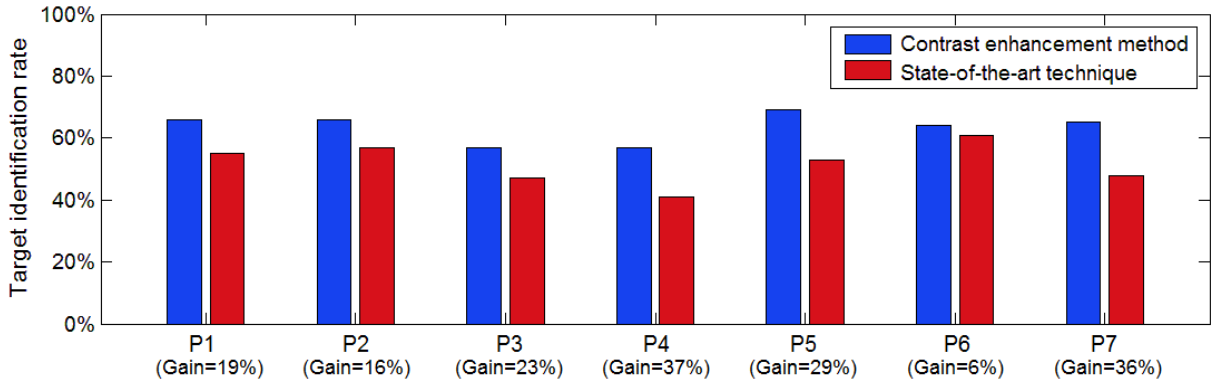


Figure 3.13. Comparison of target identification rates obtained with the proposed contrast enhancement method and with the state-of-the-art tracking techniques for the 7 lung cancer patients of the database. The percentage gains in tumour detection on contrast-enhanced images are reported for each patient.

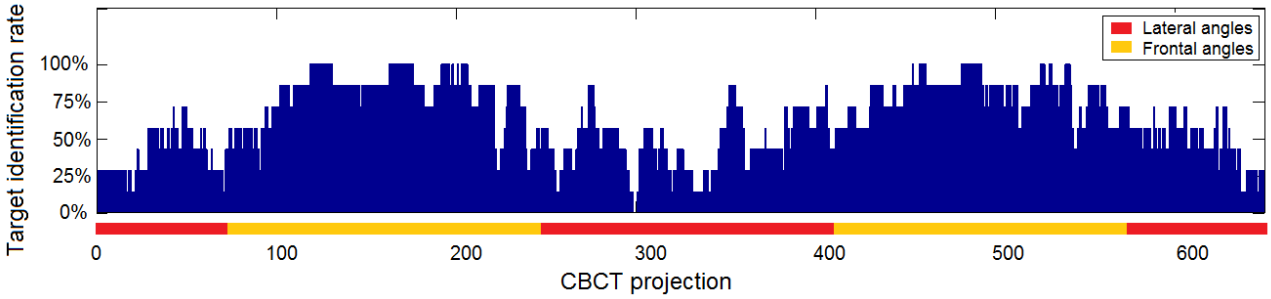


Figure 3.14. Percentage of successful tumour detections computed over the entire patient database for each CBCT projection angle.

3.3.2 Tumour contrast-to-noise ratio

The ability of the proposed contrast-enhanced tracking method in increasing tumour region visibility in CBCT projection was assessed through the contrast-to-noise ratio metric. As depicted in the following equation, the tumour CNR was computed as the difference in the mean pixel intensity between the region *T* included in the tumour contours identified on the original and contrast-enhanced projections and the surrounding region of background *B* in the selected image ROI:

$$CNR = \frac{\text{mean}(T) - \text{mean}(B)}{\text{std}(B)}. \tag{3.4}$$

The intensity difference was divided by the image noise, computed as the standard deviation of the background region.

Table 3.7 reports the comparison of the contrast-to-noise ratio computed on the CBCT projections in which tumour was successfully detected with the contrast-enhanced method and with the state-of-the-art tracking technique. The median CNR values estimated for the 7 lung cancer patients vary from 1.3 to 1.7 in the contrast-enhanced images, while a CNR range between 0.5 and 0.8 was measured when applying the conventional method on the original projections. As depicted in Table 3.7, the median value of the CNR percentage gains proved to be 136%, varying from 71% to 194%. According to the obtained results, the proposed contrast enhancement technique allows a significant increase of the image contrast between the lung lesions and the surrounding anatomical structures. The non-parametric Wilcoxon rank-sum test was applied to verify the statistically significantly difference between the CNR values computed with the novel contrast-enhanced method and with the conventional state-of-the-art tracking technique, obtaining a p-value lower than 0.001.

Table 3.7. Tumour contrast-to-noise ratio computed for the CBCT projections included in the database, in which lung tumour was successfully detected with the contrast enhancement method and with the state-of-the-art tracking technique. The CNR percentage gains in contrast-enhanced images is also reported for each patient. CNR results are expressed as median value [25^{th} – 75^{th} percentile].

Patient	CNR with contrast-enhanced method	CNR with state-of-the-art technique	CNR gain (%)
P1	1.6 [1.3 – 1.9]	0.6 [0.5 – 1.0]	136 [74 – 204]
P2	1.6 [1.3 – 1.9]	0.6 [0.3 – 0.9]	166 [83 – 404]
P3	1.3 [1.1 – 1.5]	0.8 [0.6 – 1.0]	71 [30 – 126]
P4	1.7 [1.4 – 2.0]	0.5 [0.4 – 0.7]	228 [148 – 417]
P5	1.5 [1.3 – 1.7]	0.5 [0.3 – 0.8]	194 [98 – 458]
P6	1.3 [1.1 – 1.6]	0.6 [0.4 – 0.9]	116 [55 – 231]
P7	1.5 [1.2 – 1.8]	0.6 [0.4 – 0.9]	133 [88 – 207]

3.3.3 2D/3D tracking accuracy

Several research groups have evaluated the accuracy of tumour tracking applications by means of comparison with the manually defined ground-truth, obtained by visually locating lung lesions on kV or MV images [Lewis 2010, Rottmann 2010]. In our patient cases, however, tumours were not visible on most CBCT projections, thus hindering the possibility of manual identification of target ground-truth. This limitation pushes towards the application of strategies for the assessment of tracking discrepancies with respect to the corresponding information on tumour localization coming from the 3D CBCT volume. Two alternative approaches were developed for the quantitative evaluation of the tumour tracking accuracy in 2D and 3D provided by the proposed contrast enhancement method. Both approaches require lung lesion segmentation in the 3D cone-beam CT volume reconstructed from the analyzed 2D projections. Since in our patient database tumour contours are clinically available only on the MIP CT image, we performed contour propagation by non-rigidly registering the MIP with the CBCT volume. B-spline deformable registration algorithms were applied [Sharp 2010], using the normalized mutual information as similarity metric. The ITV contours delineated on the MIP image were propagated according to the computed deformation vector field. The propagated contours were used to extract the binary tumour volume and the respective centroid from the reconstructed 3D CBCT image. The tumour volume extracted from the cone-beam CT represents the mean target position weighted over all breathing cycles performed by the patient during the acquisition of CBCT projections. Since the developed

strategies for the evaluation of the contrast-enhanced tracking method are based on the reconstructed CBCT tumour volume, we could assess only the accuracy in the estimation of the mean tumour position on the 2D projections and in the 3D space, called respectively 2D and 3D tracking accuracy. Since the ground-truth information on the tumour trajectory was not available, the accuracy in the measurement of the range and shape of tumour motion could not be evaluated.

For the assessment of the 2D tracking accuracy, the tumour centroid extracted from the cone-beam CT volume was projected at each angular position of the CBCT scan, thus obtaining the reference baselines of target trajectories along the horizontal and vertical image directions. These baselines represent the mean positions assumed by the tumour centroid during the whole CBCT scan acquisition. Flexmap corrections are applied to the obtained baselines to compensate for the discrepancy between the radiation isocenter and the mechanical isocenter of the CBCT system. The reference target baselines were compared to the mean tumour positions estimated from the 2D target trajectories identified with the proposed tracking method on the contrast-enhanced projections. A moving average filter with a window size of 100 samples was applied to the tumour trajectories detected on the CBCT images, thus obtaining the estimated target baselines. The 2D tracking error was defined as the median difference between the reference and the estimated respiratory tumour baselines computed along the horizontal and vertical image directions.

An example of tumour centroid trajectories identified in the contrast-enhanced CBCT image sequence of a patient in the database is shown in Figure 3.15. Breathing motion is evident both along the horizontal and vertical image dimensions. Horizontal dimension corresponds to the projection of AP and ML components of tumour motion, while vertical dimension represents tumour motion projection along SI direction. The reference baselines obtained from the projection of the 3D CBCT tumour centroid are also depicted in Figure 3.15. A sinusoidal trend is superimposed to the target trajectory and baseline along the horizontal direction, since the AP and ML components of tumour motion are projected along a dimension orthogonal to the axis of rotation of the gantry, which changes continuously with the projection angle. For both image directions, the 2D tumour trajectories and baselines are also influenced by flexmaps.

Table 3.8 reports the 2D tracking accuracy in estimating the horizontal and vertical tumour trajectories on cone-beam CT images with the contrast enhancement method and with the state-of-the-art techniques for each lung cancer patient of the database. The median [25^{th} – 75^{th} percentile] value of the absolute 2D tracking errors on the contrast-enhanced CBCT projections was 1.9 [1.5 – 2.7] mm along the horizontal image direction and 1.9 [1.1 – 2.7] mm along the vertical image direction. When state-of-the-art tracking techniques were applied, 2D tracking errors of 2.2 [1.9 – 2.7] mm and 1.7 [1.2 – 2.9] mm were found for the horizontal and vertical tumour trajectories,

respectively. The correlation between the 2D tracking accuracy and the amplitude of respiratory tumour motion for the two image dimensions was also investigated. Breathing amplitude was computed as the difference between the 95th and 5th percentiles of the distribution of the horizontal and vertical tumour positions identified in the contrast-enhanced images, subtracting baseline trends obtained through moving average filters. As depicted in Table 3.8, the range of the obtained respiratory amplitudes was 7.0–14.3 mm for the horizontal image dimension and 4.1–12.0 mm for the vertical direction. For 5 out of 7 patients, the projection of AP and ML components of respiratory tumour motion showed higher amplitudes than the projected target motion along SI direction. No significant correlation was found between tumour motion amplitudes and absolute 2D tracking errors, obtaining Pearson linear coefficients of 0.6 and 0.1 for the horizontal and vertical image directions, respectively.

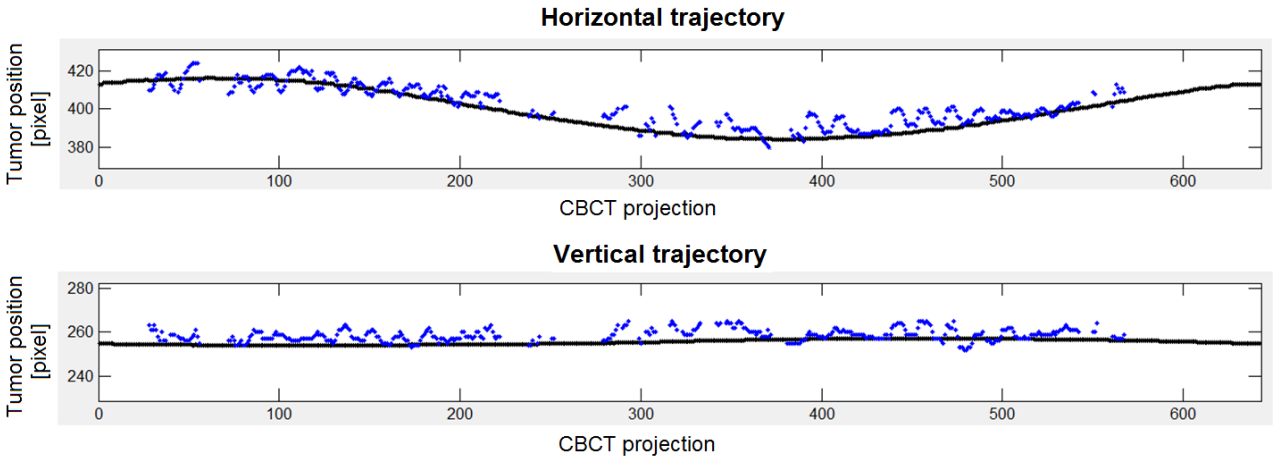


Figure 3.15. Two-dimensional tracking results obtained for the cone-beam projections of patient P1. Blue points represent the horizontal and vertical trajectories of the tumour centroid identified with the proposed tracking technique on the contrast-enhanced images. Black points corresponds to the 2D projected coordinates of the lesion centroid segmented in the 3D CBCT volume.

The 3D tracking accuracy was evaluated by comparing the 3D coordinates of the tumour centroid extracted from the 3D CBCT volume and the 3D tumour position reconstructed from the 2D trajectories obtained from contrast-enhanced CBCT projections. For each successful target detection, a projection image was created keeping only the tumour template in the identified position (Figure 3.16a). For each projection angle, the tumour template is shifted in both directions according to the corresponding flexmap corrections. The Plastimatch implementation of the Feldkamp-Kress-Davis filtered back-projection algorithm [Sharp 2009] was applied to reconstruct the 3D tumour volume from the created projection images. The geometry parameters of the

employed CBCT system that need to be specified to the FDK program include the source-to axis distance, the source-to-imager distance and the spatial resolution of the reconstructed volume. For the Elekta Synergy system, the output 3D CBCT image has a size of 410×410×264 pixel along the ML, AP and SI directions, respectively, with a resolution of 1 mm/pixel in each direction. An example of tumour volume reconstructed with the FDK program from the 2D tumour template projections is shown in Figure 3.16b. Threshold-based segmentation algorithms were applied to extract the tumour centroid from the reconstructed volume.

Table 3.8. Amplitude of respiratory tumour motion along the horizontal and vertical image dimensions and 2D tracking errors in estimating target trajectory on CBCT projections with contrast-enhanced and state-of-the-art methods.

Patient	**Breathing amplitude (mm)**		**2D tracking errors for contrast-enhanced method (mm)**		**2D tracking errors for state-of-the-art technique (mm)**	
	Horizontal direction	**Vertical direction**	**Horizontal direction**	**Vertical direction**	**Horizontal direction**	**Vertical direction**
P1	9.8	6.1	-1.9	-2.6	-1.9	-1.8
P2	14.3	9.9	-1.4	-1.1	-2.2	1.7
P3	8.7	4.1	-1.6	-3.0	-1.9	-2.0
P4	7.0	5.5	-2.6	-0.4	-2.8	1.2
P5	10.1	12.0	-2.7	-1.9	-2.6	2.0
P6	11.0	6.0	-1.0	1.1	-1.7	1.2
P7	8.0	8.6	-2.7	-2.7	-3.5	0.4

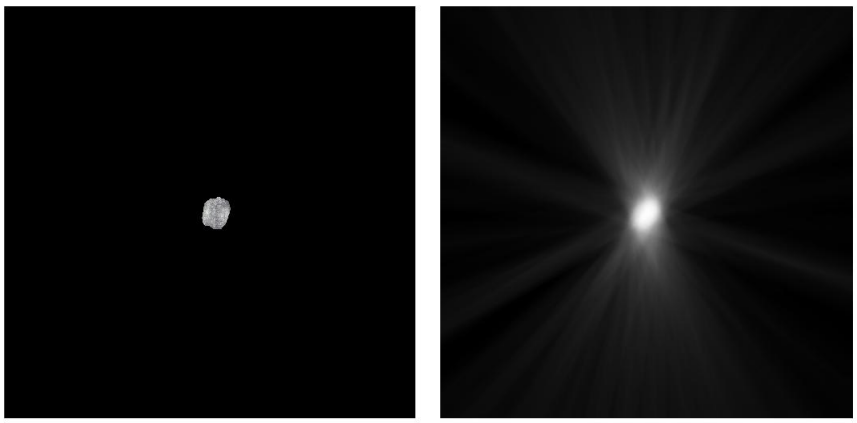


Figure 3.16. 3D tumour volume (Panel b) reconstructed with the FDK filtered back-projection algorithm from the 2D images including only the tumour template in the position identified with the contrast-enhanced method (Panel a).

The 3D tracking error was computed as the difference between the 3D coordinates of the tumour centroid extracted from the 3D CBCT image and the 3D position of the target centroid obtained from the FDK reconstructed volume. Table 3.9 lists for each patient of the database the results of the 3D tracking accuracy in estimating tumour centroid position using the contrast-enhanced and the state-of-the-art tracking techniques. The contrast enhancement method showed median 3D errors of 0.5 [0.3 – 0.7] mm, 0.3 [0.2 – 1.15] mm and 0.4 [0.15 – 1.1] mm along the ML, AP and SI directions. 3D tumour tracking errors obtained with conventional methods were significantly higher, measuring 1.3 [0.7 – 1.5] mm, 1.5 [1.3 – 2.5] mm and 2.3 [1.3 – 2.7] mm for ML, AP and SI coordinates, respectively. The results concerning the 2D and 3D tracking accuracy obtained for conventional state-of-the-art methods are comparable with the tracking performance reported in literature by previously published works [Lewis 2010, Hugo 2010]. 3D errors computed with and without enhancement proved to be statistically significantly different for AP and SI tumour coordinates, as obtained with a Wilcoxon paired signed-rank test (p-value lower than 0.05). These results demonstrate a relevant improvement in tracking accuracy with the proposed contrast enhancement approach with respect to state-of-the-art techniques.

Table 3.9. 3D tracking errors in estimating tumour centroid position with contrast-enhanced and state-of-the-art methods along ML, AP and SI directions.

Patient	3D tracking errors for contrast-enhanced method (mm)			3D tracking errors for state-of-the-art techniques (mm)		
	ML direction	AP direction	SI direction	ML direction	AP direction	SI direction
P1	0.7	0.1	0.2	1.9	-0.07	1
P2	0.5	-1.6	1.4	1.3	-3.7	2.7
P3	0.04	0.3	0.05	0.4	1.5	0.5
P4	0.6	-0.7	2.2	-1.3	1.1	2.9
P5	1.9	-0.08	0.1	1	1.4	2.6
P6	0.5	-0.3	0.8	1.6	1.7	1.6
P7	-0.1	1.7	0.4	-0.4	3.2	2.3

3.4 Deformable surface registration

The accuracy of the developed deformable registration method, described in Section 2.3.1, in tracking external breathing motion from markerless surface imaging data was assessed in 5 male healthy volunteers. The VisionRT system available at the Centre Léon Bérard was employed for

surface data acquisition. The registration error was estimated based on the known position of multiple surface control landmarks. As shown in Figure 3.17a, 11 black star-shaped markers were placed on different portions of the thorax and abdomen of the subjects. The identification of the ground-truth position of the surface control points exploited the texturing capabilities of the VisionRT system, providing the grey-level representation of the reconstructed meshes (Figure 3.17b). As described in Section 1.4.2.2, textured information is available only for static mesh acquisitions performed with the AlignRT program. Due to this technical limitation of the employed optical system, registration errors could be assessed only on statically acquired surfaces. The designed experimental testing of the developed registration algorithm covered however the whole range of surface motion. In fact, for each of the 5 subjects lying in supine position on the treatment couch, the thoraco-abdominal surface was statically acquired with the AlignRT system at three different respiratory phases, sampling the breathing cycle at the extreme and intermediate positions. The gating functionality provided by the AlignRT operation mode was exploited to gate surface acquisition at the desired phases of the respiratory cycle, represented by the inhale peak, exhale peak and an arbitrarily chosen intermediate position. Two subjects repeated the experiments twice, resulting in 7 full datasets for method evaluation. For each acquisition session, the three surfaces corresponding to different respiratory phases were pair-wise registered, acting alternatively as source and target mesh. This allows to increase the number of registrations available as validation data set, leading to 6 registrations per subject experiment, for a total of 42 registrations.

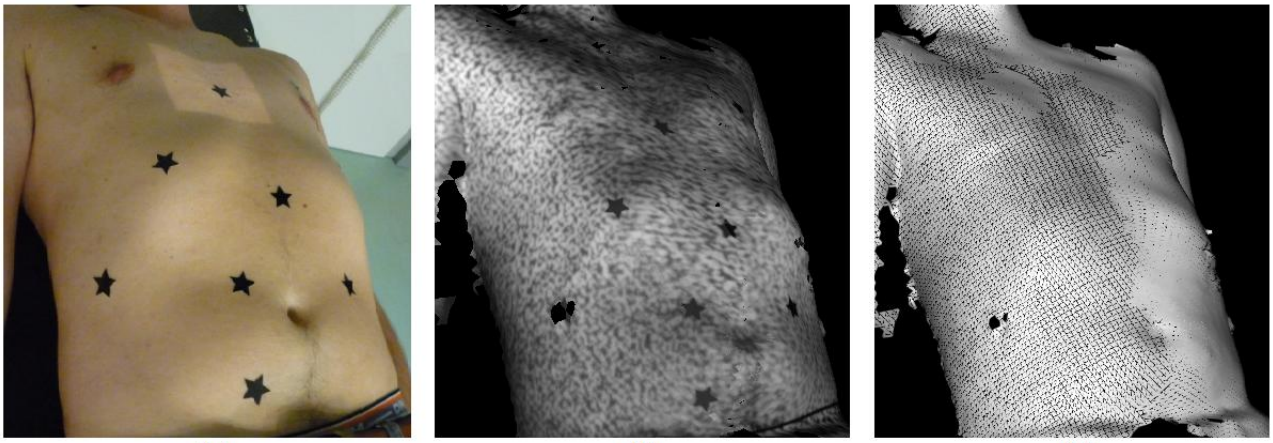


Figure 3.17. Star-shaped black markers placed on the thoraco-abdominal surface of a test subject (Panel a) and corresponding textured mesh acquired with the AlignRT optical system, showing the structured light pattern projected on the subject surface (Panel b). The integrated surface model obtained by merging data from both imaging pods is also shown in the figure (Panel c).

An example of the triangulated mesh captured by the AlignRT system for one of the test volunteers is shown in Figure 3.17c. In the selected operation mode, the integrated surface model is obtained by merging data captured from both imaging pods, thus covering a full 180° of field of view. The accuracy of the implemented registration algorithm was quantified both in terms of residual surface distance between the deformed source and the corresponding target mesh, and in terms of landmark localization error based on the known position of the surface control points. The obtained results were compared to the performance of rigid surface registration based on a standard ICP algorithm [Besl 1992], to assess the effective gain of using a deformable approach for registering thoraco-abdominal surfaces influenced by breathing motion. All computations regarding surface processing and accuracy evaluation were based on the Insight Toolkit (ITK) [Ibáñez 2005] and Visualization Toolkit (VTK) [Schroeder 2006] libraries.

The computational cost of the implemented deformable registration algorithm was evaluated for all the acquisition datasets, assessing the correlation with the mean number of vertices in the subject surfaces reconstructed with the optical system. For each acquisition session, Table 3.10 reports the number of surface vertices averaged over the three acquired breathing phases and the mean time associated to a single iteration of the registration algorithm in the outer loop. The computational performance of the developed application was evaluated using a 2.53 GHz Intel Core 2 Duo processor. The CPU time required per outer iteration ranged between 6.1 and 11.9 sec, and proved to be linearly correlated with the number of vertices in the registered meshes. The Pearson correlation coefficient between the computational time and the vertex number proved to be 0.9, with a p-value lower than 0.01.

Table 3.10. Computational performance of the implemented deformable surface registration algorithm. The acquisitions of the two subjects who performed the test session twice are identified by A1−A2 and M1−M2.

Acquisition	Mean number of points per surface	Mean computational time per outer loop iteration (sec)
A1	9922	9.5
A2	9896	8.8
JA	9820	8.8
JO	10854	10.1
MA	14457	11.9
M1	8427	6.1
M2	8410	6.8

The most time-consuming step of the implemented deformable surface registration algorithm is the Cholesky factorization of the linear system matrix, which is performed at each outer iteration, when the point correspondence weights are updated. This limitation derives from the adopted global regularization approach, which involves that each vertex has an influence on every other vertex of the mesh, resulting in a very large linear system that is costly to solve. Multi-resolution strategies and/or a more local regularization approach, using for example a fixed-size vertex neighbourhood or GPU-based techniques, could potentially provide a significant reduction of the computational cost of the algorithm, which does not currently provide real-time performance.

3.4.1 Residual surface distance

The accuracy of the implemented deformable registration algorithm is evaluated by means of residual surface distance, which represents the Euclidean distance between each point of the deformed source mesh and the closest point on the corresponding target surface. The residual surface distance was computed over the entire dataset of 42 pairs of meshes acquired from the 5 subjects at three different breathing phase. All points of the thorax and abdomen of the subjects reconstructed by the optical surface imaging system are taken into account for the estimation of the residual surface distance. A critical aspect that may negatively influence the accuracy of surface registration applied to the thoraco-abdominal area is the low spatial frequency of the considered surfaces due to the lack of geometrical features. Mesh registration was therefore performed without selecting any region of interest, since the decrease of spatial constraints due to the smaller patch size can lead to surface fitting degeneration, thus reducing the registration accuracy. For each subject acquisition, the 95^{th} percentile of the residual surface distance after performing deformable registration is reported in Table 3.11. Surface distances in the initial condition and after applying rigid mesh registration are also listed in the Table. Before applying mesh registration, the thoraco-abdominal surface distances ranged between 3.9 and 7.5 mm. The 95^{th} percentile of the residual distances obtained with the rigid registration algorithm varied from 3.4 to 8.1 mm. After applying the developed deformable approach, the residual surface distances did not exceed 1.20 mm.

Figure 3.18 shows the surface overlap computed for different breathing phases of a test subject before and after deformable registration. The greater surface distances are localized along the central line that merges the surface models reconstructed by the two imaging pods of the AlignRT optical system. The merging process generates disconnected components near the mesh boundaries, resulting in higher registration errors. Optimized surface stitching algorithms can be used to increase mesh connectivity, thus improving the regularization step of the registration.

Table 3.11. 95th percentile of the residual surface distances obtained with the implemented deformable registration algorithm, compared to the performance of rigid mesh registration and to the initial surface distances.

Acquisition	Residual surface distance (mm)		
	Initial condition	Rigid registration	Deformable registration
A1	7.5	8.1	0.8
A2	7.2	6.1	0.8
JA	4.1	3.4	0.5
JO	5.6	3.7	0.4
MA	3.9	3.4	1.1
M1	6.7	4.4	1.2
M2	7.5	4.7	1.1

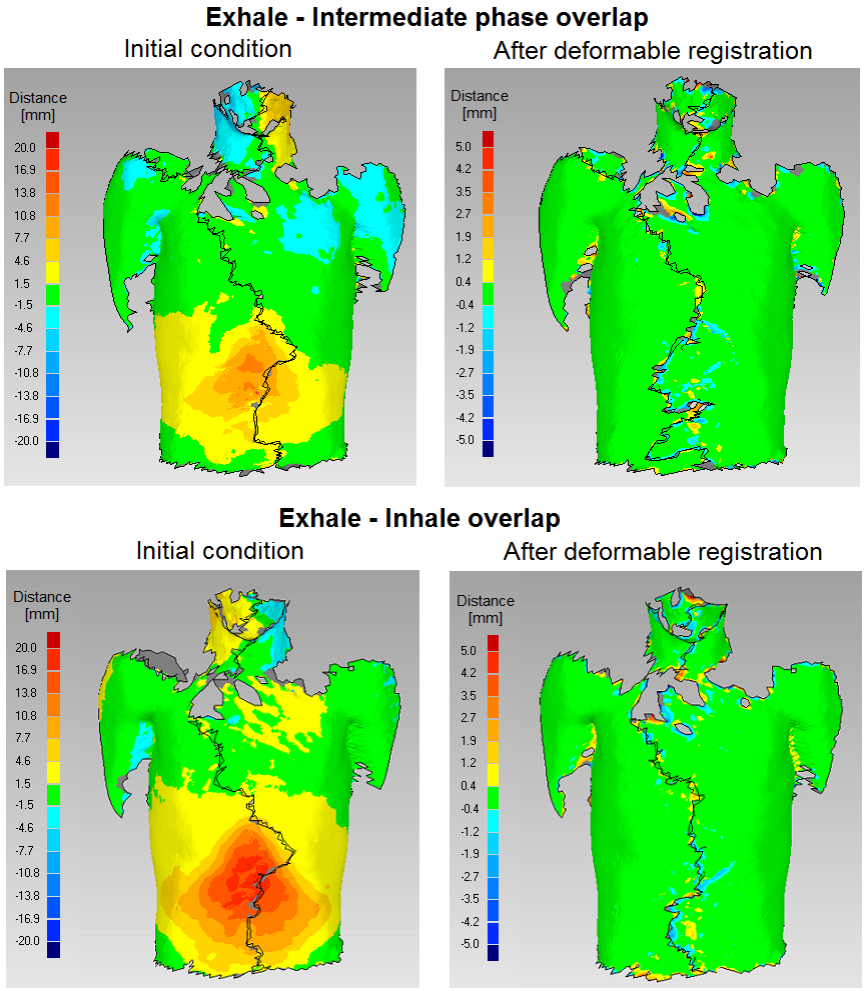


Figure 3.18. Surface overlap at different breathing phases before and after applying the developed deformable registration algorithm. The colour-based intensities represent the surface distance in the initial condition and after deformable registration. As can be noticed in the figure, the developed registration algorithm is also able to recover head rotations.

3.4.2 Landmark localization error

The registration error based on the known position of multiple surface control landmarks was also estimated to quantify the accuracy of the developed deformable mesh registration algorithm. The texture information acquired with the AlignRT system was used for the visualization of the control points and for the identification of the respective ground-truth positions. Each vertex of the triangular faces in the acquired textured meshes is parameterized by 3D spatial coordinates defining the surface topology and by 2D coordinates for texture information. The 2D texture coordinates specify for each surface vertex the position of the corresponding 2D point in the texture photo captured by the optical camera. The grey-level intensity of the specified 2D point is used to colour the corresponding 3D surface vertex in the textured mesh. As shown in Figure 3.17a, the control markers placed on the thorax and abdomen of the subjects are represented by black stars with five tips. The 3D positions of the marker's tips were manually selected on the acquired textured surfaces, and the centroids of each marker were computed by averaging the corresponding tips. Depending on the marker location for the different subjects, the number of visible marker's tips that could be identified in all three breathing phases varied from 26 to 40, with an average number of about 36. The corresponding number of marker centroids ranged between 8 for two subject acquisitions, 10 for three acquisitions, and 11 for two acquisitions. For instance, for the textured surface depicted in Figure 3.17b, the marker on the right-part of the abdomen could not be tracked due to the presence of holes in the reconstructed mesh.

The measured errors in corresponding landmark localization may be partly related to the intrinsic imprecision of the optical imaging system in 3D surface reconstruction and to the inaccuracies in the identification of marker positions on the textured meshes. An analysis of the intra- and inter-operator variabilities in the manual selection of the tips of the star-shaped markers placed on the thoraco-abdominal surface of the test subjects was performed. In order to assess the intra-operator variability, the manual clicking of the marker's tips on a reference textured surface was repeated 10 times by a single operator. The intra-operator variability was computed as the difference between the tip coordinates selected by the operator in the first test and the tip positions identified in the other 9 tests. Figure 3.19a shows the resulting intra-operator differences related to the repeated clicking of the marker's tips on the reference textured surface. The variability in the identification of the marker's centroids, computed by averaging the coordinates of the corresponding tips, is also depicted (Figure 3.19b). The 75^{th} percentile of the intra-operator differences obtained for all tests proved to be 0.4 mm for tip selection and 0.2 mm for centroid identification.

The inter-operator variability was instead assessed by comparing the positions of the marker's tips identified by 5 different operators on the three surfaces acquired for a test subject. The coordinates of the tips identified by the first operator on the three tested surfaces were taken as reference to compute the inter-operator differences. The obtained variability in the selection of marker's tips and centroid performed by the different operators is reported in Figure 3.19c and in Figure 3.19d, respectively. The 75th percentile of the resulting differences was 0.8 mm in case of tip clicking and 0.5 mm in case of centroid identification. The centroid coordinates defined by the 5 operators did not prove to be significantly different, by performing a Kruskal-Wallis non-parametric test with a p-value equal to 1. According to the obtained results, we used the positions of the centroids of the star-shaped surface markers to evaluate the accuracy of the developed registration algorithm, since they feature lower intra- and inter-operator variabilities with respect to the localization of the marker's tips. Due to the high reproducibility in centroid identification between different operators, the registration errors were computed by considering as ground-truth the centroid positions identified by a single operator.

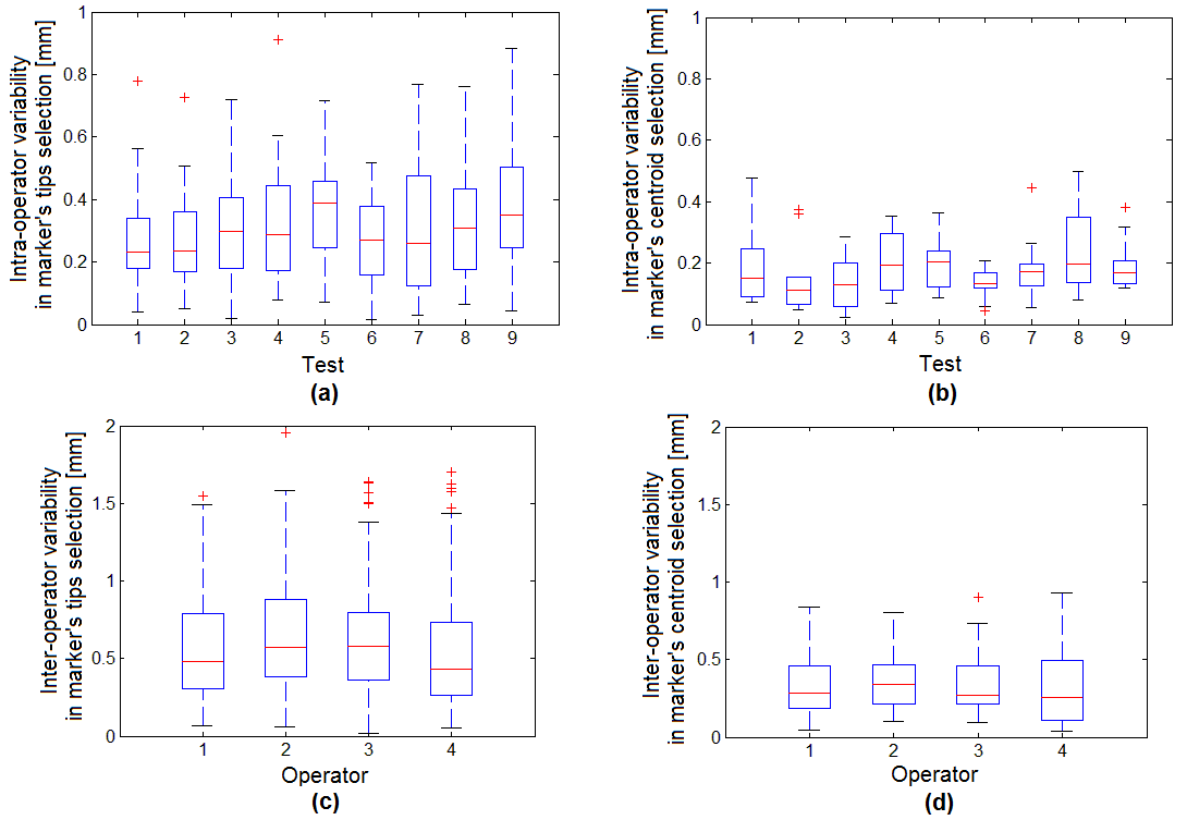


Figure 3.19. Intra- and inter-operator variability in the manual selection of the star-shaped marker tips on the textured surfaces (Panels a and c) and in the identification of the star centroids from the corresponding tips (Panels b and d). Boxplots depict the 25th, 50th (median) and 75th percentiles of the difference distribution. Whiskers extend from both sides of the box up to 100% of the quartile range and symbols (+) denote outliers.

The landmark localization error was computed as the difference between the displacement of the control points selected on the textured surfaces, considered as ground-truth, and the motion estimated through the developed deformable registration algorithm. In order to account for the discretization of the acquired meshes, the landmarks selected on the source and target surfaces were projected on the respective meshes. The real marker motion was computed from the projected points, whereas the estimated marker motion was obtained by considering the three neighbouring vertices of the source mesh triangle that includes the projected landmarks. The displacements of the neighbouring vertices derived from deformable mesh registration were linearly interpolated to estimate the landmark motion. A sensitivity analysis of the main structural parameters of the developed deformable surface registration algorithm was performed by analyzing the landmark localization errors computed over the entire acquisition dataset as a function of the following variables:

- start and final values of the stiffness factor α;
- number of iteration steps in the outer loop;
- threshold δ for the convergence of the inner iterative process.

The results of the sensitivity analysis for the considered parameters are reported in Figure 3.20. The landmark localization error, expressed as the median error in the identification of the surface marker position for all test subjects, was computed by varying one parameter at a time. As depicted in Figure 3.20, the registration errors resulting from the sensitivity analysis ranged between 1.61 and 1.69 mm. None of the tested parameters of the registration algorithm showed a significant influence on the overall registration accuracy. Slightly worse results were obtained for low numbers of outer iterations and for high final values of the stiffness factor, both contributing to a reduced ability of the surface to deform, namely to recover the breathing-induced deformation.

The outcomes of the developed algorithm corresponding to the lowest landmark localization errors were analyzed in detail. These results were obtained by using 25 outer iteration steps, with a stiffness factor ranging from 150 to 1 and a convergence threshold of 1/100. The distance between the surface control points in the initial condition and after applying rigid and deformable registration algorithms are depicted in Table 3.12 for all subject acquisitions. Before applying surface registration, the median value of the marker distances ranged between 2.5 and 5.3 mm. The median errors in landmark localization obtained with the rigid registration algorithm varied from 2.1 to 6.4 mm, with a median inter-quartile range of 3.4 mm. After applying the developed deformable approach, the median registration errors did not exceed 2.3 mm, with a median inter-quartile range of 1.6 mm. According to the obtained results, the performance of the developed deformable approach was superior to rigid procedures. For all except one of the subject

acquisitions, rigid surface registration yielded an increase in the distance between corresponding landmarks on the source and target meshes with respect to the initial condition (Table 3.12), although the overall surface distance was slightly reduced (Table 3.11). Rigid registration did not prove to be adequate when dealing with thoraco-abdominal surfaces, since it does not take into account for the deformation due to breathing motion.

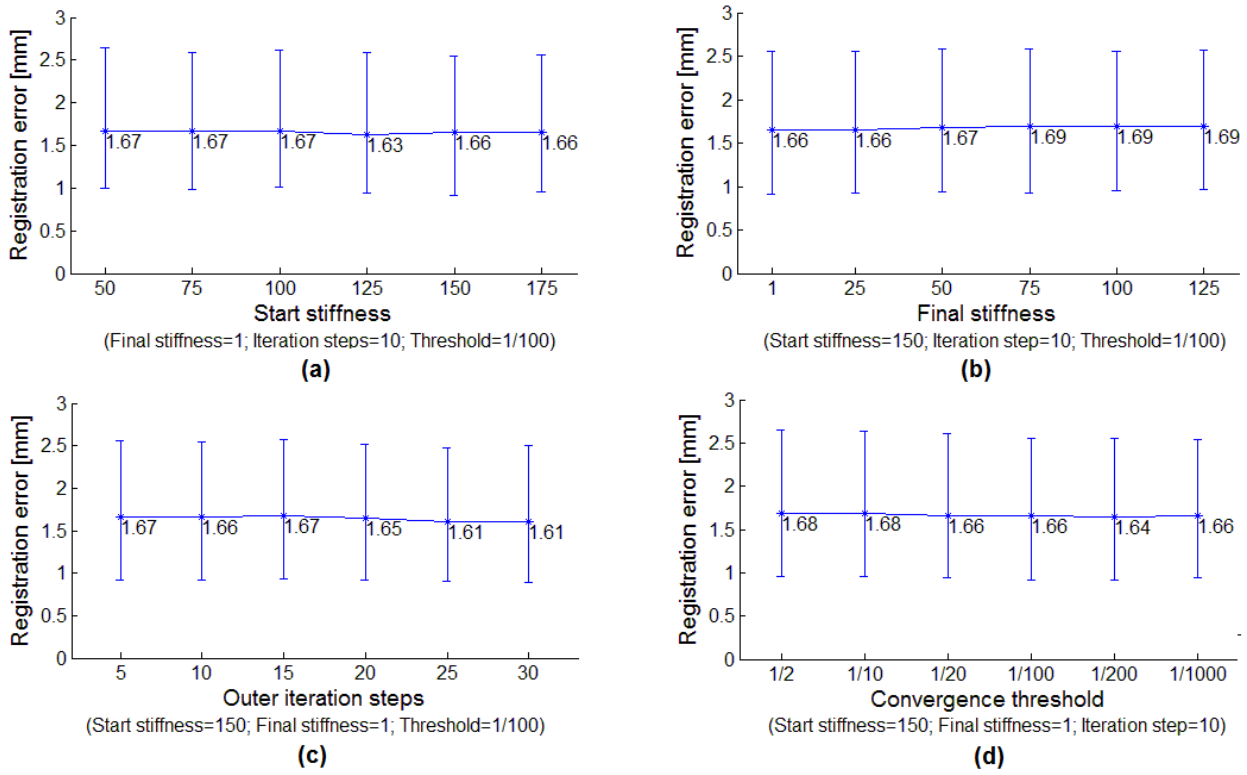


Figure 3.20. Registration errors (median ± inter-quartile range) in the identification of the marker centroids for the entire dataset of 42 registrations, as a function of the start stiffness factor (Panel a), final stiffness factor (Panel b), number of iterations in the outer loop (Panel c) and convergence threshold for the inner iteration process (Panel d).

The performance of the developed surface registration algorithm was separately evaluated for the markers located in the abdominal and thoracic regions, that were manually distinguished using the costal margin as separation line. The spatial variability of the registration accuracy in the thoraco-abdominal surfaces is illustrated in Figure 3.21. The error in landmark localization introduced by the deformable registration algorithm is separately plotted for the control points located in the thoracic and abdominal surface regions. The median value of the registration errors computed for all control points over the entire acquisition dataset was 1.6 mm, whereas it varied from 2.1 mm considering only the thoracic markers to 1.1 mm for the landmarks in the abdominal region. The error distributions estimated for the thoracic and abdominal control points proved to be statistically significantly different, by performing a Wilcoxon rank-sum test with a p-value lower than 0.01.

Table 3.12. Landmark localization errors for the implemented deformable registration algorithm, compared to the performance of rigid mesh registration and to the initial marker distances. Results are expressed as median value [25th – 75th percentile].

Acquisition	Landmark localization error (mm)		
	Initial condition	**Rigid registration**	**Deformable registration**
A1	5.3 [4.2 – 8.8]	6.4 [4.8 – 10.1]	2.1 [1.3 – 3.4]
A2	3.8 [3.0 – 7.0]	5.7 [3.9 – 7.3]	2.3 [1.6 – 3.3]
JA	2.8 [2.1 – 4.5]	2.9 [2.1 – 4.4]	1.4 [0.9 – 2.1]
JO	3.4 [1.4 – 5.2]	3.7 [3.0 – 5.4]	1.0 [0.7 – 1.9]
MA	2.5 [1.8 – 3.2]	2.1 [1.3 – 3.7]	1.4 [0.8 – 2.4]
M1	3.3 [2.1 – 4.9]	4.1 [2.4 – 6.3]	1.5 [0.9 – 2.2]
M2	4.1 [3.0 – 5.4]	4.2 [3.0 – 6.7]	1.8 [0.7 – 2.6]

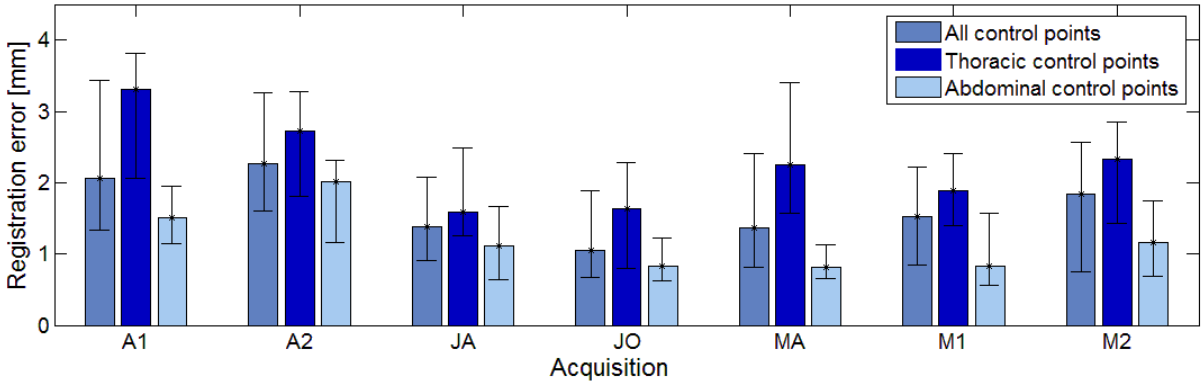


Figure 3.21. Spatial variability of the registration errors (median ± inter-quartile range), separately estimated for markers located in the thoracic and in the abdominal surface regions.

The accuracy of the implemented registration algorithm was evaluated as a function of the anatomical direction. The registration error components for each spatial direction were estimated, and the correlation between the overall registration accuracy and the direction of marker motion was established. The median error in the localization of all control points measured 0.7, 1.1 and 0.3 mm in the ML, AP an SI directions, respectively. Figure 3.22 shows the registration errors computed along each anatomical direction as a percentage of the total summed error, computed by summing all three error components. The SI component was predominant (59%) for the localization error of the thoracic markers, whereas the landmarks in the abdominal region featured a higher registration error in the ML direction (43%). The percentage error in the AP direction proved to be lower than 20% for all control points. Finally, a significant linear correlation was found between the

total registration error and the marker motion in the SI direction, obtaining a Pearson correlation coefficient of 0.7 with a p-value lower than 0.01. The correlation results between the registration accuracy and the motion direction for a representative subject's acquisition are shown in Figure 3.23.

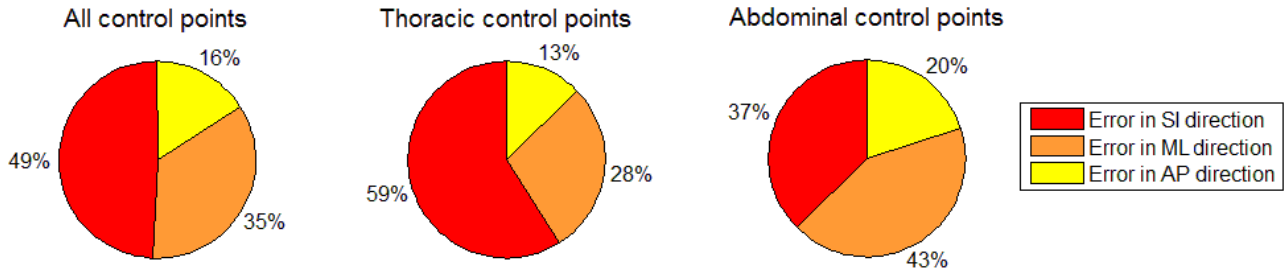


Figure 3.22. Localization errors for thoracic and abdominal control points along different anatomical directions, expressed as a percentage of the summed error.

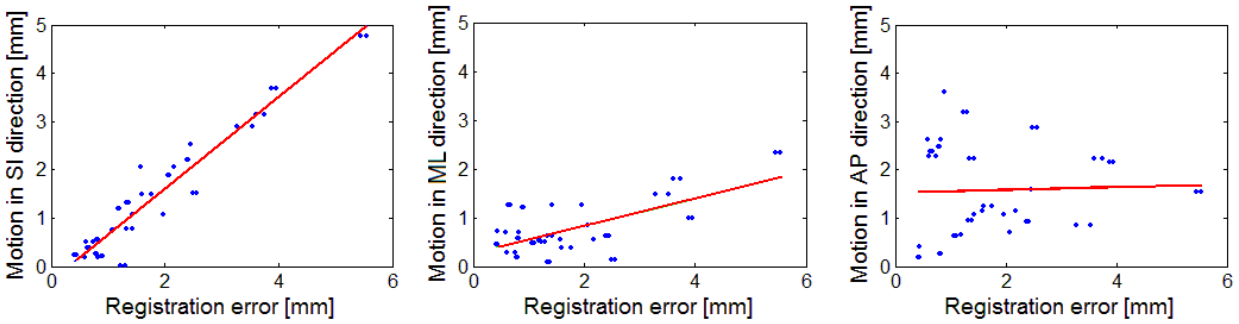


Figure 3.23. Registration errors computed for subject acquisition MA as a function of the marker motion in the three different spatial directions. The linear regression line is overlaid in red to the sample data.

The detailed analysis of the registration performance as a function of marker positions revealed that localization errors are significantly different for the thoracic and abdominal regions, featuring a higher correlation with the surface motion in the SI direction. As depicted in Figure 3.24, the deformable registration algorithm can correctly estimate the AP motion of the abdominal surface portion, but mostly fails to recover the thoracic breathing motion in the SI direction. The limited ability of the developed registration approach in estimating the SI surface motion can be partly explained by the acquisition procedure of the employed surface imaging system. Since surface motion along the SI and ML axes is normal to the direction of the projected pattern of structured light, it causes the relative slipping of the pattern with respect to the underlying surface. Displacements in SI and ML directions are consequently more difficult to capture, as confirmed by the separate error analysis for each anatomical direction (Figure 3.22). As can be noticed in figure 3.24, breathing surface motion features larger amplitude in the abdominal region along the AP

direction, whereas the thoracic motion component along the SI direction is limited. We can therefore suggest that the reduced accuracy of the implemented registration algorithm in recovering SI surface motion does not represent a critical aspect for clinical applications.

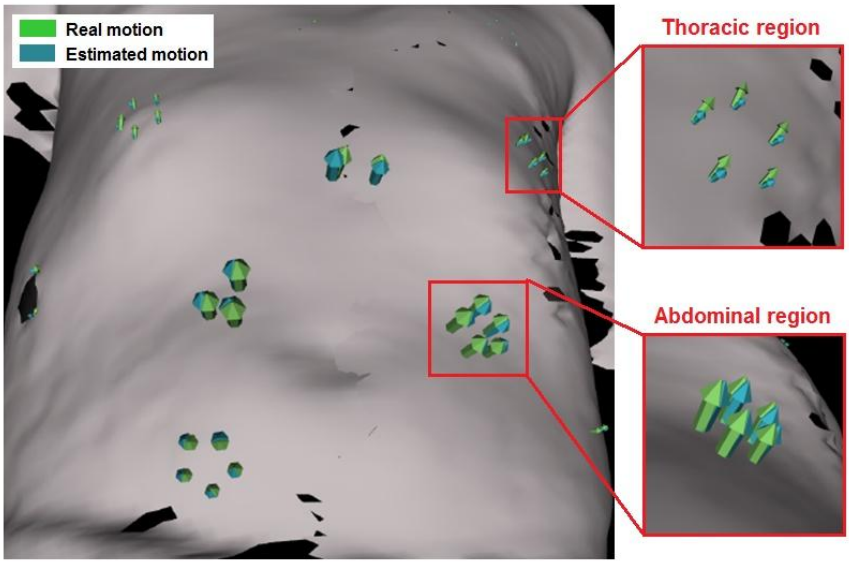


Figure 3.24. Comparison between the actual motion of the thoracic and abdominal surface regions, and the displacement estimated through deformable mesh registration.

3.4.3 Multi-dimensional breathing motion tracking

The developed deformable registration algorithm was also applied to the dynamic external surfaces of a lung cancer patient included in the database, in order to investigate the method performance in tracking the multi-dimensional breathing motion from markerless optical surface acquisitions. As explained in Section 3.1.3.3, the single-pod GateCT modality of the VisionRT system was used to continuously acquire the thoraco-abdominal surface of the selected patient. The proposed deformable registration algorithm was applied to obtain the spatial correspondence between an arbitrarily chosen reference surface, represented by the first mesh of the acquired sequence, and the following surfaces. The estimated 3D trajectories of the surface points included in the thoraco-abdominal region of interest were used to derive a multi-dimensional breathing signal. The respiratory surface motion along the three spatial directions was obtained by averaging the individual coordinates of the extracted corresponding surface points. Principal component analysis was applied for each direction and the first principal component scores associated to each surface point were used to evaluate the surface regions that mostly contribute to the respiratory motion extracted along the different directions.

Figure 3.25 shows the multi-dimensional breathing signal extracted from the dynamic thoraco-abdominal surfaces of the selected patient using the implemented deformable surface registration algorithm. The signals depicted in the lower panels of the figure represent the mean surface motion in the three spatial directions, obtained by averaging the SI, ML and AP coordinates of the corresponding surface points generated by the algorithm. The contribution of each surface point to the respiratory signals in the different anatomical directions is represented in the upper panels of the figure. This contribution is computed by projecting in the principal component space the surface point motion along the different directions. Figure 3.25 demonstrated that surface motion in the AP direction may result in the apparent displacement along the other two directions. In fact, the PCA analysis incorrectly suggests that the contribution of surface points to the SI breathing motion is mainly focused in the abdominal region, whereas ground-truth motion occurs in the AP direction. As previously explained, the orientation of the structured-light projector with respect to the patient results in the sliding of the incident pattern when the surface moves in the AP direction.

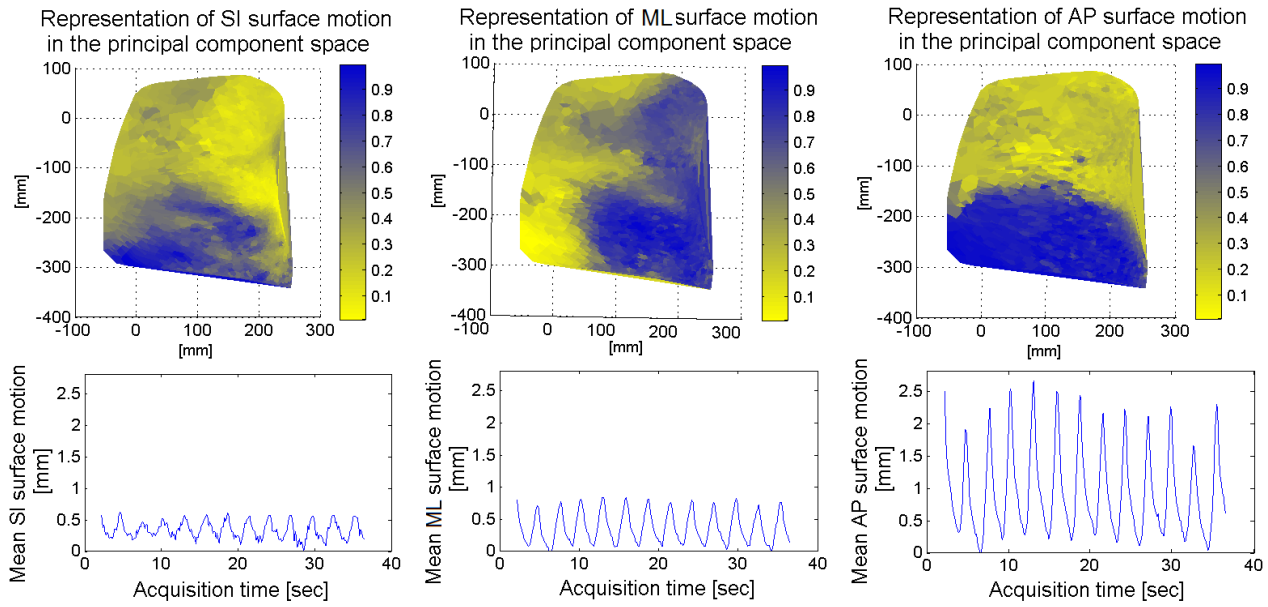


Figure 3.25. The lower panels depict the mean breathing motion of the patient's thoraco-abdominal surface along SI, ML and AP directions, computed by averaging the 3D coordinates of the registered surface points. The upper panels show the spatial representation of the multi-dimensional surface motion in the principal component space, depicting for each surface point the first principal component score obtained by applying PCA analysis to the point trajectories in the three directions. Surface points with high scores (blue colour) strongly contribute to the breathing motion in that specific direction.

The estimated motion of the thoracic and abdominal surface regions was also compared to the respiratory signal provided by the GateCT system. As described in Section 1.4.2.2, this signal is computed as the mean AP trajectory of the surface points included in a circular region of interest selected by the user. Figure 3.26 shows the comparison of the respiratory signal acquired with the GateCT system and the AP motion estimated with deformable registration for three surface points selected in the thoracic and abdominal regions. The signal of the central abdominal point correctly follows the GateCT signal, which was obtained by choosing the surface patch on the central part of the abdomen. The difference between the GateCT signal and the trajectories of the thoracic and lateral abdominal points shows the potential variability of the external surface motion. While the GateCT system provides the respiratory information on a limited surface patch, the proposed method yields the 3D motion of any points on the thoraco-abdominal surface. This approach allows to capture the patient-specific respiratory patterns of the thorax and abdomen, which generally feature different amplitude and phase shifts.

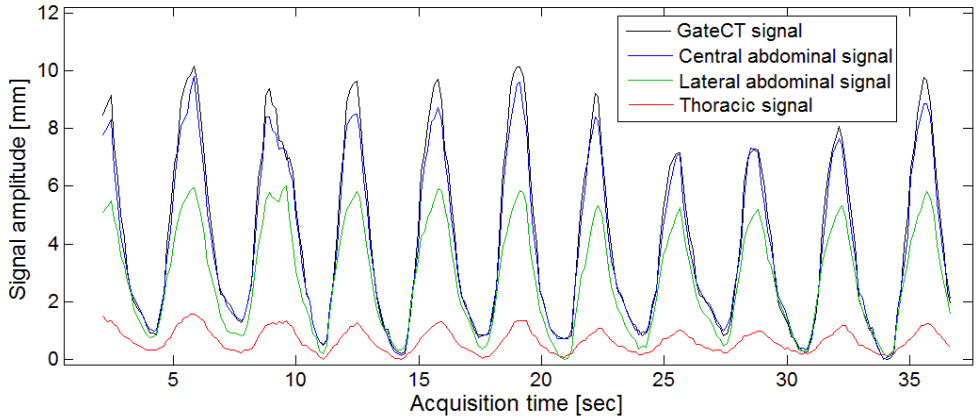


Figure 3.26. Comparison between the acquired GateCT signal and the estimated AP motion of three surface points selected on the thorax and on the central and lateral areas of the abdomen.

3.5 Surrogate breathing signal

As described in Section 2.3.2, we investigated and compared different approaches for summarizing the breathing trajectories of all surface points of the thoraco-abdominal region into a single respiratory surrogate signal, used for tumour tracking application. The tested techniques include principal component analysis (Section 2.3.2.1), k-means clustering (Section 2.3.2.2) and self-organizing map (Section 2.3.2.3). A specific dataset of 5 healthy subjects was used for the experimental testing of the different synthesis approaches. The selection of the optimal technique for the extraction of the breathing surrogate signal from the external surface displacement was

based on the comparison of the obtained signals with a reference variable that can reliably describe the intra-fraction target motion due to breathing. The diaphragm motion was selected as reference variable, since it was proved to have an accurate and reproducible correlation with lung tumour motion [Cerviño 2009a].

The analyzed dataset included 5 male subjects, undergoing the simultaneous acquisition of the external surface displacement and of the diaphragm motion during free-breathing and supine posture conditions. The surface motion was acquired by means of a marker-based optical tracking system. 42 passive reflective markers with 6 mm diameter were placed on the anterior portion of the thorax and abdomen of the test subjects, according to the spatial configuration adopted in the opto-electronic plethysmography (OEP), as depicted in Figure 3.27. OEP method is based on the non-invasive measurement of the displacement of a finite number of surface points by means of opto-electronic systems to assess the absolute chest wall volumes and their variations in the rib cage and abdomen, as described in Section 1.4.2 [Cala 1996, Aliverti 2001]. The grid configuration of the passive markers employed in the present application consists of 7 horizontal rows arranged circumferentially between the level of the clavicles and the anterior-superior iliac spine. Along the horizontal rows, the markers were arranged in 5 vertical columns, with an additional bilateral column in the midaxillary line (Figure 3.27). The 3D trajectories of the passive markers placed on the subject's skin were reconstructed with the ELITE opto-electronic localizer (BTS Engineering, Milan, Italy). The system is composed by two infrared TV cameras, which operates at an acquisition frame rate of 50 Hz, synchronized with coaxial infrared flashing light-emitting diodes. A dedicated parallel processor performs real-time pattern recognition to compute with high accuracy the three-dimensional coordinates of the surface markers, as previously explained in Section 1.4.2.1.

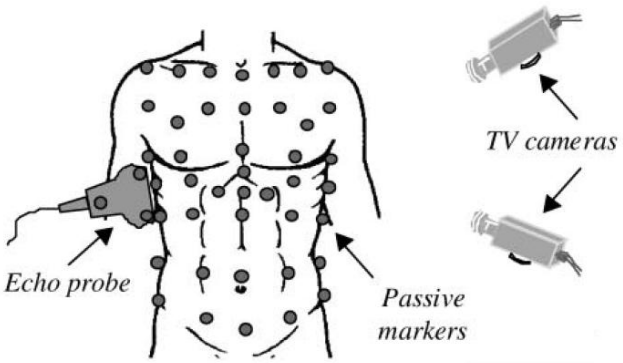


Figure 3.27. Schematic representation of the experimental setting for the synchronized acquisition of the external surface motion based on passive markers placed on the subject's thorax and abdomen, and of the internal diaphragm motion by using a markerized echo probe [Aliverti 2003].

For the 5 subjects of the considered dataset, the displacement of the passive markers placed on the thoraco-abdominal surface was acquired simultaneously with the non-invasive ultrasound imaging of the diaphragm motion [Aliverti 2003]. The ultrasonography of the diaphragm was performed with the Aloka Echo Camera, equipped with a linear ultrasound probe of 128 mm size and operating at a frequency of 3.5 MHz. As shown in Figure 3.27, the echo probe was aligned approximately axially to the subjects and placed between the two passive markers defining the right-lateral abdominal rib cage. During the experiments, the ultrasound probe was fixed to the subject's skin by adhesive tape and manually kept in a stable position. Three additional reflective markers were placed on the echo probe and tracked with the TV cameras of the opto-electronic localizer, to measure the probe's 3D position and orientation. The ultrasonographic images were synchronized with the opto-electronic system by generating a trigger signal. The spatial resolution of the echographic images was 320×240 pixels, while the acquisition frame rate was 10 Hz. A representative example of an echographic image obtained for a test subject is depicted in Figure 3.28. The cephalic margin of the zone of apposition was manually identified in the acquired ultrasound images for all the test subjects. As the intrapulmonary air greatly attenuates the transmission of ultrasound waves, the resulting abrupt discontinuity in the image at the level at which the diaphragm reflects from the chest wall and the lung intervenes was used to identify the zone of apposition (Figure 3.28). The two-dimensional coordinates of the selected cephalic extremity were mapped into the 3D space by using the information on the spatial localization of the echo probe derived from the opto-electronic system. The obtained 3D coordinates were projected to compute the superior-inferior displacement of the margin of the zone of apposition.

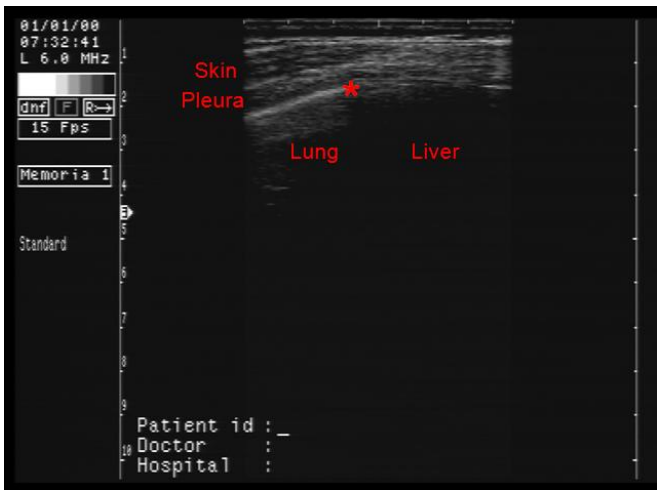


Figure 3.28. Exemplificative echographic image obtained from the right-lateral abdominal rib cage of a test subject. The red star identifies the tracked diaphragmatic point, represented by the cephalic margin of the zone of apposition.

3.5.1 Spatial variability of surface motion

From 114 to 211 ultrasound images of the diaphragm were recorded for each test subject, corresponding to on average 17 seconds of synchronized diaphragm and surface motion acquisition. The resulting number of free-breathing cycles analyzed per subject was 5. The trajectories of surface markers reconstructed by the optical system were down-sampled at the same frequency of the echographic acquisition (10 Hz). For each surface control point of the subject's thoraco-abdominal surface, a mono-dimensional distance signal is derived by computing the time-series of the distance from the lowest position assumed by the single points. As introduced in Section 2.3.2, the spatial variability of respiratory surface motion was investigated by using k-means clustering techniques. K-means algorithm was applied to the distance signals of all markers, to group surface control points according to their breathing motion pattern. The similarity metric used for k-means clustering was the Euclidean distance from the randomly initialized cluster centroids. The spatial distribution of the obtained surface clusters was analyzed for different number of clusters. For each test subject, the number of clusters *k* given in input to the k-means algorithm varied between 2 and 6. Two different cluster validity indices, represented by the Dunn and Silhouette index, were computed to determine the optimal number of clusters for surface markers, namely the parameter *k* that leads to clusters that best fit the given dataset [Ansari 2011].

Dunn validity index [Dunn 1974] attempts to identify those cluster sets that are compact and well separated, combining cluster diameters and dissimilarity between clusters to estimate the most reliable cluster number. The Dunn index is defined by the ratio between the minimal intra-cluster distance to the maximal inter-cluster distance. For any number of clusters, where c_i represent the *i*-cluster of such partition, the Dunn index *D* can be calculated with the following formula:

$$D = \frac{\min\limits_{i,j}[dist(c_i, c_j)]}{\max\limits_{l}[diam(c_l)]}, \tag{3.4}$$

where $dist(c_i, c_j) = \min\limits_{x_i \in c_i, x_j \in c_j}[d(x_i, x_j)]$ represents the smallest distance between two elements of different clusters, and $diam(c_l) = \min\limits_{x_{l_1}, x_{l_2} \in c_l}[d(x_{l_1}, x_{l_2})]$ is the largest distance of two elements from the same cluster, namely the cluster diameter. The optimal number of clusters is the one that maximizes the Dunn index.

The Silhouette validity index [Rousseeuw 1987] is based on the value of the silhouette width S_i of each i^{th}-element of the different clusters, that can be computed as:

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)}, \tag{3.5}$$

where a_i is the average dissimilarity of the i^{th} element to all other elements in the same cluster, usually estimated using a distance metric, and b_i is the minimum average dissimilarity of the i^{th} element to all elements in other clusters. Equation (3.5) results in $-1 \leq S_i \leq 1$. A value of S_i close to 1 indicates that the cluster element is assigned to the most appropriate cluster. If S_i is close to zero, it means that the cluster element is equidistant from more than one clusters, so it can be correctly assign to all of those clusters. If S_i is close to –1, the cluster element is misclassified. The best clustering corresponds to the largest overall average silhouette width for the entire dataset, namely the Silhouette index given by the average S_i for all cluster elements. Therefore, the number of clusters with the maximum Silhouette index is taken as the optimal cluster number.

Figure 3.29 depicts the values of the Dunn and Silhouette validity indices computed on the k-means clustering results of all test subjects for a number of clusters ranging from 2 to 6. The resulting optimal cluster number, corresponding to the highest value of the two indices, was two, indicating that the external surface can be mainly partitioned in two distinct regions with different breathing motion patterns. As shown in Figure 3.30a, the two clusters are mainly localized in the thoracic and abdominal compartments, usually featuring respiratory motion with different amplitude and phase shift [Aliverti 2001]. These results are confirmed also by the principal components analysis applied to the marker distance signals. For all the analyzed subjects, the median value of the data variance explained by the first two principal components (PC) was 70% (Figure 3.31). The PC scores obtained for each surface point were used to evaluate the surface regions that mostly contribute to the respiratory motion signal associated to the first two principal components. As depicted in figure 3.30b, the first PC was mainly localized on the abdominal region, showing the greater motion amplitude, whereas the second PC was predominant in the thoracic area.

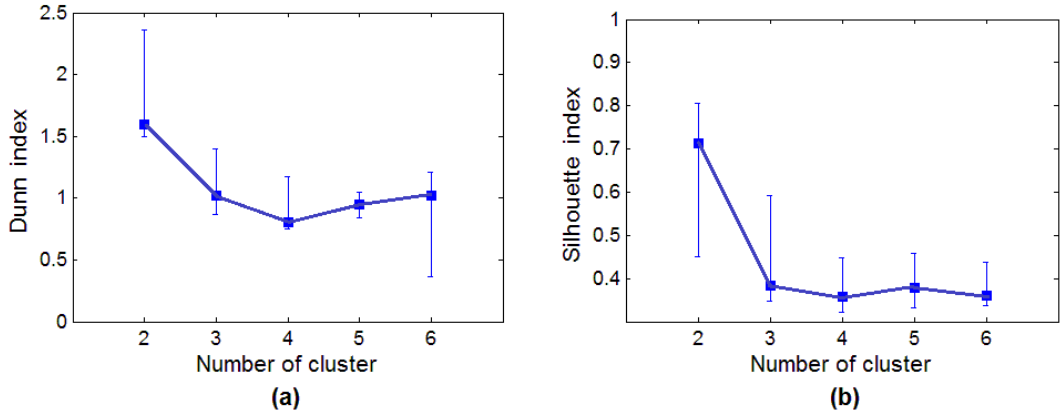


Figure 3.29. Dunn and Silhouette validity indices (median value ± inter-quartile range), computed on the clustering results of all test subjects for a number of clusters ranging from 2 to 6.

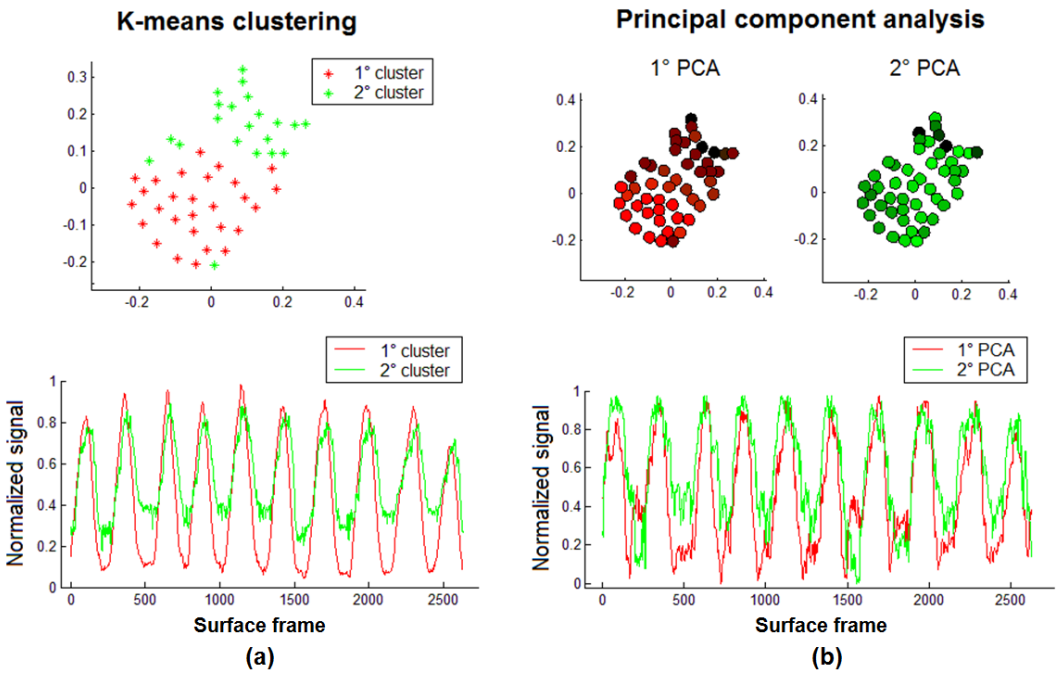


Figure 3.30. Spatial variability results of surface motion extracted from a prolonged optical acquisition of a test subject. Panel a shows the spatial distribution of the first two clusters of surface markers obtained with the k-means algorithm (upper panel) and the breathing motion signals associated to each centroid of the two clusters (lower panel). Panel b depicts the spatial distribution of the first two principal components (upper panel) and the breathing motion signals associated to each PC (lower panel).

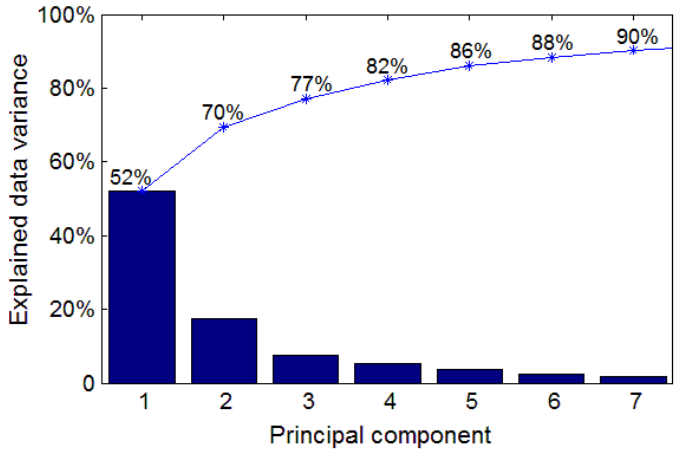


Figure 3.31. Percentage of input data variance explained by the first 7 principal components. Results are expressed as median values of the PC variance computed for each test subject. The cumulative distribution of the explained data variance is also represented.

3.5.2 Comparison with diaphragm motion

The three different methods for the synthesis of the breathing information coming from all surface control points were applied to the marker's distance signals, in order to retrieve a robust and

reliable respiratory surrogate from the external surface motion. As introduced in Section 2.3.2, for the PCA approach the overall breathing surrogate was obtained by combining the signals associated to the first two principal components as a function of their explained variance. For k-means and SOM techniques, we averaged the signals associated to the first two clusters and to the first two neurons. The breathing surrogate signals obtained with the three different methods of synthesis were compared to the reference variable, represented by the SI diaphragm motion obtained from the synchronized ultrasound imaging. Possible phase shifts between the compared signals were compensated by translating the surface-based signals by known temporal delays and finding the maximum correlation with the diaphragm signal. The breathing signals associated to diaphragm and surface motion obtained with the three different methods of synthesis for a test subject, after signal smoothing and amplitude normalization between 0 and 1, are depicted in Figure 3.32.

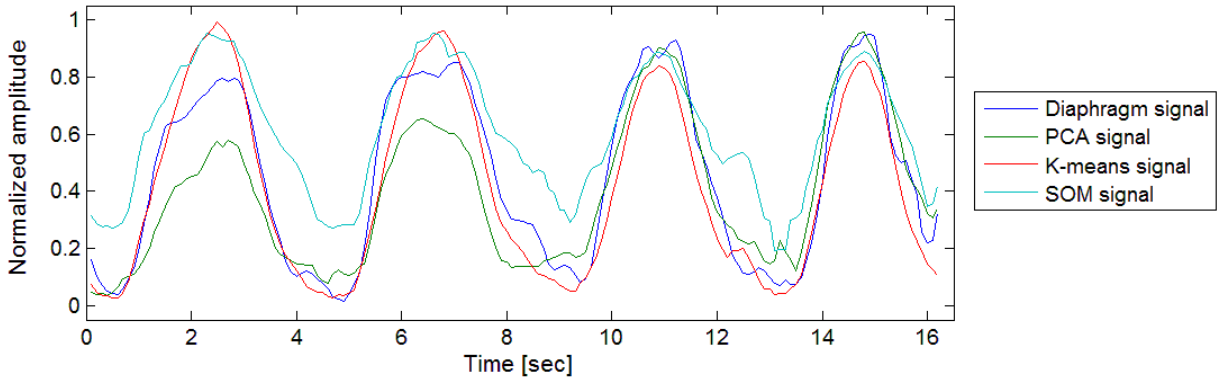


Figure 3.32. Comparison between diaphragm motion and surface-based breathing surrogate signals obtained with PCA, k-means and SOM techniques for subject S3.

The metrics employed for diaphragm and surface motion comparison are represented by the linear correlation coefficient and by the root-mean-square error (RMSE) between the normalized signals. The resulting Pearson coefficients and RMSE values obtained for the 5 test subjects are reported in Table 3.13. According to both metrics, the respiratory surrogate signals summarized with the k-means clustering method proved to better reproduce the diaphragmatic reference variable, obtaining a median value of the Pearson correlation coefficient of 0.93 and a root-mean-square error of 0.11. PCA and SOM techniques showed lower correlation, with a median Pearson coefficient of 0.90 and 0.91, and higher RMSE median values, equal to 0.15 and 0.20, respectively. Since diaphragm displacement represents a robust and reliable surrogate for lung tumour motion [Cerviño 2009a], k-means clustering was selected as the optimal algorithm for the synthesis of the surrogate breathing signal from the external optical surface acquisitions for tumour tracking applications.

Table 3.13. Pearson correlation coefficients and root-mean-square errors between the reference diaphragm motion and the normalized surrogate breathing signals obtained with the three different methods of synthesis.

Subject	Correlation coefficient			Root-mean-square error		
	PCA	K-means	SOM	PCA	K-means	SOM
S1	0.47	0.88	0.46	0.26	0.18	0.25
S2	0.95	0.93	0.93	0.08	0.10	0.10
S3	0.90	0.95	0.93	0.15	0.11	0.20
S4	0.87	0.90	0.91	0.19	0.15	0.14
S5	0.94	0.96	0.58	0.12	0.08	0.26

3.6 Surface-based tumour tracking

In this section, we report the results of the experimental evaluation of the proposed tumour tracking technique on the 7 lung cancer patients included in the database. As previously described, the testing is based on the comparison between the real tumour trajectory identified on cone-beam CT projection images and the target motion estimated from the external surface displacement using a patient-specific 4D CT breathing model. For each patient, the surrogate breathing signal was extracted from the acquired optical surfaces using the k-means clustering approach described in the previous section. The instantaneous values of the amplitude and phase parameters of the 4D CT motion model were retrieved from the computed respiratory surrogate, as described in Section 2.3.3. These parameters were used to estimate the 3D tumour trajectory during the whole surface scan from the tumour positions in the different breathing phases derived from the 4D CT planning images and updated according to the on-line information coming from the in-room CBCT imaging (see Sections 2.2.1 and 2.2.2). The illustrative results obtained for patient P5 related to the surface-based breathing surrogate, to the updated tumour positions in the 4D CT phases and to the 3D target trajectory estimated during the surface scan are depicted in Figure 2.18, Figure 2.19 and Figure 2.20, respectively.

Table 3.14 reports for each patient the comparison between the amplitude of tumour motion during planning and treatment phases. The planning target amplitudes were estimated as the difference between the maximum and minimum tumour coordinates along each spatial direction derived from the updated 4D CT breathing motion model. The treatment target amplitudes were instead computed as the difference between the 95^{th} and 5^{th} percentiles of the tumour coordinates estimated along each spatial direction with the proposed tumour tracking approach. The median value of the amplitude-based scaling factor parameter retrieved from the comparison of the external

surface motion during planning and treatment phases is also listed in Table 3.14. For all patients except one, the amplitude of surface motion recorded during treatment was higher than during planning phase, thus producing an amplification of the tumour motion amplitude. For 3 out of 7 patients, the amplitude scaling factor was superior than 0.5. For 5 patients, tumour motion amplitudes proved to be higher in the anterior-posterior directions, while 2 patients showed larger motion amplitudes along the SI direction. For 3 patients, not-negligible amplitudes were recorded also for the medio-lateral direction, with values greater than 3 mm both during planning and treatment phases.

Table 3.14. Comparison between tumour motion amplitudes along each spatial direction computed during planning and treatment phases. The median value of the amplitude scaling factor estimated from the surface-based surrogate breathing signal is also listed in the table.

Patient	**Tumour amplitude during planning (mm)**			**Tumour amplitude during treatment (mm)**			**Amplitude scaling factor**
	ML direction	**AP direction**	**SI direction**	**ML direction**	**AP direction**	**SI direction**	
P1	4.0	4.7	2.5	6.1	7.1	3.9	0.63
P2	0.4	5.2	3.0	0.8	9.1	5.8	0.53
P3	0.8	1.2	1.3	0.6	0.8	1.1	-0.06
P4	0.8	2.0	3.6	0.8	1.9	3.7	0.05
P5	3.1	7.3	6.7	3.6	8.0	7.7	0.09
P6	3.8	4.9	1.4	5.1	5.7	1.7	0.23
P7	0.3	3.6	2.6	0.3	5.9	4.5	0.54

For each cone-beam CT scan included in the database, the position of lung tumours on CBCT images was identified through the developed contrast enhancement method, described in Section 2.5.1. A manual refinement of the automatically detected target coordinates was performed, resulting in more regular tumour trajectories and in an increased target identification rates. As explained in Section 3.1.3, only the first 35 seconds of synchronized cone-beam and optical scans could be acquired without surface occlusion due to the rotation of the CBCT system. Table 3.15 lists for each patient of the database the number of CBCT projections acquired in the non-occluded period in which lung tumour could be correctly identified. These images were used for the evaluation of the accuracy of the developed tumour tracking technique. The mean number of cone-beam CT projections available per patient for the experimental method testing was 153, ranging

between 189 and 105 projections. The corresponding period of time analyzed per patient varied from 35 to 19 seconds, with a mean value of 28 seconds (Table 3.15). The number of breathing cycles that could be identified depends not only on the analyzed period of time, but also on the respiratory frequency specific for each patient. As reported in Table 3.15, the number of breathing cycles identified per patient ranged from a maximum of 9 to a minimum of 5. The duration of the respiratory cycle averaged over all cycles acquired for each patient in the non-occluded period is also listed in the Table. The measured mean cycle length ranged between 1.8 and 4.6 seconds per cycle.

Table 3.15. Number of cone-beam CT images per patient used for the assessment of the accuracy of the proposed tumour tracking method, along with the corresponding evaluated period of time and number of breathing cycles. The mean duration of the respiratory cycle is also listed for each patient.

Patient	Evaluated CBCT projections	Evaluated period of time (sec)	Evaluated breathing cycles	Mean cycle length (sec)
P1	166	30.3	9	2.8
P2	175	31.9	5	4.6
P3	105	19.2	9	1.8
P4	189	34.5	8	3.3
P5	144	26.3	5	3.5
P6	159	29.0	6	3.3
P7	135	24.6	5	3.2

3.6.1 Real and estimated target trajectories

For each of the 7 lung cancer patients of the acquired database, the 3D trajectory of the tumour centroid estimated from the surface-based breathing surrogate signal with the a priori 4D CT motion model was projected on the 2D plane of the cone-beam CT panel detector at the corresponding rotational angles specified in the acquisition file of each CBCT scan. Centroid projection was performed trough ray-tracing algorithms using the Plastimatch program also applied for the generation of the digital reconstructed radiographs. The 2D tumour trajectories projected along the horizontal and vertical dimensions of CBCT images obtained with the proposed tumour tracking method for patient P5 are depicted in Figure 3.33.

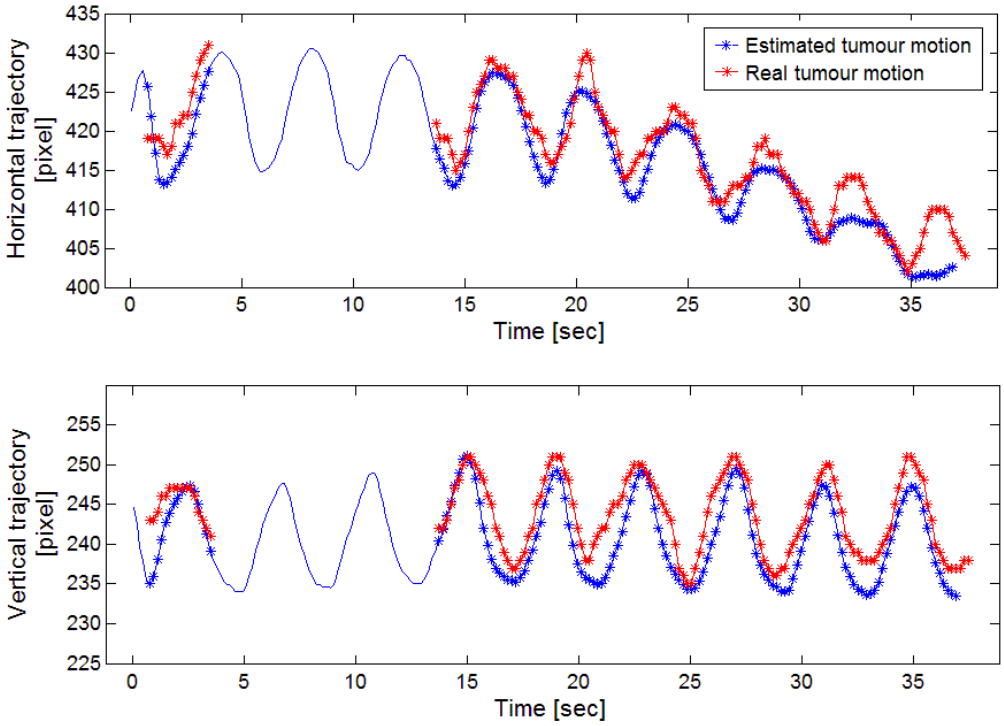


Figure 3.33. Comparison between real tumour trajectories identified in CBCT projections along the horizontal and vertical image dimensions and target trajectories estimated from surface motion for patient P5. The estimated tumour coordinates interpolated at the timestamps corresponding to the cone-beam CT images in which lung tumour could be identified are represented as blues stars.

Since the acquisition frame rates of CBCT and optical surface scans did not correspond, the 2D tumour trajectories estimated from surface motion were linearly interpolated at the acquisition timestamps corresponding to the cone-beam CT images in which lung tumour could be identified (Figure 3.33). The flexmap values at different projection angles stored in the acquisition file of the CBCT scan were added to the estimated horizontal and vertical tumour trajectories to compensate for the flexions and torsions of the cone-beam CT system. The tumour tracking accuracy was assessed by computing the difference between the real target trajectories identified in the CBCT images and the projected target trajectories estimated from the patient's surface displacement with the developed approach. The median values of the differences obtained for the tumour trajectories along the two image dimensions are listed in Table 3.16. The tracking errors obtained for all patients of the database ranged between 0.7 and 2.2 mm for the horizontal image direction and between 1.2 and 2.4 mm for the vertical image directions. The median values of the inter-quartile ranges of the tracking errors estimated for each patient proved to be 1.4 and 1.3 mm for the horizontal and vertical tumour trajectories, respectively.

Table 3.16. Tracking errors between the real and estimated tumour trajectories along the horizontal and vertical image directions. Results are expressed as median value [25^{th} – 75^{th} percentile].

Patient	Horizontal tracking error (mm)	Vertical tracking error (mm)
P1	2.2 [1.2 – 3.8]	1.6 [0.9 – 2.7]
P2	1.7 [0.8 – 3.0]	1.2 [0.6 – 1.9]
P3	1.2 [0.6 – 1.7]	2.0 [1.6 – 2.4]
P4	0.7 [0.4 – 1.0]	1.4 [0.9 – 1.9]
P5	1.7 [0.8 – 3.1]	2.4 [1.4 – 3.5]
P6	1.3 [0.8 – 1.9]	1.5 [0.7 – 2.0]
P7	1.5 [0.9 – 2.3]	1.4 [0.6 – 2.2]

For each lung cancer patient, we investigated the contribution of different breathing parameters to the estimated tracking errors. The complete respiratory cycles were extracted from the real and estimated 2D tumour trajectories according to the positions of the inspiratory peaks separately identified for the horizontal and vertical image directions. The following breathing parameters were derived for each extracted respiratory cycle:

- baseline, computed as the mean value of the tumour trajectory for each cycle;
- amplitude, obtained from the difference between the maximum and minimum value of the tumour coordinates in each cycle;
- period, estimated as the temporal difference between the acquisition timestamps associated to the two inspiratory peaks at the beginning and at the end of each cycle;
- phase, represented by the mean acquisition timestamp of the two inspiratory peaks of each cycle.

Table 3.17 reports for each patient the differences between the respiratory parameters computed for the real and estimated tumour trajectories. The median values of the absolute errors of the breathing parameters computed for both horizontal and vertical image directions are listed in the table. The phase errors were also expressed as percentage of the real period of the corresponding cycle identified from CBCT images. The median errors of the respiratory baseline and amplitude ranged between 0.7–2.2 mm and 0.9–2.6 mm, respectively. The differences in the estimation of cycle period varied from about 70 to 200 msec. The phase shifts measured for each patient were lower than 7% of the cycle length, corresponding to a maximum of 170 msec of delay.

Table 3.17. Median differences between the respiratory parameters computed for each breathing cycle identified on the real and estimated tumour trajectories.

Patient	Baseline error (mm)	Amplitude error (mm)	Period error (msec)	Phase error (msec)	Percentage phase error (%)
P1	1.6	1.3	130	110	2.9
P2	0.7	2.6	129	70	1.0
P3	1.7	1.0	71	111	6.7
P4	0.8	1.7	170	136	3.3
P5	2.2	0.9	203	170	3.7
P6	1.2	1.0	74	107	2.7
P7	1.2	1.2	198	156	5.7

The influence of the different respiratory parameters on the tumour tracking accuracy was assessed by computing the correlation between the estimated errors for the breathing parameters and for the tumour trajectories. For each breathing cycle, the tracking error was computed as the median difference between the real and estimated tumour trajectories associated to that cycle. The Pearson linear correlation coefficient was calculated between the cycle tracking errors and the corresponding differences in breathing parameter. The correlation was considered statistically significant if the Pearson coefficient was higher than 0.8 with a p-value lower than 0.1. For 5 out of 7 patients, the tracking errors proved to be significantly correlated with the baseline differences measured for tumour trajectories along at least one of the two image directions. Only for one patient, a significant correlation was found between the tracking errors and the amplitude differences, whereas no correlation was measured for the phase and period parameters. The ability of the developed tumour tracking method to robustly estimate target motion from external surface displacement was demonstrated even in presence of breathing irregularities. Figure 3.34 depicts the 2D tumour trajectories identified for patient P2, showing strong irregularities both in the respiratory amplitude and frequency. Despite the evident inter-cycle variations of the two main breathing parameters, the tumour trajectories estimated with the proposed method correctly followed the real target motion identified in CBCT projections, with median tracking errors of 1.7 and 1.2 mm for the horizontal and vertical directions, respectively.

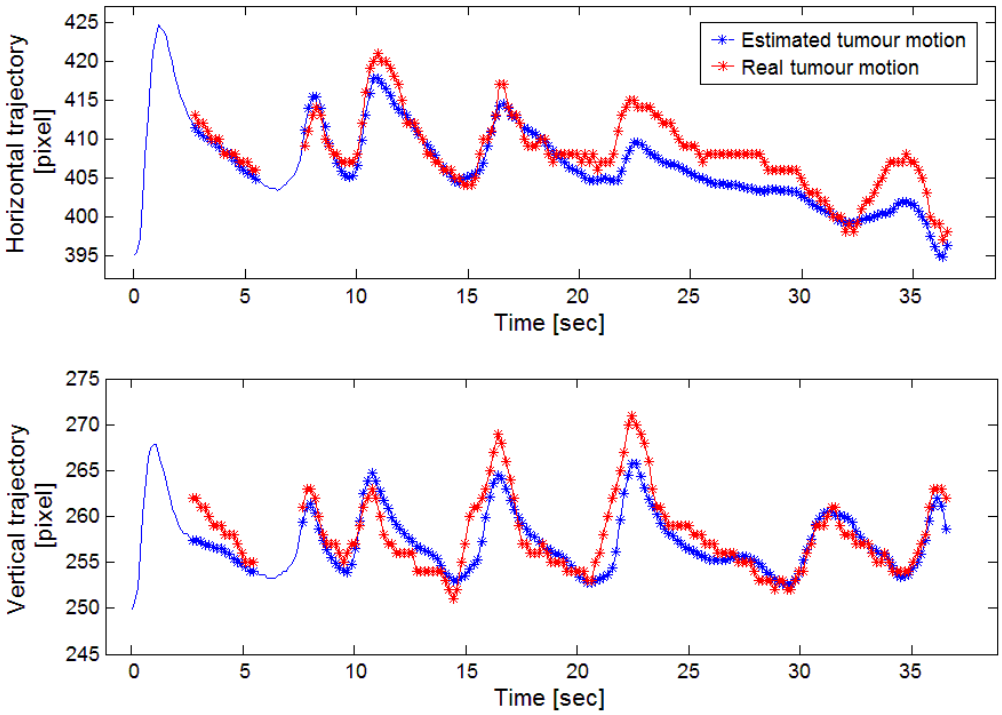


Figure 3.34. Comparison between real and estimated tumour trajectories obtained for patient P2, showing strong breathing irregularities both for the amplitude and phase parameters.

4. CONCLUSIONS

The thesis project was focused on the development and experimental testing of a tumour tracking method that allows the continuous monitoring of target position during radiotherapy treatments for the real-time compensation of intra-fraction organ motion. The proposed approach can be generally applied to all tumours located in the abdominal and thoracic districts, such as lung, liver or pancreas cancers, affected by non-negligible physiological movements mainly due to respiration. Tumour motion monitoring and mitigation strategies are required to compensate for motion-induced geometric errors, which may seriously affect the effectiveness of radiotherapy treatments causing unacceptable under-dosage of the target volume and over-dosage of the surrounding healthy tissues [Bert 2011]. The developed tumour tracking method is based on the integration of different imaging modalities, such as time-resolved planning imaging, in-room X-ray verification systems and 3D optical tracking devices. According to the proposed approach, target motion during treatment is indirectly estimated from an external respiratory surrogate, represented by the displacement of the patient thoraco-abdominal surface acquired by means of non-invasive optical surface imaging systems. The correlation between external surface motion and internal tumour position is achieved through patient-specific breathing motion models, derived directly from 4D CT planning images and updated at each treatment fraction according to the anatomical variations obtained from in-room imaging and optical surface devices. The development and experimental testing of the presented tumour tracking method involve different procedural steps and complementary algorithms, that are briefly discussed in the following paragraphs.

4.1 Synchronized internal/external motion data collection

A clinical database was specifically collected for the experimental evaluation of the accuracy of the proposed tumour tracking method, including the synchronized information on internal and external breathing motion acquired from 7 lung cancer patients. Data collected for each patient come from different imaging modalities, namely time-resolved planning CT, in-room cone-

beam CT and 3D dynamic optical surfaces captured with the non-invasive VisionRT surface imaging system. The acquired database provides in particular the synchronized breathing motion data related to the patient's external surface and internal tumour during cone-beam CT scans. An image-based method was developed to synchronize the acquisitions of CBCT projections and VisionRT surfaces. The proposed synchronization approach was assessed on a moving phantom, obtaining temporal accuracy lower than 20 msec. The collected database represents a valuable and useful tool to analyze the correlation and variability of breathing motion coming from multi-modal imaging systems, as well as for the experimental testing and comparison of different tumour tracking approaches based on external-internal correlation models. A limitation of the current database is represented by the type of included lung cancer patients. In fact, only patients with lesions located in the upper lung lobes, featuring a reduced breathing motion amplitude, could be selected for the present application. The future validation of the developed tumour tracking technique will necessarily require the experimental evaluation on lung cancer patients with lesions located also in the lower lobes, thus assessing the method accuracy in estimating tumour motion with greater amplitudes.

4.2 Contrast-enhanced tumour identification on X-ray images

We developed and investigated a contrast enhancement method for markerless lung tumour tracking in X-ray projections, exploiting the anatomical information derived from planning CT volumes. The proposed approach tested on cone-beam CT images proved to be effective in enhancing the visibility of lung lesions and reducing the overlapping effect of the surrounding anatomical structures. Results confirmed significant improvements in tumour tracking performance and spatial localization accuracy for a wide range of projection angles with respect to state-of-the-art markerless tracking techniques [Hugo 2010, Lewis 2010]. The developed algorithm allows about a 150% increase of the CBCT image contrast between lung tumours and the overlying anatomy. The rate of target identification in the contrast-enhanced cone-beam projections on the 7 analyzed patients ranged from 57% to 69%, with a 24% of increase with respect to state-of-the-art tracking techniques. The median 2D tracking accuracy proved to be lower than 2 mm both along the horizontal and vertical image directions, whereas median errors in 3D tumour localization were inferior than 0.5 mm in all spatial directions. While 2D tracking results are comparable to the outcomes of conventional state-of-the art techniques, the 3D tracking accuracy assessment accounted for a significant improvement obtained by applying the proposed contrast-enhanced method.

Despite the need of additional clinical data to validate the efficacy of contrast enhancement for improved tumour detection in X-ray images, the reported results are put forward as an evidence that the application of the proposed technique can indeed increase the accuracy of markerless lung tumour tracking and promote its applicability to a wider range of projection angles and clinical cases. The method lends itself as a clinically practicable and valuable tool for enhancing the potentiality of X-ray based image guidance in lung cancer patients. Potential benefits are related to the improvement of daily patient positioning and to the robust estimation of tumour trajectories in support of strategies for intra-fraction motion compensation. As a final remark, the described method features a low computational cost compatible with real-time performance.

4.3 4D CT breathing motion model

We created a patient-specific respiratory model that describes the tumour motion during the breathing cycle using the information contained in 4D CT imaging data acquired during treatment planning. The proposed respiratory models are estimated through deformable registration algorithms applied to the 4D CT phase images. Differently from other motion models reported in literature [Boldea 2008], we did not use a single phase as reference image for the computation of the deformation vector fields. We instead considered the mid-position CT image, which represents a time-weighted average of all 4D CT phase volumes. The MidP image was proved to contain less noise and artefacts with respect to the single 4D CT phases [Wolthaus 2008], thus increasing the accuracy and robustness of the estimated breathing motion models. We also developed a novel adaptation strategy to update the 4D CT model parameters at each treatment fraction, according to the patient's daily situation. In particular, we use the information on tumour localization derived from the reconstructed CBCT volume to adapt the target breathing baseline, while the amplitude and phase parameters were retrieved from the real-time surface displacement compared to surface motion during the planning 4D CT scan. The updated breathing parameters were used to derive the daily tumour positions for each respiratory phase using the vector fields of the 4D CT model. The proposed approach can be useful not only for tumour tracking applications based on external-internal correlation models, but also in the context of adaptive radiation therapy [Webb 2008]. The daily information on tumour motion obtained from the updated breathing model can potentially be included in the radiotherapy treatment plan, modifying for example the safety margins to cover the whole extent of target motion specific for each treatment fraction.

4.4 Deformable surface registration algorithm

A deformable surface registration algorithm was developed to extract spatial congruence from markerless optical surface acquisitions, thus yielding the three-dimensional trajectories of corresponding surface points. The implemented method was adapted from the non-rigid extension of the iterative closest point algorithm with locally affine regularization proposed by Amberg et al [Amberg 2007], adding further optimization criteria to increase the efficiency and robustness, such as the Cholesky factorization and the exclusion of bad correspondences specifically designed for the proposed application. We investigated in a group of healthy volunteers the performance of the developed mesh registration algorithm in tracking the multi-dimensional breathing motion from markerless optical imaging of the thoraco-abdominal surface. The residual distance between the deformed source and the target meshes proved to be lower than 1.20 mm for all subjects, while the median errors in the localization of surface control points was 1.61 mm. The registration accuracy was maximal in the abdominal region, where breathing motion mostly occurs, with average errors of 1.09 mm. The performance of the developed deformable approach was significantly superior to rigid registration procedures, both in terms of surface overlap and marker localization errors. The obtained results support the need for non-rigid registration when dealing with thoraco-abdominal surfaces, in order to adequately account for the deformation due to breathing motion.

The application of the developed mesh registration algorithm to real patient data demonstrated the clinical feasibility of recovering multi-dimensional breathing motion for different thoraco-abdominal regions from markerless optical surface acquisitions. The proposed method is put forward to significantly improve respiratory motion management in radiation therapy. The state-of-the-art clinical systems that currently account for breathing movement are limited to a restricted area of the body surface. Such approaches do not accurately depict the complexity of human respiration processes, involving the combined composition of thoracic and abdominal surface motion. A multi-dimensional respiratory signal accounting for different breathing patterns, as provided by the developed deformable registration approach, could potentially improve current solutions for motion-compensated treatment planning and delivery. The enhanced surface-based breathing surrogate derived from the implemented deformable mesh registration algorithm is put forward to provide potential benefits for motion artefact reduction during image acquisition and for real-time tumour tracking based on external-internal correlation models.

4.6 Surface-based surrogate signal

We implemented and compared different techniques to synthesize the breathing information coming from all points of the thoraco-abdominal surface in a single representative surrogate signal. The selected method is based on the k-means clustering technique, computing the breathing surrogate from the combination of the motion signals associated to the abdominal and thoracic surface clusters. The obtained signals proved to be highly correlated with the internal diaphragm motion, which represents a good anatomical surrogate for the prediction of lung tumour position, especially for lesions located in the lower lung lobes [Cerviño 2009a]. The developed method of synthesis can be employed to estimate a robust and reliable external surrogate from surface motion information, both for marker-based systems and for surface scanning techniques. The proposed approach can provide substantial improvements in all applications based on surface-based breathing surrogates, such as 4D CT reconstruction and motion compensation techniques, including breath-hold, respiratory gating and tumour tracking based on external-internal correlation. In the present application, the 3D trajectories of surface points are condensed in a mono-dimensional signal, computing the time-series of the distance from the lowest position assumed by the single points. As a future development, we can consider each direction of surface motion independently. The computation of a surface-based breathing surrogate for each spatial dimension can potentially provide a more accurate and robust prediction of tumour motion, allowing for anisotropic scaling factors of the amplitude parameter along the different directions of motion.

4.7 Tumour tracking algorithm

The accuracy of the developed tumour tracking technique was experimentally assessed on a clinical database acquired on 7 lung cancer patients. The tumour trajectories estimated with the proposed approach from the external surface displacement combined with the a priori patient-specific breathing motion model were compared to the real target trajectories directly identified from the synchronized cone-beam CT projections. Three representative examples of tumour tracking results are depicted in Figure 4.1. Differences between the real and estimated tumour trajectories ranged between 0.7 and 2.4 mm, with median values of 1.5 mm both along the horizontal and vertical image dimensions. The obtained tracking errors can be partly associated to the inaccuracies in localizing the lung lesions on CBCT projections. The developed contrast enhancement method for the automatic detection of tumour position on X-ray images featured median 2D tracking errors of 1.9 mm. A manual refinement of the target trajectories was introduced

to increase the tracking accuracy, which is however still limited by the intrinsic spatial resolution of CBCT projections, equal to 0.8 mm/pixel. The reported results provide an objective estimate of the accuracy and geometrical uncertainties in target localization achievable with the novel proposed approach. The obtained tumour tracking errors proved to be lower than the safety margins normally applied for the treatment of moving lesions with conventional X-ray radiotherapy. Although the validation of the proposed approach requires the experimental testing on a wider patient population, the presented work is put forward as a feasibility study for the clinical applicability of the developed surface-based tumour tracking method. The investigated technique is expected to represent an effective tool for the non-invasive compensation of intra-fraction organ motion, potentially allowing the reduction of safety margins and the increase of local tumour control whit the contemporary sparing of the surrounding normal tissues.

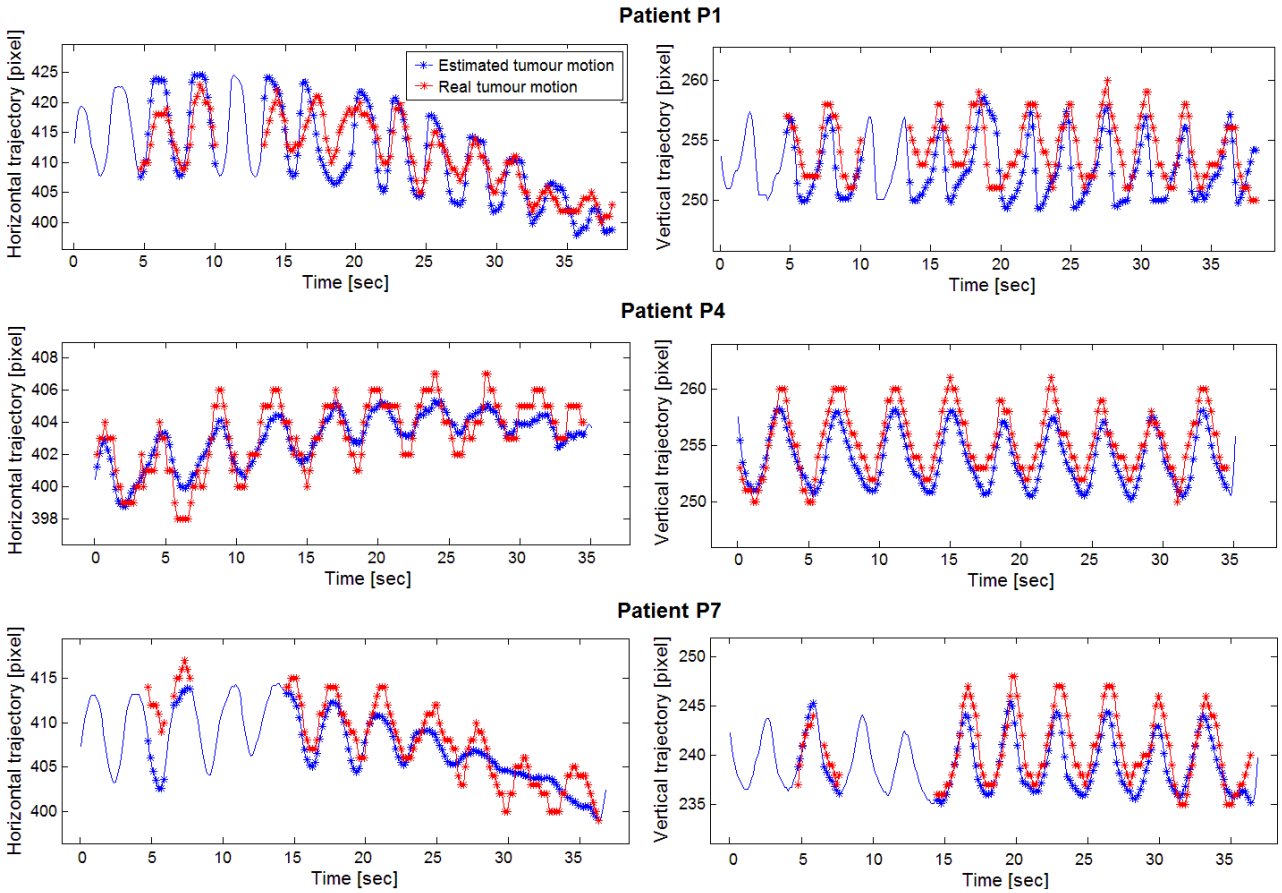


Figure 4.1. Comparison between real tumour trajectories identified in CBCT projections and target trajectories estimated from surface motion for three lung cancer patients of the database (P1, P4 and P7).

The developed approach proved to be effective in estimating target motion from the external surface displacement even in presence of irregularities in the breathing parameters, as depicted in Figure 3.34. The influence of the different respiratory parameters on the tumour tracking accuracy

was assessed for each patient of the database. Median errors of the breathing baseline and amplitude computed for each respiratory cycle did not exceed 2.2 and 2.6 mm, respectively. The measured phase shifts proved to be lower than 7% of the cycle length, corresponding to a maximum of 170 msec of delay. The obtained errors in the identification of the phase of the breathing cycles can be partly related to the temporal resolution of cone-beam CT scans, requiring about 180 msec for the acquisition of a single CBCT image. In most patients of the database, the tracking errors in the identification of tumour trajectories proved to be correlated with the errors in the estimation of the breathing baseline. For 5 out of 7 patients, a Pearson correlation coefficient higher than 0.8 was found between the tracking errors and the baseline differences. In the presently developed tumour tracking technique, only the instantaneous values of the breathing amplitude and phase parameters of the 4D CT motion model are extracted from the surface-based surrogate signal. The systematic variations of the respiratory baseline are compensated inter-fractionally, through the adaptation to the daily information on tumour localization derived from in-room CBCT imaging. A future improvement of the proposed tumour tracking method can be represented by the introduction of intra-fraction updates also of the respiratory baseline parameter, by retrieving its instantaneous values from the surrogate breathing signal obtained from the external surface displacement.

Major concerns are associated to the computational cost of the developed tumour tracking algorithm, which does not currently provide real-time performance. The most time-consuming processes are represented by the 3D surface reconstruction and by the deformable mesh registration. In the present application, we achieved surface frame rates up to 9 Hz since the 3D reconstruction is performed off-line. The time required for the real-time reconstruction of the entire body surface with the employed VisionRT optical system is about 1 second. The acquisition frame rate can be increased by reducing the size of the captured surface, loosing however important information on the spatial variability of the breathing motion patterns of thorax and abdomen. A possible solution can consist in the combination of marker-based methods and surface scanning techniques, thus taking advantage of both the fast real-time 3D localization of passive control points and of the redundancy of surface motion information accounting for inter-session deformations [Spadea 2006]. The possibility to perform real-time deformable mesh registration is also challenging, since in the current algorithm the mean time required for a single outer iteration is about 9 seconds, depending on the number of vertices in the registered meshes. Multi-resolution strategies, GPU-based techniques or a more local regularization approach could potentially provide significant speed gains in the registration algorithm. Future works will be mainly focused on the computational optimization of the developed tumour tracking method, in order to achieve the real-time

performance required for the continuous monitoring of breathing motion during radiotherapy treatments.

Compared to the current techniques clinically applied for dynamic intra-fraction tumour localization, the proposed tumour tracking method features three important advantages. The first advantage is the complete non invasivity of the approach, which allows the continuous real-time monitoring of tumour motion during the whole treatment duration. 3D target position is in fact indirectly inferred from the external surface displacement, which is acquired by means of non-invasive optical tracking systems. This differs from direct tumour tracking techniques based on radiographic imaging, which can be applied only for a limited temporal window during treatment due to the invasiveness of the employed ionizing radiations [Cui 2008, Lin 2009, Rottmann 2010]. The second advantage of the proposed method is the use of markerless 3D surface imaging systems based on stereo-photogrammetric techniques, which allows to dynamically acquire the entire thoraco-abdominal surface of the patient. The external surrogate signal is therefore derived from an extended surface portion including both thorax and abdomen, thus capturing the complex and detailed information on the different breathing patterns involved in external surface motion. The surface-based tumour tracking methods currently employed in the clinical practice, such as the CyberKnife or the Vero systems, are instead based on the acquisition of a discrete number of fiducial markers placed on the patient's skin, thus providing an under-sampled description of surface motion [Kilby 2010, Depuydt 2011]. The third advantage of the developed technique consists in the use of patient-specific adaptive correlation models, that take into account motion information derived from different time-resolved imaging modalities. Target breathing motion models are in fact derived from planning 4D CT images and updated at each treatment fraction according to patient's daily data obtained from in-room X-ray imaging systems. The current tumour tracking methods are instead based on correlation models that are not specific for each patient, such as linear/quadratic correlation, artificial neural networks and fuzzy logic [Torshabi 2010]. These approaches do not take into account motion information derived from time-resolved planning and treatment imaging, that can instead provide a valuable description of the target motion pattern during the breathing cycle.

The tumour tracking method developed in the present thesis project is mainly designed for a potential application to real-time targeting techniques in the advanced active beam scanning systems. Tumour targeting, consisting in the continuous tracking of the tumour’s changing position with the dynamic adjustment of the radiation beam, represents the most efficient and precise motion compensation technique in free-breathing conditions [Murphy 2004]. Particle therapy with active beam scanning offers excellent conditions for real-time motion compensation through tumour

targeting techniques, and extensive researches and technical developments have already been performed to guarantee the precise control of the particle beam's direction [Schardt 2010]. However, the clinical implementation of targeted treatments with scanned particles is currently limited by the inadequate accuracy and robustness of the actual target localization techniques [Rietzel 2010, Riboldi 2012]. The methodological improvements introduced in the developed tumour tracking approach, related to the use of patient-specific adaptive breathing motion models and to the redundancy of surface data, are expected to increase the precision in the estimation of intra-fraction tumour motion. The enhanced performance are put forward to guarantee the adequate tumour tracking accuracy and robustness required for the application of targeting techniques in active particle beam scanning. The expected results consist in the potential extension to extra-cranial moving tumours of the therapeutic benefits of active beam scanning, presently applied only for the treatment of lesions affected by negligible intra-fraction organ motion [Bert 2011]. This will allow to exploit the excellent geometrical selectivity and dose conformity of scanned beams to improve local target control and minimize healthy tissue complications also for tumours in the abdominal and thoracic districts affected by respiratory motion.

REFERENCES

Abdelnour AF, Nehmeh SA, Pan T. Phase and amplitude binning for 4D-CT imaging. *Phys Med Biol* 52, pp. 3515-3129 (2007).

Abdi H, Williams LJ. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2, pp. 433-459 (2010).

ACS (American Cancer Society). Cancer Facts & Figures 2012. Atlanta, Georgia (2012).

Aliverti A, Dellacà R, Pelosi P, et al. Compartmental analysis of breathing in the supine and prone positions by optoelectronic plethysmography. *Ann Biomed Eng* 29, pp. 60-70 (2001).

Aliverti A, Ghidoli G, Dellacà RL, et al. Chest wall kinematic determinants of diaphragm length by optoelectronic plethysmography and ultrasonography. *J Appl Physiol* 94, pp. 621-630 (2003).

Allen B, Curless B, Popovic Z. The space of human body shapes: reconstruction and parameterization from range scans. *ACM Trans Graph* 22, pp. 587-94 (2003).

Amaldi U, Kraft G. Radiotherapy with beams of carbon ions. *Rep Prog Phys* 68, pp. 1861-1882 (2005).

Amaldi U, Kraft G. European Developments in Radiotherapy with Beams of Large Biological Effectiveness. *J Radiat Res* 48, pp. A27-A41 (2007).

Amberg B, Romdhani S. Vetter T. Optimal step nonrigid ICP algorithms for surface registration. *Proc IEEE Conf Computer Vision and Pattern Recognition*, pp. 1-8 (2007).

Ansari Z, Azeem MF, Ahmed W. Quantitative Evaluation of Performance and Validity Indices for Clustering the Web Navigational Sessions. *WCSIT Journal* 5, pp. 217-226 (2011).

Arimura H, Egashira Y, Shioyama Y, et al. Computerized method for estimation of the location of a lung tumor on EPID cine images without implanted markers in stereotactic body radiotherapy. *Phys Med Biol* 54, pp. 665-677 (2009).

Balter JM, Wright JN, Newell LJ, et al. Accuracy of a wireless localization system for radiotherapy. *Int J Radiat Oncol Biol Phys* 61, pp. 933-937 (2005).

Baroni G, Troia A, Riboldi M, et al. Evaluation of methods for opto-electronic body surface sensing applied to patient position control in breast radiation therapy. *Med Biol Eng Comput* 41, pp. 679-688 (2003).

Baroni G, Riboldi M, Spadea MF, et al. Integration of enhanced optical tracking techniques and imaging in IGRT. *J Radiat Res* 48, pp. A61-A74 (2007).

Berbeco RI, Jiang SB, Sharp GC, et al. Integrated radiotherapy imaging system (IRIS): design considerations of tumour tracking with linac gantry-mounted diagnostic x-ray systems with flat-panel detectors. *Phys Med Biol* 49, pp. 243-255 (2004).

Berbeco RI, Neicu T, Rietzel E, et al. A technique for respiratory-gated radiotherapy treatment verification with an EPID in cine mode. *Phys Med Biol* 50, pp. 3669-3679 (2005a).

Berbeco RI, Nishioka S, Shirato H, et al. Residual motion of lung tumours in gated radiotherapy with external respiratory surrogates. *Phys Med Biol* 50, pp. 3655-3667 (2005b).

Bert C, Metheany KG, Doppke K, Chen GT. A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup. *Med Phys* 32, pp. 2753-2762 (2005).

Bert C, Saito N, Schmidt A, et al. Target motion tracking with a scanned particle beam. *Med Phys* 34, pp. 4768-4771 (2007).

Bert C, Grözinger SO, Rietzel E. Quantification of interplay effects of scanned particle beams and moving targets. *Phys Med Biol* 53, pp. 2253-2265 (2008).

Bert C, Gemmel A, Saito N, Rietzel E. Gated irradiation with scanned particle beams. *Int J Radiat Oncol Biol Phys* 73, pp. 1270-1275 (2009).

Bert C, Gemmel A, Saito N, et al. Dosimetric precision of an ion beam tracking system. *Radiat Oncol* 5, pp. 61 (2010).

Bert C, Durante M. Motion in radiotherapy: particle therapy. *Phys Med Biol* 56, pp. R113–R144 (2011).

Besl P, McKay N. A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell* 14, pp. 239-256 (1992).

Bissonnette JP, Moseley D, White E, et al. Quality assurance for the geometric accuracy of cone-beam CT guidance in radiation therapy. *Int J Radiat Oncol Biol Phys* 71, pp. S57-S61 (2008).

Boldea V, Sharp GC, Jiang SB, Sarrut D. 4D-CT lung motion estimation with deformable registration: quantification of motion nonlinearity and hysteresis. *Med Phys* 35, pp. 1008-1018 (2008).

Bookstein F. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE T Pat An Mach Int* 11, pp. 567-585 (1989).

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

Bouilhol G, Ayadi M, Rit S, et al. Is abdominal compression useful in lung stereotactic body radiation therapy? A 4DCT and dosimetric lobe-dependent study. *Phys Med*, in press (2012).

Bova FJ, Buatti JM, Friedman WA, et al. The University of Florida frameless high-precision stereotactic radiotherapy system. *Int J Radiat Oncol Biol Phys* 38, pp. 875-882 (1997).

Brahme A, Nyman P, Skatt B. 4D laser camera for accurate patient positioning, collision avoidance, image fusion and adaptive approaches during diagnostic and therapeutic procedures. *Med Phys* 35, pp. 1670-1681 (2008).

Briechle K, Hanebeck UD. Template matching using fast normalized cross correlation. *Proc XII Optical Pattern Recognition* 95 (2001).

Brock KK, Deformable Registration Accuracy Consortium. Results of a multi-institution deformable registration accuracy study (MIDRAS). *Int J Radiat Oncol Biol Phys* 76, pp. 583-596 (2010).

Brown WT, Wu X, Fayad F, et al. Application of robotic stereotactic radiotherapy to peripheral stage I non-small cell lung cancer with curative intent. *Clin Oncol (R Coll Radiol)* 21, pp. 623-631 (2009).

Buzurovic I, Huang K, Yu Y, Podder TK. A robotic approach to 4D real-time tumor tracking for radiotherapy. *Phys Med Biol* 56, pp. 1299-1318 (2011).

Cachier P, Ayache N. Isotropic energies, filters and splines for vectorial regularization. *J Math Imaging Vis* 20, pp. 251-265 (2004).

Cala SJ, Kenyon CM, Ferrigno G, et al. Chest wall and lung volume estimation by optical reflectance motion analysis. *J Appl Physiol* 81, pp. 2680-2689 (1996).

Cerviño LI, Chao AK, Sandhu A, Jiang SB. The diaphragm as an anatomic surrogate for lung tumor motion. *Phys Med Biol* 54, pp. 3529-3541 (2009a).

Cerviño LI, Gupta S, Rose MA, et al. Using surface imaging and visual coaching to improve the reproducibility and stability of deep inspiration breath hold for left-breast-cancer radiotherapy. *Phys Med Biol* 54, pp. 6853-6865 (2009b).

Chamberland M, Wassenaar R, Spencer B, Xu T. Performance valuation of real-time motion tracking using positron emission fiducial markers. *Med Phys* 38, pp. 810-819 (2011).

Chen J, Lee RJ, Handrahan D, Sause WT. Intensity-modulated radiotherapy using implanted fiducial markers with daily portal imaging: Assessment of prostate organ motion. *Int J Radiat Oncol Biol Phys* 68, pp. 912-919 (2007).

Cheung PCF, Sixel KE, Tirona R, Ung YC. Reproducibility of lung tumor position and reduction of lung mass within the planning target volume using active breathing control (ABC). *Int J Radiat Oncol Biol Phys* 57, pp. 1437-1442 (2003).

Chew LP. Guaranteed-Quality Mesh Generation for Curved Surfaces. *Proc IX Annual Symposium on Computational Geometry*, pp. 274-280 (1993).

Chu W, Ludewigt B, Renner T. Instrumentation for treatment of cancer using proton and light-ion beams. *Rev Sci Instrum* 64, pp. 2055-2122 (1993).

Cui Y, Dy JG, Alexander B, Jiang SB. Fluoroscopic gating without implanted fiducial markers for lung cancer radiotherapy based on support vector machines (SVM). *Phys Med Biol* 53, pp. N315-N327 (2008).

Depauw N, Seco J. Sensitivity study of proton radiography and comparison with kV and MV x-ray imaging using GEANT4 Monte Carlo simulations. *Phys Med Biol* 56, pp. 2407-2421 (2011).

Depuydt T, Verellen D, Haas O, et al. Geometric accuracy of a novel gimbals based radiation therapy tumor tracking system. *Radiother Oncol* 98, pp. 365-372 (2011).

Dunn JC. Well separated clusters and optimal fuzzy partitions. *J Cybern* 4, pp. 95-104 (1974).

Durante M, Loeffler JS. Charged particles in radiation oncology. *Nat Rev Clin Oncol* 7, pp. 37-43 (2010).

Evans PM. Anatomical imaging for radiotherapy. *Phys Med Biol* 53, pp. R151 (2008).

Fallone BG, Murray B, Rathee S, et al. First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system. *Med Phys* 36, pp. 2084-2088 (2009).

Fang Q, Boas D. Tetrahedral mesh generation from volumetric binary and gray-scale images. *Proc IEEE International Symposium on Biomedical Imaging*, pp. 1142-1145 (2009).

Fassi A, Schaerer J, Riboldi M, et al. A novel CT-based contrast enhancement technique for markerless lung tumor tracking in X-ray projection images. *Radiother Oncol* 99, pp. S217 (2011).

Feldkamp LA, Davis LC, Kress JW. Practical cone beam algorithm. *J Opt Soc Am A* 1, pp. 612-619 (1984).

Feldmar J, Ayache N. Rigid, affine and locally affine registration of free-form surfaces. *Int J Comput Vis* 18, pp. 99-119 (1996).

Ferlay J, Shin HR, Bray F, et al. GLOBOCAN 2008 v2.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base N° 10. International Agency for Research on Cancer, Lyon, France [http://globocan.iarc.fr] (2008).

Ferrigno G, Borghese NA, Pedotti A. Pattern recognition in 3-D motion analysis. *ISPRS Photogram Remote Sensing* 45, pp. 227-246 (1990).

Ferrigno G, Carnevali P, Aliverti A, et al. Three-dimensional optical analysis of chest wall motion. *J Appl Physiol* 77, pp. 1224-1231 (1994).

Fischer B, Modersitzki J. Ill-posed medicine - an introduction to image registration. *Inverse Problems* 24 (2008).

Ford EC, Mageras GS, Yorke E, et al. Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging. *Int J Radiat Oncol Biol Phys* 52, pp. 522-531 (2002).

Furukawa T, Inaniwa T, Sato S, et al. Design study of a raster scanning system for moving target irradiation in heavy-ion radiotherapy. *Med Phys* 34, pp. 1085-1097 (2007).

Furukawa T, Inaniwa T, Sato S, et al. Moving target irradiation with fast rescanning and gating in particle therapy. *Med Phys* 37, pp. 4874-4879 (2010).

Gabor D. Theory of communications. *Journal of Institution of Electrical Engineers* 93, pp. 429-457 (1946).

Gassa F, Biston MC, Malet C, et al. The stereotactic body radiation therapy: initiation and clinical program. *Cancer Radiother* 10, pp. 444-450 (2006).

Gianoli C, Riboldi M, Spadea MF, et al. A multiple points method for 4D CT image sorting. *Med Phys* 38, pp. 656-667 (2011).

Giraud P, Yorke E, Jiang S, et al. Reduction of organ motion effects in IMRT and conformal 3D radiation delivery by using gating and tracking techniques. *Cancer Radiother* 10, pp. 269-282 (2006).

Greco C, Wolden S. Current status of radiotherapy with protons and light ion beams. *Cancer* 109, pp. 1227-1238 (2007).

Grözinger SO, Bert C, Haberer T, et al. Motion compensation with a scanned ion beam: a technical feasibility study. *Radiat Oncol* 14, pp. 3-34 (2008).

Haack H, Kunth W, Kollhoff R, et al. Introduction of a breath gating system in a clinical 342 environment: a practical description of our experiences. *Proc World Congress on Medical Physics and Biomedical Engineering* 25, pp. 748-749 (2009).

Haberer T, Becher W, Schardt D, et al. Magnetic scanning system for heavy ion therapy. *Nucl Instrum Methods A* 330, pp. 296-305 (1993).

Hanley J, Debois MM, Mah D, et al. Deep inspiration breath hold technique for lung tumors: the potential value of target immobilization and reduced lung density in dose escalation. *Int J Radiat Oncol Biol Phys* 45, pp. 603-611 (1999).

Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. *Appl Stat* 28, pp. 100-108 (1979).

Herman MG. Clinical use of electronic portal imaging. *Semin Radiat Oncol* 15, pp. 157-167 (2005).

Hiraoka M, Matsuo Y, Nagata Y. Stereotactic body radiation therapy (SBRT) for early-stage lung cancer. *Cancer Radiother* 11, pp. 32-35 (2007).

Hoisak JD, Sixel KE, Tirona R, et al. Correlation of lung tumour motion with external surrogate indicators of respiration. *Int J Radiat Oncol Biol Phys* 60, pp. 1298-1306 (2004).

Hoogeman M, Prévost JB, Nuyttens J, et al. Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. *Int J Radiat Oncol Biol Phys* 74, pp. 297-303 (2009).

Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2009 (Vintage 2009 Populations), National Cancer Institute, Bethesda, MD [http://seer.cancer.gov] (2011).

Hsu A, Miller NR, Evans PM, et al. Feasibility of using ultrasound for real-time tracking during radiotherapy. *Med Phys* 32, pp. 1500-1512 (2005).

Hu S, Hoffman EA, Reinhardt JM. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. *IEEE Trans Med Imaging* 20, pp. 490-498 (2001).

Huang NE, Shen Z, Long SR, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond A* 454, pp. 903-995 (1998).

Hughes S, McClelland J, Tarte S, et al. Assessment of two novel ventilatory surrogates for use in the delivery of gated/tracked radiotherapy for non-small cell lung cancer. *Radiother Oncol* 91, pp. 336-341 (2009).

Hugo GD, Liang J, Yan D. Marker-free lung tumour trajectory estimation from a cone beam CT sinogram. *Phys Med Biol* 55, pp. 2637-2650 (2010).

Ibáñez L, Schroeder W, Ng L, et al. The ITK Software Guide, Second Edition, Updated for ITK version 2.4. New York: Kitware (2005).

ICRU Report 62. Prescribing, Recording and Reporting Photon Beam Therapy (supplement to ICRU Report 50). International Commission on Radiation Units and Measurements, Bethesda, MD (1999).

Imura M, Yamazaki K, Shirato H, et al. Insertion and fixation of fiducial markers for setup and tracking of lung tumors in radiotherapy. *Int J Radiat Oncol Biol Phys* 63, pp. 1442-1447 (2005).

Jaffray DA, Drake DG, Moreau M, et al. A radiographic and tomographic imaging system integrated into a medical linear accelerator for localization of bone and soft-tissue targets. *Int J Radiat Oncol Biol Phys* 45, pp. 773-789 (1999).

Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys* 53, pp. 1337-1349 (2002).

Jaffray DA. Image-guided radiation therapy: From concept to practice. *Semin Radiat Oncol* 17, pp. 243-244 (2007).

Johnson U, Landau D, Lindgren-Turner J, et al. Real Time 3D Surface Imaging for the Analysis of Respiratory Motion during Radiotherapy. *Int J Radiat Oncol Biol Phys* 60, pp. S603-S604 (2004).

Jolliffe IT. Principal Component Analysis, second edition. *Springer* (2002).

Kalender WA. X-ray computed tomography. *Phys Med Biol* 51, pp. R29-R43 (2006).

Kamino Y, Takayama K, Kokubo M, et al. Development of a four-dimensional image-guided radiotherapy system with a gimbaled X-ray head. *Int J Radiat Oncol Biol Phys* 66, pp. 271-78 (2006).

Kanoulas E, Aslam JA, Sharp GC, et al. Derivation of the tumor position from external respiratory surrogates with periodical updating of the internal/external correlation. *Phys Med Biol* 52, pp. 5443-5456 (2007).

Keall PJ, Kini VR, Vedam SS, Mohan R. Motion adaptive x-ray therapy: a feasibility study. *Phys Med Biol* 46, pp. 1-10 (2001).

Keall PJ, Kini VR, Vedam SS, Mohan R. Potential radiotherapy improvements with respiratory gating. *Australas Phys Eng Sci Med* 25, pp. 1-6 (2002).

Keall PJ, Todor AD, Vedam SS, et al. On the use of EPID-based implanted marker tracking for 4D radiotherapy. *Med Phys* 31, pp. 3492-3499 (2004).

Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys* 33, pp. 3874-3900 (2006).

Kilby W, Dooley JR, Kuduvalli G, et al. The CyberKnife Robotic Radiosurgery System in 2010. *Technol Cancer Res Treat* 9, pp. 433-452 (2010).

Kim DJ, Murray BR, Halperin R, Roa WH. Held-breath self-gating technique for radiotherapy of non-small-cell lung cancer: a feasibility study. *Int J Radiat Oncol Biol Phys* 49, pp. 43-49 (2001).

Koch N, Liu HH, Starkschall G, et al. Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part I-correlating internal lung motion with skin fiducial motion. *Int J Radiat Oncol Biol Phys* 60, pp.1459-1472 (2004).

Kohonen T. Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics* 43, pp. 59-69 (1982).

Koto M, Miyamoto T, Yamamoto N, et al. Local control and recurrence of stage I non-small cell lung cancer after carbon ion radiotherapy. *Radiother Oncol* 71, pp. 147-156 (2004).

Krämer M, Scholz M. Treatment planning for heavy-ion radiotherapy: calculation and optimization of biologically effective dose. *Phys Med Biol* 45, pp. 3319-3330 (2000).

Kraft G. The radiobiological and physical basis of radiotherapy with protons and heavier ions. *Strahlenther Onkol* 166, pp. 10-13 (1990).

Kubo HD, Wang L. Compatibility of Varian 2100C gated operations with enhanced dynamic wedge and IMRT dose delivery. *Med Phys* 27, pp. 1732-1738 (2000).

Kwa SL, Lebesque JV, Theuws JC, et al. Radiation pneumonitis as a function of mean lung dose: an analysis of pooled data of 540 patients. *Int J Radiat Oncol Biol Phys* 42, pp. 1-9 (1998).

Kybic J, Unser M. Fast Parametric Elastic Image Registration. *IEEE T Im Proc* 12, pp. 1427-1442 (2003).

Lagerwaard FJ, Haasbeek CJ, Smit EF, et al. Outcomes of risk-adapted fractionated stereotactic radiotherapy for stage I non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 70, pp. 685-692 (2008).

Lambert J, Suchowerska N, McKenzie DR, Jackson M. Intrafractional motion during proton beam scanning. *Phys Med Biol* 50, pp. 4853-4862 (2005).

Lehmann J, Perks J, Semon S, et al. Commissioning experience with cone-beam computed tomography for image-guided radiation therapy. *J Appl Clin Med Phys* 8, pp. 21-36 (2007).

Lewis JH, Li R, Watkins WT, et al. Markerless lung tumor tracking and trajectory reconstruction using rotational cone-beam projections: A feasibility study. *Phys Med Biol* 55, pp. 2505-2522 (2010).

Lewis JP. Fast Normalized Cross-Correlation. *Vision Interface* 95, pp. 120-123 (1995).

Lin T, Li R, Tang X, et al. Markerless gating for lung cancer radiotherapy based on machine learning techniques. *Phys Med Biol* 54, pp. 1555-1563 (2009).

Liu HH, Balter P, Tutt T, et al. Assessing respiration-induced tumor motion and internal target volume using four-dimensional computed tomography for radiotherapy of lung cancer. *Int J Radiat Oncol Biol Phys* 68, pp. 531-540 (2007).

Low DA, Nystrom M, Kalinin E, et al. A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. *Med Phys* 30, pp. 1254-1263 (2003).

Lu W, Low DA, Parikh P J, et al. Comparison of spirometry and abdominal height as four-dimensional computed tomography metrics in lung. *Med Phys* 32, pp. 2351-2357 (2005).

Lu W, Parikh PJ, Hubenschmidt JP, et al. A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT. *Med Phys* 33, pp. 2964-2974 (2006).

Lu HM, Brett R, Sharp G, et al. A respiratory-gated treatment system for proton therapy. *Med Phys* 34, pp. 3273-3278 (2007).

Lujan AE, Larsen EW, Balter JM, Ten Haken RK. A method for incorporating organ motion due to breathing into 3D dose calculations. *Med Phys* 26, pp. 715-720 (1999).

MacQueen JB, Some Methods for classification and Analysis of Multivariate Observations. *Proc V Berkeley Symposium on Mathematical Statistics and Probability* 1, pp. 281-297 (1967).

Mageras GS, Yorke E. Deep inspiration breath hold and respiratory gating strategies for reducing organ motion in radiation treatment. *Semin Radiat Oncol* 14, pp. 65-75 (2004).

Mah D, Hanley J, Rosenzweig KE, et al. Technical aspects of the deep inspiration breath-hold technique in the treatment of thoracic cancer. *Int J Radiat Oncol Biol Phys* 48, pp. 1175-1185 (2000).

Maintz JB, Viergever MA. A survey of medical image registration. *Med Image Anal* 2, pp. 1-36 (1998).

Mattes D, Haynor DR, Vesselle H. Nonrigid multimodality image registration. *Medical Imaging 2001: Image Processing*, pp. 1609-1620 (2001).

McClelland JR, Blackall JM, Tarte S, Chandler AC, et al. A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. *Med Phys* 33, pp. 3348-3358 (2006).

Meeks SL, Tomé WA, Willoughby TR, et al. Optically guided patient positioning techniques. *Semin Radiat Oncol* 15, pp. 192-201 (2005).

Meyer J, Wilbert J, Baier K, et al. Positioning accuracy of cone-beam computed tomography in combination with a HexaPOD robot treatment table. *Int J Radiat Oncol Biol Phys* 67, pp. 1220-1228 (2007).

Minohara S, Kanai T, Endo M, et al. Respiratory gated irradiation system for heavy-ion radiotherapy. *Int J Radiat Oncol Biol Phys* 47, pp. 1097-1103 (2000).

Moore C, Lilley F, Sauret V, et al. Opto-electronic sensing of body surface topology changes during radiotherapy for rectal cancer. *Int J Radiat Oncol Biol Phys* 56, pp. 248-258 (2003).

Moser T, Fleischhacker S, Schubert K, et al. Technical performance of a commercial laser surface scanning system for patient setup correction in radiotherapy. *Phys Med* 27, pp. 224-232 (2011).

Muralidhar KR, Murthy PN, Mahadev DS, et al. Magnitude of shift of tumor position as a function of moderated deep inspiration breath hold: An analysis of pooled data of lung patients with active breath control in image-guided radiotherapy. *J Med Phys* 33, pp. 147-153 (2008).

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

Murphy MJ, Balter J, Balter S, et al. The management of imaging dose during image-guided radiotherapy: report of the AAPM Task Group 75. *Med Phys* 34, pp. 4041-4063 (2007).

Negoro Y, Nagata Y, Aoki T, et al. The effectiveness of an immobilization device in conformal radiotherapy for lung tumor: reduction of respiratory tumor movement and evaluation of the daily setup accuracy. *Int J Radiat Oncol Biol Phys* 50, pp. 889-898 (2001).

Neicu T, Shirato H, Seppenwoolde Y, Jiang SB. Synchronized moving aperture radiation therapy (smart): average tumour trajectory for lung patients. *Phys Med Biol* 48, pp. 587-598 (2003).

Nestle U, Weber W, Hentschel M, Grosu AL. Biological imaging in radiation therapy: role of positron emission tomography. *Phys Med Biol* 54, pp. R1-R25 (2009).

Newhauser WD, Koch NC, Fontenot JD, et al. Dosimetric impact of tantalum markers used in the treatment of uveal melanoma with proton beam therapy. *Phys Med Biol* 52, pp. 3979-3990 (2007).

Nuyttens JJ, van der Voort van Zyp NC, Praag J, et al. Outcome of four-dimensional stereotactic radiotherapy for centrally located lung tumors. *Radiother Oncol* 102, pp. 383-387 (2012).

Ohara K, Okumura T, Akisada M, et al. Irradiation synchronized with respiration gate. *Int J Radiat Oncol Biol Phys* 17, pp. 853-857 (1989).

Orban de Xivry J, Janssens G, Bosmans G, et al. Tumour delineation and cumulative dose computation in radiotherapy based on deformable registration of respiratory correlated CT images of lung cancer patients. *Radiother Oncol* 85, pp. 232-238 (2007).

Orecchia R, Zurlo A, Loasses A, et al. Particle Beam Therapy (Hadrontherapy): Basis for Interest and Clinical Experience. *Eur J Cancer* 34, pp. 459-468 (1998).

Orecchia R, Fossati P, Rossi S. The National Centre for Oncological Hadron Therapy: Status of the Project and Future Clinical Use of the Facility. *Tumori* 95, pp. 169-176 (2009).

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

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

Pavlovic M. Beam-optics study of the gantry beam delivery system for light-ion cancer therapy. *Nucl Instrum Methods Phys Res A* 399, pp. 439-454 (1997).

Pedersen AN, Korreman S, Nyström H, Specht L. Breathing adapted radiotherapy of breast cancer: reduction of cardiac and pulmonary doses using voluntary inspiration breath-hold. *Radiother Oncol* 72, pp. 53-60 (2004).

Pedroni E, Bacher R, Blattmann H, et al. The 200MeV proton therapy project at the Paul Scherrer Institute: conceptual design and practical realisation. *Med Phys* 22, pp. 37-53 (1995).

Pedroni E, Bearpark R, Böhringer T, Coray A. The PSI Gantry 2: A second generation proton scanning gantry. Z *Med Phys* 14, pp. 25-34 (2004).

Pepin EW, Wu H, Sandison GA, et al. Site-specific volumetric analysis of lung tumour motion. *Phys Med Biol* 55, pp. 3325-3337 (2010).

Phillips MH, Pedroni E, Blattmann H, et al. Effects of respiratory motion on dose uniformity with a charged particle scanning method. *Phys Med Biol* 37, pp. 223-234 (1992).

PTCOG. Particle Therapy Cooperative Group: an organization for those interested in proton, light ion and heavy charged particle radiotherapy [http://ptcog.web.psi.ch] (2011).

Raaymakers BW, Raaijmakers AJ, Lagendijk JJ. Feasibility of MRI guided proton therapy: magnetic field dose effects. *Phys Med Biol* 53, pp. 5615-5622 (2008).

Remouchamps VM, Letts N, Vicini FA, et al. Initial clinical experience with moderate deep inspiration breath hold using an active breathing control device in the treatment of patients with left-sided breast cancer using external beam radiation therapy. *Int J Radiat Oncol Biol Phys* 56, pp. 704-715 (2003).

Riboldi M, Gierga DP, Chen GT, Baroni G. Accuracy in breast shape alignment with 3D surface fitting algorithms. *Med Phys* 36, pp. 1193-1198 (2009).

Riboldi M, Orecchia R, Baroni G. Real-time tumour tracking in particle therapy: technological developments and future perspectives. *Lancet Oncol* 13, pp. e383-e391 (2012).

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

Rietzel E, Bert C. Respiratory motion management in particle therapy. *Med Phys* 37, pp. 449-460 (2010).

Rit S, Wolthaus JW, van Herk M, Sonke JJ. On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion. *Med Phys* 36, pp. 2283-2296 (2009).

Rit S, Pinho R, Delmon V, et al. VV, a 4D slicer. *Proc Int Workshop on Pulmonary Image Analysis*, pp. 171-175 (2011).

Rohde GK, Aldroubi A, Dawant BM. The adaptive bases algorithm for intensity-based nonrigid image registration. *IEEE Trans Med Imaging* 22, pp. 1470-1479 (2003).

Romei M, Mauro AL, D'Angelo MG, et al. Effects of gender and posture on thoraco-abdominal kinematics during quiet breathing in healthy adults. *Respir Physiol Neurobiol* 172, pp. 184-191 (2010).

Rottmann J, Aristophanous M, Chen A, et al. A multi-region algorithm for markerless beam's-eye view lung tumor tracking. *Phys Med Biol* 55, pp. 5585-5598 (2010).

Rousseeuw P. Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *J Comput Appl Math* 20, pp. 53-65 (1987).

Ruan D, Fessler JA, Balter JM. Mean position tracking of respiratory motion. *Med Phys* 35, pp. 782-792 (2008).

Ruan D, Fessler JA, Balter JM, Keall PJ. Real-time profiling of respiratory motion: baseline drift, frequency variation and fundamental pattern change. *Phys Med Biol* 54, 4777-4792 (2009).

Ruppert J. A Delaunay Refinement Algorithm for Quality 2-Dimensional Mesh Generation. *Journal of Algorithms* 18, pp. 548-585 (1995).

Saito N, Bert C, Chaudhri N, et al. Speed and accuracy of a beam tracking system for treatment of moving targets with scanned ion beams. *Phys Med Biol* 54, pp. 4849-4862 (2009).

Sarrut D. Deformable registration for image-guided radiation therapy. *Z Med Phys* 16, pp. 285-297 (2006a).

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

Sarrut D, Delhay B, Villard PF, et al. A comparison framework for breathing motion estimation methods from 4-D imaging. *IEEE Trans Med Imaging* 26, pp. 1636-1648 (2007).

Schaerer J, Fassi A, Riboldi M, et al. Multi-dimensional respiratory motion tracking from markerless optical surface imaging based on deformable mesh registration. *Phys Med Biol* 57, pp. 357-373 (2012).

Schardt D, Elsässer T, Schulz-Ertner D. Heavy-ion tumor therapy: Physical and radiobiological benefits. *Rev Mod Phys* 82, pp. 383-425 (2010).

Schlosser J, Salisbury K, Hristov D. Telerobotic system concept for real-time soft-tissue imaging during radiotherapy beam delivery. *Med Phys* 37, pp. 6357-6367 (2010).

Schöffel PJ, Harms W, Sroka-Perez G, et al. Accuracy of a commercial optical 3D surface imaging system for realignment of patients for radiotherapy of the thorax. *Phys Med Biol* 52, pp. 3949-3963 (2007).

Schreibmann E, Chen GT, Xing L. Image interpolation in 4D CT using a BSpline deformable registration model. *Int J Radiat Oncol Biol Phys* 64, pp. 1537-1550 (2006).

Schroeder W, Martin K, Lorensen B. The Visualization Toolkit: An Object-Oriented Approach To 3D Graphics, 4th Edition. New York: Kitware (2006).

Schulz-Ertner D, Jäkel O, Schlegel W. Radiation therapy with charged particles. *Semin Radiat Oncol* 16, pp. 249-259 (2006).

Schweikard A, Glosser G, Bodduluri M, et al. Robotic motion compensation for respiratory movement during radiosurgery. *Comput Aided Surg* 5, pp. 263-277 (2000).

Seco J, Robertson D, Trofimov A, Paganetti H. Breathing interplay effects during proton beam scanning: simulation and statistical analysis. *Phys Med Biol* 54, pp. N283-N294 (2009).

Segars W. Development and application of the new dynamic NURBS-based cardiac-torso (NCAT) phantom. PhD thesis, The University of North Carolina at Chapel Hill (2001).

Seppenwoolde Y, Shirato H, Kitamura K, et al. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol Biol Phys* 53, pp. 822-834 (2002).

Seppenwoolde Y, Berbeco RI, Nishioka S, et al. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Med Phys* 34, pp. 2774-2784 (2007).

Seregni M, Pella A, Riboldi M, et al. Real-time tumor tracking with an artificial neural networks-based method: A feasibility study. *Phys Med*, in press (2011).

Shackleford JA, Kandasamy N, Sharp GC. On developing B-spline registration algorithms for multi-core processors. *Phys Med Biol* 55, pp. 6329-6351 (2010).

Sharp GC, Jiang SB, Shimizu S, Shirato H. Prediction of respiratory tumour motion for real-time image-guided radiotherapy. *Phys Med Biol* 49, pp. 425-440 (2004).

Sharp GC, Li R, Wolfgang J, et al. Plastimatch - An Open Source Software Suite for Radiotherapy Image Processing. *Proc XVI Int Conf on the Use of Computers in Radiation Therapy* (2009).

Sharp GC, Peroni M, Li R, et al. Evaluation of plastimatch B-Spline registration on the EMPIRE10 dataset. *Proc XIII Int Conf on Medical image computing and computer-assisted intervention* (2010).

Sharpe MB, Moseley DJ, Purdie TG, et al. The stability of mechanical calibration for a kV cone beam computed tomography system integrated with linear accelerator. *Med Phys* 33, pp. 136-144 (2006).

Sherouse GW, Novins K, Chaney EL. Computation of digitally reconstructed radiographs for use in radiotherapy treatment design. *Int J Radiat Oncol Biol Phys* 18, pp. 651-658 (1990).

Shewchuk JR. Delaunay refinement algorithms for triangular mesh generation. *Comput Geom* 22, pp. 21-74 (2002).

Shioyama Y, Tokuuye K, Okumura T, et al. Clinical evaluation of proton radiotherapy for non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 56, pp. 7-13 (2003).

Shirato H, Shimizu S, Kunieda T, et al. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *Int J Radiat Oncol Biol Phys* 48, pp. 1187-1195 (2000).

Shirato H, Seppenwoolde Y, Kitamura K, et al. Intrafractional tumor motion: lung and liver. *Semin Radiat Oncol* 14, pp. 10-18 (2004).

Shirato H, Onimaru R, Ishikawa M, et al. Real-time 4-D radiotherapy for lung cancer. *Cancer Sci* 103, pp. 1-6 (2012).

Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med Phys* 12, pp. 252-255 (1985).

Smith N, Meir I, Hale G, et al. Real-Time 3D Surface Imaging for Patient Positioning in Radiotherapy. *Int J Radiat Oncol Biol Phys* 57, pp. S187 (2003).

Smith RP, Bloch P, Harris EE, et al. Analysis of interfraction and intrafraction variation during tangential breast irradiation with an electronic portal imaging device. *Int J Radiat Oncol Biol Phys* 62, pp. 373-378 (2005).

Sonke JJ, Zijp L, Remeijer P, van Herk M. Respiratory correlated cone beam CT. *Med Phys* 32, pp. 1176-1186 (2005).

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

Spadea MF, Baroni G, Riboldi M, et al. Patient set-up verification by infrared optical localization and body surface sensing in breast radiation therapy. *Radiother Oncol* 79, pp. 170-178 (2006).

Stevens CW, Munden RF, Forster KM, et al. Respiratory-driven lung tumor motion is independent of tumor size, tumor location, and pulmonary function. *Int J Radiat Oncol Biol Phys* 51, pp. 62-68 (2001).

Stock M, Kontrisova K, Dieckmann K, et al. Development and application of a real-time monitoring and feedback system for deep inspiration breath hold based on external marker tracking. *Med Phys* 33, pp. 2868-2877 (2006).

Sweeney RA, Arnold W, Steixner E, et al. Compensating for tumor motion by a 6-degree-of-freedom treatment couch: is patient tolerance an issue? *Int J Radiat Oncol Biol Phys* 74, pp. 168-171 (2009).

Tarte S, McClelland J, Hughes S, et al. A Non-Contact Method for the Acquisition of Breathing Signals that Enable Distinction Between Abdominal and Thoracic Breathing. *Radiother Oncol* 81, pp. S209 (2006).

Timmerman R, Paulus R, Galvin J, et al. Stereotactic body radiation therapy for inoperable early stage lung cancer. *JAMA* 303, pp. 1070-1076 (2010).

Torshabi AE, Pella A, Riboldi M, Baroni G. Targeting accuracy in real-time tumor tracking via external surrogates: a comparative study. *Technol Cancer Res Treat* 9, pp. 551-562 (2010).

Van der Weide L, Van Sörnsen de Koste JR, Lagerwaard FJ. Analysis of carina position as surrogate marker for delivering phase-gated radiotherapy. *Int J Radiat Oncol Biol Phys* 71, 1111-1117 (2008).

Van Herk M. Errors and margins in radiotherapy. *Semin Radiat Oncol* 14, pp. 52-64 (2004).

Van Hulle MM. Topology-preserving map formation achieved with a purely local unsupervised competitive learning rule. *Neural Networks* 10, pp. 431-446 (1997).

Vedam SS, Keall PJ, Kini VR, Mohan R. Determining parameters for respiration-gated radiotherapy. *Med Phys* 28, pp. 2139-2146 (2001).

Vedam SS, Keall PJ, Kini VR, et al. Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol* 48, pp. 45-62 (2003).

Vedam SS, Keall PJ, Docef A, et al. Predicting respiratory motion for four-dimensional radiotherapy. *Med Phys* 31, pp. 2274-2283 (2004).

Veninga T, Huisman H, Van Der Maazen RW, Huizenga H. Clinical validation of the normalized mutual information method for registration of CT and MR images in radiotherapy of brain tumors. *J Appl Clin Med Phys* 5, pp. 66-79 (2004).

Vikström J, Hjelstuen MH, Mjaaland I, Dybvik KI. Cardiac and pulmonary dose reduction for tangentially irradiated breast cancer, utilizing deep inspiration breath-hold with audio-visual guidance, without compromising target coverage. *Acta Oncol* 50, pp. 42-50 (2011).

Villard PF, Beuve M, Shariat B, et al. Simulation of lung behaviour with finite elements: influence of bio-mechanical parameters. *Proc Int Conf on Medical Information Visualisation - Biomedical Visualisation*, pp. 914 (2005).

Von Siebenthal M, Székely G, Gamper U, et al. 4D MR imaging of respiratory organ motion and its variability. *Phys Med Biol* 52, pp. 1547-1564 (2007).

Wang LT, Solberg TD, Medin PM, Boone R. Infrared patient positioning for stereotactic radiosurgery of extracranial tumors. *Comput Biol Med* 31, pp. 101-111 (2001).

Wang HK, Lu TW, Liing RJ, et al. Relationship between chest wall motion and diaphragmatic excursion in healthy adults in supine position. *J Formos Med Assoc* 108, pp. 577-586 (2009).

Webb S. Adapting IMRT delivery fraction-by-fraction to cater for variable intrafraction motion. *Phys Med Biol*, pp. 53 1-21 (2008).

Wilson RR. Radiological use of fast protons. *Radiology* 47, pp. 487-491 (1946).

Wink N, Panknin C, Solberg TD. Phase versus amplitude sorting of 4D-CT data. *J Appl Clin Med Phys* 7, pp. 77-85 (2006).

Wolthaus JW, Sonke JJ, van Herk M, Damen EM. Reconstruction of a time-averaged midposition CT scan for radiotherapy planning of lung cancer patients using deformable registration. *Med Phys* 35, pp. 3998-4011 (2008).

Wong W, Sharpe MB, Jaffray DA, Kini VR, Robertson JM, Stromberg JS, Martinez AA. The use of active breathing control (ABC) to reduce margin for breathing motion. *Int J Radiat Oncol Biol Phys* 44, pp. 911-919 (1999).

Wu MT, Chang JM, Chiang AA, et al. Use of quantitative CT to predict postoperative lung function in patients with lung cancer. *Radiology* 191, pp. 257-262 (1994).

Wu Z, Rietzel E, Boldea V, et al. Evaluation of deformable registration of patient lung 4DCT with subanatomical region segmentations. *Med Phys* 35, pp. 775-781 (2008).

Wulf J, Hädinger U, Oppitz U, et al. Stereotactic radiotherapy of extracranial targets: CT-simulation and accuracy of treatment in the stereotactic body frame. *Radiother Oncol* 57, pp. 225-236 (2000).

Wulf J, Baier K, Mueller G, Flentje MP. Dose-response in stereotactic irradiation of lung tumors. *Radiother Oncol* 77, pp. 83-87 (2005).

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

Yamamoto T, Langner U, Loo BW. Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. *Int J Radiat Oncol Biol Phys* 72, pp. 1250-1258 (2008).

Yan H, Yin FF, Zhu GP, et al. The correlation evaluation of a tumor tracking system using multiple external markers. *Med Phys* 33, pp. 4073-4084 (2006).

Yan H, Zhu G, Yang J, et al. The investigation on the location effect of external markers in respiratory-gated radiotherapy. *J Appl Clin Med Phys* 9, pp. 57-68 (2008).

Yang D, Lu W, Low DA, et al. 4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling. *Med Phys* 35, pp. 4577-4590 (2008).

Yang Y, Zhong Z, Guo X, et al. A novel markerless technique to evaluate daily lung tumor motion based on conventional cone-beam CT projection data. *Int J Radiat Oncol Biol Phys* 82, pp. e749-e756 (2012).

Yorke E, Rosenzweig KE, Wagman R, Mageras GS. Interfractional anatomic variation in patients treated with respiration-gated radiotherapy. *J Appl Clin Med Phys* 6, pp. 19-32 (2005).

Zenklusen SM, Pedroni E, Meer D. A study on repainting strategies for treating moderately moving targets with proton pencil beam scanning at the new Gantry 2 at PSI. *Phys Med Biol* 55, pp. 5103-5121 (2010).

Zhang Q, Pevsner A, Hertanto A, et al. A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Med Phys* 34, pp. 4772-4781 (2007).

Zhu C, Byrd RH, Lu P, Nocedal J. Algorithm 778. L-BFGS-B: Fortran subroutines for Large-Scale bound constrained optimization. *ACM Trans Math Softw* 23, pp. 550-60 (1997).

Zitova B, Flusser J. Image registration methods: a survey. *Image and Vision Computing* 21, pp. 977-1000 (2003).