Tumor tracking based on correlation models in scanned ion beam therapy: an experimental study

Tumor tracking based on correlation models in scanned ion beam therapy: an experimental study

M Seregni[1]**, R Kaderka**[3]**, G Fattori**[1]**, M Riboldi**[1,2]**, A Pella**[1]**,
A Constantinescu**[3]**, N Saito**[3]**, M Durante**[3,4]**, P Cerveri**[1,2]**, C Bert**[3,5]
and G Baroni[1,2]

[1] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano,
P.zza Leonardo da Vinci 32, I-20133 Milano, Italy
[2] Bioengineering Unit, Centro Nazionale di Adroterapia Oncologica, Strada privata Campeggi,
I-27100 Pavia, Italy
[3] GSI Helmholtzzentrum für Schwerionenforschung, Planckstr 1, D-64291 Darmstadt, Germany
[4] Technische Universität Darmstadt, Institut für Festkörperphysik, Hochschulstraße 6,
D-64289 Darmstadt, Germany
[5] Department of Radiation Oncology, University Hospital Erlangen, Universitätsstr 27, D-91054,
Erlangen, Germany

Abstract

Accurate dose delivery to extra-cranial lesions requires tumor motion
compensation. An effective compensation can be achieved by real-time tracking
of the target position, either measured in fluoroscopy or estimated through
correlation models as a function of external surrogate motion. In this work, we
integrated two internal/external correlation models (a state space model and an
artificial neural network-based model) into a custom infra-red optical tracking
system (OTS). Dedicated experiments were designed and conducted at GSI
(*Helmholtzzentrum für Schwerionenforschung*). A robotic breathing phantom
was used to reproduce regular and irregular internal target motion as well as
external thorax motion. The position of a set of markers placed on the phantom
thorax was measured with the OTS and used by the correlation models to infer
the internal target position in real-time. Finally, the estimated target position
was provided as input for the dynamic steering of a carbon ion beam. Geometric
results showed that the correlation models transversal (2D) targeting error was
always lower than 1.3 mm (root mean square). A significant decrease of the
dosimetric error with respect to the uncompensated irradiation was achieved
in four out of six experiments, demonstrating that phase shifts are the most
critical irregularity for external/internal correlation models.

(Some figures may appear in colour only in the online journal)

1. Introduction

In external beam radiation therapy, intra-fractional organ motion needs to be properly managed in order to preserve the effectiveness of the treatment and to avoid severe side effects (Langen and Jones 2001). For this purpose, different motion mitigation techniques have been proposed (Keall *et al* 2006, Bert and Durante 2011, Riboldi *et al* 2012). Among these, tumor tracking (Keall *et al* 2001) can be regarded as an effective strategy, since it allows the patient to breathe freely while the beam is dynamically steered to track the tumor motion along its whole trajectory. This strategy requires real-time motion monitoring: the internal target position can be directly measured by means of fluoroscopy (Shirato *et al* 1999, 2000, Shimizu *et al* 2001). This strategy, however, has limited application due to the relevant non-therapeutic imaging dose delivered to the patient (Shirato *et al* 2004). As a surrogate of the internal target motion, the movement of the external body surface (typically the chest or the abdominal wall) can be monitored by measuring the position of a single surface marker (Ford *et al* 2002), a set of surface markers (Baroni *et al* 1998) or of the whole surface (Schaerer *et al* 2012). In addition, hybrid strategies that combine the localization of multiple markers with surface acquisition were also proposed (Riboldi *et al* 2004). However, the temporal variability of the correlation between surrogates and the actual target motion may lead to potential inaccuracies. For this reason, techniques involving high frequency surrogate detection and low frequency internal monitoring were proposed (Schweikard *et al* 2000) and are clinically applied for x-ray radiotherapy in the Synchrony™ Respiratory Tracking System (Accuray Inc., Sunnyvale, CA) (Kilby *et al* 2010). This system couples a real-time surrogate detection, performed by optically localizing three infra-red (IR) LEDs placed on the patient's thorax or abdomen, with internal target localization obtained by stereoscopic episodic x-ray imaging. A linear or polynomial model is then used to learn the patient's fraction-specific internal/external correlation function and to estimate the target position in real-time during irradiation.

In principle, tumor tracking can be regarded as the most effective motion mitigation technique, since it allows full compensation of the target motion combined with continuous irradiation. Such advantages recommend its application in treatments demanding high accuracy, such as in particle therapy. This technique shows superior geometrical selectivity and increased relative biological effectiveness (at least for carbon ion beams) with respect to conventional photon radiotherapy (Kraft 2000, Durante and Loeffler 2010), providing a highly conformal dose deposition on the target volume and enhanced effectiveness on radiotherapy-resistant tumors.

A tumor tracking system for scanned carbon ions beams was developed at GSI (*Helmholtzzentrum für Schwerionenforschung,* Darmstadt, Germany) and integrated into the Therapy Control System (TCS) (Grözinger *et al* 2008, Saito *et al* 2009, Bert *et al* 2010). The system relies on scanning magnets to apply transversal compensation (i.e. on the plane perpendicular to the beam). Longitudinal compensation, which is required in order to take into account for motion-induced path length variations (Mori *et al* 2007, 2008), is performed by a dedicated range shifter, which is based on a double PMMA (polymethylmetacrilate) wedge driven by fast linear motors.

In the framework of the ULICE (Union of Light Ion Centres in Europe) project, experimental activities were planned to investigate the integration of optical motion monitoring with the beam tracking system developed at GSI. A robotic breathing phantom, consisting of a human-like artificial thorax (polyvinyl carbonate bones surrounded by rubber) coupled with a six degrees-of-freedom robotic arm, which simulates the target motion, was developed at GSI (Steidl *et al* 2012). The phantom is capable of producing either regular (i.e. stationary in space and time) or irregular motion patterns (i.e. non-stationary in space and/or time).

In addition, dedicated optical motion monitoring software was developed at Politecnico di Milano (Fattori *et al* 2012). This software was designed to provide real-time monitoring and tracking of IR-reflective markers placed on the phantom's thorax and robotic arm, in two main working modalities.

- External motion phase detection: in this modality, the software monitors only the thorax motion and provides to the TCS the current motion phase, as defined in the 4D treatment planning (Bert and Rietzel 2007). Then, the TCS uses a pre-calculated look up table (LUT) to obtain the correction vector as a function of the motion phase (Saito *et al* 2009).
- Real-time tracking: the software directly provides transversal compensation vectors to the TCS, either by optically measuring the position of the target (direct tracking) or by estimating its position through dedicated algorithms implementing internal/external correlation models. In either case, longitudinal compensation is still computed through motion phase detection and the pre-calculated LUT.

In the framework of these activities, the present work is focused on the development and experimental validation of the internal/external correlation models implemented into the optical motion monitoring software. Previous studies (Torshabi *et al* 2010) reported that more complex strategies than linear or polynomial models are required to achieve an accurate estimation of the internal motion based on external surrogates. For this reason, we chose to implement two different models that can be alternatively selected by the motion monitoring software:

- A state space model (SSM), based on the state augmentation approach proposed by (Ruan *et al* 2008). This model, characterized by a reduced computational cost, was selected to test the technical feasibility of the integration of correlation models into a framework for optically-guided tumor tracking.
- An artificial neural network (ANN) model, consisting of a nonlinear mathematical model already proposed for temporal prediction in applications related to radiation therapy (Sharp *et al* 2004, Murphy and Pokhrel 2009). The ANN model, characterized by specific generalization capabilities, was selected to test tumor tracking in experimental conditions focused on the simulation of clinical-like procedures.

The purpose of the work is to describe the experimental activities conducted at GSI to validate tumor tracking based on internal/external correlation models for a scanned carbon ion beam. Results concerning the geometric and dosimetric accuracy of the implemented tumor tracking models will be presented.

2. Materials and methods

In this section, the two implemented correlation models will be first introduced, and then the experiments, involving the application of model-based tumor tracking in phantom studies, will be described.

2.1. General workflow for real-time tumor tracking

The overall workflow for real-time tracking through correlation models is reported in figure 1 and described in the following.

Two input signals were provided to the correlation model:

- the external motion signal, i.e. the displacement of external surface markers along the anterior–posterior (AP) anatomical direction;
- the internal motion signal, i.e. the position of the target.
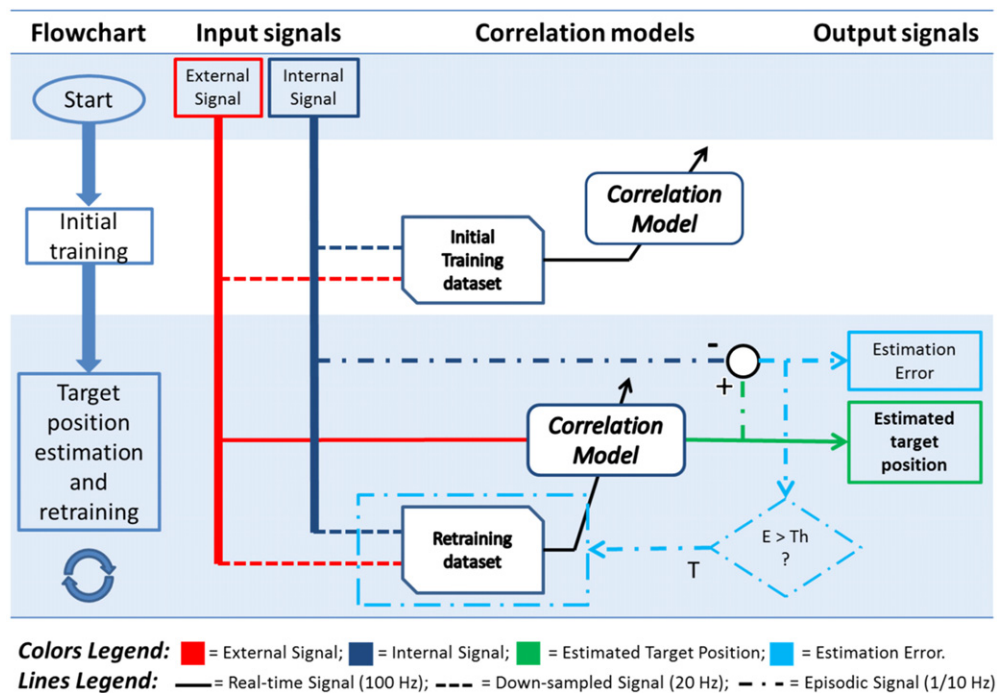
Figure 1. General workflow for real-time tracking by means of correlation models. In the left column a flowchart describes the two main working conditions of the correlation software. For each of these conditions, the inputs and the outputs of the correlation models, as well as their operations, are graphically reported. As reported in the legend, different colors are used to distinguish among different signals, whereas line style is used to represent the signal's frequency.

Both these signals were provided by the optical motion monitoring software in real-time (100 Hz). The correlation software, however, was able to simulate high-frequency and episodic imaging acquisitions by down-sampling the internal signal to 20 Hz and to 1/10 Hz, respectively.

The first step, which was required to perform tumor tracking by means of the correlation model, was the collection of a training dataset, i.e. a set of synchronized samples of the internal and of the external signals. In order to obtain detailed information on the signals dynamics, this dataset was stored simulating a high-frequency (20 Hz) imaging acquisition.

As soon as the training set was collected, the correlation model (either the SSM or the ANN) was trained. Such training provided the optimization of the model parameters and resulted in the learning of the internal/external correlation function from the information stored into the training dataset. Typical training time was less than 10 ms for the SSM and less than 3 s for the ANN.

As the training was completed, the model estimated the internal target position in real-time, i.e. in correspondence to each sample of the external signal (100 Hz). In order to check the estimation accuracy, episodic imaging was simulated by comparing the estimated target position with the actual one, which was obtained by localizing an additional reflective marker (figure 3) placed in proximity of the target but visible by the two TV Cameras. Due to the high rigidity of the target holder and to the absence of rotational motion, the reflective marker experienced the same motion of the target. As a consequence, the correlation error

could be reliably calculated. If this error exceeded a user-defined threshold, the retraining procedure was triggered: a new training dataset was acquired and the model's parameters were re-optimized. The retraining procedures were executed without stopping the real-time target position estimation and without interfering with other time-critical processes such as markers localization.

Finally, in order to compensate for delays in calculation and application of the correction vectors, time prediction was implemented into the main optical motion monitoring software by means of polynomial interpolation (Fattori *et al* 2012). The delay was quantified in 14.6 ms for transversal corrections and was mainly due to optical tracking system (OTS) processing and communication. For longitudinal correction, an additional 34.1 ms delay was considered to compensate also for the latency of the range shifter.

2.2. Correlation models

Two different correlation models were implemented for internal motion estimation: a SSM and an ANN.

2.2.1. State space model. The SSM was designed to correlate the external markers motion ($E(t)$) with the internal target motion ($T(t)$) by means of a linear function. In order to take into account the hysteretic target motion, the internal target position was correlated with both synchronized and delayed (by time lags $\tau/2$ and τ) external markers positions (Ruan *et al* 2008). The time lag τ was set to 1.5 s, which was half of the phantom respiratory period.

In this specific implementation, the internal signal relied on the position of the reflective marker placed on the target, whereas the external signal was the AP motion of a single external marker.

The internal motion was estimated according to the following equations:

$$\boldsymbol{T}(t) = A\boldsymbol{f}(t) \tag{1}$$

$$\boldsymbol{f}(t) = \begin{bmatrix} E(t) \\ E(t - \tau/2) \\ E(t - \tau) \\ 1 \end{bmatrix}, \tag{2}$$

where A is the correlation matrix. The training of the SSM consisted in the calculation, by means of least square minimization (LSM), of the elements of matrix A (equations (3) and (4)):

$$F = \begin{bmatrix} \boldsymbol{f}'_1 \\ \vdots \\ \boldsymbol{f}'_N \end{bmatrix}, \qquad T = \begin{bmatrix} \boldsymbol{T}'_1 \\ \vdots \\ \boldsymbol{T}'_N \end{bmatrix}, \tag{3}$$

$$\hat{A} = T'F(FF')^{-1} \tag{4}$$

where the couple (F, T) represents a training dataset containing N samples of the external state vector $\boldsymbol{f}(t)$ and of the internal signal $\boldsymbol{T}(t)$.

2.2.2. Artificial neural network model. Artificial neural networks are nonlinear computational models whose main advantage is the potential ability to learn complex functions without the need of any *a-priori* information. In the so-called 'supervised training', a dedicated algorithm is used to iteratively optimize the neural parameters (inter-synaptic weights and
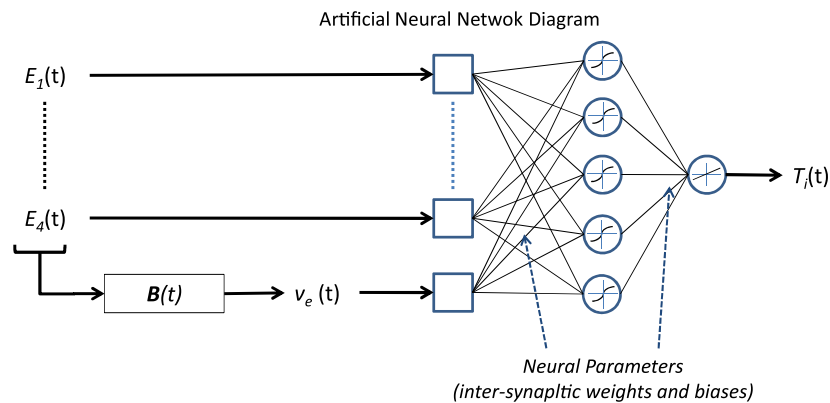
Figure 2. Diagram of the ANN used to correlate the internal and the external motion. Inputs signals, neural parameters and output signals can be observed. $E_i(t)$ indicates the external surrogate input, $B(t)$ the buffer used to compute the external velocity ($v_e(t)$). $T_i(t)$ represents the output of the ANN, i.e. the internal target position along the *i*th coordinate.

biases). This is obtained by minimizing a cost function expressing, for a given set of inputs, the difference between the actual outputs of the net and the desired outputs, as defined in the training dataset.

In our specific application, ANNs were used to learn the internal/external correlation and to estimate the internal target position as a function of the position of the external markers. We implemented an individual ANN for each motion coordinate (cranio-caudal (CC) and AP). Each ANN was configured as follows: (figure 2):

- Five input signals, which were the AP components of four external markers ($\boldsymbol{E}(t)$) along with the velocity $v_E(t)$ of their center of mass along the AP direction. In order to compute $v_E(t)$, a 'first in first out' buffer ($\boldsymbol{B}(t)$) was continuously updated and the velocity was obtained as the slope of the line fitting the points collected in the buffer. This fifth signal incorporated first order dynamic information into the correlation model, thus solving the ambiguity in the internal/external correlation function caused by hysteretic target motion (Ernst *et al* 2011, Langner and Keall 2009).
- A single hidden layer composed of five hidden units having a hyperbolic tangent activation function.
- One output unit with linear activation function, producing the estimation of the target position along one motion coordinate ($T_i(t)$).

The training was based on back-propagation, which was implemented through BFGS (Broyden–Fletcher–Goldfarb–Shanno) quasi-Newton method. Two alternative and concurrent termination criteria were defined:

- Number of iterations $\geqslant 120$
- root mean square (RMS) of the difference between the actual and the desired output $\leqslant 10^{-3}$.

This strategy was selected as a trade-off between accuracy and training time, the latter being usually less than 1 s and always less than 3 s in our experiments.

2.3. Experimental protocol

2.3.1. Systems involved in the experiment. The main systems involved in the presented experimental activity were:

- GSI Therapy Control System (Grözinger *et al* 2008, Saito *et al* 2009). The system provides transversal and longitudinal beam tracking, relying on scanning magnets and on a range shifter. Scanning magnets deflect the beam along the vertical and the horizontal direction, with an accuracy of 0.16 mm and a maximum scanning speed of 1 cm ms^{-1} (Haberer *et al* 1993). The range shifter consists in a double PMMA wedge controlled by fast servo motors. The range of the traversing ions is modulated by moving the wedges together or apart. For a 5 mm WE (Water Equivalent) shift, the correction is performed within 16 $\pm$ 2 ms and with an accuracy of 1 mm WE (for an irradiation time of 10 ms per raster point). Details concerning this device are reported in Saito *et al* (2009).
- OTS, based on TV cameras equipped with IR illumination apparatus (SMART DX-100, BTS Bioengineering, Italy). The system is able to localize markers at a frame rate of 100 Hz with a nominal accuracy lower than 0.2 mm in 4 $\times$ 3 $\times$ 3 m^3 working volume. Two calibrated TV-Cameras (sensor resolution 640 $\times$ 480 pixels) were utilized in the experiments.
- Custom software was integrated into the OTS (Fattori *et al* 2012) to perform real-time motion monitoring of the markers and to provide correction vectors for beam tracking to the TCS. The correlation models were integrated into the main motion monitoring software. In order to preserve time-critical processes involving markers localization and tracking, the ANN's training algorithm, which required significant computational resources and relatively long time (1–3 s), was executed on a dedicated machine (general purpose laptop equipped with a 2.2 GHz quad-core processor and 6 GB RAM) directly connected to the OTS workstation. Such a concern was negligible for the SSM training, since the LSM method was less computational intensive and required a much shorter execution time ($<$10 ms).
- Robotic breathing phantom for respiratory motion simulation (Steidl *et al* 2012). The system is able to simulate both the patient's body surface motion (by means of an artificial deformable thorax) and the internal tumor motion, obtained with six degrees-of-freedom industrial robotic arm. A single controller was responsible to drive both the stepper motor, actuating the thorax, and the robotic arm.
- Dose measuring equipment (Steidl *et al* 2012). The tool attached to the robotic arm end-effector consisted in a PMMA detector head, providing slots for 20 ionization chambers (PTW PinPoint® PTW Freiburg, Freiburg, Germany). Considering the beam eye view (BEV), the chambers where spaced by 10 mm in the horizontal direction and by 12 mm in the vertical direction. Chambers were read out by PTW MULTIDOS® Electrometer (PTW Freiburg, Freiburg, Germany) according to the protocol proposed by Karger *et al* (1999).

2.3.2. Experimental set-up. The experimental set-up is depicted in figure 3 and described hereafter:

The TV-Cameras of the OTS system were mounted on tripods and positioned alongside of the beam line, about 2.5 m high with respect to the floor. System calibration was performed immediately before the experiment, obtaining a 3D reconstruction error equal to 0.26 $\pm$ 0.25 mm (mean $\pm$ std. dev.) in a volume that included the artificial thorax and the detector head. Four reflective markers were attached to the phantom's thorax (figure 3): three markers were placed along its caudal end (diaphragm, E1, E2, E3) and one at about 10 cm from the diaphragm (E4). A fifth marker (T) was placed on the detector head.

The robotic phantom was positioned on a dedicated basement that was rigidly fixed to the four degrees-of-freedom treatment couch installed in the GSI Cave-M. In the nominal treatment position, the beam was delivered laterally with respect to the phantom's thorax
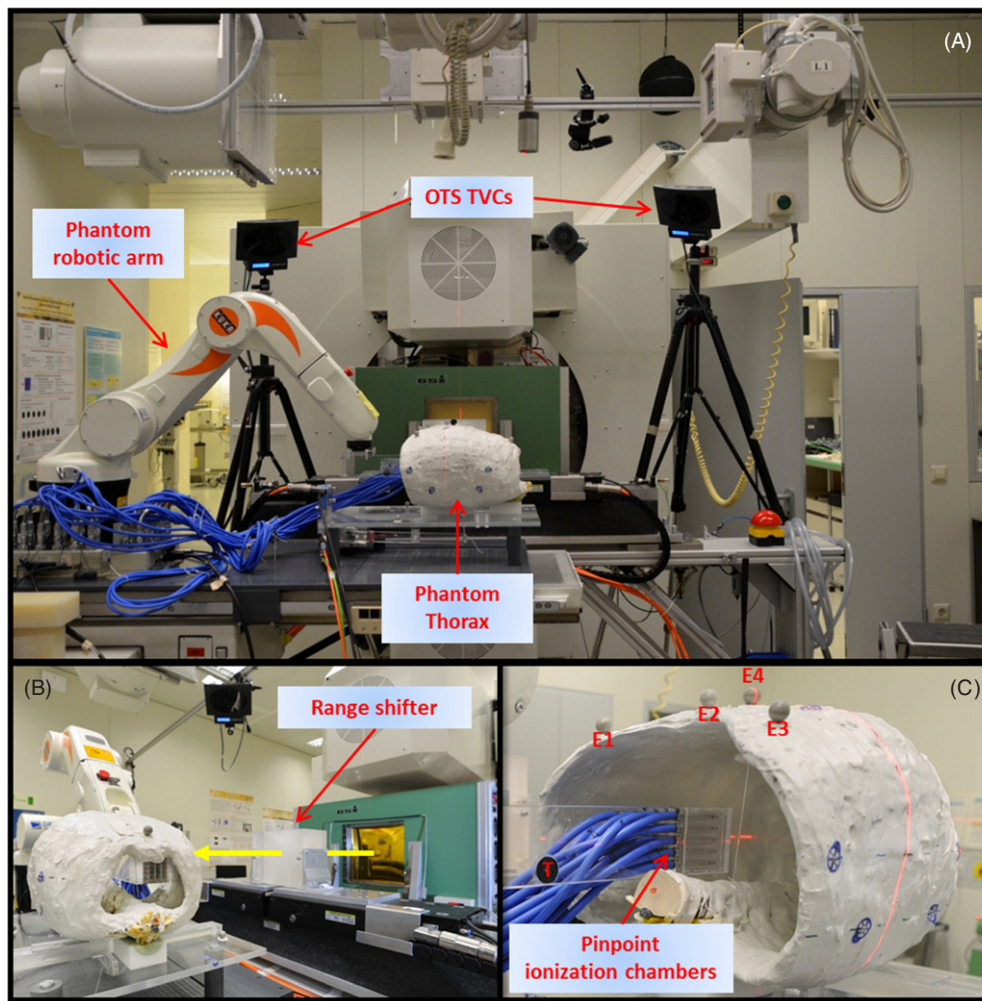
Figure 3. Experimental set-up built into GSI Cave-M. The two OTS TV-Cameras, the phantom robotic arm, and the phantom thorax are highlighted in panel A. In panels B and C the four thorax IR reflective markers (E1, . . . , E4) are labeled. The fifth IR marker (T) placed on the detector head is also labeled in panel C. The beam direction (yellow arrow) and the double wedge range shifter are pictured in panel B, whereas the PMMA detector head containing the 20 pinpoint ionization chambers can be observed in panel C. Pictures B and C: courtesy of G Otto (GSI).

(figure 3). Alignment to this position was achieved relying on isocentric laser lines (figure 3, panel C) and on the optical motion monitoring software, which allowed to perform stereotactic positioning (Baroni *et al* 2007).

2.3.3. Experiments description. The robotic phantom was able to simulate respiratory motion in three different conditions: regular motion, baseline drift and phase shift. Motion features are reported in table 1 for each of these conditions.

Internal motion consisted in a 3D sinusoidal motion (figure 4). Its amplitude and period were comparable to those observed for real lung tumors (Seppenwoolde *et al* 2002). Moreover,
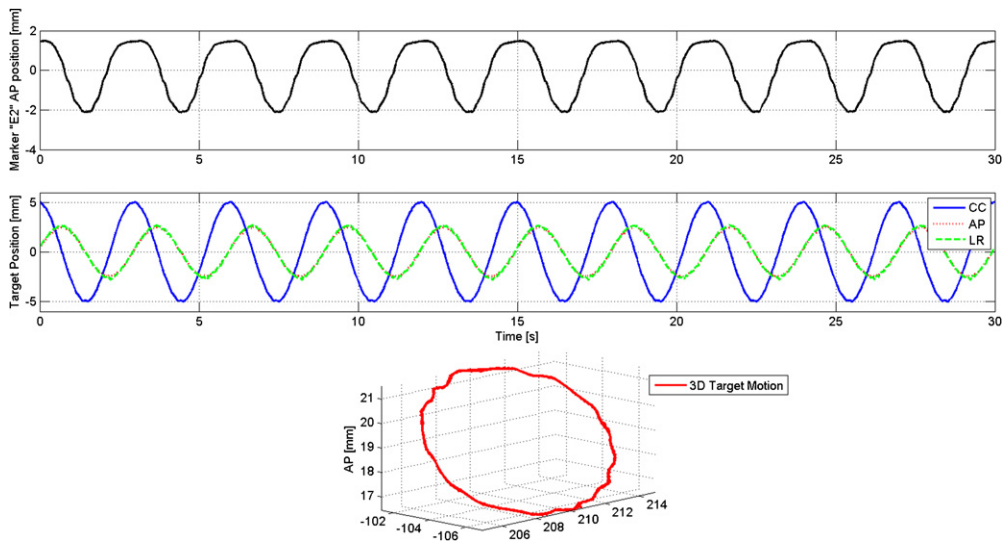
Figure 4. Upper panel: thorax marker 'E2' AP motion. Central panel: target motion along the three anatomical directions (CC, AP, LR). Lower panel: 3D target motion, hysteresis can be observed.

Table 1. Phantom's motion parameters for regular and irregular (baseline drift and phase shift) motion. Motion amplitudes are reported with respect to the anatomical directions of the artificial thorax (CC = cranio-caudal; AP = anterior–posterior; LR = lateral).

Motion type	Thorax motion period (s)	Target motion period (s)	Target motion amplitude (CC, AP, LR) (mm)	Target motion baseline drift (CC, AP, LR) (mm s^{-1})
Regular motion	3	3	10, 5, 5	0, 0, 0
Baseline drift	3	3	10, 5, 5	0.033, 0, 0
Phase shift	3	2.975	10, 5, 5	0, 0, 0

hysteretic trajectory was simulated by applying a constant 90° phase shift between the CC motion and the motion along the other two anatomical directions.

Since the external motion was generated by a cyclic expansion and contraction of the artificial thorax, its motion amplitude was not directly controllable but resulted from elastic deformations, which were imposed by means of a string pulled by a stepper motor. Figure 4 reports a 30 s trace of the internal and external motion of the phantom measured with the OTS.

As reported in table 1, baseline drift was simulated as linear motion (i.e. with constant velocity) applied to the CC motion of the target. Phase shift between the thorax and the target was achieved by reducing the target motion period by 0.025 s with respect to the thorax. As a result, a linear phase shift of 1 °s^{-1} was obtained.

For the treatment planning, a 4DCT with eight phases of the thorax was acquired in a Siemens Biograph mCT scanner. A $55 \times 55 \times 50$ mm^3 PMMA detector head was employed for the experiment to act as target and added in the 4DCT. The PMMA block contained pinpoint ionization chambers spread over a volume of $36 \times 40 \times 30$ cm^3 in four depths and five rows. Using the 4D version of the treatment planning software TRiP (Bert and Rietzel 2007, Richter *et al* 2013) a dose of 1 Gy in a volume of $35 \times 35 \times 35$ mm^3 of the PMMA block was planned.

Table 2. Experimental protocol summary. For each test, the phantom motion type and the correlation model used for target position estimation are reported.

Test name	Motion type	Correlation model
R_SSM	Regular motion	SSM
R_ANN	Regular motion	ANN
BD_SSM	Baseline drift	SSM
BD_ANN	Baseline drift	ANN
PS_SSM	Phase shift	SSM
PS_ANN	Phase shift	ANN

The planning target volume (PTV) covered the lower three rows of ionization chambers; the fourth row was on the edge of the PTV. This row was particularly sensitive to localization errors since the dose on the field edge drops in the order of 5–15% per mm distance to the edge (Kaderka *et al* 2012). The remaining four chambers were positioned outside the PTV.

Treatment planning was based on the max inhale phase of the phantom motion. This plan served for the static irradiation and was the basis for the LUTs needed for beam tracking. These LUTs contain vectors for lateral and longitudinal position compensation in every motion phase relative to the max inhale position.

Before the tracking experiments, the reference dose was measured in a dedicated test (static irradiation) where the target was stationary at the isocenter and no beam tracking was performed. As described above, this static irradiation was performed in the max inhale phase. Moreover, in order to quantify the dosimetric effects of uncompensated dose delivery, the so-called interplay pattern was acquired by irradiating the moving target (regular motion was used) without performing tumor tracking.

Different experiments were designed to assess the performance of the internal/external correlation models (table 2).

The following procedure was performed:

(1) The moving phantom was activated.
(2) Before irradiation, a training dataset was collected. It consisted in 60 samples acquired with a time period of 50 ms (20 Hz), thus a full respiratory cycle was sampled.
(3) The correlation model (SSM or ANN) was trained relying on the acquired training set.
(4) The trained model was used to estimate the target position on the transversal plane (perpendicular to the beam).

The SSM correlation model was implemented as three independent models, each one receiving as input the vertical (AP) position of one of the three considered thorax markers (E1, . . . , E3). A unique output was then obtained averaging the outputs of the three models. The fourth marker (E4) showed a reduced range of motion with respect to the other markers. As consequence, in order to further reduce complexity, the number of independent models was limited to three.

The ANN correlation model consisted in two independent ANNs: the first one estimating the target position along the CC component, the second one along the AP component. Both received the same inputs signals, as described in section 2.2.

The SSM implementation was designed to assess the technical feasibility of the integration of correlations models in the framework of real-time optical-based beam tracking, thus no specific focus was given to the simulation of clinical-like procedures. Conversely, the ANN model was implemented aiming also to simulate procedures closer to the clinical reality. As a consequence, different retraining strategies were designed for the SSM and
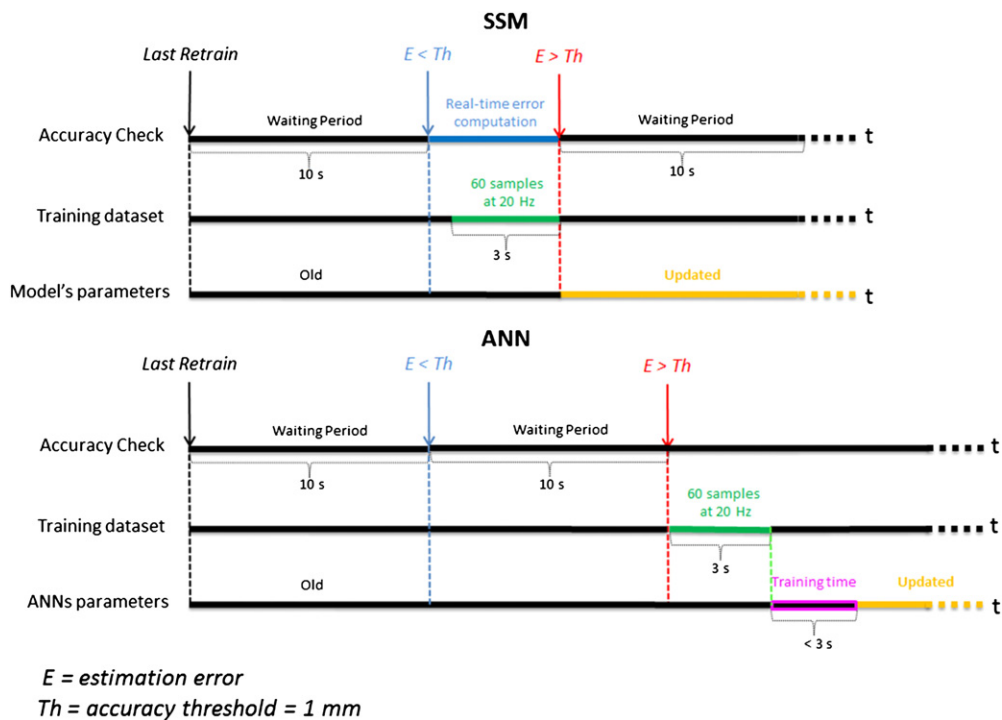
Figure 5. Schematic representation of the different retraining strategies adopted for the SSM (upper panel) and the ANN (lower panel). For each model, timings concerning the accuracy checks (first row), the acquisition of the training datasets (second row) and the update of the model's parameters (third row) are reported. The time point labeled 'E < Th' indicates that an accuracy check was performed, but the measured error (E) was below the 1 mm threshold (Th). On the contrary, the time point labeled 'E > Th' represents an accuracy check where an above-threshold error was measured, thus triggering the retraining procedure.

the ANN model, as described in figure 5: the SSM model was updated at relatively high frequency in order to quickly adapt its parameters to any variation of the internal/external correlation function, whereas the ANN model was retrained at lower frequency to test its generalization capabilities and to reduce the frequency of control acquisitions.

The longitudinal correction was always performed relying on motion phase detection and the pre-calculated LUT. The motion phase was computed in real-time by the dedicated optical motion monitoring software, which was capable to perform phase- or amplitude-based binning of the external markers AP motion (Fattori *et al* 2012). For all the presented experiments, eight phase-based motion phases were used.

(5) After the end of the irradiation, the tumor tracking software was disabled and the phantom stopped. Finally, the dose measured by the ionization chambers was analyzed.

2.4. Data analysis

2.4.1. Target tracking accuracy. For each test performed, the log file generated by the tumor tracking software was elaborated. Specifically, the estimation accuracy of the correlation model was quantified by calculating sample-by-sample the 2D estimation error, i.e. the distance (on the transversal plane) between the estimated and the actual target position, as measured in

Table 3. Geometric target tracking accuracy in the experiments involving the correlation models.

Test name	Number of retraining	Test duration (s)	2D error RMS (mm)	2D error median (mm)	2D error IQR (mm)	2D error 95th percentile (mm)
R_SSM	13	263	0.42	0.34	0.23	0.70
R_ANN	4	255	0.61	0.27	0.32	1.26
BD_SSM	21	276	0.51	0.41	0.32	0.89
BD_ANN	8	276	1.03	0.78	0.71	1.86
PS_SSM	25	277	0.62	0.49	0.40	1.10
PS_ANN	14	283	1.23	0.87	0.80	2.19

real-time by the OTS. The error distribution was characterized by calculating its median value, inter-quartile range (IQR, quantifying data variability) and 95th percentile. In order to summarize the performance of the correlation model, the RMS of the distribution was also computed.

2.4.2. Dosimetric measurements. For each ionization chamber, the absolute percentage difference between the dose measured during the tracking experiments and the reference (static) dose was computed. Then, the distribution of the dose differences on the pinpoints was characterized with median value, IQR and 95th percentile. Among the 20 installed pinpoint chambers, only the 16 included in or on the edge of the PTV were considered for this analysis.

Non-parametric statistics (Friedman test) was applied to the absolute dosimetric differences measured in the beam tracking and interplay experiments. A post-hoc multiple comparison test was then performed. For all tests, a significance level of $p = 0.01$ was considered in the analysis.

3. Results

3.1. Target tracking accuracy

Results concerning the geometric accuracy of the correlation models are reported in table 3.

It can be observed that the phase shift caused the highest geometric RMS errors for both the SSM and the ANN, being respectively 48% and 101% higher than the error measured during regular motion. For each motion pattern, the RMS error was lower for the SSM than for the ANN: 31% lower for the regular motion, 51% for the baseline drift and 50% for the phase shifts. However, the number of retraining performed was higher for the SSM than for the ANN.

As an example, figure 6 reports a 60 s trace for each motion type, where the estimated target position is computed by means of the ANN model.

The effect of the retraining procedure can be appreciated, especially in presence of irregular motion (figure 6). It can be seen that, after the retraining was performed, the baseline drift (middle panel) and the phase shift (the lower panel) were compensated, thus improving the estimation accuracy.

3.2. Dosimetric measurements

The distributions of the absolute dosimetric differences (with respect to the static irradiation) measured for each pinpoint ionization chamber are reported in figure 7.
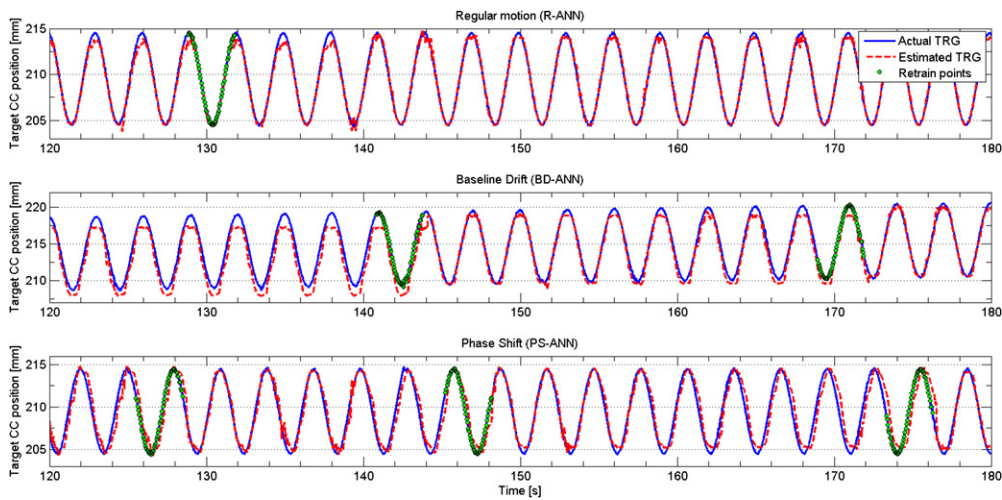
Figure 6. Actual versus ANN-estimated target position along the CC component. Green dots represent the samples included into the retraining datasets. Upper panel: regular motion. Middle panel: baseline drift. Lower panel: phase shift. The effects of the retraining procedure during irregular motion are highlighted with dotted circles.
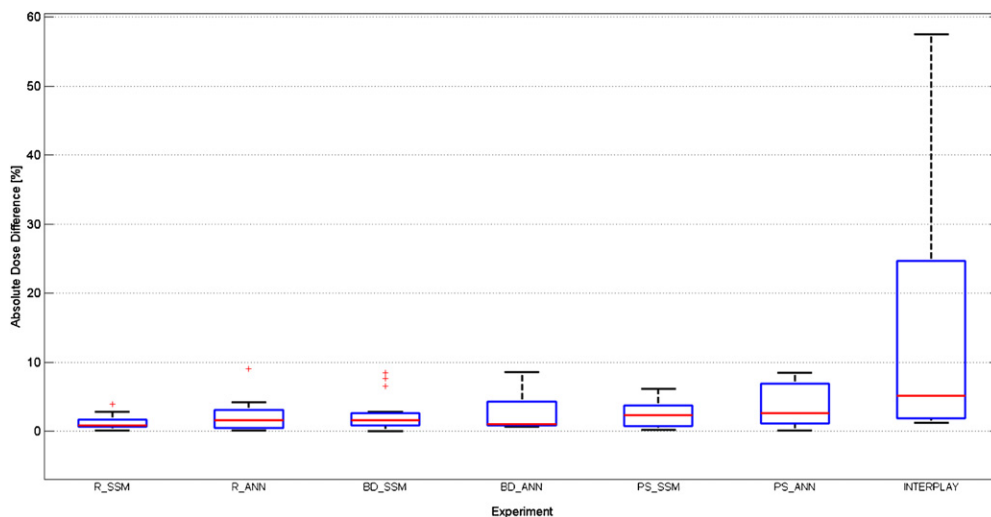
Figure 7. Boxplot of the absolute percentage dose differences measured on the 16 pinpoints in the different experiments. Red lines represent the median values, blue boxes the inter-quartile range, black dashed lines the minimum–maximum range and red crosses represent the outliers.

Considering the comparison between the two correlation models, dosimetric measurements confirmed the geometric results: for each motion patterns, the adoption of the SSM resulted in lower dosimetric differences with respect to the ANN-based model.

The uncompensated irradiation (interplay) resulted in a median dosimetric difference two times larger with respect to the worst (in terms of tracking accuracy) compensated irradiation (PS_ANN) and six times larger than the best compensated irradiation (R_SSM).

These findings were confirmed by the non-parametric statistical analysis (Friedman test), which highlighted a statically significant difference among the dose differences measured in the experiments reported in table 3 ($p < 10^{-2}$). The post-hoc multiple comparison test showed statistically significant differences between the interplay experiment and every test involving beam tracking with regular and baseline drift motion. When internal/external phase shift was simulated, statistically significant differences were not observed with respect to any other experiment.

Considering the total dose absorbed by the target, we measured a –9.5% difference (with respect to static irradiation) in the interplay experiment. When tumor tracking was applied, total dose differences were limited between –0.55% (PS_ANN) and +2.31% (BD_ANN). These results were confirmed by the measurements of mean values and standard deviations of the signed dose differences: mean values were between –0.3% and +2.3% with standard deviation increasing (min = 1.48%, max = 4.71%) as the target localization error increased.

4. Discussion

In this work, tumor tracking based on external surrogates and correlation models was experimentally tested in dedicated irradiations performed with scanned carbon ion beams at GSI.

4.1. Experimental protocol

In these experiments, the workflow required for 4D dose delivery was reproduced (Riboldi *et al* 2012): 4D treatment planning was performed on data acquired by means of 4DCT scan (Bert and Rietzel 2007), the breathing phantom was positioned on the treatment couch under optical guidance (Baroni *et al* 2007) and, finally, it was irradiated applying dynamic active scanning (Bert *et al* 2010) relying on the optical detection of the external surrogate motion coupled with the real-time estimation of the internal target position by means of correlation models (Torshabi *et al* 2010, Seregni *et al* 2012).

Two different correlation models were implemented (SSM and ANN) aiming at achieving the two main objectives of this work:

 (1) to prove the technical feasibility of the real-time integration of correlation models into a complete system for carbon ion beam tracking;
 (2) to perform tumor tracking based on correlation models simulating clinical procedures.

To meet the first requirement, the SSM was selected due to its simpler mathematical expression and its computationally faster implementation with respect to ANN. The SSM was based on a linear state-augmented correlation function, including 12 parameters. These parameters were efficiently computed by applying the least square method, whose solution required less than 10 ms on the SMART DX-100 workstation and did not interfere with time-critical processes required for real-time optical motion monitoring. However, the low complexity of the SSM implies reduced generalization capabilities. This means that, in order to achieve a given tracking accuracy, the SSM was expected to require higher retraining frequency with respect to more complex models, such as the ANN. The workflow for external surrogate-based real-time tumor tracking (figure 1, section 2.1) was successfully implemented with the SSM (section 2.3.3). Thus, the feasibility of the integration of correlation models in the framework of optically-guided ion beam tracking was experimentally proved.

However, the experimental protocol allowed the SSM to be retrained at any time if an above-threshold error was detected and at least 10 s were elapsed since the last retraining.

In addition, the retraining dataset was extracted from the signal preceding the time point when the retaining was triggered (figure 5). Both these features implicitly relied on real-time information about the internal tumor position, which, in a clinical environment, could be obtained by fluoroscopy (Shirato *et al* 1999, 2000). However, imaging dose delivered to the patients usually limits the application of this strategy in clinical practice.

In order to fix this limitation and to simulate more clinically feasible tumor tracking protocols, the ANN-based model was implemented and experimentally tested. The ANN is a more complex model with respect to the SSM, being characterized, in this specific implementation, by 36 parameters (30 neural weights plus 6 biases, as reported in figure 2). Moreover, the training procedure was performed by an iterative algorithm (BFGS quasi-Newton method) based on the computation of the first and of the approximated second derivative of the cost function. For this reason, the training of the ANN required a dedicated machine, in order not to interfere with time-critical processes running on the SMART DX-100 workstation. The increased complexity, however, can be considered as a cost to be paid for two key features of the ANN. First, the ANN does not require any *a-priori* hypothesis about the correlation function that has to be modeled. On the contrary, the SSM was based on the hypothesis that the internal/external correlation function could be efficiently modeled by a first order polynomial. Second, the ANN has potentially superior generalization capabilities with respect to the SSM. This means that, in principle, the ANN is able to achieve better accuracy than the SSM estimating data that were not part of the training dataset (Torshabi *et al* 2010). Both these features can be regarded as useful advantages in presence of inter- and intra-patient variability of the internal/external correlation function, as observed in real patients (Seppenwoolde *et al* 2002, Nishioka *et al* 2008, Redmond *et al* 2009).

In the presented experimental activity, the ANN model was tested with a protocol designed to be closer to the clinical practice with respect to SSM. For this purpose, control imaging was simulated with a period of 10 s, allowing the retraining procedure to be triggered only in correspondence to these control acquisitions. Moreover, when retraining was triggered, a new dataset was acquired in the 3 s following (and not preceding) the control acquisition. These strategies did not require real-time information on the internal target position, thus they simulated a tumor tracking procedure based on real-time external motion monitoring coupled with low frequency internal motion monitoring, as implemented, for instance, in the Synchrony system (Kilby *et al* 2010).

4.2. Geometric and dosimetric results

Considering the geometric tracking accuracy of the SSM and of the ANN model (table 2), the SSM achieved better accuracy than the ANN in every performed test. This result can be regarded as a direct consequence of the different experimental protocols that were discussed above: more information about the internal target position was provided to the SSM with respect to the ANN, as confirmed by the number of retraining performed by the two models. In addition, the total time required to update the model was considerably shorter in the SSM than in the ANN, resulting in a faster adaptation of its parameters.

For instance, in the baseline drift experiments, the ANN performed about the 38% of the retraining performed by the SSM: this resulted in a 49% higher RMS. The retraining procedure proved to be effective in improving the estimation accuracy, as reported in figure 8 for the ANN model.

The estimation error in the CC component (where the linear baseline drift was applied) increased almost linearly until a retraining procedure was triggered. This behavior, which was also present during phase shift, can be justified observing that the breathing phantom generated
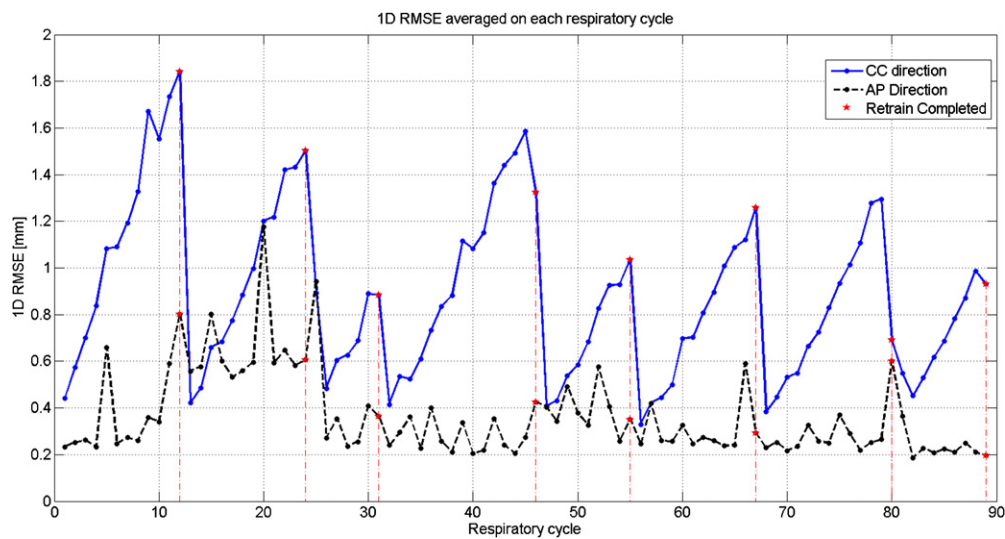
Figure 8. RMS error (averaged on each respiratory cycle) of the estimation performed by the ANN model in presence of baseline drift (BD_ANN). Red dots highlight the respiratory cycles where a retraining procedure was completed.

an irregular internal motion coupled with an always stationary external motion, resulting in a non-stationary correlation function, whose parameters were continuously changing as a function of time. Due to the slow variation of the internal motion (table 1), the correlation function could be considered stationary for short time periods, thus the models were able to learn the function described into training datasets collected over a period of 3 s. After the training procedure, the estimation error increased because the models were driven by the stationary external motion. When the retraining was triggered, the new correlation function was learned and the estimation accuracy was improved.

The reported tracking accuracy can be considered comparable with results presented by Hoogman *et al* (2009). In this work, the authors evaluated the accuracy of the Synchrony Respiratory Tracking System on a database of 44 lung cancer patients: results showed that the standard deviation of signed correlation error was up to 1.9 mm in CC and LR directions and up to 2.5 mm in AP direction.

The dosimetric differences measured when tumor tracking was performed (except for the tests PS_SSM and PS_ANN) showed a statistically significant difference with respect to the measurements concerning the interplay experiment. This result proved that tumor tracking based on correlation models can be regarded as an effective motion mitigation strategy for highly conformal treatments, such as those delivered by means of carbon ion beams. The experiments PS_SSM and PS_ANN resulted in dose differences that were not statistically significant from those measured in the interplay experiment. Phase shift, indeed, was the motion irregularity that caused the largest geometric and dosimetric errors.

Previous works (Bert *et al* 2010) quantified the dosimetric accuracy of the GSI ion beam tracking system, by measuring the dose absorbed by an array of 24 ionization chambers. A dose difference of $0.3 \pm 1.5\%$ (relative to static irradiation) was measured during a 1D sinusoidal motion directly measured with a laser position sensor. The larger dose differences reported in the present study can be justified both by the more complex target motion that was simulated and by the introduction of correlation models replacing direct target localization.

Lastly, the results of this work provided an insight on the expected dosimetric errors in correspondence to both limited and large target localization errors, the first being measured in the experiments with correlation models, the second in the interplay pattern.

4.3. Technical features of the correlation models

Different input configurations were used in the two proposed models: in the SSM, three independent models were implemented, each one driven by a single external marker. This configuration, which reproduces the current state of art of clinical tracking systems (Kilby *et al* 2010), was chosen due to the reduced complexity and increased redundancy. In the ANN model, signals from different markers were merged into a single input vector, thus exploiting enhanced information in the estimation of the internal/external correlation function.

The ANNs architecture and training strategies were designed to avoid overfitting, i.e. the loss of generalization capability coupled with high sensitivity to noise. For this purpose, a relatively simple architecture was selected and specific training termination criteria were defined (section 2.2). The main disadvantage associated with this design was the sporadic occurrence of limited ($\leqslant$1 mm) estimation errors even during regular motion, as it can be observed in figure 6 (upper panel). However, this could be considered a feasible price to be paid for preserving generalization capabilities.

Both models proved to be able to cope with hysteretic target trajectory. As known, hysteresis is a typical characteristic of the internal respiratory motion (Seppenwoolde *et al* 2002, Pepin *et al* 2010) and causes an ambiguity in the internal/external correlation function. To solve this ambiguity, either two different correlation models are used (an expiratory model and an inspiratory model) (Kilby *et al* 2010, Ernst *et al* 2011) or temporal dependency is explicitly or implicitly included into a single correlation model. We chose to follow this second approach since it allowed a reduction of the overall training time and did not introduce a dependency of the estimation accuracy on the respiratory phase detection. Temporal dependency was implemented into the SSM by performing state augmentation (Ruan *et al* 2008), thus relating the current target position with both the current and the delayed external surrogate. In the ANN model, temporal dependency was implicitly included by providing, as additional input, the external surrogate velocity, which carried information about the first order dynamic of the system.

4.4. Limitations and future developments

The main technical limitation of the proposed tumor tracking technique consisted in the requirement of a training dataset containing 60 samples acquired at 20 Hz. This relatively high sample frequency was required in order to confine the acquired points in a reduced time interval (3 s). In order to overcome this limitation, it is possible either to reduce the number of training samples or to spread them over longer time periods.

In this work, the correlation models were used to estimate the internal target position only for transversal (CC and AP) motion compensation. However, it can be observed form figure 4 that the internal target showed exactly the same motion amplitude and period in the AP and LR components (Steidl *et al* 2012). As a consequence, the estimation of LR motion form the external surrogate is not expected to introduce any further challenge with respect to the estimation of the AP motion. Longitudinal tracking was achieved relying on real-time motion phase detection and a pre-computed LUT containing the required range shift for each motion phase (Saito *et al* 2009). The LUT was obtained as part of the 4D treatment planning (Bert and Rietzel 2007). The presented dosimetric differences were due

to the combined effects of both transversal and longitudinal inaccuracies in dose deposition. Although these two different contributions cannot be separated, a dedicated experiment proved that the main source of dosimetric errors was the transversal target localization error. Two different irradiations were performed without using correlation models (direct tracking): in the first irradiation, longitudinal compensations were applied with the range shifter and a signed dose difference (with respect to the static irradiation) of $-0.61 \pm 2.18\%$ (mean $\pm$ SD) was measured. In the second irradiation, longitudinal compensations were not applied, resulting in a $-0.31 \pm 2.70\%$ dose difference. Only a slight increase of the standard deviation was measured in the second test. As a consequence, the dosimetric errors reported in this work can be attributed largely to transversal target localization errors.

5. Conclusions

In this work we presented an experimental activity concerning the application of tumor tracking based on internal/external correlation models in ion beam therapy. The proposed study extends previous works concerning internal/external correlation models (Seregni *et al* 2012) and optical motion monitoring (Fattori *et al* 2012) since the integration of these strategies is presented and validated with specific experiments.

The experimental activity reproduced the entire workflow required for 4D dose delivery in particle therapy and showed that technical feasibility of the integration of correlation models into the framework of optically-guided beam tracking. Results showed that the application of these motion compensation strategies allowed to significantly decrease the dosimetric errors measured in an uncompensated dose delivery. The presence of phase shifts between the internal and the external motion caused the highest dosimetric errors. Therefore, future work should be focused on the development of strategies dedicated to the detection and the compensation of this respiratory motion feature.

Acknowledgments

References

Baroni G, Ferrigno G and Pedotti A 1998 Implementation and application of real-time motion analysis based on passive markers *Med. Biol. Eng. Comput.* **36** 693–703
Baroni G, Riboldi M, Spadea M F, Tagaste B, Garibalidi C, Orecchia R and Pedotti A 2007 Integration of enhanced optical tracking techniques and imaging in IGRT *J. Radiat. Res.* **48** A61–74
Bert C and Durante M 2011 Motion in radiotherapy: particle therapy *Phys. Med. Biol.* **56** R113–44
Bert C, Gemmel A, Saito N, Chaudhri N, Schardt D, Durante M, Kraft G and Rietzel E 2010 Dosimetric precision of an ion beam tracking system *Radiat.Oncol.* **5** 61
Bert C and Rietzel E 2007 4D treatment planning for scanned ion beams *Radiat. Oncol.* **2** 24
Durante M and Loeffler J S 2010 Charged particles in radiation oncology *Nature Rev. Clin. Oncol.* **7** 37–43
Ernst F, Bruder R, Schlaefer A and Schweikard A 2011 Correlation between external and internal respiratory motion: a validation study *Int. J. Comput. Assist. Radiol. Surg* **7** 483–92
Fattori G *et al* 2012 Integration of optical tracking for organ motion compensation in scanned ion-beam therapy *Med. Phys.* **39** 3889

Ford E, Mageras G, Yorke E, Rosenzweig K E, Wagman R and Ling C C 2002 Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging *Int. J. Radiat. Oncol. Biol. Phys.* **52** 522–31

Grözinger S O, Bert C, Haberer T and Kraft Rietzel E 2008 Motion compensation with a scanned ion beam: a technical feasibility study *Radiat. Oncol.* **3** 34

Haberer T, Becher W, Schardt D and Kraft G 1993 Magnetic scanning system for heavy ion therapy *Nucl. Instrum. Methods* A **330** 296–305

Hoogeman M J, Prevost B, Nuyttens J, Poll J, Levendag P and Heumen B 2009 Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files *Int. J. Radiat. Oncol. Biol. Phys.* **74** 297–303

Kaderka R, Schardt D, Durante M, Berger T, Ramm U, Licher J and La Tessa C 2012 Out-of-field dose measurements in a water phantom using different radiotherapy modalities *Phys. Med. Biol.* **57** 5059–74

Karger C P, Jaekel O and Hartmann G H 1999 A system for three-dimensional dosimetric verification of treatment plans in intensity-modulated radiotherapy with heavy ions *Med. Phys.* **26** 2125–32

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

Keall P J, Mageras G S, Balter J M, Emery R S, Forster K M and Jiang S B 2006 The management of respiratory motion in radiation oncology: report of AAPM Task Group 76 *Med. Phys.* **33** 3874–900

Kilby W, Dooley J R, Kuduvalli G, Sayeh S and Maurer C R Jr 2010 The cyberknife robotic radiosurgery system in 2010 *Technol. Cancer Res. Treat.* **9** 433–52 PMID: 20815415

Kraft G 2000 Tumor therapy with heavy charged particles *Prog. Part. Nucl. Phys.* **45** S473–544

Langen K M and Jones D T L 2001 Organ motion and its management *Int. J. Radiat. Oncol. Biol. Phys.* **50** 265–78

Langner U W and Keall P J 2009 Accuracy in the localization of thoracic and abdominal tumors using respiratory displacement, velocity, and phase *Med. Phys.* **36** 386–93

Mori S, Chen G T and Endo M 2007 Effects of intrafractional motion on water equivalent path length in respiratory gated heavy charged particle beam radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **69** 308–17

Mori S, Wolfgang J, Lu H M, Schneider R, Choi N C and Chen G T Y 2008 Quantitative assessment of range fluctuations in charged particle lung irradiation *Int. J. Radiat.Oncol.* **70** 253–61

Murphy M J and Pokhrel D 2009 Optimization of an adaptive neural network to predict breathing *Med. Phys.* **36** 40–47

Nishioka S *et al* 2008 Exhale fluctuation in respiratory-gated radiotherapy of the lung: a pitfall of respiratory gating shown in a synchronized internal/external marker recording study *Radiother. Oncol.* **86** 69–76

Pepin E W, Wu H, Sandison G A, Langer M and Shirato H 2010 Site-specific volumetric analysis of lung tumour motion *Phys. Med. Biol.* **55** 3325–37

Redmond K J, Song D Y, Fox J L, Zhou J, Rosenzweig C N and Ford E 2009 Respiratory motion changes of lung tumors over the course of radiation therapy based on respiration-correlated four-dimensional computed tomography scans *Int. J. Radiat. Oncol. Biol. Phys.* **75** 1605–12

Riboldi M, Baroni G, Orecchia R and Pedotti A 2004 Enhanced surface registration techniques for patient positioning control in breast cancer radiotherapy *Technol. Cancer Res. Treat.* **3** 51–58 PMID: 14750893

Riboldi M, Orecchia R and Baroni G 2012 Real-time tumour tracking in particle therapy: Technological developments and future perspectives *Lancet Oncol.* **13** e383–91

Richter D, Schwarzkopf A, Trautmann J, Krämer M, Durante M, Jäkel O and Bert C 2013 Upgrade and benchmarking of a 4D treatment planning system for scanned ion beam therapy *Med. Phys.* **40** 051722

Ruan D, Fessler J A, Balter J M, Berbeco R I, Nishioka S and Shirato H 2008 Inference of hysteretic respiratory tumor motion from external surrogates: a state augmentation approach *Phys. Med. Biol.* **53** 2923–36

Saito N, Bert C, Chaudhri N, Gemmel A, Schardt D, Durante M and Rietzel E 2009 Speed and accuracy of a beam tracking system for treatment of moving targets with scanned ion beams *Phys. Med. Biol.* **54** 4849–62

Schaerer J, Fassi A, Riboldi M, Cerveri P, Baroni G and Sarrut D 2012 Multi-dimensional respiratory motion tracking from markerless optical surface imaging based on deformable mesh registration *Phys. Med Biol.* **57** 357–73

Schweikard A, Glosser G, Bodduluri M, Murphy M J and Adler J R 2000 Robotic motion compensation for respiratory movement during radiosurgery *Comput. Aided Surg.* **5** 263–77

Seppenwoolde Y, Shirato H, Kitamura K, Shimizu S, Van Herk M, Lebesque J V and Miyasaka K 2002 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** 822–34

Seregni M, Cerveri P, Riboldi M, Pella A and Baroni G 2012 Robustness of external/internal correlation models for real-time tumor tracking to breathing motion variations *Phys. Med. Biol.* **57** 7053–74

Sharp G C, Jiang S B, Shimizu S and Shirato H 2004 Prediction of respiratory tumour motion for real-time image-guided radiotherapy *Phys. Med. Biol.* **49** 425–40

Shimizu S *et al* 2001 Detection of lung tumor movement in real-time tumor-tracking radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **51** 304–10 PMID: 11567803

Shirato H, Shimizu S, Shimizu T, Nishioka T and Miyasaka K 1999 Realtime tumor-tracking radiotherapy *Lancet* **353** 1331

Shirato H *et al* 2000 Physical aspects of a real-time tumor tracking system for gated radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **48** 1187–95

Shirato H *et al* 2004 Feasibility of synchronization of real-time tumor-tracking radiotherapy and intensity-modulated radiotherapy from viewpoint of excessive dose from fluoroscopy *Int. J. Radiat. Oncol.* **60** 335–41

Steidl P, Richter D, Schuy C, Schubert E, Haberer T, Durante M and Bert C 2012 A breathing thorax phantom with independently programmable 6D tumour motion for dosimetric measurements in radiation therapy *Phys. Med. Biol.* **57** 2235–50

Torshabi A E, Pella A, Riboldi M and Baroni G 2010 Targeting accuracy in real-time tumor tracking via external surrogates: a comparative study *Technol. Cancer Res. Treat.* **9** 551–62 PMID: 21070077