Real-time motion-adaptive-optimization (MAO) in TomoTherapy

Article *in* Physics in Medicine and Biology · July 2009

DOI: 10.1088/0031-9155/54/14/003 · Source: PubMed

7 authors, including:

Weiguo Lu
148 PUBLICATIONS **2,372** CITATIONS
SEE PROFILE

Mingli Chen
University of Texas Southwestern Medical Center
86 PUBLICATIONS **937** CITATIONS
SEE PROFILE

Kenneth Ruchala
Sun Nuclear Corporation
149 PUBLICATIONS **3,090** CITATIONS
SEE PROFILE

Quan Chen
University of Kentucky
173 PUBLICATIONS **1,892** CITATIONS
SEE PROFILE

IOP PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **54** (2009) 4373–4398 doi:10.1088/0031-9155/54/14/003

Real-time motion-adaptive-optimization (MAO) in TomoTherapy

Weiguo Lu[1,4], Mingli Chen[1], Kenneth J Ruchala[1], Quan Chen[1], Katja M Langen[2], Patrick A Kupelian[2] and Gustavo H Olivera[1,3]

[1] TomoTherapy Inc., 1240 Deming Way, Madison, WI, USA
[2] MD Anderson Cancer Center-Orlando, Orlando, FL, USA
[3] University of Wisconsin-Madison, 1300 University Avenue, Madison, WI, USA

E-mail: wlu@tomotherapy.com

Received 3 November 2008, in final form 7 May 2009
Published 23 June 2009
Online at stacks.iop.org/PMB/54/4373

Abstract

IMRT delivery follows a planned leaf sequence, which is optimized before treatment delivery. However, it is hard to model real-time variations, such as respiration, in the planning procedure. In this paper, we propose a negative feedback system of IMRT delivery that incorporates real-time optimization to account for intra-fraction motion. Specifically, we developed a feasible workflow of real-time motion-adaptive-optimization (MAO) for TomoTherapy delivery. TomoTherapy delivery is characterized by thousands of projections with a fast projection rate and ultra-fast binary leaf motion. The technique of MAO-guided delivery calculates (i) the motion-encoded dose that has been delivered up to any given projection during the delivery and (ii) the future dose that will be delivered based on the estimated motion probability and future fluence map. These two pieces of information are then used to optimize the leaf open time of the upcoming projection right before its delivery. It consists of several real-time procedures, including 'motion detection and prediction', 'delivered dose accumulation', 'future dose estimation' and 'projection optimization'. Real-time MAO requires that all procedures are executed in time less than the duration of a projection. We implemented and tested this technique using a TomoTherapy® research system. The MAO calculation took about 100 ms per projection. We calculated and compared MAO-guided delivery with two other types of delivery, motion-without-compensation delivery (MD) and static delivery (SD), using simulated 1D cases, real TomoTherapy plans and the motion traces from clinical lung and prostate patients. The results showed that the proposed technique effectively compensated for motion errors of all test cases. Dose distributions and DVHs of MAO-guided delivery approached those of SD, for regular and irregular respiration with a peak-to-peak amplitude of 3 cm, and for medium and large

[4] Author to whom any correspondence should be addressed.

0031-9155/09/144373+26$30.00 Printed in the UK

prostate motions. The results conceptually proved that the proposed method is applicable for real-time motion compensation in TomoTherapy delivery. Extension of the method to real-time adaptive radiation therapy (ART) that compensates for all kinds of delivery errors was proposed. Further validation and clinical implementation is underway.

(Some figures in this article are in colour only in the electronic version)

Nomenclature

$\boldsymbol{B}_{i,j}$	planned beamlet dose distribution of projection i and leaf j with unit intensity
$\boldsymbol{w}$	sinogram of leaf fluence (leaf open time) indexed by the projection number i and leaf number j, $\boldsymbol{w} = \{w_{i,j}\}$
$\hat{\boldsymbol{w}}$	planned sinogram obtained through optimization in the planning procedure and emphasized by hat
$\tilde{\boldsymbol{w}}$	delivery sinogram adjusted from the planned sinogram according to tumor motion and emphasized by tilde
δ	couch movement per gantry rotation
P	number of projections per gantry rotation
$\boldsymbol{u} = (u_x, u_y, u_z)$	tumor motion (displacement); the upper index in $\boldsymbol{u}^i$ indicates the motion during the ith projection. In this paper, the word 'motion' is used as a synonym of 'displacement' with units in 'mm'

For the readers' convenience, some abbreviations used in this paper are listed below.

MAO	motion-adaptive optimization
MAD	motion-adaptive delivery, the technique presented in Lu (2008a, 2008b)
MD	planned delivery for a moving target (i.e. regular delivery without motion compensation)
SD	planned delivery for a static target (i.e. the scenario assumed during treatment planning)

1. Introduction

State-of-the-art IMRT delivery follows a planned leaf sequence. The leaf sequence is optimized in treatment planning, which assumes a certain treatment configuration, including patient setup, anatomy and the physiological state. Though there is much data on the mobility of tumors during radiotherapy treatment (Webb 2006a, Langen and Jones 2001), it is difficult to model real-time treatment configurations, e.g. patient respiration, accurately during the planning procedure. Deviation of delivery conditions from planning conditions results in sub-optimal dose distributions. In particular, IMRT delivery, which is dynamic in nature, to moving tumors could have the consequence of dosimetric deviations between the planed and delivered dose distributions. Accurate modeling or even long term prediction of tumor motion is still an infeasible task due to the complexity of intra-fraction motion, which is not only very patient specific but also has random components. However, real-time motion compensation is promising because it does not rely much on the *a priori* knowledge or assumption about intra-fraction motion. The reported real-time motion compensation methods are mainly tracking

based, and in effect, they try to put the same intensity of radiation beam on the same position in the tumor reference frame at the same time as what was planned. They are implemented through linac tracking (Nuyttens *et al* 2006, Murphy 2004), MLC tracking (Keall *et al* 2001, 2006, Neicu *et al* 2003) or couch tracking (D'Souza *et al* 2005), and can be characterized as hardware solutions. Papiez *et al* (1994, 1999, 2002, 2005, 2007), Papiez and Rangaraj (2005), Papiez and Timmerman (2008), Papiez and Langer (2006), Papiez (2004), McMahon *et al* (2007a, 2007b), Webb and Binnie (2006) and Webb (2006b) incorporated tumor motion into the dynamic MLC leaf velocity optimization. These methods are regarded as software approaches. Tracking-based methods intend to fully compensate for motion effect instantly once motion is detected. Such schemes are considered open loop tracking methods, because they do not explicitly take into account the compensation errors from hardware limitations and/or prediction. The open-loop tracking methods put great demands on hardware, such as the velocity and position accuracy of MLC, linac or the couch, etc, as well as on the accuracy of motion prediction.

In our previous work (Lu 2008a, 2008b), we presented a real-time motion-adaptive delivery (MAD) technique to compensate for relatively regular respiratory motion in Topo/TomoTherapy^{SM} delivery. MAD is a software approach that requires no hardware changes of TomoTherapy® machines (Mackie *et al* 1999). It uses out-of-order execution of the planned sinogram to compensate for the intra-fraction motion. In effect, MAD puts the *same intensity* of radiation beam on the *same position* in the tumor reference frame, similar to tracking-based methods, but *not at the same time* as planned. However, the MAD technique is still an open-loop approach. It does not account for previously uncompensated errors. One basic assumption of the MAD technique is that each tumor voxel has approximately equal chance to be irradiated during delivery as planned regardless how the tumor moves. This assumption is valid for relatively regular respiration, but may not be for highly irregular respiration or other kinds of motion.

We realized that none of the compensation is perfect and there are remaining errors. Such errors, if systematic, will accumulate. We need a negative feedback system to correct for cumulative errors. In this paper, we propose a close-loop negative feedback system for IMRT delivery. The proposed strategy, which termed as 'real-time motion-adaptive-optimization (MAO)-guided radiotherapy delivery', incorporates real-time optimization into the radiation delivery procedure and originally aims at motion compensation in TomoTherapy delivery. However, we believe, it will open up a new paradigm in IMRT optimization and delivery.

2. Workflow

Figure 1 illustrates the workflow for real-time MAO-guided TomoTherapy delivery. In this workflow, we use a regular TomoTherapy treatment planning system (TPS) to optimize a planned sinogram, but no motion margin is imposed in the planning procedure. Our approach does not alter the couch or gantry speed, or jaw positions. That is, during radiation delivery, the couch and gantry move in a constant speed, and jaws are in fixed positions, just as planned. The tumor position is real time detected/updated via some surrogates (Shirato *et al* 2000a, Keall *et al* 2004, Zhang *et al* 2003, Balter *et al* 1995, Willoughby *et al* 2006, Kupelian *et al* 2007, Shchory *et al* 2008) or directly via treatment beams (Lu *et al* 2006c). Our proposed motion management is independent of motion detection methods. A 4D dose calculator is used to accumulate the delivered dose up to the last delivered projection in real time. The dose to be delivered in future projections (beyond the upcoming projection) is estimated. Based on the predicted tumor position for the upcoming projection and accumulated delivered dose, the leaf open time for the upcoming projection is optimized in real time to account for both

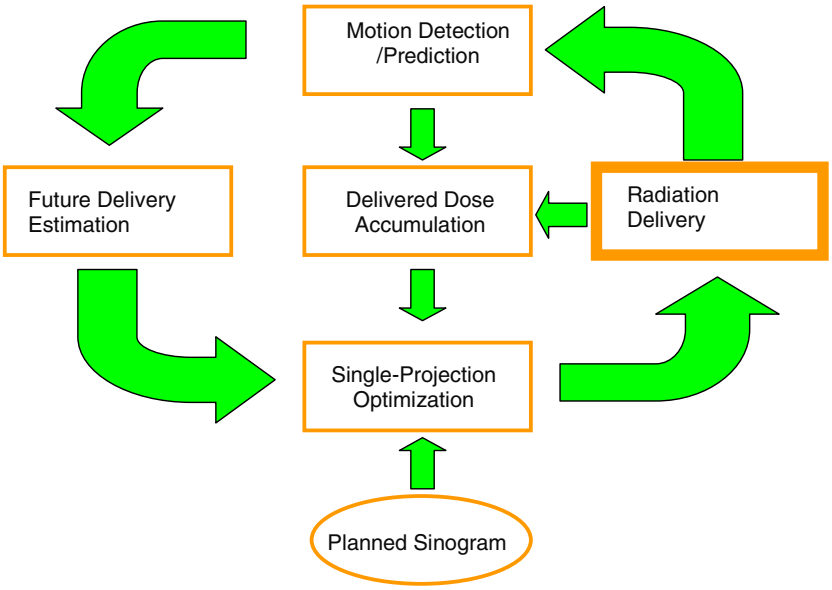

Figure 1. Workflow for MAO-guided delivery in TomoTherapy. Treatment planning uses the regular TomoTherapy planning system except that no motion margin is needed. The planned sinogram, accumulation of delivered dose, estimation of future dose and the prediction of the tumor position are fed into the optimization process to output the leaf open time for the upcoming projection. This is considered a closed loop, because the output of the optimizer is sent to the machine as the instruction to deliver dose, and the delivered dose is real time accumulated by a 4D dose calculator and sent back (as input) to the optimizer.

accumulated delivery errors and future dose estimation. The optimized leaf open time is used to control the bMLC that modulates the radiation delivery for the upcoming projection.

This workflow models the radiation delivery procedures as a negative feedback system. To implement this workflow in real time, the whole procedure as illustrated in figure 1, except the offline planning procedure, must be executed in time less than a projection. The projection time can be regarded as the temporal resolution of the real-time MAO technique, because we only perform optimization once per projection. The minimum projection time for the TomoTherapy system is about 200 ms, which corresponds to 51 projections per 10 s gantry rotation.

3. Theory

3.1. Plan optimization

IMRT plan optimization can generally be formulated as a constrained nonlinear optimization problem

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \mathbf{F}(\boldsymbol{B}\boldsymbol{w}) \tag{1}$$

subject to the constraints

$$\boldsymbol{B}\boldsymbol{w} \in \mathcal{D} \qquad \text{and} \qquad \boldsymbol{w} \in \mathcal{W}, \tag{2}$$

where $\mathbf{F}$ is the planning objective, $\mathcal{D}$ is the space of permissible dose distributions, which satisfy, for example, the minimal dose, maximal dose and DVH requirements, and $\mathcal{W}$ is

the space of feasible fluence maps that are deliverable by hardware of the therapy machine. Typically, the objective function $\mathbf{F}$ can be expressed as a weighted sum of multiple objective functions

$$\mathbf{F} = \sum_s a_s \mathbf{F}_s. \tag{3}$$

The product $\boldsymbol{Bw}$, indicating the summation of $\boldsymbol{B}_{i,j} w_{i,j}$ for all pairs (i, j) of the projection number i and leaf number j, is the overall 3D dose distribution as a result of delivering $\boldsymbol{w}$. The fluence map $\boldsymbol{w}$ can be directly or indirectly converted into delivery instructions for the machine. An example of direct conversion of the fluence map $\boldsymbol{w}$ into delivery instructions is the TomoTherapy system, where the fluence map describes the leaf open time for every projection. Indirect conversions, like step-and-shoot and dynamic MLC-based delivery, convert fluence maps into tens to hundreds of segments or several leaf sequences. IMRT optimization is usually a time-consuming process that takes minutes to hours even with state-of-the-art computers. The following are the contributing factors that make IMRT optimization a lengthy process:

(1) Different components of the objective function as in equation (3) interact, and better tradeoffs can be achieved by users adjusting the weights a_s. Therefore, it may involve significant user manipulations to drive the optimization toward clinical objectives.
(2) It is categorized as a very large scale nonlinear optimization problem. The problem size is reflected in the size of the set of beamlets $\{\boldsymbol{B}_{i,j}\}$ and in the number of voxels. The number of unknowns, i.e. the number of beamlets (= number of projections $\times$ number of leaves), can be as large as thousands for conventional IMRT and tens of thousands or even hundreds of thousands for TomoTherapy optimization. Each distribution $\boldsymbol{B}_{i,j}$ consists of dose values for the voxels in a 3D volume, and the number of voxels is typically on the order of millions.
(3) The set of beamlet dose distributions $\{\boldsymbol{B}_{i,j}\}$ is quite dense, not sparse. Theoretically, each beamlet has contributions to all voxels in the 3D volume because of scatter. Even if we only consider the primary radiation, each voxel is directly irradiated by $51/0.3 = 170$ beamlets, given that there are 51 projections per rotation and a typical pitch of 0.3 is used in the case of TomoTherapy. That is, the optimizer has to find out tradeoffs among those beamlets to minimize the objective function, and it usually requires hundreds of iterations for results to converge.

To be focused, in the following discussion, we only study the TomoTherapy delivery mode. TomoTherapy delivery is projection based. That is, the fluence map $\boldsymbol{w}$ is organized into a sinogram of beamlet intensities

$$\boldsymbol{w} = \{w_{i,j}\}, \tag{4}$$

where i is the projection index and j is the leaf index. Note that the projections, used in their wider sense, are accrued by rotations, and the total number of projections is the number of projections per rotation multiplied by the number of rotations. Typically, there are hundreds to thousands of projections in each sinogram.

3.2. Motion-encoded beamlets

A set of beamlet dose distributions $\{\boldsymbol{B}_{i,j}\}$ associated with a static patient is off-line pre-calculated before plan optimization. Same as for plan optimization, the beamlet set $\{\boldsymbol{B}_{i,j}\}$ is essential in accumulating dose and updating the fluence map for real-time optimization. But, when the tumor moves during delivery, the radiation source may no longer be at the planned positions in the tumor reference frame. Therefore, the original beamlet set $\{\boldsymbol{B}_{i,j}\}$ is no longer valid for describing the patient dose distribution and revisions of the beamlet set are necessary.

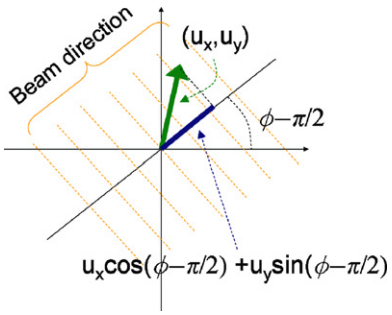

Figure 2. Illustration of projection of motion $\boldsymbol{u}$ along the leaf direction $(\cos(\phi - \pi/2), \sin(\phi - \pi/2), 0)$, where ϕ is the angle of source direction.

We call such revised beamlets 'motion-encoded beamlets' (Lu *et al* 2006a). Re-calculating beamlets from scratch in real time is infeasible. In the following, we describe approximations of motion-encoded beamlets that are modified from the original beamlets in real time.

Suppose, at the time of delivering the ith projection, tumor motion is given by $\boldsymbol{u} = (u_x, u_y, u_z)$. The modification of beamlets can be decomposed into the longitudinal and transversal directions. Let us first consider the longitudinal direction. For convenience, the source position is referenced relative to the tumor position unless otherwise stated. At the ith projection, the planned source position is $(i/P) \cdot \delta$, where δ is couch movement per gantry rotation. With motion $\boldsymbol{u} = (u_x, u_y, u_z)$, the source position changes to $(i/P) \cdot \delta - u_z$. The beamlet of this new source position can be approximated by a linear interpolation of the original beamlets of two nearest projections at the same gantry angle

$$\boldsymbol{B}'_{i,\cdot}(\boldsymbol{x}) = (1-\alpha)\boldsymbol{B}_{i-mP,\cdot}(\boldsymbol{x}_1) + \alpha \boldsymbol{B}_{i-(m+1)P,\cdot}(\boldsymbol{x}_2) \tag{5}$$

for every point $\boldsymbol{x} = (x, y, z)$ in the space, where $u_z/\delta = m + \alpha$ for some integer m and $0 \leqslant \alpha < 1$, and $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are defined as

$$\boldsymbol{x}_1 = (x, y, z + \alpha\delta) \qquad \text{and} \qquad \boldsymbol{x}_2 = (x, y, z - (1-\alpha)\delta). \tag{6}$$

Similar to longitudinal modification, transversal modification that accounts for transversal motion can also be obtained via linear interpolating the original beamlets of two nearest leaves

$$\tilde{\boldsymbol{B}}_{i,j}(\boldsymbol{x}) = (1-\beta)\boldsymbol{B}'_{i,j-l}(\boldsymbol{x}_3) + \beta \boldsymbol{B}'_{i,j-l-1}(\boldsymbol{x}_4), \tag{7}$$

where $\boldsymbol{u}_{\phi-\pi/2}/a = l + \beta$ for some integer l and $0 \leqslant \beta < 1$, a is the leaf width, ϕ is the projection angle, $\phi = (i/P)2\pi$, and $\boldsymbol{u}_{\phi-\pi/2}$ is the projection of motion $\boldsymbol{u}$ along the leaf direction $(\cos(\phi - \pi/2), \sin(\phi - \pi/2), 0)$ as illustrated in figure 2. Here, the coordinate shifts $\boldsymbol{x}_3$ and $\boldsymbol{x}_4$ are defined as

$$\begin{aligned} \boldsymbol{x}_3 &= (x, y, z) + \beta a \cdot (\cos(\phi - \pi/2),\ \sin(\phi - \pi/2), 0) \\ \boldsymbol{x}_4 &= (x, y, z) - (1-\beta) a \cdot (\cos(\phi - \pi/2),\ \sin(\phi - \pi/2), 0). \end{aligned} \tag{8}$$

Armed with equations (5) and (7) and assuming the leaf fluence for the ith projection is $\{w_{i,j}\}$, the motion-encoded dose $\boldsymbol{d}_i$ for the ith projection should be

$$\boldsymbol{d}_i = \sum_j \tilde{\boldsymbol{B}}_{i,j} w_{i,j}. \tag{9}$$

3.3. Real-time MAD

We reported the real-time MAD technique for TopoTherapy/TomoTherapy in previous papers (Lu 2008a, 2008b). The MAD technique provides an initial guess of fluence and a reasonable model for future dose estimation in our MAO strategy, which will be discussed in later sections. We briefly summarize the MAD technique here.

At any projection, if the tumor motion is $\boldsymbol{u} = (u_x, u_y, u_z)$, which is equivalent to the source moving by $-\boldsymbol{u}$ relative to the tumor from the planned trajectory, then we would use an interpolated leaf open time so that the dose of that projection is closest to what was intended at that position. Let $\hat{\boldsymbol{w}} = \{\hat{w}_{i,j}\}$ denote the planned sinogram where i is the projection index and j is the leaf index. When referenced to the couch, the source position parameterized by the projection index i traces out a helix on a cylindrical surface. If the source is shifted by $-u_z$ in the longitudinal direction for projection i, then the new leaf fluence $\tilde{\boldsymbol{w}} = \{\tilde{w}_{i,j}\}$ can be calculated by linear interpolating the planned sinogram $\hat{\boldsymbol{w}}$. That is

$$\tilde{w}_{i,j} = (1-\alpha)\cdot \hat{w}_{i-mP,j} + \alpha \cdot \hat{w}_{i-(m+1)P,j}, \tag{10}$$

where $u_z/\delta = m + \alpha$ for some integer m and $0 \leqslant \alpha < 1$.

The transversal displacement of the source is further compensated for by shifting the leaf fluence with necessary corrections such as the cone effect and the inverse square, as detailed in Lu (2008a, 2008b).

3.4. Real-time MAO

The time-consuming plan optimization, usually taking minutes to hours, returns an optimal feasible solution $\hat{\boldsymbol{w}}$ so that the dose distribution $\boldsymbol{d}^{\text{plan}} = \boldsymbol{B}\hat{\boldsymbol{w}}$ best meets the clinician's objectives. However, real-time patient motion is not, or not even possible to be, modeled in the planning procedure. It is expected that the planned, highly conformal dose distribution will be smeared out due to real-time motion. Real-time MAO-guided delivery incorporates optimization into the radiation delivery procedure. In fact, it can be regarded as a negative feedback system that self-corrects cumulative errors resulting from previously delivered projections.

Real-time MAO optimizes the leaf open time for only the upcoming projection. Suppose the planned sinogram consists of N projections and immediately before delivery of the nth projection (the upcoming projection), we want to optimize

$$\tilde{w}_{n,\cdot} = \arg\min_{w_{n,\cdot}} \|\boldsymbol{d}^{\text{delivery}} - \boldsymbol{d}^{\text{plan}}\| \tag{11}$$

subject to

$$0 \leqslant w_{n,\cdot} \leqslant w_{\max}, \tag{12}$$

where the norm $\|\|$ is a quadratic norm, $w_{\max}$ is the maximum projection time fixed for each delivery plan, $\boldsymbol{d}^{\text{delivery}}$ is defined as

$$\boldsymbol{d}^{\text{delivery}} = \boldsymbol{d}_- + \boldsymbol{d}_n + \boldsymbol{d}_+ \tag{13}$$

$\boldsymbol{d}_-$ is the accumulated dose from already delivered projections 1 to $n-1$, $\boldsymbol{d}_n$ is the dose to be optimized for the nth projection delivery and $\boldsymbol{d}_+$ is the future dose to be delivered in projections $n+1$ to N. Here dose distributions $\boldsymbol{d}_-$, $\boldsymbol{d}_n$ and $\boldsymbol{d}_+$ are evaluated in the same reference frame as in the planning procedure. In this formula, we impose only the requirements that the delivery dose ($\boldsymbol{d}^{\text{delivery}}$) be as close to the planned dose ($\boldsymbol{d}^{\text{plan}}$) as possible and the sinogram be deliverable without other constraints, such as DVH constraints, because other constraints are already implemented in deriving $\boldsymbol{d}^{\text{plan}}$ in plan optimization.

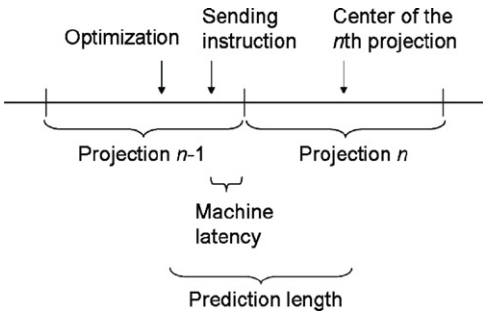

Figure 3. Real-time MAO time line. Optimization for the nth projection occurs in the $(n-1)$th projection. Once optimization is done, the result is sent to the machine for delivery instruction before the nth projection begins.

Figure 3 illustrates the time line for real-time MAO procedures. Optimization for the nth projection occurs during the delivery of the $(n-1)$th projection. The tasks in optimization also include updating the accumulated dose $\boldsymbol{d}_-$ by adding to it last delivered dose of projection $n-1$ and estimating future dose $\boldsymbol{d}_+$ of projection $n+1$ to the end. At the time of optimization, tumor motion is known for projection $n-1$, but motion of projection n needs to be predicted. In the following, we will detail how each task—*delivered dose accumulation*, *future dose estimation* and *single-projection optimization*, is done assuming we are optimizing the nth projection.

3.4.1. Delivered dose accumulation. Assume we want to optimize the nth projection. We have already calculated the accumulated dose $\boldsymbol{d}_-$ of projections 1 to $n-2$. We need to update the accumulated dose $\boldsymbol{d}_-$ by adding to it last delivered dose of the $(n-1)$th projection $\boldsymbol{d}_{n-1}^{\text{delivered}}$. How is $\boldsymbol{d}_{n-1}^{\text{delivered}}$ calculated? We need to know the delivered fluence map $\tilde{w}_{n-1,j}$ and the motion-encoded-beamlets. The delivered fluence map $\tilde{w}_{n-1,j}$ is just the optimized result for the $(n-1)$th projection. However, we only have the planned beamlets, which assume no tumor motion. At the time of delivering the $(n-1)$th projection, the beamlets are different from the planned ones due to tumor motion $\boldsymbol{u}^{n-1}$. The motion-encoded beamlets are therefore calculated according to equations (5) and (7). The dose delivered in projection $n-1$ is

$$\boldsymbol{d}_{n-1}^{\text{delivered}} = \sum_j \tilde{\boldsymbol{B}}_{n-1,j}\tilde{w}_{n-1,j} \tag{14}$$

and the delivered dose is accumulated simply as

$$\boldsymbol{d}_- := \boldsymbol{d}_- + \boldsymbol{d}_{n-1}^{\text{delivered}}. \tag{15}$$

3.4.2. Future dose estimation. Future dose estimation is a tough task because we have two unknowns: (1) tumor motion and (2) the fluence map (leaf open time) for all future projections. If we knew the whole motion trace of future projections, then the MAD strategy described in section 3.3 provides a good candidate for future dose estimation. However, in reality, we do not know the whole motion trace beforehand. A natural amendment is to replace motion by its probability density function (PDF) in the MAD strategy and express future dose in terms of a statistical expectation. The PDF of motion can be easily estimated using the previous trace and it is expected that PDF change slowly for future projections. The expected future dose based on the MAD strategy can be partially off-line pre-calculated. We state the procedure here and defer the proof until the appendix.

Step 1: Calculate the planned projection dose $\boldsymbol{d}_i^{\text{plan}}(\boldsymbol{x})$ for every projection i by adding up the contributions from all leaves

$$\boldsymbol{d}_i^{\text{plan}}(\boldsymbol{x}) = \sum_j \boldsymbol{B}_{ij}(\boldsymbol{x})\hat{w}_{ij}. \tag{16}$$

Step 2: Calculate the cumulated planned dose $\boldsymbol{g}_0(k, \boldsymbol{x})$ up to projection k for all $k \in \{1, 2, \ldots, N\}$

$$\boldsymbol{g}_0(k, \boldsymbol{x}) = \boldsymbol{g}_0(k-1, \boldsymbol{x}) + \boldsymbol{d}_k^{\text{plan}}(\boldsymbol{x}). \tag{17}$$

In particular, $\boldsymbol{g}_0(N, \boldsymbol{x})$ denotes the total planned dose from all projections, where N is the total number of projections

$$\boldsymbol{g}_0(N, \boldsymbol{x}) = \sum\nolimits_{i=1}^{N} \boldsymbol{d}_i^{\text{plan}}(\boldsymbol{x}). \tag{18}$$

Step 3: Calculate the shifted cumulated planned dose $\boldsymbol{g}(n, \boldsymbol{x}, u_z)$ for a shift $-u_z$ of the source trajectory in the longitudinal direction

$$\boldsymbol{g}(n, \boldsymbol{x}, u_z) = (1-\alpha)\boldsymbol{g}_0(n - mP, \boldsymbol{x}_0) + \alpha \boldsymbol{g}_0(n - (m+1)P, \boldsymbol{x}_1), \tag{19}$$

where $u_z/\delta = m + \alpha$, $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ are defined as those in (6). Similarly, the shifted total planned dose is calculated as

$$\boldsymbol{g}(N, \boldsymbol{x}, u_z) = (1-\alpha)\boldsymbol{g}_0(N, \boldsymbol{x}_0) + \alpha \boldsymbol{g}_0(N, \boldsymbol{x}_1). \tag{20}$$

Note that equation (20) holds because the total number of projections N is chosen to cover the range of tumor motion so that the planned sinogram is zero beyond projection $N - (m+1)P$.

Step 4: Calculate the expectation $\bar{\boldsymbol{g}}(n, \boldsymbol{x})$ of $\boldsymbol{g}(n, \boldsymbol{x}, u_z)$ with respect to the PDF $\mathbf{p}(u_z)$ of motion u_z

$$\bar{\boldsymbol{g}}(n, \boldsymbol{x}) = \int_{\underline{u}_z}^{\bar{u}_z} \boldsymbol{g}(n, \boldsymbol{x}, u_z)\mathbf{p}(u_z)\mathrm{d}u_z. \tag{21}$$

Here the PDF $\mathbf{p}(u_z)$ can be obtained from the previous motion trace $\left\{u_z^{n-i}, \ldots, u_z^{n-1}\right\}$ corresponding to, for example, 20 s of data, and we can use the notation $\mathbf{p}\left(u_z \middle| u_z^{n-i}, \ldots, u_z^{n-1}\right)$ for $\mathbf{p}(u_z)$ to indicate such dependency.

Step 5: Calculate the expected future dose $\boldsymbol{d}_+$.

$$\boldsymbol{d}_+(\boldsymbol{x}) = \bar{\boldsymbol{g}}(N, \boldsymbol{x}) - \bar{\boldsymbol{g}}(n, \boldsymbol{x}). \tag{22}$$

Note that steps 1 and 2 are calculated off-line.

Future dose estimation may be improved by using more *a priori* information. For example, dose of the projection immediately after the upcoming projection can be estimated using better probability distribution for the tumor position. Better future dose estimation could improve the results of single-projection optimization.

3.4.3. Single-projection optimization. Now we have calculated $\boldsymbol{d}_-$ and $\boldsymbol{d}_+$, we are ready to optimize the nth projection. The dose of projection n is calculated as

$$\boldsymbol{d}_n = \sum_j \tilde{\boldsymbol{B}}_{n,j} w_{n,j}, \tag{23}$$

where $\tilde{\boldsymbol{B}}_{n,j}$ is the motion-encoded beamlet given in (5) and (7), where the tumor motion $\boldsymbol{u}^n$ is obtained from prediction based on the motion trace from projection 1 to $n-1$ and to be described in the following section. $w_{n,\cdot}$ is the leaf open time to be optimized. There are typically less than 20 active leaves involved in this optimization. The beamlets of those leaves have no primary dose overlap, because they are in the same projection. The desired dose distribution $\tilde{\boldsymbol{d}}_n$ according to (11) and (13) is

$$\tilde{\boldsymbol{d}}_n = \max(\boldsymbol{d}^{\text{plan}} - (\boldsymbol{d}_- + \boldsymbol{d}_+), 0). \tag{24}$$

If the dose distributions of beamlets do not overlap, then the fluence map $\tilde{\boldsymbol{w}}_{n,\cdot} = \{\tilde{w}_{n,j}\}$ can be solved in one step:

$$\tilde{\boldsymbol{d}}_n = \sum_j \tilde{\boldsymbol{B}}_{n,j} \tilde{w}_{n,j} \Rightarrow \tilde{w}_{n,j} = \frac{\langle \tilde{\boldsymbol{B}}_{n,j}, \tilde{\boldsymbol{d}}_n \rangle}{\langle \tilde{\boldsymbol{B}}_{n,j}, \tilde{\boldsymbol{B}}_{n,j} \rangle}. \tag{25}$$

Note that $\langle \tilde{\boldsymbol{B}}_{n,j}, \tilde{\boldsymbol{B}}_{n,j'} \rangle = 0$ for $j \neq j'$ because they do no overlap. Here, dose distributions are regarded as vectors with the dose value at each voxel being a component, and the notation $\langle \cdot, \cdot \rangle$ denotes the inner product of two vectors.

However, because of radiation penumbra and scatter contribution, we found that the following ratio updating scheme (Shepard *et al* 2000) from the kth to $(k+1)$th iteration is more robust:

$$w_{n,j}^{(k+1)} = r_{n,j}^{(k)} w_{n,j}^{(k)}, \tag{26}$$

where

$$r_{n,j}^{(k)} = \frac{\langle \tilde{\boldsymbol{d}}_n, \tilde{\boldsymbol{B}}_{n,j} \rangle}{\langle \tilde{\boldsymbol{d}}_n^{(k)}, \tilde{\boldsymbol{B}}_{n,j} \rangle}. \tag{27}$$

The fluence obtained from the MAD approach is used as the initial guess $w_{n,\cdot}^{(0)}$. And we set $w_{n,j}$ equal to $w_{\max}$ if it exceeds $w_{\max}$ to ensure feasible leaf open time. Typically, the optimization converges within ten iterations.

3.5. Motion prediction

For any real-time reaction system, prediction is needed because of latency between information and action. In the case of real-time MAO for TomoTherapy, the total latency consists of the time from motion detection to MAO calculation, to sending delivery instruction, to the actual delivery. Assuming the projection time is 200 ms, the total latency, or the prediction length, is less than 300 ms, which is considered a short-term prediction. In addition, the MAO-guided delivery is a negative feedback system that can self-correct prediction errors in later projections. Therefore, prediction accuracy is not in high demand and a simple linear prediction based on autoregressive modeling (Sharp *et al* 2004) is found to work well in general for this application.

Let $\boldsymbol{x}_i$ denote the tumor position of the ith time sample. Linear prediction based on autoregressive modeling represents tumor motion as a linear combination of the past samples

$$a_1 x_{m+1} + a_2 x_{m+2} + \cdots + a_n x_{m+n} = x_{m+n+k} \qquad \forall\, m, \tag{28}$$

where n is the model order, k is the prediction length and a_i's are the coefficients that are calculated based on training data. We choose the model order n to correspond to approximately

2 s of motion data. The model order n would depend on the sampling rate, e.g. n is 60 for the sampling rate of 30 Hz. The beginning 40 s of the motion trace is used as training data and the coefficients are adaptively updated every 20 s. According to our experiments on several real breathing curves, the accuracy of linear prediction can achieve about 0.5 mm in root-mean-square errors for real respiration motion with the peak-to-peak amplitude of 3 cm and latency of 300 ms.

We also need to predict PDF of tumor motion for future dose estimation (21). But we simply use ~20 s of past data to approximate PDF of future tumor motion. The accuracy of future dose estimation is again not very crucial due to the self-correcting feature of a close-loop system.

4. Implementation

We implemented the whole algorithm in C++ programming language and integrated it with the TomoTherapy delivery system in the research bunker of TomoTherapy Inc. Information about 'system integration and experimental validation' was presented in ASTRO2008 (Lu *et al* 2008b) and details will be published in a forthcoming paper. For a TomoTherapy plan with the jaw width of 1.05 cm, pitch of 0.3, CT image of size 128 × 128 × 96, and voxel size of 4 × 4 × 2 mm^3, the whole algorithm of real-time MAO, including motion prediction, delivered dose accumulation, future dose estimation and optimization for a single projection, took about 100 ms when using a dual processor workstation with the clock speed of 3 GHz and memory usage of about 1.5 GB. Such performance is sufficient for real-time delivery modification in TomoTherapy.

5. Results

We used both simulated and clinical data to evaluate real-time MAO-guided TomoTherapy delivery. The same motion-encoded dose calculation engine as described in Lu (2008b) was used to calculate the delivered dose. All doses were calculated and DVHs were evaluated in the same reference frame as in the planning. We compared doses of the following three different delivery conditions.

(1) SD: regular delivery with a stationary tumor
(2) MD: regular delivery with a moving tumor
(3) MAO: motion-adaptive-optimization guided delivery with the same tumor motion as in 2.

In addition, we also calculated dose of the MAD technique for comparison.

5.1. Simulated data

For binary MLC-based IMRT, such as TomoTherapy or TopoTherapy (Gonzalez *et al* 2006) delivery, the effects of intra-fraction motion are more pronounced in the longitudinal direction, i.e. the Z-direction or the couch motion direction, than in the transverse direction. Our study first focused on longitudinal motion and then extended to 3D motion. For all simulations, we used the jaw width of 1.05 cm, gantry period of 10 s, projection per rotation (PPR) of 51 and pitch of 0.3. These delivery parameters correspond to the worst case scenario for Topo/TomoTherapy delivery without motion compensation.

We used the 1D simulation as given in Lu (2008a) to study the effects of longitudinal motion. Because the motion was only longitudinal, the results applied to both TopoTherapy and TomoTherapy delivery modes. The simulation setups and dose calculation schemes were

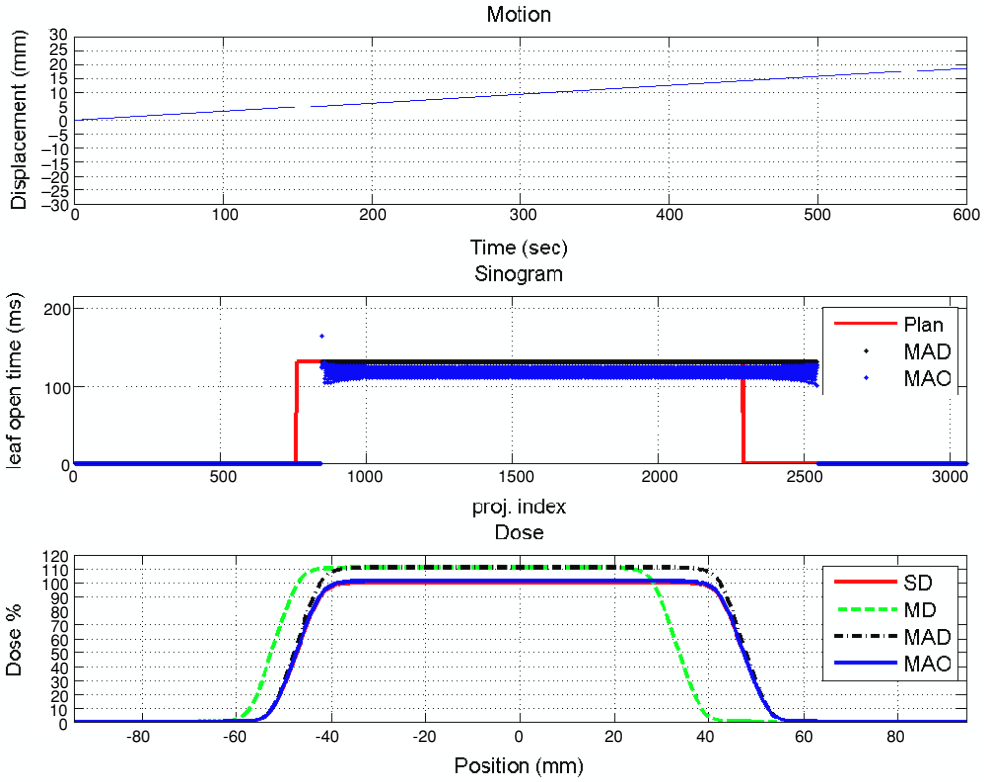

Figure 4. Simulation for real-time MAD and MAO with a given 1D rectangle-shaped fluence map. Top panel: simulated motion of constant speed, 10% of the couch speed. Middle panel: the planned sinogram (red solid line), MAD sinogram (black dot) and MAO sinogram (blue dot) for the tumor motion in the top panel. Bottom panel: comparison of static dose (red solid line, delivered by the planned sinogram for the stationary tumor), MD dose (green-dashed line, delivered by the planned sinogram for the moving tumor), MAD dose (black dot-dashed line, delivered by the MAD sinogram for the moving tumor) and MAO dose (blue solid line, delivered by the MAO sinogram for the moving tumor).

the same as given in Lu (2008a). We compared results of MAD- and MAO-guided deliveries. A uniform dose distribution was intended for an 8-cm-long target. The projection time was 196 ms (= 10 s per rotation/51 projections per rotation). The planned sinogram was a simple rectangle-shaped profile, corresponding to the leaf open time of 131 ms (= 196 ms/MF) in the tumor region and 0 elsewhere. Here, MF (= 1.5) denotes the modulation factor used in the simulation.

Figures 4–6 illustrate the results. In all those figures, the top panels show the longitudinal motion traces, the middle panels show the planned and delivery sinograms and the bottom panels compare doses of SD, MD, MAD and MAO.

Figure 4 simulates the motion of a constant speed (10% of the couch speed). This kind of motion arises when either the tumor moves downstream the radiation beam, i.e. against the couch moving direction, or the couch moves 10% slower than planned. Without compensation, such a motion will result in (1) a part of the target missing and normal tissue overdosed because of the 'dynamic' misalignment and (2) a part of the target approximately 10% overdosed because it spends ∼10% more time than planned under the radiation beam.

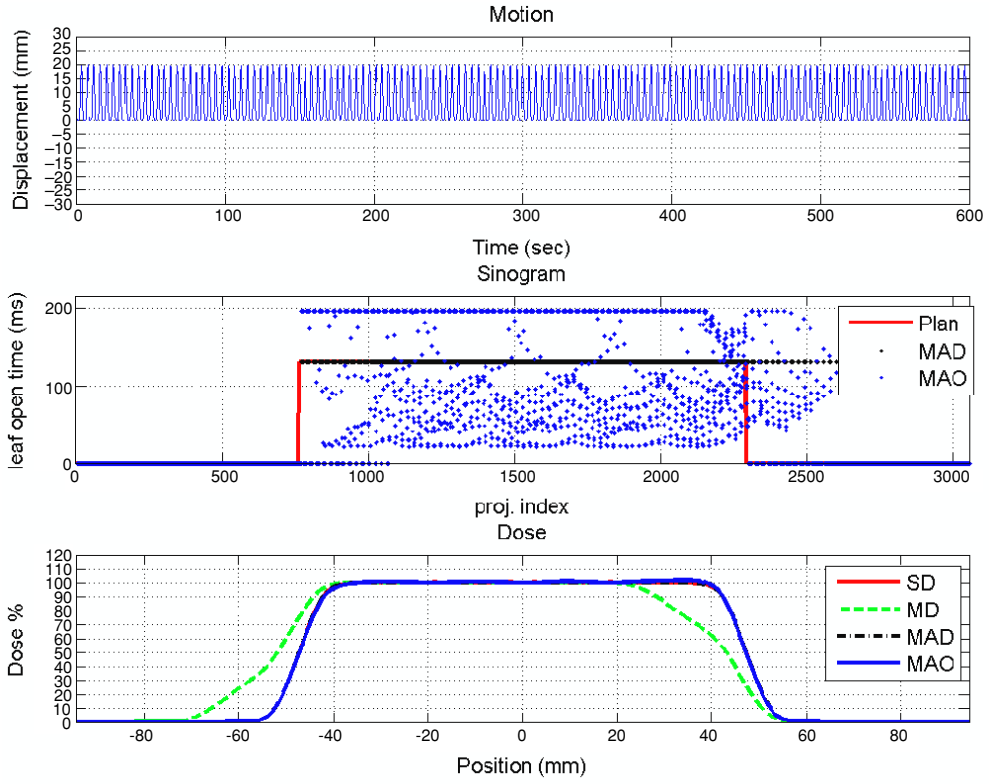

Figure 5. Similar to figure 4, but the motion in the top panel is simulated Lujan-type regular respiration $b + A\cos^{2n}(\pi t/T)$ with $T = 4.3$ s, $A = 20$ mm, $b = 0$ and $n = 3$.

These effects are clearly illustrated by MD doses (green dashed-line in the bottom panels). The MAD dose aligns well with the planned dose (SD) on the target boundary, but it is as 'hot' as the MD dose inside the target. By rearranging the delivery sequence, MAD ensures that radiation is delivered to the right locations but it does not modify beam intensity to compensate for the downstream motion of the tumor. The MAD sinogram, shown as 'dots' in the middle panel of figure 4, has the same intensity level as planned. If we knew the whole motion trace beforehand and did global re-optimization for all projections, we would get the sinogram with intensity 10% lower than planned in the tumor region. The MAO technique re-optimizes the beam intensity to compensate for the cumulative motion-induced error at every projection. As revealed in the middle panel of figure 4, the MAO sinogram oscillates at 10% below the planned sinogram. The oscillation indicates the robustness of the negative feedback system of the MAO method. At each projection, the MAO algorithm compensates for the cumulative errors from all previous projections. Such compensation is by no means perfect due to the limitation of the delivery configuration (e.g. maximum leaf open time per projection) and the uncertainty of prediction and future estimation. (Although in the simulation, the whole motion trace and future delivery were known beforehand, we still followed the workflow described in the 'theory' section. That is, motion was predicted for the upcoming projection and future delivery was estimated based on PDF of previous projections.) But both the compensation error and prediction error can be corrected by the upcoming projections. Through the negative

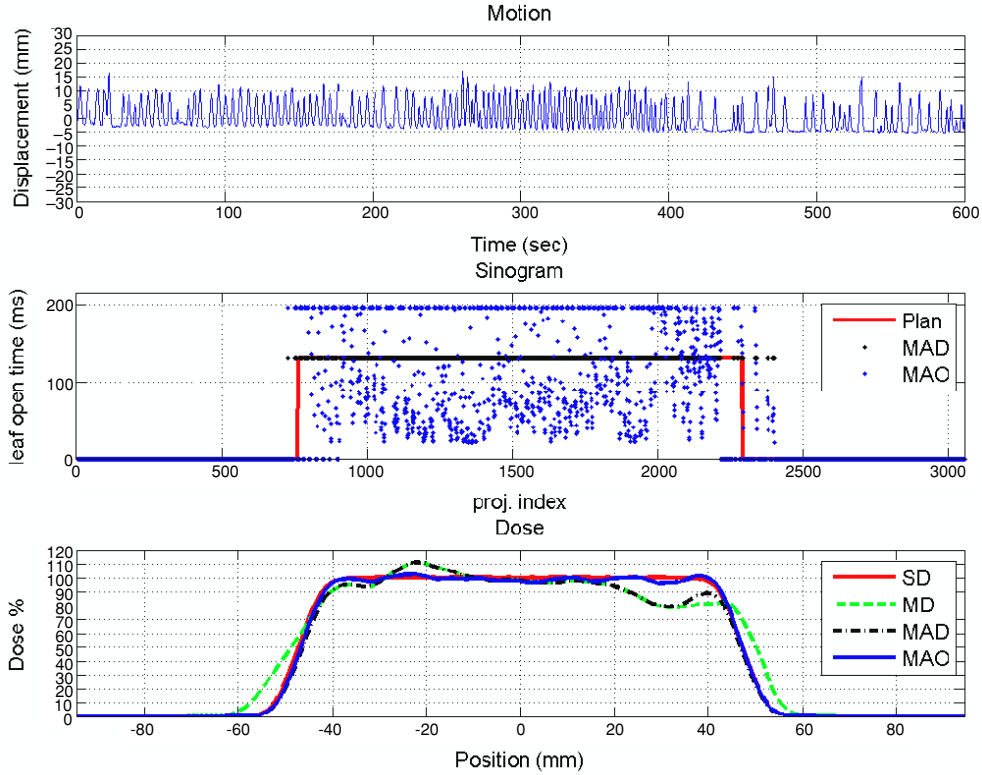

Figure 6. Similar to figure 4 but for a respiratory trace (top panel) measured from a real lung cancer patient.

feed back system, the MAO dose shown in the bottom panel perfectly matches the planned dose (SD).

Figure 5 is for the respiratory motion of the Lujan type (Lujan *et al* 1999) with a peak-to-peak amplitude of 2 cm and period of 4.3 s. As illustrated in the bottom panel, the MD dose deviates significantly from the SD dose in both the front and back ends of the dose profile because of the respiratory motion. On the other hand, both the MAD and MAO methods compensate for the motion, and their respective doses match the SD dose very well. Although the MAD and MAO doses are almost identical, their sinograms are dramatically different (middle panel of figure 5). The MAD sinogram has the same flat shape as the planned, while the MAO sinogram shows large oscillation. Different sinograms may give almost identical doses, because the set of beamlets is highly redundant. For PPR of 51 and pitch of 0.3, each voxel will receive direct irradiation from $51/0.3 = 170$ projections. Such redundancy implies the existence of multiple solutions for the same objective and feasibility of compensation for cumulative errors using the greedy-based optimization scheme of MAO. Note that in the case of perfectly regular motion as in the above example, MAD may perform better than MAO. This is expected, because in MAD there is no estimation step that would add uncertainty.

Figure 6 shows the result of a very irregular respiration with the peak-to-peak amplitude of 2 cm measured from a lung cancer patient. The combination of the large amplitude of very irregular respiration, small field size (1.05 cm) and fast gantry rotation (10 s) of TomoTherapy

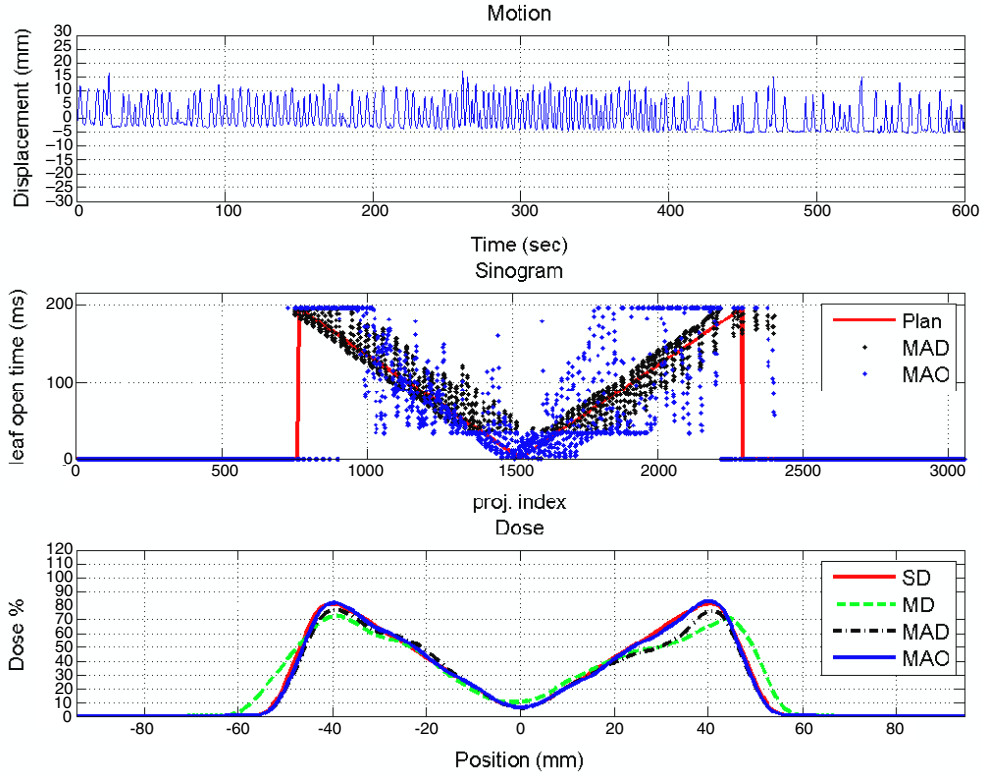

Figure 7. Similar to figure 6 but for an M-shaped plan sinogram.

delivery results in high tumor dose non-uniformity in the MD dose as shown in the bottom panel. The MAD dose again matches the SD dose well on the tumor boundary, but has the same non-uniformity as the MD dose in the inner region of the tumor. On the other hand, the MAO method reduces edge blurring and tumor dose non-uniformity effects due to irregular respiration to within 3% difference from the SD dose. Figure 7 shows the results of the same irregular respiration as in figure 6 but for an M-shaped intensity map, which simulates a highly intensity-modulated plan. It shows that MAO dose matches the SD dose well even with such a highly modulated plan.

MAD failed for the types of motion in figures 4 and 6 because it is based on the assumption that *each tumor voxel has the same chance to be irradiated as what was planned*. Both cases, the motion of constant velocity (figure 4) and very irregular respiration (figure 6), violate that assumption substantially. The real-time MAO method, however, is not based on such an assumption, and therefore works well for arbitrary motion.

5.2. Clinical data

The real-time MAO-guided delivery technique can be applied to any kind of motion. Here we present the results of two kinds of motion: the respiratory motion and the prostate motion.

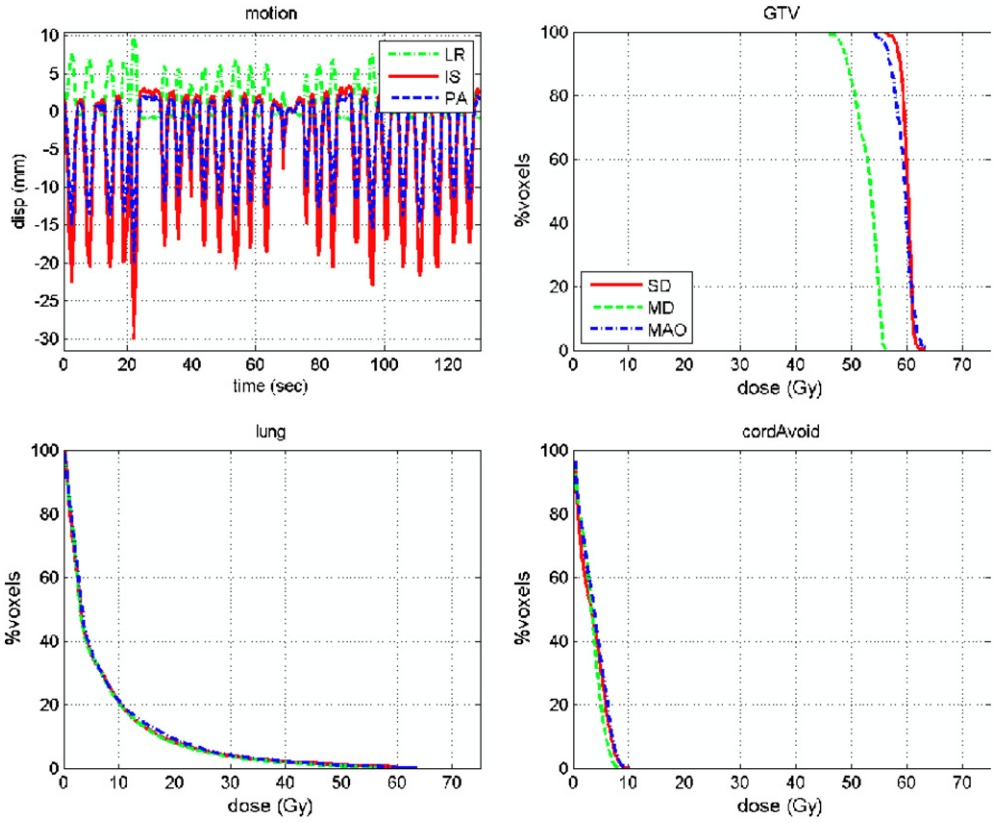

Figure 8. DVH comparisons of different delivery methods, SD (solid line), MD (dashed line) and MAO (dot-dashed line), for a lung cancer patient under TomoTherapy treatment. The top left panel shows the respiration motion trace with the peak-to-peak amplitude of about 30 mm in the SI direction, 20 mm in the AP direction and 10 mm in the LR direction. The three other panels compare DVHs of GTV, the lung and spinal cord with avoidance region.

5.2.1. Respiratory motion. We studied retrospectively a lung cancer case with a tumor size of approximately 2 cm in the inferior part of the lung. A TomoTherapy plan was optimized using TomoTherapy® HiArt® II TPS with the jaw width of 2.5 cm and pitch of 0.3. The optimization used the GTV as the target without any motion margin. Because we did not have the respiration data for that patient, two different respiration traces of the spirometer signals from other lung cancer patients were used to simulate the tumor motion of the studied patient. The results of these two different respiration traces are similar, so we only present one of them (figures 8 and 9). The measured spirometer signals were 1D only and provided the relative amplitude and phase information. We scaled the signal amplitude so that the range of its lower 10% to upper 10% roughly corresponds to 3 cm in the SI direction, 2 cm in the AP direction and 1 cm in the LR direction. These motions are close to the maximum respiratory motion reported in the literature. The motion traces are shown in the top left panel of figure 8. The treatment planning is based on the maximum expiration phase, which is used as the reference phase for all dose mappings. The planned dose distribution is shown in the top row of figure 9. We calculated the MD and MAO dose and compared them with the SD dose. We also compared their DVHs.

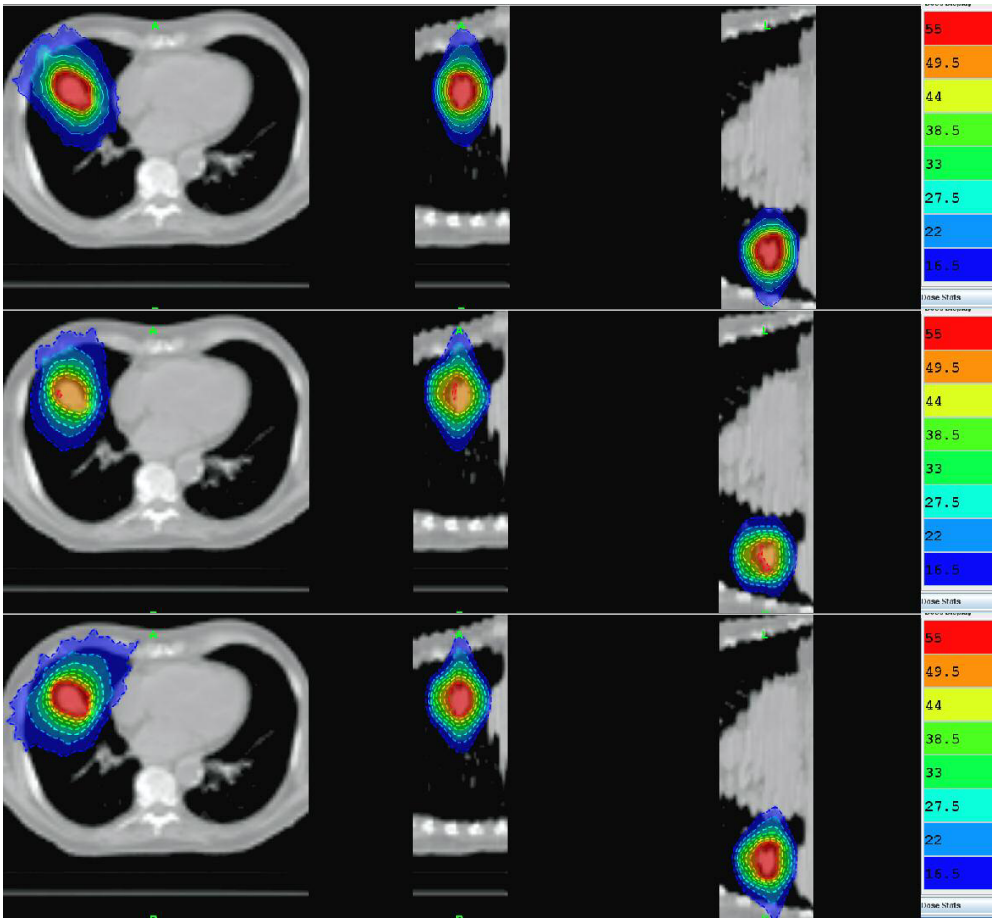

Figure 9. Dose distributions for the case shown in figure 8. The top row shows the transversal (T), sagittal (S) and coronal (C) views of the SD dose distribution, the middle row, TSC views of the MD dose distribution, and the bottom row, TSC views of the MAO dose distribution. The units of isodose lines are Gy.

The motion trace and DVH comparisons for different delivery modes are shown in figure 8. There are significant cold spots in the MD dose. Note that this is an extreme case because the GTV is only about 2 cm and the motion is as large as 3 cm. In addition, without motion margin in the treatment plan, it is expected that significant cold spots will show up because the tumor may move out of the radiation field. But even with no motion margin, the DVHs of MAO match the planned DVHs (SD) very well with negligible cold spots. The results indicate that the real-time MAO technique is an effective way to reduce the margin for treatment of a small lung tumor that undergoes significant respiratory motion.

Figure 9 compares the dose distribution in the transversal (T), sagittal (S) and coronal (C) views for the motion trace given in figure 8. The top rows are the TSC views of the SD (planned) dose distribution; the middle row, the TSC views of the MD dose distribution, and the bottom row, the TSC views of the MAO dose distribution. They provide similar information as that of DVHs. Without motion compensation, the GTV dose is significantly less than the planned dose as shown in the isodose levels. With MAO-guided delivery, the

isodose lines in the high dose region (surrounding the tumor) are very similar to those of the SD dose, though there is some discrepancy in the low dose region. Such discrepancy is mainly due to scatter contribution, which is not fully modeled in MAO because of memory limitation. Also note that in all calculations, we assume a rigid-body motion, which is quite valid for a small tumor, but not for the lung or the spinal cord. Therefore, the results of the lung dose and spinal cord dose should be read with caution, though we would expect that the difference is minimal when the tumor is small and far away from the spinal cord.

5.2.2. Prostate motion. We studied retrospectively the same prostate patient cases as reported in Langen *et al* (2008a, 2009). The intra-fraction prostate motions were real-time tracked via the electromagnetic signals of a four-dimensional localization system of Calypso®. These motions were classified into small, medium or large prostate displacements, which correspond to amplitudes of less than 3 mm, between 3 and 5 mm or larger than 5 mm and account for more than 85%, ∼10% or ∼3% of all tracked data, respectively, as described in Langen *et al* (2008b). The same optimized TomoTherapy plans as in Langen *et al* (2008a, 2009) with the jaw width of 2.47 cm, pitch of 0.287 and gantry rotation period of 29–31 s were used.

For all small prostate motions, the MAO technique works well (not presented in this paper).

Figure 10 shows the results of a medium prostate motion. The prostate moved between 3 mm and 6 mm in the SI direction for a significant amount of time. Both hot and cold spots appeared in the prostate DVH of MD compared with that of SD, because the tumor moved both upstream (against the radiation source motion) and downstream (following the radiation source motion) in the SI direction, as indicated by the SI motion trace in the top left panel. The upstream motion caused parts of the prostate receiving lower-than-planned dose, while the downstream motion caused other parts higher-than-planned dose. The MAO technique was able to compensate for both motions as illustrated in the DVH plots. The prostate DVH of MAO approached that of SD well. In addition, the hot spots in the bladder were corrected by MAO.

Figures 11 and 12 show the results for a large prostate motion. The top row of figure 12 shows the planned (SD) dose distribution in TSC views, the middle row, the MD dose distribution, and the bottom row, the MAO dose distribution. The prostate moved between 5 mm and 10 mm in both the SI and AP directions for most delivery time. Such large upstream SI motion caused significant cold spots in the prostate dose distribution, detected by the prostate DVH (figure 11) and dose distribution of MD (the middle row of figure 12). On the other hand, the MAO technique compensated for such large motion quite well. The MAO iso-dose lines (the bottom row of figure 12) were very similar to the planned dose (the top row of figure 12) and the DVHs of MAO approached those of SD (figure 11). Note that the hot spots in the rectum were corrected by MAO as well.

The above examples of prostate motions were effectively compensated for by real-time MAO. On the other hand, such motion effects are usually quite mild and quickly washed out after five fractions of delivery (Langen *et al* 2008a, 2009,). A word of caution is that here we assumed rigid-body motion for all studies. Such assumption is arguable for the prostate and definitely not suitable for the bladder and the rectum. Therefore, both the MD and MAO results presented here should be read only as proof of concept rather than clinical guidance.

6. Discussion and directions for future work

Treatment planning is based on information about the commissioning data of the delivery machine and the patient anatomy such as CT. Because the goal of treatment delivery is

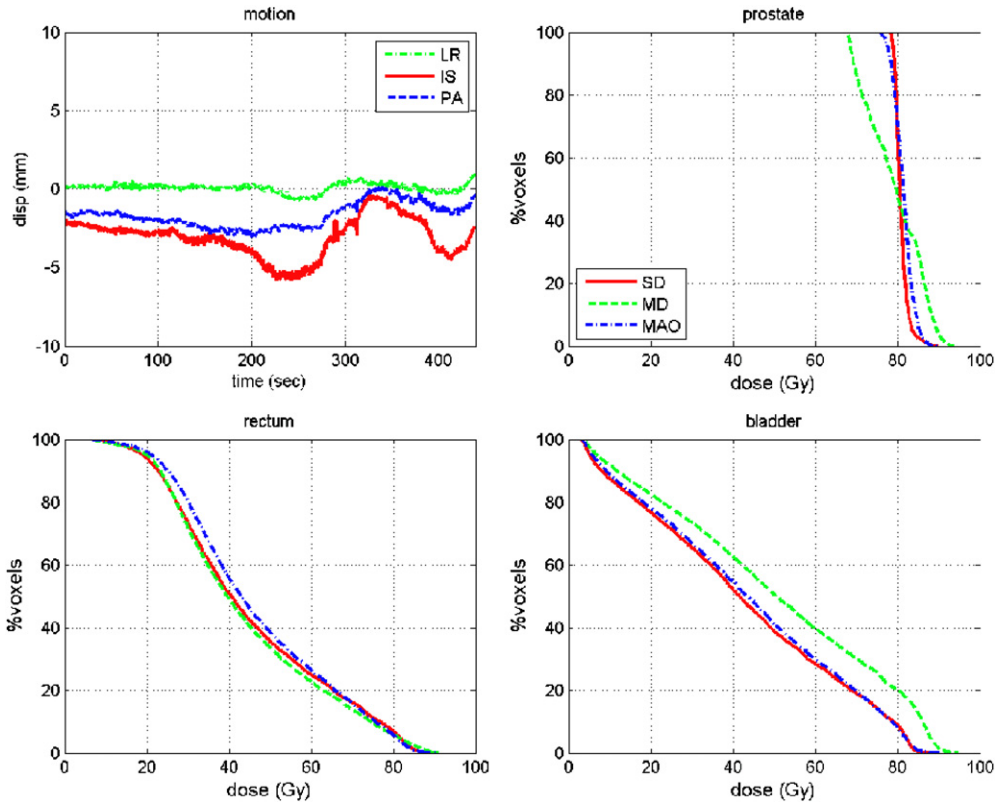

Figure 10. DVH comparisons of different delivery methods, SD (solid line), MD (dashed line) and MAO (dot-dashed line), for a prostate cancer patient under TomoTherapy treatment. The top left panel shows the prostate motion traces in the LR, SI and PA directions. The remaining panels compare DVHs of the prostate, rectum and bladder.

to reproduce the treatment plan as accurately as possible, the delivery procedure should be accurately modeled in treatment planning. However, real-time changes, such as tumor motion, are hard to accurately model in advance. The state-of-the-art radiation delivery is an open-loop procedure. It tries to reproduce the planned procedures step-by-step, but lacks the mechanism to deal with the error occurred and accumulated in each step. The proposed real-time MAO-guided radiotherapy changes the delivery scheme from a conventional open-loop system to a closed-loop system with negative feedback. In this sense, the framework of MAO can be easily extended to a more general scheme, real-time adaptive radiation therapy (ART), as illustrated in figure 13. ART (Yan *et al* 1997a, 1997b) generally refers to the concepts of using feedback during the course of radiation therapy to improve future treatment. Feedback can be used for off-line adaptive processes (Lu *et al* 2006b, Yan *et al* 2005) or on-line (Olivera *et al* 2005, Ahunbay *et al* 2008, Wu *et al* 2008) processes. Off-line ART refers to processes when the patient is not being treated, such as in between treatment fractions, whereas on-line ART refers to processes when the patient is on the treatment couch but right before delivery of the treatment beam. Both off-line and on-line ART are to compensate for the inter-fractional changes. Real-time ART, however, is to correct intra-fractional, or real time, generated errors, whether they are due to patient motion or random machine variations, such as linac output

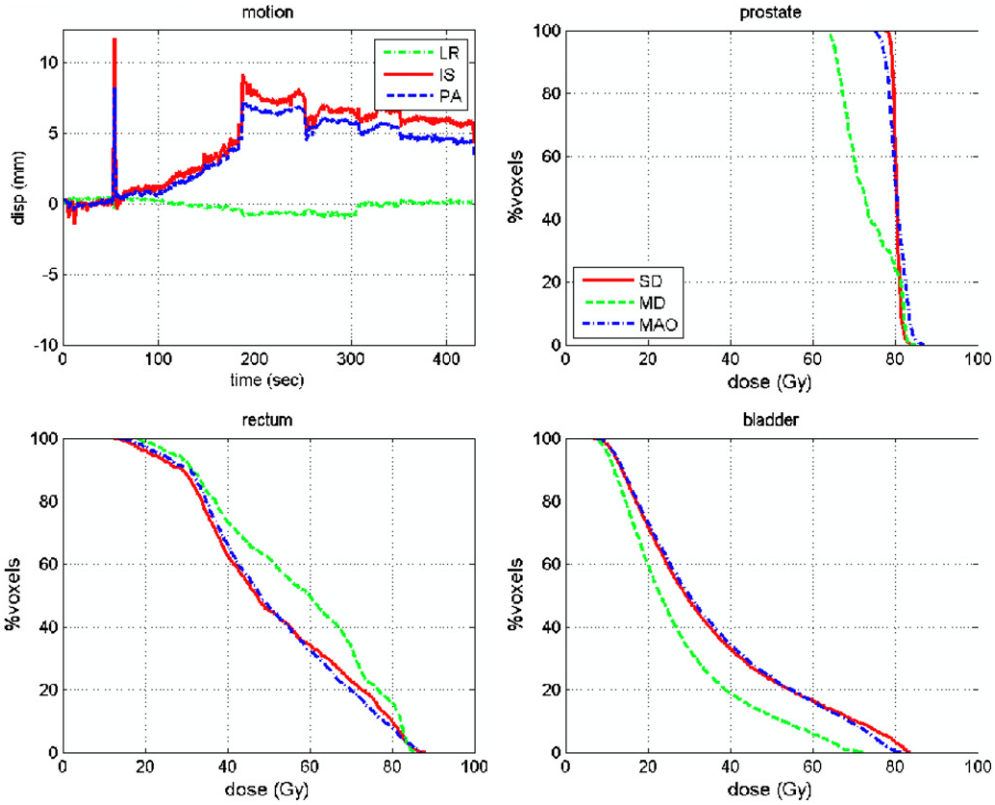

Figure 11. Similar to figure 10 but for a different prostate patient undergoing different intra-fraction prostate motion.

changes, leaf open errors, gantry rotation errors, couch motion errors, etc. Real-time ART imposes high demand on error detection and system response. Real-time dose reconstruction will also be a critical component for real-time ART. The real-time optimization workflow we proposed in this paper should be able to accomplish real-time ART for TomoTherapy delivery, provided we establish reliable implementation of motion detection, delivery verification and dose reconstruction (Kapatoes *et al* 2001).

Although we only studied the MAO technique for Tomotherapy delivery, the proposed real-time optimization scheme can be extended to other delivery modalities, provided the radiation delivery can be 'discretized' in time with sub-second resolution (system response time) and motion-encoded dose can be calculated in real time. The key component is to pre-calculate beamlets that cover regions of possible tumor motion. That is, all beamlets that cover PTV (including motion margin) need to be pre-calculated for efficient motion-encoded dose calculation, even though the radiation beam portal in initial treatment plan only covers the CTV (not including motion margin). The implementation loop would still be valid for other delivery modality: (1) real-time detecting and cumulating errors; (2) real-time optimizing for the upcoming 'projection', taking into account the cumulated errors and future prediction; (3) delivering the real-time optimized projection. The loop is (1) → (2) → (3) → (1) → (2) → (3) → ....

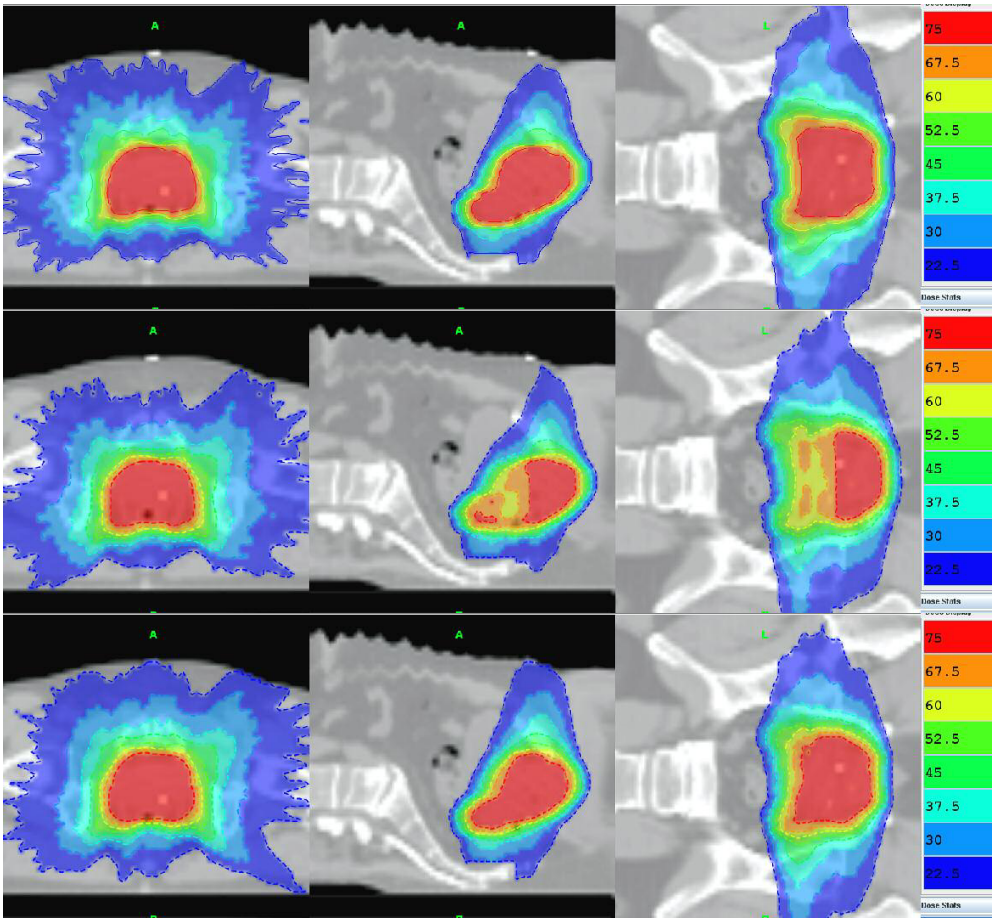

Figure 12. Dose distributions for the case shown in figure 11. The top row shows the transversal (T), sagittal (S) and coronal (C) views of the SD dose distribution, the middle row, TSC views of the MD dose distribution, and the bottom row, TSC views of the MAO dose distribution. The units of isodose lines are Gy.

Intra-fractional motion is challenging for conventional fractionated IMRT. But it is even more challenging for stereotactic body radiation treatment (SBRT) and hypofractionated therapy (Papiez and Timmerman 2008), where precise target positioning is critical. SBRT is increasingly being used in certain lung cancers because of its rate of local control (McGarry *et al* 2005). Hypofractionation for prostate cancer has the potential of therapeutic gain as well as economic advantage (Ritter *et al* 2006). Both the lung and prostate demonstrate significant intra-fractional motion, though their motions are very different in character, amplitude and frequency. Intensity-modulated proton (or other heavy particles) therapy (Flynn *et al* 2008, Muzik *et al* 2008, Thorwarth *et al* 2008, Oelfke and Bortfeld 2003, Yeboah and Sandison 2002), of course, demands most critical motion management. The proposed real-time MAO technique can potentially be applied in all those three therapy modes.

The proposed MAO technique still falls in the category of 'compensation' that regards motion as 'error' to be corrected. Motion is a challenge, yet motion is also a chance. Just as we can take advantage of inter-fractional variations of the tumor-OAR configuration to achieve

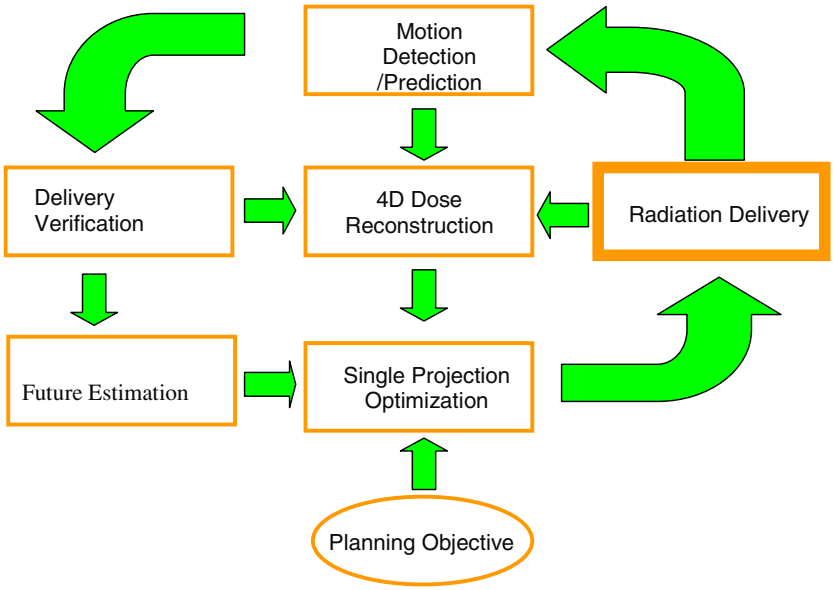

Figure 13. Workflow for real-time ART. In this workflow, both motion errors and machine errors are detected and predicted in real time. A 4D dose reconstruction engine accumulates the delivered dose in real time. The reconstructed dose, together with future estimation, is used to drive the real-time optimizer that optimizes the leaf open time for the upcoming projection.

better therapeutic gain through 'adaptive fractionation therapy' (Lu *et al* 2008a, Chen *et al* 2008), we can potentially take advantage of intra-fraction motion to achieve a 'better-than-planned' delivery. This possibility was demonstrated by Papiez's group (Papiez *et al* 2007) using 4D DMLC delivery to minimize the OAR dose. We believe that a more advanced real-time optimization scheme should be able to offer a superior delivery, which is not achievable via any plan based on static delivery. Such a scheme may require more sophisticated algorithms and a powerful computer. It definitely needs further exploration.

In this paper, we assumed that we know tumor motion in real time. However, whether we can really detect the internal motion in real time remains a hot research topic, since none of the available techniques is 100% reliable. We refer the readers to the review paper by Webb 2006a about the pros and cons of all kinds of motion detection techniques. As discussed in the previous papers (Lu 2008b, 2008a), we would suggest that an internal-fiducial-based method (Shirato *et al* 2000b, Chen *et al* 2001, Seppenwoolde *et al* 2002, Rietzel *et al* 2004, Willoughby *et al* 2006, Kupelian *et al* 2007, Shchory *et al* 2008), or direct treatment beam based method (Lu *et al* 2006c) be used for real-time motion detection, because the MAO technique presented in this paper relies on the direct position information of the tumor in real time.

In this paper, we assumed that tumor motion could be regarded as a rigid-body shift and OAR moves in the same way as the tumor. This is a good approximation for a small tumor and for OAR that is close to the tumor. Motion compensation is more significant for small tumors than for large tumors, and OAR close to the tumor is usually of more concern than those far away from the tumor. Therefore, such assumptions are valid for the majority of tumor motion cases. For other tumor motion, such as rotation or general deformation, it requires different formulae to calculate motion-encoded beamlets. The problem is then how

to faithfully delineate the deformation in real time. This may involve an ultra-fast deformable registration algorithm (Samant *et al* 2008, Lu *et al* 2004), or using pre-calculated deformation maps that are based on some 4D images such as 4DCT (Lu *et al* 2006c).

7. Conclusion

We presented a novel technique for real-time motion adaptive optimization within current TomoTherapy® hardware. The method re-optimizes leaf open time for upcoming projections in real time, in response to observed target motion. MAO is well suited to correct regular and irregular breathing-type motion as well as for drift and random motion observed for prostate patients. Calculations based on simulated and clinical data were used to validate the MAO approach. Validation measurements are underway.

Acknowledgment

The authors would like to acknowledge Dr Low of Washington University-St Louis for providing the respiration traces used in this paper.

Appendix

In this appendix, we derive the formulae for future dose estimation given in section 3.4.2. Our future dose estimation is based on MAD for fluence and PDF of motion.

Because the transversal displacement of the source can be compensated for by shifting the leaf fluence with necessary corrections such as the cone effect and the inverse square, we only need to deal with the longitudinal displacement. As described in section 3.3, if the source is shifted by $-u_z$ in the longitudinal direction for projection i, then the new leaf fluence $\tilde{\boldsymbol{w}} = \{\tilde{w}_{i,j}(u_z)\}$ can be calculated by the linear interpolation of the planned sinogram $\hat{\boldsymbol{w}}$. That is

$$\tilde{w}_{i,j}(u_z) = (1-\alpha)\cdot\hat{w}_{i-mP,j} + \alpha\cdot\hat{w}_{i-(m+1)P,j,} \tag{29}$$

where $u_z/\delta = m+\alpha$ for some integer m and $0 \leqslant \alpha < 1$. Similarly, the beamlets can be approximated by linearly interpolating the planned beamlets as in (5)

$$\tilde{\boldsymbol{B}}_{i,j}(\boldsymbol{x}) = (1-\alpha)\boldsymbol{B}_{i-mP,j}(\boldsymbol{x}') + \alpha\boldsymbol{B}_{i-(m+1)P,j}(\boldsymbol{x}''), \tag{30}$$

where $\boldsymbol{x}' = (x, y, z+\alpha\delta)$ and $\boldsymbol{x}'' = (x, y, z-(1-\alpha)\delta)$. Therefore the new dose for projection i is

$$\begin{aligned}\tilde{\boldsymbol{d}}_i(\boldsymbol{x}) &= \sum_j \tilde{\boldsymbol{B}}_{i,j}(\boldsymbol{x})\tilde{w}_{i,j} \\ &= (1-\alpha)\cdot\sum_j \tilde{\boldsymbol{B}}_{i,j}(\boldsymbol{x})\hat{w}_{i-mP,j} + \alpha\cdot\sum_j \tilde{\boldsymbol{B}}_{i,j}(\boldsymbol{x})\hat{w}_{i-(m+1)P,j} \\ &\approx (1-\alpha)\boldsymbol{d}_{i-mP}(\boldsymbol{x}') + \alpha\cdot\boldsymbol{d}_{i-(m+1)P}(\boldsymbol{x}'').\end{aligned} \tag{31}$$

Note the third line in equation (31) uses the approximations $\tilde{\boldsymbol{B}}_{i,j}(\boldsymbol{x}) \approx \boldsymbol{B}_{i-mP,j}(\boldsymbol{x}')$ and $\tilde{\boldsymbol{B}}_{i,j}(\boldsymbol{x}) \approx \boldsymbol{B}_{i-(m+1)P,j}(\boldsymbol{x}'')$. The new dose $\tilde{\boldsymbol{d}}_i(\boldsymbol{x})$ is in fact a function of tumor motion. Let us use $\tilde{\boldsymbol{d}}_i\left(u_z^i, \boldsymbol{x}\right)$ to denote its dependency on the tumor motion u_z^i of the ith projection. If motion for future projections is known, then future dose after projection n can be calculated as

$$D_+(\boldsymbol{x}) = \sum_{i=n+1}^{N} \tilde{\boldsymbol{d}}_i\left(u_z^i, \boldsymbol{x}\right). \tag{32}$$

However, motion for future projections is not known in real life. Assuming the PDF's of u_z^i are available and the same for all $i > n$, denoted by $\mathbf{p}(u_z)$, then future dose can be estimated using the expectation of equation (32)

$$\begin{aligned} \boldsymbol{d}_+(\boldsymbol{x}) &= \langle D_+(\boldsymbol{x}) \rangle \\ &= \left\langle \sum_{i=n+1}^{N} \tilde{\boldsymbol{d}}_i\left(u_z^i, \boldsymbol{x}\right) \right\rangle \\ &= \left\langle \boldsymbol{g}\left(N, \boldsymbol{x}, u_z\right) - \boldsymbol{g}(n, \boldsymbol{x}, u_z) \right\rangle \\ &= \bar{\boldsymbol{g}}(N, \boldsymbol{x}) - \bar{\boldsymbol{g}}(n, \boldsymbol{x}) \end{aligned} \tag{33}$$

where $\boldsymbol{g}$ and $\bar{\boldsymbol{g}}$ are defined in section 3.4.2.

References

Ahunbay E E, Peng C, Chen G P, Narayanan S, Yu C, Lawton C and Li X A 2008 An on-line replanning scheme for interfractional variations *Med. Phys.* **35** 3607–15

Balter J M, Sandler H M and Lam K 1995 Measurement of prostate movement over the course of routine radiotherapy using implanted markers *Int. J. Radiat. Oncol. Biol. Phys.* **31** 113–8

Chen M, Lu W, Chen Q, Ruchala K and Olivera G 2008 Adaptive fractionation therapy: II. Biological effective dose *Phys. Med. Biol* **53** 5513–25

Chen Q S, Weinhous M S, Deibel F C, Ciezki J P and Macklis R M 2001 Fluoroscopic study of tumor motion due to breathing: facilitating precise radiation therapy for lung cancer patients *Med. Phys.* **28** 1850–6

D'Souza W D, Naqvi S A and Yu C X 2005 Real-time intra-fraction-motion tracking using the treatment couch: a feasibility study *Phys. Med. Biol.* **50** 4021–33

Flynn R T, Bowen S R, Bentzen S M, Rockwell Mackie T and Jeraj R 2008 Intensity-modulated x-ray (IMXT) versus proton (IMPT) therapy for theragnostic hypoxia-based dose painting *Phys. Med. Biol.* **53** 4153–67

Gonzalez V J, Buchholz D J, Langen K M, Olivera G H, Chauhan B, Meeks S L, Ruchala K J, Haimerl J, Lu W and Kupelian P A 2006 Evaluation of two tomotherapy-based techniques for the delivery of whole-breast intensity-modulated radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **65** 284–90

Kapatoes J M, Olivera G H, Ruchala K J, Smilowitz J B, Reckwerdt P J and Mackie T R 2001 A feasible method for clinical delivery verification and dose reconstruction in tomotherapy *Med. Phys.* **28** 528–42

Keall P J, Cattell H, Pokhrel D, Dieterich S, Wong K H, Murphy M J, Vedam S S, Wijesooriya K and Mohan R 2006 Geometric accuracy of a real-time target tracking system with dynamic multileaf collimator tracking system *Int. J. Radiat. Oncol. Biol. Phys.* **65** 1579–84

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, Todor A D, Vedam S S, Bartee C L, Siebers J V, Kini V R and Mohan R 2004 On the use of EPID-based implanted marker tracking for 4D radiotherapy *Med. Phys.* **31** 3492–9

Kupelian P *et al* 2007 Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **67** 1088–98

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

Langen K M, Lu W, Ngwa W, Willoughby T R, Chauhan B, Meeks S L, Kupelian P A and Olivera G 2008a Correlation of dosimetric effect and intrafraction motion during prostate treatments delivered with helical tomotherapy *Phys. Med. Biol.* **53** 7073–86

Langen K M, Lu W, Ngwa W, Willoughby T R, Chauhan B, Meeks S L, Kupelian P A and Olivera G 2009 Dosimetric effect of prostate motion during helical tomotherapy plans *Int. J. Radiat. Oncol. Biol. Phys.* at press (corrected proof, available online 21 February 2009)

Langen K M, Willoughby T R, Meeks S L, Santhanam A, Cunningham A, Levine L and Kupelian P A 2008b Observations on real-time prostate gland motion using electromagnetic tracking *Int. J. Radiat. Oncol. Biol. Phys.* **71** 1084–90

Lu W 2008a Real-time motion-adapted-delivery (MAD) using binary MLC: I. Static beam (topotherapy) delivery *Phys. Med. Biol.* **53** 6491–511

Lu W 2008b Real-time motion-adapted-delivery (MAD) using binary MLC: II. Rotational beam (tomotherapy) delivery *Phys. Med. Biol.* **53** 6513–31

Lu W, Chen M, Chen Q, Ruchala K and Olivera G 2008a Adaptive fractionation therapy: I. Basic concept and strategy *Phys. Med. Biol.* **53** 5495–511

Lu W, Mauer C, Chen M, Ruchala K, Zhang J, Lucas D, Chen Q and Olivera G 2008b Experimental validation for real-time motion adapted optimization (MAO) guided delivery *Int. J. Radiat. Oncol. Biol. Phys. (ASTRO2008)* **72** s26

Lu W, Chen M, Olivera G H, Ruchala K and Mackie T R 2004 Fast free-form deformable registration via calculus of variations *Phys. Med. Biol.* **49** 3067–87

Lu W, Olivera G, Chen Q, Chen M, Pan Y, Schnarr E and Ruchala K 2006a Motion encoded beamlets for optimization and evaluation in four-dimensional (4D) radiotherapy *Med. Phys. (AAPM2006)* **33** 2042

Lu W, Olivera G H, Chen Q, Ruchala K J, Haimerl J, Meeks S L, Langen K M and Kupelian P A 2006b Deformable registration of the planning image (KVCT) and the daily images (MVCT) for adaptive radiation therapy *Phys. Med. Biol.* **51** 4351–74

Lu W, Ruchala K J, Chen M-L, Chen Q and Olivera G H 2006c Real-time respiration monitoring using the radiotherapy treatment beam and four-dimensional computed tomography (4DCT)-a conceptual study *Phys. Med. Biol.* **51** 4469–95

Lujan A E, Larsen E W, Balter J M and Ten Haken R K 1999 A method for incorporating organ motion due to breathing into 3D dose calculation *Med. Phys.* **26** 715–20

Mackie T R, Balog J, Ruchala K, Shepard D, Aldridge S, Fitchard E, Reckwerdt P, Olivera G, McNutt T and Mehta M 1999 Tomotherapy *Semin. Radiat. Oncol.* **9** 108–17

McGarry R, Papiez L, Williams M, Whitford T and Timmerman R 2005 Stereotactic body radiation therapy of early-stage non-small-cell lung carcinoma: phase I study *Int. J. Radiat. Oncol. Biol. Phys.* **63** 1010–5

McMahon R, Papiez L and Rangaraj D 2007a Dynamic-MLC leaf control utilizing on-flight intensity calculations: a robust method for real-time IMRT delivery over moving rigid targets *Med. Phys.* **34** 3211–23

McMahon R, Papiez L and Sandison G 2007b Addressing relative motion of tumors and normal tissue during dynamic MLC tracking delivery *Australas. Phys. Eng. Sci. Med.* **30** 331–6

Murphy M J 2004 Tracking moving organs in real time *Semin. Radiat. Oncol.* **14** 91–100

Muzik J, Soukup M and Alber M 2008 Comparison of fixed-beam IMRT, helical tomotherapy, and IMPT for selected cases *Med. Phys.* **35** 1580–92

Neicu T, Shirato H, Seppenwoolde Y and Jiang S B 2003 Synchronized moving aperture radiation therapy (SMART): average tumour trajectory for lung patients *Phys. Med. Biol.* **48** 587–98

Nuyttens J J, Prevost J B, Praag J, Hoogeman M, Van Klaveren R J, Levendag P C and Pattynama P M 2006 Lung tumor tracking during stereotactic radiotherapy treatment with the CyberKnife: marker placement and early results *Acta Oncol.* **45** 961–5

Oelfke U and Bortfeld T 2003 Optimization of physical dose distributions with hadron beams: comparing photon IMRT with IMPT *Technol. Cancer Res. Treat.* **2** 401–12

Olivera G H, Mackie T R, Ruchala K, Lu W, Jeswani S, Aoyama H, Kapatoes J and Chen Q 2005 Helical tomotherapy process *Japan. J. Med. Phys.* **25** 39–63

Papiez L 2004 DMLC leaf-pair optimal control of IMRT delivery for a moving rigid target *Med. Phys.* **31** 2742–54

Papiez L, DesRosiers C and Moskvin V 2002 Very high energy electrons (50–250 MeV) and radiation therapy *Technol. Cancer Res. Treat.* **1** 105–10

Papiez L and Langer M 2006 On probabilistically defined margins in radiation therapy *Phys. Med. Biol.* **51** 3921–39

Papiez L, McLellan J, Sandison G A, Sawchuk S, Lu X and Battista J J 1994 Inclusion of energy straggling in a numerical method for electron dose calculation *Med. Phys.* **21** 1591–8

Papiez L, McMahon R and Timmerman R 2007 4D DMLC leaf sequencing to minimize organ at risk dose in moving anatomy *Med. Phys.* **34** 4952–6

Papiez L, Montebello J, DesRosiers C and Papiez E 1999 The clinical application of dynamic shielding and imaging in moving table total body irradiation *Radiother. Oncol.* **51** 219–24

Papiez L and Rangaraj D 2005 DMLC leaf-pair optimal control for mobile, deforming target *Med. Phys.* **32** 275–85

Papiez L, Rangaraj D and Keall P 2005 Real-time DMLC IMRT delivery for mobile and deforming targets *Med. Phys.* **32** 3037–48

Papiez L and Timmerman R 2008 Hypofractionation in radiation therapy and its impact *Med. Phys.* **35** 112–8

Rietzel E, Rosenthal S J, Gierga D P, Willet C G and Chen G T 2004 Moving targets: detection and tracking of internal organ motion for treatment planning and patient set-up *Radiother. Oncol.* **73** (Suppl. 2) S68–72

Ritter M A, Forman J D, Petereit D G, Kupelian P A, Wang D, Walker W, Fowler J F, Chappell R J and TÃme W A 2006 Dose-per-fraction escalation for localized prostate cancer—a multi-institutional phase I/II trial *Int. J. Radiat. Oncol. Biol. Phys.* **66** s11

Samant S S, Xia J, Muyan-Ozcelik P and Owens J D 2008 High performance computing for deformable image registration: towards a new paradigm in adaptive radiotherapy *Med. Phys.* **35** 3546–53

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

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

Shchory B T, Schifter D, Lichtman R, Neustadter D and Corn B 2008 Static and dynamic tracking accuracy of a novel radioactive tracking technology for target localization and real time tracking in radiation therapy *Med. Phys.* **36** 2719

Shepard D M, Olivera G H, Reckwerdt P and Mackie T R 2000 Iterative approaches to dose optimization in tomotherapy *Phys. Med. Biol.* **45** 69–90

Shirato H *et al* 2000a Four-dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor *Int. J. Radiat. Oncol. Biol. Phys.* **48** 435–42

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

Thorwarth D, Soukup M and Alber M 2008 Dose painting with IMPT, helical tomotherapy and IMXT: a dosimetric comparison *Radiother. Oncol.* **86** 30–4

Webb S 2006a Motion effects in (intensity modulated) radiation therapy: a review *Phys. Med. Biol.* **51** R403–25

Webb S 2006b Quantification of the fluence error in the motion-compensated dynamic MLC (DMLC) technique for delivering intensity-modulated radiotherapy (IMRT) *Phys. Med. Biol.* **51** L17–21

Webb S and Binnie D M 2006 A strategy to minimize errors from differential intrafraction organ motion using a single configuration for a 'breathing' multileaf collimator *Phys. Med. Biol.* **51** 4517–31

Willoughby T R *et al* 2006 Target localization and real-time tracking using the Calypso 4D localization system in patients with localized prostate cancer *Int. J. Radiat. Oncol. Biol. Phys.* **65** 528–34

Wu Q J, Thongphiew D, Wang Z, Mathayomchan B, Chankong V, Yoo S, Lee W R and Yin F F 2008 On-line re-optimization of prostate IMRT plans for adaptive radiation therapy *Phys. Med. Biol.* **53** 673–91

Yan D, Lockman D, Martinez A, Wong J, Brabbins D, Vicini F, Liang J and Kestin L 2005 Computed tomography guided management of interfractional patient variation *Semin. Radiat. Oncol.* **15** 168–79

Yan D, Vicini F, Wong J and Martinez A 1997a Adaptive radiation therapy *Phys. Med. Biol.* **42** 123–32

Yan D, Wong J, Vicini F, Michalski J, Pan C, Frazier A, Horwitz E and Martinez A 1997b Adaptive modification of treatment planning to minimize the deleterious effects of treatment setup errors *Int. J. Radiat. Oncol. Biol. Phys.* **38** 197–206

Yeboah C and Sandison G A 2002 Optimized treatment planning for prostate cancer comparing IMPT, VHEET and 15 MV IMXT *Phys. Med. Biol.* **47** 2247–61

Zhang T, Keller H, O'Brien M J, Mackie T R and Paliwal B 2003 Application of the spirometer in respiratory gated radiotherapy *Med. Phys.* **30** 3165–71