Title

Fast leaf-fitting with generalized underdose/overdose constraints for real-time MLC tracking

Journal

Medical Physics, 43(1)

ISSN

0094-2405

Authors

Moore, Douglas
Ruan, Dan
Sawant, Amit

Publication Date

2015-12-31

DOI

10.1118/1.4938586

Peer reviewed

Fast leaf-fitting with generalized underdose/overdose constraints for real-time MLC tracking

Douglas Moore[a)]
Department of Radiation Oncology, UT Southwestern Medical Center, Dallas, Texas 75390

Dan Ruan
Department of Radiation Oncology, University of California, Los Angeles, California 90095

Amit Sawant
Department of Radiation Oncology, UT Southwestern Medical Center, Dallas, Texas 75390

(Received 12 August 2015; revised 7 December 2015; accepted for publication 10 December 2015; published 31 December 2015)

Purpose: Real-time multileaf collimator (MLC) tracking is a promising approach to the management of intrafractional tumor motion during thoracic and abdominal radiotherapy. MLC tracking is typically performed in two steps: transforming a planned MLC aperture in response to patient motion and refitting the leaves to the newly generated aperture. One of the challenges of this approach is the inability to faithfully reproduce the desired motion-adapted aperture. This work presents an optimization-based framework with which to solve this leaf-fitting problem in real-time.
Methods: This optimization framework is designed to facilitate the determination of leaf positions in real-time while accounting for the trade-off between coverage of the PTV and avoidance of organs at risk (OARs). Derived within this framework, an algorithm is presented that can account for general linear transformations of the planned MLC aperture, particularly 3D translations and in-plane rotations. This algorithm, together with algorithms presented in Sawant *et al.* ["Management of three-dimensional intrafraction motion through real-time DMLC tracking," Med. Phys. **35**, 2050–2061 (2008)] and Ruan and Keall [Presented at the 2011 IEEE Power Engineering and Automation Conference (PEAM) (2011) (unpublished)], was applied to apertures derived from eight lung intensity modulated radiotherapy plans subjected to six-degree-of-freedom motion traces acquired from lung cancer patients using the kilovoltage intrafraction monitoring system developed at the University of Sydney. A quality-of-fit metric was defined, and each algorithm was evaluated in terms of quality-of-fit and computation time.
Results: This algorithm is shown to perform leaf-fittings of apertures, each with 80 leaf pairs, in 0.226 ms on average as compared to 0.082 and 64.2 ms for the algorithms of Sawant *et al.*, Ruan, and Keall, respectively. The algorithm shows approximately 12% improvement in quality-of-fit over the Sawant *et al.* approach, while performing comparably to Ruan and Keall.
Conclusions: This work improves upon the quality of the Sawant *et al.* approach, but does so without sacrificing run-time performance. In addition, using this framework allows for complex leaf-fitting strategies that can be used to account for PTV/OAR trade-off during real-time MLC tracking.
 [http://dx.doi.org/10.1118/1.4938586]

Key words: real-time, DMLC, tracking, leaf-fitting algorithm

1. INTRODUCTION

The objective of external-beam radiotherapy is to deliver an ablative dose to the tumor target while attempting to minimize the dose delivered to healthy tissues and organs at risk (OARs). In thoracic and abdominal radiotherapy, motion of internal structures within the patient inevitably results in geometric misalignment between the planned and the delivered beams. One approach for handling such discrepancies is to follow or "track" the motion of the tumor target in real-time and relocalize the beam accordingly. This approach, termed as real-time tracking, has been implemented using various clinical and investigational strategies. Clinically deployed strategies include the Cyberknife system (Accuray, Sunnyvale, CA) which is comprised of a light-weight 6 MV Linac mounted on a robotic arm, and the Vero system (Mitsubishi, Japan, marketed in the US by Brainlab) which utilizes a gimbal-mounted Linac.[1–4] Investigational systems are based on moving the treatment couch[5] or a dynamic multileaf collimator (MLC) to perform continuous aperture adaptation in real-time.[6–11] In the MLC tracking approach, an independent patient monitoring system provides real-time position information to the Linac control system at a sufficiently high temporal frequency (typically 10–40 Hz). Based on this information, the control system transforms the treatment aperture in order to compensate for motion-induced deviations from the planned state. The MLC leaves are then refitted to the newly obtained aperture to approximately reproduce the desired fluence described by the initial planned aperture and finally passed on to the machine for delivery.

The process of refitting is simple when only rigid translations in the direction of MLC leaf travel are assumed.[7,12] However, the complexity of the refitting process increases progressively, particularly for highly modulated intensity modulated radiotherapy (IMRT) fields, when accounting for motion perpendicular to leaf travel, higher-order motion (e.g., rotations), the relative motion of the target with respect to OARs (e.g., spinal cord and heart), or relative motion of multiple tumor targets. In such situations, faithfully reproducing the transformed aperture is nontrivial. For example, Fig. 1(a) depicts two square apertures. These are transformed from their original shapes due to motion-induced translations and in-plane rotations, as illustrated in Fig. 1(b). In this case, refitting the MLC leaves to the transformed apertures will inevitably result in regions of underdose to the tumor target and/or overdose to normal tissue and organs at risk. While such underdose/overdose can rarely be completely eliminated, it is nevertheless useful to explore real-time leaf-fitting techniques that attempt to minimize these errors in the presence of complex motion.

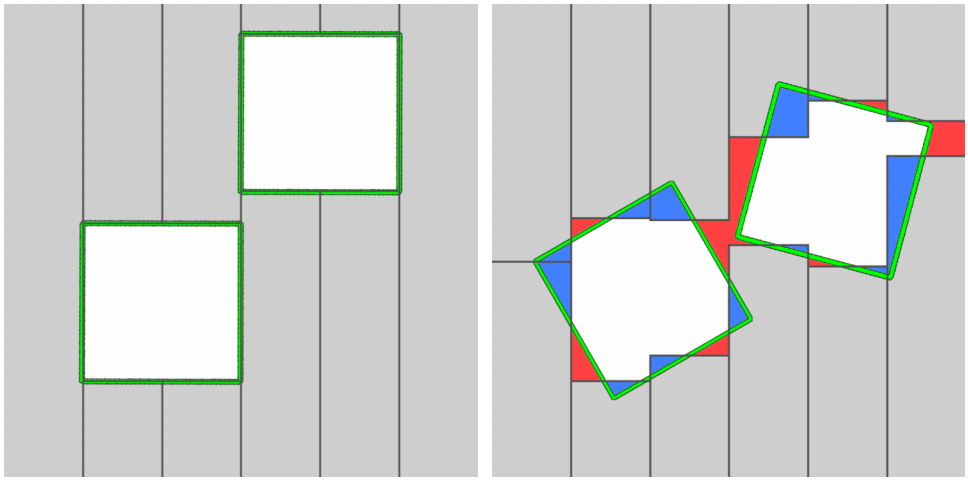
Leaf-fitting algorithms can easily grow computationally intensive and therefore contribute nontrivially to the overall system latency, the time between observed motion and the hardware response. System latencies have been reported to be between 100 and 220 ms for tracking systems based on commercially available hardware.[10,11,13] In order to account for this latency, one must introduce position estimation algorithms. These algorithms are inherently limited in their capabilities as they have but a finite amount of noisy information on which to base their predictions. As the look-ahead time for prediction decreases, the computational and predictive performance of such algorithms tends to improve.[14] By implementing fast leaf-fitting algorithms, without sacrificing fitting accuracy, the system latency may be reduced, thus improving prediction and overall system performance.

This work presents a method for rapidly and optimally fitting MLC leaf configurations to transformed apertures in the beam's-eye-view (BEV). The introduction of a generalized cost-density function over the treatment field is proposed. The leaf configuration can then be determined from the density function via an optimization principle. This approach is based on the work of Ruan,[15] henceforth referred to as Ruan 2011, but generalizes to allow for arbitrarily complex fitting criteria thus admitting Ruan 2011 as a special case. The problem is restricted to the special case of homogeneous cost-density over each 2D region of interest (ROI) in the isocentric plane as seen from the BEV with polygonal apertures assumed. The associated algorithm, referred to as piecewise, together with two comparison algorithms Ruan 2011 and midleaf,[7] is presented. Each algorithm was evaluated in terms of computation time and quality-of-fit of the resulting "treatment aperture" during tracking simulations based on eight patient-derived lung traces, each exhibiting six degree-of-freedom (6DoF) motion, and eight lung IMRT plans.

2. METHODS

This work focuses on the construction of leaf configurations that attempt to reproduce the fluence of a planned aperture f in the presence of some anatomical motion T from the BEV perspective. In principle, one would obtain the motion-compensated aperture as the composition of the planned aperture with the motion $f \circ T$. However, this is often unobtainable due to hardware constraints such as finite leaf widths and leaf-motion trajectories. The objective is thus to obtain an optimal leaf configuration subject to all hardware constraints.

Assume that there are N leaf-pairs with widths $\{\Delta_i\}_{i=1}^N$, and that the coordinate system is oriented so that the y-direction is aligned with the direction of leaf motion and the x-direction runs perpendicularly from the BEV perspective, as illustrated in Fig. 2. Let $\{x_i\}_{i=1}^N$ denote the x-coordinate of the left edges of the leaf-pairs so that $x_{i+1} = x_i + \Delta_i$ and the full aperture is covered by the N leaf pairs. The aim is to find the leaf



(a) (b)

FIG. 1. Rotational motion makes it impossible, except in exceptional circumstances, to faithfully fit MLC leaves to the transformed aperture. (a) Two planned shapes (green) that can be perfectly reproduced by the MLC leaves. (b) Same two shapes under relative motion (in-plane rotations and translations). Note the regions of overdose (red, outside of the shape) and underdose (blue, inside the shape) in (b).

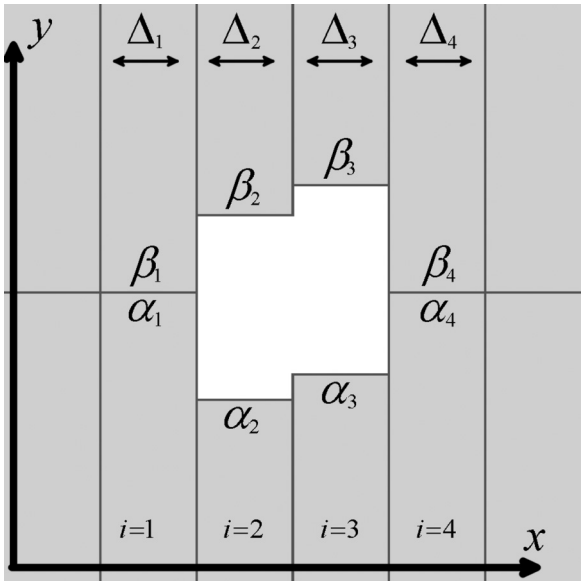


FIG. 2. The x- and y-axes are chosen to run parallel with and perpendicular to the leaf motion, respectively. The axes and examples of leaf-indices, leaf-widths, and leaf positions are depicted.

positions $A=\{\alpha_i,\beta_i\}_{i=1}^{N}$ that deliver a fluence closest to the ideal aperture $f\circ T$, where α_i and β_i are positions of the top and bottom leaf of the ith pair, respectively. It is assumed that $0\le\alpha_i\le\beta_i\le h$, where h is the maximum leaf position as specified by the hardware.

2.A. Generic formulation

We express the problem as an optimization. We define at each point (x,y) in the BEV plane two positive values, $\sigma_u(x,y)$ and $\sigma_o(x,y)$, respectively, representing the cost of underdosing or overdosing that point. Refer to σ_u (σ_o) as the underdose (overdose) cost-density function. The objective function over the leaf pairs is then defined as

$$\Phi(A)=\sum_{i=1}^{N}\Phi_i(\alpha_i,\beta_i)=\sum_{i=1}^{N}(\Phi_{u,i}(\alpha_i,\beta_i)+\Phi_{o,i}(\alpha_i,\beta_i)), \tag{1}$$

where Φ is the total cost, Φ_i is the leaf cost, and $\Phi_{u,i}$ ($\Phi_{o,i}$) is the underdose (overdose) cost of fitting the ith leaf pair. The underdose and overdose costs are defined as follows:

$$\Phi_{u,i}(\alpha,\beta)=\left(\int_0^{\alpha}dy+\int_{\beta}^{h}dy\right)\int_{x_i}^{x_{i+1}}dx\sigma_u(x,y), \tag{2}$$

$$\Phi_{o,i}(\alpha,\beta)=\int_{\alpha}^{\beta}dy\int_{x_i}^{x_{i+1}}dx\sigma_o(x,y). \tag{3}$$

One can intuitively understand $\Phi_{u,i}$ ($\Phi_{o,i}$) as the area obstructed (unobstructed) by the ith leaf, weighted by the local underdose (overdose) cost-density. Due to the independent nature of the leaf pairs, it is reasonable to assume that there are no constraints between any two given pairs. For that reason, optimization of Φ amounts to optimizing each Φ_i independently.

It is desirable to find a simpler form for Φ_i. First, the expression [Eq. (2)] can be written more suggestively as

$$\Phi_{u,i}(\alpha,\beta)=\left(\int_0^{h}dy-\int_{\alpha}^{\beta}dy\right)\int_{x_i}^{x_{i+1}}dx\sigma_u(x,y). \tag{4}$$

Note that the total cost for the ith leaf can be expressed as

$$\Phi_i(\alpha,\beta)=\int_{\alpha}^{\beta}dy\int_{x_i}^{x_{i+1}}dx[\sigma_o(x,y)-\sigma_u(x,y)]+\int_0^{h}dy\int_{x_i}^{x_{i+1}}dx\sigma_u(x,y). \tag{5}$$

Since the second term is constant in α and β and does not affect the optimization, it is possible to drop it all together and take the total cost to be

$$\Phi_i(\alpha,\beta)=\int_{\alpha}^{\beta}dy\int_{x_i}^{x_{i+1}}dx[\sigma_o(x,y)-\sigma_u(x,y)]. \tag{6}$$

Define the net cost-density function $\sigma(x,y)\equiv\sigma_o(x,y)-\sigma_u(x,y)$ so that the leafwise objective function takes the final form,

$$\Phi_i(\alpha,\beta)=\int_{\alpha}^{\beta}dy\int_{x_i}^{x_{i+1}}dx\sigma(x,y)=\int_{\alpha}^{\beta}dy\,\lambda_i(y). \tag{7}$$

It is evident that at points where the underdose cost is greater than the overdose cost, σ takes a negative value and is either positive or zero otherwise. Knowing this, one can see that $\Phi_i(\alpha,\beta)$ represents the cost of dosing the region $y\in[\alpha,\beta]$ such that if that underdose cost outweighs the overdose cost, the net cost is negative. In this way, it is easy to represent the trade-off between dosing the target and dosing an OAR—regions with high underdose and overdose penalties, respectively.

The optimal upper and lower positions for the ith leaf-pair arise as solutions to

$$\lambda_i(y)=0,\quad \frac{d\lambda_i}{dy}>0, \tag{8}$$

where the former condition ensures that the solution is an extremum while the latter distinguishes the solution as a minimum.

Without knowing the precise form of σ, and hence λ_i, it is not possible to show whether or not a minimum will exist at all much less that there is a global minimum. However, given that the MLC leaf positions are limited in resolution, there are finitely many leaf configurations, and thus, at least one minimum exists in this discrete limit. In the event that multiple minima exist, then any selection criteria may be used to choose among them.

2.B. Quasi-homogeneous cost-density

There is no single objective mechanism for selecting the cost-density function σ. The decision will depend on the physician, based on patient characteristics, disease state, and the tissue types involved. For illustration purposes, this work aims to provide one method of defining σ over the treatment field wherein the motion of each ROI is monitored and has a prespecified homogeneous cost-density over its area. Note

that, in the present context, the term ROI is used to mean 2D regions in the isocentric plane from the BEV.

For the remainder of this section, suppose that there are M ROIs, Ω_i for $i = 1 \ldots M$, each with homogeneous cost-density σ_i over its volume. Targets are considered to be ROIs that receive fluence and are defined with $\sigma_i < 0$ while OARs are considered as ROIs that should be avoided and carry $\sigma_i > 0$. It is also assumed that the background tissue (e.g., healthy lung) carries a constant cost-density of $\sigma_0 \geq 0$. To define the total cost-density function, it is convenient to define ROI-supported cost-density functions: $\sigma_i(x,y) = \sigma_i g_i(x,y)$, where

$$g_i(x,y) = \begin{cases} 1, & (x,y) \in \Omega_i \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

with i indexing the ROI. Piecing the ROI-supported cost-densities together gives the full cost-density function

$$\sigma(x,y) = \sigma_0 + \sum_{i=1}^{M} \sigma_i g_i(x,y). \tag{10}$$

In the event that only one ROI exists, namely, the target, this density function gives rise to the same objective function as Ruan[15] with homogeneous dose coefficients related as $\sigma_0 = \lambda_o$, $\sigma_1 = -(\lambda_u + \lambda_o)$.

Performing the optimization on Eq. (7) with Eq. (10) gives

$$\sigma_0 \Delta_i = -\sum_{j=1}^{M} \sigma_j \lambda_{ij}, \tag{11}$$

where, for future brevity, λ_{ij} is defined as

$$\lambda_{ij} = \int_{x_i}^{x_{i+1}} dx g_j(x,y). \tag{12}$$

2.C. Algorithmic formulation

The results of Secs. 2.A and 2.B yield a path toward implementing an efficient leaf-fitting strategy which, in principle, can adapt to highly complex dosimetric requirements. However, the greatest limitation to performance comes from the computation of the integral [Eq. (12)]. The following three observations provide us with the necessary simplifications to make this approach viable:

- The shapes generated during planning, due to the use of MLC configurations, are polygons.
- At present, the only types of motion that can be reliably monitored with clinically available motion monitoring systems are 3D translations and 3D rotations.
- Rotations out of the isocentric plane can be considered as deformations from the BEV perspective. Tracking such motion is outside the scope of our present discussion because current treatment planning and motion monitoring systems do not explicitly provide adequate information about the dosimetric impact of such complex motion in real-time.

The first point allows us to consider shapes simply in terms of the boundaries: sequences of line segments. The second and third points together require that the problem be restricted, for now, to 3D translations and in-plane rotations, and thus ensure that the shapes will remain polygons after accounting for motion.

Since the shapes are polygons with finitely many vertices, the y-axis can be broken into regions within which no vertex appears. We will refer to this process as *segmenting* and the regions as *segments*. Each segment then has two bounding y-values and the expression [Eq. (12)] is linear in y for each $0 \leq i < N$ and $1 \leq j < M$. An example of this is demonstrated in Fig. 3.

The piecewise algorithm for computing the optimal leaf configuration is then straightforward (Fig. 4). Solve Eq. (11) for each segment (and each leaf) taking the solutions as trial leaf positions. If no solution is found for a given segment or the solution lies outside of the segment, save the boundaries of the segment as trial solutions. Once this process has been carried

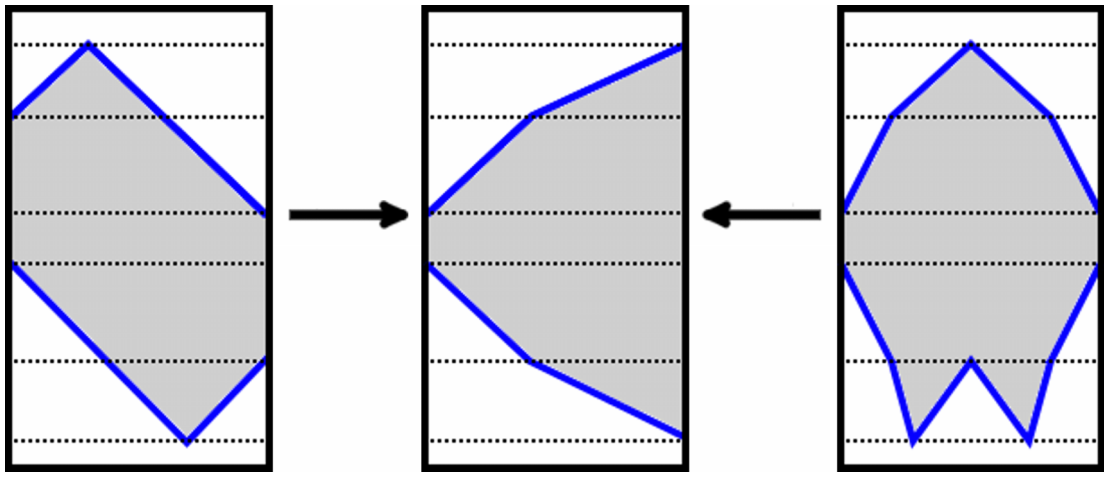


FIG. 3. Process of segmenting and reshaping the target aperture within a single leaf path. The dotted lines represent the segment boundaries. Note that this reshaping process ensures that the integral along the horizontal axis does not change and makes value of the integral evident from the linear form of the boundary. All three shapes (a)–(c) produce the same optimal leaf positions; however, (b) is more conducive to fast computation.

```
for i = 1 ... N do
    Segment the shapes
    Σ ← {0, h}
    for each segment do
        x_i(y) ← Σ_j σ_j λ_ij (y)
        y_trial ← Solve(σ_0 Δ_i + x_i(y) = 0)
        if y_trial = { } ∨ y_trial ∉ segment then
            (a, b) ← Bounds(segment)
            Σ ← Σ ∪ {a, b}
        else
            Σ ← Σ ∪ y_trial
        end if
    end for

    Sort(Σ)
    Unique(Σ)

    (α̂_i, β̂_i, Φ̂_i) ← (0,0, Φ_i(0,0))
    for α ∈ Σ do
        for β ∈ {σ|σ ∈ Σ, σ > α} do
            φ ← Φ_i(α, β)
            if φ < Φ̂_i then
                (α̂_i, β̂_i, Φ̂_i) ← (α, β, φ)
            end if
        end for
    end for

    (α̂_i, β̂_i, Φ̂_i)
end if
```

FIG. 4. The piecewise algorithm for optimally fitting MLC leaves to fields with multiple ROIs each weighted homogeneously. For each leaf pair, the shapes are segmented as in Fig. 3 and then the expression for the boundary is used to compute trial solutions.

out for each segment, it loops over all admissible combinations of top- and bottom-leaf positions and computes the cost for each. The final solution is the pair of positions corresponding to the lowest cost.

The midleaf approach determines leaf positions as positions at which the shape-edge and the leaf-edge intersect at the midpoint of the leaf, hence "Mid-Leaf." In the event that multiple solutions are found, the lower and upper leaf positions are chosen as the smallest and largest solutions, respectively. See Fig. 5(a) for details. It is important to note that this algorithm is only clearly defined if there is one ROI. This approach is computationally similar to our proposed algorithm, Fig. 4, in its selection of trial solutions, but the two algorithms differ in the selection of the final solution.

The Ruan 2011 algorithm is closer in spirit to the proposed approach, yet the two differ in both the method of computation and the assumptions made regarding treatment preferences. In this approach, underdose and overdose cost-densities, $\sigma_u(x,y)$ and $\sigma_o(x,y)$, respectively, are selected. The cost of fitting the ith leaf is then given by

$$\Psi_i(\alpha_i,\beta_i) = \int_{\alpha_i}^{\beta_i} dy\left[c_{o,i}(y) - c_{u,i}(y)\right], \tag{13a}$$

$$c_{u,i} = \int_{x_i}^{x_{i+1}} dx\sigma_u(x,y)h(x,y), \tag{13b}$$

$$c_{o,i} = \int_{x_i}^{x_{i+1}} dx\sigma_o(x,y)[1-h(x,y)], \tag{13c}$$

with

$$h(x,y) = \begin{cases} 1, & (x,y) \in \text{transformed plan aperture opening} \\ 0, & \text{otherwise} \end{cases}. \tag{14}$$

Ruan *et al.* showed that finding the extrema of Eq. (13a) amounted to computing the sets,

$$C_i^{\alpha} = \{y | I(c_i(y+\delta))I(-c_i(y-\delta)) = 1\}, \tag{15a}$$

$$C_i^{\beta} = \{y | I(c_i(y-\delta))I(-c_i(y+\delta)) = 1\}, \tag{15b}$$

with δ small, $c_i = c_{u,i} - c_{o,i}$, and $I(x) = \begin{cases} 1, & x > 0 \\ 0, & \text{otherwise} \end{cases}$. Essentially, the Ruan algorithm explicitly computes the derivatives of Eq. (13a) enforcing convexity at each solution. The complete algorithm is illustrated in Fig. 5(b). There are three points of note regarding the Ruan approach. First, the notation used is highly compact and represents a large amount of computation—at each value of y, a numeric derivative $c_i(y)$ is computed, not to mention the actual value of $c_i(y)$. Second, the objective function [Eq. (13a)] is similar to the objective function used in our proposed algorithm [Eq. (7)]. In fact, when there is only one ROI, and the cost density used in the proposed approach is chosen as in Eq. (11), the two algorithms *functionally* coincide. However, computationally, the approaches differ substantially. In particular, the time complexity of our proposed approach runs linearly with the number of shape segments while Ruan 2011 varies with the total number of potential leaf positions. To give some concept of scale, the number of shape edges will typically be ~10, while the total number of potential leaf positions in the case of the Elekta Agility 160 MLC is on the order of 40 000. Third, because of the inclusion of $h(x,y)$ in Eqs. (13a) and (13b), Ruan 2011 will always attempt to faithfully reproduce the aperture regardless of what critical structures may come into the irradiation field.

2.D. Performance evaluation

Each of the three algorithms was directly implemented in c++ and integrated into an in-house MLC tracking platform. Eight step-and-shoot lung IMRT plans using the Agility 160 MLC (Elekta, Crawley, United Kingdom) were taken from the PINNACLE (Ref. 3) treatment planning system (Philips, the Netherlands) and used to generate a set of 5832 target apertures. To generate the apertures, each plan was subjected to eight 6DoF motion traces during a tracking simulation. These traces were extracted offline via kilovoltage intrafraction monitoring (KIM) system for three lung cancer patients at Northern Sydney Cancer Centre, Royal North Shore Hospital, Sydney, Australia.[16] A set of three combinations of underdose/overdose costs were chosen for evaluation, $\{(\sigma_u,\sigma_o)\}$ = $\{(0.75,0.25),(0.25,0.75),(0.5,0.5)\}$, representing preference toward treating the PTV, sparing healthy tissue and equal balance, respectively. Using these sets of apertures and weights, each algorithm was employed to generate a treatment

```
for i = 1 ... N do
    Segment the shape
    Σ ← { }
    for each segment do
        x_i(y) ← Σ_j σ_j λ_ij(y)
        y_trial ← Solve(λ_ij(y) = ½(x_i + x_{i+1}))
        if y_trial ⊆ segment then
            Σ ← Σ ∪ y_trial
        end if
    end for

    (α̂_i, β̂_i) ← (0,0)
    if Σ ≠ { } then
        Sort(Σ)
        (α̂_i, β̂_i) ← (First(Σ), Last(Σ))
    end if

    (α̂_i, β̂_i)
end if
```

(a)

```
for i = 1 ... N do
    Compute c_{u,i}, c_{o,i}, and c_i = c_{u,i} − c_{o,i}
    C_i^α = {y : I(c_i(y + δ))I(−c_i(y − δ)) = 1}
    C_i^β = {y : I(c_i(y − δ))I(−c_i(y + δ)) = 1}

    (α̂_i, β̂_i, Ψ̂_i) ← (0,0, Ψ_i(0,0))
    for α_i ∈ C_i^α do
        for β_i ∈ {β | β ∈ C_i^β, β > α_i} do
            ψ ← Ψ_i(α_i, β_i)
            if ψ < Ψ̂_i then
                (α̂_i, β̂_i, Ψ̂_i) ← (α_i, β_i, ψ)
            end if
        end for
    end for

    (α̂_i, β̂_i, Ψ̂_i)
end if
```

(b)

FIG. 5. The two leaf fitting algorithms used for comparison against the algorithm presented in Fig. 4. (a) The midleaf algorithm which is only defined for a single ROI and returns only the selected leaf positions. (b) The Ruan 2011 algorithm which computes optimal leaf positions by explicitly calculating derivatives of the cost function Ψ_i [Eq. (13a)] and returns the leaf positions and computed cost value.

aperture, 17 496 apertures in total. The performance of each algorithm was quantified in terms of two metrics: computation time—measured with a software timer—and quality-of-fit.

The quality-of-fit of a treatment aperture was computed by determining the areas of underdose and overdose—compared to the ideal aperture $(f \circ T)$—and combining them as a weighted sum using the underdose/overdose costs as weights as in Eq. (1). This was carried out independently of the fitting algorithm in question to ensure that the results were directly comparable. In comparing qualities-of-fit, it is important to note that different implementation details can result in small differences in leaf positions—on the order of 10 μm—that produce small variations in the final quality of fit. For this reason, two quality-of-fit scores were considered equivalent if they were within 1.0%.

3. RESULTS

The computation time statistics are summarized in Table I which presents the minimum, average, and maximum time observed per aperture fitting. The piecewise algorithm performed comparably to the midleaf algorithm. Note that the performance of the piecewise and Ruan 2011 approaches was independent of the underdose/overdose cost used.

TABLE I. The time clock-time used by each algorithm to perform a full aperture fitting.

	Time per fitting (ms)		
Algorithm	Minimum	Average	Maximum
Midleaf	0.020	0.082	0.195
Piecewise	0.063	0.226	0.527
Ruan 2011	52.1	64.2	86.3

For each underdose/overdose cost value and pair of algorithms, the quality-of-fit scores for each of the 5832 treatment apertures were determined and compared (see Table II). When the underdose and overdose costs differ significantly, the piecewise and Ruan 2011 algorithms produced significantly improved fitting over the midleaf approach while performing comparably to one another in terms of this metric. In the case that the underdose and overdose costs were equal, all three algorithms performed similarly.

As a final comparison, the quality-of-fit averaged over the entire treatment simulation for each algorithm and underdose/overdose cost was computed. This average was compared to the average of the midleaf quality-of-fit, see Table III.

As expected from Fig. 5, piecewise and Ruan 2011 produce approximately the same average quality-of-fit regardless of underdose/overdose costs and perform significantly better than the midleaf approach when the underdose/overdose costs are not balanced.

TABLE II. The quality-of-fit was compared for each pair of algorithms. The percent of fittings (out of a total of 5832) that were better, equivalent (up to a 1% difference), and worse are presented. Note the similarity between the performance of the piecewise and Ruan 2011 approaches.

		Percent of fittings		
Underdose/overdose	Algorithm	Better	Within 1%	Worse
0.75/0.25	Piecewise vs midleaf	64.7	35.1	0.223
	Piecewise vs Ruan 2011	5.01	95	0.034
	Ruan 2011 vs midleaf	60.9	36.4	2.78
0.50/0.50	Piecewise vs midleaf	4.12	95.6	0.24
	Piecewise vs Ruan 2011	0.89	99	0.137
	Ruan 2011 vs midleaf	4.18	94.9	0.943
0.25/0.75	Piecewise vs midleaf	68.5	31.5	0
	Piecewise vs Ruan 2011	1.41	98.6	0
	Ruan 2011 vs midleaf	64.7	35.2	0.103

TABLE III. The average quality-of-fit of the piecewise and Ruan 2011 algorithms with each underdose/overdose cost. The average quality-of-fit is presented relative to the total cost generated by the midleaf algorithm.

Underdose/overdose cost	Algorithm	Average quality-of-fit relative to midleaf
0.75/0.25	Piecewise	0.878
	Ruan 2011	0.879
0.5/0.5	Piecewise	0.985
	Ruan 2011	0.985
0.25/0.75	Piecewise	0.834
	Ruan 2011	0.834

4. DISCUSSION

The introduction of a generalized cost-density of the treatment field allows for complex fitting criteria ranging from PTV/OAR trade-off to restricting the treatment to motion within a predefined spatial window. With a particular choice of density function, the piecewise approach is equivalent to that of Ruan 2011. However, by not building the pretreatment aperture into the framework in general, this approach can consider real-time MLC tracking in situations where the preplanned aperture does not capture the desired objective.

In addition to the general framework, a particular implementation in which the motion-adapted pretreatment aperture and the background tissue each have homogeneous cost was evaluated. The results described in Sec. 3 suggest that the piecewise and Ruan 2011 approaches are interchangeable in terms of quality-of-fit, with both substantially outperforming the midleaf approach. However, based on computation time, the midleaf and piecewise algorithms dominated. It is therefore reasonable to conclude that a strong balance of performance in terms of both metrics is achieved by the piecewise approach.

Given the observations in Sec. 2.C regarding the functional equivalence of the proposed algorithm piecewise and Ruan 2011, it is noteworthy to point out that in practice, they do not perform identically in terms of quality of fit, see Table II. The underlying cause of this difference is attributed to the number of floating-point computations performed in each, specifically there are far fewer in piecewise than in Ruan 2011. This has two manifestations. First, equality comparison becomes tenuous at best, hence the concession that "equality" be defined as a 1% difference or less. Second, when there are two open fields in a single leaf path—which can happen often in IMRT apertures under rotations—it is possible that one algorithm chooses to block one open field while the other algorithm blocks the second open field. This can result in a significant difference (>1%) between the final qualities of fit when accumulated over all of the leaves.

4.A. The use of cost-densities and prospective improvements

It is interesting to consider how the asymmetric underdose/overdose weights impact the treatment. To begin with, assume that there is only a single ROI and the healthy background. Then, the limits of the ratio $\kappa = \sigma_u/\sigma_o$ describe how piecewise and Ruan 2011 will behave as one favors the two extremes: $\kappa \to 0$ yields maximum tissue sparing while $\kappa \to \infty$ produces maximum target coverage. As more ROIs are introduced, the weighting schemes grow more complex, and the distinction between the piecewise and Ruan 2011 becomes apparent.

One possibility is to have a ROI with a high overdose cost compared to the underdose cost of the target. This is represented in the piecewise algorithm as $\sigma_2 + \sigma_1 > 0$ with $\sigma_1 < 0 < \sigma_2$ such that σ_2 is an OAR while σ_1 is a target. If this is the case, in regions where the OAR is in beam line with the target, the OAR will be blocked by a closed leaf-configuration. This essentially states that avoiding the OAR is more important than treating the target in that instance. When the target and OAR move out of alignment, the leaf configuration will open to more faithfully reproduce the initial planned aperture. Such a scheme is not possible in the Ruan 2011 formalism due to the introduction of the $h(x,y)$ step function described in Sec. 2.C. An example of this is illustrated in Fig. 6 wherein an underdose distribution Fig. 6(a) and an overdose distribution Fig. 6(b) are combined to produce the resulting cost density in Ruan 2011 and piecewise frameworks, Figs. 6(c) and 6(d), respectively. The target step function $h(x,y)$ results in treatment of the tumor regardless of transient structures whereas this may not occur in our approach. Figure 7 exhibits a simulation of tracking a target as it transits a high-overdose region using the Ruan 2011 definition of cost density [Figs. 7(a)–7(c)] and our generalized approach [Figs. 7(d)–7(f)] demonstrating that one could more aggressively spare OARs in principle at the expense of underdosing regions of the target. This underdose can be accounted for in principle by employing time-dependent cost-density functions and would likely result in slightly longer treatment times. This discussed briefly at the end of Sec. 4.A. Note that the proposed approach can reproduce that of Ruan 2011, but not vice versa. This follows from the fact that cost described in Eq. (13a) is a special case of proposed cost Eq. (7). This suggests that our generalized framework described in Sec. 2.A admits a richer solution space than the restricted set proposed in Ruan 2011.

Another point of note is that while piecewise and Ruan 2011 both resemble the approach taken by Tacke *et al.* in explicitly weighting relative underdose and overdose,[8] unlike Tacke, our generic approach and Ruan 2011 both readily admit inhomogeneous underdose/overdose weighting, and handle nontranslational adaptation by construction. The value in this weighting scheme is that target depth, ROI thickness, etc., could potentially be incorporated. For example, the cost of treating an ROI may increase as the (line-of-sight) thickness from the BEV perspective increases—essentially one can include dose-related heuristics in the selection of the cost-densities.

On a final note, one may also introduce time-dependent (or MU-dependent) cost-density functions. This would allow at least two further accommodations. First, it would admit the possibility of adjusting the weighting scheme based on past performance. If a region was underdosed at a previous time-point, the algorithm may be able to account for this to overdose at a later time-point. The second possibility is that

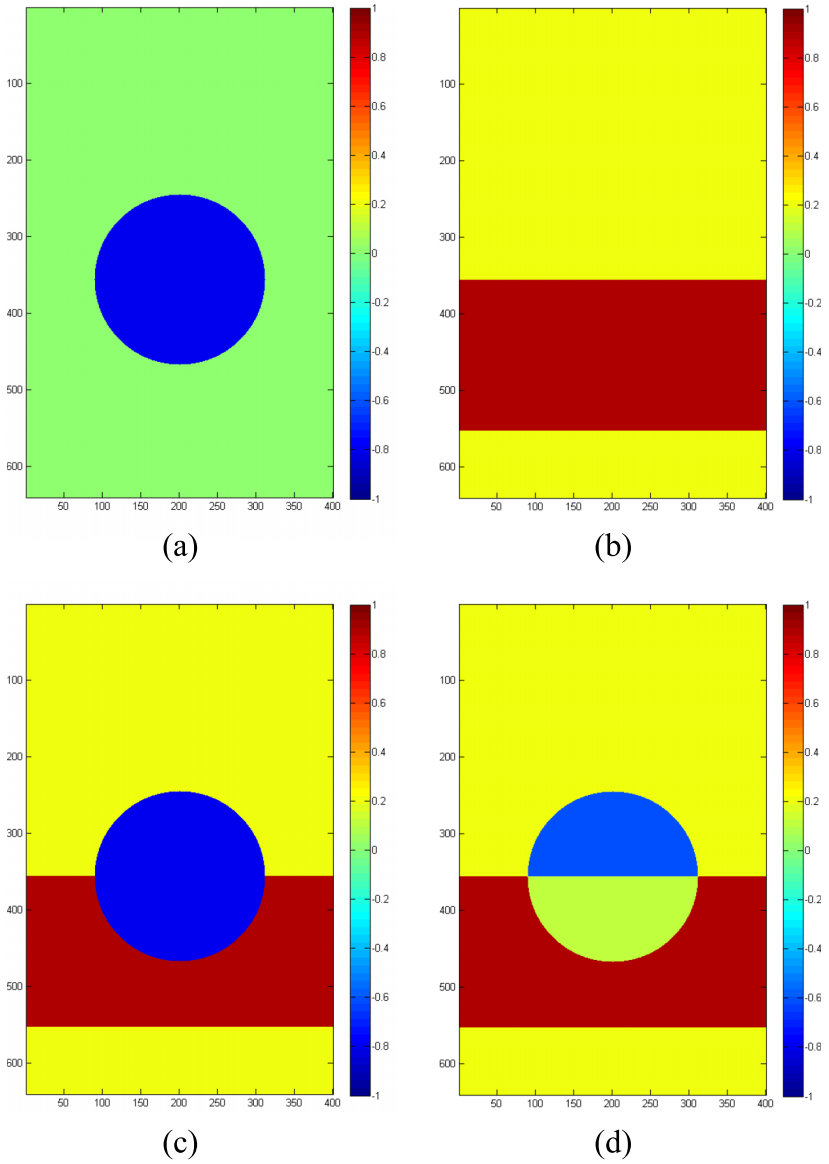


FIG. 6. An example of combining underdose (a) and overdose (b) cost density distributions in Ruan 2011 (c) and the proposed generalized framework (d). In the case of the Ruan 2011 solution, regions of the field corresponding to target area always receive a final cost density indicative of a target without regard to transient OAR. This is indicated by homogeneous cost density within the circular target area in (c). In the proposed framework, this is not necessarily the case. The cost densities in a region are additive such that the final cost density is representative of the cumulative intent of the plan. This is depicted as the half-circular region of positive cost density (d) representing a region in which the overdose cost outweighs the underdose cost associated with the circular target.

using this time dependence, one could incorporate motion trajectories to *prospectively* adjust the weighting scheme. For example, if the target is expected to pass in front of an OAR at a future time-point, then it could be beneficial to adjust the current cost-density to account for that future underdose.

4.B. The effects of assumptions on performance

The restriction to polygonal shapes can be relaxed without much computational cost as long as, at least on the scale of a leaf-width, the boundary can be expressed in a closed form. An example of this might be the projection of ROI contours onto the BEV plan. In this situation, the algorithm presented herein does not change. The only change in performance will arise due to the differences in the computational complexity of solving Eq. (11) without the assumption of linear boundaries. It is unlikely that would be a significant loss. Recall, however, that MLC leaves have finite resolution in both coordinate directions. Because of this, one can always approximate a BEV projected shape as a polygon at least at the resolution of the MLC. Alternatively, more robust methods such as splines may be employed.

With regard to the number of ROIs, it was noted previously that the complexity of the piecewise algorithm goes as $O(N)$, with N the number of shape segments in the leaf-path. In a realistic application, one might expect at most 5 ROIs with each ROI contributing at most ten segments to a given leaf-path for a total of $N \simeq 50$. This would result in full-aperture computations on the order of milliseconds.

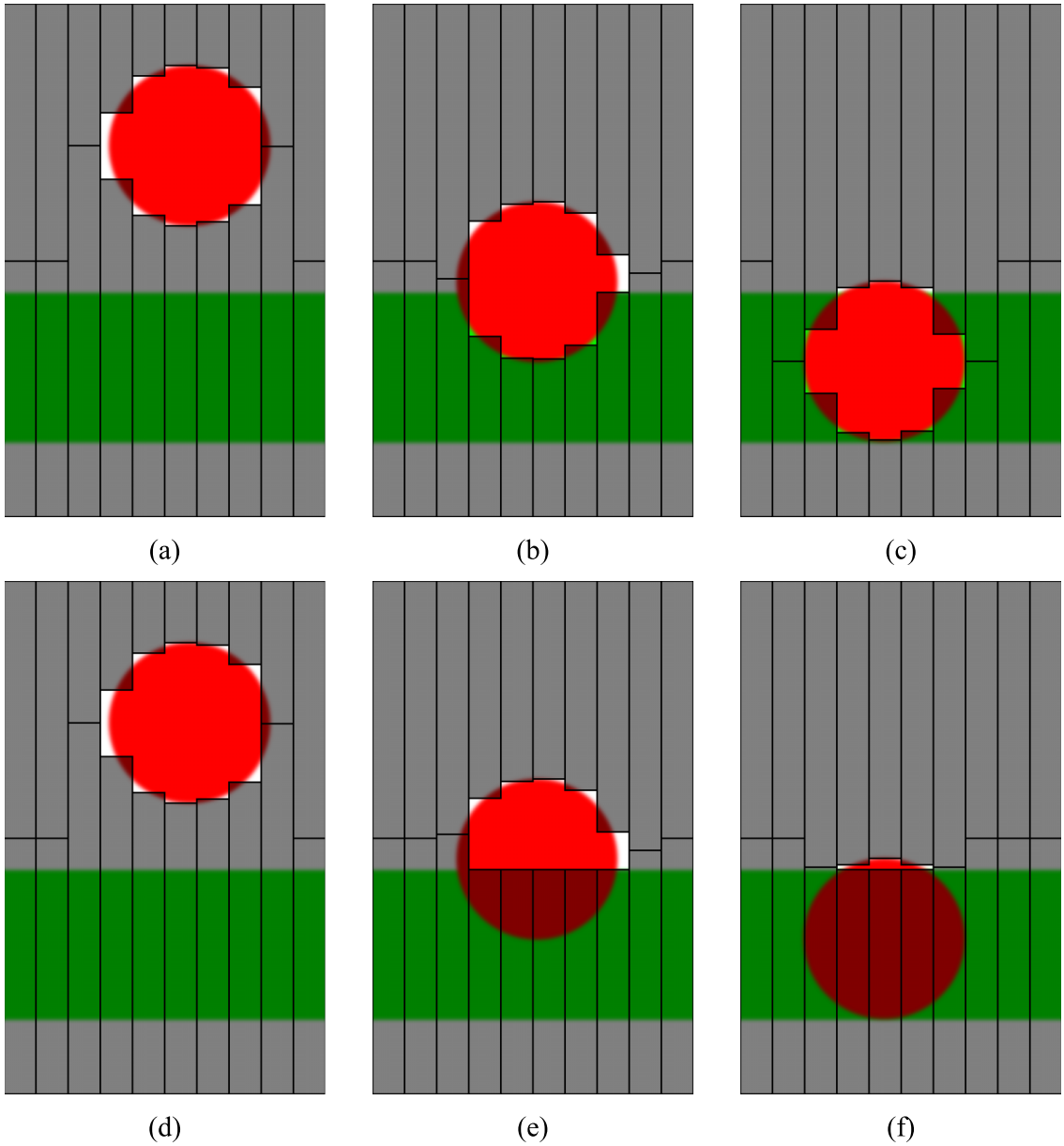


FIG. 7. Schematic illustration of the target (red) and an OAR (green) tracked by a MLC aperture. In the case of the Ruan 2011 framework [(a)–(c)], both partial (b) and complete (c) overlap of the target and OAR results in dose to the OAR (overdose) because the target is always treated. In contrast, the generalized framework described herein [(d)–(f)] can spare the OAR by closing the leaves in regions of target/OAR overlap [(e) and (f)]. Note that there always exists a choice of underdose/overdose costs for which the proposed generalized approach reduces to the Ruan 2011 solution.

A final concern is whether the number of leaves involved in the aperture fitting impacts performance. In this work, the computations were performed for all leaves regardless of whether they could reasonably contribute to the final open field. One should expect that by skipping leaves that are clearly not involved in the aperture formation, Ruan 2011 would show a greater performance improvement than either midleaf or piecewise. This follows from the fact that in midleaf and piecewise, the performance of computing the positions of an uninvolved leaf is comparable to testing for leaf-involvement, while for Ruan 2011, computing leaf positions of such leaves is much more complex. There are many improvements that can be made to the Ruan 2011 algorithm to decrease the work required in computing the set C^{α} and C^{β}. However, this work was focused on both the generalized framework used to construct piecewise as well as the relative performance between naïve implementations of the various algorithms.

4.C. Prospects and future development

The prospects for this approach are varied. It has already been noted that improved runtime performance may result in a reduction in prediction errors by decreasing the end-to-end system latency.[14] However, if the current level of prediction error is acceptable, one option is to exchange this improvement in prediction for more complex real-time processing, e.g., on-the-fly dose calculations. While the additional processing may increase the system latency, the debt is paid for by the

improved leaf-fitting performance. Such improvements open the window for more advanced tracking techniques that would have otherwise been infeasible due to hardware and predictive constraints.

An avenue of improvement that both piecewise and Ruan 2011 readily support, as mentioned by Ruan,[15] is the addition of leaf-velocity constraints. The idea is to maintain two sets of trial solutions, Σ^{α} and Σ^{β}, one for the lower leaf and one for the upper leaf. Trial solutions are appended to these sets in accordance with what is achievable based on the current leaf positions and the maximum admissible leaf velocities. The algorithm then proceeds in much the same way as in Fig. 4, distinguishing between lower and upper trial leaf positions. When considering leaf-velocities, one must decide how to handle situations in which the leaves cannot achieve fitting that is as good as what might be possible with infinitely fast leaves. In this approach, the mechanism for handling such a case could be to introduce a threshold value for the quality-of-fit (or fitting cost)—normalized in some way—such that if a fitting fails to meet that threshold value, the treatment beam is paused. The advantage of this approach is that if only a small number of leaves fail to reach their desired positions, they may not significantly impact the overall quality of fit and, therefore, the treatment need not be paused.

Additionally, it is conceivable that one may incorporate "generalized leaves," moving components such as the Agility 160 MLC's diaphragms which must also be considered when generating optimal MLC configurations. It would be prudent to also consider alternative mechanisms for defining the cost-density function in the isocentric plane from the BEV perspective, possibly from dosimetric considerations in the patient volume.

To summarize, this approach admits a wealth of possible strategies to leaf-fitting that potentially allow for indirect control of the quality of MLC tracking in real-time, and admits, in most clinically relevant situations, a fast, efficient, and high-quality implementation.

5. CONCLUSION

A framework for constructing real-time MLC leaf-fitting algorithms that explicitly optimize for underdose/overdose regions was described, and a particular implementation was presented for the use case of multiple targets and OARs, each with uniform underdose weighting. The resulting algorithm was compared to two standard algorithms: midleaf and Ruan 2011. Our approach was found to perform leaf-fittings for apertures with 80 leaf pairs in an average of 0.226 ms while accounting for both 3D translations and in-plane rotations. This performance was comparable to that of the midleaf approach (0.082 ms) but was roughly two orders of magnitude faster than the naïve implementation of Ruan 2011 (64.2 ms). At the same time, the piecewise algorithm generates leaf configurations with qualities-of-fit comparable to those from Ruan 2011, thereby showing significant improvement over the midleaf approach when underdose and overdose are unequally weighted.In this way, we improve upon the quality of the midleaf, but do so without sacrificing run-time performance.

ACKNOWLEDGMENTS

The authors would like to thank Paul Keall's group at The University of Sydney and Jeremy Booth at the Royal Northshore Hospital, Sydney, for providing the lung motion traces used in this study. Additionally, the authors acknowledge funding support from Elekta Limited and NIH Grant No. R01CA169102.

[a]Author to whom correspondence should be addressed. Electronic mail: douglas.moore@utsouthwestern.edu

[1]J. S. Kuo, C. Yu, Z. Petrovich, and M. I. J. Apuzzo, "The CyberKnife stereotactic radiosurgery system: Description, installation, and an initial evaluation of use and functionality," Neurosurgery **53**, 1235–1239 (2003).

[2]D. W. Andrews, G. Bednarz, J. J. Evans, and B. Downes, "A review of 3 current radiosurgery systems," Surg. Neurol. **66**, 559–564 (2006).

[3]Y. Kamino, K. Takayama, M. Kokubo, Y. Narita, E. Hirai, N. Kawawda, T. Mizowaki, Y. Nagata, T. Nishidai, and M. Hiraoka, "Development of a four-dimensional image-guided radiotherapy system with a gimbaled x-ray head," Int. J. Radiat. Oncol., Biol., Phys. **66**, 271–278 (2006).

[4]T. Depuydt, D. Verellen, O. Haas, T. Gevaert, N. Linthout, M. Duchateau, K. Tournel, T. Reynders, K. Leysen, M. Hoogeman, G. Storme, and M. De Ridder, "Geometric accuracy of a novel gimbals based radiation therapy tumor tracking system," Radiother. Oncol. **98**, 365–372 (2011).

[5]W. D. D'Souza, S. A. Naqvi, and C. X. Yu, "Real-time intra-fraction-motion tracking using the treatment couch: A feasibility study," Phys. Med. Biol. **50**, 4021–4033 (2005).

[6]P. J. Keall, H. Cattell, D. Pokhrel, S. Dieterich, K. H. Wong, M. J. Murphy, S. S. Vedam, K. Wijesooriya, and R. Mohan, "Geometric accuracy of a real-time target tracking system with dynamic multileaf collimator tracking system," Int. J. Radiat. Oncol., Biol., Phys. **65**, 1579–1584 (2006).

[7]A. Sawant, R. Venkat, V. Srivastava, D. Carlson, S. Povzner, H. Cattell, and P. Keall, "Management of three-dimensional intrafraction motion through real-time DMLC tracking," Med. Phys. **35**, 2050–2061 (2008).

[8]M. B. Tacke, S. Nill, A. Krauss, and U. Oelfke, "Real-time tumor tracking: Automatic compensation of target motion using the Siemens 160 MLC," Med. Phys. **37**, 753–761 (2010).

[9]A. Krauss, S. Nill, M. Tacke, and U. Oelfke, "Electromagnetic real-time tumor position monitoring and dynamic multileaf collimator tracking using a Siemens 160 MLC: Geometric and dosimetric accuracy of an integrated system," Int. J. Radiat. Oncol., Biol., Phys. **79**, 579–587 (2011).

[10]M. F. Fast, S. Nill, J. L. Bedford, and U. Oelfke, "Dynamic tumor tracking using the Elekta Agility MLC," Med. Phys. **41**, 111719 (5pp.) (2014).

[11]M. Glitzner, S. P. M. Crijns, B. D. de Senneville, J. J. W. Lagendijk, and B. W. Raaymakers, "On the suitability of Elekta's Agility 160 MLC for tracked radiation delivery: Closed-loop machine performance," Phys. Med. Biol. **60**, 2005–2017 (2015).

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

[13]A. Sawant, R. L. Smith, R. B. Venkat, L. Santanam, B. C. Cho, P. Poulsen, H. Cattell, L. J. Newell, P. Parikh, and P. J. Keall, "Toward submillimeter accuracy in the management of intrafraction motion: The integration of real-time internal position monitoring and multileaf collimator target tracking," Int. J. Radiat. Oncol., Biol., Phys. **74**, 575–582 (2009).

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

[15]D. Ruan and P. J. Keall, in *Presented at the 2011 IEEE Power Engineering and Automation Conference (PEAM)* (2011).

[16]C. Y. Huang, J. N. Tehrani, J. A. Ng, J. Booth, and P. Keall, "Six degrees-of-freedom prostate and lung tumor motion measurements using kilovoltage intrafraction monitoring," Int. J. Radiat. Oncol., Biol., Phys. **91**, 368–375 (2015).