Published in final edited form as:

Phys Med Biol. 2010 February 7; 55(3): 883–902. doi:10.1088/0031-9155/55/3/022.

A two-stage sequential linear programming approach to IMRT dose optimization

Hao H Zhang¹, Robert R Meyer², Jianzhou Wu³, Shahid A Naqvi¹, Leyuan Shi⁴, and Warren D D'Souza¹

¹Department of Radiation Oncology, University of Maryland School of Medicine, Baltimore, MD

²Computer Sciences Department, University of Wisconsin, Madison, WI

³Department of Radiation Medicine, Oregon Health and Science University School of Medicine, Portland, OR

⁴Department of Industrial and Systems Engineering, University of Wisconsin, Madison, WI

Abstract

The conventional IMRT planning process involves two stages in which the first stage consists of fast but approximate idealized pencil beam dose calculations and dose optimization and the second stage consists of discretization of the intensity maps followed by intensity map segmentation and a more accurate final dose calculation corresponding to physical beam apertures. Consequently, there can be differences between the presumed dose distribution corresponding to pencil beam calculations and optimization and a more accurately computed dose distribution corresponding to beam segments that takes into account collimator-specific effects. IMRT optimization is computationally expensive and has therefore led to the use of heuristic (e.g., simulated annealing and genetic algorithms) approaches that do not encompass a global view of the solution space. We modify the traditional two-stage IMRT optimization process by augmenting the second stage via an accurate Monte-Carlo based kernel-superposition dose calculations corresponding to beam apertures combined with an exact mathematical programming based sequential optimization approach that uses linear programming (SLP). Our approach was tested on three challenging clinical test cases with multileaf collimator constraints corresponding to two vendors. We compared our approach to the conventional IMRT planning approach, a direct-aperture approach and a segment weight optimization approach. Our results in all three cases indicate that the SLP approach outperformed the other approaches, achieving superior critical structure sparing. Convergence of our approach is also demonstrated. Finally, our approach has also been integrated with a commercial treatment planning system and may be utilized clinically.

1. Introduction

Conventional IMRT optimization may be summarized as follows. Given a set of target volumes, a large set of idealized pencil beam dose calculations are first performed. The number of pencil beams per beam angle may vary from 25–500. Since pencil beam calculations are time-consuming process, investigators have typically resorted to fast and approximate dose calculation methods (Bortfeld et al 1993, Deasy et al 2002, Thieke et al 2002, Wu et al

hzhan001@umaryland.edu, rrm@cs.wisc.edu, wuji@ohsu.edu, snaqvi@umm.edu, leyuan@engr.wisc.edu and wddsouza@gmail.com.

Publisher's Disclaimer: This is an author-created, un-copyedited version of an article accepted for publication in Physics in Medicine and Biology. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The definitive publisher authenticated version is available online at <http://iopscience.iop.org/0031-9155/55/3/022/>.

2003). Such approximate dose calculation methods generally do not consider collimator-specific effects. This optimization of pencil beam weights yields an idealized intensity map (I-map), which is real-valued. I-map segmentation algorithms are then applied to extract a set of deliverable beam apertures and corresponding aperture weights. Several investigators have reported on discretizations of I-maps (Langer et al 2001, Luan et al 2004, Gunawardena et al 2006) to generate deliverable apertures with corresponding weights. A final 3D dose calculation is then performed that considers collimator-specific effects for each aperture. The consequence of (1) using such approximate and fast dose calculation methods for the initial pencil beam calculation and (2) discretization of the real-valued I-map is that there arise significant differences between the putative dose distributions (and DVHs) based on idealized pencil beam calculations and the dose distributions from a final more accurate dose calculation.

To overcome such limitations, several investigators have proposed a direct-aperture optimization (DAO) approach to dose optimization (Shepard et al 2002, Earl et al 2003, Men et al 2007). DAO is an IMRT inverse planning technique which optimizes leaf positions and relative weights of the apertures simultaneously thereby, eliminating the discretization of I-maps and extraction of beam segments step in the conventional IMRT process. The user specifies the delivery angles and the number of apertures to deliver for each angle and thus, has control over the complexity of the treatment plan.

Segment weight optimization (SWO) is an alternative approach that is applied after the segmentation step of the conventional IMRT optimization process. SWO in principle uses the same aperture set obtained by conventional IMRT, then relies on a more accurate dose calculation corresponding to each beam aperture taking into account aperture specific effects and re-optimizes the weights associated with them to improve the plan quality.

Irrespective of the approach employed, most optimization tools currently available to planners use heuristic or local search methods, such as simulated annealing (Webb 1989, Mageras and Mohan 1993, Morrill et al 1995), genetic algorithms (Ezzell 1996, Wu and Zhu 2001), tabu search (Gillio 1998), and gradient-related methods (Spirou and Chui 1998, Cotrutz et al 2001, Chang et al 2002). The advantage of using such methods is speed and the handling of large-scale instances. The disadvantage is that such approaches do not necessarily guarantee global optimality. On the other hand, exact mathematical programming approaches (e.g. linear programming) are well developed in the Operations Research area, but have not been widely utilized in IMRT planning. If the problem is feasible and bounded, which is the case for the IMRT planning problem, then formulating and solving it by an exact approach guarantees the optimal solution.

In this paper, we employ the principle of SWO and propose a two-stage approach to IMRT dose optimization by using the first stage of the conventional IMRT optimization process followed by I-map segmentation as a 'warm-start' approach to initiate a second stage optimization process. This second stage in our approach utilizes aperture-specific Monte Carlo based kernel-superposition dose calculations to provide input data to an optimization process. In our approach, the dose optimization is formulated as a linear program (LP). Due to problem size constraints, we then use a sequential linear programming (SLP) approach to solve the corresponding dose optimization problem by dividing the problem into a set of inter-connected sub-problems, each of which is solved using LP while maintaining a global view of the full problem. We tested our approach on three clinical cases (two head and neck and one whole pelvis/prostate) and compared it with the conventional IMRT optimization approach, a DAO approach and SWO approach. Tests were performed by considering delivery constraints for multileaf collimators from two vendors (Elekta Oncology Systems, Crawley, UK and Varian Medical Systems, Palo Alto, CA). The convergence of our approach was also investigated.

2. Methods and Materials

2.1. Two-stage sequential linear programming approach

The two-stage approach in this work employs a warm-start approach and involves (1) the generation of an initial set of beam segments from the first stage of the conventional IMRT process, (2) calculation of dose distributions corresponding to these beam apertures by considering collimator-specific effects, (3) formulation of the dose optimization problem as a LP and (4) solution of the dose optimization problem by dividing the problem into several connected sub-problems and solving each sub-problem one at a time. The first stage of our approach involves (1) and (2) above while the second stage of our approach involves (3) and (4) above.

2.1.1. Generation of initial set of beam segments—Conventional IMRT optimization was first performed in the Pinnacle³ planning system (Philips Medical Systems, Cleveland, OH). The resulting idealized intensity map for each beam was segmented (leaf-sequencing) in Pinnacle³ to yield a set of deliverable apertures based on MLC constraints corresponding to each linear accelerator vendor (Elekta and Varian).

2.1.2. Beam segment-specific dose calculations—The dose distribution corresponding to each individual aperture from Pinnacle³ was calculated using the Monte Carlo Kernel Superposition (MCKS) dose calculation code (Naqvi et al 2003). MCKS explicitly allows sampling of individual photons from the head and propagates them through the exact collimator structure and the patient. The explicit photon sampling and modeling of physical collimator structure allows MCKS to accurately account for collimator specific effects such as MLC leaf curvature, tongue-and-groove, and beam hardening. MCKS allows sampling from a variety of photon sources. For convenience, we utilized the dual-source model (Naqvi et al 2001,2003) for the Elekta machine. For the Varian machine, we used a phase-space source (Naqvi et al 2005) obtained from a prior EGSnrc/BEAMnrc simulation (Rogers et al 1995) of the physical geometry of the Varian 21EX from the tungsten target to a plane just above the highest secondary collimator. The MCKS code has been benchmarked against the EGSnrc/dosxyznrc code (Naqvi et al 2003). MCKS, has more accurate modeling of photons incident on the patient than deterministic convolution superposition codes, and does not require explicit fluence and kernel hardening corrections. In this paper, MCKS is used to obtain 3D dose matrices for each beam aperture of each beam angle to be used in subsequent optimization.

2.1.3. Subdivision of dose optimization problem and sequential optimization—The overall size of the dose optimization problem, which includes the dose matrices associated with all beam apertures combined with target and critical structure voxel constraints, makes it difficult to solve as a single problem using an exact approach such as an LP (described in 2.1.4). Table 1 shows the size of the problem in the three cases considered including number of voxels, number of constraints and number of variables. As a result, the overall problem is divided into sub-problems. Each sub-problem only considers the apertures corresponding to a single beam angle. The optimization of aperture weights for the current angle is performed by considering the full set of voxels and the dose contributions from other apertures (fixed to values from the previous sub-problems). This is analogous to a block-Gauss-Seidel approach (Quarteroni et al 2007), in which the weights corresponding to beam apertures for a given beam angle are optimized a block at a time.

To accelerate the convergence of our approach we employed a warm-start. A variety of approaches, such as class solutions (De Gersem et al 2001, Arráns et al 2003, Beaulieu et al 2004) have previously been considered for warm-starts. While numerous warm-start approaches can be considered, in this study, we investigated two warm-start approaches for

the initial set of aperture weights derived from the conventional IMRT planning approach in the Pinnacle³ planning system: (1) Pinnacle³ beam weights obtained from the first stage of the optimization in section 2.1.1 and (2) equal dose contributions from each beam angle. While both approaches resulted in identical final solutions, the first warm-start approach converged faster and so we will focus on this approach exclusively in this work. The sequential approach was initiated by optimizing the weights for a set of apertures corresponding to the first angle subject to the weighted dose contributions from the remaining apertures. As seen from Figure 1, in the initial iteration the aperture weights of Beam 1 are optimized subject to constraints with the assumption that the contributions from other beams are fixed (aperture weights fixed to the warm-start values). The optimization of a set of aperture weights for a given beam angle is referred to as an iteration. Next, the LP to optimize the aperture weights of the second angle is solved taking into account the newly computed weights for the apertures in the first angle and the warm-start weights for the remaining angles as seen from Figure 1. This process is continued until the weights of the apertures corresponding to all involved beam angles have been optimized (referred to as a sweep). After one complete sweep, which involves solving as many LPs as there are beam angles, a new set of aperture weights for all angles is obtained. We use seven angles to further illustrate the sequential optimization approach (see Fig 2). The seven sub-problems are referred to as $SP_1, \dots, SP_7$. The set of aperture weights in SP_1 are denoted by w_i for $i \in \theta_1$. Conceptually, suppose the dose contributions from each angle are denoted by $D_1, \dots, D_7$. In each sweep of seven angles, we have seven iterations. In each iteration k , where $k = 1, \dots, 7$, we solve the sub-problem SP_k assuming a dose contribution D_r from each angle $r \neq k$. This process is repeated for each of the seven sub-problems so that all apertures are considered in each sweep.

Successive iterations and sweeps are performed using the most recent aperture weights. The optimization is terminated when convergence is achieved. Convergence was determined by calculating the change in aperture weights between successive sweeps. The criteria for convergence were a maximum change in aperture weights w_i of less than 1% between sweeps. This ensures that no aperture weight changes significantly between sweeps.

2.1.4. Formulation of dose optimization problem as LP—Each sub-divided problem was formulated as an LP and solved using the simplex method. An LP problem consists of constraints in the form of a collection of linear inequalities (and, optionally, equations) on a number of real variables and a given linear function (on these real variables) which is to be maximized or minimized. The explicit definition of constraints provides greater control of the optimization model and thereby reduces the need for the trial-and-error often used approach in conventional IMRT planning. The LP was formulated in the General Algebraic Modeling System (GAMS Development Corporation, Washington, DC), a high-level modeling system for mathematical programming and optimization and solved with an integrated high-performance commercial solver CPLEX (ILOG Inc., Sunnyvale, CA). We introduce i as an index of apertures and j as an index of voxels. When solving each sub-problem, the LP dose optimization objective (or cost function) is to minimize a weighted sum of doses delivered to the voxels that violate the thresholds related to dose-volume histogram (DVH) constraints. The treatment region is divided into regions: region O includes organs-at-risk (OARs), region T includes the planning target volume. Input data includes beam-aperture-specific dose matrices B corresponding to each delivery aperture of each utilized angle.

DVH constraints are handled through the introduction of penalty variables that account for target and OAR underdose and overdose as relevant. $b^{TARL}, b^{TARU}, b^{OAR}$ and b^{SPC} specify the desired dose thresholds for each region (the superscripts mean target, organ-at-risk, and any special region) at which DVH constraints are desired. $x_j, j \in T$ is a vector of underdoses/overdoses (doses below the prescription level or over certain level (hot spot) for target voxels) delivered to the target region. The variable $x_j, j \in O$ is a vector of overdoses (doses over a

certain critical level) delivered to the OARs. S is a subset of O , which corresponds to a special region such as the spinal-cord voxels, in which hard upper-bound dose constraints instead of DVH constraints can be specified. w_i is the weight (to be optimized) for aperture i of beam angle set θ . α^{TAR_L} , α^{TAR_U} and α^{OAR} are the penalty weights assigned for underdosing, overdosing the target and overdosing the OARs (different weights may be set for different OARs). The overall problem from which the sub-problems are derived is thus:

$$\text{Minimize } \alpha^{TAR_L} \sum_{j \in T} x_j + \alpha^{TAR_U} \sum_{j \in T} x_j + \alpha^{OAR} \sum_{j \in O} x_j$$

$$\text{subject to } \sum_{i \in \theta} w_i B_{i,j} + x_j \geq b^{TAR_L} \quad \text{for } j \in T \quad (1)$$

$$\sum_{i \in \theta} w_i B_{i,j} \leq x_j + b^{TAR_U} \quad \text{for } j \in T \quad (2)$$

$$\sum_{i \in \theta} w_i B_{i,j} \leq x_j + b^{OAR} \quad \text{for } j \in O \quad (3)$$

$$\sum_{i \in \theta} w_i B_{i,j} \leq b^{SPC} \quad \text{for } j \in S \quad (4)$$

$$w_i, x_j \geq 0 \quad (5)$$

Each sub-problem k is derived from the overall problem by fixing intensity terms w_i , $i \notin \theta_k$ (beams weights other than beam k) so that the problem formulation for the sub-problem becomes:

$$\text{Minimize } \alpha^{TAR_L} \sum_{j \in T} x_j + \alpha^{TAR_U} \sum_{j \in T} x_j + \alpha^{OAR} \sum_{j \in O} x_j$$

$$\text{subject to } \sum_{i \in \theta_k} w_i B_{i,j} + x_j \geq b^{TAR_L} - \sum_{i \notin \theta_k} w_i B_{i,j} \quad \text{for } j \in T \quad (6)$$

$$\sum_{i \in \theta_k} w_i B_{i,j} \leq x_j + b^{TAR_U} - \sum_{i \notin \theta_k} w_i B_{i,j} \quad \text{for } j \in T \quad (7)$$

$$\sum_{i \in \theta_k} w_i B_{i,j} \leq x_j + b^{OAR} - \sum_{i \notin \theta_k} w_i B_{i,j} \quad \text{for } j \in O \quad (8)$$

$$\sum_{i \in \theta_k} w_i B_{i,j} \leq b^{SPC} - \sum_{i \notin \theta_k} w_i B_{i,j} \quad \text{for } j \in S \quad (9)$$

$$w_i, x_i \geq 0 \quad (10)$$

2.2. Test Cases

We tested the SLP approach on three challenging clinical cases (two head and neck and one prostate case, both involving primary tumors and nodal volumes) which are routinely treated with IMRT in our clinic. In the first head and neck case, the target volume consisted of a primary supraglottic tumor and a high risk nodal volume (**Case 1**). In the second head and neck case, the target volume consisted of a tonsillar primary tumor, high risk nodal volume and low risk nodal volume (**Case 2**). In both cases, a simultaneous-integrated boost (SIB) was selected as the mode of delivery, i.e., different dose prescriptions for primary tumor and nodal volume for the same number of treatment fractions. The OARs included the brainstem, parotid glands, spinal cord and anterior oral mucosa. We used previously published constraints for head and neck IMRT (Lee et al 2007) in the optimization process with a dose prescription of 70 Gy to the primary tumor, 59.4 Gy to the high risk nodal volume and 50.4 Gy to the low risk nodal volume. In the prostate case, the target volume consisted of the prostate and the pelvic lymph nodes (**Case 3**). The critical structures included the rectum, bladder and small bowel. The dose volume constraints were derived from a national cooperative group trial, Radiation Therapy Oncology Group (RTOG 0126), with a dose prescription of 72 Gy (45 Gy to the whole pelvis including prostate and nodal volume and 27 Gy boost to the prostate alone). Tables 2–4 summarize the dose constraints for the cases considered in this work.

2.3. Plan Comparisons

In this work, the two-stage SLP approach was compared with 3 other planning approaches: (1) “conventional” IMRT approach, (2) Direct Machine Parameter Optimization (DMPO), an example of a DAO approach and (3) segment weight optimization (SWO). All of the above approaches were executed in the Pinnacle³ planning system. The “conventional” planning approach involved a fast dose calculation corresponding to idealized pencil beams, pencil beam optimization followed by the generation of real-valued intensity-maps. These I-maps were then discretized and the leaf sequencing algorithm in the planning system was applied to generate a set of deliverable shapes. DMPO is an example of a DAO approach consisting of directly optimizing the aperture shapes and weights. In our execution of DMPO, the maximum number of aperture shapes allowed for each beam angle was set to the same number as that obtained from the “conventional” approach. This was done to prevent artificial degradation in plan quality due to limiting the number of apertures. SWO involved re-optimizing the aperture weights obtained from the “conventional” approach in which Pinnacle³ uses a gradient-based method for the optimization algorithm (Gill et al 1998).

Although the underlying principle in SLP is similar to SWO, there are two differences between the SWO employed in the Pinnacle³ planning system and the SLP approach. First, while both the planning system and our approach calculate aperture-specific dose distributions to begin the second stage of the optimization process (see section 2.1), the planning system uses

convolution/superposition while we use the MCKS code for dose calculation. Second, once the aperture-specific dose calculations are generated, the planning system uses a gradient search algorithm while we use the SLP approach described above to optimize the weights of the apertures. To enable a fair comparison, the final dose distributions for all of the optimization approaches were calculated using the MCKS dose algorithm. The four optimization approaches were tested on three cases described above using the same constraints summarized in Tables 2–4 for the 120-leaf MLC corresponding to Varian and the 80-leaf MLC corresponding to Elekta. Delivery-specific constraints for both vendors were considered.

3. Results

The results summarized below are based on using the weights of the apertures generated by the conventional planning approach in Pinnacle³ as the warm-start for both the SWO and SLP approaches. While we did experiment with a warm-start in which we assumed equal contribution from beams, SLP converged to the same solution but with different speed of convergence (see 3.5 for details). Computation times corresponding to each approach are reported in Table 5, and were based on the performance on a SunBlade 2500 machine.

3.1. Delivery Efficiency

Table 5 summarizes the number of apertures and monitor units (MUs) for the four approaches compared in this work. Even though the apertures used to initiate the SLP approach were those obtained from the “conventional” approach, the final number of apertures with non-zero weights following LP-based optimization was less than the initial number. This phenomenon was observed for all 3 cases and for both Elekta and Varian MLCs: the SLP approach yielded the lowest number of apertures. We also observed that following SLP-based optimization, the aperture weights (and as a result monitor units) changed substantially from the initial Pinnacle³ derived weights used in the warm-start. Fig. 3 shows the change in monitor units between the conventional IMRT plans (corresponding to the warm-start) and SLP plans. The weights associated with some of apertures following SLP was zero, thereby, contributing to a reduction in the number of apertures compared to the warm-start.

3.2. Test Case 1

3.2.1. Elekta Results—For Case 1, the conventional plan and SWO plan used 159 apertures (7 angles), the DMPO plan used 133 apertures and the SLP plan used 113 apertures (46 of the original apertures yielded weights of zero). Fig. 4 shows a comparison of the target and OAR DVHs resulting from the four optimization approaches. All plans were normalized so that 95% of PTV (primary) received 70 Gy. The SLP plan outperformed the other three plans in general. We quantify the plan quality by comparing the percentage of differences of values (percent of volume exceeded the DVH constraints and maximum doses received by critical structures, such as spinal cord, brainstem, vocal cord) across different approaches calculated at the constraint levels. The SLP optimization approach resulted in an improvement of 32–39% for the right parotid (RP), 19–39% for the left parotid (LP) and 29–41% for mucosa compared with the other optimization approaches. The SLP approach also resulted in the lowest maximum cord dose and lowest maximum brainstem dose. These improvements were achieved with 15–29% less apertures but 8–31% more MUs.

3.2.2. Varian Results—The same case was also tested on Varian machine: the conventional plan, DMPO plan and SWO plan all used 141 apertures (7 angles), the SLP plan used 107 apertures. Fig. 5 shows a comparison of the DVHs resulting from the four optimization approaches. The SLP plan still outperformed the other three plans most of the time. The SLP optimization approach resulted in an improvement of 9–34% for the right parotid (RP), 1–34% for the left parotid (LP) and 12–38% for mucosa. The SLP approach still resulted in the lowest

maximum cord dose and lowest maximum brainstem dose. These improvements were achieved with 24% less apertures but 3–15% more MUs.

3.3. Test Case 2

3.3.1. Elekta Results—For Case 2, the conventional plan and SWO used 357 apertures (9 angles), the DMPO plan used 339 apertures and the SLP plan used 274 apertures. The SLP approach achieved 19–23% aperture number reduction. Fig. 6 shows a comparison of DVHs resulting from the four optimization approaches for this more challenging case. All plans were normalized so that 95% of PTV (primary) received 70 Gy. The SLP optimization approach resulted in similar DVHs for the right parotid (maximum of 7% difference), an improvement of 4–22% for the left parotid compared with the other optimization approaches. For the vocal cord, the SLP improved 12–48% compared with conventional and SWO plan, but 20% worse than DMPO plan. The SLP approach also resulted in the lowest maximum cord and vocal cord doses. For this case, the MUs used by SLP approach were within the range of the other approaches.

3.3.2. Varian Results—The conventional plan and SWO used 172 apertures (9 angles), the DMPO plan used 168 apertures and the SLP plan used 143 apertures. The SLP approach achieved more than 15% aperture reduction without a significant increase in MUs. Fig. 7 shows a comparison of DVHs. The SLP approach resulted in lower volumes receiving more than 55 Gy for the right parotid, an improvement of 2–41% for the left parotid. For the vocal cord, the SLP plan is about 30% better than conventional plan, but 30% worse than DMPO and SWO plan. The SLP approach also resulted in the lowest maximum cord and vocal cord doses.

3.4. Test Case 3

3.4.1. Elekta Results—For Case 3, both the conventional plan and SWO plan used 85 apertures (7 angles, initial plan and boost plan together), the DMPO plan used 82 apertures and the SLP plan used 67 apertures. Fig. 8 shows a comparison of the prostate and OAR DVHs. The DVHs were derived from summing the doses corresponding to the initial and boost plans, both of which were normalized such that 100% of the prostate received the initial and boost prescription doses. At higher dose levels, the SLP approach spared rectum and small bowel better than the other approaches. The SLP approach used about 6–22% more MUs than the other approaches, but 18–22% less apertures.

3.4.2. Varian Results—The conventional plan and the SWO plan both used 135 apertures, the DMPO plan used 125 apertures and the SLP plan used 118 apertures. Fig. 9 shows a comparison of DVHs resulting from the four optimization approaches. The SLP approach protected the rectum and small bowel better than the other approaches. For this case, the SLP approach also used the least number of MUs among all approaches.

3.5. Clinical convergence of the SLP process

We demonstrate the convergence of the approach by answering two questions (1) Does the convergence of the SLP approach depend on the initial points (different warm-start approaches) and (2) Does the convergence of the SLP approach depend on the order of the beam angles considered in the sequential optimization process? We use Fig. 10 to answer the questions. Note here we used one case for demonstration, other cases behaved the same. Fig. 10a shows that with two different starting points, the method converged to the same solution with different speed. This means using different initial aperture weights only affects the speed of the convergence of our approach and will not affect the final plan quality. We also tested three different orders in which beam angles were selected and their corresponding apertures optimized: sequential order (clockwise selection), reverse order (counter-clockwise selection)

and random order. Fig. 10b shows that after 10 sweeps, there was little difference in the objective function values. Fig. 10c further shows that there is minimal clinical difference in the convergence of the DVHs as a function of the beam angle order (there was also minimal difference in the DVHs with different initial points) after 10 sweeps. Fig. 10d and Fig. 9e show the convergence of DVHs of the right parotid and spinal cord, respectively.

4. Discussion

The purpose of this work was to develop better IMRT plans by replacing the conventional SWO stage of the IMRT optimization process. We demonstrate that following the generation of a reasonable set of apertures derived from the Pinnacle³ planning system, highly accurate input data can be provided to an exact LP-based optimization process using aperture-specific Monte-Carlo-based dose calculations. The proposed exact optimization approach, SLP, solves the corresponding dose optimization problem by dividing it into a set of smaller interconnected sub-problems, each of which is solved using LP while maintaining a global view of the full problem. Up to 48% improvement of volume receiving restricted dose was achieved compared with conventional IMRT, DMPO and SWO performed in the Pinnacle³ planning system when considering both Elekta and Varian MLCs. The overall dose distribution was greatly improved, and the SLP approach also using fewer apertures and comparable MUs when compared with the other approaches.

When the IMRT planning problem is solved using local search or heuristic methods, such as gradient methods, simulated annealing, or genetic algorithms, a key issue is that the search is likely to be trapped in local minima (Deasy 1997, Wu and Mohan 2002, Llacer et al 2003). One of the main conclusions of these investigations is that DV constraints may make the feasible space nonconvex and this may lead to multiple minima. It has been demonstrated that even when multiple minima do exist for some specially designed cases, some of these minima are very close to each other in cost and the resulting treatment plans are practically identical (Wu and Mohan 2002, Llacer et al 2003). Our approach overcomes these concerns by formulating the dose optimization problem as an LP and obtaining the global solution via an exact math-programming approach. Thus, when formulated appropriately, the planner need not be concerned with local minima. Our approach guarantees global optimality, which will return a high quality treatment plan.

In this research, the initial set of apertures were generated from conventional IMRT plans. Using these apertures, quicker convergence of SLP was achieved and provided integration of our approach with a commercial planning system. With these reasonable apertures, good quality treatment plans were obtained together with improved delivery efficiency. The disadvantage is the additional computation time corresponding to the needed conventional plan before execution of the SLP approach. Other approaches, such as anatomy-based (De Gerssem et al 2001, Beaulieu et al 2004) and class solutions (Arráns et al 2003, Mott et al 2004) have been proposed to generate reasonable aperture sets. These aperture sets may be applied repeatedly for patients with the same stage of disease and anatomical site. Although these methods have not been widely applied in clinics, the advantage is the quick generation time (usually, seconds). Our method can take advantage of these approaches to reduce total computation time thereby, eliminating the need for the generation of the conventional IMRT plan. In future work, we aim to identify approaches that might provide a reasonable set of initial apertures, so that the computational efficiency is improved without compromising treatment plan quality.

In order to simplify IMRT plans while improving delivery and radiation efficiency, some researchers have proposed direct aperture optimization (DAO), which directly optimizes weights and leaf positions with a predetermined number of MLC shapes permitted for each

beam (Keller-Reichenbecher 1999, Shepard et al 2002, Jiang et al 2005, Van Asselen et al 2006). DAO is a one-step optimization process that eliminates the step of leaf sequencing in two-step optimization and allows users to have direct control of the complexity of IMRT plans. In an initial study using their in-house software, Shepard et al (2002) demonstrated the feasibility and advantages of using DAO on a head and neck case and a prostate case. Compared with traditional two-step optimization, they found that DAO resulted in a significant reduction in both the number of apertures and the number of MUs (Shepard et al 2002). However, they only considered cases consisting of the primary tumor with no nodal volumes.

In this work, we set the maximum number of allowable apertures in the DMPO plans to be the same as the number obtained from the conventional IMRT process. Our results show that DMPO did not decrease the number of apertures compared with the conventional IMRT approach for one of the six cases studied here. However, it should be pointed out that the actual shapes of the apertures generated by DMPO differ from those generated by Pinnacle in our “conventional” plans. SWO on the other hand used the same apertures generated by conventional IMRT and re-optimized the aperture weights using a gradient search algorithm. It was observed that the SWO approach did not reduce the apertures compared to conventional IMRT. As in SWO, SLP also started with the same apertures as conventional IMRT. However, SLP used the least apertures in all six cases. This reduction resulted from the assignment of zero weights to some of the apertures during LP optimization, thereby eliminating some of the apertures in the conventional IMRT process. Thus, SLP was able to obtain high quality plans while improving delivery efficiency.

In this study the DMPO plans did not in general result in a reduction in aperture number compared with the conventional IMRT plans while the SLP plans resulted in the lowest number of apertures. We attribute the differences in our results to those of Shepard et al with respect to the lack of reduction of the number of apertures in the direct aperture approach to the complex cases studied (primary tumor plus nodule volumes) compared to the cases reported in Shepard et al. In this study, we found that SLP used comparable amount of MUs as the other approaches. Compared with conventional IMRT, SLP used 88%–118% of MUs that conventional IMRT used. SLP used 81%–144% of MUs that DMPO used and 84%–108% of MUs that SWO used.

5. Conclusion

We have demonstrated that for the complex clinical cases considered in this work, the SLP approach can yield very desirable plans with adequate coverage of the PTV and in general superior sparing of critical structures when compared with the conventional IMRT, DMPO, and SWO approaches. These results were obtained with comparable delivery efficiency compared with the other approaches considered.

Acknowledgments

This work was supported in part by grants from the NSF DMI 0355567, 0400294 and NIH/NCI CA130814.

References

- Arráns R, Gallardo MI, Roselló J, Sánchez-Doblado F. Computer optimization of class solutions designed on a beam segmentation basis. *Radiother. Oncol* 2003;69:315–321. [PubMed: 14644491]
- Beaulieu F, Beaulieu L, Tremblay D, Lachance B, Roy R. Automatic generation of anatomy-based MLC fields in aperture-based IMRT. *Med. Phys* 2004;31:1539–1545. [PubMed: 15259658]
- Bedford JL, Webb S. Constrained segment shapes in direct-aperture optimization for step-and-shoot IMRT. *Med. Phys* 2006;33:944–958. [PubMed: 16696471]
- Bortfeld T, Schlegel W, Rhein B. Decomposition of pencil beam kernels for fast dose calculations in three-dimensional treatment planning. *Med. Phys* 1993;20:311–318. [PubMed: 8497215]

- Chang SX, Cullip TJ, Rosenman JG, Halvorsen PH, Tepper JE. Dose optimization via index-dose gradient minimization. *Med. Phys* 2002;29:1130–1146. [PubMed: 12094983]
- Cotrutz C, Lahanas M, Kappas K, Baltas D. A multiobjective gradient-based dose optimization algorithm for external beam conformal radiotherapy. *Phys. Med. Biol* 2001;46:2161–2175. [PubMed: 11512617]
- De Gersem W, Claus F, De Wagter C, De Neve W. An anatomy based beam segmentation tool for intensity-modulated radiation therapy and its application to head-and-neck cancer. *Int. J. Radiat. Oncol. Biol. Phys* 2001;51:849–859. [PubMed: 11699497]
- Deasy JO. Multiple local minima in radiotherapy optimization problems with dose-volume constraints. *Med. Phys* 1997;24:1157–1161. [PubMed: 9243478]
- Deasy JO, Wickerhauser MV, Picard M. Accelerating Monte Carlo simulations of radiation therapy dose distributions using wavelet threshold de-noising. *Med. Phys* 2002;29:2366–2373. [PubMed: 12408311]
- Earl MA, Shepard DM, Naqvi S, Li XA, Yu CX. Inverse planning for intensity-modulated arc therapy using direct aperture optimization. *Phys. Med. Biol* 2003;48:1075–1089. [PubMed: 12741503]
- Ezzell GA. Genetic and geometric optimization of three-dimensional radiation therapy treatment planning. *Med. Phys* 1996;23:293–305. [PubMed: 8815371]
- Gill, PE.; Murray, W.; Saunders, MA.; Wright, MH. User's guide for NPSOL 5.0: A Fortran package for nonlinear programming (Technical Report SOL 86–2). 1998.
- Gillio JP. Tabu search: an automatic treatment planning algorithm for conformal radiotherapy. *Med. Phys* 1998;25:796.
- Gunawardena ADA, D'Souza WD, Goadrich LD, Meyer RR, Sorensen KJ, Naqvi SA, Shi L. A difference-matrix metaheuristic for intensity map segmentation in step-and-shoot IMRT delivery. *Phys. Med. Biol* 2006;51:2517–2536. [PubMed: 16675867]
- Jiang Z, Earl MA, Zhang GW, Yu CX, Shepard DM. An examination of the number of required apertures for step-and-shoot IMRT. *Phys. Med. Biol* 2005;50:5653–5663. [PubMed: 16306659]
- Keller-Reichenbecher MA, Bortfeld T, Levegrun S, Stein J, Preiser K, Schlegel W. Intensity modulation with the 'step and shoot' technique using a commercial MLC: a planning study. *Int. J. Radiat. Oncol. Biol. Phys* 1999;45:1315–1324. [PubMed: 10613328]
- Langer M, Thai V, Papiez L. Improved leaf sequencing reduces segments or monitor units needed to deliver IMRT using multileaf collimators. *Med. Phys* 2001;28:2450–2458. [PubMed: 11797948]
- Lee N, Mechalakos J, Puri DR, Hunt M. Choosing an Intensity-Modulated Radiation Therapy technique in the treatment of head-and-neck cancer. *Int. J. Radiat. Oncol. Biol. Phys* 2007;68:1299–1309. [PubMed: 17241750]
- Llacer J, Deasy JO, Bortfeld TR, Solberg TD, Promberger C. Absence of multiple local minima effects in intensity modulated optimization with dose-volume constraints. *Phys. Med. Biol* 2003;48:183–210. [PubMed: 12587904]
- Luan S, Wang C, Hu XS, Naqvi SA, Yu CX, Lee CL. A new MLC segmentation algorithm/software for step-and-shoot IMRT delivery. *Med. Phys* 2004;31:695–707. [PubMed: 15124986]
- Mageras GS, Mohan R. Application of fast simulated annealing to optimization of conformal radiation treatments. *Med. Phys* 1993;20:639–647. [PubMed: 8350815]
- Men C, Romeijn HE, Taskin ZC, Dempsey JF. An exact approach to direct aperture optimization in IMRT treatment planning. *Phys. Med. Biol* 2007;52:7333–7352. [PubMed: 18065842]
- Mohan R, Arnfield M, Tong S, Wu Q, Siebers J. The impact of fluctuations in intensity patterns on the number of monitor units and the quality and accuracy of intensity modulated radiotherapy. *Med. Phys* 2000;27:1226–1237. [PubMed: 10902551]
- Morrill SM, Lam KS, Lane RG, Langer M, Rosen II. Very fast simulated reannealing in radiation therapy treatment plan optimization. *Int. J. Radiat. Oncol. Biol. Phys* 1995;31:179–188. [PubMed: 7995750]
- Mott JH, Livsey JE, Logue JP. Development of a simultaneous boost IMRT class solution for a hypofractionated prostate cancer protocol. *Br. J. Radiol* 2004;77:377–386. [PubMed: 15121701]
- Naqvi SA, Earl MA, Shepard DM. Convolution/superposition using the Monte Carlo method. *Phys. Med. Biol* 2003;48:2101–2121. [PubMed: 12894973]

- Naqvi SA, Sarfaraz M, Holmes T, Yu C, Li XA. Analysing collimator structure effects in head-scatter calculations for IMRT class fields using scatter raytracing. *Phys. Med. Biol* 2001;46:2009–2028. [PubMed: 11474941]
- Naqvi S, D'Souza W, Earl M, Ye S, Shih RP, Li XA. Using a photon phase-space source for convolution/superposition dose calculations in radiation therapy. *Phys. Med. Biol* 2005;50:4111–4124. [PubMed: 16177534]
- Quarteroni, A.; Sacco, R.; Saleri, F. *Numerical mathematics*. Springer; 2007.
- Rogers DWO, Faddegon BA, Ding GX, Ma CM, Wei J, Mackie TR. BEAM: A Monte Carlo code to simulate radiotherapy treatment units. *Med. Phys* 1995;22:503–524. [PubMed: 7643786]
- Shepard DM, Earl MA, Li XA, Naqvi S, Yu C. Direct aperture optimization: a turnkey solution for step-and-shoot IMRT. *Med. Phys* 2002;29:1007–1018. [PubMed: 12094970]
- Spirou SV, Chui CS. A gradient inverse planning algorithm with dose-volume constraints. *Med. Phys* 1998;25:321–333. [PubMed: 9547499]
- Thieke C, Nill S, Oelfke U, Bortfeld T. Acceleration of intensity-modulated radiotherapy dose calculation by importance sampling of the calculation matrices. *Med. Phys* 2002;29:676–681. [PubMed: 12033562]
- Van Asselen B, Schwarz M, Van Vliet-Vroegindeweij C, Lebesque JV, Mijnheer BJ, Damen EM. Intensity-modulated radiotherapy of breast cancer using direct aperture optimization. *Radiother. Oncol* 2006;79:162–169. [PubMed: 16712992]
- Webb S. Optimisation of conformal radiotherapy dose distributions by simulated annealing. *Phys. Med. Biol* 1989;34:1349–1370. [PubMed: 2682694]
- Wu Q, Djajaputra D, Lauterbach M, Wu Y, Mohan R. A fast dose calculation method based on table lookup for IMRT optimization. *Phys. Med. Biol* 2003;48:N159–N166. [PubMed: 12870589]
- Wu Q, Mohan R. Multiple local minima in IMRT optimization based on dose-volume criteria. *Med. Phys* 2002;29:1514–1527. [PubMed: 12148734]
- Wu X, Zhu Y. An optimization method for importance factors and beam weights based on genetic algorithms for radiotherapy treatment planning. *Phys. Med. Biol* 2001;46:1085–1099. [PubMed: 11324953]

Figure 1.
Illustration of the first two iterations of the sequential linear programming process

Figure 2.
Illustration of the second stage of the SLP dose optimization method

Figure 3.

Absolute MU difference corresponding to each aperture between the *warm-start* conventional Pinnacle³ plan and the SLP plan (a) Head and neck Case 1 tested on Elekta MLC (b) Head and neck Case 2 tested on Elekta MLC (c) Prostate case tested on Elekta MLC (d) Head and neck Case 1 tested on Varian MLC (e) Head and neck Case 2 tested on Varian MLC (f) Prostate case tested on Varian MLC

Figure 4.

DVHs comparing 4 approaches and corresponding to (a) primary tumor PTV, nodes PTV and right parotid (RP), (b) nodes CTV, left parotid (LP) and brainstem and (c) primary tumor CTV, mucosa and cord in the head and neck case 1 with Elekta MLC constraints

Figure 5.

DVHs comparing 4 approaches and corresponding to (a) primary tumor PTV, nodes PTV and right parotid (RP), (b) nodes CTV, left parotid (LP) and brainstem and (c) primary tumor CTV, mucosa and cord in the head and neck case 1 with Varian MLC constraints

Figure 6. DVHs comparing 4 approaches and corresponding to (a) primary tumor PTV, high-risk PTV, low-risk PTV and vocal cord, (b) primary tumor CTV, right parotid (RP) and cord (c) high-risk CTV, low-risk CTV and left parotid (LP) in the head and neck case 2 with Elekta MLC constraints

Figure 7.

DVHs comparing 4 approaches and corresponding to (a) primary tumor PTV, high-risk PTV, low-risk PTV and vocal cord, (b) primary tumor CTV, right parotid (RP) and cord (c) high-risk CTV, low-risk CTV and left parotid (LP) in the head and neck case 2 with Varian MLC constraints

Figure 8.

DVHs comparing 4 approaches and corresponding to (a) bladder and bowel, (b) prostate and rectum in the whole pelvis/prostate case with Elekta MLC constraints

Figure 9.

DVHs comparing 4 approaches and corresponding to (a) bladder and bowel, (b) prostate and rectum in the whole pelvis/prostate case with Varian MLC constraints

Figure 10.

Illustration of convergence of the SLP approach with (a) different starting points (b) same starting point but different aperture orders (c) comparison of final DVHs of RP and spinal cord with different aperture orders (d) convergence of RP DVH with increasing sweep counts (e) convergence of spinal cord DVH with increasing sweep counts

Table 1

Summary of the problem sizes

	Head and neck 1	Head and neck 2	Whole pelvis/prostate
Voxel number	54,768 (6,616,764) [†]	68,452 (4,633,087)	69,534 (4,239,360)
Number of constraints	77,021	137,670	139,068
Elekta			
Data size	1.8GB	2.8GB	674MB
Number of variables	75,899	136,752	139,156
Varian			
Data size	2.6GB	2.6GB	1.0GB
Number of variables	75,881	136,567	139,179

[†] Number of voxels were constrained on (total number of voxels)

Table 2

Summary of dose and dose-volume constraints for head and neck case 1

% volume of LP receiving ≥26 Gy	% volume of LP receiving ≥ 32 Gy	% volume of RP receiving ≥ 26 Gy	% volume of RP receiving ≥ 32 Gy	% volume of mucosa receiving ≥ 35 Gy	Maximum cord dose (Gy)	Maximum brainstem dose (Gy)
66	33	66	33	50	45	54

Table 3

Summary of dose and dose-volume constraints for head and neck case 2

% volume of LP receiving ≥ 26 Gy	% volume of LP receiving ≥ 32 Gy	% volume of RP receiving ≥ 26 Gy	% volume of RP receiving ≥ 32 Gy	% volume of vocal cords receiving ≥ 30 Gy	Maximum vocal cord dose (Gy)	Maximum cord dose (Gy)
66	33	66	33	50	73.5	45

Table 4

Summary of dose and dose-volume constraints for whole pelvis/prostate case

Whole Pelvis				
% volume of bladder receiving ≥ 36 Gy	% volume of bladder receiving ≥ 42 Gy	% volume of rectum receiving ≥ 33 Gy	% volume of rectum receiving ≥ 40 Gy	% volume of bowel receiving ≥ 15 Gy
50	25	50	25	30
Prostate				
% volume of bladder receiving ≥ 29 Gy	% volume of bladder receiving ≥ 33 Gy	% volume of rectum receiving ≥ 27 Gy	% volume of rectum receiving ≥ 30 Gy	
50	25	50	25	

Table 5

Summary of number of apertures, monitor units and computation times for three cases

	Head and neck 1			Head and neck 2			Whole pelvis/prostate		
	# Aper.	MUs	CPU(min.)	# Aper.	MUs	CPU(min.)	# Aper.	MUs	CPU(min.)
Elekta									
Conventional	159	1050	17	357	2368	38	85	1195	37
DMPO	133	821	20	339	1979	40	82	1114	27
SWO	159	1095	52	357	2494	98	85	1277	62
SLP	113	1187	44	274	2083	78	67	1359	71
Varian									
Conventional	141	701	21	172	1213	41	135	1550	39
DMPO	141	783	36	168	1546	50	125	1646	39
SWO	141	804	69	172	1265	110	135	1778	65
SLP	107	829	47	143	1250	106	118	1512	84