

Genetic algorithm based deliverable segments optimization for static intensity-modulated radiotherapy

Yongjie Li^{1,2}, Jonathan Yao² and Dezhong Yao¹

¹ School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, People's Republic of China

² Topslane Inc., Pleasant Hill, CA 94523, USA

E-mail: dyao@uestc.edu.cn

Received 28 March 2003, in final form 28 August 2003

Published 30 September 2003

Online at stacks.iop.org/PMB/48/3353

Abstract

The static delivery technique (also called step-and-shoot technique) has been widely used in intensity-modulated radiotherapy (IMRT) because of the simple delivery and easy quality assurance. Conventional static IMRT consists of two steps: first to calculate the intensity-modulated beam profiles using an inverse planning algorithm, and then to translate these profiles into a series of uniform segments using a leaf-sequencing tool. In order to simplify the procedure and shorten the treatment time of the static mode, an efficient technique, called genetic algorithm based deliverable segments optimization (GADSO), is developed in our work, which combines these two steps into one. Taking the pre-defined beams and the total number of segments per treatment as input, the number of segments for each beam, the segment shapes and weights are determined automatically. A group of interim modulated beam profiles quickly calculated using a conjugate gradient (CG) method are used to determine the segment number for each beam and to initialize segment shapes. A modified genetic algorithm based on a two-dimensional binary coding scheme is used to optimize the segment shapes, and a CG method is used to optimize the segment weights. The physical characters of a multileaf collimator, such as the leaves interdigitation limitation and leaves maximum over-travel distance, are incorporated into the optimization. The algorithm is applied to some examples and the results demonstrate that GADSO is able to produce highly conformal dose distributions using 20–30 deliverable segments per treatment within a clinically acceptable computation time.

1. Introduction

Intensity-modulated radiotherapy (IMRT) has received increasing attention because of its advantage of producing highly three-dimensionally conformal dose distributions to the target, while minimizing the dose to organs at risk (OARs) and normal tissues. Compared with the traditional three-dimensional conformal radiation therapy (3D-CRT) that conforms only two dimensionally to the target projection in the plane perpendicular to the beam orientation, IMRT can significantly improve the outcome of radiation therapy. A lot of effort has been taken to make IMRT implementation easy through increasing the automatization of the beam setup, shortening the optimization time of inverse planning, enhancing the ability of dose verification and so on (Webb 2000). Some commercial treatment planning systems, such as PeacockTM and Corvus (Nomos Corporation, Sewickley, PA, USA), Helios (Varian Oncology Systems, Palo Alto, CA, USA), ARTP (Topslane Inc., Pleasant Hill, CA, USA), have provided IMRT tools for routine clinical treatment.

Traditionally, IMRT treatment planning considers beam profile optimization and beam delivery as two separate independent processes. After a time-consuming optimization of modulated beam profiles under the constraint of an objective function, a leaf-sequencing tool is followed to translate these beam profiles into deliverable structures if a multileaf collimator (MLC) is used for delivery. A series of static segments with uniform intensity but different shapes will be generated if the static delivery mode is adopted. Currently, the static delivery mode has been widely used because of the advantages of simple delivery and easy quality assurance, compared with the sliding window mode.

However, a drawback of the static mode is the prolonged treatment time, because large numbers of segments will be needed if the modulated beam profiles have multiple maxima and minima. In order to overcome this, many efforts have been made (Bortfeld *et al* 1994, Xia and Verhey 1998, Siochi 1999, Que 1999, Crooks *et al* 2002). However, there is not a single algorithm that can be most efficient for all treatments and intensity maps (Que 1999, Crooks *et al* 2002), the reduction of treatment time of the static mode is still not satisfying. A general rule is that the number of segments increases approximately linearly as the number of intensity levels increases, and the number of segments required for each beam is two or three times the number of levels in the intensity map (Xia and Verhey 1998, Que 1999). The large number of segments results in a large monitor unit (MU) to cGy ratio and significant uncertainties in leakage, head scatter and tongue-and-groove effects (Que 1999, Shepard *et al* 2002).

Another weakness of this two-step process is, if some segments with very small enclosed shape and low monitor unit (MU) are neglected in order to shorten the treatment time and decrease the effect of accumulated scatter and tongue-and-groove, there are always differences existing between the optimal profiles and the delivered profiles. The differences will be enlarged up to 5.7% when the physical limitations of MLC, such as leaf interdigitation and maximum over-travel distance, must be taken into account for some types of MLC (Seco *et al* 2002). These differences may degrade the conformity of the dose distribution of the target. The IMRT planning has to be repeated if the differences are large enough.

In order to improve the agreement between the optimal and deliverable profiles, several efforts have been made. Webb *et al* (1998, 2001), Alber and Nusslin (2000) and Spirou *et al* (2001) have tried to generate smoother profiles using different techniques. Their results show that a smoother beam profile needs fewer deliverable segments and can decrease the difference between the optimal and deliverable profiles. But these profile-smoothing techniques may potentially degrade the conformity of dose distribution, the reason being the presence of the modulation which yields the conformity, and in most cases, smoothing the profiles will

slightly or severely degrade the modulation of profiles and then degrade the conformity of dose distribution (Spirou and Chui 1998, Spirou *et al* 2001, Webb *et al* 1998, 2001). Cho and Marks (2000), Alber and Nusslin (2001), Siebers *et al* (2002) and Seco *et al* (2002) incorporated the physical limitations of MLC into the optimization process, through which beam profiles with higher deliverability were generated.

In order to overcome the key shortcoming of prolonged treatment time of the static IMRT, another type of research work concentrating on direct optimization of the static segments has also been performed. As early as 1991, Webb introduced the concept of two-weights-per-field when he studied the 3D conformal treatment planning for radiation fields defined by a MLC. He used two weights per field to spatially modulate the intensity across the field when both the target volume and volume containing OARs were in the line of sight. The two weights were optimized using a simulated annealing algorithm.

Woudstra and Storchi (1999) developed a technique for simultaneous optimization of segmented intensity modulated beams and beam orientations. They defined three simple segments for each of 36 beam orientations with the help of beam's eye view (BEV) of volumes. Then these 108 segment weights were optimized and a biological score was used to determine which beams with the three segments survived. Tervo and Kolmonen (2000) incorporated the delivery constraints to directly optimize MLC leaf positions and delivery times for static IMRT using an iterative Cimmino algorithm. Bednarz *et al* (2002) used several simple rules to define some deliverable segments, normally 70 to 111 segments per case, and the mixed-integer programming (MIP) method was used to optimize the weights of the pre-defined segments. Shepard *et al* (2002) developed a turnkey solution for static IMRT called direct aperture optimization (DAO). In DAO the leaf positions and aperture (segment) weights were optimized simultaneously using a simulated annealing algorithm. DAO took the number of apertures per beam as input. The MLC physical constraints were incorporated in the optimization process. Their results demonstrated that DAO could produce highly conformal step-and-shoot treatment plans using only three to five apertures per beam direction.

The genetic algorithm (GA), as a powerful global optimization approach, has been introduced to solve the radiation optimization problems. Ezzell (1996) used the GA to find an optimal combination of external beams. Langer *et al* (1996) used the GA to optimize the beam weights for treatment of abdominal tumours. Yu and Schell (1996) chose the GA for the optimization of prostate implants. Wu and Zhu (2000) proposed a technique to optimize the beam directions and weights using a mixed-encoding GA. Zhang *et al* (2001) used a hybrid algorithm called guided evolutionary simulated annealing (GESA), which combined the GA and SA together, to optimize the gamma knife treatment planning. Despite the drawback of extensive computation similar to simulated annealing, the potential of the GA is very attractive.

In this paper, a new efficient technique, called genetic algorithm based deliverable segments optimization (GADSO), is developed to combine the two steps of static IMRT into a single step. GADSO is concerned with the simultaneous determination of (1) the number of segments per beam, (2) the weights and (3) the shapes of the deliverable segments within a clinically acceptable computation time. In GADSO a group of interim modulated beam profiles quickly calculated using a conjugate gradient (CG) method are used to determine the number of segments for each beam and to initialize segment shapes based on the leaf-sequencing algorithm of Xia and Verhey (1998). Taking the pre-defined beam setup and the total number of segments per case as input, the segment weights and shapes are then optimized iteratively. Here we use a two-dimensional binary coding GA to optimize the segment shapes. For each group of segments with new shapes, their weights are optimized using the CG method. Also, the physical constraints of MLC are incorporated into the optimization. In order to test the proposed technique, some examples are used, and the results demonstrate that GADSO is

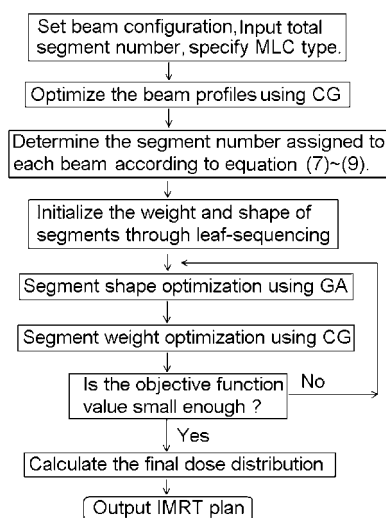


Figure 1. The flow diagram of the GADSO algorithm.

able to produce a highly conformal dose distribution using 20 to 30 deliverable segments per case within a clinically acceptable optimization time.

2. Materials and methods

2.1. General ideas

Considering the complex relations between beam profiles and dose distribution, we think it is not always a good idea to assign equal numbers of segments for each beam, or pre-define some segments for each beam according to some simple rules. Here we will use a group of interim modulated beam profiles (1) to determine the number of segments for each beam according to the complexity of profiles and (2) to initialize segment shapes by applying a leaf-sequencing operation to the profiles. The optimization of these interim modulated profiles will be quickly implemented using a CG method. In order to simplify the inverse problem and make the optimization computationally feasible for clinical practice, the segment weights and shapes will be treated as two separate groups of variables and they are optimized iteratively using CG and GA. For the convenience of computation, each segment will be divided into grids and coded using a two-dimensional (2D) binary scheme. Also, an immunity operation will be coupled into GA in order to improve the optimization performance. Figure 1 shows the flow diagram of the presented GADSO algorithm.

2.2. Determination of initial segments

In this section two issues are determined: (1) the number of segments assigned to each beam and (2) the initial segment shapes. In the previous works of other researchers, the numbers of segments for each beam were either set to be equal, or determined according to some simple rules. The segment shapes were initialized either to match the BEV of the target, or to match one part of the BEV of the target and (or) OARs according to the overlapping status of BEVs of volumes. Here a novel idea is presented. Both of them will be determined according to an interim intensity profile of each beam. The interim beam profiles are optimized using a

CG method within a very short time. Then a profile complexity factor (PCF) is defined to evaluate the complexity of the beam profile and to determine the number of segments of each beam. Then a segment quality factor (SQF) is used to determine which segments will be left to participate in the optimization.

2.2.1. Optimization of interim beam profiles. The optimization of modulated beam profiles has been studied by many researchers and a variety of algorithms have been published. These algorithms basically fall into two classes: (1) deterministic algorithm, such as gradient-based algorithm (including steepest-descent, conjugate gradient and so on), maximum likelihood method and linear programming, and (2) stochastic algorithm, such as simulated annealing (SA) and GA. The stochastic algorithms have the powerful global searching ability, but the drawback of excessive computation time limits their clinical application. Gradient-based algorithms are fast, but have the potential to get trapped in local minima (Deasy 1997). Here we will use a gradient-based algorithm and the local minima problem will be ignored, for the beam profiles are just used to initialize the segments and a global algorithm will be followed for the refined optimization.

The gradient-based algorithm used here is a conjugate gradient (CG) method similar to that adopted by Spirou and Chui (1998), which can be written as

$$\vec{x}^{(k+1)}(\lambda) = \vec{x}^{(k)} + \lambda \cdot \vec{h}^{(k+1)} \quad (1)$$

$$\vec{h}^{(k+1)} = -\nabla F_{\text{obj}}(\vec{x}^{(k+1)}) + \frac{[\nabla F_{\text{obj}}(\vec{x}^{(k+1)}) - \nabla F_{\text{obj}}(\vec{x}^{(k)})] \cdot \nabla F_{\text{obj}}(\vec{x}^{(k+1)})}{\nabla F_{\text{obj}}(\vec{x}^{(k)}) \cdot \nabla F_{\text{obj}}(\vec{x}^{(k)})} \cdot \vec{h}_k \quad (2)$$

where k means the k th iteration, $\vec{x}^{(k)}$ is the beam profile vector of the k th iteration, $\nabla F_{\text{obj}}(\vec{x}^{(k)})$ is the gradient of the objective function at the point $\vec{x}^{(k)}$, λ is the step size of the k th iteration, which can be easily calculated using a one-dimensional linear optimization. $F_{\text{obj}}(\vec{x}^{(k)})$ is the objective function defined as

$$F_{\text{obj}}(\vec{x}^{(k)}) = \alpha \cdot F_{\text{OAR}}(\vec{x}^{(k)}) + \beta \cdot F_{\text{PTV}}(\vec{x}^{(k)}) \quad (3)$$

$$F_{\text{OAR}}(\vec{x}^{(k)}) = \sum_{i=1}^{N_{\text{OAR}}} \sum_{j=1}^{NT_i} \delta_j \cdot w_j \cdot (d_j(\vec{x}^{(k)}) - p_j)^2 \quad (4)$$

$$F_{\text{PTV}}(\vec{x}^{(k)}) = \gamma \cdot \sum_{j=1}^{NT_{\text{PTV}}} \delta_j \cdot w_j \cdot (d_j(\vec{x}^{(k)}) - p_j)^2 - \eta \cdot \sum_{j=1}^{NT_{\text{PTV}}} \delta_j \cdot w_j \cdot \left(d_j(\vec{x}^{(k)}) - p_j - d_j(\vec{x}^{(k)}) \cdot \log \left(\frac{d_j(\vec{x}^{(k)})}{p_j} \right) \right) \quad (5)$$

$$d_j(\vec{x}^{(k)}) = \sum_{m=1}^{N_{\text{ray}}} a_{jm} \cdot \vec{x}_m^{(k)} \quad (6)$$

where $F_{\text{OAR}}(\vec{x}^{(k)})$ is the term associated with all OARs, $F_{\text{PTV}}(\vec{x}^{(k)})$ is the term associated with the target, N_{OAR} is the total number of OARs, NT_i is the point number in the i th OAR, NT_{PTV} is the point number in the target, $\delta_j = 1$ when the j th point dose in the volume breaks the constraints, else $\delta_j = 0$, w_j is the weight of the j th point, d_j is the calculated dose of the j th point in the volume, p_j is the prescribed dose of the j th point in the volume, α , β are the regularizing factors that control the importance between target and OARs, γ , η are the importance factors of the first and second terms in equation (5), N_{ray} is the total ray (also

called pencil beam) number, a_{jm} is the dose deposited on the j th point from a unit weight of the m th ray, $\bar{x}_m^{(k)}$ is the intensity of the m th ray.

The second term in equation (5) is derived from the work of Wu and Zhu (2002) and used to measure the homogeneity of dose distribution in the target through the entropy information. According to information theory, maximizing entropy is equivalent to minimizing information in a system, i.e. to maximize the homogeneity of dose distribution.

In our objective function both dose and dose–volume constraints are embedded. A dose constraint for the target is expressed as ‘all doses should be higher than $D_{D_{\min}}$ and lower than $D_{D_{\max}}$ ’, and a dose–volume constraint for the target is expressed as ‘no more than $V_{\min}\%$ of the target volume can absorb doses lower than $D_{DV_{\min}}$ ’. For the OARs, a dose constraint is denoted as ‘all doses should be lower than $D_{D_{\max}}$ ’, and a dose–volume constraint is denoted as ‘no more than $V_{\max}\%$ of the volume can absorb doses higher than $D_{DV_{\max}}$ ’. The dose–volume constraints are realized through the penalty of those points breaking $V_{\max}\%$ or $V_{\min}\%$ after sorting the doses in ascending or descending order.

2.2.2. Determination of the number of segments for each beam. A profile complexity factor (PCF) is carefully designed by us to determine how many segments should be assigned to each beam. The profile complexity factor of the i th beam, PCF_i , is given by

$$PCF_i = \frac{SN_i}{\sum_{k=1}^{N_{\text{beam}}} SN_k} \cdot \frac{\sum_{m=1}^{SN_i} MU_{im}}{MUU_i} \quad (7)$$

where SN_i is the total number of segments of the i th beam after leaf sequencing, N_{beam} is the total number of beams, MU_{im} is the number of monitor units of the m th segment of the i th beam, MUU_i is the maximum non-normalized intensity value of the i th beam profile. The first term in equation (7) represents the fraction of the segment number of the i th beam profile to the total segment number of all the N_{beam} profiles. The second term in equation (7) means the inverse of the MU efficiency (Webb *et al* 1998) (see details in section 3.1). From equation (7) we can see that a beam profile with more trial segments and less MU efficiency corresponds to a more complex profile, and then a higher PCF is obtained.

The leaf-sequencing algorithm is based on the works of Xia and Verhey (1998). Their algorithm, aiming at minimizing the number of segments for a given intensity modulated beam and maximizing the irradiation area for each segment, has been proved to be most frequently the algorithm that needs the fewest MLC segments (Que 1999, Crooks *et al* 2002).

In order to adequately reflect on the complexity of beam profiles, a high number of different intensity levels (20 in this paper) is used for leaf sequencing, the results from which are called ‘trial segments’. After PCF of all profiles is calculated, the remaining number of segments for the i th beam, SRN_i , is calculated by

$$SRN_i = NS_{\text{total}} \cdot \frac{PCF_i}{\sum_{k=1}^{N_{\text{beam}}} PCF_k} \quad (8)$$

where NS_{total} is the total number of segments per case, a parameter specified by the planner. We think that the profile with more complexity needs more trial segments. From equations (7) and (8) we can see that the beam profile with more trial segments has higher PCF, and then a higher SRN is obtained, hence more remaining segments will be assigned to this beam.

2.2.3. Initialization of segment shapes. A parameter, called the segment quality factor (SQF), is defined by us to determine which segment will be selected to participate in the optimization. The segment quality factor of the i th segment, SQF_i , is defined as

$$SQF_i = \alpha_S \cdot \frac{S_i}{S_{\max}} + \alpha_W \cdot \frac{W_i}{W_{\max}} \quad (9)$$

where S_i is the area enclosed by the i th segment, S_{\max} is the maximum area of all segments of this beam, W_i is the weight of the i th segment, W_{\max} is the maximum weight of all segments of this beam, α_S and α_W are the weighting factor of the segment area and the weight contributing to the SQF, respectively. We use SQF to evaluate the quality of a segment, that is, whether this segment is suitable for treatment with MLC. Considering the physical characters of MLC, such as transmission and penumbra, we think a segment with larger enclosed area and greater weight (MU) is a better one.

With α_S and α_W in equation (9), we can adjust the importance of segment area and weight in the segment quality factor (SQF). For most MLCs, a significant error may be introduced when too many segments with small shape are delivered because of the transmission and penumbra of MLC (Shepard *et al* 2002). Also, segments with small MUs are not encouraged because (1) round-off errors can become significant, and (2) a medical linac's dose delivery accuracy may be less compared to larger MU segments (Que 1999). So the determination of α_S and α_W is mainly according to clinical limitations. If segment MU has greater influence on the delivery accuracy for a specified linac, we can set α_S a little higher than α_W , for example, set α_S to be 0.8 and α_W 0.2. Similar consideration is suitable for the effect of segment area.

After SRN is calculated for each beam, a leaf-sequencing operation will be done iteratively for each beam profile. The intensity level varies from lower to higher, until the number of segments obtained from leaf sequencing is equal to or greater than the assigned number of segments for each beam. In most cases, the number of segments from the leaf-sequencer is not always exactly equal to the specified number. It is all right if they are equal; if not, more segments are needed and the redundant segments with little SQF will be discarded after being sorted according to SQF.

During the leaf sequencing, MLC physical constraints, such as maximum leaf over-travel distance, the leaf interdigitation limitation and minimum leaf gap, are applied and no undeliverable segment is allowed.

2.3. Iterative optimization of segment weights and shapes

Considering the complex relations among segment weights and shapes and dose distributions, the segment weights and shapes are treated as two separate groups of variables and they are optimized iteratively in order to simplify the inverse problem and make the optimization computationally feasible for clinical practice. This idea is similar to that of Pugachev *et al* (2000), who treated the optimization of beam orientations and beam profiles separately because of the huge search space when these two types of coupled variables are involved in the optimization. Analogously, the segment weights and shapes are coupled with each other, requiring segment weight optimization for every new segment shape. Here we use a modified GA to optimize the segment shapes. For each group of segments with new shapes, their weights are optimized using the CG method.

2.3.1. Segment shape optimization using the GA. GA is a global optimization technique that simulates the natural process of evolution in which the fittest solutions survive after numerous generations of natural selection and genetics (Wu and Zhu 2000). Usually, an unknown variable in a solution is represented as genes in a chromosome, and a solution with many variables is represented as a chromosome (also named individual). This process is called coding. Fitness is defined for each individual to evaluate its goodness. During the optimization process, GA keeps a group of candidate individuals (solutions) in each generation. Through genetic operations, such as selection, crossover and mutation, those individuals with better

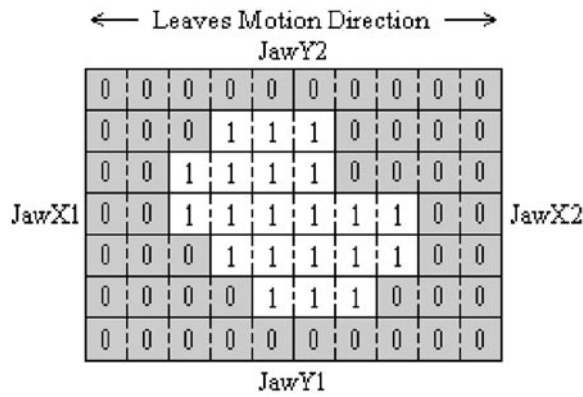


Figure 2. An illustration of the 2D binary coding scheme.

fitness will be propagated to the next generation with higher probability. After numerous generations of genetics, the individual with the best fitness among the last population is regarded as the optimal solution.

Normally, the GA adopts a binary coding scheme to represent solutions as chromosomes in a generation because this scheme has been well developed and is convenient for some applications. Here we use a 2D binary chromosome to encode the solution instead of the popular 1D mode. The 2D binary coding scheme is shown in figure 2. Along the direction of motion of the leaves the area enclosed by collimation jaws is divided into small grids. A grid covered by MLC is represented as ‘0’, else as ‘1’. It will be shown that this 2D binary coding scheme is simple and very convenient for the optimization of segment shapes, and no decoding process is needed for the computation of individual fitness.

Our GA optimization begins with NP_{size} groups of initial individuals (here an ‘individual’ is a group of segments), and in each group there are NS_{total} segments. NP_{size} is the population size. The quality of each individual is evaluated by a fitness value, and the purpose of optimization is to find the individual (i.e. a group of segments) with maximum fitness. The fitness value is calculated by

$$\text{Fitness}(\vec{s}) = F_{\max} - F_{\text{obj}}(\vec{s}) \quad \vec{s} = (s_1, s_2, \dots, s_{NS_{\text{total}}}) \quad (10)$$

where F_{\max} is a rough estimation of the maximum value of the objective function, \vec{s} is a group of segments whose weights and shapes need to be optimized. Both F_{\max} and $F_{\text{obj}}(\vec{s})$ are calculated using equations (3)–(6). Simply, F_{\max} can be determined by setting a value that is a little greater than the maximum objective value among all the groups of trial segments, which makes sure that all the fitness values are positive, a requirement by the selection operation mentioned below. It should be noted that the value of F_{\max} has no effect on the GA optimization, because there is a minus sign before $F_{\text{obj}}(\vec{s})$ in equation (10) and the minimum of $F_{\text{obj}}(\vec{s})$ always corresponds to the maximum of Fitness_i . Here N_{ray} in equation (6) is replaced by NS_{total} , and a_{jm} means the dose deposited on the j th point from a unit weight of the m th segment, $\vec{x}_m^{(k)}$ is the weight (i.e. MU) of the m th segment.

To any two randomly selected parent individuals (segment groups), a crossover operation will be applied to the two corresponding segments of these two groups according to a specified crossover probability. Then a mutation operation on the two children segments will be done according to a mutation probability.

A crossover operation is illustrated in figure 3. The crossover positions X_c and Y_c are selected randomly, and two of the four divided parts, such as the parts A1 and B1 of ‘Parent 1’,

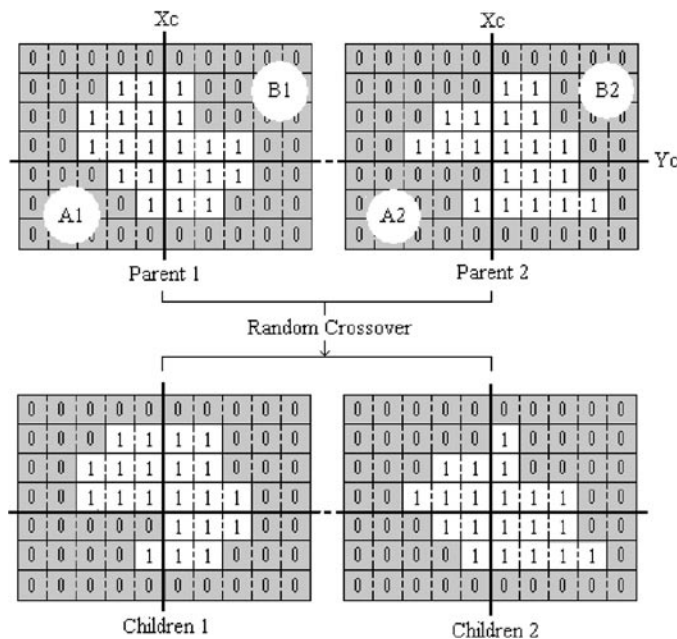


Figure 3. An illustration of the crossover operation. Parts A1 and B1 are randomly selected from the four parts. Part A1 is exchanged with part A2, and part B1 is exchanged with part B2.

are then randomly selected. Then the selected parts A1 and B1 are exchanged with the corresponding parts A2 and B2, respectively. This crossover operation is applied to each pair of selected parents, and two children will be produced for each pair of parents. The two new children will be selected into the new generation if their fitness values are greater than those of their parents, else their parents will be reserved.

If the mutation can be done according to the mutation probability calculated using equation (13), this operation will be applied to a randomly selected leaf of the two children individuals (figure 4). Then a choice is randomly made to determine whether this leaf will be closed or open. The grid A will be set to be '1' for being open (figure 4(a)), or grid B will be set to be '0' for being closed (figure 4(b)).

2.3.2. Selection of GA parameters. The selection of parameters in the GA, such as population size, crossover probability and mutation probability, is an important issue for the optimization. Population size determines how many individuals (groups of segments) are in one generation. A small size makes GA only search a small part of the solution space and has the potential for premature convergence. On the other hand, a large size will make the genetic evolution very slow and the computation time will be unacceptable. Also, crossover probability and mutation probability can influence the evolution. A higher crossover probability can make the convergence quicker, but one too high may potentially lead the GA to be premature. Mutation probability is mainly determined according to the population size and the length of chromosome. Higher population size and longer chromosome normally need lower mutation probability. Too high a mutation probability has the potential to make GA evolution unstable.

The determination of population size has been theoretically studied (Goldberg 1989), but all three of these parameters are mostly empirically selected in the engineering application (Yu *et al* 1997, Wu and Zhu 2000). Normally, crossover probability should be high, about 0.8

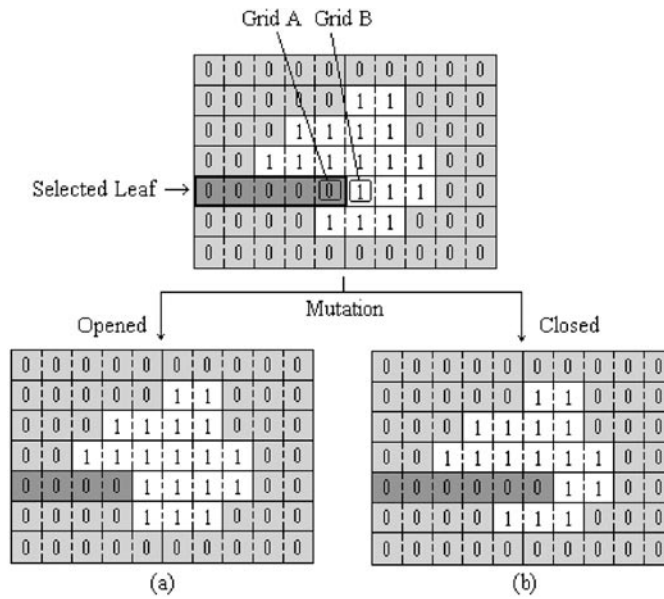


Figure 4. An illustration of mutation operation. A leaf is randomly selected and this leaf will be randomly determined to be open (a) or closed (b).

to 0.97, and mutation probability should be very low, about 0.005 to 0.02 (Goldberg 1989). As for the population size, Wu and Zhu (2000) doubled the number of variables for a moderately sized problem. We adopt this idea and find it suitable for our GADSO.

2.3.3. Fitness scaling. In order to enhance the performance of the genetic process, the following formulae are adopted to scale the fitness of each individual

$$\text{Fitness}_i = \frac{\exp(\text{Fitness}_i / T_k)}{\sum_{i=1}^{N_{Psize}} \exp(\text{Fitness}_i / T_k)} \tag{11}$$

$$T_{k+1} = T_0 \cdot \lambda^k \tag{12}$$

where Fitness_i is the fitness of the i th individual in the k th generation, λ is a factor between (0, 1), normally 0.90 to 0.99, T_0 is the parameter called initial temperature. The method in equation (11) is also called Boltzmann fitness scaling (de la Maza and Todor 1993). The idea of equation (12) is from the simulated annealing algorithm (Szu and Hartley 1987). Fitness scaling aims to overcome premature convergence at an earlier state of the evolution, or the stalling status at a later state of the evolution.

Premature convergence is a problem often discovered in the GA with small population (Goldberg 1989). At the start of a small population GA search it is common to have a few extraordinary individuals with very high fitness in a population of mediocre individuals. These extraordinary individuals would take over a significant proportion of the population in a single generation. This would in turn lead to premature convergence and the optimal solution may not be found. The opposite phenomenon is apparent at a later stage of genetic progress. The average fitness of the population becomes very close to the maximum fitness of the population. This leads to the best members getting the same number of copies for reproduction as the mediocre members, which then stalls the genetic evolution.

With equations (11) and (12), the few extraordinary chromosomes are scaled down while the lowly weighted members of the population get scaled up. Also the difference of fitness values would be scaled up at the later stage because of the decreased temperature, which could make those best individuals more outstanding.

2.3.4. Incorporation of MLC constraints and immunity operation. After both crossover and mutation operation, the MLC physical constraints, such as maximum over-travel distance, leaves interdigitation limitation and minimum leaves gap, will be checked, and any operation making the new segment break the constraints will be discarded and the operation will be repeated until a deliverable segment is generated. For example, Siemens MLC does not allow leaf interdigitation and a minimum leaf gap of 0.2 cm is needed. Then incorporation of MLC constraints into optimization is a necessary measure.

Also, an immunity operation after mutation is applied in order to enhance the deliverable performance of segments. We think that a segment with greater weight (MU) and more regular shape is more suitable for delivery. These immunity operations include: (1) a segment with very small enclosed area being replaced by one of its parent with higher fitness, and (2) a segment with very irregular shape having a higher probability to become regular. This 'higher probability' will be realized in the crossover and mutation operation of the next generation.

2.3.5. Optimization of segment weights using CG. After each loop of optimization of segment shapes is finished, a segment weights optimization will be performed using a CG method mentioned in section 2.2.1. It should be noted that, before the segment weights and shapes optimization, a_{jm} in equation (6) needs to be recalculated and it represents the dose deposited on the j th point from a unit weight of the m th segment, and $\vec{x}_m^{(k)}$ represents the weight of the m th segment in the k th iteration.

2.4. Dose calculation

A key problem of probabilistic methods, such as the GA, is the extensive computation, which should be paid more attention for clinical use. In fact, the maximum computation time in GADSO is spent on the calculation of dose distributions, which has been carefully designed and coded.

In our study, the dose calculation is based on a pencil-beam-based three-dimensional full scatter convolution (FSC) algorithm, in which both the scatter effects and tissue heterogeneity correction are involved. Before the optimization, the doses at each point contributed by each pencil beam, called the dose deposition matrix, are calculated for all volumes. To speed up the calculation, the following measures are adopted. (1) Only the doses of those pencil beams in the BEV of the target are calculated and the MLC leaves are allowed to be open only in the target BEV, (2) the deposition matrices are indexed so that only a very small number of point doses that exceed a specified threshold value will participate in the dose calculation because these matrices are very sparse and (3) after each change in leaf position in the crossover and mutation operation, the new deposition matrix of the corresponding segment is rapidly calculated just by adding the dose contributions of the changing leaf to (when the leaf is open) or subtracting them from (when the leaf is closed) the old deposition matrix. With the optimized segments, a delivered dose calculation is performed using the non-indexed deposition matrices at the end of the whole optimization.

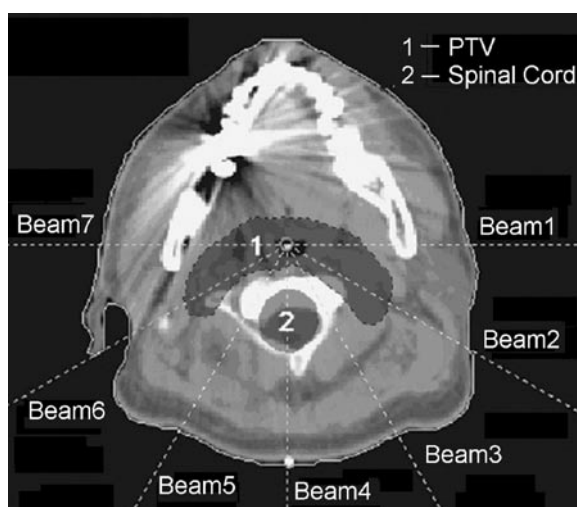


Figure 5. The simulated case of nasopharynx cancer consisting of a spinal cord (OAR) surrounded by a C-shaped planning target volume (PTV). The dotted lines indicate the central axis of the beams directed at the isocentre.

3. Results

In order to verify the presented technique, first a comparison between two-step optimization results and the GADSO results is made based on a simulated phantom case of nasopharynx cancer. Then a simulated prostate phantom case is used to (1) analyse how many segments are needed for a complex case, and (2) compare the results between the plan of equal numbers of segments per beam and different numbers of segments per beam. A Siemens MLC with 29 pairs of leaves is used for all the test cases. The leaf thickness of the MLC is 1.0 cm at the isocentre plane. The interdigitation of leaves is not allowed and the maximum over-travel distance is limited to 10.0 cm for this MLC, also a minimum gap of 0.2 cm is needed between a leaf and its opposing leaf or the leaves adjacent to its opposing leaf. All these MLC physical limitations are incorporated into the GADSO optimization. All the tests are done on a personal computer (AMD athlon(tm) MP 1800+, 1.53 GHz, 512 MB RAM).

3.1. Comparison with conventional two-step optimization results

The comparison is done based on a simulated case of nasopharynx cancer shown in figure 5, which consists of a spinal cord surrounded by a C-shaped planning target volume (PTV). Seven equispaced lateral- and posterior-oblique coplanar 6 MV photon beams are used. The prescription dose of 70 Gy to the PTV is set to 100%, and the maximum dose irradiated to the spinal cord is limited to 50 Gy. The CT slice intervals are 0.5 cm and the size of the voxel volume is set to 0.2 cm × 0.2 cm × 0.5 cm for both the spinal cord and the PTV. The pencil beam size is 0.5 cm × 0.5 cm at the isocentre plane.

First, the plan is optimized using a pencil-beam-based optimization tool embedded in the ARTP treatment planning system (TPS), which is based on a modified conjugate gradient (CG) method. The optimization is terminated when $(F_{\text{obj}}(\vec{x}^{(k)}) - F_{\text{obj}}(\vec{x}^{(k-1)})) / F_{\text{obj}}(\vec{x}^{(k)})$ is less than 0.001 after the k th iteration. After the optimization a leaf-sequencing operation is

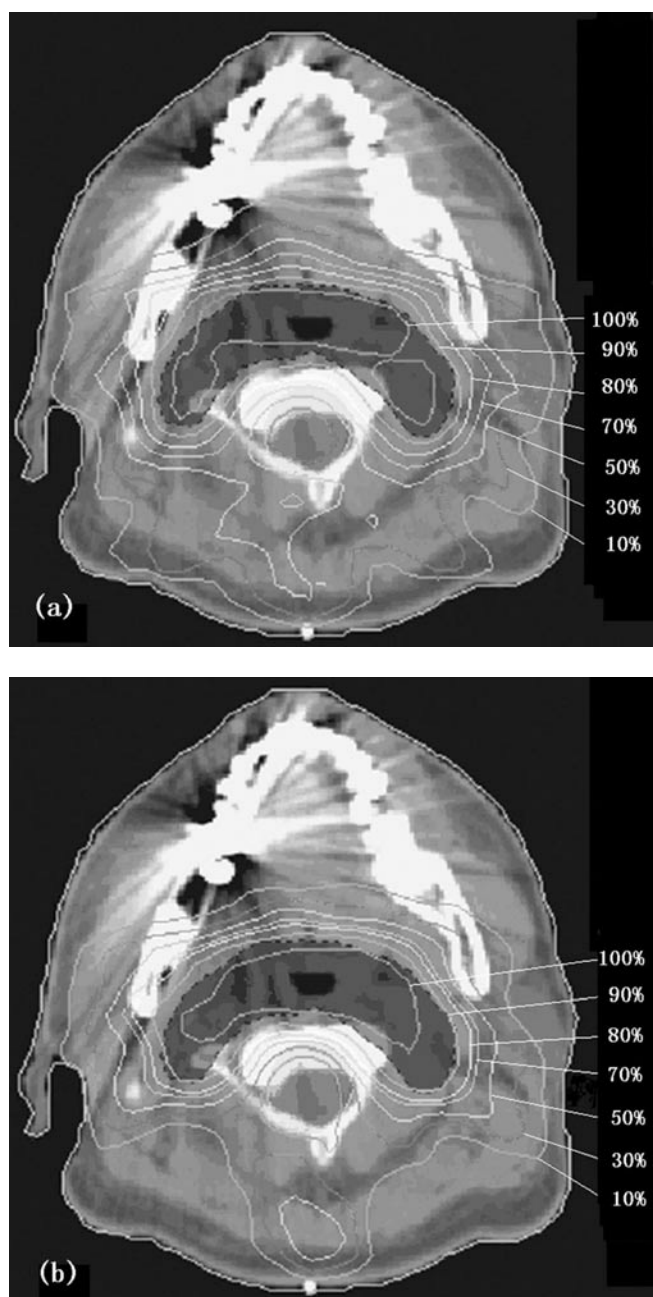


Figure 6. Deliverable dose distributions of (a) pencil-beam-based optimization and (b) GADSO.

applied to the modulated beam profiles with ten intensity levels. Then the delivered doses are calculated using these sequenced segments and the dose distribution is shown in figure 6(a).

For the convenience of comparison of pencil-beam-based optimization and GADSO, a leaf-sequencing operation with three intensity levels is applied to a copy of the above optimization result, which leads to 34 segments.

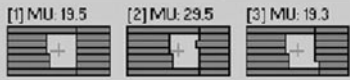
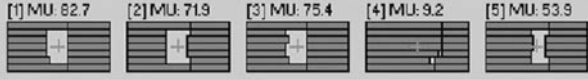



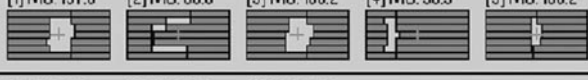

	Beam1
	Beam2
	Beam3
	Beam4
	Beam5
	Beam6
	Beam7

Figure 7. Segment list of GADSO.

The same plan is optimized using the GADSO algorithm, in which NS_{total} , the total segment number per case, is set to 34, the same as the segment number when the intensity level is 3 for the pencil-beam-based optimization. Crossover probability is 0.90, mutation probability is 0.01 and the population size NP_{size} is empirically set to be 60. For a fair comparison later, the grid size (i.e. the step size along the direction of motion of the leaves) for GADSO is 0.5 cm, the same as the pencil size of the two-step optimization. The optimization is stopped when no better solution is found in five successive generations. The final (delivered) dose distributions are shown in figure 6(b). All the segments are listed in figure 7, all of which match the MLC physical constraints.

The dose–volume histograms (DVHs) are compared in figure 8. Figure 8(a) shows the DVH comparison between GADSO and pencil-beam-based optimization with ten intensity levels, in which there is no significant difference existing between the two plans. Some detailed comparisons are listed in table 1, in which the maximum MU in profile means the maximum intensity value in a beam profile, and the MU efficiency for a beam is defined as the ratio of the maximum MU in a beam profile to the delivered number of MUs (Webb *et al* 1998). From the table we can see that the MUs needed to deliver the prescribed dose distributions are significantly reduced from 33 220 to 19 210, and of course, the delivery efficiency of GADSO is improved highly, compared with that of pencil-beam-based optimization. The optimization time (including the time spent on calculation of the dose deposition matrices and the final doses) used for pencil-beam-based optimization and GADSO are 4 min 21 s and 8 min 35 s, which are clinically acceptable.

Figure 8(b) shows the DVH comparison between GADSO and pencil-beam-based optimization with three intensity levels, from which we can clearly see that GADSO gives a better result than pencil-beam-based optimization when they use the same number of segments, 34.

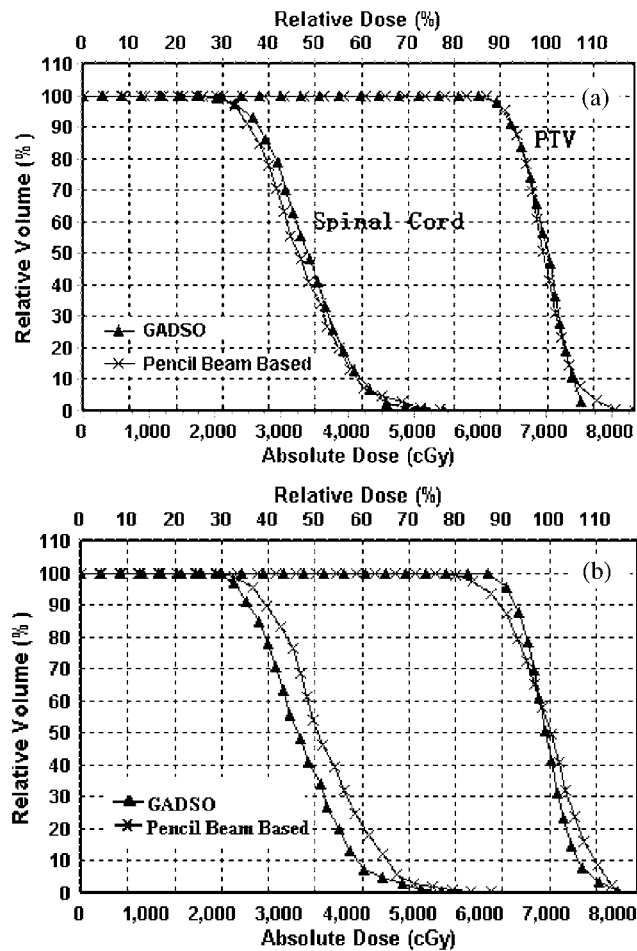


Figure 8. DVHs of GADSO and pencil-beam-based optimization when the intensity level number is 10 (a) and 3 (b). As for (b), GADSO and pencil-beam-based optimization have the same number of segments, 34.

Table 1. A comparison of the results between GADSO and pencil-beam-based optimization (PBO). The number of MUs in the table is the total MU instead of the fractional MU.

	Segment number		Beam MU		Max MU in profile		MU efficiency (%)	
	PBO	GADSO	PBO	GADSO	PBO	GADSO	PBO	GADSO
Beam 1	10	3	3 650	680	2 278	680	62.4	100.0
Beam 2	14	5	4 670	2 930	2 745	2 930	58.8	100.0
Beam 3	15	7	6 090	4 300	2 768	2 775	45.5	64.5
Beam 4	12	5	4 270	1 940	3 047	1 409	71.4	72.6
Beam 5	16	6	5 710	2 710	2 480	2 117	43.4	78.1
Beam 6	14	5	5 820	4 520	2 424	4 369	41.6	96.7
Beam 7	9	3	3 010	2 130	3 010	2 130	100.0	100.0
Total	90	34	33 220	19 210	18 752	16 410	56.4	85.4

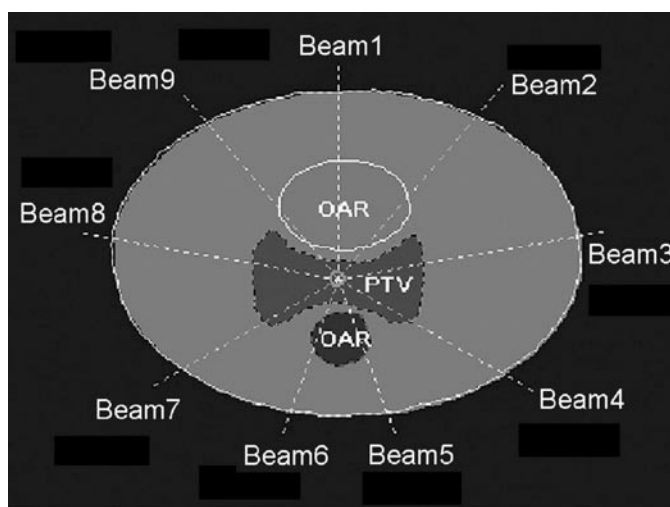


Figure 9. A simulated case that includes a prostate (PTV) with a concave outline with adjoining organs at risk (OARs), bladder and rectum.

Table 2. Prescribed constraints for the volumes in the simulated prostate case.

Volume	Dose constraints	Dose–volume constraints
PTV	Min 70 Gy, Max 75 Gy	
Bladder	Max 60 Gy	No more than 40 Gy in 50% volume
Rectum	Max 60 Gy	No more than 40 Gy in 30% volume

3.2. How many segments are needed per case?

The key purpose of this work is to find a step-and-shoot IMRT plan that can produce acceptable dose distributions using as small a number of segments as possible. Here a comparison work is done to investigate how many segments are enough for a moderately complicated case shown in figure 9. The simulated case includes a prostate (PTV) with a concave outline with adjoining organs at risk (OARs), bladder and rectum. Nine equispaced coplanar 6 MV photon beams are used to irradiate the PTV. The constraints imposed on each volume are listed in table 2. Dose prescription to the PTV is set to 73 Gy, which is normalized to 100%. The CT slice intervals are 0.5 cm, the size of the voxel volume is set to 0.3 cm × 0.3 cm × 0.5 cm for both the PTV and the rectum, and 0.5 cm × 0.5 cm × 0.5 cm for the bladder because of the larger volume than the PTV and the rectum. The pencil beam size is 0.5 cm × 0.5 cm at the isocentre plane.

The GADSO optimization is run using 25, 30, 35 and 40 total segments per case and the corresponding DVHs are shown in figure 10. The optimization times for these runs are 6 min 10 s, 8 min 25 s, 13 min 18 s and 17 min 36 s, respectively. From the DVHs we can see that the result is not clinically acceptable when the total number of segments is 25, which is such a small number that some of the nine beams only have one or two segments. But the results for 30 and 35 total segments are good enough and all the constraints are matched well for all the OARs, and have a much better PTV dose distribution than that for 25 segments. When the total number of segments increases to 40, there are no significant improvements for the DVHs except for longer computation times because of the increase of optimization

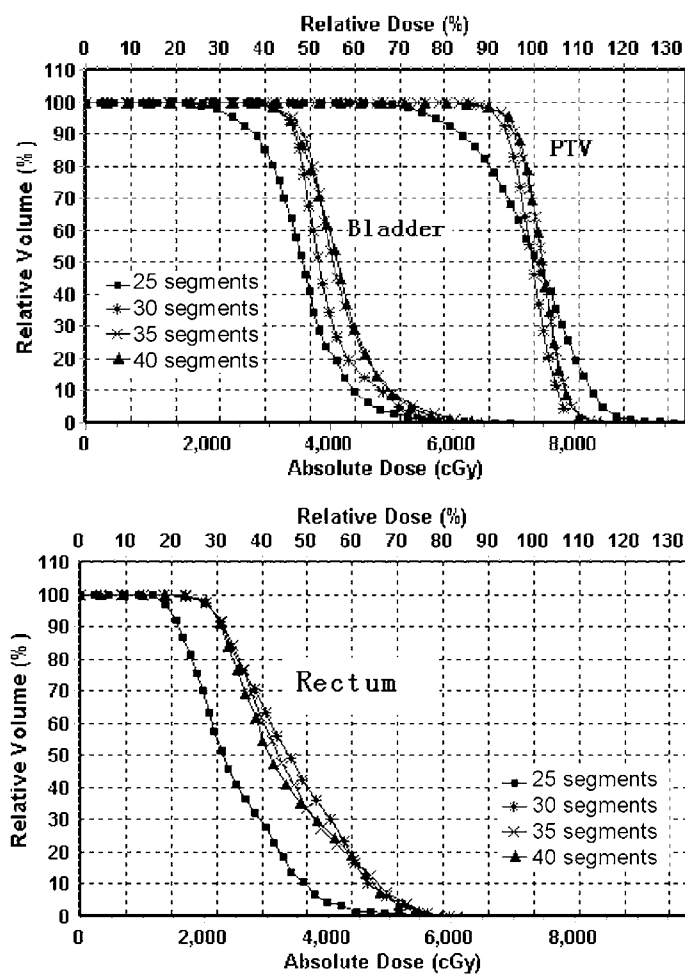


Figure 10. The DVH curves of the tests when GADSO optimization is run using 25, 30, 35 and 40 total segments per case.

variables, for example, the individual number in one generation. The dose distribution on the slice containing the isocentre is shown in figure 11 when the total number of segments is 35.

3.3. Different segment numbers per beam versus equal segment numbers per beam

For the convenience of clinical use and research work, another input mode expressed as 'segment number per beam' is also provided besides the input of total segment number per case. Here the test case described in section 3.2 is used again and two plans are optimized, i.e. plan A uses 36 as the input of the total segment number per case, and plan B takes 4 as input of the segment number per beam for this nine-beam plan. The DVHs curves are shown in figure 12.

From the DVHs we can see that a more uniform dose distribution for PTV and lower doses for OARs are obtained by different segment numbers per beam than those obtained by

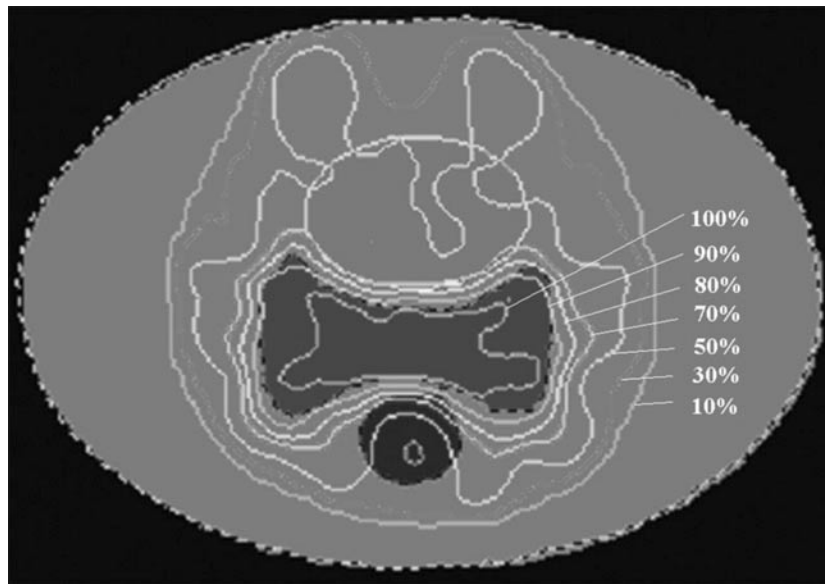


Figure 11. The dose distribution on the slice containing the isocentre when the number of total segments is 35.

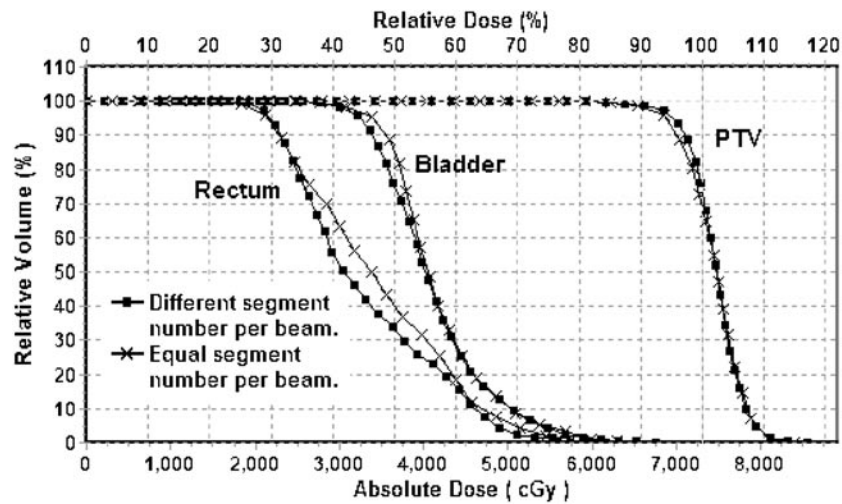


Figure 12. DVHs of the results of different segment numbers and equal segment numbers per beam.

equal segment numbers per beam. A comparison of the two plans is listed in table 3. It is clear that the segment numbers for beams 3, 4, 7 and 8 are smaller than those for other beams, the reason is that these beams need less modulation because no OAR was irradiated directly by these four beams. By assignment of suitable segment numbers to different beams according to their complexity, a smaller MU number is needed than that for equal segment numbers per beam (reduced from 23 698 to 23 483).

Table 3. A comparison between two plans that have different segment numbers per beam (plan A) and equal segment numbers per beam (plan B). The number of MUs in the table is the total MU instead of the fractional MU.

	Segment number		Beam MU	
	Plan A	Plan B	Plan A	Plan B
Beam 1	5	4	3 960	3 742
Beam 2	5	4	2 520	2 670
Beam 3	3	4	2 786	3 015
Beam 4	3	4	2 280	2 631
Beam 5	5	4	3 195	2 988
Beam 6	5	4	2 527	2 502
Beam 7	3	4	1 591	1 681
Beam 8	2	4	2 166	2 251
Beam 9	5	4	2 458	2 218
Total	36	36	23 483	23 698

4. Discussion and conclusions

Aiming to combine the two-step process of the conventional static IMRT into a single-step one, this paper presented a method called GADSO for simultaneous determination of the number of segments for each beam, the segment shapes and weights. A group of interim modulated beam profiles were used to determine the number of segments for each beam according to the complexity of the modulated beam profiles and to initialize the shapes of segments by applying a leaf-sequencing operation to the profiles. Then the segment weights and shapes were treated as two separate groups of variables and optimized iteratively using CG and GA. For the convenience of computation, each segment was divided into grids and encoded using a two-dimensional (2D) binary scheme. This coding scheme has been proved to be perspicuous and very suitable for optimization of the shapes of segments using GA, for it can provide an effective crossover and mutation operation and no decoding operation is needed for the calculation of individual fitness.

The aim of this work is mainly to shorten the treatment time by reducing the segment number without degrading the dose distribution, which is obtained by (1) the adaptively selected suitable segment number for each beam, and (2) the optimization of the weight and shape of segments. It should be pointed out that, though the selected segment number assigned to each beam may not be optimal, the final result is still superior or at least equivalent to that by assigning an equal number of segments to each beam, which has been demonstrated in the test cases.

According to our experiences, 30 to 35 segments per IMRT plan are enough for most of the complicated cases. It is a drawback of GADSO that trial-and-error work may be needed to select a suitable total number of segments for a more complicated case if the treatment time is strictly considered.

Instead of setting an equal number of segments for each beam (Webb 1991, Woudstra and Storchi 1999, Shepard *et al* 2002), or setting different numbers of segments just according to some simple rules (Woudstra and Storchi 1999, Bednarz *et al* 2002), the determination of segment number for each beam and the initialization of the weight and shape of segments are adaptively done according to an interim intensity map for each beam. A suitable segment number for each beam makes it possible to produce good dose conformity using a smaller total number of segments. A suitable initial segment shape makes it possible to implement the GA optimization using shorter iteration times.

It should be noted that the profile complexity factor (PCF) and the segment quality factor (SQF) defined in equations (7) and (9) are designed according to our experience. They act well in GADSO, but the forms of these two equations may not be the best ones. It is a difficult thing to previously determine how many segments should be used in one beam direction, because we can know if they are suitable numbers only when the dose distribution is calculated. Maybe it is a good idea to combine our strategies together with those rules experientially proposed by other researchers (Bednarz *et al* 2002).

One of the main advantages of GADSO is to generate a feasible result within a clinically acceptable computation time, even using an optimization flow based on a time-consuming GA algorithm, which is realized mainly by (1) a suitable initialization of the first generation, (2) the adaptation of an effective fitness scaling technique, (3) a carefully designed dose calculation engine and (4) optimization of segment weights using a time-saving CG method.

As for the problem of being potentially trapped in local minima by using a local search method such as CG, a lot of studies have been done on it (Deasy 1997, Borgers and Quinto 1999, Wu and Mohan 2002, Llacer *et al* 2003). One of the main conclusions is that dose-volume constraints make the feasible space possibly nonconvex and therefore possibly lead to multiple minima. Fortunately, it has been demonstrated that even when multiple minima do exist for some specially designed cases, those minima are very close to each other in cost and the resulting treatment plans are practically identical, no clear evidence having been found that multiple minima present impediments to finding a suitable IMRT solution (Wu and Mohan 2002, Llacer *et al* 2003). Even though these limited studies cannot be considered conclusive, the results do increase our confidence in the validity of gradient methods in IMRT optimization.

The test results of a simulated nasopharynx cancer prove that GADSO can produce results competing with the pencil-beam-based optimization within an acceptable computation time. The results of the second test show that 30–35 segments are enough for a moderately complicated case. Also, those plans with different segment numbers per beam can produce better results than those with equal segment numbers per beam, which was demonstrated by the third test case. All the test results show that the presented algorithm is efficient and feasible and can play a useful role in the static IMRT, even for conformal radiotherapy (CRT) when the segment number is reduced to as little as one per beam.

We think that it is a very difficult thing for GADSO to generate a dose distribution obviously better than that generated from pencil-beam-based optimization normally using more segments. Our next work is to get a better dose distribution by only increasing a very few numbers of segments based on the current GADSO algorithm. Ideally, the total segment number could be automatically selected by the computer within a clinically acceptable computation time.

Acknowledgments

This work is supported by a grant from NSFC of China, grant no 90208003, a grant from Doctor Training Fund of the Ministry of Education (MOE), China and a grant from TRAPOYT of China. This work is also partially supported by Topslane Inc. The authors are grateful to Wenyan Chen, Lin Feng and Dongmei Lu of Topslane for their helpful discussions and assistance. Also, the authors would like to thank the anonymous reviewers for their valuable comments.

References

- Alber M and Nusslin F 2000 Intensity modulated photon beams subject to a minimal surface smoothing constraint *Phys. Med. Biol.* **45** 49–52

- Alber M and Nusslin F 2001 Optimization of intensity modulated radiotherapy under constraints for static and dynamic MLC delivery *Phys. Med. Biol.* **46** 3229–39
- Bednarz G *et al* 2002 The use of mixed-integer programming for inverse treatment planning with pre-defined field segments *Phys. Med. Biol.* **47** 2235–45
- Borgers C and Quinto E T 1999 On the non-uniqueness of optimal radiation treatment plans *Inverse Problems* **15** 1115–38
- Bortfeld J, Kahler D L, Waldron T J and Boyer A L 1994 X-ray field compensation with multileaf collimators *Int. J. Radiat. Oncol. Biol. Phys.* **28** 723–30
- Cho P S and Marks R J 2000 Hardware sensitive optimization for intensity modulated radiotherapy *Phys. Med. Biol.* **45** 429–40
- Cotrutz C, Lahanas M, Kappas C and Baltas D 2001 A multi-objective gradient-based dose optimization algorithm for external beam conformal radiotherapy *Phys. Med. Biol.* **46** 2161–75
- Crooks S M, McAven L F, Robinson D F and Xing L 2002 Minimizing delivery time monitor units in static IMRT by leaf-sequencing *Phys. Med. Biol.* **47** 3105–16
- de la Maza M and Todor B 1993 *Proc. 5th Int. Conf. on Genetic Algorithms* (San Mateo, CA: Morgan Kaufmann) pp 124–31
- Deasy J O 1997 Multiple local minima in radiotherapy optimization problems with dose–volume constraints *Med. Phys.* **24** 1157–61
- Ezzell G A 1996 Genetic and geometric optimization of three-dimensional radiation therapy treatment planning *Med. Phys.* **23** 293–305
- Goldberg D E 1989 *Genetic Algorithms in Search, Optimization, and Machine Learning* (Reading, MA: Addison-Wesley)
- Hristov D, Stavrev P, Sham E and Fallone B G 2002 On the implementation of dose–volume objectives in gradient algorithms for inverse treatment planning *Med. Phys.* **29** 848–56
- Langer M *et al* 1996 A generic genetic algorithm for generating beam weights *Med. Phys.* **23** 965–71
- Leichtman G S, Aita A L and Goldman H W 2000 Automated gamma knife dose planning using polygon clipping and adaptive simulated annealing *Med. Phys.* **27** 154–62
- Lessard E and Pouliot J 2001 Inverse planning anatomy-based dose optimization for HDR-brachytherapy of the prostate using fast simulated annealing algorithm and dedicated objective function *Med. Phys.* **28** 773–8
- Llacer J 1997 Inverse radiation treatment planning using the dynamically penalized likelihood method *Med. Phys.* **24** 1751–64
- Llacer J, Deasy J O, Bortfeld T R, Solberg T D and Promberger C 2003 Absence of multiple local minima effects in intensity modulated optimization with dose–volume constraints *Phys. Med. Biol.* **48** 183–210
- Pugachev A B, Boyer A L and Xing L 2000 Beam orientation optimization in intensity-modulated radiation treatment planning *Med. Phys.* **27** 1238–45
- Que W 1999 Comparison of algorithms for multileaf collimator field segmentation *Med. Phys.* **26** 2390–6
- Rosen I I *et al* 1991 Treatment plan optimization using linear programming *Med. Phys.* **18** 141–52
- Seco J, Evans P M and Webb S 2002 An optimization algorithm that incorporates IMRT delivery constraints *Phys. Med. Biol.* **47** 899–915
- Shepard D M, Earl M A, Li X A, Naqvi S and Yu C 2002 Direct aperture optimization: a turnkey solution for step-and-shoot IMRT *Med. Phys.* **29** 1007–18
- Siebers J V, Lauterbach M, Keall P J and Mohan R 2002 Incorporating multi-leaf collimator leaf sequencing into iterative IMRT optimization *Med. Phys.* **29** 952–9
- Siochi R A C 1999 Minimizing static intensity modulation delivery time using an intensity solid paradigm *Int. J. Radiat. Oncol. Biol. Phys.* **43** 671–80
- Spirou S V and Chui C S 1998 A gradient inverse planning algorithm with dose-volume constraints *Med. Phys.* **25** 321–33
- Spirou S V, Fourniew-Bidoz N, Yang J, Chui C-S and Ling C C 2001 Smoothing intensity-modulated beam profiles to improve the efficiency of delivery *Med. Phys.* **28** 2105–12
- Szu H and Hartley R 1987 Fast simulated annealing *Phys. Lett. A* **122** 157–62
- Tervo J and Kolmonen P 2000 A model for the control of a multileaf collimator in radiation therapy treatment planning *Inverse Problems* **16** 1875–95
- Webb S 1989 Optimization of conformal radiotherapy dose distributions by simulated annealing *Phys. Med. Biol.* **34** 1349–70
- Webb S 1991 Optimization by simulated annealing of three-dimensional conformal treatment planning for radiation fields defined by a multileaf collimator *Phys. Med. Biol.* **36** 1201–26
- Webb S 2000 *Intensity-Modulated Radiation Therapy* (Bristol: Institute of Physics Publishing)
- Webb S 2001 A simple method to control aspects of fluence modulation in IMRT planning *Phys. Med. Biol.* **46** 187–95

- Webb S, Convey D J and Evans P M 1998 Inverse planning with constraints to generate smoothed intensity-modulated beams *Phys. Med. Biol.* **43** 2785–94
- Woudstra E and Storchi P R M 1999 Simultaneous optimization of segmented intensity modulated beams and beam orientations *Radiother. Oncol.* **51** S45
- Wu Q and Mohan R 2002 Multiple local minima in IMRT optimization based on dose–volume criteria *Med. Phys.* **29** 1514–27
- Wu X and Zhu Y 2000 A mixed-encoding genetic algorithm with beam constraint for conformal radiotherapy treatment planning *Med. Phys.* **27** 2508–16
- Wu X and Zhu Y 2002 A maximum-entropy method for the planning of conformal radiotherapy *Med. Phys.* **28** 2241–6
- Wu X, Zhu Y and Luo L 2000 Linear programming based on neural networks for radiotherapy treatment planning *Phys. Med. Biol.* **45** 719–28
- Xia P and Verhey L J 1998 Multileaf collimator leaf sequencing algorithm for intensity modulated beams with multiple static segments *Med. Phys.* **25** 1424–34
- Xing L *et al* 1999 Estimation theory and model parameter selection for therapeutic treatment plan optimization *Med. Phys.* **26** 2348–58
- Yu Y and Schell M C 1996 A genetic algorithm for the optimization prostate implants *Med. Phys.* **23** 2085–91
- Yu Y, Schell M C and Zhang J B Y 1997 Decision theoretic steering and genetic algorithm optimization: application to stereotatic radiosurgery treatment planning *Med. Phys.* **24** 1742–50
- Zhang P *et al* 2001 Optimization of gamma knife treatment planning via guided evolutionary simulated annealing *Med. Phys.* **28** 1746–52