

A two-stage programming approach to fluence map optimization for intensity-modulated radiation therapy treatment planning

H. Rocha¹, J.M. Dias^{1,2}, B. C. Ferreira^{3,4} and M. C. Lopes^{3,4}

¹ INESCC, Coimbra, Portugal

² FEUC, Coimbra, Portugal

³ IPOC-FG, Coimbra, Portugal

⁴ I3N, Aveiro, Portugal

Abstract— The fluence map optimization (FMO) problem is one of the most studied problems in intensity-modulated radiation therapy treatment planning. Although many approaches have shown to yield good solutions to the FMO problem, the optimal solutions obtained ensure that the resulting treatment is the best possible with respect to the weighting parameters of the formulation used. Since the ‘optimal’ weighting scheme is unknown, the choice of the weight parameters is typically a long trial-and-error process until a satisfactory solution is achieved. Moreover, for selecting the best irradiating directions, it is not clear how traditional trial-and-error parameter tuning should be incorporated or managed. A two-stage programming approach is proposed to reduce the dependency of the optimal solutions on the weight parameters and simultaneously improve the overall plan quality. This approach is yet another step towards automated generation of treatment plans which will result in breakthrough developments in radiation therapy care.

Keywords— Intensity-modulated Radiation Therapy, Fluence Map Optimization, Inverse planning.

I. INTRODUCTION

Cancer is one of the most significant health problems worldwide and was recently identified by the World Health Organization as one of the four leading threats to human health and development. Radiation therapy, along with surgery and chemotherapy, is one of the main treatment modalities for cancer, used for more than half of the cancer patients, either to cure the disease or to palliate symptoms. It is generally used as a local treatment, irradiating tissues with proven or suspected disease with ionizing radiation, based on the fact that cancerous cells are focused on fast reproduction and are not as able to repair themselves when damaged by radiation as healthy cells. Therefore, the goal of the treatment is to deliver enough radiation to kill the cancerous cells but not so much that jeopardizes the ability of the surrounding healthy cells to survive.

Typically, radiation is generated by a linear accelerator and

several beams of ionizing radiation, sent at different incidence directions, pass through the patient. The intersection of these beams is centered at the tumor attempting to sterilize all cancer cells while the surrounding healthy organs and tissues receive radiation from some but not all radiation beams and may thus be spared. An important type of radiation therapy is intensity-modulated radiation therapy (IMRT), a modern technique where the radiation beam is modulated by a multileaf collimator. Multileaf collimators (MLC) enable the transformation of the beam into a grid of smaller beamlets of independent intensities. Beamlets do not exist physically. Their existence is generated by the movement of the leaves of the MLC that block part of the beam during portions of the delivery time. The MLC has movable leaves on both sides that can be positioned at any beamlet grid boundary.

In the inverse planning of the radiation therapy treatment plan, for a prescribed plan, a correspondent set of parameters (beams and fluences) is algorithmically computed in order to fulfill the prescribed doses and restrictions. A common way to solve the IMRT optimization problems is to use a beamlet-based approach leading to a large-scale programming problem. Due to the complexity of the whole optimization problem, the treatment planning is typically divided into three smaller problems which can be solved sequentially: beam angle optimization problem, fluence map optimization problem and realization problem. The first problem consists in finding the minimum number of beams and corresponding directions that satisfy the treatment goals using optimization algorithms [1, 2]. After deciding what beam angles should be used, a patient will be treated using an optimal plan obtained by solving the fluence map (or intensity) optimization problem - the problem of determining the optimal beamlet weights for the fixed beam angles. Many mathematical optimization models and algorithms have been proposed for the intensity problem, including linear models [3, 4], mixed integer linear models [5, 6], nonlinear models [7, 8], and multiobjective models [9, 10]. After an acceptable set of fluence maps is produced, one must find a suitable way for delivery (realization problem). Typically, beamlet intensities are discretized and one

of the many existing techniques ([3, 11]) is used to construct the apertures and intensities that approximately match the intensity maps previously determined. However, plan's quality deterioration must be prevented when reproducing the optimized intensity maps [12, 13].

The objective of this paper is to present a two-stage programming approach for the fluence map optimization (FMO) problem. This new approach aims at improving the fulfilment of dose prescription limits for structures other than tumors in a first stage while proper tumor irradiation is aimed in a second stage. Another important goal that we kept in mind during the development of this approach is the reduction of the dependency of the optimal solutions on the weight parameters and a consequent improved usefulness in beam angle optimization. The benefits of this approach are discussed using a retrospective treated case of head-and-neck tumor at the Portuguese Institute of Oncology of Coimbra. The paper is organized as follows. In the next section we describe the FMO problem. Section 3 briefly presents the two-stage programming approach proposed. Section 4 presents the obtained results. In the last section we have the conclusions.

II. FLUENCE MAP OPTIMIZATION PROBLEM

For optimization purposes, radiation dose distribution deposited in the patient, measured in Gray (Gy), needs to be assessed accurately. Each structure's volume is discretized in voxels (small volume elements) and the dose is computed for each voxel using the superposition principle, i.e., considering the contribution of each beamlet. Typically, a dose matrix D is constructed from the collection of all beamlet weights, by indexing the rows of D to each voxel and the columns to each beamlet, i.e., the number of rows of matrix D equals the number of voxels (V) and the number of columns equals the number of beamlets (N) from all beam directions considered. Therefore, using matrix format, we can say that the total dose received by the voxel i is given by $\sum_{j=1}^N D_{ij}w_j$, with w_j the weight of beamlet j . Usually, the total number of voxels considered reaches the tens of thousands, thus the row dimension of the dose matrix is of that magnitude. The size of D originates large-scale problems being one of the main reasons for the difficulty of solving the FMO problem.

Most of the FMO models in the literature belong to a class of constrained optimization models such that an objective function is optimized while meeting dose requirements. A variety of criteria may be considered to be included in the objective function, leading to many different objective functions. It is beyond the scope of this study to discuss which formulation of the FMO problem is preferable. Romeijn et al. [9] demonstrated that most of the treatment plan evalu-

ation criteria proposed in the medical physics literature are equivalent to convex penalty function criteria when viewed as a multi-criteria optimization problem. Here, we will use a convex penalty function voxel-based nonlinear model [7]. The conclusions drawn regarding this particular model embedded in the proposed two-stage programming strategy are valid also if different FMO formulations are considered. In this model, each voxel is penalized according to the square difference of the amount of dose received by the voxel and the amount of dose desired/allowed for the voxel. This formulation yields a quadratic programming problem with only linear non-negativity constraints on the fluence values [4]:

$$\min_w \left[\underline{\lambda}_i \left(T_i - \sum_{j=1}^N D_{ij}w_j \right)_+^2 + \bar{\lambda}_i \left(\sum_{j=1}^N D_{ij}w_j - T_i \right)_+^2 \right]$$

$$s.t. \quad 0 \leq w_j \leq w^{max}, j = 1, \dots, N,$$

where T_i is the desired dose for voxel i , $\underline{\lambda}_i$ and $\bar{\lambda}_i$ are the penalty weights of underdose and overdose of voxel i , w^{max} is the maximum beamlet intensity allowed and $(\cdot)_+ = \max\{0, \cdot\}$. This nonlinear formulation implies that a very small amount of underdose or overdose may be accepted in clinical decision making, but larger deviations from the desired/allowed doses are decreasingly tolerated [7].

The optimal solutions obtained ensure that the resulting treatment is the best possible with respect to the weighting parameters (λ) used. Since it is impossible to attribute effective clinical meaning to the weight parameters, the 'optimal' weighting scheme is unknown and the choice of the weights is typically a long trial-and-error process until a satisfactory solution is achieved. Furthermore, for beam angle optimization it is not clear how traditional trial-and-error parameter tuning should be incorporated or managed. A two-stage programming approach is proposed next, to reduce the dependency of the optimal solutions on the weight parameters and at the same time trying to improve the overall plan quality.

III. TWO-STAGE APPROACH TO THE FLUENCE MAP OPTIMIZATION PROBLEM

Treatment plan optimization is inherently a multi-criteria process. However, typically, the FMO problem is modeled as a weighted sum function where constraints are often implemented as objectives. We propose the following two-stage programming approach, using the FMO problem formulation described in the previous Section, that can be used for most of the formulations present in the literature and/or used in clinical practice:

Initialization Choose $w^{max} > 0$. Set $T_i, i = 1, \dots, V$ according to the prescribed doses.

First stage Set $w_j^0 = w^{max}, j = 1, \dots, N$. Set $\underline{\lambda}_i = \bar{\lambda}_i = 0$ for tumors and $\underline{\lambda}_i = 0, \bar{\lambda}_i = 1$ for the remaining structures. Obtain the optimal solution of (1), w^{opt} .

Second stage Set $w_j^0 = 0, j = 1, \dots, N$. Set $\underline{\lambda}_i = \bar{\lambda}_i = 1$ for tumors and $\underline{\lambda}_i = \bar{\lambda}_i = 0$ for the remaining structures. Obtain the optimal solution of (1), for $0 \leq w_j \leq w_j^{opt}$ instead of $0 \leq w_j \leq w^{max}, j = 1, \dots, N$.

After setting the maximum beamlet intensities allowed (w^{max}), the first stage only considers structures other than tumors in an optimization loop that starts with the fluences equal to the maximum fluence allowed. The fluence intensities are then iteratively optimized (decreased) until the minimum of the first stage is (fast and easily) achieved. Zero is naturally the optimal objective function value of the first stage and the optimal fluences (w^{opt}) correspond to the maximum beamlet intensities allowed that fulfill the dose requirements for all structures but tumors. In a second stage, the fluences are optimized only for tumors considering the previously optimized fluences as upper limits.

IV. NUMERICAL TESTS AND DISCUSSION

A clinical example of a retrospective treated case of head-and-neck tumor at the Portuguese Institute of Oncology of Coimbra is used to test the proposed approach. The patient CT set and delineated structures were exported via Dicom RT to a freeware computational environment for radiation therapy research. In general, the head-and-neck region is a complex area to treat with radiation therapy due to the large number of sensitive organs in this region (e.g. eyes, mandible, larynx, oral cavity, etc.). For simplicity, in this study, the organs at risk (OARs) used for treatment optimization were limited to the spinal cord, the brainstem and the parotid glands.

The tumor to be treated plus some safety margins is called planning target volume (PTV). For the head-and-neck case in study it was separated in two parts with different prescribed doses: PTV1 and PTV2. The prescription dose for the target volumes and tolerance doses for the OARs considered in the optimization are presented in Table 1.

Our tests were performed on a 2.66Ghz Intel Core Duo PC with 3 GB RAM. In order to facilitate convenient access, visualization and analysis of patient treatment planning data, the computational tools developed within MATLAB and CERR [14] (computational environment for radiotherapy research) were used as the main software platform to embody

Table 1: Prescribed doses for all the structures considered for IMRT optimization.

Structure	Mean dose	Max dose	Presc. dose
Spinal cord	–	45 Gy	–
Brainstem	–	54 Gy	–
Left parotid	26 Gy	–	–
Right parotid	26 Gy	–	–
PTV1	–	–	70.0 Gy
PTV2	–	–	59.4 Gy
Body	–	80 Gy	–

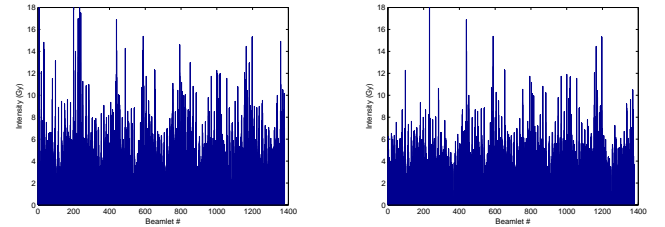


Fig. 1: Optimal fluence intensities of first stage – left and optimal fluence intensities of second stage – right.

our optimization research and provide the necessary dosimetry data to perform optimization in IMRT. The dose was computed using CERR's pencil beam algorithm (QIB). To address the convex nonlinear formulation of the FMO problem in (1) we used a trust-region-reflective algorithm (*fmincon*) of MATLAB 7.4.0 (R2007a) Optimization Toolbox.

Usually, in head-and-neck cancer cases, patients are treated with 5 to 9 equispaced beams in a coplanar arrangement. Considering 7 equispaced beams, the total number of beamlets for the tested case is 1378 and the optimal solutions of stages one and two are depicted in Fig. 1.

Typically, results are judged by depicting their cumulative dose-volume histogram (DVH) or by a variety of metrics. A metric usually used for plan evaluation is the volume of PTV that receives 95% of the prescribed dose. Typically, 95% of the PTV volume is required. The occurrence of coldspots, less than 93% of PTV volume receives the prescribed dose, and the existence of hotspots, the percentage of the PTV volume that receives more than 110% of the prescribed dose, are other measures usually used to evaluate the target coverage. Mean and/or maximum doses of OARs are usually displayed to verify organ sparing.

The results considering 7 equispaced beams for the two-stage solution, denoted *TWO*, and for the single stage solution, denoted *ONE*, are presented in Tables 2 and 3. In terms of targets coverage we can verify that both *TWO* and *ONE* treatment plans obtained satisfactory target coverage num-

Table 2: Target coverage obtained by treatment plans.

Target coverage	ONE	TWO
PTV1 at 95 % volume	67.225 Gy	67.175 Gy
PTV1 % > 93% of Rx (%)	99.650	99.625
PTV1 % > 110% of Rx (%)	0.000	0.000
PTV2 at 95 % volume	57.975 Gy	57.925 Gy
PTV2 % > 93% of Rx (%)	97.844	97.619
PTV2 % > 110% of Rx (%)	15.691	15.855

Table 3: OARs sparing obtained by treatment plans.

OAR	Mean Dose (Gy)		Max Dose (Gy)	
	ONE	TWO	ONE	TWO
Spinal cord	–	–	45.475	44.875
Brainstem	–	–	51.725	52.125
Left parotid	25.398	25.133	–	–
Right parotid	25.132	24.997	–	–

bers. However, *TWO* manage to fulfill the required dose limits for all OARs while *ONE* fail to do so.

V. CONCLUSION

A new programming approach to the FMO problem that reduces the dependency of the optimal solution on the weight parameters was proposed and tested successfully using a clinical head-and-neck cancer case. This two-stage approach guarantees organ sparing in a first stage and attempts proper tumor coverage in a second stage which can be useful also for a beam angle optimization process. Each beam angle set has guaranteed organ sparing and the best beam angle set should correspond to the one that presents better tumor coverage. This weight parameter independence might be useful as well for automated generation of treatment plans towards replacement of the current manual trial-and-error planning, which will result in breakthrough developments in radiation therapy care.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by QREN under Mais Centro (CENTRO-07-0224-FEDER-002003) and FEDER funds through the COMPETE program and Portuguese

funds through FCT under project grant PTDC/EIA-CCO/121450/2010. This work has also been partially supported by FCT under project grant PEst-OE/EEI/UI308/2014. The work of H. Rocha was supported by the European social fund and Portuguese funds from MCTES.

REFERENCES

1. Rocha H, Dias JM, Ferreira BC, Lopes MC. Beam angle optimization for intensity-modulated radiation therapy using a guided pattern search method *Phys Med Biol*. 2013;58:2939–2953.
2. Rocha H, Dias JM, Ferreira BC, Lopes MC. Selection of intensity modulated radiation therapy treatment beam directions using radial basis functions within a pattern search methods framework *J Glob Optim*. 2013;57:1065–1089.
3. Romeijn HE, Ahuja RK, Dempsey JF, Kumar A. A column generation approach to radiation therapy treatment planning using aperture modulation *SIAM J Optim*. 2005;15:838–862.
4. Romeijn HE, Ahuja RK, Dempsey JF, Kumar A, Li J. A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning *Phys Med Biol*. 2003;48:3521–3542.
5. Lee EK, Fox T, Crocker I. Integer programming applied to intensity-modulated radiation therapy treatment planning *Ann Oper Res*. 2003;119:165–181.
6. Preciado-Walters F, Langer MP, Rardin RL, Thai V. Column generation for IMRT cancer therapy optimization with implementable segments *Ann Oper Res*. 2006;148:65–79.
7. Aleman DM, Glaser D, Romeijn HE, Dempsey JF. Interior point algorithms: guaranteed optimality for fluence map optimization in IMRT *Phys Med Biol*. 2010;55:5467–5482.
8. Spirou S, Chui CS. A gradient inverse planning algorithm with dose-volume constraints *Med Phys*. 1998;25:321–333.
9. Romeijn HE, Dempsey JF, Li J. A unifying framework for multi-criteria fluence map optimization models *Phys Med Biol*. 2004;49:1991–2013.
10. Thieke C, Kufer KH, Monz M, Scherrer A, Alonso F, Oelfke U, Huber PE, Debus J, Bortfeld T. A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation *Radiother Oncol*. 2007;85:292–298.
11. Alber M, Meedt G, Nusslin F, Reemtsen R. On the degeneracy of the IMRT optimization problem *Med Phys*. 2002;29:2584–2589.
12. Rocha H, Dias JM, Ferreira BC, Lopes MC. Discretization of optimal beamlet intensities in IMRT: a binary integer programming approach *Math Comput Model*. 2012;55:1969–1980.
13. Rocha H, Dias JM, Ferreira BC, Lopes MC. Combinatorial optimization for an improved transition from fluence optimization to fluence delivery in IMRT treatment planning *Optimization*. 2012;61:969–987.
14. Deasy JO, Blanco AI, Clark VH. CERR: A Computational Environment for Radiotherapy Research *Med Phys*. 2003;30:979–985.

Author: Humberto Rocha
 Institute: INESC-Coimbra
 Street: Rua Antero de Quental 199
 City: Coimbra
 Country: Portugal
 Email: hrocha@mat.uc.pt