Future Generation Computer Systems 81 (2018) 280–290

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Application of optimization model with piecewise penalty to intensity-modulated radiation therapy

Caiping Guo [a,b,c], Pengcheng Zhang [a,b], Liyuan Zhang [a,b], Zhiguo Gui [a,b,*], Huazhong Shu [d]

[a] National Key Laboratory for Electronic Measurement Technology, North University of China, Taiyuan 030051, China
[b] Key Laboratory of Instrument Science & Dynamic Measurement Technology, North University of China, Taiyuan 030051, China
[c] Electronic Engineering Department, Taiyuan Institute of Technology, Taiyuan 030008, China
[d] Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China

HIGHLIGHTS

- New maximum-dose-based and gEUD-based quadratic sub-scores are proposed.
- New sub-scores overcome the shortcoming of semi-deviation.
- New sub-scores solve the problem of vanishing gradient in feasible solution space.
- New sub-scores used in radiotherapy planning can expand the search solution space.
- By expanding the solution space, the quality of radiotherapy plan is improved.

ARTICLE INFO

Article history:
Received 22 June 2017
Received in revised form 6 September 2017
Accepted 1 October 2017
Available online 23 October 2017

Keywords:
Quadratic model
Piecewise penalty function
Maximum dose
Generalized equivalent uniform dose

ABSTRACT

Purpose: Both maximum-dose-based and generalized equivalent uniform dose (gEUD)-based quadratic sub-scores, which penalize doses higher than the prescribed dose, exhibit the shortcomings of semi-deviation and a vanishing gradient in the feasible solution space. To address these drawbacks, this study proposes new sub-scores for the maximum dose criterion and the gEUD criterion. *Methods*: In new sub-scores, a dosage lower than the prescribed dose is assigned a linear penalty function, and one higher than the prescribed dose is assigned an extra quadratic penalty function. To test their efficiency, they were incorporated into a physical model and a hybrid physical–biological model, respectively, and were tested on a phantom TG119 and two types of clinic cases. The improved methods were compared with their original methods and the dose-volume (DV)-based optimization method. Additionally, the improved gEUD-based method was compared with another gEUD-based quadratic optimization method. The gradient-based optimization algorithm was applied to solve these large-scale optimization problems. *Results*: For similar or better PTV coverage, optimization based on our proposed quadratic models is capable of improving the OARs sparing. In practice, by using multiple DV constraints for each optimized structures, the DV based optimization may be able to arrive at similar plan, whereas greater trial-and-error is performed to adjust parameters of optimization model. Although the optimal prescribed dose remains unclear, at the same prescribed dose, our proposed optimization method can obtain better plan. *Conclusion*: Our proposed optimization method has the potential to expand the solution space and improve the quality of radiotherapy plan.

1. Introduction

The purpose of radiation treatment planning is to ensure the conformal and homogeneous irradiation of the planning target volume (PTV) while protecting normal tissue (NT) and organs at risk (OARs). This can be achieved through intensity-modulated radiation therapy (IMRT) by modulating fields from several directions of the beam [1].

The key technique and main task of IMRT is to devise an acceptable plan through the solution to an inverse problem. The process of the solution is defined as inverse planning, the key link of which is to solve intensity distributions of external beams that determine the quality of radiotherapy. Furthermore, in inverse planning, the

* Correspondence to: National Key Laboratory for Electronic Measurement Technology and Key Laboratory of Instrument Science & Dynamic Measurement Technology, College of information and Communication Engineering, North University of China, No. 3, Xueyuan Road, Taiyuan 030051, China.
E-mail address: gzgtg@163.com (Z. Gui).

https://doi.org/10.1016/j.future.2017.10.003
0167-739X/© 2017 Elsevier B.V. All rights reserved.

quality of an optimized treatment plan is affected by the optimization objective function and the optimization algorithm [2].

The objective function is an important index for the optimization and evaluation of treatment planning because it is a tool to assess a radiotherapy plan as well as the connection between ray parameters and output dose distributions. Moreover, it influences the searching ability of the optimization method. The optimized criteria applied to the objective function are either physical factors (maximum dose, minimum dose, uniform dose, and dose-volume constraint) or biological indices (tumor control probability (TCP), normal tissue complication probability (NTCP), and generalized equivalent uniform dose (gEUD)) [3]. Due to the uncertainty associated with them, the TCP and NTCP models have not been utilized in radiotherapy inverse planning systems [4]. gEUD-based optimization, however, has been investigated by several scholars [4–14] and clinically utilized in the Varian system [15]. In these studies, most optimization models, based on physical criteria or gEUD, were transformed into linear or quadratic models. Wu et al. [16] proposed some dose-based quadratic optimization models by penalizing the mean-squared deviations of the various doses or DV constraints predefined by the planner. Dirscherl et al. [12] proposed a gEUD-based quadratic model that penalized the square of the absolute difference between the actual dose and the prescribed dose. To penalize gEUD doses higher than the prescribed gEUD0, Raysearch Laboratories AB SS [15] proposed an optimization model of the relative deviation square. Mihailidis et al. [10] and Lee et al. [14] adopted gEUD-based linear optimization models penalizing relative deviation between the actual and the prescribed gEUD.

Gradient-based optimization algorithms have been utilized in commercial inverse planning systems (HELIOS and Pinnacle) due to their speed [17]. Nevertheless, when gradient algorithms are used to solve the relative deviation-based optimization models mentioned above [10,14–16], the problems of semi-deviation and the vanishing gradient, described below, arise. These may lead to the loss of better solutions in the feasible solution space, which can result in lower doses to OARs without any sacrifice of PTV treatment goals. Using the gEUD-based optimization model [15] shown in Eq. (1) as an example illustrates the shortcomings of this kind of function. $H(\cdot)$ is a step function, gEUD is the actual dose, and gEUD0 is the prescribed dose. Then,

$$f(gEUD) = H(gEUD - gEUD_0)\left(\frac{gEUD - gEUD_0}{gEUD_0}\right)^2 \tag{1}$$

(1) Semi-deviation penalties [18]. The organ constrained by the objective function (1) is assigned a quadratic penalty if and only if $gEUD \geq gEUD_0$; otherwise, the penalty is zero. In this case, reducing the dose to gEUD0 is the only incentive of using the optimization algorithm, even though a better plan with a lower gEUD may well be attainable without sacrificing other treatment goals.

(2) Vanishing gradient in the feasible solution space. If the gEUD delivered to the organ controlled by Eq. (1) is less than or equal to gEUD0, the function is equal to zero, which implies zero gradients in the interior of the feasible solution space. Assuming a feasible point in the solution space, it is difficult for a gradient-based optimization algorithm to predict the step length.

The above problems limit the search capability of the gradient algorithm. We attempt to resolve these problems by a dose prescription of zero to the OARs, whereas the conflict between PTV coverage and OAR sparing tends to exacerbate. The improvement to the linear model by Mihailidis et al. [10] and Lee et al. [14] will be reported elsewhere by us. Our work in this study seeks to propose and assess improved quadratic sub-scores for maximum dose sub-score [16] and gEUD sub-score [15], to solve the two problems described above. To verify their efficiency, the new sub-scores were incorporated into a maximum dose-based physical model and a gEUD-based hybrid physical–biological model, respectively, and were tested on three types of cases. The two models were used as optimization objective functions of the fluence map optimization (FMO) inverse problem. The large-scale, constrained optimization problems were then solved by a gradient-based optimization algorithm (L-BFGS) [19]. To avoid non-physical solutions, the square roots of the beamlet weights were considered optimized variables [20].

In the following sections, we describe in details the materials and methods contained in the proposed method. We then show our experimental results, and finally discuss the results and future direction of research.

2. Materials and methods

2.1. Improved optimization models

The new maximum-dose-based quadratic sub-score and the gEUD-based quadratic sub-score were incorporated into a physical model and a hybrid model, respectively, and were tested on three types of testing cases. The optimization models used in our work were defined as the weighted sum of sub-scores for all organs under the optimization.

(1) Improved maximum dose-based physical model

In the maximum-dose-based model, maximum-dose-based sub-score was applied to minimize the dose delivered to the OAR, along with a mean dose-based sub-score to control the dose to the PTV. A typical maximum-dose-based physical quadratic model is defined as

$$\min_{x\geq 0} \sum_{\sigma\in C} \zeta^{\sigma} \frac{1}{N_\sigma} \sum_{j\in\upsilon_\sigma} (\omega_j x - D_{\max}^{\sigma})_+^2 + \sum_{\vartheta\in T} \xi^{\vartheta} \frac{1}{N_\vartheta} \sum_{j\in\upsilon_\vartheta} (\omega_j x - D_{mean}^{\vartheta})^2. \tag{2}$$

The term $\frac{1}{N_\sigma}\sum_{j\in\upsilon_\sigma}(\omega_j x - D_{\max}^{\sigma})_+^2$ is the original maximal dose sub-score, and the operator $(\cdot)_+$ is equal to step function $H(\cdot)$. C and T represent the collection of OARs and PTV, respectively, υ_σ denotes the set of voxels in OAR σ, and υ_ϑ denotes the set of voxels in PTV ϑ. The number of voxels in the OAR and the PTV are represented by N_σ and N_ϑ, respectively. ζ^σ and ξ^ϑ are the weighting factors, representing the clinical significance of the corresponding sub-objective functions, and are determined by trial and error. $D_{\max}$ and D_{mean} represent the prescribed maximal dose to the OAR and the prescribed mean dose to the target, respectively. ω_j, computed using CERR's pencil beam algorithm (QIB), is the jth row of dose deposition matrix $\boldsymbol{W}$, and $\boldsymbol{x}$ is the optimized vector of beamlet weight (i.e., fluence element).

To solve above mentioned problems, an extra linear penalty is introduced to the maximum-dose-based sub-score. The improved physical model with piecewise penalty function is

$$\min_{x\geq 0} \sum_{\sigma\in C} \zeta^{\sigma} \frac{1}{N_\sigma} \sum_{j\in\upsilon_\sigma} (\omega_j x + (\omega_j x - D_{\max}^{\sigma})_+^2) + \sum_{\vartheta\in T} \xi^{\vartheta} \frac{1}{N_\vartheta} \sum_{j\in\upsilon_\vartheta} (\omega_j x - D_{mean}^{\vartheta})^2 . \tag{3}$$

In Eq. (3), $\frac{1}{N_\sigma}\sum_{j\in\upsilon_\sigma}(\omega_j x + (\omega_j x - D_{\max}^{\sigma})_+^2)$ is the improved maximum-dose-based sub-score. For critical structures σ, doses lower than the prescribed dose $D_{\max}$ are given a linear penalty $\omega_j x$, whereas doses higher than $D_{\max}$ are given an extra quadratic penalty. By doing so, the problems, the semi-deviation penalty and the vanishing gradient in the feasible solution space, are all solved.

(2) Improved gEUD-based hybrid physical–biological model