

文章编号: 1007-4627(2019)01-0096-08

基于 GPU 加速的质子调强放疗鲁棒优化器

徐遥¹, 裴曦^{1,†}, 刘红东¹, 霍万里¹, 周解平², 徐榭^{1,3}

(1. 中国科学技术大学物理学院, 合肥 230026;

2. 中国科学技术大学附属第一医院西区, 合肥 230031;

3. 伦斯勒理工学院, 美国 纽约 12180)

摘要: 开发一种基于图形处理器 (GPU) 加速的质子调强放疗鲁棒优化器, 用于减小质子束射程不确定性和靶区定位偏差对质子放疗的影响。建立的鲁棒优化模型使用的目标函数包括 9 种边界剂量目标, 分别是: 无偏差情况、2 种射程偏差 (偏长与偏短)、6 种摆位不确定性 (前后、侧向、上下入射方向各 2 种正负偏差)。首先靶区和危及器官的剂量贡献矩阵使用笔形束算法计算得到, 然后使用共轭梯度法优化目标函数让其满足约束条件, 这两部分均采用 GPU 加速。头颈部、肺部和前列腺三个临床病例被用来检测本优化器的性能表现。与传统基于计划靶区 (PTV) 的质子调强放疗计划相比, 鲁棒优化器能够优化出对射程不确定性和摆位误差更加不敏感的治疗计划, 让靶区实现了高剂量均匀性的同时危及器官 (OARs) 也得到了更好的保护。经过 100 次迭代, 三个病例的优化时间均在 10 s 左右。该结果证明了基于 GPU 加速的质子调强放疗鲁棒优化器能够在短时间内设计出高鲁棒性的质子治疗计划, 从而提高质子放射治疗的可靠性。

关键词: 图形处理器; 射程不确定性; 摆位不确定性; 质子调强放疗; 鲁棒优化方法

中图分类号: R73 **文献标志码:** A **DOI:** 10.11804/NuclPhysRev.36.01.096

1 引言

得益于质子束射程末端陡峭的剂量梯度, 质子调强治疗 (IMPT) 能够提供比光子治疗更加适形的靶区剂量分布的同时, 也减小了靶区周围器官的辐射损伤^[1]。在 IMPT 计划设计时, 根据射束能量与布拉格峰位置的关系, 靶区所在区域被划分为一定数量的能量层。而每个能量层上的扫描点的权重会被进一步调整以满足剂量目标和约束条件^[2]。由于布拉格峰后沿处的剂量迅速下降, IMPT 对质子射程不确定和病人摆位误差非常敏感。在 IMPT 中, 很多因素都可能会对最终的剂量分布造成影响, 比如根据 CT 值计算的阻止本领误差、摆位误差、解剖学变化等等^[3-5]。在光子治疗中, 通常在临床靶区 (CTV) 基础上向外扩展一定距离来获得计划靶区 (PTV), 从而减少计划的不确定性, 但该方法在 IMPT 中有效性较差^[6]。

为了解决上述问题, 鲁棒优化被引入用来补偿 IMPT 中会干扰最终剂量分布的不确定性。鲁棒优化的主要思想是在逆向优化阶段考虑这些不确定性, 已有多篇文献实现 IMPT 的鲁棒优化, 并且与

传统的基于 PTV 的优化技术相比, 取得了更好的效果。Unkelbach 等^[4]提出了线性规划方法以处理射程不确定性, 使用该方法得到的计划对射程变化不敏感; Pflugfelder 等^[7]提出了最差情况优化方法, 用来处理多种不确定性的情况。Fredriksson 等^[8]使用了最小最大鲁棒优化模型来同时处理射程与摆位误差, 并且在测试病例中取得了比传统优化方法更好的结果。Li 等^[9]证实鲁棒优化与基于 PTV 的优化相比能够减轻肺部病例在 IMPT 下对摆位不确定性和病人解剖变化的敏感度。Fredriksson 等^[10]比较了 3 种基于最差情况的鲁棒优化方法 (体素最差情况、目标函数最差情况和复合最差情况), 发现在所有情况下, 并没有哪一种明显占优的模型。Zaghian 等^[11]比较了最差模型与最小最大模型的鲁棒优化方法后发现: 对于前列腺与颅底癌症患者而言, 最差模型对不确定性更不敏感; 对于头颈部病例来说, 最小最大模型效果更好。IMPT 中的机会约束方法能够很好地补偿不确定性, 该方法使用线性规划方法, 在没有采用 MPI 并行方式时需要数个小时的 CPU 计算时间^[12]。与最差和最小最大鲁棒优化不同, 本文所提出的优化器的目标函数包括了所有不确定情况下的剂量分

收稿日期: 2018-12-12; 修改日期: 2019-02-26

基金项目: 国家重点研发计划资助项目 (2017YFC0107500); 国家自然科学基金资助项目 (11575180); 安徽省重点研究与开发计划项目 (1804a09020039); 安徽省自然科学基金项目 (1908085MA27)

作者简介: 徐遥 (1993-), 男, 浙江杭州人, 在读硕士研究生, 从事医学物理研究; E-mail: yaoxu@mail.ustc.edu.cn

† 通信作者: 裴曦, E-mail: xpei@ustc.edu.cn.

布, 适应面更加广泛。同时因为只是在目标函数中增加剂量分布的数目, 其计算速度要快于基于线性规划的方法。

为了使鲁棒优化适合临床实际情况 (物理上可能存在的偏差), 本文提出了包含 9 种边界剂量目标的鲁棒优化模型: 分别是: 无偏差情况、2 种射程偏差、6 种摆位不确定性 (前后、侧向、上下入射方向偏差各 2 种)。优化的目标函数包括了这 9 种情况的剂量分布, 即每一次迭代均同时优化这 9 种剂量分布。同时优化 9 种情况的剂量分布使得计算量大大增加, 因此本文使用了 GPU 来加速计算, 并且使用了 3 个临床病例来测试其有效性。

2 方法

本文建立的质子调强放疗不确定性模型基于标准优化函数, 函数中包含质子束射程不确定性和靶区定位误差。采用共轭梯度 (CG) 方法优化不确定性模型, 利用 GPU 加速矩阵向量乘法运算。通过比较基于 GPU 的鲁棒优化结果与传统的基于 PTV 的优化结果, 评价了基于 GPU 的 IMPT 快速鲁棒优化器的有效性。本文所采用的剂量计算引擎由课题组内部研发, 采用的是广泛应用于商业治疗计划系统 (TPS) 中的笔形束算法 (Pencil Beam)^[13], 并利用了 GPU 进行了加速^[14]。

2.1 不确定性模型

传统的基于 PTV 的 IMPT 优化模型与光子调强放疗 (IMRT) 一样, 都使用二次型目标函数来表示优化的目标与约束^[15]。该目标函数如下所示:

$$\min F(\omega) = \sum_{i \in \text{PTV}} \rho_{\text{PTV}} H(D_i) (D_i - D_{\text{PTV}})^2 + \sum_{i \in \text{OARs}} \rho_{\text{OARs}} H(D_i) (D_i - D_{\text{OARs}})^2, \quad (1)$$

$$D_i = \sum_{j=1}^m \phi_{i,j} \cdot \omega_j, \quad \omega_j \geq 0, \forall j \in m, \quad (2)$$

式中: PTV 表示计划靶区; OARs 表示危及器官; i 表示体素序号; j 表示扫描点序号; m 表示扫描点的总数。 ρ_{PTV} 与 ρ_{OARs} 分别表示计划靶区以及危及器官的权重。 D_i 表示体素 i 的剂量; $\phi_{i,j}$ 表示扫描点 j 对体素 i 的剂量贡献; ω_j 表示扫描点 j 的权重。 D_{PTV} 和 D_{OARs} 分别为用户设置的计划靶区与危及器官的剂量目标。当 D_i 满足剂量目标时, 选择算子 $H(D_i) = 0$, 否则 $H(D_i) = 1$ 。 $\phi_{i,j}$ 会被射程与摆位不确定性影响, 并最终影响计划靶区与危及器官的剂量分布。

在实施 IMPT 计划时, 最终剂量分布会受到射程不确定、摆位误差和解剖学变化等影响, 这些影响因素可分为相对于原始计划的射程变化与病人摆位变化两类。与传统的计划设计不同, 鲁棒优化事先考虑了这些因素的影响, 主要问题在于考虑的因素越多, 计算时间就越长。因此, 在当前计算能力的限制下, 只能对实际模型进行简化。根据现有报道, IMPT 鲁棒优化一般采用两种措施对实际情况进行简化。其中一个简化方法是离散化射束射程变化与摆位变化, 只挑选最极端的情况。本研究也采用了类似的简化手段。另一个简化手段是在迭代过程中更新目标函数时, 只考虑各个不确定性模型中最差的剂量分布。由于做了这些简化, 我们只能得到相对较为保守的计划。本研究的目标函数考虑了所有不确定性模型的剂量分布, 而不仅仅是最差情况的剂量分布, 因此能够做更全面的分析。同时为了应对增加的计算量, 我们使用了 GPU 来加速优化过程。

本文所提出的不确定性模型见式 (3), 它包括了 9 种边界剂量目标: 2 种反映射程不确定性; 6 种摆位不确定性, 前后入射方向 (anteroposterior), 侧向入射方向 (lateral), 上下入射方向 (superior-inferior) 各 2 种; 1 种反映标准情况。

$$\min F_{\text{robust}}(\omega_j) = \sum_{k \in R} \sum_{i \in \text{CTV}} \rho_{\text{CTV}}^k H(D_i^k) (D_i^k - D_{\text{CTV}})^2 + \sum_{k \in R} \sum_{i \in \text{OARs}} \rho_{\text{OARs}}^k H(D_i^k) (D_i^k - D_{\text{OARs}})^2, \quad (3)$$

$$D_i^k = \sum_{j=1}^m \phi_{i,j}^k \cdot \omega_j, \quad (4)$$

$$\omega_j \geq 0, \forall j \in m, \quad (5)$$

式中: R 表示所有射程与摆位不确定性的集合; k 表示其中一种不确定性。 ρ_{CTV}^k 与 ρ_{OARs}^k 表示在第 k 中不确定性下, 临床靶区和危及器官的权重, 通常情况下都等于 1。 D_i^k 表示第 k 种不确定性下第 i 个体素的剂量, $\phi_{i,j}^k$ 表示第 k 种不确定性下射束 j 对体素 i 的剂量贡献。与常规优化相同, 最终目标都是通过优化 ω_j 来最小化目标函数值。临床治疗时具体情况是无法预测的, 我们所能做的是设计出在这 9 种不确定性下最优的计划。

与其他文献相同, 我们做出了以下假设: (1) 所有射束都会受到相同射程不确定性的影响; (2) 射程与摆位不确定性的最大值在治疗时保持不变。

2.2 鲁棒优化方法

本文采用共轭梯度法来求解式 (3), 其流程如图 1 所示。

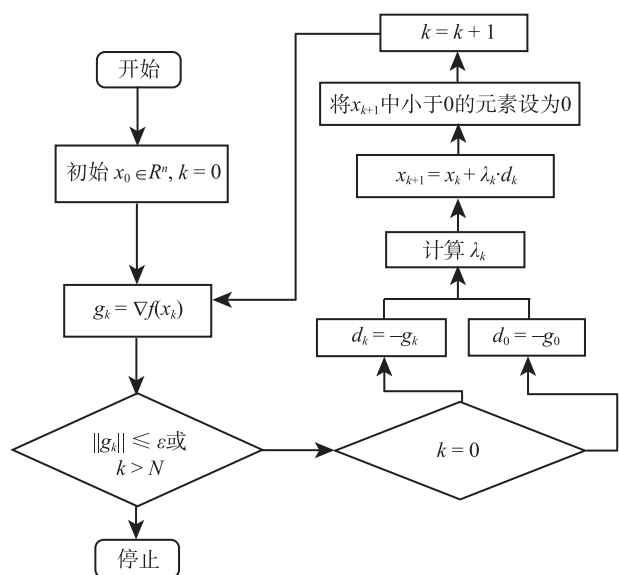


图 1 共轭梯度法流程图

在图 1 中, x_0 是初始解, k 表示当前迭代次数, g_k 表示目标函数在 x_k 处的梯度值, ϵ 表示容许的误差, 本文采用 0.1, N 表示最大迭代次数. d 表示搜索方向, λ 表示搜索步长. β_k 的计算方法为 Polak-Ribiere-Polyak 方法^[16], 即 $\beta_k = g_k^T (g_k - g_{k-1}) / \|g_{k-1}\|^2$.

使用共轭梯度法来求解最优权重时需要解决两个问题: (1) 如何处理小于 0 的权重; (2) 如何处理剂量体积约束条件. 共轭梯度法常用来求解无约束优化问题, 每一次迭代生成的权重 x_{k+1} 可能包含小于 0 的元素, 但实际上射束的权重不可能为负值. 我们所采用的解决方案是在每一次迭代后, 将 x_{k+1} 中小于 0 的元素设为 0, 但这可能会使得优化更快地陷入停滞状态. 当优化陷入停滞状态时, 将 β_k 设为 0, 使共轭梯度法退化为梯度下降法, 有助于脱离停滞状态. 在处理类似于“ $D_{60}=5000 \text{ cGy}$ ”的剂量体积约束时, 首先对所有体素的剂量进行排序, 找出剂量在前 60% 的体素, 再筛选出其中剂量小于 5000 cGy 的部分, 将这些体素的选择

算子 H 设为 1, 其余体素的选择算子设为 0.

2.3 鲁棒优化加速

对鲁棒优化的 CPU 串行版本进行测试后, 发现矩阵向量乘法占用的计算时间最多.

鲁棒优化需要储存 9 种不确定模型下的剂量贡献矩阵, 占用的内存非常大, 矩阵向量乘法所需要的时间也很长. 而由于质子的布拉格峰特性, 某一个射束只会对少数体素产生明显的剂量贡献, 剂量贡献矩阵中大部分元素都接近于 0. 因此, 采用行压缩稀疏矩阵格式 (Compressed Sparse Row Format) 来储存剂量贡献矩阵能够节省内存空间并提高计算速度. 行压缩稀疏矩阵采用了三个数组来表示稀疏矩阵, 分别存储了有效元素、列序号以及压缩后的行偏移量, 具体例子见图 2.



图 2 行压缩系数矩阵

为了进一步提高系数矩阵向量乘法 (SpMV) 的速度, 需要采用并行加速. 针对 CPU, 我们基于 Intel TBB 实现了多线程版本的 SpMV. 针对 GPU, 采用了 Nathan Bell 提出的 SpMV 算法^[17]. 在 CPU 多线程版本中, 1 个线程处理矩阵的 1 行; 在 GPU 版本中, 1 个 warp 内的 32 线程处理矩阵的 1 行. 使用实际病例的剂量贡献矩阵测试后, 发现 CPU 与 GPU 版本均能接近内存或显存的带宽极限, 并且 Nathan Bell 所提出的方法要快于 NVIDIA 官方提供的 cuSPARSE 库. 在优化过程中, 计算密集部分例如剂量计算以及优化过程均在 GPU 上进行, CPU 只负责管理病人数据以及处理剂量贡献矩阵, 流程见图 3.

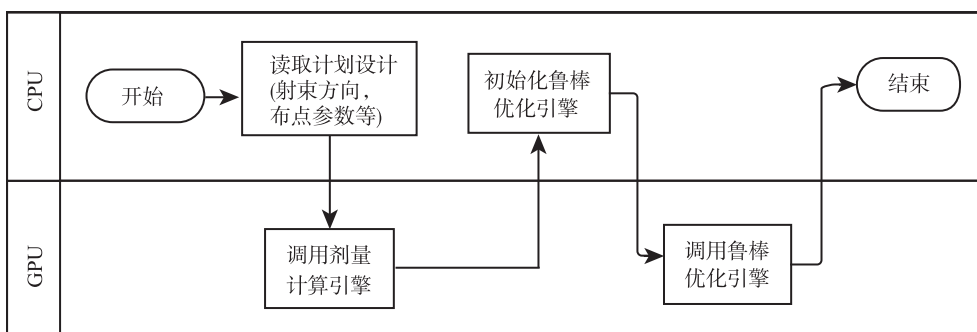


图 3 优化流程图

2.4 测试数据

本文使用 3 个病例测试了鲁棒优化器的有效性, 其中包括 1 例头颈部病例, 1 例肺部和 1 例前列腺病例, 各计划的靶区和危及器官 (OARs) 的剂量限值见表 1。

表 1 三个病例的剂量目标与约束

头颈部	剂量/cGy
靶区	$D_{\text{mean}} = 7600$
脑干	$D_{\text{max}} = 5400$
腮腺	$D_{50} < 3000$
脊椎	$D_{\text{max}} < 4500$
视神经	$D_{\text{max}} < 5000$
脑垂体	$D_{\text{max}} < 5000$
颞叶	$D_{\text{max}} < 6000$
肺部	
靶区	$D_{\text{mean}} = 7500$
食道	$D_{60} < 7000$
脊椎	$D_{\text{max}} < 4500$
肺	$D_{20} < 2000$
肺	$D_5 < 4000$
心脏	$D_{50} < 2500$
心脏	$D_{25} < 4500$
肝脏	$D_{40} < 3000$
肾脏	$D_{18} < 2000$
前列腺	
靶区	$D_{\text{mean}} = 7600$
直肠	$D_{40} < 4000$
膀胱	$D_{30} < 4000$
股骨头	$D_5 < 5000$
睾丸	$D_{\text{max}} < 200$
脊椎	$D_{\text{max}} < 4500$

在上述 3 个病例中治疗计划的制定均采用了 IMPT 技术。其中, 在头颈部病例的治疗计划中, 3 个射束角度分别为 0° , 90° 和 270° , 摆位不确定性为 ± 3 mm, 射程不确定性为射束射程的 $\pm 3\%$ 。对于肺部病例的治疗计划, 将射束的角度分别为 60° 和 320° , 摆位不确定性为 ± 5 mm, 射程不确定性为射束标准射程的 $\pm 3\%$ 。对于前列腺病例的治疗计划, 两个射束的角度分别为 90° 和 270° , 摆位不确定性和射程不确定性与肺部病例相同。每个能量层中, 扫描点在射野方向观视 (BEV) 中以 5 mm 的间距分布。我们将能量层之间的间隔设置为 80% 布拉格峰宽度。然后, 在横截面上建立 $2 \text{ mm} \times 2 \text{ mm}$ 剂量网格, 表 1 展示了剂量目标与约束。头颈部, 肺部和前列腺病例的靶区体积分别为 53.02, 98.33 和 104.04 cm^3 , 相应的总体素数分别为 7645, 239 932 和 34 890, 扫描点数分别为 23 657, 18 729 和 7 112。基于 PTV 的传统优化方法中所使用的 PTV 相对于 CTV 的外扩距离与常规 IMRT 计划相同, 并且剂量约束是针对 PTV 来设置的。本研究中, 我们在鲁棒优化时引入

不确定性模型, 不需要将 CTV 扩展成 PTV, 因此所有剂量目标都是针对 CTV 设置的^[18]。另外在将 OAR 的剂量贡献矩阵稀疏化时, 阈值设置为 10^3 , 即忽略小于矩阵最大值千分之一的元素。对于靶区, 该阈值设置为 10^{-5} 。

值得注意的是, 在常规的 IMPT 计划中并没有考虑任何的不确定性, 这使得其剂量分布不能反映出不确定性所带来的影响。因此, 为了比较基于 PTV 的优化方式和鲁棒优化方式, 我们计算了这两种优化方式设计出的计划在相同不确定模型下的剂量分布。所有测试均在同一测试平台上进行, 其配置为 Intel Core i7-6700 3.40GHz CPU, 64.0GB 内存, 以及 NVIDIA Titan V GPU。鲁棒优化器所有代码均使用 Visual Studio 2017 与 CUDA 10 编译。编译 CPU 代码时, 使用了 O2 优化级别; 编译 GPU 代码时, 使用了 O3 优化级别。

3 结果

本文使用了 3 个病例 (头颈癌、肺癌和前列腺癌) 对前述的鲁棒优化器进行了验证。分别使用传统的基于 PTV 的优化器和基于 CTV 的鲁棒优化器对治疗计划进行优化。通过计算每种方法下最差的剂量分布来量化鲁棒性, 进而对两种方法的结果做出比较。正常情况与各种不确定性下, 相应的感兴趣结构的剂量体积直方图 (DVHs) 经过标称后展示。DVH 的淡蓝色阴影展示了考虑不确定性后治疗计划的鲁棒性, 阴影范围越大, 意味着计划的敏感性越强, 鲁棒性越差。

图 4 显示了头颈部癌症病例的 CTV、脑干和腮腺的 DVH。图 5 显示肺部病例 CTV、食管和脊髓的 DVH。图 6 显示前列腺病例 CTV、膀胱和直肠的 DVH。基于 PTV 的传统计划和基于 CTV 的鲁棒优化计划的剂量约束和优化参数都相同。与基于 PTV 的传统计划相比, 鲁棒优化的 CTV 的 DVH 阴影更加狭窄, 表明鲁棒优化计划对参数的不确定性更加不敏感。此外, CTV 的 DVH 更加陡峭, 鲁棒优化计划的最大剂量也更小。这说明该鲁棒优化器既可以实现靶区内更高的剂量, 又可以实现对正常组织更好的保护。

4 讨论

与光子放射治疗相比, 质子束的物理性质 (陡峭的布拉格峰) 决定了质子剂量分布对治疗参数的不确定性更加敏感。在 IMPT 中, 治疗计划通过优化每个质子束的多个能量层上的扫描点的权重来获得靶区和危及器官的理想剂量分布。

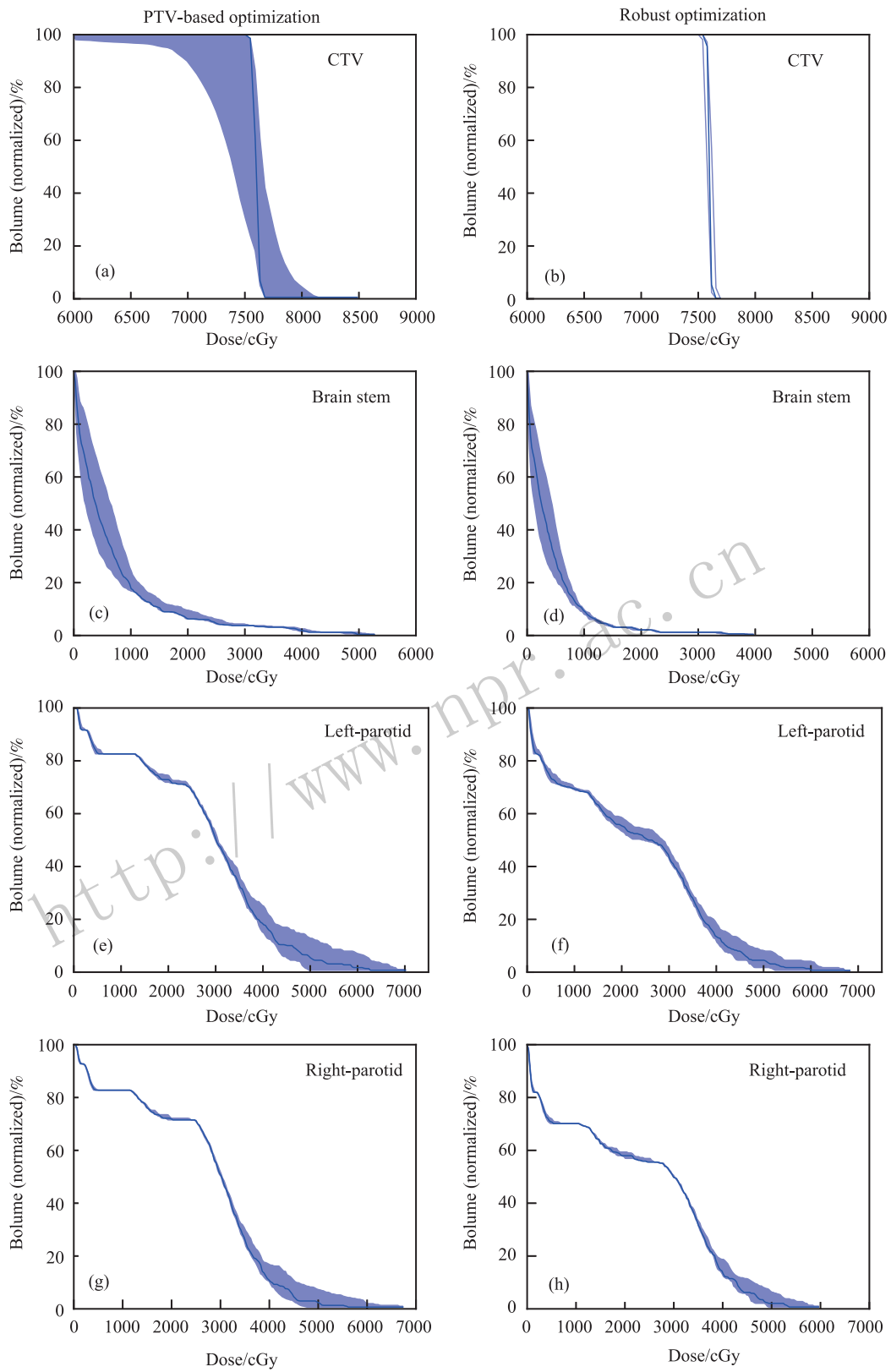


图 4 (在线彩图) 考虑不确定性的头颈部 DVH
实线表示无偏差的剂量曲线。图(a), (c), (e) 和 (g) 分别是基于 PTV 治疗计划的 CTV、脑干、左腮腺和右腮腺的 DVH。图(b), (d), (f) 和 (h) 是基于鲁棒优化治疗计划的 CTV、脑干、左腮腺和右腮腺的 DVH。

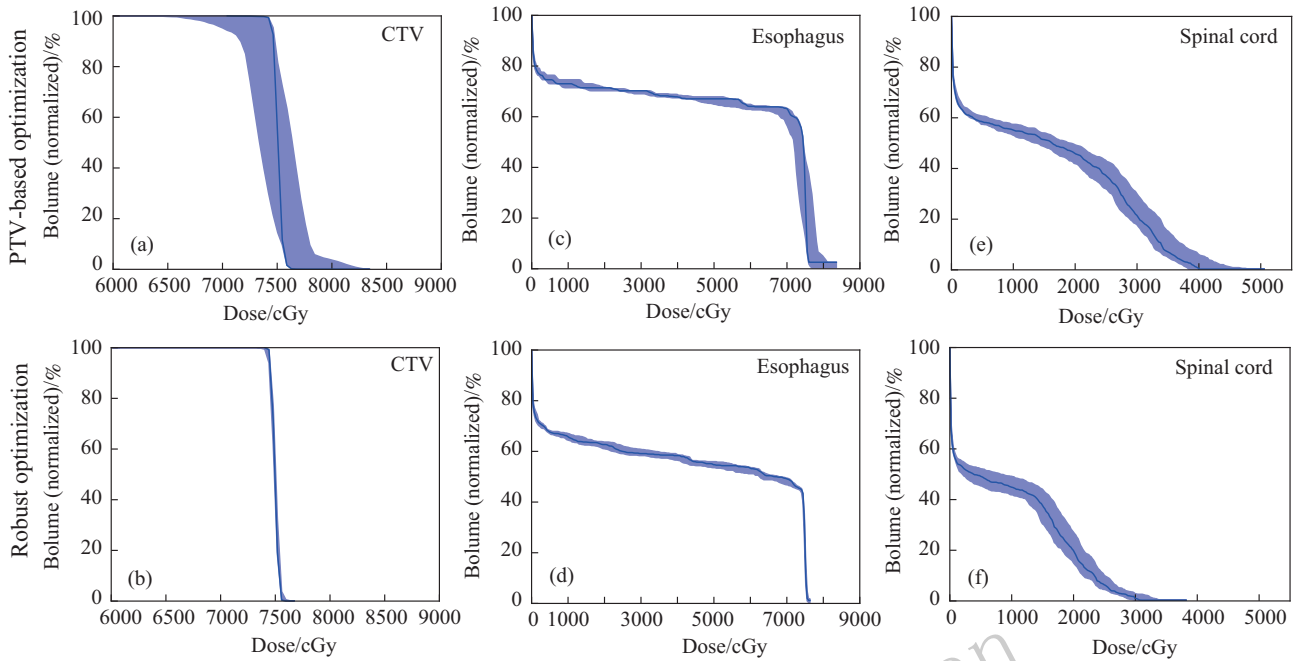


图 5 (在线彩图) 考虑不确定性的肺部 DVH
 实线表示无偏差的剂量曲线。图(a), (c) 和 (e) 分别是基于 PTV 治疗计划的 CTV、食管和脊髓的 DVH。图(b), (d) 和 (f) 分别是基于鲁棒优化治疗计划的 CTV、食管和脊髓的 DVH。

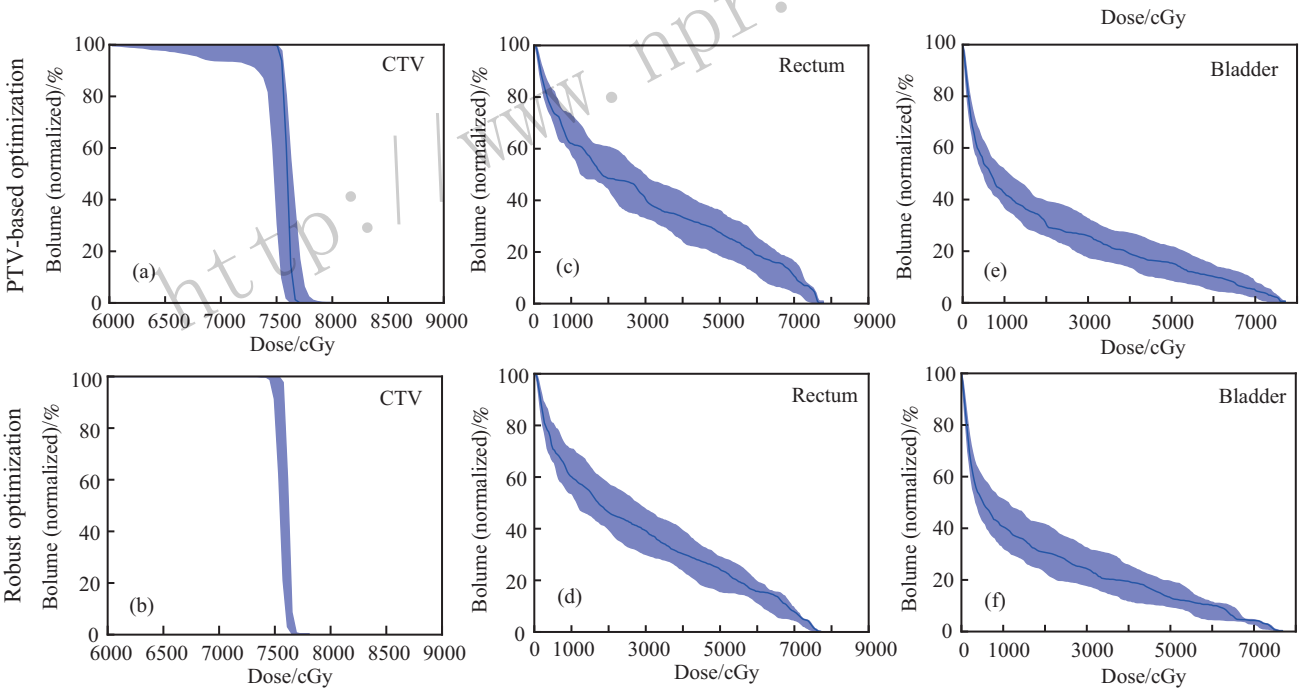


图 6 (在线彩图) 考虑不确定性的前列腺 DVH
 实线表示无偏差的剂量曲线。图(a), (c) 和 (e) 分别是基于 PTV 治疗计划的 CTV、直肠和膀胱的 DVH。图(b), (d) 和 (f) 是基于鲁棒优化治疗计划的 CTV、直肠和膀胱的 DVH。

在评估质子放射治疗计划时, 治疗参数不确定性的影响是不容忽视的。在本文中, 使用了考虑射程和摆位不确定性得到的最差剂量分布作为评估质子放射治疗计划的鲁棒性的指标。本文提出了一种基于 GPU 的快速

鲁棒优化器, 它考虑了标准射程、增加射程和缩短射程以及摆位的不确定性。

对于头颈部病例(如图 4 所示), 当基于 PTV 的计划和鲁棒计划都未考虑摆位和射程的不确定性时,

CTV 的平均剂量能够达到 7 600 cGy。但是, 在考虑摆位和射程不确定性时, 基于 PTV 的计划中 CTV 的最小剂量和最大剂量分别变为 6 641.28 和 8 485.72 cGy; 而在鲁棒计划中, 剂量与初始计划是相对一致的, 分别是 7 493.30 和 7 694.21 cGy。脑干的剂量要小于 5 000 cGy, 鲁棒计划中脑干的最大剂量为 3 951.75 cGy, 满足剂量约束; 而基于 PTV 的计划中脑干的最大剂量为 5 268.80 cGy, 不满足约束条件。对于腮腺、脊髓和垂体, 基于 PTV 的计划和鲁棒计划都可以满足约束条件。

对于肺部病例(如图 5 所示), 当基于 PTV 的计划和鲁棒计划都未考虑摆位和射程的不确定性时, PTV 的最小剂量和最大剂量分别为处方剂量的 98.9% 和 101.3%。但是, 在考虑摆位和射程不确定性时, 基于 PTV 的计划中 CTV 的最小剂量和最大剂量分别为 6 364.81 和 8 343.21 cGy; 而在鲁棒计划中, 相应的剂量与初始计划相对一致, 分别为 7 304.67 和 7 673.41 cGy。由于食管和靶区重叠, 食管的剂量很难满足剂量约束。显然, 与基于 PTV 的计划相比, 鲁棒计划可以实现较低的食管剂量。脊髓的剂量要求不超过 4 500 cGy, 鲁棒计划中脊髓的最大剂量 3 820.46 cGy, 满足剂量限制, 而在基于 PTV 的计划中脊髓的剂量不满足要求。

对于前列腺病例(如图 6 所示), 当不考虑摆位和射程的不确定性时, PTV 的最小剂量和最大剂量大约为处方剂量的 98.8% 和 101.2%。在考虑摆位和射程不确定性时, 基于 PTV 的计划中 CTV 的最小剂量和最大剂量分别为 6 011.33 和 7 988.27 cGy, 而在鲁棒计划中, 相应的剂量与初始计划相对一致, 分别为 7 291.19 和 7 809.95 cGy。基于 PTV 的计划中直肠剂量略微超过直肠的期望剂量, 而鲁棒计划完全能满足直肠的剂量约束。关于膀胱和远离 PTV 的其他正常器官, 两个计划都可以剂量限制。

鲁棒优化不确定模型包含了 9 种边界条件, 计算量是普通优化的 9 倍。如果采用 CPU 串行实现, 计算时间较长, 头颈部病例, 肺部病例和前列腺病例进行 100 次迭代的优化时间分别为 87, 140, 50 s。如果使用 CPU 多线程实现, 头颈部病例、肺部病例和前列腺病例进行 100 次迭代的优化时间分别为 30, 47 和 17 s。在 GPU 上运行矩阵向量乘法和其他耗时的计算, 可以大大减少优化时间。头颈部病例、肺部病例和前列腺病例进行 100 次迭代的优化时间分别为 11, 10 和 7 s。

影响质子鲁棒优化计划计算速度的因素主要包括靶区和危及器官的体素总数、质子笔形束的数量、

GPU 和 CPU 的计算能力、优化参数等等。当剂量贡献矩阵超过 GPU 全局内存容量时, CPU 和 GPU 之间的频繁数据传输成为提升性能的主要瓶颈。随着 GPU 性能的发展, 质子鲁棒优化的计算速度肯定会有所提升。

5 结论

本文提出的鲁棒优化器可以极大地提高质子计划的鲁棒性, 补偿射程和摆位的不确定性。与基于 PTV 的计划相比, 快速鲁棒优化器可以提高靶区的高剂量均匀性, 同时能够保护危及器官。针对靶区不在高度异质区域情况下, 虽然不确定性模型只有 9 个边界剂量目标, 但是在随机的摆位和射程不确定性的情况下也可以获得良好的结果。基于 GPU 平台, 可以在十余秒内完成鲁棒优化过程, 极大地提高了优化器的计算速度。以后我们将会把本文所使用的不确定性模型与 GPU 加速方法扩展到解剖结构的定位和变形、呼吸运动等研究中。

参考文献:

- [1] LOMAX A. *Physics in Medicine and Biology*, 1999, **44**(1): 185.
- [2] ARES C, HUG E B, LOMAX A J *et al. International Journal of Radiation Oncology Biology Physics*, 2009, **75**(4): 1111.
- [3] LOMAX A J, BÖRINGER T, BOLSI A, *et al. Medical Physics*, 2004, **31**(11): 3150.
- [4] UNKELBACH J, BORTFELD T, MARTIN B C, *et al. Medical Physics*, 2008, **36**(1): 149.
- [5] PARODI K, ENGHARDT W, HABERER T. *Physics in Medicine and Biology*, 2002, **47**(1): 21.
- [6] MEYER J, BLUETT J, AMOS R, *et al. International Journal of Radiation Oncology Biology Physics*, 2010, **78**(2): 428.
- [7] PFLUGFELDER D, WILKENS J J, OELFKE U. *Physics in Medicine and Biology*, 2008, **53**(6): 1689.
- [8] FREDRIKSSON A, FORSGREN A, HÅRDEMARK B. *Medical Physics*, 2011, **38**(3): 1672.
- [9] LI H, ZHANG X, PARK P, *et al. Radiotherapy and Oncology*, 2015, **114**(3): 367.
- [10] FREDRIKSSON A, BOKRANTZ R. *Medical Physics*, 2014, **41**: 081701.
- [11] ZAGHIAN M, CAO W, LIU W, *et al. Journal of Applied Clinical Medical Physics*, 2017, **18**(2): 15.
- [12] AN Y, LIANG J, SCHILD S E, *et al. Medical Physics*, 2017, **44**(1): 28.
- [13] HONG L, GOITEIN M, BUCCIOLINI M, *et al. Physics in Medicine & Biology*, 1996, **41**(8): 1305.
- [14] GU X, CHOI D, MEN C, *et al. Physics in Medicine and Biology*, 2009, **54**(20): 6287.
- [15] OELFKE U, BORTFELD T. *Medical Dosimetry*, 2001, **26**(2): 113.
- [16] POLYAK B T. *USSR Computational Mathematics and Mathematical Physics*, 1969, **9**(4): 94.

- [17] BELL N, GARLAND M. Efficient sparse matrix-vector multiplication on CUDA[R]. Nvidia Technical Report NVR-2008-004, Nvidia Corporation, 2008.
- [18] LIU W, LI Y, LI X *et al.* *Medical Physics*, 2012, **39**: 3089.

Robust Optimizer for Intensity Modulated Proton Therapy based on GPU

XU Yao¹, PEI Xi^{1,†}, LIU Hongdong¹, HUO Wanli¹, ZHOU Jieping², XU Xie^{1,3}

(1. School of Physical Sciences, University of Science and Technology of China, Hefei 230026, China;

2. The First Affiliated Hospital of USTC West District, Hefei 230031, China;

3. Rensselaer Polytechnic Institute, Troy 12180, USA)

Abstract: This paper describes the development of a fast robust optimization tool that takes advantage of the GPU technologies. The objective function of the robust optimization model considered nine boundary dose distributions—two for \pm range uncertainties, six for \pm set-up uncertainties along anteroposterior (A-P), lateral (R-L) and superior–inferior (S-I) directions, and one for nominal situation. The nine boundary influence matrices were calculated using an in-house dose engine for proton pencil beams of a finite size, while the conjugate gradient method was applied to minimize the objective function. The GPU platform was adopted to accelerate both the proton dose calculation algorithm and the conjugate gradient method. Three clinical cases—one head and neck cancer case, one lung cancer case and one prostate cancer case—were investigated to demonstrate the clinical significance of the proposed robust optimizer. Compared with conventional planning target volume (PTV) based IMPT plans, the proposed method was found to be conducive in designing robust treatment plans that were less sensitive to range and setup uncertainties. The three cases showed that targets could achieve high dose uniformity while organs at risks (OARs) were under better protection against setup and range errors. The run times for the three cases were around 10 s for 100 iterations. The GPU-based fast robust optimizer developed in this study can serve to improve the reliability of traditional proton treatment planning by achieving a high level of robustness in a much shorter time.

Key words: GPU; range uncertainty; setup uncertainty; IMPT; robust optimization

Received date: 12 Dec. 2018; **Revised date:** 26 Feb. 2019

Foundation item: National Key R&D Program of China (2017YFC0107500); National Natural Science Foundation of China (11575180); Anhui Provincial Key R&D Program (1804a09020039); Anhui Provincial Natural Science Foundation (1908085MA27)

† **Corresponding author:** PEI Xi, E-mail: xpei@ustc.edu.cn.