Interactive Dose Shaping - efficient strategies for
CPU-based real-time treatment planning

P Ziegenhein, C P Kamerling and U Oelfke
German Cancer Research Center (DKFZ), Department of Medical Physics in Radiation Oncology, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
E-mail: p.ziegenhein@dkfz.de

Abstract. Conventional intensity modulated radiation therapy (IMRT) treatment planning is
based on the traditional concept of iterative optimization using an objective function specified
by dose volume histogram constraints for pre-segmented VOIs. This indirect approach suffers
from unavoidable shortcomings: i) The control of local dose features is limited to segmented
VOIs. ii) Any objective function is a mathematical measure of the plan quality, i.e., is not able
to define the clinically optimal treatment plan. iii) Adapting an existing plan to changed patient
anatomy as detected by IGRT procedures is difficult.

To overcome these shortcomings, we introduce the method of Interactive Dose Shaping (IDS)
as a new paradigm for IMRT treatment planning. IDS allows for a direct and interactive
manipulation of local dose features in real-time. The key element driving the IDS process is a
two-step Dose Modification and Recovery (DMR) strategy: A local dose modification is initiated
by the user which translates into modified fluence patterns. This also affects existing desired
dose features elsewhere which is compensated by a heuristic recovery process.

The IDS paradigm was implemented together with a CPU-based ultra-fast dose calculation and
a 3D GUI for dose manipulation and visualization. A local dose feature can be implemented via
the DMR strategy within 1-2 seconds. By imposing a series of local dose features, equal plan
qualities could be achieved compared to conventional planning for prostate and head and neck
cases within 1-2 minutes.

The idea of Interactive Dose Shaping for treatment planning has been introduced and first
applications of this concept have been realized.

1. Introduction
The IMRT planning process traditionally consists of an iterative Newton-based optimization of
an objective function which is defined from dose volume constraints of pre-segmented volumes
of interest [2]. This approach is one of the most popular methods to generate an IMRT plan
in clinical environments. However, it suffers from various inherent shortcomings. First, the
locality of dose prescription is too coarse. Local dose features can only be restrained to pre-
segmented VOIs. Eliminating undesired hotspots in the therapy plan is difficult, controlling the
shape of an individual isodose line is almost impossible. Second, the trade-off between planning
goals in the objective function is expressed by a set of scalar, relative weighting factors. The
meaning of these factors is clinically not intuitive and the planner has to find the right set
of parameters through a time intensive trial-and-error approach. Third, using the traditional
optimization method, it is difficult to rapidly adapt already optimized treatment plans to a
changed patient anatomy as it is required in adaptive radiation therapy (ART). In addition
to these clinical drawbacks, the method of the iterative Newton based optimization as used in radiation therapy also has several disadvantages from the computational point of view. In [8] it is claimed that the optimization method is a memory bound problem when a pre-calculated dose influence matrix \((D_{ij})\) is used. The runtime of the optimization is determined by the transfer speed of the dose influence data between the memory of the planning system to the arithmetic unit (e.g. CPU or GPU). It could be shown that the arithmetic performance potential of modern computer systems cannot be exploited by the optimization method. Since the (parallel) arithmetic computation speed of modern computers develops faster than memory bandwidth, one has to conclude that the Newton-based IMRT plan optimization is not well suited for modern throughput oriented computer systems as used for instance in [5]. Furthermore, the iterative nature of the optimization method prevents an efficient implementation on distributed memory systems.

In this work, an alternative therapy planning paradigm called Interactive Dose Shaping (IDS) is introduced. It is designed to overcome the clinical shortcomings of the traditional IMRT plan optimization. In contrast to [3] the presented interactive planning method does not employ an objective function which is driving a Newton-based optimization but a so called Dose Modification and Recovery strategy which is introduced in this work.

2. Materials & Methods

2.1. The Concept of Interactive Dose Shaping

The development of the IDS paradigm is driven by three main key features. First, the new planning technique involves the therapist more into the planning process. In contrast to the traditional method which consist of an automated optimization block, the therapist should be enabled to interactively control the plan generation process by imposing a series of local dose features into the plan. This paradigm shift facilitates the transition of clinical knowledge from the therapist to the plan generating algorithm. Second, using IDS it should be possible to adapt pre-optimized therapy plans to a changed patient geometry easily and third, the IDS engine has to be implemented by compute-bound algorithms to guarantee proper performance scaling on modern computational architectures.

In order to realize these three features, the IDS concept employs a two-step Dose Modification and Recovery (DMR) strategy. In the dose Modification step, the IDS engine imposes a local, desired dose feature into the treatment plan as requested by the therapist. The request can be formulated by interactive graphical tools via a 3D user interface in the planning framework. The desired local dose feature is implemented into the plan by manipulating the relevant fluence amplitudes of the beam setup. Changing the fluence amplitudes of the beams does not only impose the intended dose feature but also affects other areas of the plan. These unintended dose variations are compensated in a subsequent Recovery step. The Recovery identifies affected areas outside the desired locality and generates a manipulation pattern for the fluence amplitudes which re-installs already established dose features without eradicating the foregoing dose Modification step. Depending on the complexity of the clinical plan, 10 to 30 locations have to be recovered in order to compensate for unwanted, unintentional dose changes outside the Modification area. The initial plan on which the DMR process is applied can either be a pre-optimized plan (in order to impose small alterations due to geometrical changes of the patient) or an unmodulated plan to start from scratch as shown in the example in figure 1.

The Dose Modification and Recovery strategy is explained in detail in section 2.3. As mentioned before, using the \(D_{ij}\) matrix approach for the dose calculation is a memory-bound problem and therefore not suited for the interactive planning paradigm. It is replaced by an on-the-fly dose calculation method which is described in section 2.2. The IDS engine is embedded into a new treatment planning framework which provides a powerful three-dimensional graphical representation of the plan together with a set of tools which provide an interface for implementing
the interactive dose Modification. Some aspects of the dose Modification interface and parts of this framework is described in [4]

2.2. Ultra-fast, low-latency dose calculation
It is necessary to re-calculate the dose distribution after imposing a dose Modification and after each Recovery step due to two reasons: On the one hand, the effect of the dose Modification must be visible to the subsequent Recovery module. On the other hand, the Recovery locations are chosen sequentially based on the effect of the foregoing Recovery. Thus, 11 to 31 dose calculations (one dose Modification and 10 to 30 Recovery steps) have to be performed for each DMR process within a real-time constraint (1-2 seconds).

To solve this task an ultra-fast CPU-based dose calculation algorithm is introduced based on the pencil beam convolution technique with singular value decomposed kernels according to [1]:

$$D_f(x_p, y_p, d) = \sum_{i=1}^{3} D'_i(d) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi'(x, y) f(x, y) \times w_i(x - x_p, y - y_p) dx dy$$  

(1)

The dose $D$ in the radiological depth $d$ at the point $(x_p, y_p)$ relative to the central axis is calculated as a sum over the depth dependent kernel components $D'_i(d)$ multiplied with the convolution of the fluence $\Psi'(x, y)$, the modulation (transmission) function $f(x, y)$ and the kernel weight $w_i(x - x_p, y - y_p)$. Usually, the convolution is taken out in frequency domain via a fast Hartley transform [1] [6]. For the use in IDS the convolution is calculated in a discretized spatial domain which has three main advantages: i) The DMR strategy imposes local modifications of the fluence amplitudes. In spatial domain these can be considered as relatively small additional terms to the convolution. ii) The discrete convolution in spatial domain consists of simple multiply add operations which are highly optimized in hardware on modern CPUs. iii) By controlling the radius of the finite pencil beam kernel weight $w_i(x - x_p, y - y_p)$ the trade-off accuracy/time can be individually selected for each region of interest.

2.3. The Dose Modification and Recovery strategy
The Dose Modification and Recovery process as a central part of the IDS engine is shown in figure 1a). The IDS process begins with a request from the user who specifies a local dose Modification to the plan (first green box). The request is realized by a set of interactive tools which provide a variety of Modification means. A sophisticated request which can be used for clinical planning is for instance the manipulation of a three-dimensional isodose surface which is explained in [4]. For the scope of this paper which is intended to demonstrate the concepts of IDS, a much simpler user request is considered: The user demands a dose Modification of only one voxel of the plan. An example of the voxel Modification tool is depicted in figure 2 on a phantom setup containing a horse-shoe shaped target and a spherical OAR in close proximity. The figure shows a screen shot from our newly designed planning framework with the voxel selecting tool enabled (the tool proposes groups of 4x4x4 voxels which are shown as color-washed cubes). On the example of requesting a dose Modification on the encircled voxel in figure 2 the DMR process is described in the following while the change of the dose pattern throughout the process is shown in figure 1 b), c) and d) on one transversal slice of the planning setup.

The initial plan on which the user request the dose Modification is shown in figure 1b). The beam fluences are not modulated and the shown isodose line denotes a prescribed dose to the target volume. The location of the Modification is marked by a white dot in figure 1c) and 1d). The user requests a dose reduction of 30% in that voxel $\tilde{m}$. The first step of the DMR strategy is to impose the request into the plan. To accomplish that, the dose Modification module (first blue box in Figure 1a) imposes a manipulation of the fluence map $\Delta f_M$ so that the requested
Already imposed plan features are re-installed by a Recovery process performing a series of automated compensating dose modifications. The requested Dose Modification is imposed into the treatment plan, inevitably affecting existing desired plan features elsewhere.

Figure 1. a) The process of the Dose Modification and Recovery strategy (blue boxes) and the relation to the graphical user interface (green boxes). b) The initial planning state of the phantom case. c) After the dose Modification of one voxel (white dot). d) After the Recovery process.

In a second step the undesired change on the dose distribution outside the area of the Modification to \( \vec{m} \) is compensated by the Recovery. The Recovery module (second blue box in figure 1a) identifies locations \( \vec{r} \) which have been affected unintentionally by the dose Modification and provides one adaptation \( \Delta f_k \) to the fluence map in each recovery step \( k \) to re-install the initial feature in \( \vec{m} \) is installed. The result is shown in Figure 1c). The dose Modification not only imposes the desired local dose feature \( D^{(f + \Delta f_M)}(\vec{m}) \) but also affects the global dose distribution, that means there is another voxel \( \vec{r} \) so that \( D^{(f + \Delta f_M)}(\vec{r}) \neq D(f(\vec{r})) \). In this example the effect on the target volume is very distinct: The isodose line in figure 1c) reveals the undesirable cold spots in the target volume as a side effect of the foregoing fluence manipulation.
dose distribution:

\[ D^{(f_1 + \Delta f_m + \Delta f_R)}(\vec{r}) \approx D^f(\vec{r}) \]  

(2)

For that, the Recovery strategy prefers fluence amplitudes manipulations \( \Delta f_R \) that have a minimal effect on the previously requested dose feature in \( \vec{m} \). The effect of the Recovery process is shown in figure 1d). On the one hand, the dose distribution in the target is almost fully recovered while on the other hand, the requested dose Modification in voxel \( \vec{m} \) is still in place. After the Recovery process is finished, the resulting 3D dose distribution is presented to the user for evaluation. The whole DMR process can be done in real-time (about 1-2 seconds) on a single processor desktop PC. This guarantees the interactive character of the IDS concept and provides the therapist with an immediate feedback on the requested dose Modification.

3. Results

The concepts of the Interactive Dose Shaping paradigm containing the DMR strategy, the ultra-fast dose calculation module, several dose Modification tools and a three-dimensional graphical representation of the patient geometry/dose have been implemented into a new treatment planning framework.

The implementation of the DMR strategies and the adapted dose calculation method is fast enough to allow for interactive planning of clinical 3D treatment plans within real-time. The runtime of the dose calculation is in the order of 20 ms for typical nine-beam plan.

In [4] the planning result of a prostate case is discussed using the 3D isodose manipulation tool for IDS. A planning result for a more complex intracranial case is shown in figure 3(b). The target dose was prescribed to 6000 cGy. The maximum dose to the left eye was set to 1000 cGy while 2000 cGy was allowed to the right eye due to the close proximity to the tumor. The IDS planning process took less than 2 minutes by imposing various local dose Modifications on a nine-beam photon setup. This planning time includes user interactions and the computational runtime of the DMR.

4. Discussion

The short runtime of the DMR implementation and of the adapted dose calculation method shows that the IDS method is feasible to realize a new interactive planning method from a computational point of view. In contrast to the traditional Newton-based optimization method using pre-calculated dose influence data, IDS is able to exploit the parallel arithmetic capabilities of modern CPU-based hardware which enables scaling to high performance computing and especially future architectures.
The first results presented in this paper and in [4] demonstrate that the IDS planning method is able to derive plans with a similar quality than achieved by traditional methods.

5. Conclusion
The IDS planning paradigm has the potential to become an new interactive planning method which can be used in the clinical routine some day. Although, the concepts are still under development we could demonstrate that the new method is accurate and fast enough to allow for interactive planning which directly includes the therapist into the plan generation process.

References
[1] Bortfeld T 1993 Decomposition of pencil beam kernels for fast dose calculations in three-dimensional treatment planning Med. Phys. 20 311
[2] Bortfeld T 2006 IMRT: a review and preview Phys. Med. Biol. 51 R363
[3] Craft D L, Hong T S, Shih H A and Bortfeld T R 2012 Improved Planning Time and Plan Quality Through Multicriteria Optimization for Intensity-Modulated Radiotherapy I. J. of Rad. Onc. Bio. Phys. 82 e83 - e90
[4] Kamerling C P, Ziegenhein P, Oelfke U 2013 A 3D isodose manipulation tool for interactive dose shaping J. Phys.: Conf. Ser. presented at ICCR 2013, not yet published
[5] Men C, Gu X, Choi D, Majumdar A, Zheng Z, Mueller K and Jiang S 2009 GPU-based ultrafast IMRT plan optimization Phys. Med. Biol. 54 6565-73
[6] Siggel M, Ziegenhein P, Nil1 S and Oelfke U 2012 Boosting runtime-performance of photon pencil beam algorithms for radiotherapy treatment planning Physica Medica 28(4) 273-80
[7] Ziegenhein P, Kamerling C P, Bangert M, Kunkel J and Oelfke U 2013 Performance-optimized clinical IMRT planning on modern CPUs Phys. Med. Biol. 58 3705-15
[8] Ziegenhein P, Kamerling C P, Heinrich H and Oelfke U 2012 SU-C-213AB-01: Application of Ultra-Fast Dose Calculation in Real-Time Interactive Treatment Planning Med. Phys. 39 3598

---

Figure 3. IDS planning results on a complex intracranial patient case. a) The geometrical patient setup. b) The DVH comparing the traditionally Newton based optimization (dashed curves) and using IDS (solid curves)