A Difference-Matrix Metaheuristic for Intensity Map Segmentation in Step-and-Shoot IMRT Delivery

Athula Gunawardena†‖, Warren D'Souza‡, Laura Goadrich§, Robert Meyer‖, Kelly Sorensen‖, Shahid A. Naqvi‡, and Leyuan Shi§

† Department of Mathematics and Computer Sciences, University of Wisconsin-Whitewater, 800 West Main Street, Whitewater, WI, USA
‡ Department of Radiation Oncology, University of Maryland, School of Medicine, 22 South Greene Street, Baltimore, MD, USA
§ Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, USA
‖ Department of Computer Sciences, University of Wisconsin-Madison, Madison, WI, USA

Abstract. At an intermediate stage of radiation treatment planning for IMRT, most commercial treatment planning systems for IMRT generate intensity maps that describe the grid of beamlet intensities for each beam angle. Intensity map segmentation of the matrix of individual beamlet intensities into a set of MLC apertures and corresponding intensities is then required in order to produce an actual radiation delivery plan for clinical use. Mathematically, this is a very difficult combinatorial optimization problem, especially when mechanical limitations of the MLC lead to many constraints on aperture shape, and setup times for apertures make the number of apertures an important factor in overall treatment time. We have developed, implemented, and tested on clinical cases a metaheuristic (that is, a method that provides a framework to guide the repeated application of another heuristic) that efficiently generates very high-quality (low aperture number) segmentations. Our computational results demonstrate that the number of beam apertures and beam-on time in the treatment plans resulting from our approach is significantly smaller than the corresponding values for treatment plans generated by the heuristics embedded in a widely-used commercial system. We also contrast the excellent results of our fast and robust metaheuristic with results from an exact method, branch-and-cut, which attempts to construct optimal solutions, but, within clinically acceptable time limits, generally fails to produce good solutions, especially for intensity maps with more than five intensity levels. Finally, we show that in no instance is there a clinically significant change of quality associated with our more efficient plans.

1. Introduction

Treatment planning in IMRT usually involves a two-step process in which (1) the dose within the tumor and critical structures is first optimized without consideration of the delivery specifics and (2) the intensity map (matrix) from each beam orientation is then segmented to produced segments (apertures) by taking into account the physical delivery constraints of the MLC. When segmenting the intensity map, one must consider both the number of shapes and the beam-on time. At our institution IMRT treatments are planned and delivered using the Corvus, v 4.0 (Nomos Corporation, Cranberry Township, PA) treatment planning system and Elekta SL 20 (Elekta Oncology Systems, Crawley, UK) linear accelerators. The leaf-sequencing algorithm in the Corvus planning system tends to produce a large number of beam delivery segments for each beam orientation (20-70) depending on the number of modulation intensity levels and the complexity of the case. There exists an inter-segment delay of approximately 6 seconds on an Elekta SL linear accelerator. Coupled with a large number of delivery segments, this can significantly increase the treatment time to approximately 45 min or longer. This has the potential to increase patient discomfort and spurious intra-fraction patient motion as well as decrease clinical throughput.

The leaf sequencing or segmentation problem described above is a particularly difficult combinatorial problem for which a variety of heuristic and "exact" approaches have been proposed. Boyer (Boyer & Yu 1999), Evans (Evans, Hansen & Swindell 1997), Galvin (Galvin, Chen & Smith 1993), Bortfeld (Bortfeld, Kahler, Waldron & Boyer 1994), Xia and Verhey (Xia & Verhey 1998), and Siochi (Siochi 1999) discuss several heuristic methods that have been developed for this problem. Based on computational comparisons in Que (Que 1999), the Xia-Verhey approach, which focuses on decompositions based on intensities that are powers of 2, performs relatively well with respect to these alternatives in terms of number of segments. On the other hand, the Bortfeld, et al. approach tends to yield larger numbers of segments, but smaller beam-on times. In our comparisons below, we present the results obtained using the commercial treatment planning system Corvus, which uses an heuristic approach to segmentation. For the case in which only beam-on time is considered, an interesting exact network flow approach to segmentation was developed recently by Boland, et al. (Boland, Hamacher & Lenzen 2004). Luan, et al. (Luan, Wang, Hu, Naqvi, Yu & Lee 2004) also investigate the use of network flow techniques, but focus on minimizing aperture count. Taking into account both aperture count and beam-on time, Langer (Langer, Thai & Papiez 2001) derived an exact approach based on the solution via branch-and-cut of an integer programming model and compared his results with those of the heuristic methods of Bortfeld (Bortfeld et al. 1994), Boyer and Yu (Boyer & Yu 1999), and Xia and Verhey (Xia & Verhey 1998). Since Langer's method is based on the "exact" procedure of branch-and-cut, it yields (given sufficient computing time, which as will be seen below, may be well beyond the time available to construct a treatment plan) the minimum number of segments subject to certain delivery constraints, including a

minimum beam-on time constraint, and hence he is able to demonstrate significant improvements relative to heuristic approaches. Unfortunately, in our experience, exact approaches based on Langer's technique failed to produce useful results when additional constraints were added to Langer's basic delivery constraints in order to take into account the more complex delivery constraints that are required to model certain MLCs. (These additional constraints and the corresponding results are described below.) The resulting problems are simply too difficult for branch-and-cut methods to tackle even when allocated several hours of computing time. The metaheuristic that we develop here coordinates the search processes arising from the use of multiple starting points for a base heuristic (which is related to a method of Engel (Engel 2003) but takes into account complex delivery constraints). It employs a composite scoring function based on a weighted sum of the number of shape matrices in the intensity map decomposition and the beam-on time. This problem is harder than the beam-on time problem, because it involves binary-valued step functions that have value 1 for a sub-aperture that is used (with positive radiation intensity) and value 0 otherwise.

Results with our metaheuristic are compared with leaf segments produced by the heuristics included in the Corvus treatment planning system (versions 4.0 and 5.0). Additionally, we benchmark our approach against a branch-and-cut based exact approach described by Langer et al. Further, we compare the resulting 3-D dose distributions resulting from our metaheuristic-generated leaf sequences to the original dose distributions using an in-house developed and previously described dose calculation engine.

2. Methods and Materials

We start with an example (see table 1) showing a 5-intensity-level map for a clinical case. This matrix of relative intensities must be decomposed into a integer-weighted sum of binary matrices, each of which must obey a set of constraints dictated by the physical mechanisms available to control the aperture of the MLC (details of these constraints are discussed below).

There are numerous types of IMRT machines currently in clinical use, with rather different physical constraints that determine the allowable leaf positions (and, hence, corresponding shape matrices) that are possible for that device. These machine types also have widely varying setup times for transitions between shapes, and for certain machine types these setup times can dominate the radiation delivery time (termed beam-on time), a factor that we will consider below.

2.1. The DMM Metaheuristic for Intensity Map Segmentation

Our research has focused on the development of a *metaheuristic* (Difference Matrix Metaheuristic (*DMM*), described in detail below) that guides the repeated applications of a *base heuristic* (*bDif*, described below) in the solution of a segmentation problem.

Table 1. Sample 5-intensity-level map.

Angle 55^o							
0	0	80	100	100	80	40	0
0	80	100	80	60	100	100	40
0	80	60	60	60	80	40	40
0	100	60	60	60	60	100	60
60	60	80	80	80	80	80	0
20	40	20	20	40	80	20	0
0	100	60	80	100	100	100	0
0	40	80	100	80	80	0	0
0	0	60	100	40	0	0	0

Roughly speaking this metaheuristic combines multi-start (the use of multiple carefully selected "starting points"), local search, and coordination mechanisms that focus the search of the heuristic in more promising areas of the solution space.

We start by describing the base heuristic *bDif*, which performs segmentations by sucessively extracting apertures that are constructed by applying "greedy" row-wise local searches. The criteria used in shape construction are based on the elements of "difference matrices." The difference matrix with respect to a given matrix consists of "forward" differences of successive columns (for technical reasons, zero end columns are added to the left and right borders of the original matrix before the difference computation). Algebraically, if A denotes an $m\times n$ matrix, then the difference matrix D_A is the $m\times(n+1)$ matrix of *column-wise forward differences* with entries $d_{i,j} = a_{i,j}-a_{i,j-1}$, where $a_{i,-1} = 0$ and $a_{i,n} = 0$. The *Diff* operator is defined by the relation $D_A = \mathit{Diff}(A)$. Note that the difference process is invertible in the sense that the original matrix can be computed from the difference matrix, so no information is lost in the construction of the difference matrix. In the algorithmic discussion below we focus on difference matrix processes, but note that the linearity and invertibility of the *Diff* operator implies that "segmentation" of a difference matrix is equivalent to segmentation of the corresponding intensity matrix. (Previous research in this area (Engel 2003) has laid the theoretical groundwork for difference-matrix methods by demonstrating that these methods can produce minimum beam-on time segmentations in the absence of leaf collision constraints.) In the algorithmic descriptions below, D will denote a generic difference matrix that corresponds to either$\mathit{Diff}(A)$, where A is the original intensity map or to a difference matrix obtained by applying *Diff* to a residual intensity map obtained by extracting one or more weighted shape matrices from A. Using this notation, the input and output of the overall *DMM* procedure are as follows:

Algebraic descriptions (via constraints that use binary variables) of the delivery constraints of the Elekta MLC determine the conditions that shape matrices S_k must satisfy and are given below in the context of the discussion of an alternative approach based on an integer programming model. Our base heuristic *bDif* enforces these

Input: An $m \times n$ intensity matrix $A = (a_{i,j})$ comprised of nonnegative integers. (The column indices of A are given by $0, \ldots, n-1$.)

Output: K binary aperture shape matrices S_k and positive integers α_k $(k = 1, \ldots, K)$ giving corresponding beam-on times for the apertures. The apertures obey the delivery constraints of the Elekta MLC and the weight-shape pairs satisfy $\sum_{k=1}^{K} \alpha_k S_k = A$.

constraints by checking deliverability as each row is added to the current shape.

The following functions play key roles in the ***base heuristic bDif***:

$\boldsymbol{Count(D)} =$ the number of nonzero entries in D.
$\boldsymbol{Freq(D, \alpha)} =$ the frequency (number of appearances) of $\pm\alpha$ in D.

An important underlying concept of *bDif* is that clusters of 0 entries in a difference matrix correspond to "flat" portions of the corresponding intensity map that suggest good shapes. Thus, *bDif* focuses on the construction of shapes whose extractions produce small values of the *Count* function. The choice of the intensity values α used in the *bDif* segmentation process is determined in part by considering values of α in descending order of $Freq(D, \alpha)$.

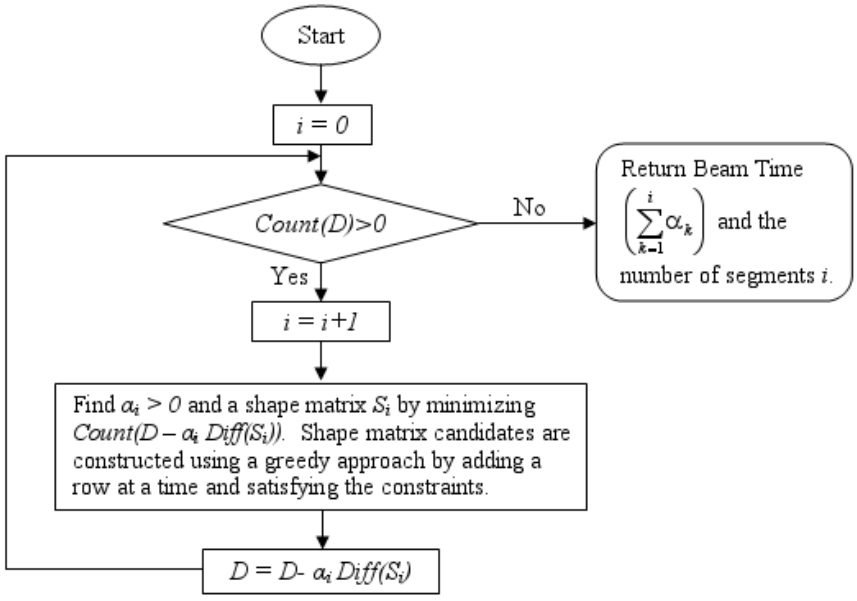


Figure 1. The main procedures of the base heuristic *bDiff*.

The following additional properties are considered in the procedures used to construct shapes:

To begin the *bDif* shape construction process with an initial row, we consider each nonzero row i in D and generate one or more row shapes based on the application of heuristics that focus on the maintenance or improvement of the value of the Bound

$$
\begin{aligned}
M_i &= \text{the 1-norm (sum of absolute values) of row } i \text{ in } D \\
N &= \{i | M_i > 0\} = \text{the indices of non-zero rows of } D \\
Max(D) &= \max\{M_i | i \in N\} = \text{the norm of the largest row in } D \\
Min(D) &= \min\{M_i | i \in N\} = \text{the smallest non-zero row norm in } D \\
Bound(D) &= \max(D) - \min(D) = \text{a measure of the variation of the entries of } D
\end{aligned}
$$

measure. We then construct a pool of 2-row shapes by adding second rows that satisfy the aperture constraints. We limit the pool size by keeping at most the C (the method parameter C is discussed further below) shapes giving the lowest values for $Count(D - \alpha Diff(S))$. Repeating this local search procedure, we continue to add one row at a time to the current partial shapes. At the end of this process, we select a completed shape from the pool and extract it from the residual intensity matrix using an intensity value as described above. The shape generation and extraction procedure is then repeated until the intensity map has been segmented by *bDif*.

In order to broaden the search for a good segmentation, the metaheuristic *DMM* (as indicated in the flow chart below) employs *bDif* within a multi-start procedure (described below) and then coordinates the corresponding multiple local searches by comparing results and focusing further search effort on promising regions that contain "good" shapes. This is done by employing a promise index $\Pi(S, \alpha, D)$ defined by a shape matrix S and a corresponding beam time α and a difference matrix D as follows: apply *bDif* to segment the residual matrix $D - \alpha Diff(S)$ and define $\Pi(S, \alpha, D) = \sum_{k=1}^{K'} \alpha_k + 7 * K'$ where K' is the number of shapes in the corresponding segmentation of $D - \alpha Diff(S)$. Low values of Π therefore correspond to "good" shapes, and a Π value of 0 indicates that a segmentation has been completed by using shape S.

DMM uses this promise index in the row-wise construction (as described in the flow chart below) of an $R \times C$ Shape Table (ST), which accumulates the results of searches for good segmentations and at completion contains the results with the smallest number of shapes in its last row. Since each element in each row corresponds to a shape used in a segmentation, the row parameter R (set to 85 in our runs) sets an upper bound on the number of shapes in the segmentation. The column parameter C determines the number of partial segmentations that will be simultaneously developed. (The value of C plays an important role in determining the run time of the procedure, and for a run time of approximately 3 minutes (on an 800MHz Linux computer) per beam angle, the following C values were used for 5,10, and 100-intensity-level maps: 40,20,2.) The elements of ST consist of tuples given by $\{S, D - \alpha Diff(S), \Pi(S, \alpha, D), previous\_index\}$, where *previous_index* is a pointer to the previous ST entry (if any) that gave rise to residual difference matrix D. Thus, following the pointers back up from a shape that completes a segmentation gives the remaining shapes and intensities in the segmentation. The initial row of ST corresponds to the C multiple starting points and is generated by setting D to $Diff(A)$, and the C shape matrices for that row are chosen to approximately minimize $Count(D - \alpha Diff(S))$. To determine the value of α used for an ST entry, we consider the

smallest v (a parameter in the heuristic) distinct elements in absolute value in $Diff(A)$, apply $bDif$ to each $D - \alpha Diff(S)$, and then select the best result (according to the α value). (In our results, the following values of v were used for 5,10, and 100-intensity-level maps: 2,4,15.) We refer to this set of values of α as the intensity pool.Successive rows of ST are generated by selecting tuples via pools of residual difference and shape matrices. The residuals $D - \alpha Diff(S)$ from the current row form the difference pool. The first shape extracted by $bDif$ from each of these residuals is placed in the current shape pool. The *(S, α, D)* tuples for the next row are generated by forming all possible tuples from the three pools, and then selecting from this set the *(S, α, D)* tuples that yield the best C values of the promise function Π. These tuples, augmented by *previous_index* pointers to appropriate entries in the preceding row, form the tuples of the next row. Note that the case $C = 1$ produces one segmentation that is identical to a *bDif* segmentation. Once a row of ST has been completed that contains at least one Π value of 0, *DMM* then terminates by using *previous_index* values to backtrack (starting at an entry with Π value of 0 that has minimum beam time) through ST to obtain the final segmentation.

The following figure illustrates a shape table and a corresponding set of pointers. In this example, segmentations (not shown) using shapes S_1 and S_3, while initially promising, were later discarded after improved segmentations using shapes S_2 and S_4 proved to be superior. The best segmentation thus occurs in one of the four chains of length 5 shown in the figure. Each of these segmentations has 5 segments (the smallest number of segments of any of the segmentations generated by the search process), so *DMM* selects a chain with minimum corresponding beam-on time from this set.

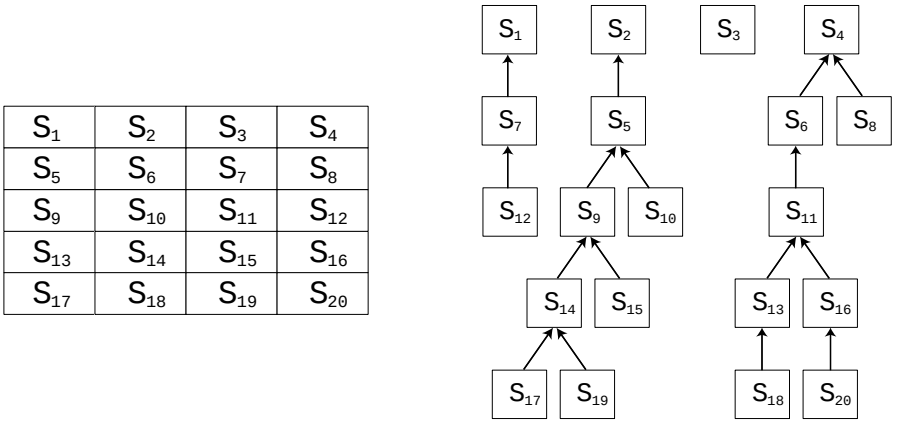


Figure 2. An example of a shape table and a corresponding search tree structure.

2.2. An "Exact" Branch-and-Cut Approach

We compare our metaheuristic with the heuristic used in Corvus and with an MIP approach based on a model proposed in Langer et al, (Langer et al. 2001) (extending his model via additional constraints, since the MLC constraints of the Elekta are more

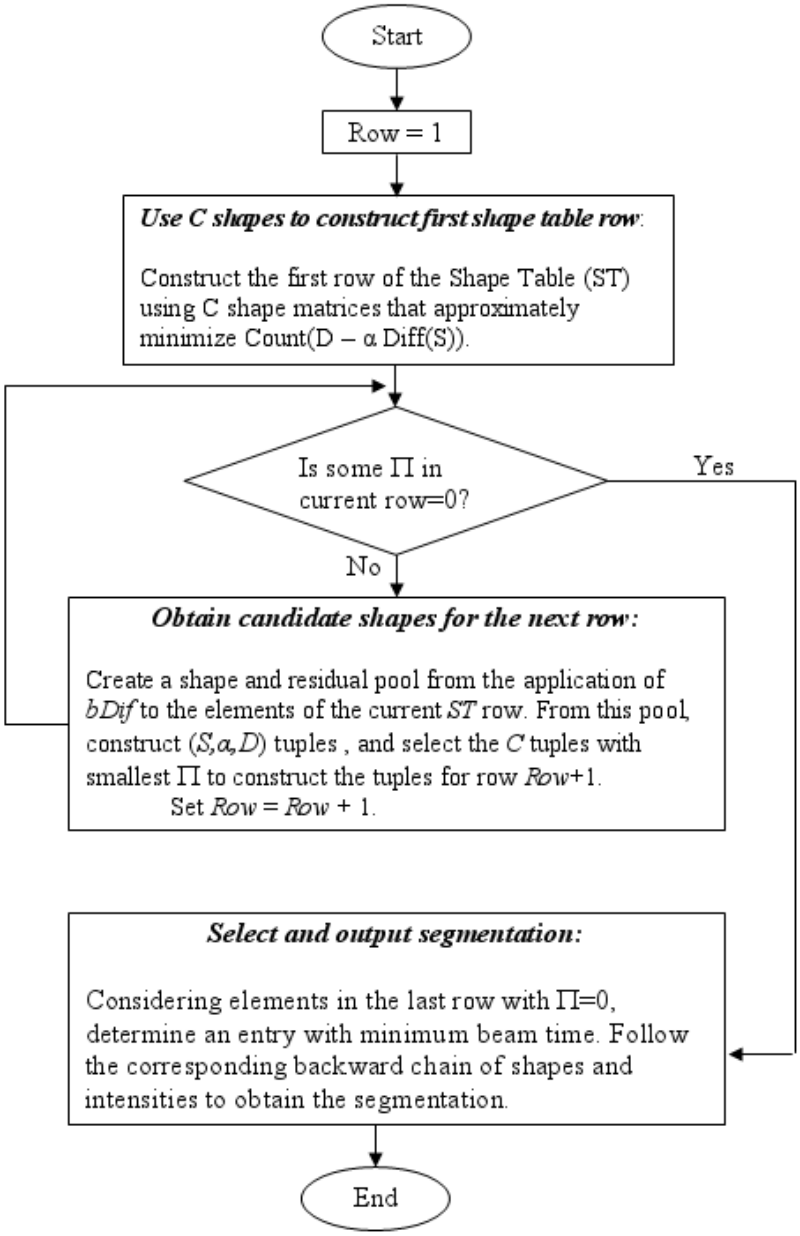


Figure 3. The main procedures of the *DMM* metaheuristic for segmentation.

complex than those considered by Langer). In the latter, 5-intensity-level inputs with values 0, 20, ..., 100 are scaled by dividing by 20 to produce matrices with entries 0, ..., 5 and 10-intensity-level inputs are similarly scaled by dividing by 10. Binary dose variables $d[i, j, t]$ are used to specify if radiation is delivered at bixel i,j at time t . Thus, if such a variable is 1 in a 5-intensity-level case, the scaling dictates that radiation will be delivered for 20 seconds from that bixel.

Langer suggested a two-phase method in which beam-on time is minimized in the first phase, and the number of segments is minimized (subject to a minimum beam-on time constraint) in the second phase. Unfortunately, the additional constraints (as

described below) required to model the Elekta make this MIP approach very problematic for intensity maps with 10 or more intensity levels. For difficult clinical cases, this approach yields relatively poor results even for 5-intensity-level maps. Since the *BC* method is seldom able to produce or verify (within the two hours of computing time that we allowed) an optimal solution for the problem of minimizing the number of shapes (which is more difficult than minimizing beam-on time), it is best regarded in the context of these problems as a heuristic method. The software used to construct and solve the MIP model was the AMPL modeling language combined with the CPLEX branch-and-cut solver (a leading commercial solver), running on 800MHz Linux machines, and the results of the implementation are presented below.

As proposed by Langer, we use binary variables $l[i, j, t]$ that are set to 1 to indicate bixels $[i, j]$ covered by a left leaf at time t, binaries $r[i, j, t]$ for the right leaves, and $d[i, j, t]$ for uncovered bixels, leading to the constraint

$$l[i, j, t] + r[i, j, t] + d[i, j, t] = 1.$$

Standard leaf position constraints may then be modeled using these variables (see Langer (Langer et al. 2001)). In order to model some special properties of the Elekta MLC, we needed to augment Langer's model to include several additional types of constraints, some of which we describe in detail below in order to provide some insight into modeling issues.

Mono-shape constraints require that each aperture consist of only one "connected" shape matrix which satisfies the leaf collision constraints. In order to model these constraints binary variables $delivery[i, t]$ are introduced. A variable of this type is 1 if and only if radiation is being delivered in row i at time t. Therefore, the binary variable *delivery* is forced to be 0 if the i^{th} row is not being used at time t. These properties are enforced by the following constraints, where $Cols$ denotes the number of columns in the intensity map:

$$delivery[i, t] \leq \sum_j d[i, j, t] \leq delivery[i, t] * Cols.$$

Binary variables *drop* are used (beginning with row 2) to determine if the preceding row in a shape is non-zero and the current row is 0. This will allow the determination of the end row of a mono-shape. The constraints below will force *drop* to 1 in this case.

$$delivery[i - 1, t] - delivery[i, t] \leq drop[i, t].$$

Similarly, the binary variable *jump* is used to determine if the previous row is 0 and the current row is non-zero. This will signify the beginning of a mono-shape. The constraints below will force *jump* to 1 in this case:

$$delivery[i, t] - delivery[i - 1, t] \leq jump[i, t].$$

To ensure that the mono-shape constraint holds, there can only be one row where a mono-shape begins. This leads to the constraint:

$$\sum_i jump[i, t] \leq 1.$$

Finally, the following constraint ensure that once the mono-shape ends, all subsequent rows are 0:

$$delivery[i+1,t] \leq 1 - \sum_{I=2,,i} drop[I,t].$$

The following "diagonal" leaf collision constraints are required by the Elekta, but note that the algebraic model must be written so that these leaf constraints will be applied to the leaf positions only in the case of adjacent nonzero rows (in which case the RHS of the constraints will be 1, so that the leaf position variables on the LHS cannot both have value 1):

$$l[i+1,j,t] + r[i,j+1,t] \leq 3 - delivery[i+1,t] - delivery[i,t].$$

$$l[i-1,j,t] + p[i,j+1,t] \leq 3 - delivery[i-1,t] - delivery[i,t].$$

Tongue and groove constraints as specified in Langer's paper were not used in the results reported below, since CPLEX could not determine integer feasible solutions of these more complex problems in twelve hours of computing time.

2.3. Clinical Test Cases and MU calculation

We tested our metaheuristic intensity map segmentation algorithm and the BC implementation on three clinical IMRT cases: (i) head and neck, (ii) pancreas and (iii) prostate. All plans were devised so that the dose per fraction was 1.8 Gy. The head and neck plan consisted of 5 beam angles: 55^o, 165^o, 245^o, 290^o and 350^o. The prescription dose was 45 Gy and the critical structures were the parotid glands, and the spinal cord. The pancreas case plan consisted of 7 beam angles: 0^o, 51^o, 103^o, 154^o, 206^o, 257^o and 308^o. The prescription dose was 63 Gy and the critical structures were the right kidney, spinal cord. The prostate case plan consisted of 6 beam angles: 35^o, 80^o, 135^o, 225^o, 280^o and 325^o. The prescription dose was 75.6 Gy and the critical structures were the bladder and rectum. Plans were designed with 5, 10 and 100 intensity levels and in versions 4.0 and 5.0 of the Corvus treatment planning system. The dose and dose-volume constraints for the critical structures in each of the cases are summarized in Table 3. Plans were normalized so that 97.5% of the head and neck CTV received the prescribed dose while 95% of the CTV in the pancreas and prostate cases received the prescribed dose.

Following dose optimization in the Corvus treatment planning system, the idealized intensity maps corresponding to each beam angle were extracted along with the Corvus MLC controller leaf sequence files for the Elekta SL20 accelerator for each plan. The number of segments obtained from the Corvus (versions 4 and 5), difference matrix metaheuristic and *BC* implementation were compared. Given the relative weighting of each segment within the leaf sequence for a single beam, we calculated the relative "beam-on time" for the Corvus leaf sequence. The absolute MUs were obtained directly from the treatment plan. The idealized intensity maps were then input into our *DMM*

algorithm to generate a new leaf sequence and the leaf segments were output in the format required by the MLC controller file. Using the relative weighting of each segment, we calculated the relative "beam-on time" for the *DMM* leaf sequence. To calculate the absolute MUs required with our leaf sequence, we calculated the ratio between the relative "beam-on time" obtained from the Corvus leaf sequence and that obtained from our leaf sequence for each beam and then multiplied the absolute plan MUs obtained from Corvus by this ratio. An absolute MU comparison is only provided between Corvus 4.0 and the *DMM* algorithm for purposes of brevity.

We compared the realistic dose maps obtained from both the Corvus leaf sequence and the leaf sequence obtained with our approach. The calculation was performed using a previously described convolution/superposition-based dose calculation algorithm (Naqvi, Earl & Shepard 2003) on a 30x30 cm^2 water phantom and at a depth of 2 cm. Following a comparison of the dose maps from each beam angle, we performed a 3-D dose calculation on patient CT data sets with the same convolution/superposition dose calculation algorithm and using both the Corvus leaf sequence and the metaheuristic leaf sequence in order to investigate if there was any degradation in the dose distribution.

Table 2. Summary of dose and dose-volume constraints for the head/neck, pancreas and prostate cases. The same constraints were used in both Corvus version 4.0 and 5.0.

Head/Neck				
Structure	Limit dose (Gy)	% volume > limit dose	Minimum dose	Maximum dose
Parotid (L)	30	20	9	40
Parotid (R)	15	20	1	23
Cord	30	25	2	40

Pancreas				
Structure	Limit dose (Gy)	% volume > limit dose	Minimum dose	Maximum dose
Kidney (R)	25	5	15	35
Cord	20	5	5	30

Prostate				
Structure	Limit dose (Gy)	% volume > limit dose	Minimum dose	Maximum dose
Bladder	45	15	10	70
Rectum	45	50	15	70

3. Results

The results are organized as follows. We first describe the comparison between our metaheuristic leaf sequencing algorithm and the Corvus results. The number of segments and MUs are compared for the plans obtained from version 4.0 of Corvus, while only the number of segments are compared (for purposes of brevity and relevance) for the plans

obtained with version 5.0. Following this comparison we describe the results obtained from the *BC* implementation.

3.1. Corvus, version 4.0 and DM metaheuristic comparison

Tables 3, 4 and 5 show a beam-by-beam comparison of the number of segments and MUs for the head and neck, pancreas and prostate cases respectively obtained with plans generated in Corvus, v4.0. The results are shown for the plans with 5, 10 and 100 intensity levels and are denoted by Corvus4 for the results obtained from v. 4.0 in Corvus, *DMM* for our metaheuristic algorithm, BC30 and BC120 for the *BC* implementation. The decreases in the total number of segments when our leaf sequence is applied in comparison with the leaf sequence available in Corvus4 was 81%, 69% and 50% for 5, 10 and 100 intensity levels in the head and neck case. The corresponding decreases in the number of MUs were 40%, 45% and 32% as a function of the number of intensity levels. The decreases in the number of segments for the pancreas case were 69%, 73% and 44% for 5, 10 and 100 intensity levels. The MUs were decreased by 25%, 30% and 30%, respectively. For the prostate case, the number of segments were decreased by 81%, 69% and 49% as a function of intensity levels. The *DMM* metaheuristic consistently produces high quality (and sometimes optimal) segmentations within 3 minutes, and, in all cases, produces a plan with a number of segments that is smaller than the number produced by Corvus 4.0 and Corvus 5.0.

3.2. Comparison of an "exact" method with DMM

While the branch-and-cut (*BC*) approach, given a sufficiently large amount of computing time (about 2 hours per beam angle), can occasionally produce lower cardinality segmentations than three minute runs of our metaheuristics for 5-intensity-level maps, *BC* is "brittle" in the sense that it sometimes fails to produce any solutions for 5-intensity-level cases (the DNR notation in the tables below stands for "Did Not Run", indicating that no feasible solution was obtained within the time allowed), and *BC* generally fails for 10 (or higher)-intensity-level cases. On the other hand, the *DMM* metaheuristic consistently produces high quality (and sometimes optimal) segmentations within 3 minutes, and, in all cases, produces a plan with a number of segments that is smaller than the number produced by Corvus 4.0 and Corvus 5.0. The column headings *BC*30 and *BC*120 indicate the branch-and-cut method with 30 minutes and 120 minute time limits. *DMM* was allowed a time limit of approximately three minutes. Since Corvus does not allow segmentation to be performed as a separate task, it is difficult to assign a time to the Corvus runs. See also figure 4 for a graphical comparison of segment counts.

Table 6, 7 and 8 summarize the comparison between number of segments produced by Corvus 5.0 and our metaheuristic. The reduction in the number of segments obtained with our algorithm in the head and neck case was 34% 42% and 28% for 5, 10 and 100 intensity levels. In the pancreas case, the decrease in the number of segments with our

Table 3. Method comparisons for a **prostate** case. (DNR indicates that a feasible solution could not be found within the time limit.)

Number of Segments					**Number of Segments**					**Number of Segments**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120
35	41	7	10	10	35	41	12	DNR	DNR	35	48	25	DNR	DNR
80	22	4	4	4	80	32	11	18	15	80	45	19	DNR	DNR
135	40	7	12	12	135	42	13	DNR	DNR	135	49	23	DNR	DNR
225	31	6	9	5	225	33	12	18	18	225	45	28	DNR	DNR
280	23	4	4	4	280	25	6	15	15	280	34	17	DNR	DNR
325	35	8	10	10	325	33	10	DNR	DNR	325	47	24	DNR	DNR
Beam-On-Time					**Beam-On-Time**					**Beam-On-Time**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120
35	346	180	200	200	35	367	260	DNR	DNR	35	405	239	DNR	DNR
80	186	100	100	100	80	334	180	180	150	80	220	160	DNR	DNR
135	321	160	240	240	135	402	240	DNR	DNR	135	290	192	DNR	DNR
225	375	140	180	180	225	415	200	180	180	225	406	280	DNR	DNR
280	224	120	120	120	280	224	120	150	150	280	233	144	DNR	DNR
325	430	220	200	200	325	391	180	DNR	DNR	325	295	220	DNR	DNR

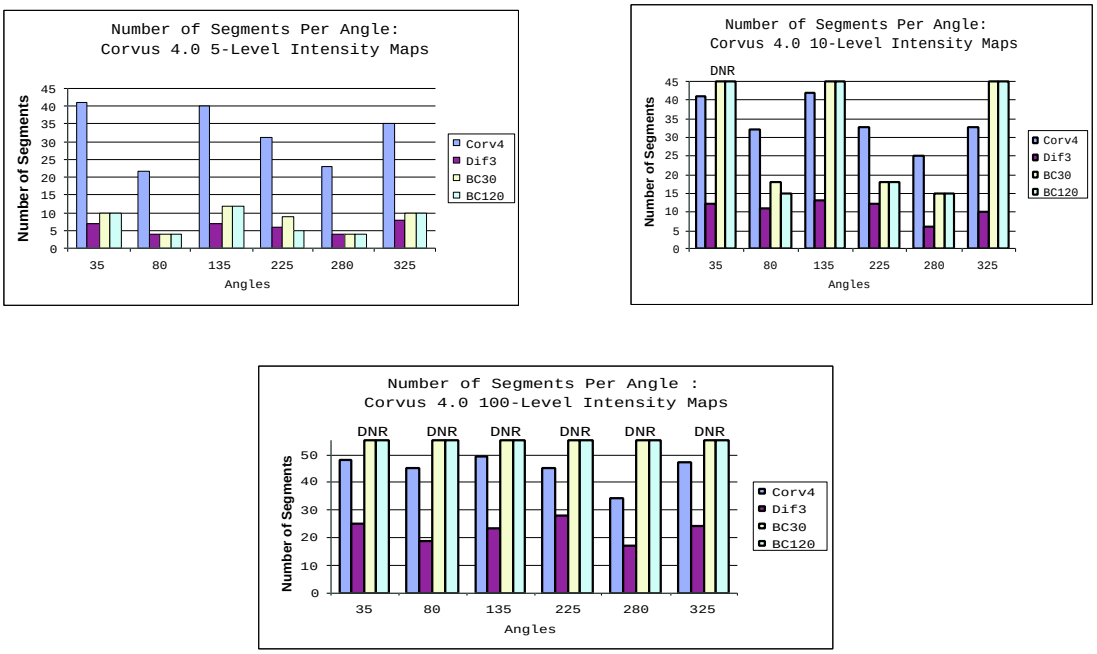


Figure 4. Comparison of the number of segments for the Corvus 4.0 **prostate** intensity maps.

Table 4. Method comparisons for a challenging **head and neck** case. (DNR indicates that a feasible solution could not be found within the time limit.)

Number of Segments					**Number of Segments**					**Number of Segments**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120
55	44	6	9	7	55	49	13	23	23	55	55	31	DNR	DNR
165	43	10	DNR	DNR	165	41	12	18	18	165	56	28	DNR	DNR
245	35	9	8	6	245	47	13	DNR	DNR	245	58	25	DNR	DNR
290	33	7	9	9	290	44	8	17	17	290	50	24	DNR	DNR
350	46	8	DNR	10	350	50	13	DNR	DNR	350	62	32	DNR	DNR
Beam-On-Time					**Beam-On-Time**					**Beam-On-Time**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120
55	382	160	180	140	55	391	190	230	230	55	387	233	DNR	DNR
165	292	300	DNR	DNR	165	278	210	180	180	165	313	255	DNR	DNR
245	381	200	160	140	245	377	300	DNR	DNR	245	236	173	DNR	DNR
290	342	180	180	180	290	308	160	170	170	290	274	158	DNR	DNR
350	347	200	DNR	200	350	479	220	DNR	DNR	350	436	303	DNR	DNR

Table 5. Method comparisons for a difficult **pancreas** case. (DNR indicates that a feasible solution could not be found within the time limit.)

Number of Segments					**Number of Segments**					**Number of Segments**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120
0	62	20	DNR	DNR	0	74	19	DNR	DNR	0	96	51	DNR	DNR
51	62	15	DNR	DNR	51	81	18	DNR	DNR	51	90	50	DNR	DNR
103	45	21	DNR	14	103	54	18	DNR	DNR	103	69	39	DNR	DNR
154	51	16	DNR	DNR	154	67	22	DNR	DNR	154	81	57	DNR	DNR
206	63	26	DNR	DNR	206	88	21	DNR	DNR	206	97	65	DNR	DNR
257	45	10	DNR	DNR	257	59	16	DNR	DNR	257	82	38	DNR	DNR
308	53	8	DNR	8	308	63	10	DNR	16	308	75	37	DNR	DNR
Beam-On-Time					**Beam-On-Time**					**Beam-On-Time**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120	Angle	Corv4	DMM	BC30	BC120
0	526	440	DNR	DNR	0	370	320	DNR	DNR	0	481	408	DNR	DNR
51	541	340	DNR	DNR	51	580	340	DNR	DNR	51	542	400	DNR	DNR
103	488	440	DNR	280	103	380	300	DNR	DNR	103	421	264	DNR	DNR
154	474	360	DNR	DNR	154	380	400	DNR	DNR	154	477	421	DNR	DNR
206	674	580	DNR	DNR	206	430	410	DNR	DNR	206	685	420	DNR	DNR
257	392	220	DNR	DNR	257	250	250	DNR	DNR	257	423	206	DNR	DNR
308	349	160	DNR	160	308	150	140	DNR	160	308	306	189	DNR	DNR

Table 6. : Method comparison for a **prostate** case. Dif generates the optimal solution in those cases in which BC is able to establish an optimal solution. (DNR indicates that a feasible solution could not be found within the time limit.)

Number of Segments					**Number of Segments**					**Number of Segments**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv5	DMM	BC30	BC120	Angle	Corv5	DMM	BC30	BC120	Angle	Corv5	DMM	BC30	BC120
35	7	4	4	4	35	24	11	DNR	15	35	33	23	DNR	DNR
80	6	5	5	5	80	16	9	14	14	80	36	21	DNR	DNR
135	6	4	4	4	135	17	11	18	18	135	37	26	DNR	DNR
225	8	5	5	5	225	20	10	DNR	DNR	225	37	24	DNR	DNR
280	7	4	4	4	280	19	7	12	12	280	32	23	DNR	DNR
325	6	4	4	4	325	24	10	18	18	325	33	27	DNR	DNR

Table 7. Method comparisons for a challenging **head and neck** case. These results illustrate that the BC approach has difficulty with more complex cases.

Number of Segments					**Number of Segments**					**Number of Segments**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv5	DMM	BC30	BC120	Angle	Corv5	DMM	BC30	BC120	Angle	Corv5	DMM	BC30	BC120
55	11	8	7	7	55	22	15	DNR	DNR	55	36	28	DNR	DNR
165	16	12	DNR	DNR	165	23	12	DNR	DNR	165	48	33	DNR	DNR
245	14	9	DNR	9	245	19	13	DNR	14	245	35	29	DNR	DNR
290	11	8	6	6	290	16	8	7	6	290	40	25	DNR	DNR
350	19	10	13	9	350	23	12	DNR	14	350	42	29	DNR	DNR

Table 8. Method comparisons for a difficult **pancreas** case.

Number of Segments					**Number of Segments**					**Number of Segments**				
5-Intensity-Levels					**10-Intensity-Levels**					**100-Intensity-Levels**				
Angle	Corv5	DMM	BC30	BC120	Angle	Corv5	DMM	BC30	BC120	Angle	Corv5	DMM	BC30	BC120
0	23	15	DNR	DNR	0	34	17	DNR	DNR	0	62	53	DNR	DNR
51	21	13	DNR	DNR	51	40	25	DNR	DNR	51	67	49	DNR	DNR
103	12	11	DNR	DNR	103	28	13	DNR	DNR	103	63	46	DNR	DNR
154	17	13	DNR	DNR	154	29	18	DNR	DNR	154	55	50	DNR	DNR
206	25	13	DNR	DNR	206	41	23	DNR	DNR	206	72	55	DNR	DNR
257	21	11	10	9	257	31	17	DNR	DNR	257	52	45	DNR	DNR
308	16	11	DNR	12	308	28	15	DNR	DNR	308	56	40	DNR	DNR

approach was 36%, 45%, 21% while in the prostate case, the corresponding decrease was 35%, 59% and 31% as a function of intensity levels. In general, a greater reduction was seen in the 10-intensity level cases than with 5 or 100 intensity for version 5.0. The improvement in the number of segments using our approach was less dramatic than with version 4.0. This may be attributed to the fact that the leaf-sequencing is more efficient in version 5.0 and hence the improvement achievable while still significant, is reduced.

3.3. Calculated dose comparison

Figures 5, 6 and 7 show a comparison of the calculated dose from the intensity maps in a water phantom for the head and neck, pancreas and prostate cases and a representative beam angle. The dose calculation (Naqvi et al 2003) is performed at a depth of 2 cm. As can be seen from these figures, the calculated dose maps agree well for all three cases and intensity levels. However, there are two noticeable differences between the dose maps calculated using the *DMM* and the Corvus v. 4.0 algorithms. First, tongue-and-groove effects are seen in the *DMM* leaf sequence and minimized in the Corvus leaf sequence. This is because the Corvus leaf sequence forces the leaves to move in one direction only during step-and-shoot delivery. Second, noticeable leakage is visible on the Corvus dose maps due to the partial transmission through backup Y-diaphragm present in the Elekta SL20 linac.

Despite these differences in the dose maps the quantity of real consequence is the 3-D dose distribution. Figures 8, 9, and 10 show the comparison between the 3-D dose distribution from the step-and-shoot delivery sequence using the *DMM* and Corvus v4.0 algorithms for the head and neck, pancreas and prostate cases respectively. It is seen that an overlay of the dose distributions show that the isodose lines agree very well. Minor discrepancies in the isodose line comparison were not considered clinically significant.

4. Discussion

The computational results presented above demonstrate that for a variety of difficult clinical cases the *DMM* metaheuristic is able to achieve a dramatic reduction in aperture count relative to the procedures used in the Corvus 4 commercial treatment planning software. For example, the number of apertures generated by *DMM* for the 5-intensity-level prostate intensity maps is generally less than 20% of the number generated by Corvus 4. The aperture reduction ratio is about 30% for the 10-intensity-level map and about 50% for the 100-intensity-level maps. Similar reduction ratios are obtained for the head-and-neck and pancreas cases. The beam-on times for *DMM* (as measured by summing intensities over all apertures) are also typically about 50% of those of Corvus 4. For the less differentiated (flatter and hence simpler) intensity maps associated with Corvus 5 for the same clinical cases, the *DMM* aperture count improvements are generally not as dramatic, but are approximately 50% for the 10-intensity-level prostate and 70-80% for the other levels. The results obtained for the CPLEX implementation of the branch-and-cut (*BC*) approach demonstrate that this method is capable of generating results comparable to DMM for the simplest cases (such as some of the 5-intensity-level maps), but *BC* generally fails to produce any useful results for cases of medium to high difficulty (in terms of either number of intensity levels or variability within the intensity map).

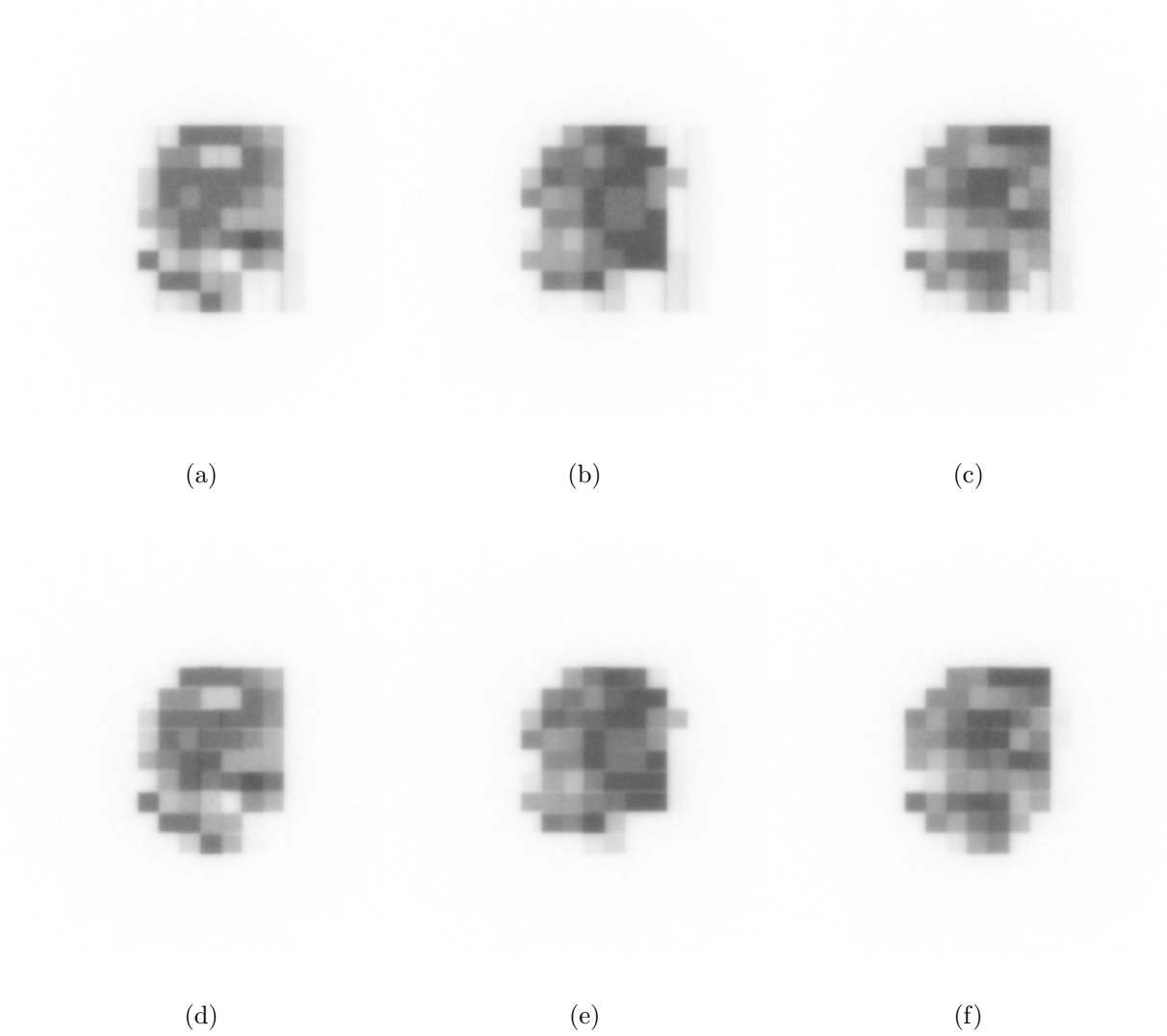


Figure 5. Head/Neck case: comparison of dose maps calculated in a water phantom at a depth of 2 cm. (a), (b) and (c) are the dose maps from Corvus, v4.0 for 5, 10 and 100 intensity levels. (d), (e) and (f) are the corresponding dose maps obtained using our metaheuristic approach

5. Conclusion

We have demonstrated that a metaheuristic based on difference matrices can outperform both commercial treatment planning systems and "exact" integer programming approaches to intensity map segmentation. Commercial systems are fast, but often produce segmentations with large numbers of apertures, whereas the branch-and-cut method for integer programming seeks an optimal segmentation, but in many cases has difficulty constructing such a result within a clinically acceptable time frame. Our difference-matrix-based heuristics are reliable and fast, and yield segmentations that are comparable to or match the best results generated by two-hour runs of the branch-and-

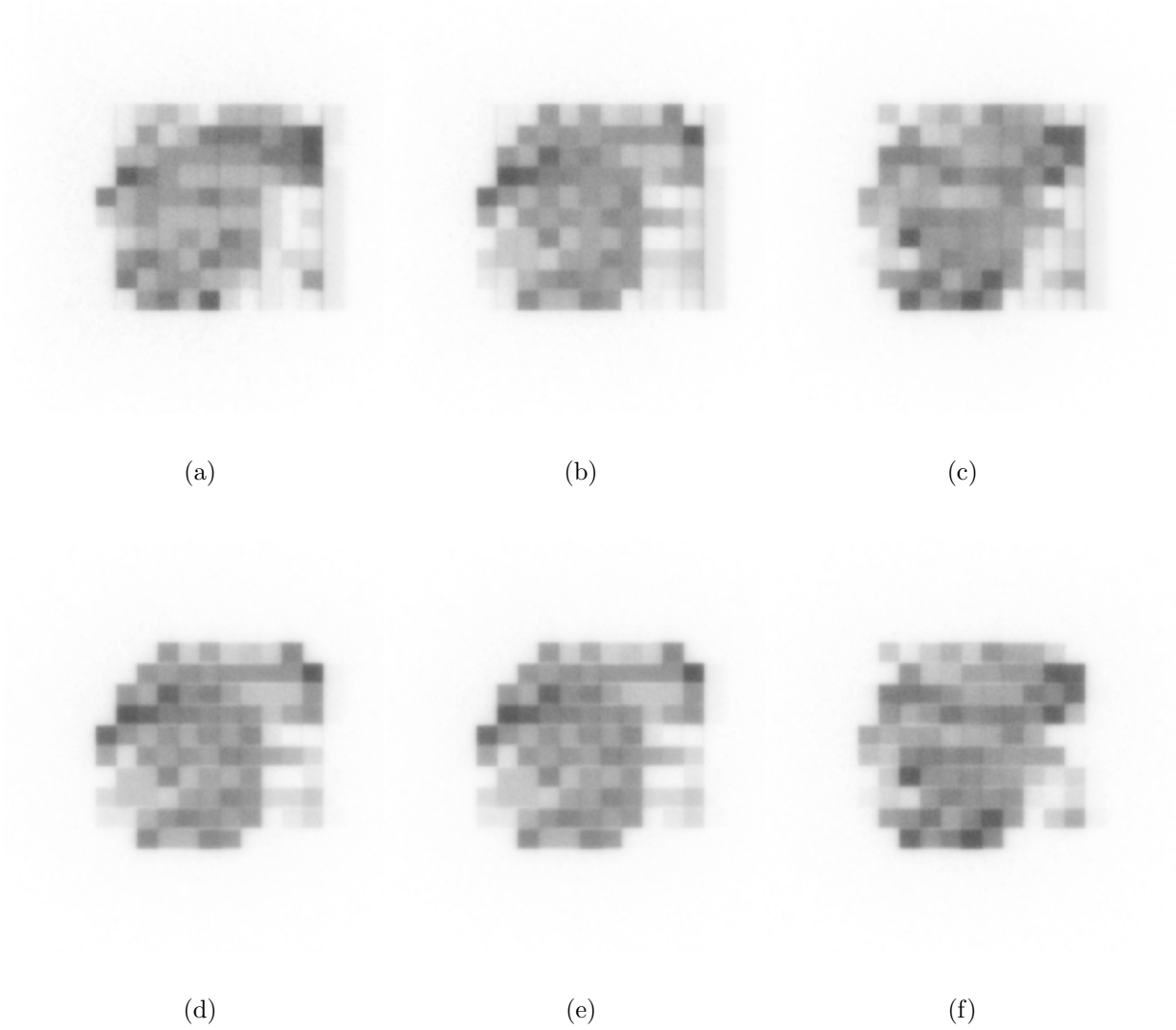


Figure 6. Pancreas case: comparison of dose maps calculated in a water phantom at a depth of 2 cm. (a), (b) and (c) are the dose maps from Corvus, v4.0 for 5, 10 and 100 intensity levels. (d), (e) and (f) are the corresponding dose maps obtained using our metaheuristic approach

cut approach. We also validate the clinical quality of the segmentations that we obtain by comparisons with the results of commercial treatment planning systems.

Acknowledgments

We would like to acknowledge the financial support from NSF grants DMI-0100220 and DMI- 0400294.

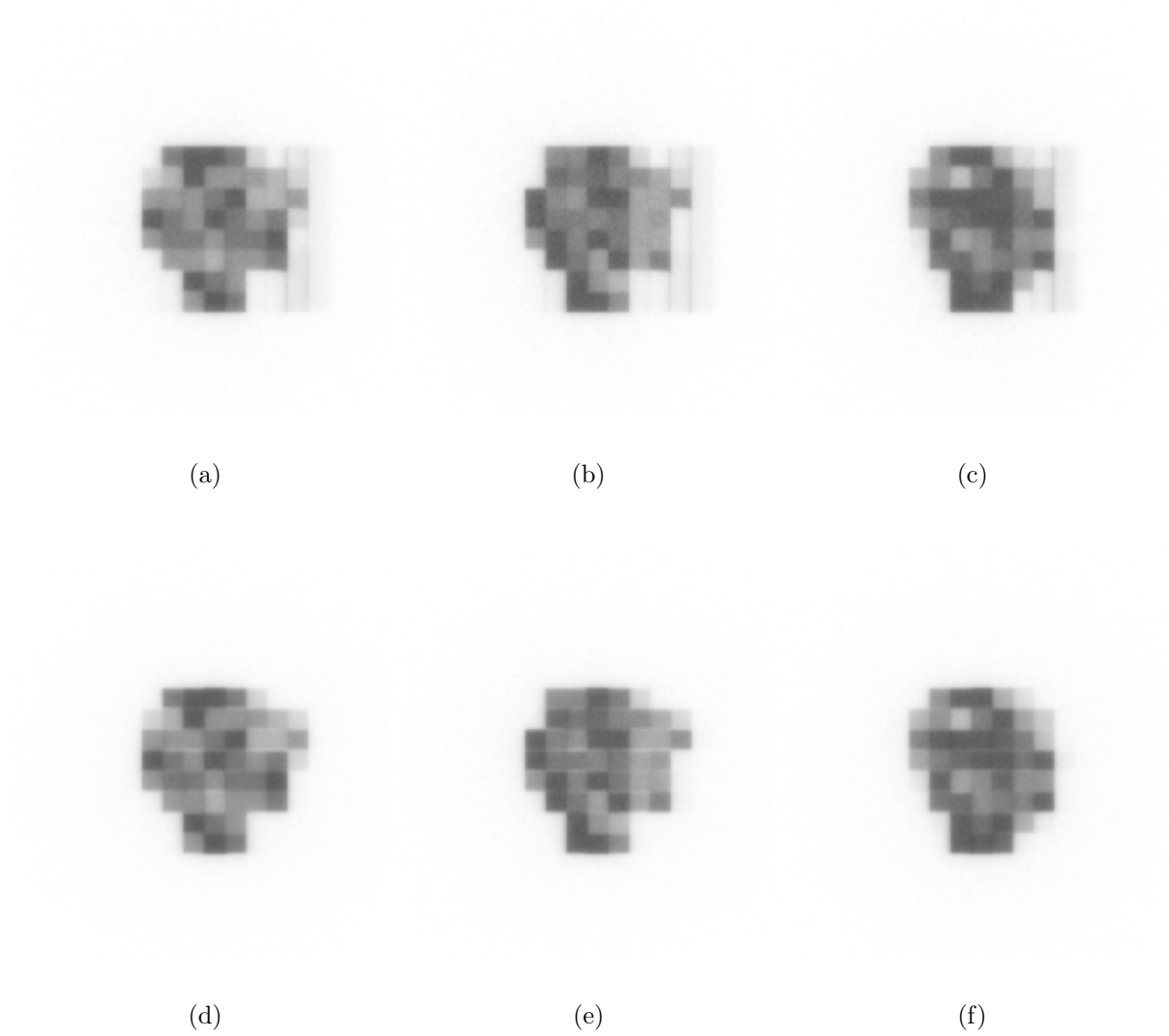


Figure 7. Prostate case: comparison of dose maps calculated in a water phantom at a depth of 2 cm. (a), (b) and (c) are the dose maps from Corvus, v4.0 for 5, 10 and 100 intensity levels. (d), (e) and (f) are the corresponding dose maps obtained using our metaheuristic approach

References

Boland, N., Hamacher, H. W. & Lenzen, F. (2004), 'Minimizing beam-on time in cancer radiation treatment using multileaf collimators', *Networks* **43**(4), 226–240.

Bortfeld, T. R., Kahler, D. L., Waldron, T. J. & Boyer, A. L. (1994), 'X-ray field compensation with multileaf collimators', *Int. J. Radiat. Oncol., Biol., Phys.* **28**, 723–730.

Boyer, A. L. & Yu, C. Y. (1999), 'Intensity-modulated radiation therapy with dynamic multileaf collimators', *Semin Radiat. Oncol.* **9**, 48–59.

Engel, K. (2003), A new algorithm for optimal multileaf collimator field segmentation, Technical Report TR D-18051, Rostock, Germany.

Evans, P. M., Hansen, V. N. & Swindell, W. (1997), 'The optimum intensities for multiple static multileaf collimator field compensation', *Medical Physics* **24**, 1147–1156.

Galvin, J. M., Chen, X. G. & Smith, R. M. (1993), 'Combining multileaf fields to modulate fluence

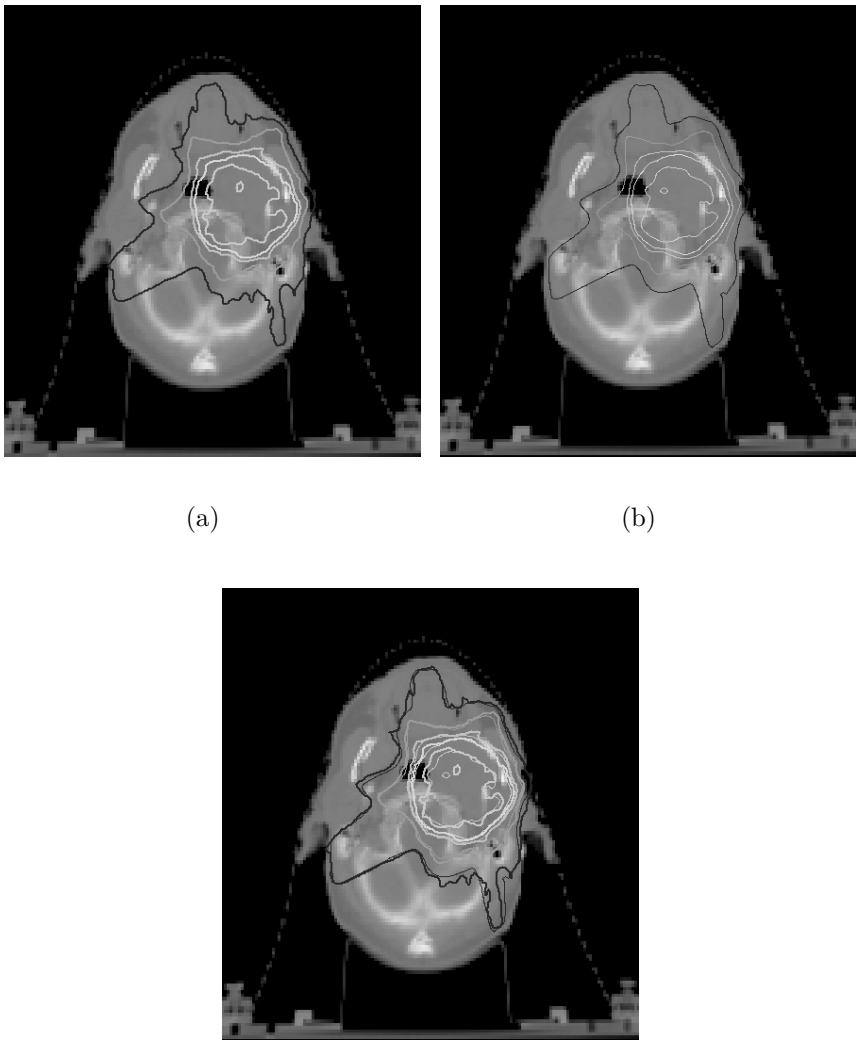


Figure 8. Head/Neck case: 3D dose distribution obtained from (a) original Corvus, v4.0 and (b) difference matrix based leaf sequence. (c) Overlay of dose distributions from Corvus and difference matrix leaf-sequences.

distributions', *Int. J. Radiat. Oncol., Biol., Phys.* **27**, 607–705.

Langer, M., Thai, V. & Papiez, L. (2001), 'Improved leaf sequencing reduces segments or monitor units needed to deliver imrt using multileaf collimators', *Medical Physics* **28**, 2450–2458.

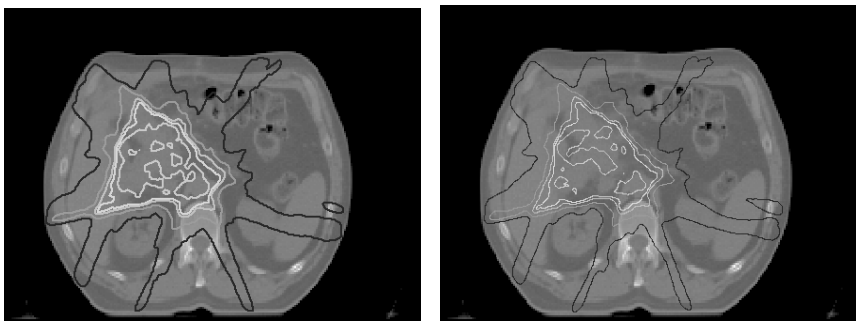
Luan, S., Wang, C., Hu, X. S., Naqvi, S. A., Yu, C. X. & Lee, C. L. (2004), 'A new mlc segmentation algorithm/software for step-and-shoot imrt delivery', *Medical Physics* **31**(4), 695–707.

Naqvi, S. A., Earl, M. A. & Shepard, D. M. (2003), 'Convolution/superposition using the monte carlo method', *Physics in Medicine and Biology* **48**(14), 2101–2121.

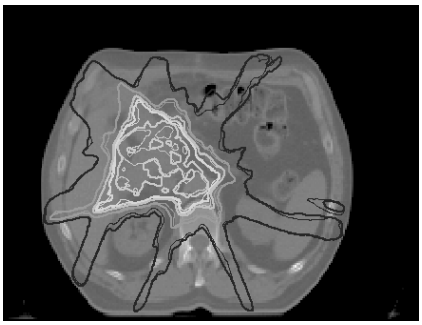
Que, W. (1999), 'Comparison of algorithms for multileaf collimator field segmentation', *Medical Physics* **26**, 2390–2396.

Siochi, R. A. (1999), 'Minimizing static intensity modulation delivery time using an intensity solid paradigm', *Int. J. Radiat. Oncol., Biol., Phys* **43**, 671–680.

Xia, P. & Verhey, L. J. (1998), 'Multileaf collimator leaf sequencing algorithm for intensity modulated



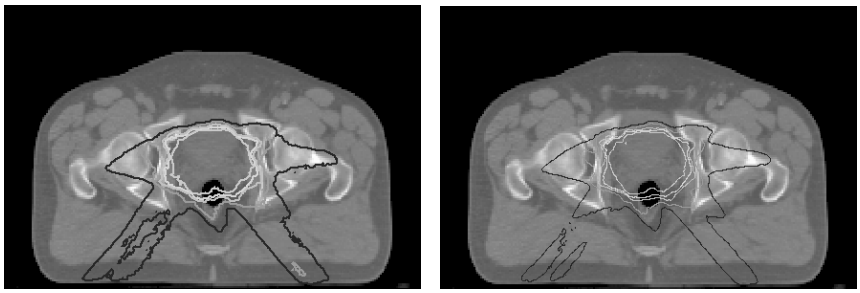
(a) (b)



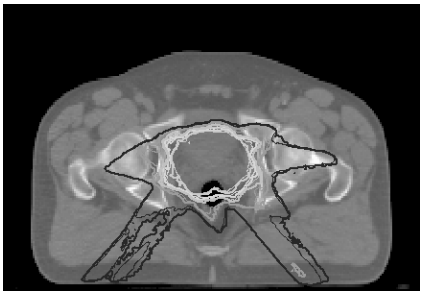
(c)

Figure 9. Pancreas case: 3D dose distribution obtained from (a) original Corvus, v4.0 and (b) difference matrix based leaf sequence. (c) Overlay of dose distributions from Corvus and difference matrix leaf-sequences.

beams with multiple static segments', *Medical Physics* **25**, 1424–1434.



(a) (b)



(c)

Figure 10. Prostate case: 3D dose distribution obtained from (a) original Corvus, v4.0 and (b) difference matrix based leaf sequence. (c) Overlay of dose distributions from Corvus and difference matrix leaf-sequences.