

Inverse treatment planning for Gamma Knife radiosurgery

D. M. Shepard^{a)}

Department of Radiation Oncology, University of Maryland School of Medicine, 22 South Greene Street, Baltimore, Maryland 21201-1595

M. C. Ferris

Computer Sciences Department, University of Wisconsin-Madison, 1210 West Dayton Street, Madison, Wisconsin 53706

R. Ove and L. Ma

Department of Radiation Oncology, University of Maryland School of Medicine, 22 South Greene Street, Baltimore, Maryland 21201-1595

(Received 28 April 2000; accepted for publication 3 October 2000)

An inverse treatment planning system for Gamma Knife radiosurgery has been developed using nonlinear programming techniques. The system optimizes the shot sizes, locations, and weights for Gamma Knife treatments. In the patient's prescription, the user can specify both the maximum number of shots of radiation and a minimum isodose line that must surround the entire treatment volume. After satisfying all of the constraints included in the prescription, the system maximizes the conformity of the dose distribution. This automated approach to treatment planning has been applied retrospectively to a series of patient cases, and each optimized plan has been compared to the corresponding manual plan produced by an experienced user. The results demonstrate that this tool can often improve the tumor dose homogeneity while using fewer shots than were included in the original plan. Therefore, inverse treatment planning should improve both the quality and the efficiency of Gamma Knife treatments. © 2000 American Association of Physicists in Medicine. [S0094-2405(00)01412-7]

Key words: Gamma Knife, inverse treatment planning, radiosurgery

I. INTRODUCTION

The Leksell Gamma Knife provides an advanced stereotactic approach to the treatment of tumors and vascular malformations within the head.¹⁻⁴ Over 100 Gamma Knife units are installed worldwide, and more than 20 000 patients are treated each year.⁵

Each patient's treatment plan is developed by a neurosurgeon working in conjunction with a radiation oncologist and a physicist. Through an iterative process of trial and error, they determine the number of shots along with their size, location, and weight. When the treatment volume is small, the treatment plan may only require one or two shots of radiation. For these cases, treatment planning is typically straightforward. The planning process, however, becomes much more complex for both irregularly shaped tumors and tumors that are large in size. For these cases, the complexity of the treatment planning process makes it difficult to take full advantage of the powerful capabilities of the Gamma Knife. Therefore, we have sought to automate the Gamma Knife treatment planning process.

We have examined a variety of techniques in hopes of developing an automated planning approach that is fast, flexible, and robust. The system must be fast, because the treatment planning process cannot begin until after the patient's stereotactic head frame is in place and MRI or CT images have been obtained. It is therefore unacceptable to have a planning procedure that takes hours to complete. Our goal has been to create an automated planning system that can

produce a treatment plan within 20 min. Flexibility in defining the treatment goals is also of key importance, because each plan must be tailored to meet the individual needs of the patient. The system must also be flexible enough to satisfy the preferences of a variety of treatment planners. Finally, the system must be robust. A robust system is one that produces a high quality solution for treatment volumes of all sizes and shapes. The ideal technique should provide a high quality result independent of the quality of the starting point used in the optimization.

A number of researchers have investigated techniques for automating the Gamma Knife treatment planning process.⁶⁻¹² One approach incorporates the assumption that each shot of radiation can be modeled as a rigid sphere. The problem is then reduced to one of geometric coverage, and a ball packing approach can be used to determine the shot locations and sizes.^{6,12} The use of a modified Powell's method in conjunction with simulated annealing has also been proposed.⁸⁻¹¹

This paper addresses a different approach where the dose distribution is modeled and a formal constrained optimization is used to determine the treatment plan. With this technique, the shot sizes, locations, and weights are optimized simultaneously. The optimization does not require the user to provide initial shot locations, and the optimization model can include dose constraints applied to both the target and the sensitive structures. The dose model used in this technique was generated using a nonparametric optimization approach.

The treatment plan optimization is based on the use of migrating shot locations and a nonlinear programming approach. The effectiveness of this technique was initially examined using a two-dimensional testing environment. A series of optimizations have also been performed using real three-dimensional patient data. This paper introduces the dose model and the optimization techniques. It also provides comparisons between optimized patient plans and the corresponding plans produced manually by a neurosurgeon.

Our constrained optimization program was developed simultaneously with a simpler algorithm using a least-squares objective function. Both approaches incorporate a migrating shot formulation. Although the least-squares model is not as extensible, it serves as a useful benchmark for evaluating the constrained optimization approach.

II. MATERIALS AND METHODS

A. Dose model

For each patient, the optimization uses an algebraic model in order to determine the dose delivered at each calculation point. Calculation points are located on a three-dimensional grid that is applied in the area of interest. In the optimization model, *S* represents the set of the shots and *W* represents the possible shot sizes (4, 8, 14, and 18 mm). The complete dose distribution can be calculated as the sum of the contributions from each shot. The dose calculation requires the location of each shot (*x_s, y_s, z_s*) and the relative weight (*a_{s,w}*) that is assigned to each. The weight is proportional to the treatment time for the given shot. In practice, this means that the dose, *D*, at a point (*x, y, z*) is

$$D(x,y,z) = \sum_{(s,w) \in S \times W} a_{s,w} D_w(x_s, y_s, z_s, x, y, z), \tag{1}$$

where *D_w(x_s, y_s, z_s, x, y, z)* is the dose delivered to the calculation point (*x, y, z*) by a shot of size *w* centered at (*x_s, y_s, z_s*).

In order to determine the form of *D_w*, the treatment planning system, Leksell GAMMAPLAN (version 5.20), was used to simulate the delivery of a shot of width *w* placed at the center a patient's stereotactic head frame. For each shot width, dose profiles in the *x*, *y*, and *z* directions were extracted. These profiles were then averaged as follows:

$$\bar{D}_w(d) = \frac{D_w(0,0,0,d,0,0) + D_w(0,0,0,0,d,0) + D_w(0,0,0,0,0,d)}{3} \tag{2}$$

The center of the patient's head frame was defined as the origin. The values that were obtained were used as data in a nonlinear parameter estimation problem.

The dose distribution for each shot of radiation was assumed to be spherically symmetric. Consequently, the dose modeling problem was reduced to one of determining a functional form for the dose delivered at a calculation point located a distance *d* from the center of the shot. The dose

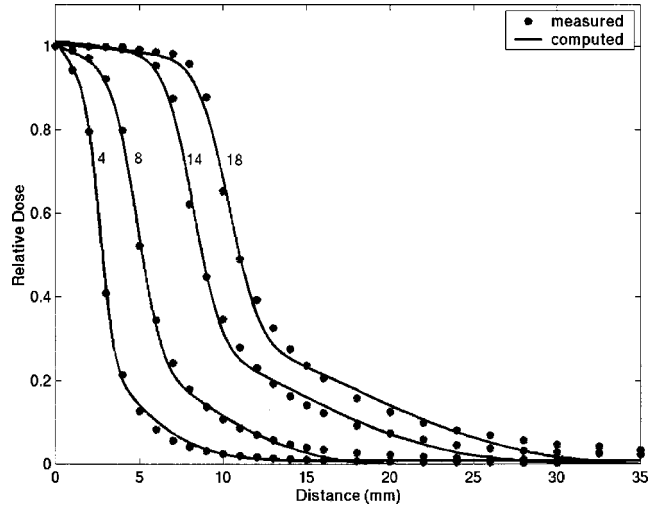


FIG. 1. A comparison between dose profiles extracted from the Leksell GAMMAPLAN system and the profiles from dose distributions predicted by the two error function model.

model we have used incorporates a sum of error functions.¹³ The error function can be thought of as a point source (*δ* function) convolved with a finely collimated aperture function. The following functional form was used with *d* representing the distance to the center of the shot:

$$\sum_{i=1}^2 \lambda_i \left(1 - \operatorname{erf} \left(\frac{d-r_i}{\sigma_i} \right) \right), \tag{3}$$

where

$$\operatorname{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2} dx. \tag{4}$$

For each of the four shot sizes, the six parameters *λ_i*, *r_i*, and *σ_i* were fit to the data via a least-squares approach. The resulting nonlinear optimization problem is

$$\min_{\lambda, r, \sigma} \left\| \bar{D}_w(d) - \sum_{i=1}^2 \lambda_i \left(1 - \operatorname{erf} \left(\frac{d-r_i}{\sigma_i} \right) \right) \right\|. \tag{5}$$

This problem was solved using the nonlinear optimization algorithm CONOPT.¹⁴⁻¹⁷ The results for all four shot sizes are shown in Fig. 1. The fit is best for the small shot sizes and decreases slightly in accuracy for the larger shot sizes. The values of the parameters that we generated are provided in Table I.

B. Constrained optimizations

Constrained optimization models were written in the modeling language of GAMS (the General Algebraic Modeling System).¹⁸ GAMS is a high-level modeling system for mathematical programming problems. It consists of a language compiler and a variety of integrated high-performance solvers. GAMS was chosen for the optimizations based on its versatility and its state-of-the-art optimization algorithms. In particular, the formulations that we describe are constrained nonlinear programming problems, and as such are solved

TABLE I. The optimized dose parameters for the dose model outlined by Eq. (3).

Shot (mm)	λ_1	r_1	σ_1	λ_2	r_2	σ_2
4	0.649 200	1.365 916	4.413 680	0.599 844	2.661 771	0.668 291
8	0.401 007	7.035 785	5.702 337	0.648 584	4.849 365	1.149 176
14	0.363 704	13.972 59	7.196 694 0	0.657 808	8.199 979	1.321 161
18	0.381 801	17.678 57	8.194 611	0.634 696	10.315 83	1.441 725

using a state-of-the-art optimization algorithm CONOPT.^{14–17} For each of the nonlinear functions in our model, CONOPT uses first-order derivative information to determine a search direction, which is then investigated using a classical line search technique.

The actual algorithm in CONOPT is based upon a generalized reduced gradient (GRG) approach to optimization. A critical feature of CONOPT is that it maintains a feasible solution throughout the optimization once a feasible solution has been obtained. The user can therefore choose to terminate the optimization prematurely with assurance that the solution provided will satisfy all of the constraints. In order to allow for very large models, all matrix operations are implemented using sparse matrix techniques. The overhead of the GRG algorithm is minimized through a variety of techniques including dynamic feasibility tolerances, reuse of Jacobians, and an efficient reinversion routine.

Each constrained optimization incorporates both an objective function and a set of constraints. The objective functions that we have tested include: (1) maximizing the conformality of the dose distribution, (2) minimizing the number of shots, and (3) minimizing the dose to a sensitive structure such as the brain stem or optic chiasm.

Each optimization also includes a series of constraints. A constraint is a condition that must be satisfied in order for a solution to be considered feasible. Constraints have been used to specify each of the following: (1) the maximum number of shots, (2) a minimum isodose line that must surround the entire tumor volume, and (3) a maximum dose for each sensitive structure.

In the first phase of the optimization, the optimizer seeks a feasible solution that satisfies the constraints. After a feasible solution is obtained, the optimizer minimizes or maximizes the objective function. If a user is unsatisfied with the optimized plan, he or she can update the treatment goals and perform an additional optimization. The solution from the first optimization should provide a high quality starting point for the second optimization. Therefore, the time requirements for the second optimization are minimized.

We have investigated two constrained optimization approaches, namely mixed integer programming and nonlinear programming. The mixed integer programming approach has been addressed elsewhere.¹⁹ This paper will describe the use of a nonlinear programming approach to optimizing Gamma Knife treatment plans.

The nonlinear programming approach that we have implemented incorporates migrating shot locations. In other words, the position of each shot can change over the course of the optimization. At each shot location (x_s, y_s, z_s) , a rela-

tive weight $(a_{s,w})$ is assigned for each of the four available shots. The dose computation includes a grid of calculation points, and each point is categorized as either target (T) or normal tissue (NT).

A key feature of the optimization approach presented here is that the user can specify a minimum isodose line that must surround the entire treatment volume. For Gamma Knife treatments, the neurosurgeon often requires that the entire target is encompassed by the 50% isodose line. (The 50% isodose line is defined relative to the maximum dose.) We model such a constraint by imposing strict lower and upper bounds on the dose allowed in the target. Namely, for all (i, j, k) in T

$$0.5 \leq D(i, j, k) \leq 1.$$

This constraint guarantees that the dose throughout the tumor is maintained between a value of 0.5 and 1. Consequently, all of the points in the tumor must receive a dose greater than 50% of the maximum tumor dose. For our Gamma Knife prescriptions, the user specifies the isodose curve that must encompass the tumor. It is only after the plan has been reviewed that the absolute dose in Gray is prescribed. Note that with this formulation it is easy to change the minimum isodose requirement by modifying the numerical values of the bounds.

The general optimization approach that we used is maximize

$$\frac{\int_{(x,y,z) \in T} D(x,y,z) dx dy dz}{\int_{(x,y,z)} D(x,y,z) dx dy dz}$$

subject to

$$D(x,y,z) = \sum_{(s,w) \in S \times W} a_{s,w} D_w(x_s, y_s, z_s, x, y, z),$$

$$0.5 \leq D(x,y,z) \leq 1, \quad \forall (x,y,z) \in T, \quad (6)$$

$$a_{s,w} \geq 0.$$

The goal of the optimization is to maximize the percentage of the integral dose that is deposited in the target subject to a constraint that the 50% isodose line must surround the target. The maximum number of shots that can be assigned nonzero weights is constrained to be less than or equal to N . In this formulation, the percentage of integral dose deposited in the target serves as a measure of the conformity of the dose distribution. Spatial integrals, represented by $\int dx dy dz$, are taken over all volumes of interest and are implemented as a simple sum over discrete points.

An important feature of our nonlinear programming approach is that the user does not need to preselect the size of each shot. In our formulation, a shot of width w is used at location s if $a_{s,w} \geq 0$. Ideally, the optimization would include a binary variable assigned to each shot of radiation. This binary variable would serve as an on/off switch to indicate whether or not a shot has been assigned a nonzero time [$H(t) = 1$ if $t > 0$ and $H(t) = 0$ when $t = 0$]. The optimization would then include a constraint that the sum of all the binary variables must be less than the maximum number of allowable shots.

The use of a true step function is disallowed in a nonlinear programming formulation. Therefore, we have approximated the step function using the following approach:

$$H(\alpha, t) = \frac{2 \arctan(\alpha t)}{\pi}, \tag{7}$$

where $\lim_{\alpha \rightarrow \infty} H(\alpha, t) = H(t)$. Therefore, the condition

$$\sum_{(s,w) \in \{1, \dots, q\} \times W} H(\alpha, t_{s,w}) \leq N \tag{8}$$

is added to the constraint set, where q represents the number of migrating shot locations. Typically, we include more migrating shot locations than the number of allowable shots ($q > N$). This increases the search space and provides improved robustness.

Based on our experience, the best speed and robustness are seen when a two-step optimization is used. An initial optimization is performed with an α value of ten. A second optimization is then performed with an α value of one hundred. The constraint on the total number of shots becomes more absolute in nature with increasing values of α . Higher α values, however, lead to a more nonlinear formulation that tends to be more difficult to solve.

C. Least-squares formulation

Our initial attempt at formulating a nonlinear migrating shot model used a simple least-squares formulation. This approach is based on a variational principle that allows explicit generation of optimization equations.

With this method, a prescribed dose $P(x, y, z)$ was set at each of the calculation points. $P(x, y, z)$ may have a complex shape. For example, one might choose to conform to areas of hypoxia in the target. N individual shot distributions are considered as a set of N non-orthonormal basis functions ψ_n , $n \in \{1, \dots, N\}$. These basis functions are functions of the shot locations as well as the spatial coordinates. Together with the relative weights (a_n), the shot locations form the set of optimization parameters. Explicitly, $\psi_n \equiv D_w(x_n, y_n, z_n, x, y, z)$. The minimization problem is defined in terms of least-squares deviation from the prescribed dose by

$$S = \int_{(x,y,z)} M \left(\sum_n a_n \psi_n - P \right)^2 dx dy dz. \tag{9}$$

A weighting factor, $M(x, y, z)$, is assigned to each region of interest (target, critical structure, etc.). The weighting factor defines the importance of matching the prescription in each region.

Variation of S with respect to the parameters (a_n, x_n, y_n, z_n) produces a set of nonlinear equations, the solutions of which define local minima,

$$\sum_m \left\{ a_m \int M(\psi_n \psi_m) \right\} = \int M(P \psi_n) dx dy dz, \tag{10}$$

$$\int M \left(\sum_n a_n \psi_n - P \right) \frac{\partial \psi_m}{\partial x_m} dx dy dz = 0,$$

$$\int M \left(\sum_n a_n \psi_n - P \right) \frac{\partial \psi_m}{\partial y_m} dx dy dz = 0, \tag{11}$$

$$\int M \left(\sum_n a_n \psi_n - P \right) \frac{\partial \psi_m}{\partial z_m} dx dy dz = 0.$$

A positivity constraint ($a_n \geq 0$) is applied to each shot weight. This set of equations applies for nonvanishing weights, and is to be solved for (a_n, x_n, y_n, z_n) . We solve this system by linearizing the spatial equations [Eq. (11)] with respect to shot locations in the neighborhood of a solution and solving them in an iterative fashion. The explicit linearized equations are not shown. The time parameters a_n are also updated at each iteration, solving Eq. (10) with the constraint ($a_n \geq 0$) imposed. General solutions will involve some zero weighted shots, in which case Eq. (10) is handled as a least-squares optimization to determine the weights a_n . Shots with persistently zero weights over several iterations are eventually excluded. Once the zero shots have been excluded, Eq. (10) can generally be inverted to determine the weights.

This least-squares formulation does not allow specification of a prescribed isodose coverage, and it is often necessary to specify a margin around the tumor volume to increase the minimum isodose that covers the target. More generally the method is limited by dependence on an explicit representation of an objective function that results in simple iterative equations. This hampers the incorporation of additional criteria in the objective function. An additional limitation is that shot widths are determined and fixed in advance. At present, this first-generation solver serves as a benchmark for evaluating the second-generation constrained solver.

III. RESULTS

A. Two-dimensional optimizations

We have created a testing environment for our models within MATLAB, using the image processing toolbox. This interactive environment allows the user to draw a two-dimensional treatment volume of any size and shape. We used a MATLAB/GAMS interface to communicate between these two software packages.²⁰

After an optimization is complete, the MATLAB/GAMS interface returns the solution to MATLAB. In MATLAB, the results are visualized and analyzed using tools such as dose

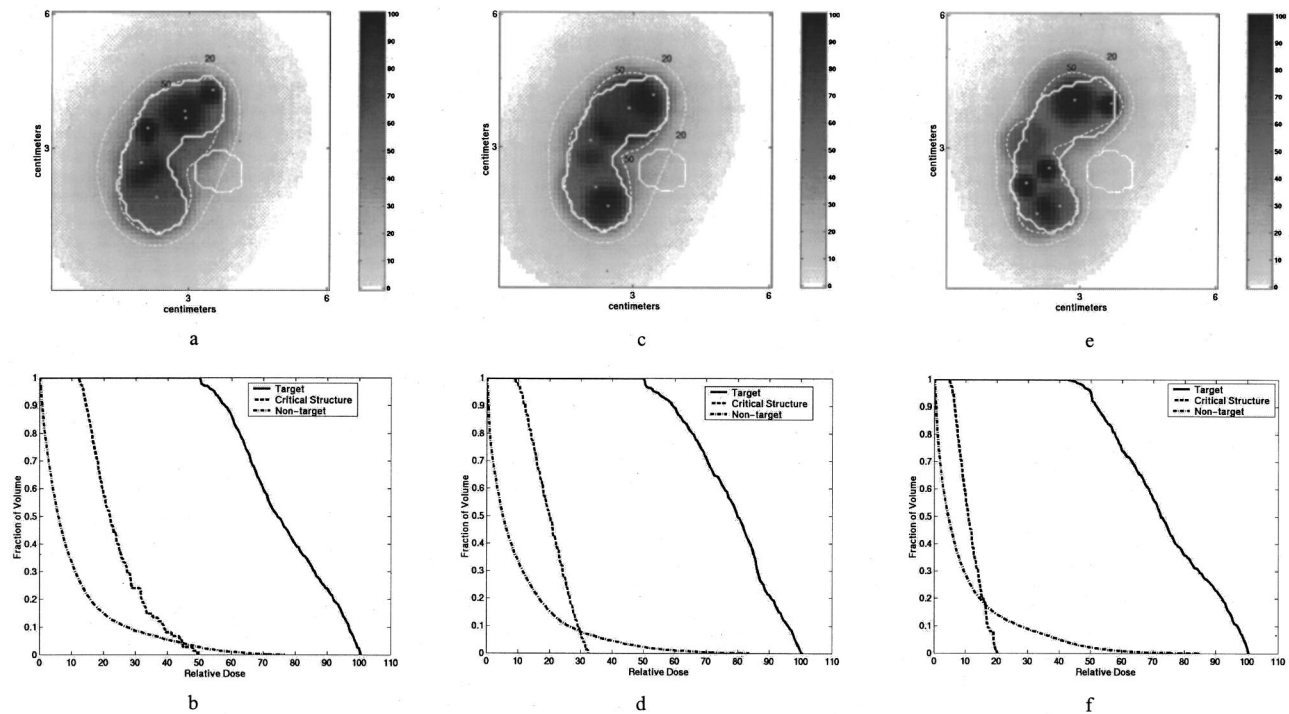


FIG. 2. Three optimizations were performed using an irregularly shaped target with a sensitive structure located in the concave region of the target. The white asterisks represent the center location of a shot. (a) The optimizer sought to maximize the percentage of dose deposited in the target while surrounding the target with the 50% isodose line and while not exceeding six shots. (c) the optimizer sought to minimize the number of shots while maintaining 55% of the integral dose within the target and while surrounding the target with the 50% isodose line. (e) The optimizer sought to minimize underdosage of the target while not exceeding 20% of the maximum dose in the sensitive structure and while not exceeding eight shots. (b), (d), and (f) The corresponding DVHs.

volume histograms (DVH). This testing environment has proven to be an excellent tool for pinpointing the advantages and disadvantages of various formulations and numerous optimization techniques.

A powerful feature of the GAMS modeling system is that the user can easily switch between a wide variety of model types, objective functions, and constraints. Consequently, we have examined numerous formulations in the two-dimensional test environment. This is illustrated in Fig. 2, which presents the results from three optimizations performed using an irregularly shaped target with a sensitive structure located within the concave region of the target. Clinically, this type of geometry is seen when a tumor is adjacent to a patient's optic chiasm or brain stem.

The first optimization performed using this setup sought to maximize the percentage of integral dose deposited within the target. A constraint ensured that the 50% isodose line surrounded the entire target. The maximum number of shots was set at six. This type of formulation assumes that the treatment planner can use the volume of the target and the irregularity of its shape in order to determine an acceptable number of shots. The optimized dose distribution and the corresponding dose volume histogram are shown in Figs. 2(a) and 2(b).

Figures 2(c) and 2(d) present the results from a second optimization in which the goal was to minimize the total number of shots. A constraint was used to ensure that the 50% isodose line covered the entire target. A second constraint required that at least 55% of the integral dose was

deposited within the target. Five shots were included in the optimized result.

Figures 2(e) and 2(f) show the results from a third optimization. The maximum number of shots was eight, and no point in the sensitive structure was allowed to exceed 20% of the maximum target dose. At least 55% of the integral dose had to be deposited in the target. Given these constraints, the optimizer sought to minimize the total underdosage within the target. In Fig. 2(e) note that the 20% isodose line closely matches the edge of the sensitive structure. Note that a less homogeneous tumor dose was accepted here to meet the difficult constraint on the dose to the sensitive structure.

The results in Fig. 2 illustrate that an optimization can be tailored to meet each patient's individual needs. For example, if a patient's tumor dose is limited due to the proximity of a sensitive structure, it is possible to maximize the tumor dose uniformity while maintaining an upper bound on the dose to a sensitive structure. Another feature of inverse planning is that one can choose to limit the number of shot sizes used in the treatment plan. This can be used to improve treatment efficiency by reducing the number of helmet changes.

The majority of the optimizations that we have performed have sought to maximize the percentage of the integral dose deposited within the target. Each optimization also specifies a minimum isodose line that must surround the target and the maximum number of shots. A few optimized results produced using this approach are illustrated in Fig. 3. The first case shown is a triangular-shaped target. It is difficult to

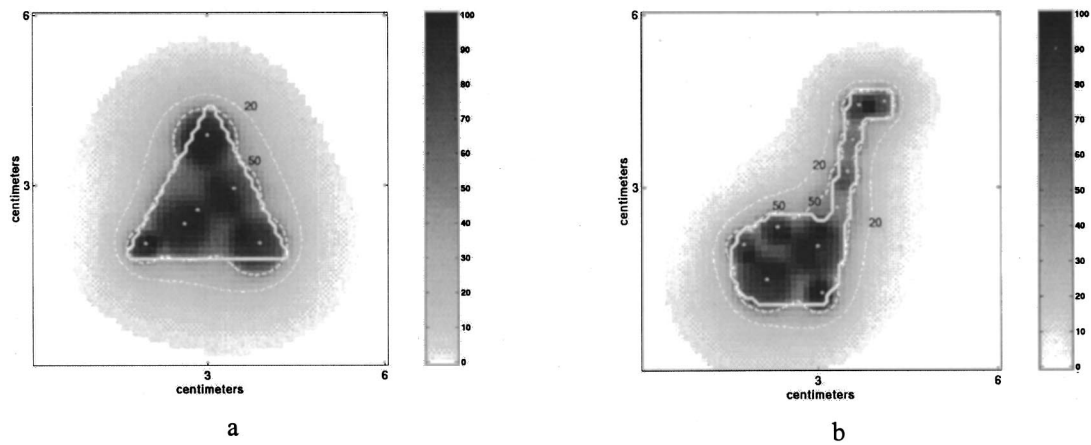


FIG. 3. Optimized dose distributions producing using the automated approach: (a) a triangular shaped target, (b) a target in the shape of a musical note.

conform to a triangle due to its sharp corners. With inverse treatment planning, however, a conformal result was produced with six shots of radiation. The second result [see Fig. 3(b)] is for an irregularly shaped target in the shape of a musical note. The long and narrow extension of this target provides an added level of difficulty, but the optimizer was able to conform tightly to the target boundary.

The nonlinear programming approach that we have used is designed to find solutions that satisfy first-order optimality conditions. Therefore, global optimality is not guaranteed. Our approach, however, has proven robust and effective for a wide range of target sizes and shapes. The robustness is primarily due to the fact that the shot locations are chosen as part of the optimization. The optimization code uses both gradient information and the constraints to aid in its search. We refer to this as a “migrating shot location approach.”

We have made several modeling enhancements in order to improve the speed of the optimizations and the quality of the final solutions. The first of these enhancements is that all four shot sizes are allowed at each migrating shot location. During the optimization, the arctangent function is used to choose the size to be applied at each location. This approach reduces the number of constraints in the model and significantly improves the computational times. In order to improve robustness, we typically use more trial shot locations than are requested in the final solution while constraining the optimization to choose the most effective locations among this set. This is again accomplished using the arctangent formulation. For example, ten possible shot locations might be included in an optimization that specifies a maximum of six shots. The optimization is constrained to choose six out of these forty possible shots (ten locations and four shots assigned at each). This large number of “migrating” shots increases the search space and enhances the robustness of the technique. For all of the results presented in this paper, random shot locations were used as the starting point for the optimization. Despite the poor quality of this initial guess, each optimization produced a feasible solution with a conformal dose distribution.

B. Patient results

Our experience with two-dimensional optimizations allowed us to choose appropriate formulations to apply to real three-dimensional patient data. The nonlinear programming approach to treatment planning has currently been applied to four patient cases. For each case, the optimized treatment plan was compared to the original treatment plan delivered to the actual patient.

Figure 4 provides a dose volume histogram comparison between a manually defined treatment plan and three optimized plans produced for the same patient. Seven shots of radiation were used in the patient’s original treatment plan. Two constrained optimizations were performed using five and six shots of radiation, respectively. A least-squares optimization was also performed using five shots. It can be seen that the optimized plans produced a more homogeneous target dose without increasing the dose to the normal tissue.

For the same patient, a series of isodose plots are shown in Fig. 5. The top row displays an axial, a coronal, and a sagittal slice taken from the seven shot treatment plan produced by the physician. The bottom row displays the corresponding slices from an optimized plan using six shots. Note that the optimized treatment plan maintained the conformity of the neurosurgeon’s plan while using fewer shots.

Table II provides a comparison between optimized and manually defined plans for four Gamma Knife patients. The treatment goals for each optimization were determined based upon the patient’s original prescription. In each case, the number of shots used in the optimized plan was equal to or less than the number used in the manual plan.

In Table II, the plans are compared based upon the radiotherapy guidelines of the Radiation Therapy Oncology Group (RTOG).²¹ The guidelines specify that a case is per protocol if the target is encompassed by 90% of the prescription isodose. In Table II, V90 indicates the percentage of the target volume covered by 90% of the prescription isodose. Table II also includes the conformity index (CI) for each plan. The conformity index is defined as the volume of the

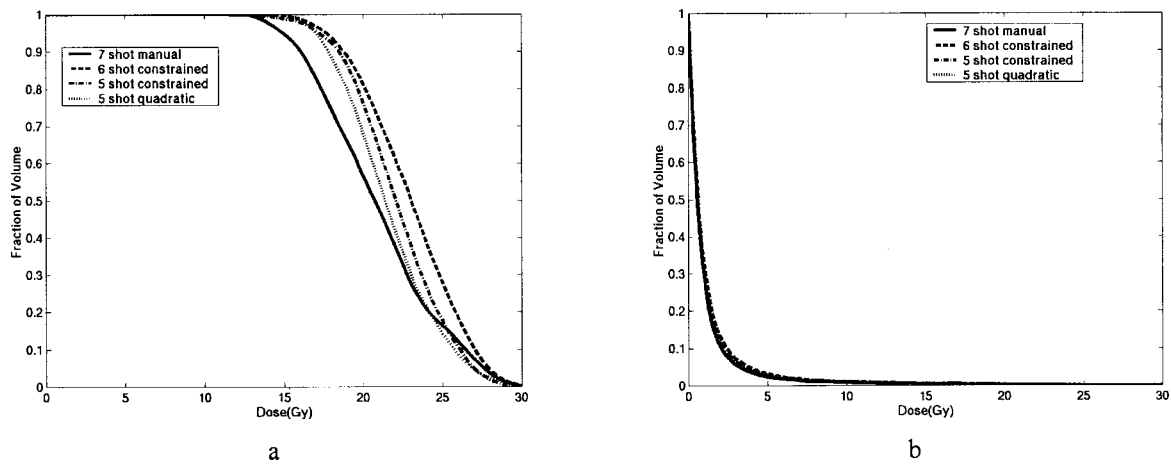


FIG. 4. (a) The solid line plots the results from a seven-shot plan produced by a neurosurgeon. A more uniform tumor dose is provided by both the six-shot and five-shot plans produced using the inverse-planning approach. Along with the six- and five-shot plans produced using a constrained optimization approach, the DVH curve for a five-shot plan using the least-squares approach is shown. (b) Note that all four plans produce very similar DVH curves for the normal tissue.

prescription isodose divided by the target volume. The RTOG considers a case to be per protocol if this ratio falls between 1.0 and 2.0. For each of the four patients, the constrained optimization produced a plan that satisfied the RTOG requirements in terms of both tumor coverage and dose conformity. Typically, tumor margins are included in

our least-squares formulation in order to achieve adequate tumor dose homogeneity. Consequently, the conformity indices for this approach may deviate from the protocol.

In defining a patient's treatment plan, a clinical judgment must be made as to the importance of normal tissue sparing as compared to that of tumor dose homogeneity. Initial

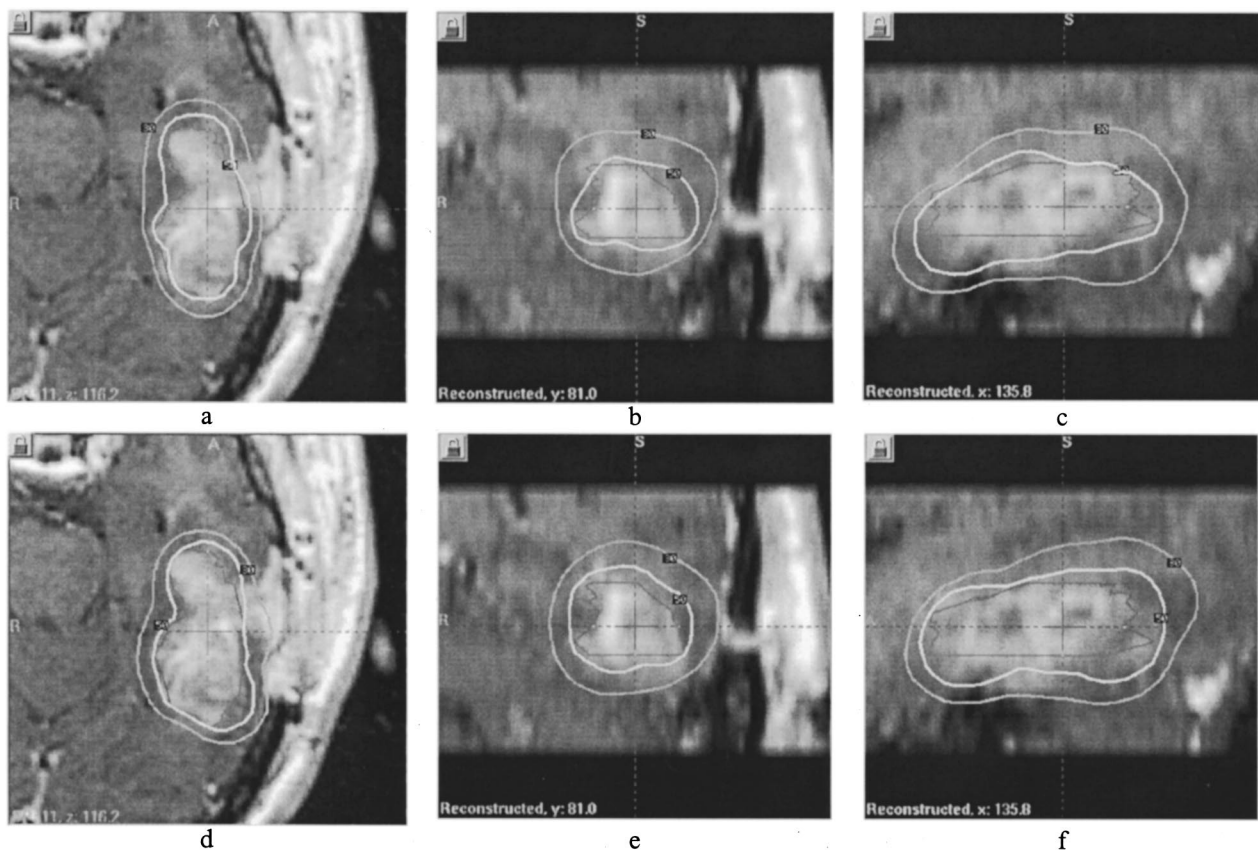


FIG. 5. (a)–(c) An axial, a coronal, and a sagittal slice, respectively, from the seven-shot dose distribution produced by the neurosurgeon. This is the same patient that was presented in Fig. 6. (d)–(f) The corresponding slices from the optimized six-shot case. In all cases, the purple/pink line outlines the tumor. The yellow line is the 50% isodose curve, and the green line is the 30% isodose curve.

TABLE II. The results for four patient cases are summarized based on V90 and conformity index. V90 indicates the percentage of the target volume covered by 90% of the prescription isodose. The conformity index is the ratio between the volume of the prescription isodose and the volume of the target.

Patient number	Target volume (cm ³)	V90			Conformity index		
		Manual	Least sq.	Constrained	Manual	Least sq.	Constrained
1	8.3	99	99	100	1.70	1.92	1.94
2	7.1	100	100	100	2.06	2.04	1.59
3	6.2	100	98	100	1.76	1.41	1.03
4	35.3	100	99	100	1.28	1.49	1.27

analysis from a recent RTOG brain metastases radiosurgery trial suggests a local control benefit to prescribing at a lower isodose line (less homogeneous distributions). This result may be due to the delivery of higher doses to the centrally located hypoxic regions of the tumor.²² With the constrained optimization approach, the minimum isodose coverage is defined by the user. For example, a patient case was optimized with the minimum target isodose constrained with a lower bound of 45%. A second optimization was performed on the same patient with a lower bound of 60%. For both plans, a dose of 18 Gy was prescribed to the target's minimum isodose line. The second plan provided a more uniform target dose, but conformity index increased from 1.73 to 1.81. There was also an increase in the volume of normal tissue receiving more than 10 Gy (17.0 vs 21.9 cm³). A key feature of inverse planning is the ease with which the prescription can be modified in order to strike the desired balance between dose homogeneity and dose conformity. In particular, specification of the optimization problem as a set of constraints proved advantageous over simply matching prescription dose by least squares. The constrained approach allows formulation of the problem in terms of clinically oriented guidelines, such as those proposed by the RTOG for plan evaluation, guaranteeing a satisfactory solution. The least-squares approach could only achieve such results by a trial and error process.

Due to variations in both skull size and target location, the dose distribution produced by a shot of radiation changes from one patient to the next. Consequently, the parameters in the dose model should be reoptimized for each patient. This reoptimization process is trivial provided that the dose profiles are known. Unfortunately, there is no straightforward approach for extracting the dose profiles from the Leksell GAMMAPLAN system. Therefore, for this initial work, the parameters in Table I were used for all of the optimizations.

All of our treatment plans were optimized on a Sparc Ultra-10 with a 330 MHz processor. The amount of time required for optimizing each patient's treatment plan varied widely over a range of 5–45 min. The speed of an optimization is highly dependent upon both the number of shots and the number of calculation points. To improve the speed, we typically run an initial optimization using a coarser grid of calculation points. The result is then used as the starting point in a second optimization using a denser grid of points. It should be noted that a more detailed analysis of inverse

planning as applied to real patients will be provided in a future article.

IV. CONCLUSIONS

An automated treatment planning system for Gamma Knife radiosurgery has been developed. This technique is based upon the use of a nonlinear programming approach in conjunction with migrating shot locations. The system simultaneously optimizes the shot sizes, locations, and weights, and it does not require any preplanning on the part of the user. The optimized dose distributions from a series of patients have been analyzed, and the results demonstrate that inverse planning is often able to improve dose conformity while reducing the total number of shots. Consequently, this tool has the potential to improve both the quality and the efficiency of Gamma Knife procedures.

V. FUTURE WORK

With refinements in the dose model, this approach to inverse planning can be further improved. This is illustrated with a DVH comparison shown in Fig. 6. The first set of DVH curves were computed using the manually defined treatment plan that included five shots of radiation. The second set of DVH curves were computed from an optimized dose distribution using five shots. In this case, the dose was calculated using the error function model described earlier. Finally, the optimized shots sizes, locations, and weights were entered into the GAMMAPLAN system, and a third set of DVH curves were obtained. It can be seen that the quality of the plan was significantly degraded due to the lack of agreement between the two dose calculations. The agreement could be improved by updating the dose model on a patient by patient basis. We are currently working on automating this procedure. We are also planning to develop a dose model based on a cubic *b*-spline fit to the data. This formulation may prove easier for the optimizer to solve. It should be noted that the patient results presented earlier in this text were all based upon the final dose calculation performed after the optimized shot locations and weights were plugged into the GAMMAPLAN system.

Future work will also focus on speeding up the optimization process. The optimizations performed for this paper began with shots of radiation placed at random locations in the target. Improving the starting point for each optimization

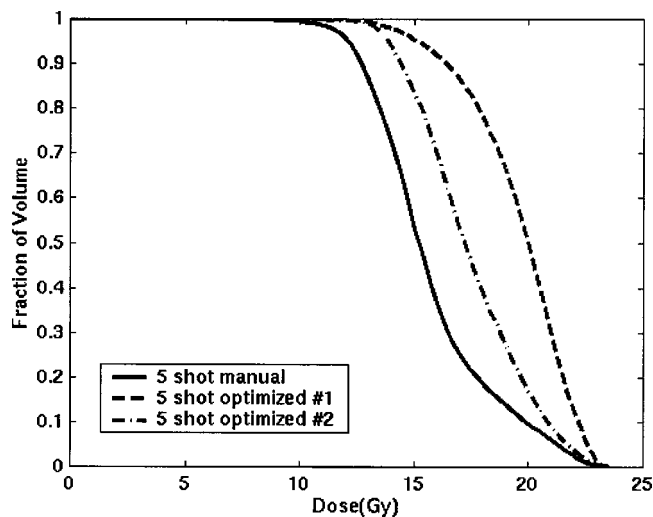


FIG. 6. The solid line on the DVH curves was computed from a patient's manually defined treatment plan that included seven shots of radiation. The dashed line on the DVH was computed from an optimized dose distribution using seven shots. In this case, the dose was calculated using the error function model described earlier. Finally, the optimized shots sizes, locations, and weights were entered into the GAMMAPLAN system, and the DVH curve represented by the dashed and dotted line was obtained. It can be seen that the quality of the plan was significantly degraded due to the lack of agreement between the two dose calculations. Reformulating the dose model on a patient by patient basis should largely remove these discrepancies.

could significantly reduce the optimization time. We are currently testing a variety of techniques in hopes of developing a means of quickly obtaining a high quality initial guess.

ACKNOWLEDGMENTS

The authors would like to thank Bruce Rakes and Elekta for their support in this research. This research was partially supported by Air Force Office of Scientific Research Grant No. F49620-98-1-0417 and National Science Foundation Grant No. CCR-9972372.

^{a)}Electronic mail: dshep001@umaryland.edu

¹J. C. Ganz, *Gamma Knife Radiosurgery* (Springer, Australia, 1997).

²A. Wu *et al.*, "Physics of gamma knife approach on convergent beams in stereotactic radiosurgery," *Int. J. Radiat. Oncol., Biol., Phys.* **18**, 941–949 (1990).

- ³J. C. Flickinger *et al.*, "Treatment planning for gamma knife radiosurgery with multiple isocenters," *Int. J. Radiat. Oncol., Biol., Phys.* **18**, 1495–1501 (1990).
- ⁴A. H. Maitz *et al.*, "Quality assurance for gamma knife stereotactic radiosurgery," *Int. J. Radiat. Oncol., Biol., Phys.* **32**, 1465–1471 (1995).
- ⁵<http://www.elekta.com>
- ⁶Q. J. Wu and J. D. Bourland, "Morphology-guided radiosurgery treatment planning and optimization for multiple isocenters," *Med. Phys.* **26**, 2151–2160 (1999).
- ⁷G. Stücklschweiger, "[Stereotaxic convergent irradiation with the gamma knife. A study of the possibilities for optimizing the dosage distribution]," *Strahlenther Onkol.* **171**, 499–509 (1995).
- ⁸H. Z. Shu *et al.*, "Treatment planning optimization by quasi-Newton and simulated annealing methods for gamma unit treatment system," *Phys. Med. Biol.* **43**, 2795–2805 (1998).
- ⁹Y. Yan *et al.*, "Clinical treatment planning optimization by Powell's method for gamma unit treatment system," *Int. J. Radiat. Oncol., Biol., Phys.* **39**, 247–254 (1997).
- ¹⁰H. Shu *et al.*, "Three-dimensional optimization of treatment planning for gamma unit treatment system," *Med. Phys.* **25**, 2352–2357 (1998).
- ¹¹L. Luo *et al.*, "Optimizing computerized treatment planning for the gamma knife by source culling," *Int. J. Radiat. Oncol., Biol., Phys.* **45**, 1339–1346 (1999).
- ¹²R. A. Stone, V. Smith, and L. Verhey, "Inverse planning for the Gamma Knife," *Med. Phys.* **20**, 865 (1993).
- ¹³P. S. Cho, H. G. Kuterdem, and R. J. Marks, "A spherical dose model for radiosurgery plan optimization," *Phys. Med. Biol.* **43**, 3145–3148 (1998).
- ¹⁴A. Drud, "Interactions between nonlinear programming and modeling systems," *Math. Program.* **79**, 99–123 (1997).
- ¹⁵A. Drud, "CONOPT: A GRG code for large sparse dynamic nonlinear optimization problems," *Math. Program.* **31**, 153–191 (1985).
- ¹⁶A. Drud, "A large scale GRG code," *ORSA J. Comput.* **6**, 207–216 (1994).
- ¹⁷A. Drud, *CONOPT—A System for Large Scale Nonlinear Optimization* (ARKI Consulting and Development A/S, Bagsvaerd, Denmark, 1996).
- ¹⁸A. Brooke, D. Kendrick, and A. Meeraus, *GAMS: User's Guide* (The Scientific Press, South San Francisco, CA, 1988).
- ¹⁹M. C. Ferris and D. M. Shepard, "Optimization of Gamma Knife radiosurgery," presented at the Proceedings of the DIMACS Workshop on Discrete Mathematical Problems and Medical Applications, Rutgers University, Piscataway, NJ, 1999.
- ²⁰M. C. Ferris, "MATLAB and GAMS: Interfacing Optimization and Visualization Software," *Mathematical Programming Technical Report No. 98-19*, 1998.
- ²¹E. Shaw *et al.*, "Radiation Therapy Oncology Group: Radiosurgery quality assurance guidelines," *Int. J. Radiat. Oncol., Biol., Phys.* **27**, 1231–1239 (1993).
- ²²E. Shaw (unpublished).