to appear: IEEE Transactions on Robotics and Automation

Planning for Camera-Guided Robotic Radiosurgery

Achim Schweikard, Mohan Bodduluri, John R. Adler

Abstract— **Radiosurgery is a treatment for brain tumors and other lesions which employs a moving beam of photon radiation. In robotic radiosurgery, the radiation source is moved by a six degree-of-freedom robotic arm. A treatment planning system for robotic radiosurgery was developed. This planning system combines collision avoidance techniques and geometric algorithms for finding appropriate robot paths and beam activation profiles. The following design goals characterize our application: (1) ability to generate conformal three-dimensional dose distributions, (2) transparency of the planning process, (3) full use of the motion flexibility provided by the robotic arm. The new planning system is in clinical use in several leading medical institutions. Results obtained from studies suggest that highly conformal distributions can be generated in a very effective way.**

Keywords— **Medical Robotics, Geometric Reasoning, Motion Planning, Collision Avoidance, Image-Guided Surgery, Treatment Planning, Radiosurgery, Radiation Therapy, Medical Physics**

Fig. 1. Robotic radiosurgical system, patient table, phantom skull, x-ray cameras

I. Introduction

In radiosurgery, brain lesions are treated with an intense beam of radiation [4], [9], [12]. During treatment, the lesion is targeted from many directions. This is done to reduce the dose in the healthy tissue, while accumulating a high dose inside the target. The radiosurgical treatment consists of several phases. First, a precise 3D map of the anatomical structures in the area of interest is constructed using computed tomography (CT) or magnetic resonance (MR). Next, a motion path for the radiation beam is determined to deliver an appropriate dose distribution (taking into account a variety of medical constraints). Finally, a jointed mechanism moves the radiation source according to this path.

A robotic radiosurgical system has been built to overcome limitations inherent in earlier systems. This system is based on an on-line x-ray vision system and a robotic arm moving a linear accelerator. The system consists of the following four components (fig. 1): (1) A compact and light-weight linear accelerator running at high frequency generates a 6 MV (cylindrical) photon beam. (2) A GM-Fanuc 420 manipulator with six degrees of freedom moves the linear accelerator during treatment. (3) CT and MR scanners acquire pre-operative images of the brain. An on-line vision system with two orthogonal x-ray cameras takes images of the patient's skull at fixed time intervals during treatment. (4) a treatment planning system computes a beam path, given the constraints placed by the desired dose distribution. The beam path must observe additional constraints, i.e., the path must be collision-free and must not obstruct surveillance and online vision subsystems.

The treatment delivery proceeds as follows. Prior to treatment, a *collimator* (lead inset for beam shaping) of appropriate size is inserted into the radiation source. It should be noted that the current system version uses conical collimators, i.e., the beam is a (nearly cylindrical) cone with very small conicity angle. Collimator sizes range from 5mm to 40mm in diameter, with increments of 2.5mm. The collimator typically remains in place throughout the entire treatment. During treatment, the linear accelerator is moved in point-to-point mode through a series of configurations. At each configuration, the beam is activated for a small time interval, while the arm is held still. The duration of activation can vary from one configuration to the next. Radiographs are taken within the time intervals between beam activation. The vision system then computes the patient's current position by correlating the images to precomputed radiographs. Small movements of the patient are compensated for by a corresponding motion of the manipulator arm.

In contrast, earlier systems use gantry constructions with two (rotational) joint axes movable under computer control. Typically, such systems are combined with a patient table with four degrees of freedom (three translational degrees of freedom and one rotation axis). This allows for executing a restricted class of motions under computer control (see e.g. [9], [12]).

For conventional systems, the patient's head is fixed with a *stereotaxic frame* (fig. 2).

Thus robotic radiosurgery offers the following advantages over earlier radiosurgical and radiotherapeutical methods: 1. The tissue does not have to be fixed in space. Besides being very painful for the patient, fixation with stereotaxic frames is only possible for the lesions in the head, so radio-

J.R. Adler is with Department of Neurosurgery, Stanford University, Stanford, CA. 94305, USA. E-mail: adler@flamingo.stanford.edu

M. Bodduluri is with Accuray Inc., Sunnyvale, CA. 94086, USA.

A. Schweikard is currently with Department of Computer Science, Technical University Munich, Orleansstr. 34, 81667 Munich, Germany. E-mail: schweika@in.tum.de

Fig. 2. For conventional radiosurgery, the patient's head must be immobilized with a frame via holes drilled directly into the skull. Image-guided radiosurgery obviates the need for rigid head fixation.

surgery with conventional systems methods was limited to the head.

2. The robotic gantry has six degrees of freedom. Thus the radiation source can be moved along arbitrary paths (i.e., translated and rotated) in space.

In addition to increased patient comfort, the first of these features allows for *fractionated* treatment, i.e., the total dose can be delivered in a series of 2-30 treatments, where only a small dose is delivered during each treatment. Fractionation is very effective in radiation therapy, but so far could not be used in radiosurgery due to the necessity of rigidly fixing the tissue in space.

In the following, we will describe the treatment planning and collision avoidance techniques developed for this robotic system. The second of the above features allows for executing arbitrary paths in space. A planning system was developed which combines motion planning and treatment planning algorithms. A main result is that highly conformal dose distributions could be generated in an effective and precise way by making full use of the system's kinematic flexibility.

II. Treatment Planning Constraints

The *direct dosimetry problem*, or forward planning problem, is to compute the dose distribution in an anatomic region given a treatment plan. It relies on a model of physical characteristics of photon beams in tissue structures of different types, measured in phantom studies (see, e.g., [8]). The *inverse dosimetry problem*, also called inverse planning problem, is to find a treatment plan whose execution will achieve a desired dose distribution.

We will first describe the main constraints in treatment

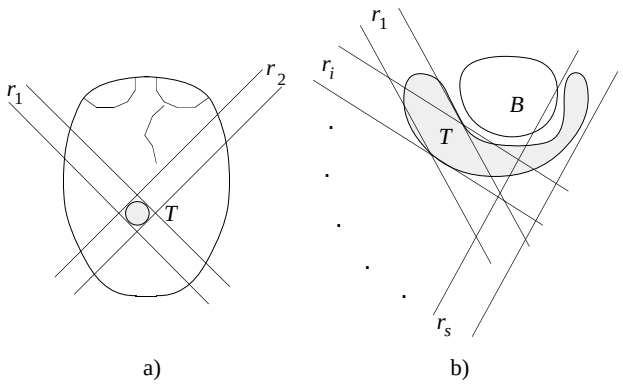

Fig. 3. Cross-firing at a tumor

planning. The objective is to find a motion suitable for treating the particular condition presented by a patient. Fig. 3-a schematically depicts an axial cross-section of the brain with a circular tumor T. If T is irradiated from only one direction r_1 (with a cylinder beam), the healthy tissue along the beam absorbs approximately the same dose as the tumor. If, instead, we use two beam directions, r_1 and r_2, the dose deposited in the tumor is approximately twice the dose in healthy tissue. Using more beam directions can lead to further improvements of the dose distribution, and a very sharp drop-off of the dose in the tissue surrounding the (spherical) tumor region can be achieved. In this type of treatment, the axes of all beams cross a single point in space, called the *isocenter*. Notice that the treated volume (i.e., the region receiving $\geq 80\%$ of the maximal dose) is spherical if the axes of all beams cross a single isocenter. However, for tumors of non-spherical shape this single-isocenter treatment procedure is problematic for the following reason: The (non-spherical) tumor region T in fig. 3-b surrounds a *critical* or particularly radiation-sensitive healthy region B. The dose in B should remain as small as possible. A single-isocenter treatment will generate a treated volume of spherical shape. Treating a sphere containing T will lead to an over-radiation of B. The beams shown in fig. 3-b do not cross a single isocenter. However, the beam directions shown in the figure equally may not be adequate for this case, since this type of beam motion will accumulate dose along the upper (concave) boundary of T, leading to an undesirable inhomogeneity of dose inside T.

An appropriate dose distribution for a given tumor should satisfy the following constraints:

1. Critical and/or radiation-sensitive healthy structures near the tumor (i.e., nerves, blood vessels, sensitive brain structures) should receive very low dose.
2. For certain cases, the dose inside the tumor should be highly uniform, i.e., the dose in the tumor should not only be lower-bounded but also upper-bounded.
3. The treated volume should cover the entire tumor and should match the tumor shape.
4. The drop-off around the treated volume should be as sharp as possible.

For current radiosurgical systems a treatment plan is determined manually by the surgeon for each individual pa-

tient. In many cases, appropriate dose distributions require a large set of beam directions and manual repositioning of the patient.

III. Related Work

The most widely used hardware systems for radiosurgery are LINAC-based gantry systems (LINAC-systems, [4], [9], [12]) and Gamma-knife systems ([2]). For LINAC-based radiosurgery the beam is moved along circular arcs in space. Gamma-knives do not move the beam source. Instead, static cobalt (^{60}CO) sources attached to a helmet generate an array of beams. As in LINAC-based radiosurgery, a stereotaxic frame is used to correlate tomographic images to the intra-operative patient position. For the Gamma-knife, the axes of all beams cross a single point in space. This results in an isocentric treatment mode similar to LINAC-procedures. The actual treatment delivery for typical LINAC treatments takes on the order of 20-30 minutes. Radiation time increases, if more than one isocenter is used, since the patient must be repositioned manually. It should be noted that the total *treatment time* for frame-based radiosurgery (as opposed to radiation time) is typically between six and seven hours. This is due to the fact that the stereotaxic frame must be visible on the tomographic images, and must therefore be put in place before acquiring the tomographic images. Treatment planning can only be carried out after the tomography has been taken.

A newer system is the NOMOS Peacock system ([7]), which employs a specific beam modulation hardware in connection with a standard LINAC gantry. This system operates on a series of two-dimensional cross-sections of the anatomic region to be treated. Treatment planning for this system is based on simulated annealing.

The inverse planning problem in conventional radiosurgery and radiation therapy has been studied extensively in the literature (see, e.g., [3], [5], [7], [11], [13], [14]). Most of the known approaches are based on LINAC treatment procedures.

Progress on exact methods for computing inverse treatment plans in two dimensions with stripe beams has been reported by Barth [3]. This approach relies on solving integral equations derived from a forward model of beam fluency. Although this approach is promising, progress regarding generalizations to the three-dimensional case has so far remained limited.

New approaches for inverse planning based on optimization have been described recently in [5] and [7], [11]. Both of the latter approaches compute plans in a fully automated non-linear optimization process, which attempts to satisfy the given constraints as closely as possible. The individual constraints have parameters defining their respective priority. Common to these approaches are fairly high computing times, so that inverse planning is typically carried out in a single step.

In our case, fully computerized optimization of distributions is problematic and is not used for the following reasons:

- The constraints describing an adequate dose distribution are conflicting. As an example, in many cases it is difficult to achieve accurate shape matching (of tumor and treated volume) *and* sharp dose drop-off around the treated volume.
- In general, non-linear optimization methods will only yield a local optimum, with no information about an overall optimum.
- The results of non-linear optimization depend on settings of internal system parameters, such as step-lengths, and momentum values. The precise way in which changes in these internal parameters will influence the resulting distribution is not easy to specify. Small changes in these values can yield largely diverging results. It is difficult to make the influence of such parameter settings and trade-offs transparent to the system user.

To address problems with optimization methods, treatment planning for our system is performed interactively by the surgeon, where each of the following three steps is carried out separately.

1. Selecting beam configurations
2. Computing beam weights, i.e., duration of activation at each configuration
3. Kinematics and path planning

The first planning step involves finding appropriate beam configurations, given the shape and the location of the tumor. New paradigms for beam selection have been developed to allow for full use of the system's kinematic flexibility while addressing safety and verifiability requirements of the application.

The second step assigns weights to each of the selected beam configurations. Linear programming methods for computing such weights in radiation therapy have been proposed by several authors (see, e.g., [14], [15]). Extensions using quadratic programming in radiation therapy are given by Menguy et al. [10], and previously by other groups [13]. We describe a framework for the application of linear programming in robotic radiosurgery, which relies on a simple description of the constraints for a desired distribution. The main advantage of linear and quadratic programming in this context is completeness. I.e., under appropriate assumptions for the physical characteristics of the beam, it can be decided exactly whether the input constraints are feasible. However, for infeasible constraints, the test must be repeated with modified constraints. In this sense, our method for computing beam weights is only semi-automatic. The computing times for each planning cycle are very short, so that the surgeon can evaluate different alternatives rapidly, thus gathering information on adequate compromises between the individual requirements. The use of internal optimization parameters can be avoided.

In the third step, the selected beams are connected by a collision-free motion of the radiation source. For conventional radiosurgery, motion paths are not changed during treatment. The described robotic system is camera-guided and patient motion is tracked by the radiation source. Therefore, the pre-computed motion path of the robot may have to be adapted *during* treatment execution.

The planned motion must be collision-free, and collision avoidance must be performed dynamically.

The following section describes beam selection methods. Section V derives a constraint set from a standardized description of the input planning problem, given previously selected beam configurations. Section VI describes methods for path planning, based on the beam directions computed in the first step. The implementation of these methods as well as results obtained from phantom studies and clinical use are summarized in section VII.

IV. Selecting Beam Configurations

a) Spherical target regions: We first assume the tumor is spherical, with center p and diameter d. We set the diameter of all beams to d. The beam configurations are then selected in the following way. The axes of all beams cross the tumor center p (fig. 4-a). p is the isocenter. We consider beam directions determined by vertices of a regular polyhedron centered at p.

The choice of orientation of this polyhedron is problematic if there are critical healthy structures in the tumor vicinity. Furthermore there is no such regular polyhedron with more than 20 vertices. A more even distribution is obtained using more directions. The methods in [16] provide point grids on spheres which are adequate for our purposes. Our planning method thus uses a large set of beams (typically $n \geq 100$) for a single sphere, and addresses the presence of critical regions not during the beam selection phase, but during the beam weighting phase.

b) Ellipsoidal target regions: While the above scheme gives both very sharp dose drop-off around a sphere-shaped lesion and sufficient homogeneity, generalizations of this scheme for treating non-spherical tumors are not obvious. One method is to use two spheres instead of a single sphere, where the two spheres are treated independently. However, it is generally difficult to cover an arbitrary tumor shape with spheres, such that the volume of overlap with the tumor environment is small and the entire tumor is covered. This is aggravated by the fact that the two spheres should not intersect, since this would yield very inhomogeneous dose distributions, i.e., high dose in the intersection region of the spheres and/or insufficient coverage of the tumor region.

A simple generalization of the above single sphere method is obtained in the following way. Instead of using a single isocenter, we place a series of isocenters evenly spaced along a line segment. (Notice the execution of such a motion requires a translational motion of the radiation source.) Each such isocenter is treated in the same way as a single isocenter (fig. 4-b). Central to this approach is the fact that each sphere receives very low dose, and many spheres are overlaid. In this way we can avoid hot spots in the intersection of two spheres.

However, if we use the same set of directions for all isocenters thus chosen, many of the beams will be parallel. Accumulated dose fall-off from such beams will cause dose inhomogeneity in the surrounding tissue.

To address this problem, we rotate the beams sets used for each isocenter by a small amount. In practice, the effect of such a rotation is equivalent to using only a small sample of beams for each isocenter, where this sample is chosen at random from a finer grid.

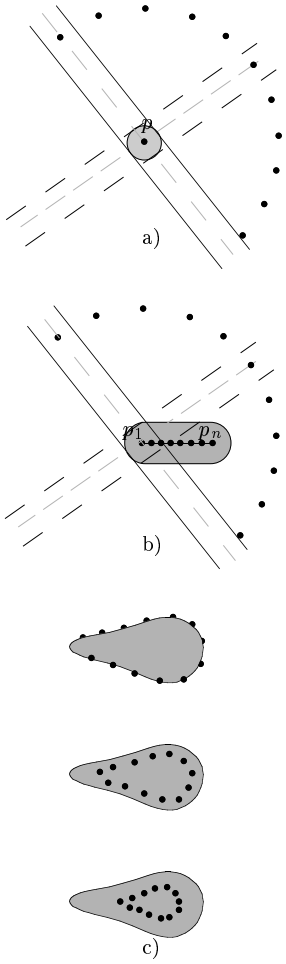

Fig. 4. Beam selection modes. a) Beam axes cross a single (common) isocenter p, and treated volume is spherical. b) Isocenter points $p_1, \ldots, p_n$ are placed on a line segment, resulting in pseudo-ellipsoidal treated volume. c) Beam selection for general target shapes, based on a grid of isocenter points on the tumor surface.

c) Target regions of general shape: The above approach is restricted to the pseudo-cylindrical shapes obtained by sweeping a sphere. A direct generalization of this scheme is the following. Using an algorithm described in [16], we compute a grid of points on the *tumor surface* This process is illustrated in the top image of fig. 4-c. A series of points is placed on the surface of the target (shown gray). Each such point is now treated as an isocenter. As above, a small sample of beams is used for each isocenter. Accumulation along certain principal directions is avoided by using a fine grid. (A finer grid of beams produces a more even distribution than a coarse grid.)

Accumulation of high dose in the tumor center is avoided by placing isocenters on the tumor surface, and using a

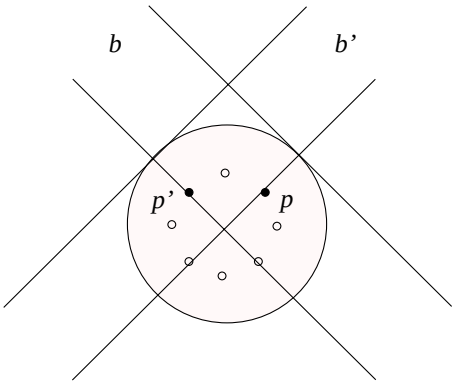

Fig. 5. Spherical target region, beams b, b' through isocenter points p, p'.

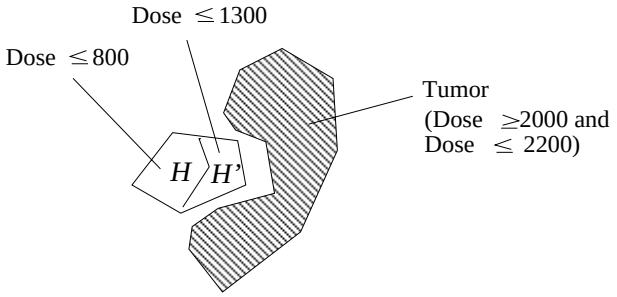

Fig. 6. Specifying distribution characteristics. To specify desired homogeneity in the tumor, upper and lower dose bounds are set. Separate thresholds are set for subregions H, H' of a critical region.

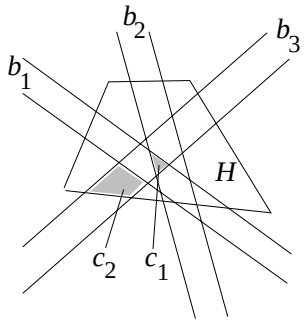

Fig. 7. Arrangement with three beams, critical region H. The beams determine a partitioning of space into cells.

sufficiently small collimator. Methods for computing beam weights in the following section are used to increase dose homogeneity inside the tumor.

The two bottom images in fig. 4-c show variants of this technique. In this case, the initial point grid in the top image has been retracted. This retraction is only done if necessary, to compensate for insufficient dose in the tumor center.

For nearly spherical tumor shapes, it is not obvious which of the selection modes a) - c) should be used. The answer to this question is surprisingly simple and gives additional motivation for the selection mode in c). We consider a circular target region in the plane. By mode a), we would place a single isocenter in the center of this region, and all beam axes cross this center (fig. 4-a). Instead, selection mode c) places isocenters along the surface of this circular region. Retracting this grid of points yields a point set as in fig. 5. Measurements show that the radiation dose fall-out in surrounding tissue depends on the beam diameter. This fall-out decreases for smaller beams. Thus, using a beam with smaller diameter will result in a decrease of fall-out in the surrounding tissue. This suggests that mode c) with a small collimator is preferable even for spherical tumors. However, for practical reasons we cannot always use mode c): Firstly, the use of a smaller beam will lead to large increase in total radiation time. Secondly, due to physical beam characteristics, the collimator diameter cannot be arbitrarily small. Thus common radiosurgical systems use only collimators of diameters ≥ 5 mm.

V. Computing Beam Weights

During the second planning step, weights are assigned to each of the previously selected beams. Let $b_1, \ldots, b_n$ be the selected beams. The weight of each beam is its activation duration. We first assume that the dose delivered at a point p by a single beam only depends on the activation duration, and not on other physical effects, such as the decrease of dose with depth in tissue. Specifically, if p is in all cylinders $b_1, \ldots, b_n$, we first assume that the dose $D(p)$ at p is given by $D(p) = w_1 + \ldots + w_n$, where w_i denotes the activation duration of beam b_i.

We specify the constraints for the distribution in the following way: The regions of interest (tumor, critical regions) are delineated manually by polygons in two-dimensional cross-sections on the control interface workstation. This results in a stack of polygons for each region. The stacks are converted into a three-dimensional representation of the corresponding region. For each region, we then enter upper (and lower) dose thresholds (fig. 6).

The beam directions selected in the previous step define an *arrangement* of cylinders in space. Maximally connected regions not intersected by any cylinder boundaries are called *cells*. The cells in this arrangement are further partitioned by the region boundaries (i.e., boundaries of target and critical regions). The regions c_1, c_2 in fig. 7 are cells in the (common) arrangement of cylinders and region boundaries. For each cell we obtain inequalities using the given the upper and lower dose thresholds. As an example, consider the cells c_1, c_2 in fig. 7. We assume an upper threshold β has been given for the critical region H. Then c_1 determines the inequality

$$w_1 + w_2 + w_3 \leq \beta,$$

while c_2 gives

$$w_3 \leq \beta.$$

The following scheme decides whether the distribution specified by delineations and dose thresholds can be achieved: We compute the inequalities corresponding to each cell. This gives rise to a system of linear inequalities.

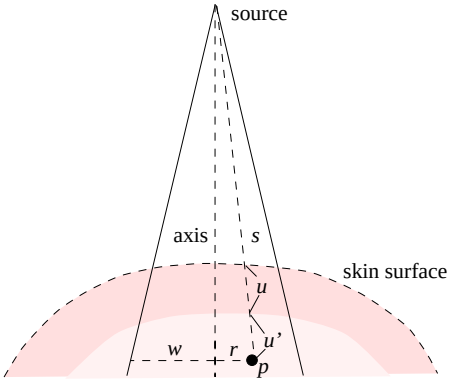

Fig. 8. Beam model for inverse planning. Source-skin distance s, off-axis distance r, radiological depth determined by tissue layer thicknesses u, u' and density ratios with respect to water, width w of orthogonal cross-section at p.

(Notice that the inequality for the cell c_2 is redundant and can be removed from this system.) The feasibility of the resulting system can be established with standard techniques. If feasible, a vector of values $w_1, \ldots, w_n$ is produced, giving dose weights for generating the desired distribution.

The above beam model only provides a coarse approximation for the physical characteristics of photon beams. To address this problem, we use a standard technique for forward dosimetric calculation ([8], pp. 182-204) to obtain a more accurate beam model for inverse planning. This model is based on a table representation of the following physical effects (fig. 8):

1. attenuation of the beam depending on radiological depth in tissue
2. decrease of dose with the distance between the source and the skin surface (source-surface distance s)
3. divergence of the beam (depending on the collimator size and the distance from the source)
4. change of dose with distance of the point p to the beam's central axis (off-axis distance r)

The radiological depth of a point p in tissue is the distance of p to the patient's skin surface, adapted in such a way that distinct densities of tissue along the beam path are taken into account. This is done by determining the thicknesses of distinct layers traversed (values u and u' in fig. 8), and multiplying these thicknesses by a factor representing density ratios with respect to water.

The source-surface-distance is nearly constant in our case, so that a single value is used (see also section VI).

The divergence at p can be assumed to depend only on the width w of the beam at the (orthogonal) cross-section containing p, a nd not on the particular collimator diameter ([8]). A single table of attenuation factors for a discrete range of widths w is used for all collimators.

We thus assume that the dose at p stemming from a single beam b_1 is proportional to the activation duration x_1 of this beam, and is given by $a_1 x_1$. a_1 is a *constant* factor representing the combined influence of the above effects at a single point p.

To determine the constraints for the weight computation, we place a regular grid of points over region of interest.

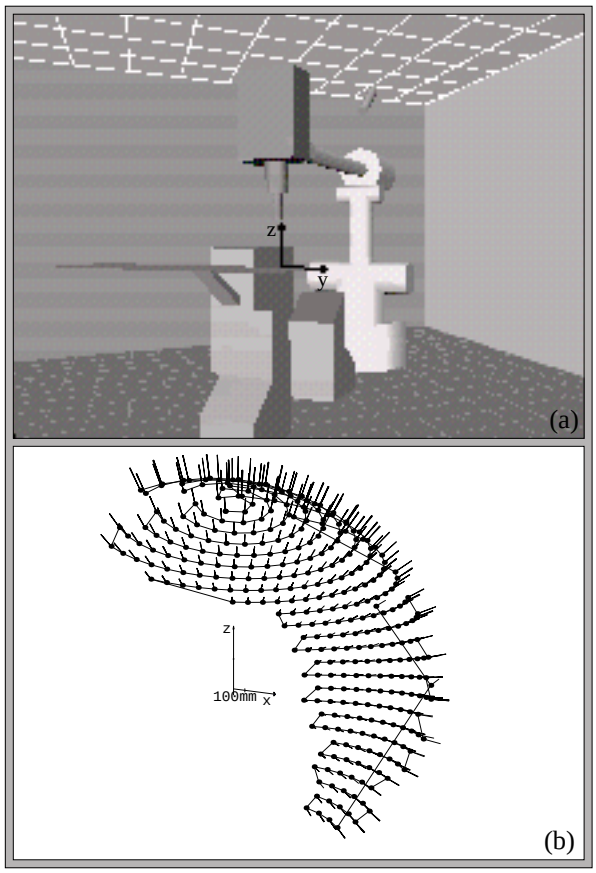

Fig. 9. Path template connecting node configurations

Each grid point p in a region with upper dose bound β determines a constraint of the form

$$a_1 x_1 + \ldots + a_n x_n \geq \beta.$$

Similarly, we can obtain constraints corresponding to lower dose thresholds.

VI. Kinematics and Collision Avoidance

The beam configurations selected by the planning system must be connected by a collision-free path of the linear accelerator. For safety reasons we use a hard-coded sequence of beam configurations, called *nodes*. A fixed motion, connecting the nodes, provides a path template. This path template is given in advance, and will not be changed during treatment planning. The computation of the template path relies on a kinematic simulation program included in the planning system. This simulator contains forward and inverse kinematics for the robot as well as collision detection, joint angle limit testing, and graphical display. Fig. 9-a shows the workspace with the base coordinate system at the upper end of the patient couch.

To determine the node points, we compute a grid of tool center points on a sphere centered at the work space origin (fig. 9-b). For each node point the beam axis crosses the work space origin.

A cylinder beam can be translated along and rotated about its axis without changing spatial placement. Thus a

cylinder has four degrees of freedom. In contrast, the robot has six degrees of freedom. Therefore the equation system for computing the robot joint angles from a given beam placement is under-determined. Valid solutions form a two-dimensional subspace of the six-dimensional joint space. However, practical considerations provide additional constraints. Firstly, the radiation source should remain at a fixed and constant distance from the workspace origin. This is due to the fact that the beam widens with depth in tissue. By remaining at an even distance from the origin, we can ensure that the beam diameter remains (nearly) constant within the target region. Secondly, the robot's joint angle ranges are limited. The range of beam directions should be as large as possible to allow for homogeneous distributions. Ideally, this range of directions covers a full sphere. A full sphere of directions cannot be reached, due to the presence of obstacles. Furthermore, reaching a large range of directions in the presence of obstacles requires complex motions of the arm. For ill-chosen angles, the limitations in joint ranges will require frequent posture changes (i.e., changes between elbow-up and elbow-down configurations of the arm, etc.).

The following scheme is used to resolve the kinematic redundancy in such a way, that posture changes along path are avoided. By convention, the $z-$axis of the effector coordinate system is aligned with the beam axis. The beam points along the positive $z-$direction of the effector system. We represent the orientation of the effector coordinate system by a 3×3 orientation matrix R, where

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}.$$

This matrix can be resolved into a matrix explicitly containing the Euler angles α, β and γ by setting

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} =$$

$$\begin{pmatrix} c_\alpha c_\beta c_\gamma - s_\alpha s_\gamma & -c_\alpha c_\beta s_\gamma - s_\alpha c_\gamma & c_\alpha s_\beta \\ s_\alpha c_\beta c_\gamma + c_\alpha s_\gamma & -s_\alpha c_\beta s_\gamma + c_\alpha c_\gamma & s_\alpha s_\beta \\ -s_\beta c_\gamma & s_\beta s_\gamma & c_\beta \end{pmatrix}$$

with the convention $s_\alpha = \sin(\alpha)$ and $c_\alpha = \cos(\alpha)$ for α, and analogous conventions for β and γ.

This matrix corresponds to $z-y-z-$ Euler angles ([6], p. 45). Specifically, γ represents the rotation about the *new* $z-$axis of the effector coordinate system previously rotated about the $z-$ and $y-$axes by α and β.

We must place the beam axis in such a way that it crosses the work space origin. The $z-$vector of the (new) effector system is the third column in the matrix R. We must thus solve this matrix equation using the third column of the matrix equation alone. The condition of being under-determined is reflected in the fact that we can compute

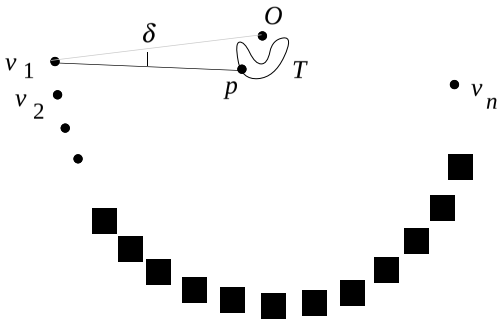

Fig. 10. Origin O, target region T, tool-center points $v_1, \ldots, v_n$ for hard-coded nodes

values for α and β from this column, but γ does not occur in this column. Indeed, γ can be chosen arbitrarily, and distinct choices for γ will result in the same $z-$vector of the effector system, but distinct placements of the x and y (effector) axes.

Fig. 9-b indicates the ordering in which the template nodes are to be traversed. Thus we connect the template nodes by a series of horizontal arcs. While moving from one node to the next node within one arc, the joint angles will be incremented or decremented by small amounts. Each node v_i in the given arc corresponds to a vector $(\alpha_i, \beta_i, \gamma_i)$ of angles, where α_i and β_i are fixed. Our goal is to choose the values γ_i in such a way that posture changes and collisions can be avoided.

We place a regular grid $\epsilon_1, \ldots, \epsilon_k$ over the range of angle values γ, i.e. the interval $[0, \pi)$. Fixing a value for γ_1 will determine a range of values for γ_2, such that the orientation $(\alpha_2, \beta_2, \gamma_2)$ is reachable without collisions or posture changes. This range clearly depends on our choice of γ_1. We thus obtain a search tree, in which the leaves are values for γ_i in the set $\epsilon_1, \ldots, \epsilon_k$, which are reachable without collisions or intersection. Best-first search is then applied in the following way. The cost of a leaf in the search tree is set to $1/c$, where c denotes the minimum distance to one of the joint angle limits. Specifically, for γ_1, we choose the value of lowest cost, and compute the range of reachable values for γ_2 with the simulation program. If this range (for γ_2) is empty, the process is restarted with γ_1 set to the value of next-to-lowest cost. Otherwise, we choose γ_2 to be the leaf with lowest cost among the available choices for γ_2, and expand the leaves corresponding to γ_3. Notice that the values γ_i are only computed once, and determine the single hard-coded template path, which is used (after small adaptation) for all treatments.

The beam directions selected by the inverse planning system are not necessarily equal to the beam directions in the template path. As noted in section IV, beam orientations are selected using a spherical grid. We can thus simply use the grid of (reachable) node points to determine the beam orientations for inverse planning. Thus each treatment beam is given by an isocenter point p and the tool center point v_i of a node (fig. 10). Given appropriate limits for the tumor size we can assume that all isocenter points p are close to the workspace origin. In contrast, the source-

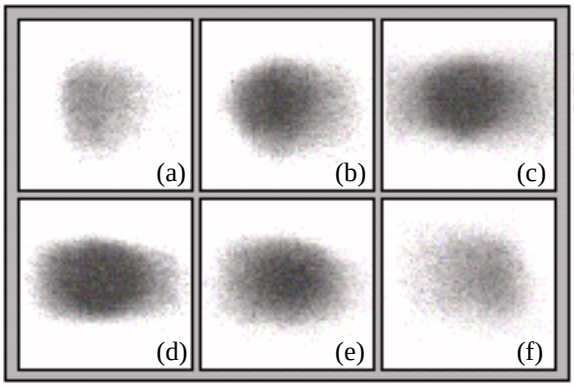

Fig. 11. Photographic film phantom for spherical distribution (axial cross-sections)

target distance for the beam (distance between p and v_i) is large. Thus we can ensure that the angular deviation δ between a node and the corresponding beam configuration during the actual treatment will remain below a fixed threshold δ_0. Furthermore, we must compensate for small patient motions during treatment. The treatment is interrupted if larger movements of the patient occur. When computing the template path, nodes causing intersections are removed. To allow for deviations of magnitude $\leq \delta_0$ from the path template, and to allow for tracking small patient motions, this intersection test uses appropriately enlarged representations of objects.

For verification, the template path in fig. 9-b was executed and a photographic film in a phantom skull (Plexiglas reconstruction of a skull, filled with water) was exposed to the generated radiation. The set-up is shown in fig. 1. An equal dose value of 20 cGy was delivered at each of 300 nodes evenly spaced along horizontal arcs. For each node, the beam axis was aimed at the work space origin. The set of reachable beams does not cover a full sphere (fig. 9-b), i.e. the nodes in this template path cover a range of surface area π when projected on a unit sphere. The restrictions in the workspace result in a penumbra along the horizontal direction in fig. 11-c, -d.

VII. Planning System

The tomographic images show the area of interest in a series of parallel 2D cross-sections, called slices. To specify a desired dose distribution, we proceed as follows. The slices are displayed of the control interface screen. The surgeon delineates regions of interest (tumor, critical regions) on the screen in each individual slice. Dose constraints are specified for each region. To prescribe a desired homogeneity, we specify both upper and lower dose constraints for this region. The desired drop-off of dose around the target volume may vary in certain areas. We thus allow for setting individual upper and lower constraints for edges of a surrounding polygon.

The delineation is done manually, and results in a stack of polygons for each region. We convert each stack into a three-dimensional representation of this region by thickening between slices. The input values for the planning

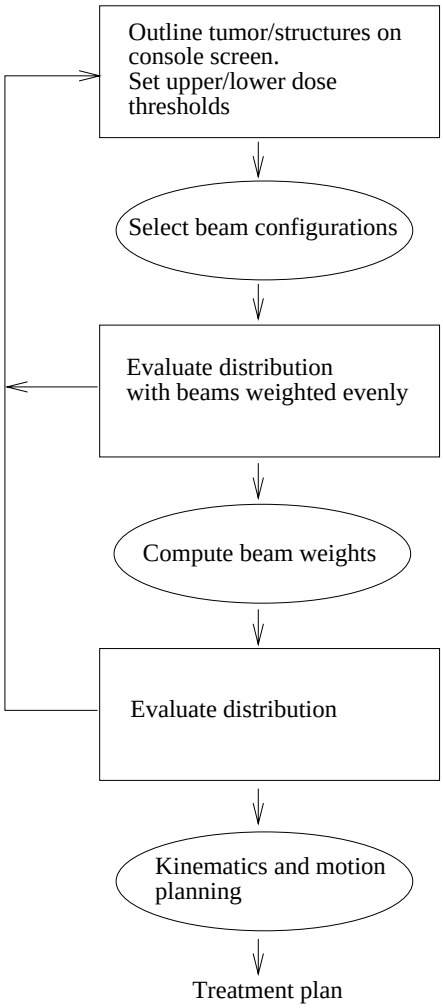

Fig. 12. Planning steps

system are: The number of beams to be used, the diameter of the collimator, delineations for the target and critical regions, and dose thresholds. The beam selection method in section IV is invoked to compute a set of beam directions corresponding to the input tumor shape and location.

The distribution is displayed and evaluated after completion of the beam selection phase, where all beams are weighted evenly.

The second planning phase computes beam weights. A set of constraints resulting from the upper and lower dose thresholds is computed with the methods in section V. The resulting set of equations is then solved with simplex-based linear programming. Fig. 12 summarizes the planning steps.

Currently, the system does not suggest changes of planning constraints if input constraints (upper/lower thresholds) are not satisfiable. However, the planning cycles for treatments with 100 beams are very short ($\leq$ 10 sec) so that dose thresholds for individual regions can be entered and tested with respect to satisfiability one by one, according to their priority in the given case. The number of necessary planning cycles depends on the planning con-

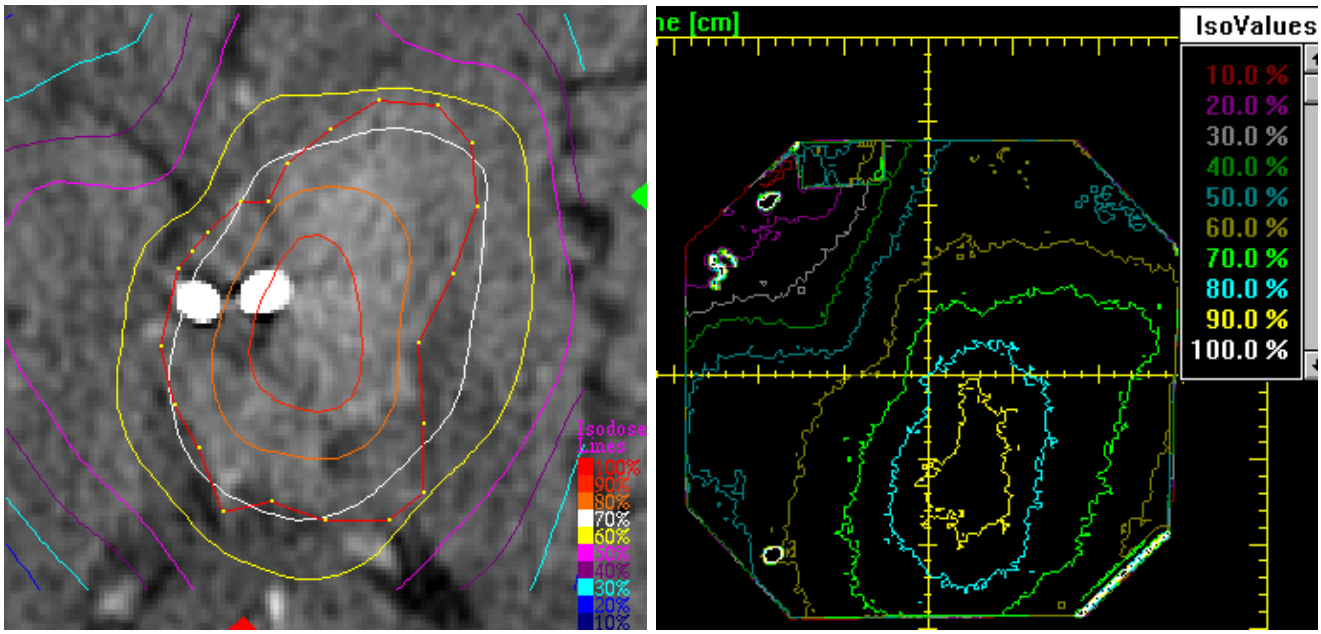

Fig. 13. Comparison of forward dose distribution calculated by the planning system versus actual radiation exposure in a water phantom. Treatment performed with general shape planning paradigm at Cleveland Clinic. Left image: patient anatomy, with isodose curves and tumor delineation. Right image: Radiation exposure in a water phantom obtained by executing planned treatment.

straints entered by the surgeon and is difficult to quantify in advance. In practice, two-three such planning cycles are typically sufficient for simple cases.

After computing beam weights, the direct dosimetry program is called to compute the dose distribution. Three types of graphical visualization are available to the surgeon for evaluating a treatment plan:

- Isodose surfaces in 3D, overlaid on a 3D tumor representation, and lines indicating beam directions.
- Color representation of dose distributions in planar cross-sections.
- Dose-volume histograms (DVH) showing the region volume (ordinate values) absorbing a dose x as a function of x (abscissa values), ([1]). Thus a dose-volume histogram correlates a dose value (given in percent of the highest dose) to the number of voxels absorbing that particular dose (see e.g. Fig. 18). A separate DVH is computed for the tumor, critical tissue, and for tissue surrounding the tumor.

The above visualizations refer to dose distributions computed in simulation. Photographic film phantoms are used for verification. A comparison between the dose computed in the planning phase and the resulting phantom is shown in Fig. 13. This figure shows a case treated with the new planning system and the general shape selection mode described above. The left image shows the anatomy in this case, together with the isodose lines for the plan computed by the planning system. The right image shows the physical distribution obtained by performing the actual treatment with a photographic film phantom.

In a test series, cases treated with frame-based stereotaxic radiosurgery (LINAC-system at Stanford Medical Center) were considered. In each case, data for the actual treatment were retrieved. Plans for the same cases were then computed for the robotic system, using the described planning techniques. Figs. 14-16 show a typical example. Fig. 14-a shows an axial cross-section of the brain for a patient presenting a large non-spherical tumor of the falx cerebri, delineated in the cross-section. Fig. 14-b shows the 3D reconstruction of the tumor outline. A regular grid of 600 points was expanded on a surface reconstruction, giving isocenter points. A lower dose threshold of 2000 cGy and an upper threshold of 2200 cGy was prescribed for the tumor region. The total computing time for one planning cycle (selecting beam configurations and computing beam weights, but excluding final forward dose calculation) was 1 minute on the SGI-Indigo control interface workstation.

Fig. 14-c shows the 80% isodose surface resulting from the computed motion. Fig. 14-d overlays the input tumor shape with the generated isodose surface to illustrate the matching between input and output shapes. For verification, the motion was executed with the robotic system while exposing a photographic film phantom (fig. 15). The phantom shows a close matching between the input tumor shape and the actual shape of the region receiving high radiation dose. The distribution for the treatment actually performed in this case with the LINAC-system was then recomputed. Fig. 16 shows the dose-volume histograms for both cases. The black curve is the curve obtained for the above plan generated with the beam selection mode c) for targets of general shape. For the conventional system, a plan with two isocenters was used. The treated volume is therefore obtained by overlaying two spheres. This results in a significantly less homogeneous distribution (gray curve in fig. 16).

It should be noted that comparisons of distribution characteristics (conformity, homogeneity and other characteristics) for different hardware systems (e.g., frame-based sys-

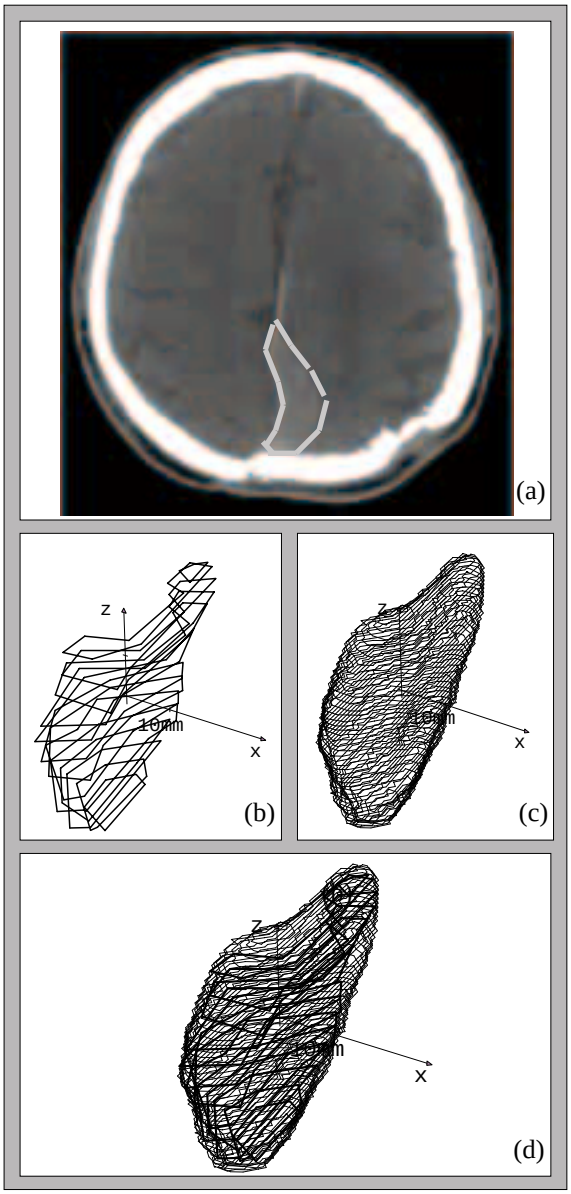

Fig. 14. Sample case (Stanford Medical Center). (a) Axial CT with tumor delineated in black (b) 3D representation of tumor delineation (c) Isodose surface (d) Isodose surface overlaid on tumor reconstruction

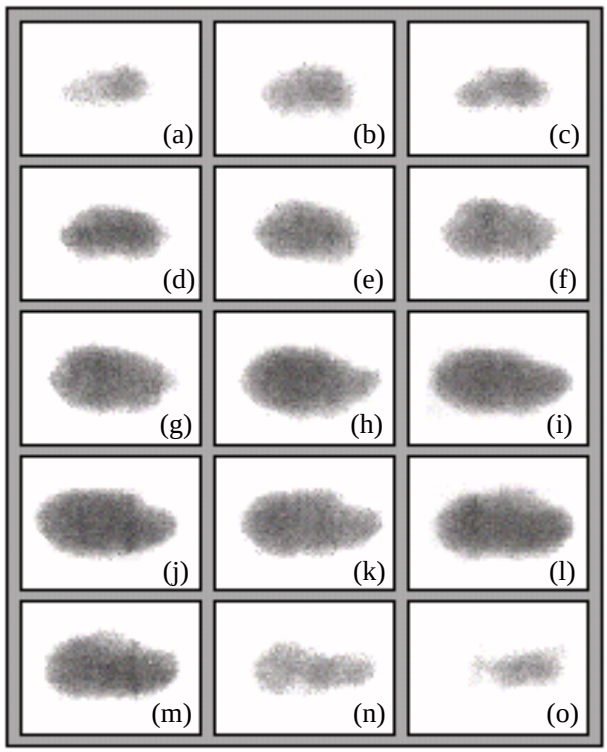

Fig. 15. Film phantom for sample case in fig. 14, showing the shape of the region receiving high dose. Film slices in the phantom are arranged in the same spatial structure as the tomographic images.

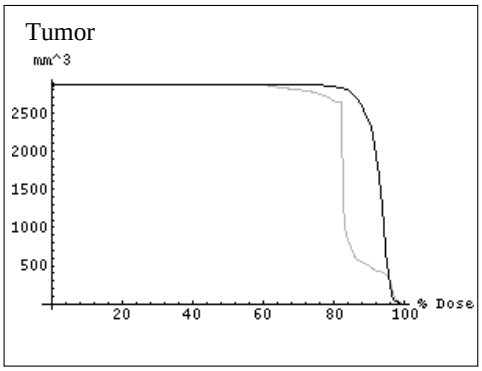

Fig. 16. Comparison of dose-volume-histograms for sample case (tumor). Gray curve: conventional treatment; black curve: general shape planning paradigm.

tems versus camera-guided systems) are problematic, since such characteristics must be considered together with treatment time. In most cases the conformity of distributions can be improved by using more complex, but also more time-consuming procedures. Improved conformity may result in less homogeneous distributions. Similarly, planning times for different systems are not directly comparable, since planning methods vary to a large extent with the complexity of the given case and procedures used in different clinics.

The following study shows typical treatment times for the new system and illustrates the trade-offs between treatment time and the achievability of planning criteria. In this case a synthetic target region (block-shaped region of size $10mm \times 10mm \times 20mm$) was used. Fig. 17 shows the distributions for plans with 25, 50 and 75 beam directions respectively. Fig. 18 shows the corresponding dose-volume histograms. Substantially higher conformity and homogeneity can be achieved, if more beams are used. However, this increases treatment time. Table 1 shows the respective treatment times for the current system version. In order to quantify the homogeneity in this table, we computed the integral over the dose-volume histograms in each case, and normalized to unit volume. Thus a perfectly homogeneous distribution (each point in the target region receiving exactly 100% of the maximal dose) corresponds to a (normalized) integral value of 100. The table shows that

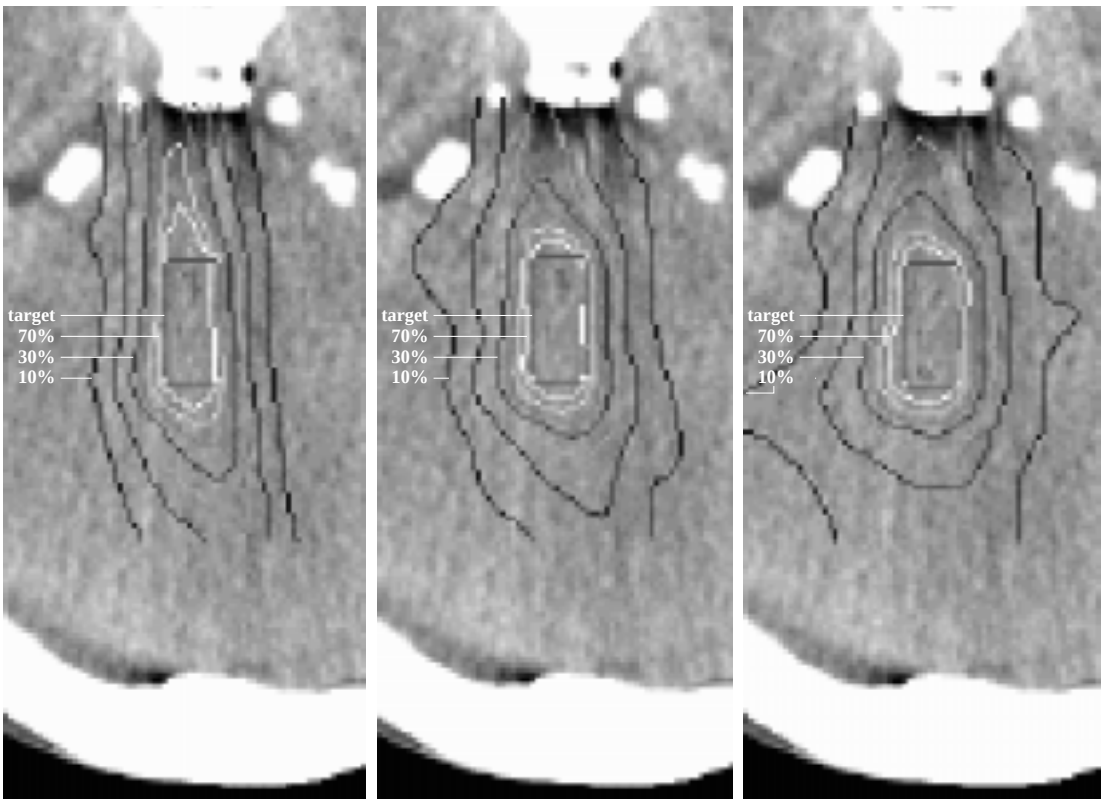

Fig. 17. Evaluation of total treatment time versus achievable dose conformity (matching between treated volume and rectangular target region), synthetic case. Images show isodose lines in percent of max. dose. a) Plan with 25 beams. b) 50 beams. c) 75 beams.

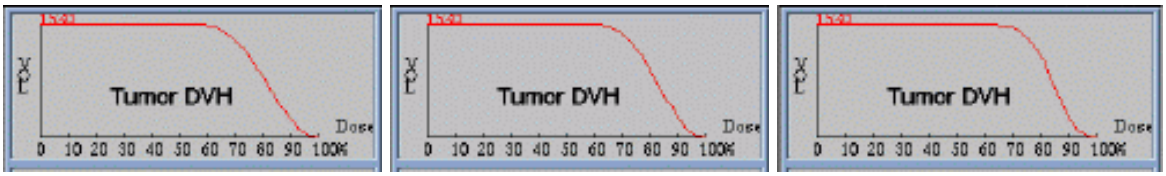

Fig. 18. Evaluation of total treatment time versus achievable dose homogeneity in the target region (synthetic case). Dose-volume histograms for target region: a) Plan with 25 beams. b) 50 beams. c) 75 beams.

higher homogeneity is achievable if the number of beams is increased. The treatment times correspond to approximately 1/3 for motion time (including communication), 1/3 image processing and 1/3 radiation time.

Number of beam directions	Treatment time	Integral (DVH for target)
25	15 min.	78.8
50	28 min.	81.7
75	40 min.	83.1

TABLE I
TREATMENT TIME VERSUS HOMOGENEITY OF THE RESULTING DOSE DISTRIBUTION (SYNTHETIC CASE).

VIII. Discussion

The new planning system is in clinical use at several leading medical institutions. Very recently, a number of treatments has been performed with the described beam selection mode for general target shapes, and the described methods for computing beam weights.

The system uses standard hardware components, and no stereotaxic frame is needed. This not only shortens treatment time, but also simplifies the overall treatment procedure. The planning system can generate highly conformal dose distributions without using special hardware for beam attenuation. Computing times for inverse planning are very short. Treatment interruption with manual repositioning of the patient can be avoided. This suggests that very complex radiosurgical treatments, including treatments previously considered impractical, can be performed in a very cost-effective way.

It is likely that the new methods will also allow for applying radiosurgery to extra-cranial tumors. This is currently not possible, since standard radiosurgical systems require rigid fixation of the anatomic site. Tumors in other parts of the body (outside the brain) are typically larger than brain tumors, and conformal treatment modes for non-spherical targets are particularly desirable.

Acknowledgments

This research was funded in part by the Sheik Enany Fund and the Lorraine Ulshafer Fund. The authors thank Rhea Tombropoulos, Jean-Claude Latombe, Don Caddes, Dan Halperin, Lydia Kavraki, Peter Kneschaurek and James Wang for comments and suggestions. We also acknowledge the help of Richard Cox from Stanford Medical Center with the implementation of forward treatment planning subsystem and beam model. Greg Glosser from Cleveland Clinic provided the example in fig. 13.

References

[1] Austin-Seymour, M. M., Chen, G. T. Y. et al.: Dose-volume-histogram analysis of liver radiation tolerance, *Intern. J. Radiation Oncology Biol. Phys.*, 12:31–35, 1986.

[2] Becker, G.; Kortmann, R., Kaulich, T.W., Duffner, F., Bamberg, M.: Gamma knife versus stereotactic linear accelerator. Utilization, clinical results and cost-benefit relations. *Radiologe* 36(4): 345-53, 1996.

[3] Barth, N. H.: An Inverse Problem in Radiation Therapy. *Intern. J. Radiation Oncology Biol. Phys.*, 18:425–431, 1990.

[4] Betty, O. O., Munari, C., and Rosler, R.: Stereotactic Radiosurgery with the Linear Accelerator: Treatment of Arteriovenous Malformations. *Neurosurgery*, 24(3):311–321, 1989.

[5] Gustafsson, A., Lind, B.K., Brahme, A. A generalized pencil beam algorithm for optimization of radiation therapy. *Med. Phys.*, 21(3):343-356, 1994.

[6] Craig, J. J.: Robotics, Mechanics and Control. Reading (Mass.): Addison-Wesley, 1982.

[7] Carol, M. P., Targovnik, H.: 3D planning and delivery system for optimized conformal therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 24, 1992.

[8] Khan, F.: The Physics of Radiation Therapy. Baltimore: Williams and Wilkins, 1984.

[9] Lutz, W., Winston, K. R., and Maleki, N.: A System for stereotactic radiosurgery with a linear accelerator. *Int. J. Radiation Oncology Biol. Phys.*, 14:373–381, 1988.

[10] Menguy, Y., Cinquin, P., et al.: External radiotherapy of prostatic carcinoma: a quadratic optimization of dose distribution. Proc. CVRMed-MRCAS, Lecture Notes in Computer Science, vol. 1205:675–684, 1997.

[11] Morrill, S.M., Lam, K.S., Lane, R.G., Langer, M., Rosen, I.I., Very fast simulated reannealing in radiation therapy treatment plan optimization. *Int J Radiat Oncol Biol Phys*, 31(1):179-188, 1995.

[12] Podgorsak, E. B., et al.: Dynamic Stereotactic Radiosurgery. *Intern. J. Radiation Oncology Biol. Phys.*, 14:115–126, 1988.

[13] Redpath, A. T., Vickery, B.,L., Wright, D. H.: A set of fortran subroutines for optimizing radiotherapy plans. *Comput. Programs Biomed.*, 5(2):158-164, 1975.

[14] Renner, W. D., O'Connor, T. P., Bermudez, N. M.: An algorithm for design of beam compensators. *Intern. J. Radiation Oncology Biol. Phys.*, 17: 227–34, 1989.

[15] Rosen, I., Lane, R. G., Morrill, S. M., Belli, J. A.: Treatment Plan Optimization using Linear Programming. *Medical Physics*, 18:141–152, 1991.

[16] Turk, G.: Re-Tiling Polygonal Surfaces. *Computer Graphics*, 26(3):55–64, 1992.