Shelling Procedure and Optimization by Simulated Annealing For Sphere Packing

Summary

We provide two models (model A and model B) for Gamma Knife treatment planning, one is a Mixed Integer Programming (**MIP**) model, and the other is an optimal sphere-packing model.

Based on dose distribution and the requirements of Gamma Knife unit, a Mixed Integer Programming (**MIP**) model is constructed and discussed.

Another model is a sphere packing approach to Gamma Knife Treatment Planning problem. Based on the theory of digital image processing and simulated annealing, an algorithm to obtain an initial sphere-packing plan is designed, and a method from the thought of simulated annealing is used to optimize the solution.

Due to the complexity of the MIP model and the large variables, the calculating would cost too much time. Many researchers have developed or have been developing other programming models. There are the same problems of time complexity. So we have to establish model B to speed up the computation and find the best or approximate best plan for the Gamma Knife Treatment.

In model B, we simulate some images from CT/MRI, and then translate them into digital-image as matrices by image processing. Then we get an initial result through **Shelling Procedure** and adjust further.

Our algorithm is recursion procedure for sphere packing. It determines the candidate location of the sphere center heuristic from largest sphere radius to the smallest one. As we know, when a ball moves inside a large close region, the track of the sphere center forms a close surface with the shape similar to the boundary of the region. We adopt heuristic method to choose one voxel as the center of the sphere. The idea of finding the largest distance between two voxels is that we should use the space efficiently. The spheres located in these two points can reduce the waste voxels near the boundary, and can spare more space for further utilization.

The solution obtained by **Shelling Procedure** must be adjusted further. From the thought of **Simulated Annealing**, we realize a program to maximum the sum of volume of spheres and to minimum the number of them.

Simulated computation shows model B satisfies the constraints of the Gamma Knife unit treatment planning and can be further studied to use in the real systems.

Background

Gamma Knife Unit[6,7,8,9,10]

The Gamma Knife is a highly specialized treatment unit that provides an advanced stereotactic approach to the treatment of tumor and vascular malformations within the head. The Gamma Knife delivers a single, high dose of gamma ray emanating from 201 Cobalt-60 unit sources. Inside a shielded treatment unit, beams from 201 cobalt-60 radioactive sources are focused so that they intersect at a certain point in space, producing an ellipsoidal region of high radiation dose referred to as a shot. (Figure 1)

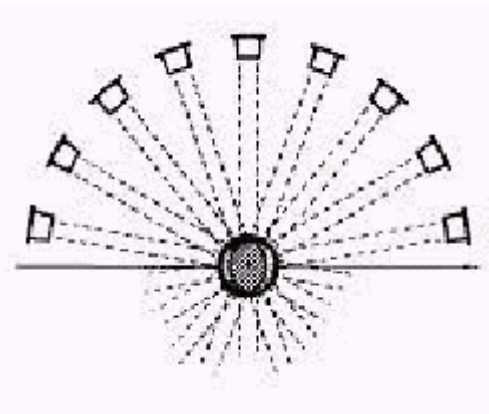

Figure 1: A shot of radiation is formed at the intersection of 201 beams

A brief history

In 1968, Professor Lars Leskell of the Karolinska Institute in Stockholm, Sweden and Professor Borge Larsson of the Gustaf Werner Institute at the University of Uppsala, Sweden developed the Gamma Knife. As far back as the 1940's, Leskell recognized the need for an instrument to target deep-seated intracranial structures without the risks of invasive open skull surgery. Currently, there are about 200 Gamma Knife machines worldwide.

Treatment Procedure

1. Fix patient’s head.
2. Use “magnetic resonance imaging” (MRI) or “computed tomography” (CT) scan the patient’s head to find the location and the volume of the tumor.
3. Develop the patient's treatment plan. Find a optimal multiple shots plan due to the irregularity and size of tumor shapes and the fact that the focusing helmets are only available in four sizes (4, 8, 14 and 18mm).
4. Deliver an efficient high dose of radiation to the target volume.

Treatment Goal

The plan aims to deliver a high dose of radiation to the intracranial target volume with minimum damage to the surrounding normal tissue. The treatment goals can vary from one neurosurgeon to the next, so a planning tool must be able to accommodate several different

requirements. Among these requirements, the following are typical, although the level of treatment and importance of each may vary.

1. A complete 50% isodose line coverage of the target volume. This means that the complete target must be covered by a dose that has intensity at least 50% of the maximum delivered dosage. This can be thought of as a “homogeneity” requirement.
2. To minimize the nontarget volume that is covered by a shot or the series of delivered shots. This requirement is clear and can be thought of as a “conformity” requirement.
3. To limit the amount of dosage that is delivered to certain sensitive structures close to the target. Such requirements can be thought of as “avoidance” requirements.

In addition to these requirements, it is also preferable to use as small number of shots as possible to reduce the treatment time.

Problem Analysis

The goal of stereotactic radiosurgery for a brain tumor is to deliver the desired dosage to the target, and only the target. This is not possible in reality. So they do the next best thing, which is to deliver enough dosage to the target, to avoid as much normal tissue as possible, and to deliver as little radiation as possible to whatever normal tissue must be affected. There are two additional important criteria–dose homogeneity and dose conformality. That is, we do not want ‘hot spots,’ which have been experimentally determined to cause complications; and we do want rapid falloff of dose levels outside the actual tumor. As for simplification, we can firstly consider the focused beams an ideal sphere solid. That means, the dose inside the sphere is homogeneous and the dose level of sphere of different size is the same. And the dose outside the sphere is 0, take D=4mm sphere for example (Figure 2).

Figure 2: diameter of 4mm sphere dose distribution curve

The figure shows that a shot merely acts upon in the spherical region; the outside of the sphere is not affected by the shot. Therefore, the target volume can be filled with several shots, and the problem can be reduced to sphere packing problem.

However, the dose distribution of a shot in practice is not a simple step function with 0 dose outside the sphere and the dose inside the sphere is not homogeneous, that is, the dose gradient exists. In this situation, we can reduce the target volume to a smaller size and we also use sphere packing plan to solve this problem.

Basic Assumptions

- 201 beams of radiation simultaneously intersect at the isocenter forming an ideal spherical dose distribution.
- A shot is a non-elastic, solid 3D sphere.
- The shot sphere diameters are the same size as the diameters of collimator (diameter of 4, 8, 14 and 18mm).
- The tumor volume is not large. The length, width and height of a tumor region are from 20mm to 40mm.
- The dose distribution is that the dose level inside the shot sphere is high enough to kill cancer cells, and the dose outside is low enough.
- A treatment plan is considered acceptable if 50% isodose curve encompasses the target.
- We consider isodose curve encompasses the target region tightly and closely.

Constructing the Models

Model A: Model Based on Mixed Integer Programming

Dose distribution model

Let x be the distance from the isocenter of a dose sphere, r be a measure of the radius of the sphere, A sum of error functions has been noted in the literature to approximate this dose distribution. Then the radius dose distribution for x can be expressed as:

$$1-\sum_{i=1}^{n} m_i \cdot erf\left(\frac{x-r_i}{\sigma_i}\right) \qquad \left(\sum_{i=1}^{n} m_i = 1,\ n=2 \text{ for simply}\right)$$

where $erf(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$.

According the above expression, we could draw the images that dose distribution (Figure 3) and effective area of dose (Figure 4) approximately.

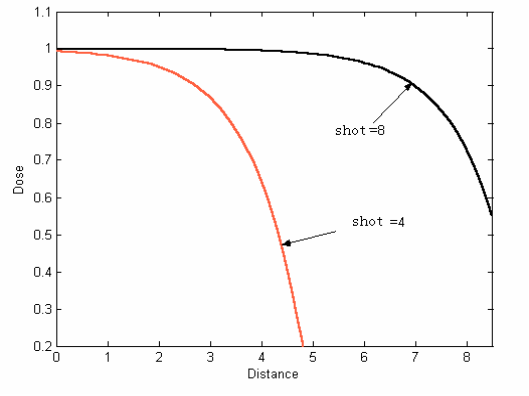

Figure 3: Dose distribution

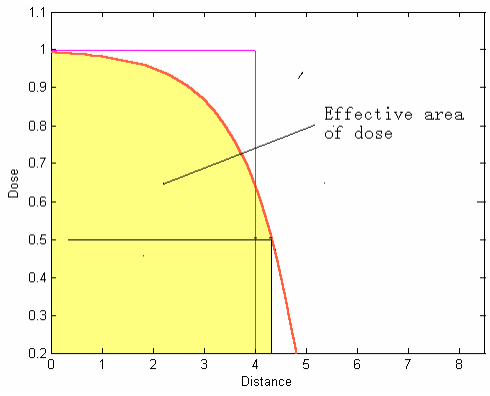

Figure 4: Effective dose distribution area

Description of Model

The complete dose distribution can be calculated as a sum of contributions from each shot delivered, once the location of the center of that shot (x_s, y_s, z_s) is known, and the length of time of delivery t_s ; w is known as four kinds of diameter. In practice, this means that for all (*i*, *j*, *k*), we restrict the shot locations to be within the target area, set shot location on grid, (we generate grid on target that values 1mm (Figure 5), and use binary variables to indicate if a pair of (shot location, shot size) is used or not. We choose a grid of possible shot location, and pre-calculate for each grid location, every pixel and each r and use the optimization algorithm to decide whether or not to use a shot as a particular location. Since the only shots that are considered are shots that lie within the target, so we could easily determine whether or not a shot lies within the target by the terms of track bounds of target. When the description of the dose is determined an optimization model can be formulated. The basic variables of the optimization include the number of shots of radiation that will be delivered, along with the coordinates of the center location of the shot (x_s, y_s, z_s), the dose of any point in the target $D_s(i, j, k)$, the time t_s that each shot is exposed, the T of bound on the length of time that a particular shot can be exposed, and ψ_s (heuristically. If use the shot *s*, then, ψ_s =1, else ψ_s =0).

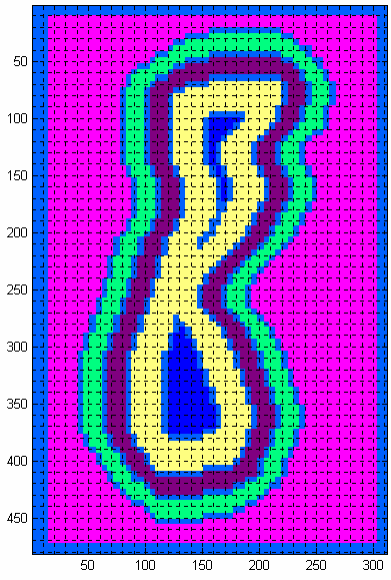

Figure 5: Grid on figure

Now, the problem is reduced to solving a MIP (Mixed Integer Programming):

$$Min: Dose(i,j,k) = \sum t_s * D_s(i,j,k)$$
$$S.t: 1 \le Dose(i,j,k) \le 2$$
$$\underline{T}\psi_s \le t_s \le \bar{T}\psi_s$$
$$\sum_{s \in S} \psi_s \le N$$
$$\psi_s \in \{0,1\}$$
$$D(x_s, y_s, z_s, i, j, k) := D_s(i,j,k)$$

Due to the complexity of the MIP model and the large variables, the calculating would cost too much time. Other programming models were constructed by many researchers [6,7,8,9,10], there were the same problems of time complexity. So we have to establish another model to simplify the problem.

Model B: Model Based on Shelling Procedure and Optimization by Simulated Annealing

Description of Model

Under the assumption that a shot is a non-elastic, solid 3D sphere, the radiosurgical treatment planning can be deduced an optimization of packing unequal spheres into a three-dimensional bounded region. Given an input (R; V; S; L), where R is a 3D bounded region, V a positive integer, S a multiset of spheres, and L a location constraint on spheres. We want to find a packing of R using the minimum number of spheres in S such that the covered volume is at least V, the location constraint L is satisfied; and the number of points on the boundary of R that are touched by spheres is maximized. Wang [14] shows that not only finding an optimal solution to the problem is computationally intractable, but also optimization of the related problems is NP-hard. Therefore, some sort of approximation is needed. The paper [2,3] proposes a model under the assumption that spheres are available with unlimited supply, the 3D bounded region is a polytope, and there are no location constraints. The model is a nonconvex optimization problem with quadratic constraints and a linear objective function. The computation complexity of the model of [3] is very high.

Shelling Procedure to produce an initial solution

Our treatment planning for a gamma knife unit is based on sphere-packing problem. We establish a shelling procedure to solve sphere-packing problems.

Before we depict our procedure, it is necessary to know the following principles:

The sequence of packing is from the sphere of the largest diameter to the smallest one.

Definition: *Considering a voxel as a three-dimensional box. The length, width and the height of each voxel is 1unit.*

The major steps of this algorithm are as follows:

Step1: Transfer the data of the boundary of the target to three-dimensional 0-1 arrays. The voxel value in the target region is 1, and the voxel value outside the region is 0.

Step2: Find a reasonable sphere center.

Pack from the largest sphere to the smallest one. First, we pack the sphere of diameter D=18mm into the region. As we know, when a small ball moves inside a large close region, the track of the sphere center forms a close surface with the shape similar to the boundary of the region. However, the shape similarity is small as the diameter increases.

The method to obtain the close surface of a sphere center is similar to the methods in [1,7].

Find two voxels on the close surface, the distance between which is the largest. We choose one voxel as the center of the sphere in a heuristic way.

The idea of finding the largest distance between two voxels is that we should utilize the space efficiently. The spheres located in these two points can reduce the waste of boundary voxels, and can spare more space for further utilization.

Step3: Remove the spherical region. Find whether the sphere of D=18mm can be packed into the remaining region, go on packing the sphere until the remaining region cannot be contain the sphere of this diameter. Repeat **step2** using a smaller sized sphere until the remaining region cannot contain any sphere of all sizes.

Step4: Find whether the total volume of all sized spheres meets the given requirements, and the sphere number is reasonable.

Step5: Store the location of centers of all the spheres and their diameters, thus form an initial solution, which is a feasible Gamma Knife treatment plan.

The solution may not be the best plan. A further optimal algorithm will be depicted in the next section.

A two-dimensional example of above procedure:

As a two-dimensional situation, the close region can be seen as a close curves. Spheres can be reduced to given sized circles. According to above algorithm, we select a close plane region arbitrarily and make a trial of above procedure; the procedure is shown in Figure 7.

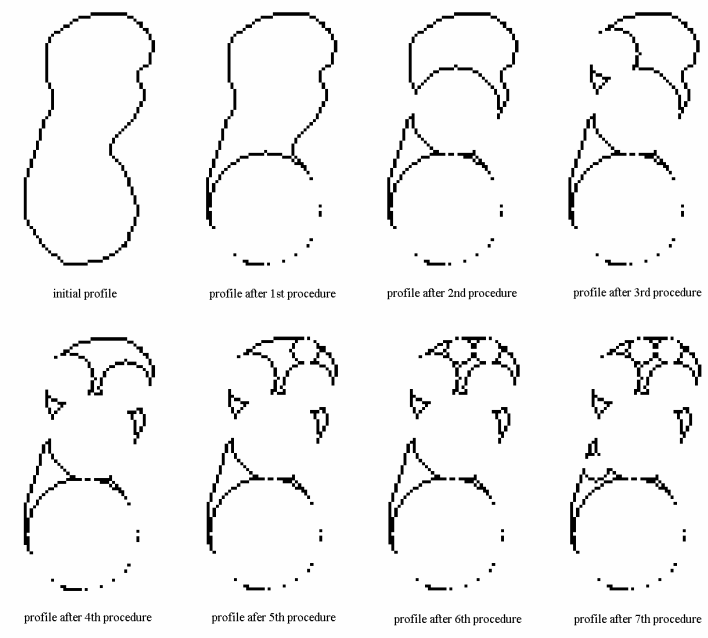

Figure 7: the procedure of circle packing using above algorithm

Procedure details of Figure 7:

1. Generate the 0-1 array of given region.
2. Find the tracks of the largest circle center (Figure 8), if the track does not exist, that means the profile cannot totally contain the circle, we may choose a smaller sized circle, and find the track again.

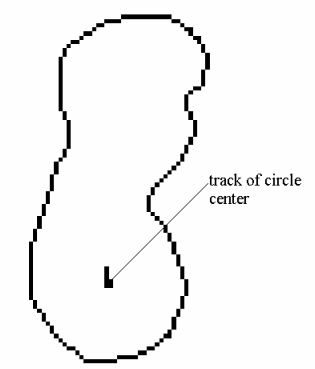

Figure 8: track of circle center

3. Then, we may find two points on the track, the distance between which is the largest. Record their locations. Select one location, place the circle, and remove the circle area from initial region; repeat the above procedure, find the largest circle or circles (given four sizes) that can be placed in the remaining region until none of the four sized circles can be placed in the remaining region; record the result.
4. Compute the area of all circles in each situation; if the computed area is more enough, such as 90% of origin area and the shot number is less, the better the result is.

Optimization by Simulated Annealing

Notice that simulated annealing method can be applied to optimize problems. The initial solution obtained by **Shelling Procedure** must be further adjusted. The objective function is using as few spheres as possible to occupy maximum space (more than 90% of original space). Denote the objective function as:

$$F_n = V - \frac{4}{3}\pi \sum_{i=1}^{N} r_i^3$$

where F_n is the volume of the residual fractions. V is the target volume. r_i is the radius of the sphere i, N is the number of spheres. The solution space is a 4-dimensional space consisting of three continuous and one discrete variable parameters. Denote one solution as:

$$S = \{(x_i, y_i, z_i, r_i), i = 1,2,\cdots,N\}$$

where (x_i, y_i, z_i) is the coordinate of the center of the sphere i. The solution space is so large that they cannot be explored exhaustively. The adjustment algorithm is as follows:

Step 1: Calculate the objective function value of the initial solution. Then we obtain F_0 and

S.

Step 2: Obtain the shift matrix W. It must be have the same degree of S.

$$W = \{(h_i, p_i, q_i, k_i), i = 1,2,\cdots, N\}$$

Where h_i, p_i and q_i can be obtained by random way, and all of these data obey the standard normal distribution from $-\varepsilon$ to ε. When the temperature t_m is high, we should set ε larger. Where k_i is the special data which will be described as follows:

Step 3: Obtain new potential solution by $S + W$.

$$S' = \{(x_i + h_i, y_i + p_i, z_i + q_i), i = 1,2,\cdots, N\}$$

Where $k_i + r_i$ must be one element of the set (0, 4,8,14,18). And $k_i + r_i$ can only be near r_i.

For example, if r_i = 18, then $k_i + r_i$ can select 14 or 18. If r_i =4, then $k_i + r_i$ can be 0, 4 or 8.

Step 4: Test if S' is the proper solution, in other words, if S' obeys the restriction: (1) the distance between each two center of the spheres is lager than or equal to the sum of the radius of these two spheres. (2) the distance from the center of any sphere to the boundary of the target volume is larger than or equal to the sphere's radius. If S' meets the restriction, continue **steps 5**, or go to **step 2**.

Step 5: Calculate the value of the new solution, then we obtain F_0'. If F_0' is smaller than F_0, in other words, S' is better than S, so let $S = S'$, $F_1 = F_0'$. Then we have the new solution, go to **step 6**.

Use **Shelling Procedure** to estimate if we can add one or more sphere to fill lacuna. Note that this way may lead to adding the dimension of the solution matrix. If F_0' is larger than F_0, then if

$$\exp\left(\frac{F_0' - F_0}{t_m}\right) > random(0,1)$$

S' is better than S, go to **step 6**. Otherwise go to **step 2**.

Step 6: When new better solution is found, if $|F_n - F_{n+1}| > \zeta$ (where ζ is a small number we preassigned.) go to **step 1**.

We should reduce the temperature if $|F_n - F_{n+1}| < \zeta$ (which means we have found the best solution at this temperature), but how to realize it? We have many ways. In our program, we select Lundy and Mess' method:

$$t_{m+1} = \frac{t_m}{1 + \beta \times t_m}$$

Where $\beta = \dfrac{t_0 - t_f}{M \times t_0 \times t_f}$

t_0 and t_f are two critical value. t_0 is the lowest temperature and t_f is the highest. Both of them have been designated at advance. M is the largest number of adjustable times. In paper [13], the iterate method is one time per each temperature, so the total of iterative times is M.

Step 7: Let $t_m = t_{m+1}$, set ε a lower value, go to step 1.

One example on 2-D is as follows (Figure 9):

(a) initial result (b) final result

Figure 9: comparison of initial result and the final result

Figure 9 (b) is adjusted from (a). The area of all circles in initial result to total area of origin shape is less than 84% and the shots number is 10, while the area of all circles in final result to total area of origin shape is more than 90% and the shot number is 9. So we get a satisfied result.

Evaluation of Our Models

Model A was established on the basis of dose distribution. We find that many optimization of Gamma Knife Radiosurgery are also based on this distribution. So Model A is available in theory. But in fact the model is a MIP problem. It generates a large enormous data, and is difficult to calculate within time variable. But in some case MIP is flexible and practical on account of that MIP could find global optimization.

Shelling Procedure and Simulated Annealing cooperating with each other made our model and method preferable. Shelling Procedure has an advantage on the initial solution: we search for the possible area of spheres' center. The area is very important to our work. All of the solutions are to be contained in this area. Based on this fact, we find better solutions quickly, but we cannot use this method to optimize the result which is produced by itself. Simulated Annealing needn't deal with the difference of every target volume. We only need a few data. Computer can treat all other business, and the efficiency can meet the requirements. And it

can improve the solution undoubted. Furthermore, we can watch the total movement course, for example, our program can realize it like a cartoon movie. The process is similar to the genuine steel annealing. Like all Simulated Annealing, how to define solution space and objective function is very hard. It is inadvisable to use Simulated Annealing to find the initial solution, the cooperation of **shelling procedure** and the Simulated Annealing is a efficient way to solve this problem.

References

1. Taeil YiI, *A Tree for a Brain Tumor*. Florida Section Newsletter. February 2001,Volume 22, Issue 2.
2. A. Sutou and Y. Dai, *A Study of the Global Optimaization Approach to Spherical Packing Problems*, Dept. of Math. and Comp. Sciences Research Reports B-361, Tokyo Institute of Technology, May 2000, submitted, revised July 2001.
3. A. Sutou and Y. Dai, *Global Optimization Approach to Unequal Sphere Packing Problems in 3D*, Journal of Optimization and Applications.
4. A.Rosenfeld and A.C.Kak. *Digital Picture Processing*, V.2, 2nd Edition, Academic Press, 1982.
5. Kenneth R. Castleman, *Digital Image Processing*, Prentice Hall 2000.
6. M. C. Ferris, *MATLAB and GAMS interfacing optimization and visualization soft ware*, Technical Report Mathematical Programming Technical Report98-19, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, 1998.
7. M. C. Ferris, J. Lim, and D. M. Shepard, *An optimization approach for Radiosurgery treatment planning*, SIAM Journal On Optimization, Forthcoming,(2002).
8. M. C. Ferris, J. Lim, and D. M. Shepard, Radiosurgery optimization via nonlinear programming, Annals of Operations Research, Forthcoming, (2002).
9. Michael C. Ferris, Jinho Lim, and David Shepard, *Radiosurgery Optimization via Nonlinear Programming*, Annals of Operations Research, vol 119, pp 247-260, 2003.
10. Jinho Lim, *Optimization in radiation treatment planning*, PhD Thesis, December 2002, University of Wisconsin - Madison.D. P. Bertsekas, Network Optimization: Continuous and Discrete Models, Athena Scientic, 1998.
11. L. Luo, H. Shu, W. Yu, Y. Yan, X. Bao, and Y. Fu, *Optimizing computerized treatment planning for the Gamma Knife by source culling*, International Journal of Radiation Oncology, Biology and Physics, 45 (1999), 1339-1346.
12. H. Z. Shu, Y. L. Yan, X. D. Bao, Y. Fu, and L. M. Luo, *Treatment planning optimization by quasi-newton and simulated annealing methods for gamma unit treatment system*, Physics, Medicine and Biology, 43 (1998), 2795-2805.
13. Lai K K, Chan J W M. *Developing a simulated annealing algorithm for the cutting stock problem*, Computers Ind. Engng, 1996.
14. J. Wang, *Packing of unequal spheres and automated radiosurgical treatment planning*, Journal of Combinatorial Optimization, 3 (1999), 453-463.