INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Development of Techniques for Optimization and Verification of Radiation Treatments

Dimitre H. Hristov

Department of Physics
McGill University, Montreal
April, 1998

A Thesis Submitted to the Faculty of Graduate Studies and Research in Partial Fulfillment of the Requirements of the Degree of Doctor of Philosophy

На Катя

Не тъгувай, животът е кратък.
Всичко тук е - и пъкъл и рай.
Всичко вземай и дай без остатък,
работи и обичай докрай.

Салис Таджер

Abstract

Algorithms for optimization and verification of radiation treatments have been developed. The first one, an *active set algorithm* for inverse treatment planning employs a conjugate gradient routine for subspace minimization in order to achieve a higher rate of convergence than the widely used constrained steepest descent method at the expense of a negligeable amount of overhead calculations and storage. The active set algorithm is found to be superior to the constrained steepest descent in terms of both its convergence properties and the residual value of the cost functions at termination. The active set approach can significantly accelerate the process of inverse treatment planning by decreasing the number of time consuming dose calculations.

The second algorithm employs a *continuous penalty function method* to solve approximately a large-scale constrained minimization problem which reflects the goal of sparing healthy tissues as much as possible while delivering the necessary tumorcidal dose. The performance of the continuous penalty function method is optimized by a numerical investigation of a few integration schemes and a pair of weighting functions which influence the performance of the method. Clinical examples are presented that illustrate possible applications of the techniques in the context of multi-objective optimization.

An image correlation based algorithm for automatic registration of pairs of portal images has also been developed. Accounting for both in-plane translations and rotations, the algorithm uses fast-Fourier-transforms and a sequential approach to speed up the registration without degrading the accuracy of the match. The technique has also been applied to the automatic registration of portal images to digitally reconstructed radiographs (DRRs) which have been modified to resemble megavoltage images. The results indicate the feasibility of this approach as a tool for treatment setup verification.

Résumé

Dans cette thése, des algorithmes servant à optimiser et à vérifier les traitements en téléradiothérapie ont été développés. Le premier, un algorithme d'ensemble actif s'avère utile dans la planification inverse des traitements. Plutôt que d'utiliser la méthode conventionelle de la descente la plus rapide à contraintes l'algorithme d'ensemble actif emploie une routine de gradient conjugué ce qui donne une convergence précose et une plus petite valeur de la fonction de pénalité lorsque la convergence est atteinte. En conséquence, en diminuant le nombre nécessaire des calculs dosimétriques, l'algorithme d'ensemble actif peut accélérer le processus de la planification inverse des traitements d'une façon significative.

Le deuxième algorithme utilise une fonction de pénalité continue afin de résoudre d'une manière approximative le problème de minimisation à contraintes en téléradiothérapie résultant du double objectif de donner à la tumeur une dose importante tout en épargnant les tissues sains autant que possible. La performance de cet algorithme est accrue par un examen numérique de plusieurs schémas d'intégration et de deux fonctions représentatives d'importance des structures anatomiques. Des exemples cliniques de l'utilisation de cette technique sont présentés.

Le dernier algorithme sert à la juxtaposition automatique de deux images portales. Celui-ci tient compte des translations et des rotations dans le plan de l'image. L'algorithme utilise des transformations rapides de Fourier et une approche séquentielle pour accélérer la juxtaposition sans nuire à son exactitude. Cet technique est aussi propice à la juxtaposition d'une image portale et d'une radiographie digitalement reconstruite qui est modifiée afin de ressembler à une image obtenue par un faisceau thérapeutique de haute énérgie. Les résultats démontrent la valeur de cette approche pour la vérification du positionnement du patient par raport au champ d'irradiation.

Original Contribution

An analysis is performed of the computational properties of treatment planning optimization with intensity modulated beams. For differentiable objective functions, the conjugate-gradient methods are identified as optimal gradient minimization techniques to tackle the large-scale inverse problem. A robust active set algorithm with a conjugate-gradient routine for subspace minimization is designed in order to account for the physically imposed non-negativity constraints on the independent variables. The theoretically expected superiority of the active set method to the widely used constrained steepest descent is confirmed numerically for two largely accepted treatment objectives and different irradiation geometries.

A novel formulation of the inverse problem is suggested which uses only target dose levels. A mathematical analysis is conducted which proves that any local solution of the resulting constrained minimization problem is a global solution. A continuous penalty function method is introduced as an approximate numerical technique to perform the large-scale constrained optimization. The applicability of the method to the particular statement of the inverse problem is proved. Several numerical integration techniques are investigated and an optimal one is identified. A procedure for the clinical use of the technique is introduced which allows autonomous determination of the target importance weight that produces a clinically acceptable target coverage.

In-plane rotation search is implemented in a correlation based portal-image registration algorithm which adopts a full calculation of correlation integrals rather than their sampling through Monte Carlo techniques. A fast Fourier transform (FFT) implementation of the Pearson correlation coefficient (PCC) is derived. Investigation is performed of the properties of the PCC with respect to in-plane rotations. These properties are consequently exploited to design a sequential search for the transformation parameters in

order to accelerate the image registration. The utility of two different correlation operators have been tested for image matching: the PCC and the normalized cross-correlation. The theoretical indication that the PCC is more robust for registration of images with linear transformations of the intensities is experimentally confirmed. The feasibility of the image correlation approach to the automatic registration of portal images to digitally reconstructed radiographs (DRRs) is also demonstrated.

Acknowledgments

I would like to express my sincere gratitude to Dr. B. Gino Fallone, my thesis supervisor, for his guidance, support and understanding throughout my tenure as a student at McGill University. Without the excellent working environment he has created the completion of this project would not have been possible.

I would also like to extend my thanks to Dr. Ervin Podgorsak, Director of the Medical Physics Unit, for giving me the opportunity to enroll in one of the best educational programs in medical physics in North America. His consideration for and attention to the graduate students are greatly appreciated.

I am grateful to Mr. Tony Falco, M.Sc., for the numerous useful suggestions and discussions on several aspects of this work. His help and support during these years will always be remembered.

I thank Mrs. Marina Olivares, M.Sc., and Mr. Michael Evans, M.Sc., for the interesting and enlightening discussions on inverse treatment planning and for their encouragement during this project.

Last, but not the least, I would like to thank Mrs. Boriana Mihova, Mr. Kiril Georgiev, Mr. Vencislav Iliev and Mr. Boiko Ivanov for the lasting warmth of their friendship, for the invigorating strength of their beliefs and for their mesmerizing power to bring life beyond the mundane routine.

The research presented in this thesis has been financially supported in part by the Medical Research Council of Canada (MT 13110). The Teaching Assistantship from the Physics Department is also gratefully acknowledged.

List of Figures

List of Tables

Contents

CHAPTER 1 Introduction

1.1 Introduction to Radiation Therapy

Cancer is the second leading cause of death in Canada,[1] accounting for 60 000 victims annually.[2] In 1997, 130 000 patients are expected to be diagnosed with cancer in this country.*[2] Even though hyperthermia,[3] immunotherapy[4] as well as other novel therapies[5, 6] have been actively explored, the majority of the cancer patients will be subject to some combination of surgery, radiotherapy and chemotherapy. Approximately half of the cancer patients will be administered radiation at some point in the course of their treatment.[7]

Radiation therapy aims at tumor eradication by means of ionizing radiations, the most commonly employed types being photons and electrons, with energies ranging from a few hundred keV to a few MeV. Among all cellular perturbations caused by the ionizing radiations,[8] the double-strand break in DNA is the predominant cause for radiation-induced cellular death.[9] The number of double-strand breaks is related to the physical quantity *absorbed dose*[10] or simply *dose* which is the energy deposited by the radiation per unit mass of material (The *SI* unit of dose is *gray* (Gy) defined as 1 Gy = 1 J/kg). With increasing dose to the tumor volume, the number of killed cancer cells increases and so does the probability of cure. The random nature of the radiation damage[11–14] determines a sigmoidal shape of the relation between the dose and the tumor control probability (TCP) which is the probability that there are no surviving clonogenic cells in the tumor (Fig. 1.1). The tumor lethal dose (dose required to achieve 95% TCP) varies between 2 000 cGy to more than 8 000 cGy depending on the tumor size, extent, type, radiosensitivity as well as on the tumor pathologic grade and differentiation.[15] The normal tissue complication probability (NTCP), which is the probability of inducing some particular complication (end-point) in a non-tumor-bearing organ(s), is a similar

* This number excludes the estimated 61 000 cases of non-melanoma skin cancer.

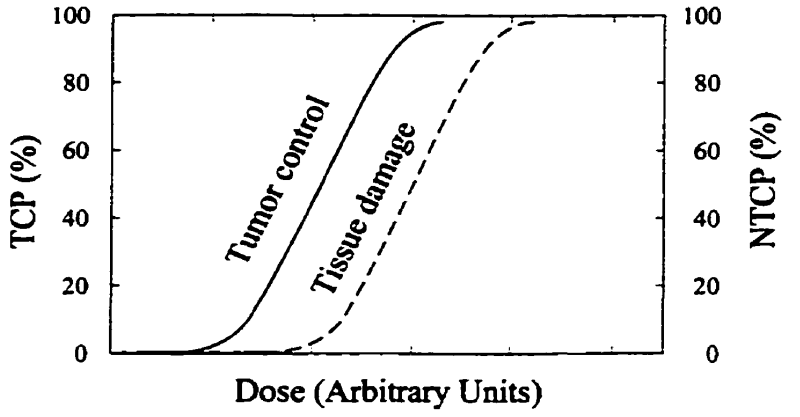

Figure 1.1. *Response curves of tumor cure (TCP) and normal tissue complications (NTCP). The mutual position of these curves varies in different clinical circumstances depending on the radiosensitivity of the tumor and the involved normal tissues.*

function of the dose (Fig. 1.1) since the same principles of radiation killing apply to both tumor and normal cells. Both the TCP and the NTCP of the involved healthy tissues are considered when the value of the dose to be deposited in the tumor volume is selected.

Radiation can be administered to a patient by sealed radioactive sources in catheters which are inserted in the tumor volume (brachytherapy) or most commonly by directed megavoltage external beams of x-rays or electrons (external beam therapy or teletherapy). A typical teletherapy treatment unit (a linear accelerator or a cobalt unit) uses a radiation source mounted on a rotating gantry capable of moving around a patient who lies on a treatment coach (Fig. 1.2). In many cases, multiple properly collimated beams from different directions can be used to deliver the necessary tumorcidal dose and to minimize the dose to the healthy structures surrounding the tumor.

The process of external beam therapy consists of several steps. First, a patient diagnosed with cancer undergoes a thorough evaluation that aims at the determination of

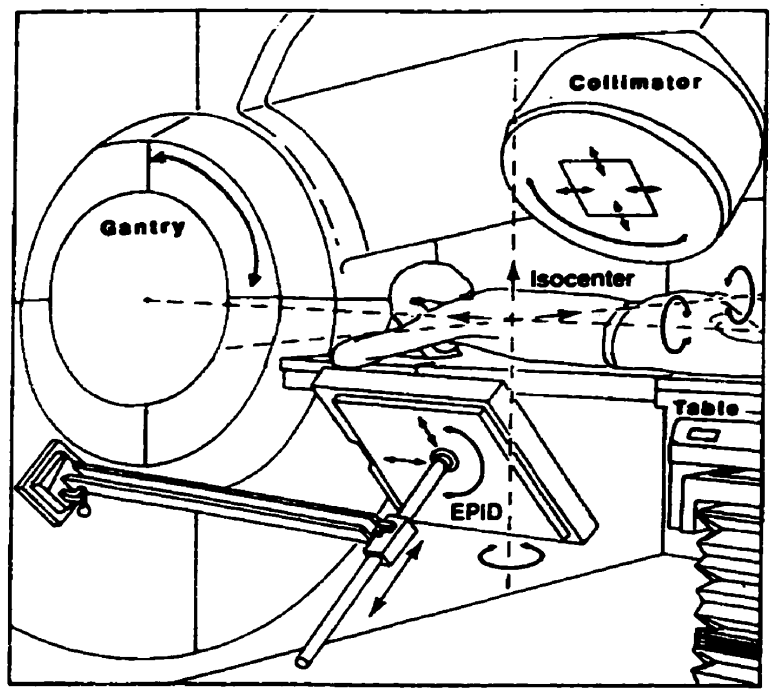

Figure 1.2. *External beam treatment unit (With modifications from Bijhold et al.*[16]*).*

the tumor volume, its extent and relationships to critical structures in the body. During this process, the radiation oncologist uses palpation, biopsies and the information from various imaging procedures such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), single photon emission tomography (SPECT), and diagnostic ultrasound, in order to define the gross tumor volume (GTV)[17] which comprises the gross extent of the malignancy. Local subclinical tumor spread is then included in the clinical target volume (CTV)[17] which is defined as the GTV plus a margin. The GTV and the CTV are based on anatomical, biological, and clinical considerations. A planning target volume (PTV)[17] is then defined as the CTV plus a

margin, which accounts for the technical aspects of the treatment such as organ motion and variations of patient's positioning. In addition to the tumor volume, critical structures to be spared during the radiation treatment can also be identified and outlined.

The CTV, the PTV as well as the outlines of the critical structures are used as input data for *treatment planning*, which is the next step in the course of radiotherapy.† A beam arrangement (number of beams, type, energy, apertures, directions and modifiers) is selected and the dose distribution within the patient is calculated and evaluated. This process continues until the resulting dose distribution achieves an adequate coverage of the target volume (CTV and PTV) and acceptable sparing of the critical structures.

To aid the process of treatment planning and verification, the patient's bony anatomy can be imaged on a simulator which is a diagnostic x-ray unit that mimics the geometry, the alignment as well as the movements of the actual treatment unit. The process of treatment simulation can be performed prior to and/or subsequent to treatment planning. During the simulation, localization of the target can be performed by referencing the position of anatomic structures relative to skin marks. An appropriate beam arrangement can be selected by examining the relative position of the various organs at different gantry angles. The shape and the position of shielding blocks can be verified by placing the blocks in position and by obtaining radiographs under the geometry in which the treatment is to occur. These radiographs also serve as reference images for the verification of the patient setup during the actual treatment. CT simulation,[18–20] which utilizes the patient's CT data and various image processing techniques is being increasingly used to perform some of the above functions of the conventional simulator. As a result of such simulation, *digitally reconstructed radiographs* (DRRs) are produced which reflect the desired setup of the patient's anatomy with respect to the radiation fields. A DRR is simulated by

† In this thesis we will frequently use the generic term 'radiotherapy' to denote external beam therapy.

projecting rays from the intended position of the radiation source through the patient's CT data (patient's anatomy) to the pixels in the plane of a virtual imaging device. Line integrals are evaluated by summing the CT values (or the linear attenuation coefficient values) along the individual rays. The values of the line integrals are assigned to the corresponding pixels to form a DRR.

The next step in the teletherapy procedure is the actual radiation delivery which is administered in small daily doses (usually fractions of 2 Gy of the total prescription dose) generally over a period of three to six weeks. During the daily treatments, a radiographic film or an electronic portal imaging device (EPID) (Fig. 1.2) are used to acquire megavoltage *portal images* with the radiation beam for *treatment verification*. The size and the shape of the radiation field as well as the proper positioning of the patient are checked by comparison of the portal images to the diagnostic portal images obtained during simulation. After the conclusion of the radiation treatment, patient's follow-up is conducted.

1.2 Conformal Therapy

1.2.1 Rationale and Potential Impact

Radiotherapy has curative potential‡ given that the primary tumor is confined to its local or local-regional site (Table 1.1). According to the data of the National Cancer Institute Surveillance, Epidemiology and End Results (SEER) program[21, 22] and the National Cancer Data Base,[23] approximately 65-72% of the cancer patients in the United States are initially diagnosed with locally confined disease without clinical evidence of distant cancer spread (metastasis). Even though two-thirds of these patients are cured after a treatment of the localized disease by surgery and/or radiation therapy,[24]

‡ Radiation can also be used with a palliative intent.

Table 1.1 *Cancers commonly treated by definitive (curative) radiation.*[8]

Hodgkin's disease
Non-Hodgkin's lymphomas
Cervical cancer
Prostate cancer
Head and neck cancers
Cancers of the central nervous system
Seminoma
Retinoblastoma
Choroidal melanoma
Unresectable lung cancer
Unresectable pancreatic cancer
Unresectable sarcoma

a significant portion still succumbs to the disease. Furthermore, 30–50% of the patients treated with curative radiation therapy fail at their primary tumor site.[25, 26] Even though these failures relate not only to treatment parameters but also to biological factors, they indicate the inability of radiotherapy to provide tumor control in certain cases due to either (i) inadequate dose delivery (insufficient dose) or (ii) geometric misses of the target.

An important question is, however, whether improvements in radiation therapy that increase local tumor control probability by addressing the above two problems will lead to increased patient survival. There has been some concern that new local control patients will ultimately fail either due to the presence of undetected micrometastases at the time of the diagnosis or due to the development of distant metastases subsequent to the primary tumor treatment.[25, 27, 28] However, several lines of evidence confirm that improved local tumor control does result in increased disease free survival.[25, 29, 30] For

example, data from several institutions indicate that, depending on both the treatment site and the tumor stage, 18% to 88% of the patients that undergo salvage surgery[§] survive for 5 years and more free of the disease. Furthermore direct correlation seems to exist between the overall incidence of metastatic disease and the local or the local-regional tumor control.[24] For instance, after a review of the outcome of 679 surgically staged (without detectable metastasis) patients with carcinoma of the prostate treated with permanent implantations of encapsulated ^{125}I sources, Fuks et *al*[31] have demonstrated that the relative risk of distant metastases subsequent to local relapse was four times greater than the risk without evidence of local failure. Other retrospective studies have reported increased metastatic occurrences after local failure in carcinoma of the breast,[32] lung,[33] rectum,[34, 35] prostate,[36, 37] uterine cervix,[38] in head and neck tumors[39] as well as in soft tissue sarcomas.[40] Therefore there are sufficient indicators to warrant the development and the evaluation of approaches that aim at improved local tumor control by assuring an adequate target coverage and/or escalating the dose to the target volume without increasing the risk to surrounding healthy tissues.[28]

Three-dimensional computer controlled conformal therapy (CCRT) is such an approach. The term conformal therapy is associated with external beam therapy to denote treatment designs that tailor a high-dose region to the target volume and simultaneously deliver low doses beyond its extent.[41–43] (Conformal therapy has been attempted since the inception of radiotherapy, but the degree of conformity has improved over the years.) By sparing more of the critical organs conformal treatments may also allow higher doses to be delivered to the tumor (dose escalation) to increase the probability of local control.[24, 28, 44] This increase has been estimated at 2.5% (median) per 1 Gy in the

§ Salvage surgery is attempted for selected patients after radiotherapy failure.

neighborhood of the 50% range (median 52 Gy) of the dose-response curve for patients with macroscopic tumors.[45]

The potential impact of the improved local tumor control on the long-term patient survival has been roughly estimated for several treatment sites. Using detailed analyses of the causes of failure, Suit[25] estimated that complete local tumor control (100%) would result in a 20% increase of the 5-year survival rate for patients with carcinoma of the uterine cervix, a 14% increase for patients with carcinoma of the oral cavity and the oro-pharynx, a 21% increase for patients with carcinoma of the colorectum and a 14% increase for patients with carcinoma of the ovaries. Given the annual rates of cancer incidence for the above sites, these figures would translate in approximately 2500 additional survivals in Canada if complete local control were achieved for these tumors.[25] (A factor of 10 has been used in the above estimate to account for the size of the population of Canada compared to that of USA.)

Yorke *et al.*[46] studied the impact of improved local control for the case of prostate cancer. They applied a model[47] for the metastasis development of prostatic carcinoma to the clinical results of a long-term follow-up study[31] in order to obtain reasonable biological parameters for the model. Under the assumption of 100% local control achieved by conformal therapy, the model predicts a 10% increase of the 5-year distant-metastasis free survival and a 35% increase of the 10-year distant-metastasis free survival for patients with prostatic carcinoma.[48] Similar trends could be expected for other treatment sites where metastasis dissemination occurs as a relatively late effect.[48] The actual gain in survival due to conformal therapy depends on the achievable local tumor control, an unknown which is currently under investigation by several clinical trials with conformal therapy for prostate cancer.[49, 50]

1.2.2 Development of Conformal Therapy

Many of the key concepts of conformal therapy were first introduced by Takahashi[41] in 1965. Takashi and co-workers used 3-D models of the tumor (based on planar tomography) to plan the treatment, orthogonal light beams to align the target with the machine isocenter, as well as "geared sectional collimators" and a mechanical control system to conform the beam shape to the shape of the target as the machine rotated about the patient. Similar early conformal techniques were developed in other centers.[51–54]

The advent of modern conformal therapy was marked by the introduction of CT and CT-assisted 3-D treatment planning in radiotherapy. Before CT scans became available, tumor volumes were ill-defined and marginal misses, especially in large tumors were frequently encountered in radiotherapy.[55–57] To avoid local relapse due to these misses treatment fields were designed to encompass the identifiable tumor plus large margins (2 cm or more) of normal tissues surrounding the target. The relatively large volume of healthy tissues irradiated in the treatment often limited dose escalation because of the restrictions imposed by normal tissue tolerance. Thus treatment designs were often restrained by the normal tissue tolerance rather than dose levels required to control the tumor. Computed tomography improved significantly the coverage of the target during radiation treatments by allowing better definition of the target volume and its relation to the surrounding healthy structures in the process of treatment planning.[55–57]

Further improvements in the target coverage were obtained with the use of high-power workstations, which allowed computer graphics,[20, 58–60] 3-D dose calculation algorithms,[61] as well as tools for the evaluation of 3-D dose distributions[62, 63] to be integrated in the treatment planning process.[64]

Concurrently with the development of 3-D conformal treatment planning new generation treatment units were designed to execute conformal plans. The first machine of this

generation, the Scanditronix MM50 racetrack microtron[65–67] is a typical example of the treatment capabilities provided by these units: megavoltage photon and electron beams of various energies, a fully computerized control system, and a multileaf collimator[68–70] (an assembly of up to 40 thin tungsten leaves on each side of the collimator which are moved in and out of the beam under computer control). It is the recent widespread availability of multileaf collimator systems (MLC) that has significantly fostered the interest in conformal therapy mostly due to the fact that an MLC opens up the possibility of improved dose conformation[65, 71–76] by computer-controlled modulation of the intensity of each beam.[65, 77, 74, 78–82] Treatments with intensity modulated beams have been initiated clinically for intracranial[76] lesions and prostate cancers.[83] Such treatments have been shown to be executable in reasonable time limits when multileaf collimation is used – 10 to 12 min for prostate cases (a dose fraction of 140 cGy delivered by 6 beams).[83]

The on-line electronic portal imaging device[84] (EPID) is another important component of the equipment required for conformal therapy.[48] The role of the EPID is not restricted only to the crucial task of providing near real time information on patient's setup.[85–89] EPID are currently becoming an important tool for evaluating target motion and set-up uncertainties[90–92] which are to be incorporated in the treatment planning process.[93, 94] Portal dose images obtained by an EPID can be compared to calculated ones.[95, 96] In addition, given patient-specific information, on-line images can further be used for reconstruction of the actual dose distributions created during the treatment.[97, 98] Thus, an EPID creates the opportunity of complete geometric and dosimetric verification of conformal radiation treatments.[96–100]

1.3 Treatment Planning Optimization

1.3.1 Rationale

Given the new and improved technical abilities that modern computer-controlled treatment units provide for radiation delivery, the treatment planning process is crucial in assuring that these features are used properly for achieving the most beneficial effect.[48, 53, 101] Ling *et al.*[48] observe that currently, despite the advent of high-speed graphics workstations, very little is done beyond the "planning by convention" whereby the same arrangement of beams is applied to patients for a given disease site followed by some manual optimization of the beam positions and weights. This process does not guarantee optimal utilization of modern radiotherapy equipment since it relies on "past clinical experience" which is non-existent for dose escalation studies or for the novel radiation distribution patterns that can be achieved with the currently available technical capabilities, e.g. intensity modulation.[48] On the other hand novel treatment designs for dose escalations that optimize few treatment parameters may require as much as 1.5 physicist-months of much trial and error.[48] Whereas such an effort can be well justified for finding generic treatment parameters for some disease sites which exhibit high inter-patient similarity (e. g. prostate and nasopharynx), it is inapplicable for cases where the tumor extent and position are highly variable (e. g. brain and lung cancers) or for cases where a large number of optimization variables is involved, e. g. intensity modulation. Thus computer-aided optimization is essential for the progress of conformal therapy.

1.3.2 Approaches

Computer-assisted treatment planning optimization can be described by the generic optimization algorithm shown in Fig. 1.3. Two distinct approaches exist that differ in their evaluation of the dose distributions. The *biological approach* which is the more

relevant one from a conceptual point of view, takes into account the architecture of the irradiated normal tissues and the corresponding dose-volume effects. (For example, very little damage to any part to the spinal cord may have lethal effect, whereas the lung remains functional even if a significant part of it is destroyed.) Models of cell and organ response[47, 102–110] are used to predict biological indexes such as NTCPs and TCP from the dose distribution. Cost functions based on these indexes or their combinations[111–113] are then used to evaluate the probability of success of the radiation treatment and to serve as objectives for the biologically based optimizations.[112, 114–119] Minimization algorithms such as constrained steepest-descent[112, 116–119] and simulated annealing[114, 115] were used with biological objectives for the selection of optimal beam[‖]-on times (weights),[112, 116–119] energies,[119] and directions.[120]

The validity, however, of the existing TCP and NTCP models is not yet well established[118, 120–122] and their clinical predictive power is unproven.[118, 123] Furthermore, significant uncertainties exist in the values of the radiobiological parameters used by these models due to the paucity of clinical data, especially at the higher dose levels.[124] For these reasons, for treatment designs, TCP and NTCPs, if considered at all, are only secondary adjuncts to dose-based and intuitive criteria such as (a) target dose homogeneity, (b) maximum tolerance doses of critical structure, and (c) dose-volume considerations.[30, 44, 125–127] (A dose-volume consideration usually specifies the percentage volume of the critical organ that can be sacrificed above certain dose level.)

The *physical approach* to treatment planning optimization takes the above criteria into account by employing dose-based objective functions such as the integral (total) dose to the target,[128, 129] the integral dose to healthy tissue,[128, 51, 130–134] the dose to the center of the target,[129, 135–137] the difference between the integral dose to the target and

‖ Including pencil beams, wedged beams, open beams, etc.

the healthy tissues,[138] target dose uniformity[132, 133, 139] as well as the sum of the squares of the residuals between the prescribed and the delivered dose.[73, 140–148] Minimization algorithms such as linear,[51, 128, 130–134, 136], mixed-integer[129] and quadratic programming techniques,[132, 142] gradient techniques,[116, 144, 149] simulated annealing,[136, 141, 143, 144, 150] iterative deconvolution and filtered backprojection,[146, 151, 152] iterative reconstruction techniques, [73, 147, 148] as well as genetic algorithms[137] were applied to physical objective functions for the selection of optimal beam weights. A combination of simulated annealing and an iterative reconstruction technique was recently used for the optimal selection of the number and the orientations of beams in intensity modulated treatments.[150]

1.3.3 Inverse Treatment Planning

Ideally, regardless of the particular objective in use, an optimization algorithm should automatically determine all the relevant parameters of the treatment (Fig. 1.3). In practice, however, the size of the solution space is prohibitive for such an attempt. For this reason the various optimization algorithms explore only certain subsets of the treatment parameters.

The inverse problem[43, 153] in conformal radiotherapy generalizes the most commonly solved problem of beam weight optimization for a set of fixed beams. The inverse problem considers the individual beam ports as being tiled in pencil beams (beamlets) with their own weights. Different sets of pencil beam weights correspond to different modulation of the beams. The dose to any point in the patient volume is then given as the superposition of the dose contributions to that point of all beamlets. A solution of the inverse problem is the modulation of the fluence (intensity) across the individual beam ports which results in an optimal plan, according to the selected physical or biological objective. Numerical or analytical procedures for solving the inverse problem are referred to as *inverse treatment planning* techniques.

Compared to uniform beams, modulated beams have the potential of creating improved conformal distributions, especially for targets with concave regions.[65, 72, 71, 73–75] The design of the beam profiles, however, is a process of considerable complexity, clearly beyond the ability of a human planner. The modulation of the beam intensities is a large-scale problem which involves a few hundred[116] to a few thousand variables (pencil beam weights).[74, 122] Furthermore, the objective function evaluation is very-time consuming due to the combined effect of the following two factors.

First, the novel dose distributions produced by intensity modulated beams generally possess "unusual" properties such as high-gradient regions and small but very low- or very high-dose domains. Since these features can be clinically significant, fine sampling of the dose distribution is important[154–156] and therefore, the number of the required calculation points per structure may have to be at least an order of magnitude larger than the number of points used in treatment planning optimization with uniform or wedged beams. (For the latter case, depending on the size of the organ, 100 to 800 randomly distributed points per structure have been found sufficient.[156, 157])

Second, Chen *et .al*[158] have demonstrated that the calculation of the dose distribution for treatment planning optimization by intensity modulation is to include lateral transport in order to avoid target underdosage and to reduce unwanted injuries to critical organs when critical structures are immediately adjacent to the target. Convolution/superposition algorithms adequately take into account lateral transport of radiation.[159–161] However, photon-dose calculations by such an algorithm for a $64\times64\times128$ grid on a DEC Alpha (3000/400) station (Digital Equipment Corp., Marlborough, MA) have been reported to exceed 5 min turnaround time for typical clinical beam arrangements.[162] Thus for thousands of sampling points, a single iteration of an inverse treatment planning procedure may require a few seconds to a few minutes on the currently available hardware in the

radiotherapy departments.

Given the large-scale nature of the inverse problem and its computational properties, the development of versatile and efficient optimization algorithms for inverse treatment planning is imperative for the advancement of intensity modulated conformal therapy. The results of our research in this direction are presented in Chapter 2.[163–165]

There is another facet related to the computational nature of inverse treatment planning. Given the current status of the radiobiological models, the clinical advancement of intensity modulated conformal therapy is based on the physical approach to inverse treatment planning which provides continuity with respect to the existing standard practice.[83] However, to make the inverse problem manageable, the clinical goal of conforming the high dose region to the target while sparing healthy tissues is cast in an optimization problem by relatively simple models. For this reason the resulting optimized plans may not be clinically acceptable. Our research effort to improve on the existing models with physical objective functions is presented in Chapter 3.[166–169]

1.4 Treatment Setup Verification

1.4.1 Rationale

Conformal radiotherapy (with or without intensity modulation) reduces the treatment margins around the CTV in order to achieve a higher TCP while maintaining low NTCPs. Therefore, high geometric accuracy in the placement of the target volume with respect to the treatment beams (ports) is crucial for the success of conformal treatments.[170, 171] The attainment of precise target coverage implies monitoring, detection and correction of field placement errors which result in geographic misses of some fraction of the target volume. Field placement errors may be caused by setup errors (improper blocks or wedges placed

in the beam port), by inaccurate patient positioning or by movement of internal organs with respect to the external skin marks used to align the patient.

Being difficult to monitor, field placement errors are the largest source of uncertainty in dose delivery[172] compared to the uncertainties resulting from variations in dosimetry, treatment unit output and gantry and treatment coach stability. Moreover, by employing radiobiological models, several studies have demonstrated that setup errors can have significant impact on both the TCP and the NTCP.[173–176] Goitein and Busse[173] have demonstrated that tumors with steep dose-effect curves (Fig. 1.1) such as supraglottic lesions are very sensitive to field placement errors. For positioning errors of ±5 mm, they have estimated TCP reduction between 12% and 40% depending on the size of the selected margins and the frequency of the errors which were assumed to be random. Assuming a constant dose distribution along the beam axis,# Brahme[174] calculated the TCP reduction as a function of the misalignment of the radiation field edge and the target border. Depending on the steepness of the dose-response curve (Fig. 1.1) some decrease of 3-7% in the TCP was evaluated for a 2 mm shift and of 9-50% for a 5 mm shift. Using a radiobiological model, for the cases of lesions encircling the brain stem or the spinal cord, Daftari *et al.*[176] evaluated that patient positioning had to be kept within 2 mm to ensure a NTCP value of 1%. By retroactively incorporating treatment positioning errors into a 3-D treatment planning system, Rudat *et al.*[171] have evaluated a resulting reduction in TCP of 2% for esophagus carcinomas and 5% for prostate carcinomas.

Qualitatively, the conclusions of these theoretical models are supported by some clinical studies. Kinzie *et al.*[177] reported results on the relapse rates for patients with Hodgkin's disease in relation to the adequacy of the field placement as inferred from portal images. Treatment setups which had consistently resulted in partial shielding

Close approximation to that assumption is the dose distribution produced by a pair of parallel opposed beams.

of the target volume or of the supraclavicular, auxiliary and/or abdominal nodes were considered inadequate. The relapse rate for patients who had been treated with improperly placed fields was 54% with 33% of these recurrences appearing in the irradiated volume. For patients without field placement errors, the relapse rate was 14% with only 7% of recurrences in the primary tumor site. Another study by White *et al.*[178] statistically significant difference in the survival rates for patients with and without major treatment protocol variations, 80% of which were due in part or completely to shielding errors as revealed by portal film analysis. The relapse rate was 69% in the group with major protocol variations and 34% in the other group.

It is evident from the existing theoretical studies that accurate beam localization with respect to the patient anatomy (in the order of 2 mm) is essential for achieving the necessary local tumor control, particularly in the case of conformal therapy with its tight margins. However, in clinical practice, for 20% of the treatment setups, the discrepancies between the intended and the actual field edge positions are in order of 10 mm.[179] An approach that can significantly improve the accuracy of the field placements is described below.

1.4.2 Portal Imaging

The goal of portal imaging is to determine whether a patient setup is performed correctly within the prescribed limits and to suggest the necessary corrective measures if the field placement errors are inadmissibly large. For this purpose, the treatment beam (port) is used to acquire a portal image of the patient in the treatment position. An EPID or a film are employed to record the image which is usually double-exposed to provide a better visualization of the spatial relationship between the patient anatomy and the treatment field (Fig. 1.4). (In a double exposure image the treatment field appears as a dark shadow cast on the previously obtained image of the patient anatomy with a

large rectangular field.) The portal image is compared to a reference one that reflects the desired placement of the treatment field with respect to the patient anatomy (Fig. 1.4). A diagnostic-quality image obtained during the conventional simulation process, a DRR produced by virtual simulation or an approved portal image from the first treatment session can serve as reference images. After visual or computer-assisted registration of the anatomy images (Fig. 1.4), a misalignment of the borders of the actual and of the desired treatment field manifests an inaccurate patient setup.

Conventionally, portal images are obtained by films.[177, 179–187] Portal films include verification films, which are slow and exposed throughout the treatment and localization films which are exposed only with a small fraction of the daily dose.[188] A definitive correlation exists between the rate of field placement errors and the frequency of portal image acquisitions.[183, 185, 189] For example, for extended mantle fields, increasing the frequency of portal film acquisitions from 3 to 18 per treatment decreases the frequency of field placement errors from 55% to 29%.[183]

Despite the efficiency of portal radiography in reducing the rate of treatment setup errors, there is a significant numbers of hospitals that have not adopted portal imaging.[187, 188] In 1989, a survey among 25 centers in Canada indicated that only 55% of the institutes acquired routinely portal images of all the patients treated with curative intent.[187] Overall, only 67% of the patients receiving radical treatments have field placement accuracy checked by portal imaging. In U.S.A., an earlier study reported that 90% of the clinics acquired portal films on the first day of the treatment for more than 75% of their patients. However, similarly to the practice in Canada, only 40% employed routine radiographs on a weekly basis (4 to 6 portal films per treatment course).[188]

Somewhat surprisingly, the above described patterns of clinical portal imaging practice are not caused solely by the off-line nature of portal film radiography. Indeed, even though EPIDs[84] can provide near real-time digital images for on-line viewing and

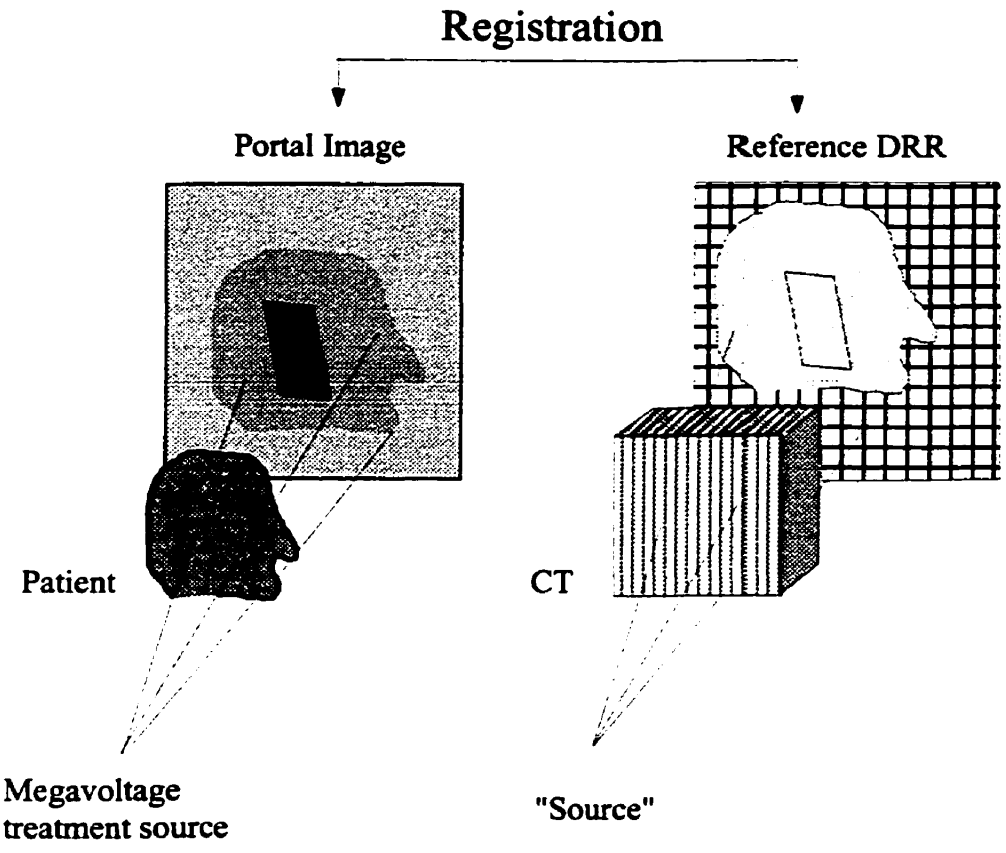

Figure 1.4. *Schematic representation of treatment setup verification by portal imaging. On the left: double exposed portal image of the treatment field and the surrounding patient anatomy. On the right: a reference image (in this particular example a DRR) of the patient anatomy and the desired placement of the treatment field. After anatomy registration, in the absence of field placement errors the borders of the actual and of the desired treatment field will be aligned.*

quantitative analysis, they have not been adopted for clinical radiotherapy as widely as initially anticipated.[190] A survey at The Third International Workshop on Electronic Portal Imaging (San Francisco, July 1994) demonstrated that among the 34 hospitals that

had installed EPIDs, 38% used them daily, 26% used them weekly, 21% did not use them often and 15% did not use them clinically.[191]

The difficulties associated with the detection and evaluation of field placement errors are regarded as the main obstacle to the wider and more frequent clinical use of portal imaging.[186, 187, 192–195] Accurate on-line manual registration of images and subsequent measurements are considered impractical since they require a few minutes to be performed.[192] For this reason, the majority of the portal images are evaluated on subjective basis. Dunscombe *et al.*[187] found that 63% of the hospitals that participated in their study relied on a visual judgement of the portal images, 25% of the hospitals used both subjective and quantitative measure, but relied predominantly on the subjective approach, 8% of the hospitals evaluated the images half quantitatively and half subjectively and only 4% of the hospitals used mostly "semi-quantitative" methods in their analysis of portal images.

The visual subjective analysis of the images, however, is not only somewhat time-consuming but prone to significant errors as well, due to the poor contrast of portal images which results from the predominance of the Compton scattering in the imaging process and from the degrading effect of scattered radiation.** For example, Herman *et al.*[194] reported a study whereby patient setup errors were estimated by visual comparison of EPID images and reference ones obtained during simulation. The patient setup was corrected if errors larger than 5 mm were detected. A further analysis of the images before and after the on-line correction revealed that more than 17% of the accepted final setups still exhibited errors exceeding the 5 mm criterion.

The above considerations indicate that the development of fast, computer-assisted methods for the detection and the evaluation of fields placement errors can contribute

** These are general properties of the image formation at megavoltage energies, independent of the actual portal image detector.

significantly to the effective use of electronic (or film) portal imaging in clinical practice.[186, 187, 192–196] In the second part of this thesis (Chapter 4), we present our contribution to the research effort[16, 197–220] in the field of image processing for treatment setup verification.

1.5 Objectives and Organization of Thesis

This thesis presents our work in two distinct areas essential for the efficient clinical use of the modern technologies for conformal radiation treatments: (i) the development of techniques for the design of intensity modulated radiation beams and (ii) the development of techniques for field placement detection and evaluation. Our objectives were:

I. to develop a versatile optimization algorithm that, under the limitations of the currently existing hardware, can be applied to various objective functions at the large-scale of the inverse problem.

II. to develop an optimization technique that can improve on the existing inverse treatment planning algorithms by exploring beam intensity modulation while providing continuity with the existing clinical practice.

III. to develop an automated anatomy image registration algorithm in order to facilitate the process of detection and evaluation of field placement errors.

In Chapter 2, the existing methods for inverse treatment planning are reviewed in order to identify somewhat optimal algorithms for treatment planning optimization with intensity modulated beams. An active set algorithm is introduced as an alternative to the widely used constrained steepest-descent method. The two methods are compared for two treatment objectives and three anatomical sites.

In Chapter 3, after an analysis of the limitations of the conventional inverse treatment planning technique, an alternative statement of the inverse problem is suggested. A

continuous penalty function method is introduced to find approximate solutions of the resulting large-scale constrained optimization problem. The performance of the technique as a function of several parameters is investigated. Examples are presented to illustrate some strategies for the clinical applications of the technique.

In the second part of this thesis, Chapter 4 describes a registration algorithm based on image correlation. Two correlation operators can be employed: (i) the normalized cross-correlation and (ii) the Pearson linear-correlation coefficient can be used. Their utility is investigated for the cases of portal-to-portal and portal-to-DRR registration.

Finally, after summarizing the features of the techniques presented in this thesis, Chapter 5 discusses some of the areas to be explored by future research.

To streamline the presentation, mathematical calculations are given in the Appendices. In Appendix A the Differential Scatter-Air Ratio dose calculation model is outlined. In Appendix B, we prove the applicability of the continuous penalty function method to the inverse problem formulated in Chapter 3. In Appendix C, we derive the Fast Fourier Transform implementation of the Pearson linear correlation coefficient.

A list of references, sorted by the order of appearance, follows each chapter. A complete bibliography is included at the end of the thesis.

Several aspects of this work have been presented at national and international meetings,[221–226] and have been published as abstracts[163, 164, 166, 168, 227] and articles in conference proceedings,[167, 169, 228, 229] and in peer-reviewed journals.[165, 230, 231]

References

1 Statistics Canada. Mortality – Summary List of Causes, 1994. Toronto, 1994. Catalogue no. 84-209-XPB.

2 National Cancer Institute of Canada. Canadian cancer statistics 1997. Toronto, 1997.

3 G. H. Hahn. *Hyperthermia and Cancer*. Plenum, New York, 1982.

4 L. J. Old. Immunotherapy for cancer. *Sci. Am.*, 275(3):136–143, 1996.

5 A. Oliff, J. B. Gibbs, and F. McCormick. New molecular targets for cancer therapy. *Sci. Am.*, 275(3):144–149, 1996.

6 J. K. Folkman. Fighting cancer by attacking its blood suply. *Sci. Am.*, 275(3):150–154, 1996.

7 A. H. Knowlton. General principles of radiation therapy. In D. S. Fishter and J. C. Marsh, editors, *Cancer Therapy*, pages 169–175. G. K. Hall Medical Publishers, Boston, 1982.

8 Allen S. Lichter and Theodore S. Lawrence. Recent advances in radiation oncology. *The New England Journal of Medicine*, 332(6):370–372, 1995.

9 M. M. Elkind. DNA damage and cell killing: cause and effect? *Cancer*, 56:2351–2363, 1985.

10 ICRU Report No. 33: Radiation Quantities and Units. Technical report, Interanational Commision on Radition Units and Measurements, Washington, DC, 1980.

11 T. R. Munro and C. W. Gilbert. The relation between tumor lethal dose and the radiosensitivity of tumor cells. *Br. J. Radiol.*, 34:246–251, 1961.

12 E. H. Porter. The statistics of dose/cure relationships for irradiated tumors. *Br. J. Radiol.*, 53:210–227, 336–345, 1980.

13 H. E. Johns and J. R. Cunningham. *The physics of radiology*, pages 682–683. Charles C. Thomas, Springfield, Illinois, 1983.

14 A. Brahme and A. K. Ågren. Optimal dose distributions for eradication of heterogeneous tumors. *Acta Oncol.*, 26:377–385, 1987.

15 P. Rubin and W. D. Siemann. Principles of radiation oncology and cancer radiotherapy. In P. Rubin, editor, *Clinical Oncology*. W.B. Saunders Company, The Curtis Center, Independence Square West, Philadelphia, PA 19106, 7th edition, 1993.

16 J. Bijhold, K. G. A. Gilhuijs, M. van Herk, and H. Meertens. Radiation field edge detection in portal images. *Phys. Med. Biol.*, 36:1705–1710, 1991.

[17] ICRU. *ICRU Report 50: Prescribing, Recording, and Reporting Photon Beam Therapy.* Bethesda, Maryland: International Commission on Radiation Units and Measurements, 1993.

[18] J. G. Galvin, A. T. Turrisi, and Cheng. Treatment simulation using a CT unit. In *Radiation Oncology Physics. AAPM Monograph No. 15.* American Institute of Physics, New York, 1986.

[19] T. Nishidai, Y. Nagata, M. Takahashi, M. Abe, N. Yamaoka, H. Ishihara, Y. Kubo, H. Ohta, and C. Kazusa. CT simulator: a new 3-D planning and simulating system for radiotherapy. *Int. J. Rad. Oncol. Biol. Phys.*, 18:499–504, 505–513, 1990.

[20] G. W. Sherouse and E. L. Chaney. The portable virtual simulator. *Int. J. Rad. Oncol. Biol. Phys.*, 21:475–482, 1991.

[21] M. H. Myers and B. F. Hankey. Cancer patient survival experience in the United States. Publication No. (NIH) 80–2148, DHHS, U. S. Goverment Printing Office, Washington, D.C., 1980.

[22] Cancer Trends. In *1987 annual cancer statistics review*, pages II.1–II.203. US Department of Health and Human Services, 1988.

[23] H. R. Menick, L. Garfinkel, and G. D. Dodd. Preliminary report of the National Cancer Data Base. *CA-A Cancer Journal for Clinicians*, 41:7–18, 1991.

[24] S. A. Leibel, C. C. Ling, G. J. Kutcher, R. Mohan, C. Cordon-Cordo, and Z. Fuks. The biological basis for conformal three-dimensional radiation therapy. *Int. J. Rad. Oncol. Biol. Phys.*, 21:805–811, 1991.

[25] H. D. Suit. Potential for improving survival rates for the cancer patient by increasing the efficacy of the treatment of the primary tumor. *Cancer*, 50:1227–1234, 1982.

[26] V.T. DeVita. Progress in cancer management. *Cancer*, 51:2401–2409, 1983.

[27] V. T. DeVita, M. Lippman, S. A. Hubbard, D. C. Idhe, and S. A. Rosenberg. The effect of combined modality therapy on local control and survival. *Int. J. Rad. Oncol. Biol. Phys.*, 12:487–501, 1986.

[28] A. S. Lichter. Three-dimensional conformal radiation therapy: a testable hypothesis. *Int. J. Rad. Oncol. Biol. Phys.*, 21:853–855, 1991.

[29] S. Dische. Radiotherapy in the nineties. *Acta Oncol.*, 31(5):501–511, 1992.

[30] H. D. Suit. Local control and patient survival. *Int. J. Radiat. Oncol. Biol. Phys.*, 23(3):653–660, 1992.

[31] Z. Fuks, S. A. Leibel, K. E. Wallner, C. B. Begg, W. R. Fair, L. L. Anderson, B. S. Hilaris, and W. F. Whitmore. The effect of local control on metastatic dissemination in carcinoma of the prostate: Long-term results in patients treated with ^{125}I implantation. *Int. J. Rad. Oncol. Biol. Phys.*, 21(3):537–47, 1991.

[32] J. Hayward and M. Galeffi. The significance of local control in the primary treatment of breast cancer. *Ann. Surg*, 122:1244–1247, 1987.

[33] C. K. Chung, J. A. Styker, M. O'Neil, and W. E. DeMuth. Evaluation of adjuvant postoperative radiotherapy for lung cancer. *Int. J. Rad. Oncol. Biol. Phys.*, 8:1877–1880, 1982.

[34] A. Vigliotti, T. A. Rich, M. M. Romsdahl, H. R. Withers, and M. J. Oswald. Postoperative adjuvant radiotherapy for adenocarcinoma of the rectum and rectosigmoid. *Int. J. Rad. Oncol. Biol. Phys*, 13(7):999–1006, 1987.

[35] S. E. Schild, J. A. Jr. Martenson, L. L. Gunderson, D. M. Ilstrup, K. K. Berg, M. J. O'Connell, and L. H. Weiland. Postoperative adjuvant therapy of rectal cancer: An analysis of disease control, survival, and prognostic factors. *Int. J. Rad. Oncol. Biol. Phys.*, 17(1):55–62, 1989.

[36] C.A. Perez, M. V. Pilepich, and F. Zivnuska. Tumor control in definitive irradiation of localized carcinoma of the prostate. *Int. J. Rad. Oncol. Biol. Phys*, 12:523–531, 1986.

[37] D. A. Kuban, A. M. El-Mahdi, and P. F. Schellhammer. Effect of local tumor control on distant metastasis and survival in prostatic adenocarcinoma. *Urology*, 30(5):420–426, 1987.

[38] C. A. Perez, R. R. Kuske, H. M. Camel, A. E. Galakatos, M. A. Hederman, M. S. Kao, and B. J. Walz. Analysis of pelvic tumor control and impact on survival in carcinoma of the uterine cervix treated with radiation therapy alone. *Int. J. Rad. Oncol. Biol. Phys*, 14(4), 1988.

[39] S. A. Leibel, C. B. Scott, M. Mohiuddin, V. A. Marcial, L. R. Coia, L. W. Davis, and Z. Fuks. The effect of local-regional control on distant metastatic dissemination in carcinoma of the head and neck: Results of an analysis from the RTOG head and neck database. *Int. J. Radiat. Oncol. Biol. Phys*, 21(3):549–556, 1991.

[40] H. D. Suit, H. J. Mankin, W. C. Wood, M. C. Gebhardt, D. C. Harmon, A. Rosenberg, J. E. Tepper, and D. Rosenthal. Treatment of the patient with stage M0 soft tissue sarcoma. *Journal of Clinical Oncology*, 6(5):854–62, 1988.

[41] S. Takahashi. Conformation radiotherapy: rotation techniques as applied to radiography and radiotherapy of cancer. *Acta Radiol. Suppl.*, (242):1–141, 1965.

[42] G. J. Kutcher, R. Mohan, S. A. Leibel, Z. Fuks, and C. C. Ling. Computer-controlled 3D conformal therapy. In *Radiation Therapy Physics*. Springer-Verlag, Berlin, Germany, 1992.

[43] S. Webb. *The physics of three-dimensional radiation therapy*. Institute of Physics Publishing, Bristol, 1993.

[44] S. A. Leibel, G. J. Kutcher, R. Mohan, L. B. Harrison, J. G. Armstrong, M. J. Zelefsky, T. J. LoSasso, C. M. Burman, G. S. Mageras, C. S. Chui, L. J. Brewster, M. E. Masterson, Y. C. Lo, C. C. Ling, and Z. Fuks. Three-dimensional conformal radiation therapy at the Memorial Sloan-Kettering Cancer Centre. *Semin. Radiat. Oncol.*, 2:274–289, 1992.

[45] P. Okunieff, D. Morgan, A. Niemierko, and H. Suit. Radiation dose-response of human tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1227–1237, 1995.

[46] E. D. Yorke, Z. Fuks, L. Norton, W. Whitmore, and C. C. Ling. Modelling the development of metastases from primary and locally recurrent tumors: comparison with a clinical data base of prostate cancer. *Cancer Res.*, 53:2987–2993, 1993.

[47] E. D. Yorke, G. J. Kutcher, A. C. Jackson, and C. C. Ling. Probability of radiation-induced complications in normal tisues with parallel architecture under conditions of uniform whole or partial organ irradiation. *Radiother. Oncol.*, 26:226–237, 1993.

[48] C. C. Ling, C. Burman, C. S. Chui, A. Jackson, G. J. Kutcher, S. A. Leibel, T. LoSasso, G. Mageras, R. Mohan, E. Yorke, and Z. Fuks. Perspectives of multidimensional conformal radiation treatment. *Radiotherapy and Oncology*, 29:129–139, 1993.

[49] S. A. Leibel, R. Heimann, G. J. Kutcher, M. Zelefsky, C. M. Burman, E. Mellian, Orazem. J., S. Kelson, R. Mohan, T. LoSasso, Y.-C. Lo, J. Wiseberg, D. S. Chapman, C. C. Ling, and Z. Fuks. Three-dimensional conformal radiation therapy in localized carcinoma of the prostate: preliminary results of a phase I dose-escalation study. *Int. J. Radiat. Oncol. Biol. Phys.*, 28(1):55–65, 1994.

[50] A. Pollack, G. Zagars, G. Starkshall, C. H. Childress, S. Kopplin, A. L. Boyer, and I. I. Rosen. Conventional vs. conformal radiotherapy for prostate cancer: preliminary results of dosimetry and acute toxicity. *Int. J. Radiat. Oncol. Biol. Phys.*, 34(3):555–564, 1994.

[51] J. Mantel, H. Perry, and J. J. Weinkam. Automatic variation of field size and dose rate in rotation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 2:697–704, 1974.

[52] M. B. Levene, P. K. Kijevski, L. M. Chin, S. M Bengt, B. E. Bjarngard, and S. Hellman. Computer-controlled radiation therapy. *Radiology*, 129:769–775, 1978.

[53] J. A. Brace, T. J. Davy, D. B. L. Skeggs, and H. S. Williams. Conformation therapy at the royal free hospital: a progress report on the tracking cobalt project. *Br. J. Radiol.*, 54:1068–1074, 1981.

[54] B. S. Proimos. Synchronous field shaping in rotational megavolt therapy. *Radiology*, 74:753–757, 1960.

[55] M. Goitein. The utility of computed tomography in radiation therapy: an estimate of outcome. *Int. J. Radiat. Oncol. Biol. Phys.*, 5:1799–1807, 1979.

[56] J. S. Donald. Computed tomography applied to radiotherapy treatment palnning: techniques and results. *Radiology*, 133:477–482, 1979.

[57] J. E. Munzenrider, M. Pilepich, J. B. Rene-Ferero, I. Tchakarova, and B. L. Carter. Use of body scanner in radiotherapy treatment planning. *Cancer*, 40:170–179, 1979.

[58] M. Goitein and M. Abrahms. Multi-dimensional treatment planning: I. Delineation of anatomy. *Int. J. Radiat. Oncol. Biol. Phys.*, 8:777–787, 1983.

[59] M. Goitein, M. Abrahms, D. Rowell, H. Pollari, and J. Wiles. Multi-dimensional treatment planning: II. Beam's eye-view back projection, and projection through CT sections. *Int. J. Radiat. Oncol. Biol. Phys.*, 9:780–797, 1983.

[60] D. L. Mc Shan, B. A. Fraass, and A. S. Lichter. Full integration of the Beam's Eye View concept into clinical treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:1485–1494, 1990.

[61] Photon Treatment Planning Collaborative Working Group. Three-dimensional dose calculations for radiation treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:25–36, 1991.

[62] R. E. Drzymala, R. Mohan, L. Brewster, J. Chu, M. Goitein, W. Harms, and M. Urie. Dose volume histograms. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:71–78, 1991.

[63] M. L. Kessler, R. K. TelHaken, B. A. Fraass, and D. L. McShan. Expanding the use and effectiveness of dose-volume histograms for 3-D treatment planning I: Integration of 3-D dose display. *Int. J. Radiat. Oncol. Biol. Phys.*, 29:1125–1131, 1994.

[64] Photon Treatment Planning Collaborative Working Group. State-of-the-art of external photon beam radiation treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:9–23, 1991.

[65] A. Brahme. Design principles and clinical possibilities for a new generation of radiation therapy equipment. *Acta Oncol.*, 26:403–412, 1987.

[66] A. Brahme and D. Reisstad. Microtrons for electron and photon radiotherapy. *IEEE Trans. Nucl. Sci,*, NS-28:1880–1883, 1981.

[67] A. Brahme, T Kraepelien, and H. Svensson. Electron and photon beams from a 50 MeV racetrack microtron. *Acta Radiol. Oncol.*, 19:305–319, 1980.

[68] J. M. Galvin, A. R. Smith, R. D. Moeller, R. L. Goodman, W. D. Powlis, J. Rubenstein, L. Solin, B. Michael, M. Needham, C. J. Huntzinger, and Kligermann. Evaluation of multileaf collimator design for a photon beam. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:789–801, 1993.

[69] J. M. Galvin, A. R. Smith, and B. Lally. Characterization of a multileaf collimator system. *Int. J. Radiat. Oncol. Biol. Phys.*, 25:181–192, 1993.

[70] T. J. Jordan and P. C. Williams. The design and performance characteristics of a multileaf collimator. *Phys. Med. Biol.*, 39:231–251, 1994.

[71] N. H. Barth. An inverse problem in radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:425–431, 1990.

[72] A. Brahme, A. J. Roos, and I. Lax. Solution of an integral equation in rotation therapy. *Phys. Med. Biol.*, 27:1221–1229, 1982.

[73] T. Bortfeld, J. Burkelbash, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied to conformation radiotherapy. *Phys. Med. Biol.*, 35:1423–1434, 1990.

[74] T. Bortfeld, D. L. Kahler, T. J. Waldron, and A. L. Boyer. X-ray field compensation with multileaf collimators. *Int. J. radiat. Oncol. Biol. Phys.*, 28:723–730, 1994.

[75] S. Webb. Optimizing the planning of intensity-modulated radiotherapy. *Phys. Med. Biol.*, 39:2229–2246, 1994.

[76] M. P. Carol. Beam modulation conformal therapy. In *International Symposium on 3D Radiation Treatment Planning and Conformal Therapy*, pages 435–445, St. Louis, Missouri, April 21–23, 1994. Proceedings.

[77] J. M. Galvin, X.-G. Chen, and R. M. Smith. Combining multileaf fields to modulate fluence distributions. *Int. J. Radiat. Oncol. Biol. Phys.*, 27:697–705, 1993.

[78] R. G. Lane, M. D. Loyd, C. H. Chow, E. Eklelunu, and I. I. Rosen. Custom beam profiles in computer-controlled radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 22:167–174, 1991.

[79] B. A. Fraass, L. Marsh, M. K. Martel, K. Forster, D. L. McShann, and K. TenHanken. Multileaf collimator-based intensity modulation for conformal therapy. *Med. Phys.*, 21:1008, 1994. Abstract.

[80] A. S. Zacarias, R. G. Lane, and I. I. Rosen. Assesment of a linear accelerator for segmented conformal radiation therapy. *Med. Phys.*, 20:193–198, 1993.

[81] P. Källman, B. K. Lind, A. Eklof, and A. Brahme. Shaping of arbitrary dose distributions by dynamic multileaf collimation. *Phys. Med. Biol.*, 33:1291–1300, 1988.

[82] S. V. Spirou and C. S. Chui. Generation of arbitrary intensity profiles by dynamic jaws or multileaf collimators. *Med. Phys.*, 21(1031-1041), 1994.

[83] C. C. Ling, C. Burman, C. S. Chui, G. J. Kutcher, S. A. Leibel, T. LoSasso, R. Mohan, T. Bortfeld, L. Reinstein, S. Spirou, X. Wang, Q. Wu, M. Zelefsky, and Z. Fuks. Conformal radiation treatment of prostate cancer using inversely-planned intensity modulated photon beams with dynamic multileaf collimation. *Int. J. Radiat. Oncol. Biol. Phys.*, 35(4):721–730, 1996.

[84] A. L. Boyer, L. Antonuk, A. Fenster, M. van Herk, H. Meertens, P. Munro, L. E. Reinstein, and J. Wong. A review of electronic portal imaging devices (EPID). *Med. Phys.*, 19(1):1–16, 1992.

[85] J. Leong and M. A. Stracher. Visualisation of internal motion within a treatment portal during a radiation treatment. *Radiother. Oncol.*, 9:153–156, 1987.

[86] M. C. Kirby and P. C. Williams. Portal imaging for the verification of breast treatments. *Radiother. Oncol.*, 22:314–316, 1991.

[87] W. de Neve, F. van den Heuvel, M. de Beukeeler, M. Coghe, L. Thon, W. de Roover, M. van Lancher, and G. Storme. Routine clinical on-line portal imaging followed by immediate field adjustment using a telecontrolled patient coach. *Radiother. Oncol.*, 24:45–54, 1992.

[88] D. J. Gladstone, M. van Herk, and L. M. Chin. Verification of lung attenuator positioning before total body irradiation using an electronic portal imaging device. *Int. J. Radiat. Biol. Oncol. Phys.*, 27(2):449–454, 1993.

[89] P. M. Evans, J. Q. Gildersleve, C. Rawlings, and W. Swindell. Technical note: the implementation of patient position correction using a megavoltage imaging device on a linear accelerator. *Br. J. Radiol.*, 66(789):833–838, 1993.

[90] M. A. Hunt, T. E. Schultheiss, G. E. Desobry, M. Hakki, and G. E. Hanks. An evaluation of setup uncertainties for patients treated to pelvic sites. *Int. J. Rad. Oncol. Biol. Phys.*, 32(1):227–233, 1995.

[91] A. Tinger, J. M. Michalski, W. R. Bosch, R. K. Valicenti, D. A. Low, and R. J. Myerson. An analysis of intratreatment and intertreatment displacements in pelvic radiotherapy using electronic portal imaging. *Int. J. Rad. Oncol. Biol. Phys.*, 34(3):683–690, 1996.

[92] E. Vigneault, J. Pouliot, J. Laverdiere, J. Roy, and M. Dorion. Electronic portal imaging device detection of radioopaque markers for the evaluation of prostate position during megavoltage irradiation: clinical study. *Int. J. Radiat. Biol. Oncol. Phys.*, 37(1):205–212, 1997.

[93] E. Fontenla, C. A. Pelizzari, and G. T. Chen. Implications of 3-dimensional target shape and motion in aperture design. *Med. Phys.*, 23(8):1431–1441, 1996.

[94] A. Bel, M. van Herk, and J. V. Lebesque. Target margins for random geometrical treatment uncertainties in conformal radiotherapy. *Med. Phys.*, 23(9):1537–1545, 1996.

[95] R. Boellaard, M. van Herk, and B. J. Mijnheer. A convolution model to convert transmission dose images to exit dose distributions. *Med. Phys.*, 24(2):189–199, 1997.

[96] T. R. McNutt, T. R. Mackie, P. Reckwerdt, and B. R. Paliwal. Modeling dose distributions from portal dose images using the convolution/superposition method. *Med. Phys.*, 23(8):1381–1392, 1996.

[97] V. N. Hansen, P. M. Evans, and W. Swindell. The application of transit dosimetry to precision radiotherapy. *Med. Phys.*, 23(5):713–721, 1996.

[98] T. R. McNutt, T. R. Mackie, P. Reckwerdt, and B. R. Paliwal. Calculation of portal dose using the convolution/superposition method. *Med. Phys.*, 23(4):527–535, 1996.

[99] M. C. Kirby and P. C. Williams. Measurement possibilities using an electronic portal imaging device. *Radiother. Oncol.*, 29(2):237–243, 1993.

[100] M van Herk, K Gilhuijs, P Williams, P. Cinquin, L. Cionini, and W. Swindell. Role of electronic imaging in high dose/high rate precision therapy. *Radioth. Oncol.*, 29:269–270, 1993.

[101] B. A. Fraass. The development of conformal radiation therapy. *Med. Phys.*, 22(11 Pt.2):1911–1921, 1995.

[102] A. B. Wolbarst, L. M. Chin, and G. K. Svensson. Optimization of radiation therapy: integral-response of a model biological system. *Int. J. Radiat. Oncol. Biol. Phys.*, 8:1761–1769, 1982.

[103] J. T. Lyman. Complication probabilities as assessed from dose-volume histograms. *Radiat. Res.*, 104:S13–S19, 1985.

[104] H. R. Withers, J. M. Taylor, and B. Maciejewski. Treatment volume and tissue tolerance. *Int. J. Radiat. Oncol. Biol. Phys.*, 14:751–759, 1988.

[105] G. J. Kutcher and C. Burman. Calculation of complication probability factors for nonuniform normal tissue irradiation: the effective volume method. *Int. J. Radiat. Oncol. Biol. Phys.*, 16:1623–1630, 1989.

[106] A. Ågren, A. Brahme, and I. Turesson. Optimization of uncomplicated control for head and neck tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 19:1077–85, 1990.

[107] P. Källman, A. Ågren, and A. Brahme. Tumour and normal tissue responses to fractionated non-uniform dose delivery. *Int. J. Radiat. Biol.*, 62(2):249–262, 1992.

[108] A. Niemierko and M. Goitein. Modelling of normal tissue response to radiation: the critical volume-model. *Int. J. Radiat. Oncol. Biol. Phys.*, 25:135–145, 1992.

[109] C. Raphael. Mathematical modelling of objectives in radiation therapy treatment planning. *Phys. Med. Biol.*, 37:1293–1312, 1992.

[110] A. C. Jackson, G. J. Kutcher, and E. D. Yorke. Probability of radiation-induced complications for normal tisues with parallel architecture subject to non-uniform irradiation. *Med. Phys.*, 20:613–625, 1993.

[111] T. E. Shultheiss and C. G. Orton. Models in radiotherapy: Definitions of decision criteria. *Med. Phys.*, 12:183–187, 1985.

[112] P. Källman, B. K. Lind, and A. Brahme. An algorithm for maximizing the probability of complication-free tumour control in radiation therapy. *Phys. Med. Biol.*, 37:871–890, 1992.

[113] N. L. Jain, M. G. Kahn, R. E. Drzymala, B. E. Emami, and J. A. Purdy. Objective evaluation of 3-d radiation treatment plans:A decision-analytic tool incorporating treatment preferences of radiation oncologists. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:321–333, 1993.

[114] R. Mohan, G. S. Mageras, B. Baldwin, L. J. Brewster, G. J. Kutcher, S. Leibel, C. M. Burman, C. C. Ling, and Z. Fuks. Clinically relevant optimization of 3-d conformal treatments. *Med. Phys.*, 19(4):933–944, 1993.

[115] G. S. Mageras and R. Mohan. Application of fast simulated annealing to optimization of conformal treatments. *Med. Phys.*, 20(3):639–647, 1993.

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

[117] A. Gustafsson, B. K. Lind, R. Svensson, and A. Brahme. Simultaneous optimization of dynamic multileaf collimation and scanning patterns or compensation filters using a generalized pencil beam algorithm. *Med. Phys.*, 22(7):1141–1156, 1995.

[118] X. Wang, R. Mohan, A Jackson, S. A. Leibel, Z. Fuks, and C. C. Ling. Optimization of intensity-modulated 3d conformal treatment plans based on biological indices. *Radiotherapy and Oncology*, 37:140–152, 1995.

[119] S. Hyödynmaa, A. Gustafsson, and A. Brahme. Optimization of conformal electron beam therapy using energy- and fluence-modulated beams. *Med. Phys.*, 23(5):659–666, 1996.

[120] S. Söderstrom, A. Gustafsson, and A. Brahme. The clinical value of different treatment objectives and degrees of freedom in radiation therapy optimization. *Radiotherapy and Oncology*, 29:148–163, 1993.

[121] S. Webb. Optimization by simulated annealing of three-dimensional conformal treatment palnning for radiation fields defined by multileaf collimator. II. Inclusion of two-dimensional modulation of the X-ray intensity. *Phys. Med. Biol.*, 37(8):1689–1704, 1992.

[122] R. Mohan, X. Wang, T. Bortfeld, A. L. Boyer, G. J. Kutcher, S. A. Leibel, Z. Fuks, and C. C. Ling. The potential and limitations of the inverse radiotherapy technique. *Radiotherapy and Oncology*, 32:232–248, 1994.

[123] A. Niemierko. Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Med. Phys.*, 24(1):103–110, 1997.

[124] A. J. van der Kogel. Radiobiological and clinical data on dose/volume effect relationships in normal tissues. *Radiother. Oncol.*, 29:127–128, 1993.

[125] A. Dutriex. Prescription, precision and decision in treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1291–1296, 1987.

[126] Goitein M. The comparison of treament plans. *Semin. Radiat. Oncol.*, 2:246–256, 1992.

[127] P. S. Hahn, S. Shalev, D. Viggars, and P. Therrien. Treatment planning for protocol-based radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:937–939, 1990.

[128] A. Niemierko. Random search algorithm (RONSC) for optimization of radiation therapy with both physical and biological end points and constraints. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:89–98, 1992.

[129] M. Langer, R. Brown, M. Urie, J. Leong, M. Stracher, and J. Shapiro. Large-scale optimization of beam weights under dose-volume restrictions. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:887–893, 1990.

[130] G. K. Bahr, J. G. Kereiakes, H. Horwitz, R. Finney, J. Galvin, and K. Goode. The method of linear programming applied to radiation treatment planning. *Radiology*, 91:686–693, 1968.

[131] L. Hodes. Semiautomatic optimization of external beam radiation treatment planning. *Radiology*, 110:191–196, 1974.

[132] A. T. Redpath, B. L. Vickery, and D. H. Wright. A new technique for radiotherapy using quadratic programming. *Phys. Med. Biol.*, 21:781–791, 1974.

[133] D. Sonderman and P. G. Abrahamson. Radiotherapy treatment design using mathematical programming models. *Oper. Res.*, 33:705–725, 1985.

[134] S. M. Morrill, R. G. Lane, J. A. Wong, and I. I. Rosen. Dose-volume considerations with linear programming. *Med. Phys.*, 18(6):1200–1210, 1991.

[135] M. Langer and J. Leong. Optimization of beam weights under dose-volume restrictions. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1255–1259, 1987.

[136] M. Langer, S. Morrill, R. Brown, M. Urie, O. Lee, and R. Lane. A comparison of mixed integer programming and fast simulated annealing for optimizing beam weights in radiation therapy. *Med. Phys.*, 23(6):957–964, 1996.

[137] M. Langer, R. Brown, S. Morrill, R. Lane, and O. Lee. A genetic algorithm for generating beam weights. *Med. Phys.*, 23(6):965–971, 1991.

[138] I. Rosen, R. G. Lane, S. M. Morrill, and J. A. Belli. Treatment plan optimization using linear programming. *Med. Phys.*, 18(2):141–152, 1991.

[139] A. Niemierko, M. Urie, and M. Goitein. Optimization of 3D radiation therapy with both physical and biological end points and constraints. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:99–108, 1992.

[140] G. Starkshall. A constrained least-square optimization method for external beam radiation therapy treatment planning. *Med. Phys.*, 11:659–665, 1984.

[141] S. Webb. Optimization of conformal radiotherapy dose distributions by simulated annealing. *Phys. Med. Biol.*, 34:1349–1370, 1989.

[142] A. Djordjevitch, D. J. Bonham, E. M. A. Hussein, J. W. Andrew, and M. E. Hale. Optimal design of radiation compensators. *Med. Phys.*, 17:397–404, 1990.

[143] S. Webb. Optimization by simulated annealing of three-dimensional, conformal treatment planning for radiation fields defined by a multileaf collimator. *Phys. Med. Biol.*, 36:1201–1226, 1991.

[144] K. J. Weeks and M. R. Sontag. 3-D dose-volume compensation using nonlinear least-squares regression technique. *Med. Phys.*, 18:474–480, 1991.

[145] T. W. Holmes and T. R. Mackie. A unified approach to the optimization of brachytherapy and external beam therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 20:859–873, 1991.

[146] T. W. Holmes, T. R. Mackie, and P. J. Reckwerdt. An iterative filtered backprojection inverse treatment planning algorithm for tomotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1215–1225, 1995.

[147] L. Xing and G. T. Y. Chen. Iterative methods for inverse treatment planning. *Phys. Med. Biol.*, 41:2107–2133, 1996.

[148] L. Xing, C. Pelizzari, F. T. Kuchir, and G. T. Y. Chen. Optimization of relative weights and wedge angles in treatment planning. *Med. Phys.*, 24(2):215–221, 1997.

[149] B. K. Lind. Properties of an algorithm for solving the inverse problem in radiation therapy. *Inverse Problems*, 6:415–426, 1990.

[150] J. Stein, R. Mohan, X.-H. Wang, Q. Wu, K. Preiser, C. Ling, and W. Schlegel. Number and orientations of beams in intensity-modulated radiation treatments. *Med. Phys.*, 24(2):149–160, 1997.

[151] A. Brahme. Optimization of stationary and moving beam therapy techniques. *Radioth. Oncol.*, 12:129–140, 1988.

[152] A. L. Boyer, G. E. Desorby, and N. H. Wells. Potential and limitations of invariant kernel conformal therapy. *Med. Phys.*, 18(4):703–712, 1991.

[153] A. Brahme. Treatment optimization using physical and radiobiological objective functions. In A. Smith, editor, *Radiation Therapy Physics*, pages 209–246. Springer, Berlin, 1995.

[154] A. Jackson, R. Mohan, and B. Baldwin. Comments on "Sampling techniques for the evaluation of treatment plans". *Med. Phys.*, (5):1375–1376, 1993.

[155] X.-Q. Lu and L. M. Chin. Further discussions on sampling methods-a response to Letters to the Editor [Med. Phys. 20, 1375-1376 (1993)] and [Med. Phys. 20, 1377-1380 (1993)]. *Med. Phys.*, (5):1375–1376, 1993.

[156] M. Langer, R. Brown, P. Kijewski, and H. Chul. The reliability of optimization under dose-volume limits. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:529–538, 1993.

[157] S. M. Morrill, , I. I. Rosen, R. G. Lane, and J. A. Belli. The influence of dose constraint point placement in optimizaed radiation therapy treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 19(2):129–142, 1990.

[158] Z. Chen, X. Wang, T. Bortfeld, R. Mohan, and L. Reinstein. The influence of scatter on the design of optimized intensity modulations. *Med. Phys.*, 22(11 Pt. 1):1727–1733, 1995.

[159] T. R. Mackie, J. W. Scrimger, and J. J. Battista. A convolution method of calculating dose for 15-MV x-rays. *Med. Phys.*, 12:188–196, 1985.

[160] A. L. Boyer and E. C. Mok. A photon dose distribution model employing convolution calculations. *Med. Phys.*, 12:169–177, 1985.

[161] R. Mohan, C. Chui, and L. Lidofsky. Differential pencil beam dose computation model for photons. *Med. Phys.*, 13:64–73, 1986.

[162] J. E. McGary and A. L. Boyer. An interactive, parallel, three-dimensional fast Fourier transform convolution dose calcualtion using a supercomputer. *Med. Phys.*, 24(4):519–522, 1997.

[163] D. H. Hristov and B. G. Fallone. Parametrization of the independent variables in treatment planning optimization. *Med. Phys.*, 23(6):1111–1112, 1996. (abstract).

[164] D. H. Hristov and B. G. Fallone. An active set algorithm for treatment planning optimization. *Med. Phys.*, 24(6):1052, 1997. (abstract).

[165] D. H. Hristov and B. G. Fallone. An active set algorithm for treatment planning optimization. *Med. Phys.*, 24(9):1445–1464, 1997.

[166] D. H. Hristov and B. G. Fallone. An approach to the inverse treatment planning technique. *Med. Phys.*, 23(5):812, 1996. (abstract).

[167] D. H. Hristov and B. G. Fallone. An approach to the inverse treatment planning technique. In *42nd Annual General Meeting of The Canadian Organization of Medical Physicists*, Vancouver, British Columbia, 20-22 June, 1996. Proceedings.

[168] D. H. Hristov and B. G. Fallone. Optimization and applications of an inverse treatment planning algorithm. *Med. Phys.*, 24(7):1204, 1997. (abstract).

[169] D. H. Hristov and B. G. Fallone. Optimization and applications of an inverse treatment planning algorithm. In *43nd Annual General Meeting of The Canadian Organization of Medical Physicists*, pages 134–136, Charlottetown, Prince Edward Islands, 10-12 July, 1996. Proceedings.

[170] J. P. Gerard. High dose high-precision in radiation oncology. *Radiother. Oncol.*, 29:167–168, 1993.

[171] V. Rudat, M. Flentje, D. Oetzel, M. Menke, W. Schlegel, and M. Wannenmacher. Influence of the positioning error on 3D conformal dose distributions during fractionated radiotherapy. *Radiother. Oncol.*, 33:157–166, 1994.

[172] A. Dutreix. When and how can we improve precision in radiotherapy. *Acta Radiol. Oncol.*, 2:275–292, 1984.

[173] M. Goitein and J. Busse. Immobilization error: some theoretical considerations. *Radiology*, **117**:407–412, 1975.

[174] A. Brahme. Dosimetric precision requirements in radiation therapy. *Acta Radiol. Oncol.*, 23:379–391, 1984.

[175] M. Goitein and T. Schultheiss. Strategies for treating possible tumor extensions: some theoretical considerations. *Int. J. Radiat. Oncol. Biol. Phys.*, **11**:1519–1528, 1985.

[176] I. Daftari, P. L. Petti, J. M. Collier, J. R. Castro, and P. Samuel. The effect of patient motion on dose uncertainty in charged particle irradiation for lesions encircling the brain stem or spinal cord. *Med. Phys.*, 18:1105–1115, 1991.

[177] J. J. Kinzie, G. E. Hanks, C. J. Maclean, and S. Kramer. Patterns of care study: Hodgkin's disease relapse rates and adequacy of portals. *Cancer*, 52:2223–2226, 1983.

[178] J. E. White, T. Chen, J. McCracken, P. Kennedy, H. G. Seydel, G. Hartman, J. Mira, M. Khan, F. Y. Durrance, and O. Skinner. The influnce of radiation therapy quality control on survival, response and sites of relapse in oat cell carcinoma of the lung. *Cancer*, 50:1084–1090, 1982.

[179] I. Rabinowitz, J. Broomberg, M. Goitein, K. McCarthy, and J. Leong. Accuracy of radiation field alignment in clinical practice. *Int. J. Radiat. Oncol. Biol. Phys.*, 11:1857–1867, 1985.

[180] B. Kihlen and B. I. Ruden. Reproducibility of field alignment in radiation therapy. *Acta Oncol.*, 13:1583–1588, 1989.

[181] S. A. Rosenthal, J. M. Galvin, J. W. Goldwein, A. R. Smith, and P. H. Blitzer. Improved methods for determination of variability in patient positioning for radiation therapy using simulation and serial portal film measurements. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:621–625, 1992.

[182] J. V. Lebesque, A. Bel, J. Bijhold, and A. A. M. Hart. Detection of systematic patient errors by portal film analysis. *Radiother. Oncol.*, 23:198, 1992.

[183] J. E. Marks, A. G. Haus, H. G. Sutton, and M. L. Griem. Localization error in the radiotherapy of Hodgkin's disease and malignant lymphoma with extended mantle fields. *Cancer*, 34:83–90, 1974.

[184] J. E. Marks and A. G. Haus. The effect of immobilization on localization error in the radiotherapy of head and neck cancer. *Clin. Radiol.*, 27:175–177, 1976.

[185] R. W. Byhardt, J. D. Cox, A. Hornburg, and G. Liermann. Weekly localization films and detection of field placement errors. *Int. J. Radiat. Oncol. Biol. Phys.*, 4:881–887, 1978.

[186] P. B. Dunscombe and K. Fox. The precision of determining compliance with prescribed fields from conventional portal films. *Br. J. Radiol.*, 62:48–52, 1989.

[187] P. B. Dunscombe and K. Fox. Portal radiography practice in Canada. *Int. J. Radiat. Oncol. Biol. Phys.*, 24:177–179, 1992.

[188] L. E. Reinstein, H. I. Amols, P. J. Biggs, R. T. Droege, A. B. Filimonov, W. R. Lutz, and S. Shalev. Radiotherapy portal imaging quality. Report #24, AAPM, New York, American Institute of Physics, 1988.

[189] J. E. Marks, A. G. Haus, H. G. Sutton, and M. L. Griem. The value of frequent treatment verification films in reducing localization error in the irradiation of complex fields. *Cancer*, 37:2755–2761, 1976.

[190] W. De Neve. An overview of correction strategies for setup errors. In *4th International Workshop on Electronic Portal Imaging*, Amsterdam, The Netherlands, 1996. Proceedings.

[191] H. Wang. Private communication.

[192] F.-F. Yin, P. Rubin, M. SChell, R. Wynn, B. West, G. Uschold, and D. Nelson. Clinical implementation of epid at the university of rochester cancer center. In *3rd International Workshop on Electronic Portal Imaging*, San Francisco, U.S.A., 1994. Proceedings.

[193] M. Goitein. The potential of real time portal imaging. In *3rd International Workshop on Electronic Portal Imaging*, San Francisco, U.S.A., 1994. Proceedings.

[194] M. G. Herman, R. A. Abrams, and R. Mayer. Clinical use of on-line portal imaging for daily patient treatment verification. *Int. J. Radiat. Oncol. Biol. Phys.*, 24:216, 1992. Suppl. 1.

[195] C. Rawlings, S. Griffits, H. Probst, and N. Howson. Clinical experience with megavoltage portal imaging. In *4th International Workshop on Electronic Portal Imaging*, Amsterdam, The Netherlands, 1996. Proceedings.

[196] P. Dunscombe. The future of portal verification. In *Joint Meeting of Canadian Association of Radiologists, Canadian Organization of Medical Physicists and the Canadian College of Physicists in Medicine*, page 3, Montreal, Quebec, June 4–7, 1995. CCPM Symposium: Future Trends in Medical Imaging.

[197] J. Leong. A digital image processing system for high energy x-ray portal images. *Phys. Med. Biol.*, 29:1527–1535, 1984.

[198] H. Meertens. Digital processing of high energy photon beam images. *Med. Phys.*, 12:111–113, 1985.

[199] G. Sherouse, J. Rosenman, H. McMurr, S. Pizer, and E. Chaney. Automatic digital contrast enhancement of radiation therapy portal films. *Int. J. Radiatian Oncology Biol. Phys.*, 13:801–806, 1987.

[200] H. Amols and T. Lowinger. An inexpensive microcomputer based port film image enhancement system. *Med. Phys. (abstract)*, 14:483, 1987.

[201] W. Powlis, I. Brikman, S. Seshardri, and P. Block. Portal radiographs: Digital enhancement of contrast. *Radiology*, 169:839–841, 1988.

[202] S. Jones and A. Boyer. Investigation of an FFT-based correlation technique for verification of radiation treatment setup. *Med. Phys.*, 18:1116–1125, 1991.

[203] J. Bijhold, K.G.A. Gilduijs, and M. van Herk. Automatic verification of radiation field shape using digital portal images. *Med. Phys.*, 19:1007–1014, 1992.

[204] I. Crooks and B.G. Fallone. Contrast enhancement of portal images by selective histogram equalization. *Med. Phys.*, 20:199–204, 1993.

[205] I. Crooks and B.G. Fallone. A novel algorithm for the edge-detection and edge-enhancement of medical images. *Med. Phys.*, **20**:993–998, 1993.

[206] K.W. Leszczynski, S. Shalev, and N.S. Cosby. The enhancement of radiotherapy verification images by an automated edge detection technique. *Med. Phys.*, 19:611–621, 1992.

[207] J.M. Balter, C.A. Pelizzare, and G.T.Y. Chen. Correlation of projection radiographs in radiation therapy using open curve segments and points. *Med. Phys.*, 19:329–334, 1992.

[208] T. Radcliffe, R. Rajapakshe, and S. Shalev. Pseudocorrelation: A fast, robust, absolute, grey-level image alignment algorithm. *Med. Phys.*, 21:761–769, 1994.

[209] H. Wang and B.G. Fallone. A robust morphological algorithm for automatic radiation field-extraction and correlation of portal images. *Med. Phys.*, 21:237–244, 1994.

[210] K. G. A. Gilhuijs and M. van Herk. An automatic in-line inspection of patient setup in radiation therapy using digital portal images. *Med. Phys.*, 20:667–677, 1993.

[211] G. X. Ding, S. Shalev, and G. Gluchev. A ρ-θ technique for treatment verification in radiotherapy and its clinical applications. *Med. Phys.*, 20:1135–1143, 1993.

[212] S. Zhou and L.J. Verhey. A robust method of multileaf collimator (mlc) leaf-configuration verification. *Phys. Med. Biol.*, 39:1929–1948, 1993.

[213] K. Eilertsen, A. Skretting, and T. L. Tennvassås. Methods for fully automated verification of patient set-up in external beam radiotherapy with polygon shaped fields. *Phys. Med. Biol.*, 39:993–1012, 1994.

[214] J. Moseley and P. Munro. A semiatomatic method for registration of portal images. *Med. Phys.*, 21(4):551–558, 1994.

[215] K. W. Leszczynski, S. Loose, and P. Dunscombe. Segmented chamfer matching for prescription-treatment image registration in radiotherapy. *Phys. Med. Biol.*, 40:83–94, 1995.

[216] H. Wang and B.G. Fallone. A mathematical model of radiation field edge localization. *Med. Phys.*, 22:1107–1110, 1995.

[217] L. Dong and A. L. Boyer. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1053–1060, 1995.

[218] D. S. Fritsch, E. L. Chaney, A. Boxwala, M. J. McAuliffe, S. Raghavan, A. Thall, and J. R. Earnhart. Core-based portal image registration for automatic radiotherapy treatment verification. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1287–1230, 1995.

[219] K. G. A. Gilhuijs, P. J. H. van de Hen, and M. van Herk. Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Med. Phys.*, 23(3):389–399, 1996.

[220] J. Cai, S.-Q. Zhou, J. Hopkinson, V. A. Saxena, and R. S. Cox. Alignment of multi-segmented anatomical features from radiation therapy images by using least square fitting. *Med. Phys.*, 23(12):2069–2075, 1996.

[221] D. H. Hristov, T. Falco, and B. G. Fallone. Automatic image correlation of megavoltage digitally reconstructed radiographs and portal images. Montreal, Quebec, June 4–7, 1995. Joint Meeting of Canadian Association of Radiologists, Canadian Organization of Medical Physicists and the Canadian College of Physicists in Medicine.

[222] B. G. Fallone, D. H. Hristov, E. B. Podgorsak, and L. Souhami. Digital portal verification of stereotactic radiosurgery. Montreal, Quebec, Sept. 13–17, 1995. Annual Meeting of The Royal College of Physicians and Surgeons of Canada.

[223] D. H. Hristov and B. G. Fallone. An approach to the inverse treatment planning technique. Vancouver, British Colimbia, 20-22 June, 1996. 42nd Annual General Meeting of The Canadian Organization of Medical Physicists.

[224] D. H. Hristov and B. G. Fallone. Parameterization of the independent variables in treatment planning optimization. In *38th Annual General Meeting of The American Association of Medical Physicists*, Philadelphia, PA, 21-25 July, 1996.

[225] D. H. Hristov and B. G. Fallone. Optimization and applications of an inverse treatment planning algorithm. In *43nd Annual General Meeting of The Canadian Organization of Medical Physicists*, Charlottetown, Prince Edward Islands, 10-12 July, 1996. J. R. Cunningham Young Investigators Symposium.

[226] D. H. Hristov and B. G. Fallone. An active set algorithm for treatment planning optimization. In *39th Annual General Meeting of The American Association of Medical Physicists*, Milwakee, Wisconsin, 27-31 July, 1996.

[227] D. H. Hristov and B. G. Fallone. Automatic image correlation of megavoltage digitally reconstructed radiographs and portal images. *Med. Phys.*, 23(5):661, 1995. (abstract).

[228] D. H. Hristov, T. Falco, and B. G. Fallone. Automatic image correlation of megavoltage digitally reconstructed radiographs and portal images. In *Joint Meeting of Canadian Association of Radiologists,*

Canadian Organization of Medical Physicists and the Canadian College of Physicists in Medicine, pages 17–18, Montreal, Quebec, June 4–7, 1995. Proceedings.

[229] D. H. Hristov and B. G. Fallone. Automatic image correlation of megavoltage digitally reconstructed radiographs and portal images. In *4th International Workshop on Electronic Portal Imaging*, pages 17–18, Amsterdam, The Netherlands, June 10–12, 1996. Proceedings.

[230] D. H. Hristov and B. G. Fallone. A grey-level image alignment algorithm for registration of portal images and digitally reconstructed radiographs. *Med. Phys.*, 23(1):75–84, 1996.

[231] D. H. Hristov and B. G. Fallone. A continuous penalty function method for inverse treatment planning. *Med. Phys.*, 25:208–223, 1998.

CHAPTER 2 An Active Set Algorithm for Inverse Treatment Planning

2.1 Introduction

2.1.1 Existing Approaches

The idea of using intensity modulated beams for conformal therapy was first proposed by Brahme *et al.*[1] who, by inverting analytically a prespecified dose distribution, solved the inverse problem for the case of a circular symmetric target with a healthy organ at its center. Their solution was then generalized for the case of targets with an axis of symmetry[2, 3] and arbitrary shape.[4] The resulting profiles, however, are generally piece-wise negative and therefore not physically realizable. Tulovsky *et al.*[5] addressed this problem for the case of circular isocentrically centered target and obtained non-negative optimal profiles for this case. However, the *analytical approach* ignores beam divergence, tissue inhomogeneities and lateral scatter, which are to be considered in order to reduce unwanted injuries to critical organs.[6]

Recent advances in computer technology have fostered considerable interest in the numerical solutions of the inverse problem for targets of arbitrary shapes. A brief description of these methods is given below.

Iterative deconvolution[7–9] is a method for inverse treatment planning similar to the filtered backprojection reconstruction of computed tomography (CT) images. In CT the 3-D distribution of the linear attenuation coefficient is reconstructed from the transmission data by filtering of the measured profiles and subsequent backprojection. The iterative filtered backprojection implementation[9] which is the most versatile of the iterative deconvolution techniques,[10] inverts this process. Within each transversal slice of the patient data, the desired dose distribution is projected along the source rays to produce an estimate of the desired profiles in that plane. These estimates are further filtered to account for lateral scatter and geometric blur. The resulting profiles are used for forward calculation of the dose distributions and the difference between the desired

and the calculated dose distributions is then used as an input for a new iteration. This procedure requires full 3-D dose calculation at each iteration. In addition, the technique is not readily applicable to non-coplanar beams since it works on slice-per-slice basis.

Iterative dose reconstruction techniques[11–13] for inverse treatment planning are analogous to the algebraic reconstruction techniques in tomographic imaging. They employ various updating schemes to obtain a dose distribution which is the best match of the desired one in least-square sense. Similarly to the iterative deconvolution techniques, iterative dose reconstruction methods are restricted to least-square dose objectives. Their advantage is, however, that the intensities of non-coplanar beams can be optimized.††

Feasibility search[14, 15] techniques for inverse treatment planning aim at the identification of feasible rather than optimum solutions of the inverse problem. A feasibility search algorithm requires a set of convex constraints that the prescribed dose must satisfy. (An example of convex constraint is the requirement that the dose to any point in a critical organ be below a certain planner-specified level.) Then a projection of the pencil beam weights onto the convex sets specified by the constraints is established. The pencil beam weights are projected alternatively among the constraint sets until convergence is achieved. If the intersection of the convex sets of the constraints is not empty, the resulting pencil beam profiles produce dose distributions which satisfy the dose constraints.

There are two drawbacks of the feasibility search techniques. First, they produce intensity modulation that satisfies the constraints (if such exists) but in case of success do not indicate how much more stringent a constrained could be. Second, they cannot be easily generalized to consider non-linear functions such as TCP and NTCPs.

Djordjevitch *et. al*[16] formulated the problem of optimal compensator design as a *quadratic programming* problem. A linear-quadratic function is minimized which com-

†† Bortfeld's technique[11] can also be considered as a gradient technique.

bines the integral dose to selected healthy points and the least-square difference between the target dose and a preset level. The linear constraints require that the dose to selected vulnerable points be below specified limits. Therefore the approach can be regarded as further generalization of the inverse planning techniques based on least-square objectives. However, quadratic programming is restricted to linear-quadratic functions and cannot be extended to nonlinear functions of dose. Furthermore, quadratic programming techniques solve sequences of linear programming problems and therefore fail to provide approximate solutions when feasible parameters that satisfy the constraints do not exist.[17, 18]

Simulated annealing[19] is a stochastic optimization method based on the Metropolis algorithm[20] for simulation of a collection of atoms in equilibrium at a given temperature T. In each step of the Metropolis algorithm an atom is given a small random displacement and the change ΔE in the energy of the system is evaluated. If $\Delta E \leq 0$ the new configuration is retained as a starting point for the new step. If $\Delta E > 0$ the configuration is accepted with a probability $P(\Delta) = \exp(-\Delta E/k_B T)$, where k_B is the Boltzmann constant. By repeating the basic step many times the thermal motion of a collection of atoms with a heat bath at temperature T is simulated. In order to find the ground state of the system one should proceed with the simulations in the manner followed by the experiments that determine the low-temperature state of a material – for example, experiments that grow single crystal from a melt. After the substance is melted, the temperature is lowered slowly, with a substantial amount of time spent in the vicinity of the freezing point.

The simulated annealing method uses this approach by substituting the cost function in place of the energy and by defining configurations in terms of the parameters to be optimized. The temperature is a control parameter in the same units as the cost function. The system is optimized by "melting" at a high temperature and then lowering

the temperature by slow decrements until the system "freezes". At each temperature the simulation must proceed long enough for the system to reach a steady state. The annealing schedule is determined by the number of rearrangements at each temperature and the temperature sequence.

The most attractive feature of the simulated annealing is that it can find the global minimum of virtually any function given a sufficiently slow cooling scheme. However, similar to the physical systems which when annealed rapidly reach a metastable state different from the ground one, the simulated annealing method is bound to produce a minimum different from the global one if the cooling is too fast. In addition, a large number of iterations is required at each temperature to thermalize the system. For example, 4 million iterations have been reported for the optimization of up to 8192 variables (128 beams with 64 beamlets) in a 3-D model (32^3 dose points) that employed a least-square dose objective.[21] A single iteration does not require a full 3-D dose calculation, but does evaluate the cost function from the 3-D dose distribution. Therefore, an equivalent number of approximately 3000 full 3-D dose calculations must be evaluated. The excessive number of dose calculations is the main factor that currently limits the applicability of the simulated annealing to the large-scale inverse treatment planning problem.

Gradient algorithms[22–26] are somewhat optimal for intensity modulation design. First, they have much better rate of convergence than the stochastic methods. Second, gradient techniques can be applied to optimization problems that are beyond the scope of dose reconstruction, feasibility searches and quadratic programming techniques. Examples of such problems are optimizations based on biological or mixed biological/physical objective functions.

Gradient methods do have their drawbacks. First, they are applicable only to

differentiable functions of the optimization variables. This requirement, however, can be met by almost all clinically significant biological or physical indexes.[26] More relevant limitation is the inability of gradient methods to avoid local minima. In this case, if the solution is not acceptable, the gradient search should be restarted with a different initial guest. Another approach is to identify proper starting points for certain beam arrangements and treatment sites and to use these for new cases.[24, 25]

2.1.2 Objectives

Among the gradient methods, the constrained steepest-descent (CSD) has been exclusively employed for inverse treatment planning.[22–26] Its main advantage is that the requirement of nonnegative pencil beam weights (nonnegativity constraint on the independent variables) can be satisfied by simply imposing them at each iteration step.[22] Such a procedure is not mathematically robust for other gradient-based approaches such as the conjugate gradient (CG) and the quasi-newton (QN) methods. These methods rely on the notion that the function to be minimized can be approximated by a quadratic form in the vicinity of a minimum. They exploit the properties of the quadratic functions to construct a self-consistent set of descent directions such that successive line minimizations along these directions lead to the minimum. However, when nonnegativity constraints are imposed, a line minimization can lead to an infeasible iterate. Truncation to zero of the negative variables creates a feasible point which is different from the line minimum. This disrupts the necessary process of successive line minimizations. On the other hand the afore mentioned methods have higher convergence rate than the steepest-descent method for unconstrained minimization problems.[27] Thus their use can potentially accelerate the design of intensity modulated beams by decreasing the number of time-consuming three-dimensional dose computations.[28, 29]

Our objective is to develop robust implementation of the CG for the purposes

of inverse treatment planning. We propose an active set method with a conjugate-gradient routine for subspace minimization in order to accommodate the non-negativity constraints. The utility of the active set method for treatment planning optimization is to be investigated by a comparison of its performance to that of the CSD. Two cost functions, different treatment geometries, relevant starting points and termination criteria are considered for the purposes of objective comparison.

2.2 Background

2.2.1 Beam Intensity Modulation in Dose Calculations

Beam intensity modulation is included in dose calculations by tiling the cross-section of each beam in small elements (pencil beams) at a certain reference distance (Fig. 2.1). The pencil beams are indexed by c, which is the number of the beam they belong to and two indices (j, k) indicating the position of the pencil beam within the beam port. The coefficient $H_{j,k,c}^{m,n,l}$ is the calculated dose contribution per unit weight from a pencil beam (j, k, c) to a voxel P at a grid location (m, n, l) . The pencil beam dose contribution depends on different physical factors: the geometry of the irradiation, the treatment modality, and patient anatomy. However, in practice, the coefficients $H_{j,k,c}^{m,n,l}$ reflect the modelling of the above physical factors by the particular dose calculation model. For instance, if scattered radiation is not considered, $H_{j,k,c}^{m,n,l}$ is zero beyond the ray line of a pencil beam, contrary to the case where the scatter is taken into account.

The inclusion of beam intensity modulation in dose calculations is done by assigning a nonnegative weight $w^{j,k,c}$ to each pencil beam. The dose $D^{m,n,l}$ to a voxel P is then:

$$D^{m,n,l} = H_{j,k,c}^{m,n,l} w^{j,k,c}. \tag{2.1}$$

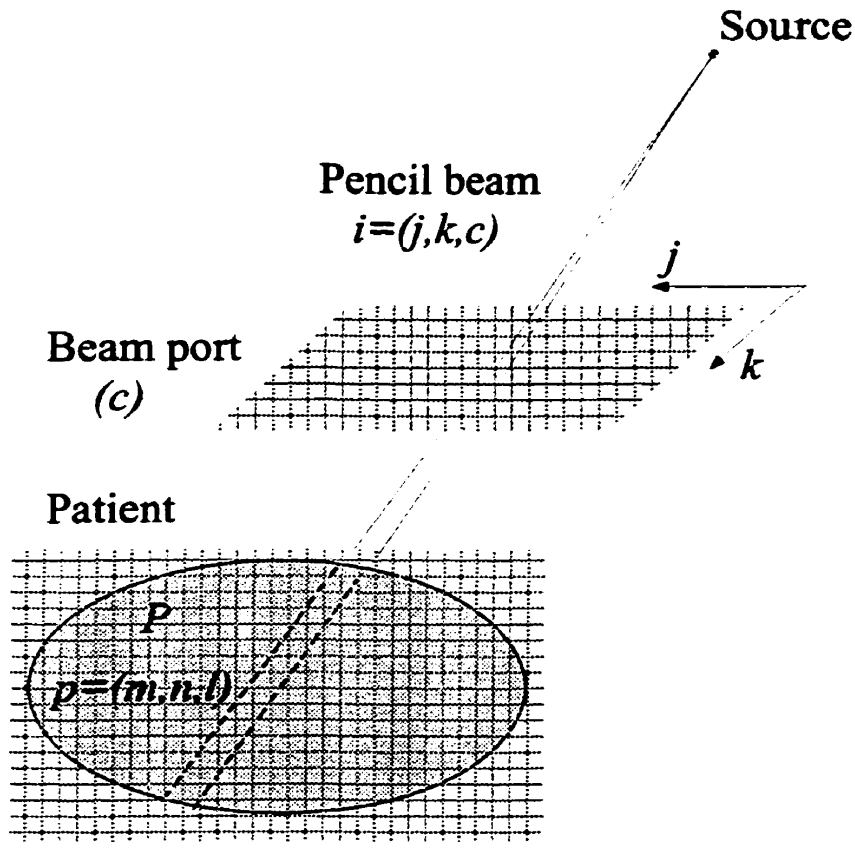

Figure 2.1. *Inclusion of modulated beam intensity in dose calculations. Only one beam is shown for clarity.*

Summation over repeated indices is assumed. If the indices are combined as shown in Fig. 2.1 we have:

$$D^p = H_i^p w^i, \qquad p = 1..N, \qquad i = 1..M \tag{2.2}$$

where N is the total number of dose calculation voxels and M is the total number of pencil beams used for optimization. In vector notation:

$$\mathbf{D} = \hat{\mathbf{H}}\mathbf{w}, \qquad \mathbf{D} = \{D^p(\mathbf{w})\}_{p=1}^{N}, \quad \hat{\mathbf{H}} = \{H_i^p\}, \quad \mathbf{w} = \{w^i : w^i \geq 0\}_{i=1}^{M} \tag{2.3}$$

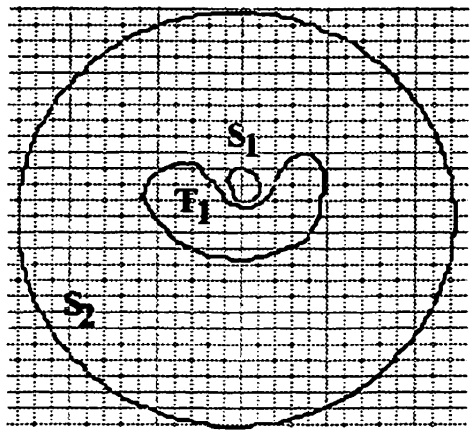

Figure 2.2. *Structure labelling. The internal structures, the tumour and the patient surface are outlined manually or semiautomatically on consecutive images obtained by computed tomography. The points of the dose calculation grid are labelled with respect to the structures S_a they belong to. Two ($a = 1, 2$) structures (S_1, S_2) and a tumour target T_1 ($b = 1$) are labelled in this particular example.*

2.2.2 Cost functions

Quadratic objective. When the goal of the optimization is specified in terms of the desired dose distribution, a quadratic objective is widely used. Let D_{S_a} be the tolerance dose of structure S_a (Fig. 2.2), where a is an index that enumerates the various anatomical structures. Let D_{T_b} be the dose to be delivered uniformly to the target T_b (Fig. 2.2), where b is an index that enumerates the target volumes to be treated. Let the total number of healthy structures be n_S and let the total number of targets to be treated be n_T. Then, the objective function to be minimized can be written as :

$$S(\mathbf{w}) = \beta \sum_{b=1}^{n_T} \left\langle (D^p(\mathbf{w}) - D_{T_b})^2 \right\rangle_{T_b} + \sum_{a=1}^{n_S} m_a \left\langle R_p (D^p(\mathbf{w}) - D_{S_a})^2 \right\rangle_{S_a} \tag{2.4}$$

$$R_p = \begin{cases} 1 & \text{if} \quad D^p(\mathbf{w}) \geq D_{S_a} \\ 0 & \text{if} \quad D^p(\mathbf{w}) < D_{S_a} \end{cases}.$$

The averages are calculated over the points of the corresponding structures. Different structures can be given different weights m_a and the overall weighting of healthy tissues versus the tumour targets is accounted for by the parameter β.

Probability of uncomplicated local tumor control. A relevant biological objective is to eradicate the tumour under nonuniform dose delivery without causing severe damage to healthy normal tissues. This goal can be incorporated by various biological models into the probability of uncomplicated local tumor control P_+. Elaborate discussion on P_+ and a underlying biological model is given by Ågren et al.[30] and Källman et al.[23, 31] In what follows we only present the corresponding analytical expressions used in the optimization.

In the model under consideration, for several target volumes, the probability of tumour control P_T is given by Poisson statistics as[31] :

$$\mathsf{P}_T(\mathbf{w}) = \prod_{b=1}^{n_T} \prod_{j=1}^{n_b} 2^{-e^{e\gamma_b\left(1-\frac{D_{j,b}(\mathbf{w})}{D_{50,b}}\right)-k_b \ln n_b}}, \tag{2.5}$$

where $D_{j,b}(\mathbf{w})$ is the dose in voxel j in target volume T_b for the current pencil beam weights $\mathbf{w}$, $D_{50,b}$ is the 50% response dose, γ_b is the maximum value of the normalized dose response gradient and e is the base of the natural logarithm. The total number of voxels in target volume T_b is n_b. The scale factor k_b accounts for the often reduced volume dependence of a heterogeneously growing tumor and has a value of unity for homogeneous tumors. The probability of injury P_I is given by

$$\mathsf{P}_I(\mathbf{w}) = 1 - \prod_{a=1}^{n_0} \left(1 - \left[1 - \prod_{j=1}^{n_a} \left[1 - \left(2^{-e^{e\gamma_a\left(1-\frac{D_{j,a}(\mathbf{w})}{D_{50,a}}\right)}} \right)^{s_a} \right]^{\frac{1}{n_a}} \right]^{\frac{1}{s_a}} \right) \tag{2.6}$$

where s_a is the relative seriality[23] of the structure S_a. The probability of uncomplicated local tumour control $\mathbf{P}_+(\mathbf{w})$ can be expressed by $\mathbf{P}_T(\mathbf{w})$ and $\mathbf{P}_I(\mathbf{w})$ as

$$\mathbf{P}_+(\mathbf{w}) = \mathbf{P}_T(\mathbf{w}) - \mathbf{P}_I(\mathbf{w}) + \delta(1 - \mathbf{P}_T(\mathbf{w}))\mathbf{P}_I(\mathbf{w}) \tag{2.7}$$

where δ is the fraction of patients with statistically independent response.[30]

Given the above, a suitable cost function to be minimized is the probability $\mathbf{P}_-(\mathbf{w})$ of obtaining severe complications and/or a recurrence :

$$\mathbf{P}_-(\mathbf{w}) = 1 - \mathbf{P}_+(\mathbf{w}) \; . \tag{2.8}$$

2.3 Gradient Optimization Algorithms

2.3.1 Optimal Algorithm Selection

Any gradient algorithm for unconstrained minimization of a smooth cost function $C(\mathbf{w})$ of the pencil beam weights $\mathbf{w}$ can be presented by the following model algorithm.[32]

Algorithm U (*Model algorithm for n-dimensional unconstrained optimization*). Let $\mathbf{w}_k$ be the current estimate of the minimum point $\mathbf{w}^*$.

U1 [Test for convergence]. If the conditions for convergence are satisfied, the algorithm terminates with $\mathbf{w}_k$ as the solution.

U2 [Compute a search direction]. Compute a non-zero n vector $\mathbf{d}_k$, the direction of search.

U3 [Compute a step length]. Compute a positive scalar α_k such that $C(\mathbf{w}_k + \alpha_k \mathbf{d}_k) < C(\mathbf{w}_k)$

U4 [Update the estimate of the minimum]. Set $\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{d}_k$ and go back to step U1.

The various gradient algorithms for unconstrained optimization differ mostly in the way the descent direction $\mathbf{d}_k$ is selected, which itself determines the rate of convergence

Table 2.1 *Characteristics of gradient unconstrained optimization algorithms. The number of the independent variables (pencil beam weights) is M. Operations relate to the number of floating point operations to be performed for the calculation of a descent direction $\mathbf{d}_k$ once the gradient of the cost function is determined. The exact numbers vary slightly with the specific implementation.*

	Gradient Methods		
	Steepest descent	Conjugate gradients	Quasi-newton
Storage	M	few M	M^2
Operations	M	few M	few M^2
Convergence	linear	linear-superlinear	superlinear

towards the minimum (the minimizer) $\mathbf{w}^*$ of $C(\mathbf{w})$. Better selection of $\mathbf{d}_k$ is associated with increasing overhead computations and memory storage (Table 2.1). In a typical problem of intensity modulation the number of the pencil beams is in the order of 10^4 ($M \sim 10^4$) and the number of the calculation points is in the order of 10^5 ($N \sim 10^5$). Given the dose calculation equation $\mathbf{D} = \hat{\mathbf{H}}\mathbf{w}$, the number of floating point operations needed for a single evaluation of the dose distribution is $2MN$. Therefore, a QN method requires a number of overhead calculations (Table 2.1) comparable to that needed for the evaluation of the dose distribution. The large amount of storage required for the QN method (few hundred Mbytes) also hampers its utility for optimization of intensity modulated beams. On the other hand, the CG are computationally comparable to the QN methods[33] but at the expense of a negligible increase in the overhead calculations and memory (Table 2.1). Therefore, a CG implementation for inverse treatment planning can be a viable alternative to the CSD method.

It must be emphasized that the above considerations are strictly valid for unconstrained optimization. For the case of constraint optimization, the properties of the

various methods associated with the overhead calculations and the storage requirements remain in the same proportionality, but their convergence properties are affected by the nature of the constraints and the objective (cost) function. For this reason, an active set implementation of the non-negativity constraints in the CG is proposed below and its performance is compared to that of a CSD implementation.

2.3.2 Constrained Steepest Descent (CSD) method

In an unconstrained steepest descent method the descent direction $\mathbf{d}_k$ (Step U2 of the model algorithm) is selected to be the opposite to that of the local gradient of the cost function, that is $\mathbf{d}_k = -\mathbf{g}_k = -\nabla_{\mathbf{w}} C(\mathbf{w})|_{\mathbf{w}=\mathbf{w}_k}$. The step length (Step U3 of the model algorithm) α_k can be heuristically selected or computed by minimizing $C(\mathbf{w})$ along the ray $\mathbf{w}_k + \alpha_k \mathbf{d}_k$.

An inclusion of the positivity constraints $w_k^i = (\mathbf{w}_k)^i \geq 0$ is done by the following modifications. The descent direction is calculated as:

$$\begin{aligned} d_k^i &= (\mathbf{d}_k)^i = 0 \quad \text{if} \quad w_k^i = 0 \quad \text{and} \quad -(\mathbf{g}_k)^i \leq 0 \\ d_k^i &= (\mathbf{d}_k)^i = -(\mathbf{g}_k)^i \quad \text{otherwise.} \end{aligned} \tag{2.9}$$

The step length α_k is determined by minimizing $C(\mathbf{w})$ along the ray $\mathbf{w}_k + \alpha_k \mathbf{d}_k$. However, whenever a constraint ($w_k^i = (\mathbf{w}_k)^i \leq 0$) is encountered, w_k^i is set to zero and the minimization continues along a new "bent" vector with components $max\left[0, w_k^i + \alpha_k d_k^i\right]$. This implementation is a generalization of the widely used "step and truncate" approach to the CSD.[22–26] A pseudo code statement of the algorithm is given in Fig. 2.3 and an extensive discussion on the properties of this algorithm is provided in the original work by McCormick.[34]

Algorithm CSD

(*Constrained steepest descent minimization of a smooth function* $C(\mathbf{w})$.)

Let $\mathbf{w}_k$ be the current estimate of the minimum point $\mathbf{w}^*$ and
$\mathbf{g}_k = \nabla_{\mathbf{w}} C(\mathbf{w})|_{\mathbf{w}=\mathbf{w}_k}$ be the gradient of the cost function.
WHILE the conditions for convergence are not satisfied [*Test* **(T1)** *for convergence*]
DO

BEGIN kth iteration
IF $(w_k^i = 0$ and $-(\mathbf{g}_k)^i \leq 0)$ THEN $d_k^i = (\mathbf{d}_k)^i = 0$
ELSE $d_k^i = (\mathbf{d}_k)^i = -(\mathbf{g}_k)^i$ [*Compute a feasible search direction* $\mathbf{d}_k$.];
BEGIN [*Compute a step length by a line minimization.*]
Compute a positive scalar α_k by minimizing $C(\mathbf{w})$ along the ray $\mathbf{w}_k + \alpha_k \mathbf{d}_k$. Whenever a constraint $(w_k^i = (\mathbf{w}_k)^i \leq 0)$ is encountered, set w_k^i to zero and continue the minimization continues along a new "bent" vector with components $max\left[0, w_k^i + \alpha_k d_k^i\right]$.
END
$(\mathbf{w}_{k+1})^i = max\left[0, w_k^i + \alpha_k d_k^i\right]$ [*Update the iterate*];
END

Figure 2.3. *CSD pseudo code.*

2.3.3 Active Set Method with Conjugate Gradient Subspace Minimization (AS)

We have introduced an active set method for inverse treatment planning which accounts for the simple bounds on the independent variables ($w^i \geq 0$). The algorithm is based on the following notion. Let I (active set) be the set of the indexes of these components w^{*i} of the solution vector $\mathbf{w}^*$ that are to be kept at the boundaries. For the variables w^{*i} that do not belong to the active set ($i \notin I$) the corresponding components of the function gradient are equal to zero ($\mathbf{g}^i(\mathbf{w}^*) = 0$). On the other hand for the variables w^{*i} in the active set ($i \in I$) the corresponding components of the function gradient have to be positive ($\mathbf{g}^i(\mathbf{w}^*) > 0$). Suppose that at the solution $\mathbf{w}^*$ a variable w^{*i} in the active set ($i \in I$) existed such that $\mathbf{g}^i(\mathbf{w}^*) < 0$. In this case a CSD iteration (Sec. (2.3.2)) would create a new feasible point with a smaller value of the objective function by removing the variable w^{*i} from its bound. However, that contradicts the fact that $\mathbf{w}^*$ is a constrained minimum. Therefore the assumption of $\mathbf{g}^i(\mathbf{w}^*)$ being smaller than zero for variables in the active set is incorrect.

If I were known *a priori*, then one could fix the corresponding variables w^{*i}, $i \in I$ to their boundaries and perform unconstrained optimization in the subspace of the remaining variables by using, for instance, a CG method. However, since the set I is not known, at each iteration step, the active set methods aim at developing a prediction I_k of the correct active set I. Such a prediction is based on the properties of the active set I at the solution. However, these properties are tested at the current iteration point and therefore the prediction I_k of the correct active set could be wrong. Hence, active set methods also include a procedure for testing whether the prediction is correct and altering it if not. In our implementation, after the update of the active set, a descent direction is calculated as in the CSD. An essential feature of the active set methods is that all iterates are feasible. A statement of the algorithm is given below along with a pseudo code outline in Fig. 2.4.

Algorithm AS (*Algorithm for n-dimensional constrained optimization subject to positivity constraints*). Let $\mathbf{w}_k$ be the current estimate of the minimum point $\mathbf{w}^*$ and $\mathbf{g}_k = \nabla_{\mathbf{w}} C(\mathbf{w})|_{\mathbf{w}=\mathbf{w}_k}$ be the gradient of the cost function.

A1 [Test for convergence (**T1**)]. If the conditions for convergence are satisfied, the algorithm terminates with $\mathbf{w}_k$ as the solution.

A2 [Update the active set I_k]. Compute I_k as the set of the indexes j such that $w_k^j = 0$ and $-g_k^i = -(\mathbf{g}_k)^i \leq 0$.

A3 [Compute a feasible search direction]. Compute a descent direction $\mathbf{d}_k$ such that $d_k^i = 0$ if $i \in I_k$ and $d_k^i = -g_k^i$ otherwise.

A4 [Compute a step length] Compute $\bar{\alpha}$, the maximum non-negative step along $\mathbf{d}_k$. Compute α_k minimizing along $\mathbf{d}_k$ or its "bent" version if constraints are encountered (see Sec. 2.3.2).

A5 [Update the iterate]. Set $w_{k+1}^i = max\left[0, w_k^i + \alpha_k d_k^i\right]$.

If $\alpha_k < \bar{\alpha}$ (feasible step) go to step A6 otherwise go to step A1.

A6 [Conjugate gradient iteration]

1. [Update the iterate] Set $w_{k+1}^i = w_k^i + \alpha_k d_k^i$.
2. [Decide which logic to perform (**T2**)] If minimization in the current subspace (indexes i not in I_k) is to be terminated, go to Step A1. Otherwise, perform a CG update of the descent direction $\mathbf{d}_k$ in the current subspace and go to Step A4.

The CG update of the descent direction $\mathbf{d}_k$ is different from $-\mathbf{g}_k$, used by the CSD. It is determined by the last descent direction $\mathbf{d}_{k-1}$, the last function gradient $\mathbf{g}_{k-1}$ and the local downhill gradient $\mathbf{g}_k$ by[27]

$$\mathbf{d}_k = -\mathbf{g}_k + \frac{(\mathbf{g}_k - \mathbf{g}_{k-1})\cdot\mathbf{g}_k}{\mathbf{g}_{k-1}\cdot\mathbf{g}_{k-1}} \mathbf{d}_{k-1} \,. \tag{2.10}$$

The steepest descent direction $-\mathbf{g}$ is used as a first member of a CG sequence. For the case of a quadratic function, a CG update of the descent direction guarantees that

Algorithm AS (*Active set algorithm for constrained minimization subject to nonnegativity constraints*). Let $\mathbf{w}_k$ be the current estimate of the minimum point $\mathbf{w}^*$ and $\mathbf{g}_k = \nabla_{\mathbf{w}} C(\mathbf{w})|_{\mathbf{w}=\mathbf{w}_k}$ be the gradient of the cost function.

WHILE the conditions for convergence are not satisfied [*Test* **(T1)** *for convergence.*]
DO

BEGIN kth iteration
Update the active set I_k as the set of the indexes j such that:
$w_k^j = 0$ and $-g_k^i = -(\mathbf{g}_k)^i \leq 0$;
IF $i \in I_k$ THEN $d_k^i = 0$
ELSE $d_k^i = -g_k^i$ [*Compute a feasible search direction* $\mathbf{d}_k$];
SUBSPACE_MINIMIZATION = TRUE;
DO

BEGIN
Compute $\bar{\alpha}$, the maximum non-negative step along $\mathbf{d}_k$;
BEGIN [*Compute a step length by a line minimization.*]
Compute a positive scalar α_k by minimizing $C(\mathbf{w})$ along the ray $\mathbf{w}_k + \alpha_k \mathbf{d}_k$. Whenever a constraint $(w_k^i = (\mathbf{w}_k)^i \leq 0)$ is encountered, set w_k^i to zero and continue the minimization continues along a new "bent" vector with components $max[0, w_k^i + \alpha_k d_k^i]$.
END
$(\mathbf{w_{k+1}})^i = max[0, w_k^i + \alpha_k d_k^i]$ [*Update the iterate*] ;
IF the step was feasible $(\alpha_k < \bar{\alpha})$ THEN
IF minimization in the current subspace (indexes i not in I_k) is to be terminated [test **(T2)**]
THEN SUBSPACE_MINIMIZATION = FALSE;
ELSE perform a CG update of the descent direction $\mathbf{d}_k$ in the current subspace;
ELSE SUBSPACE_MINIMIZATION = FALSE;
END

UNTIL SUBSPACE_MINIMIZATION is TRUE;
END

Figure 2.4. *AS pseudo code*

the minimization along $\mathbf{d}_k$ results in an iteration estimate $\mathbf{w}_k$ which also minimizes the cost function with respect to the previous direction $\mathbf{d}_{k-1}$. That is not true for any function $C(\mathbf{w})$. Nevertheless, the rationale for using the CG is the fact that $C(\mathbf{w})$ can be approximated by a guadratic form in the vicinity of a minimizer $\mathbf{w}^*$.

The change in the objective function from iteration to iteration is used for the logic test **T2** and the termination test **T1**. They both indicate termination if

$$|C(\mathbf{w}_{k-1}) - C(\mathbf{w}_k)| < \tau(1 + |C(\mathbf{w}_k)|) \tag{2.11}$$

where τ is a user-specified parameter. The merit of the test given by Eq. (2.11) is that it is not unduly stringent even when $|C(\mathbf{w}_k)|$ is considerably smaller than unity. Simultaneously for large absolute values of the cost function the above test [Eq. (2.11)] turns into the widely used termination test

$$|C(\mathbf{w}_{k-1}) - C(\mathbf{w}_k)| < \tau|C(\mathbf{w}_k)| \tag{2.12}$$

that examines the relative change in the objective function from iteration to iteration. In order to insure that the correct active set is selected at the solution, **T1** also requires that the last descent direction be a CSD one and a feasible step be performed along that direction. The termination criterion (**T1**) is used for the CSD method as well. The use of a single termination criterion based only on the change of the objective function cannot rule out the possibility of termination at a point which is not a local minimum. On the other hand, it does not involve any additional overhead calculations which become considerable when, as is the case in inverse treatment planning, the number of variables is large.

The prediction I_k of the correct active set I is updated under two circumstances: either the conjugate gradient minimization in the current variable subspace (indexes i not in I_k) has been accomplished (test **T2**) or a constraint in the current subspace has been

encountered. In the first case, after the subspace minimization, a update of the active set is needed since there may be variables w^i in the current active set which are no longer bound (active) by the inequality $\mathbf{g}^i(\mathbf{w}) > 0$ at the current point. In the second case, the incorrectness of the active set is conspicuous and a new prediction is to be made.

2.3.4 Performance Evaluation

The performance of a minimization algorithm depends on the cost function, on the selected performance measures, on the initial search point, as well as, on the termination criteria. For this reason it is important to identify a relevant starting point and termination criteria, as well as, performance measures for the evaluation of the optimization routines. We compared the performance of the CSD and the CG by the residual value of the cost function at termination, the total number of iterations (dose calculations) and the number of main iterations. A main iteration comprises a selection of the descent direction and a single line minimization. Both the CSD and the AS need a single dose calculation and a single evaluation of the function gradient for the selection of the descent direction and a few dose calculations per line minimization. At each main iteration, the AS method requires additional floating point operations for the evaluation of the maximum feasible step $\bar{\alpha}$ and the CG descent direction. These amount to few M and present a negligible calculation overhead (Sec. 2.3.1). The number of main iterations is an indicator of the efficiency of the CSD and the CG in the selection of the descent direction. The number of main iterations (gradient calculations) is a somewhat more objective measure than the total number of iterations since the latter depends on the parameters of the routine that selects the step along the current descent direction. For the ideal case, if a single step were performed to the line minimum, the total number of iterations would be equal to the number of main iterations.

The value of the objective function at termination is of primary concern when

$P_{-}(\mathbf{w})$ is used for the optimization for it reflects directly the utility of the achieved dose distribution. However, when the quadratic objective $S(\mathbf{w})$ is employed, a smaller residual value may not readily result in significantly better dose distribution from a clinical point of view. For this reason, we also examined the dose distributions obtained by the AS and the CSD.

2.4 Method

2.4.1 Target Volumes and Organs at Risk

Three different cases were used for the comparison of the optimization algorithms. Case A is a simulated concave target with an organ at risk in a cylindrical phantom (Fig. 2.5). Case B is a solid prostate cancer with organs at risk being bladder, rectum and normal tissue stroma (Fig. 2.6). Case C represents a target volume located in the head and neck region with organs at risk being spinal cord and normal tissue stroma (Fig. 2.7). The relative tumor density is assumed 90% in the target volume. Due to the uncertainty in the values of the radiobiological parameters we used reasonable estimates based on published data.[23, 25, 35–37] The values of the radiobiological parameters and those of the dose levels for the physical objective are given in Table 2.2. The fraction δ of patients with statistically independent tumour and normal responses has been set to zero. The different sets of weights assigned to the various anatomical structures for the case of the quadratic objective are given in Table 2.3. For all cases, homogeneous tumors are assumed.

2.4.2 Irradiation Geometries and Pencil Beams

In our particular implementation, the Differential-Scatter-Air-Ratio (DSAR) dose calculation model[38] of a 3D treatment planning system (GRATIS, Sherouse Inc.) was

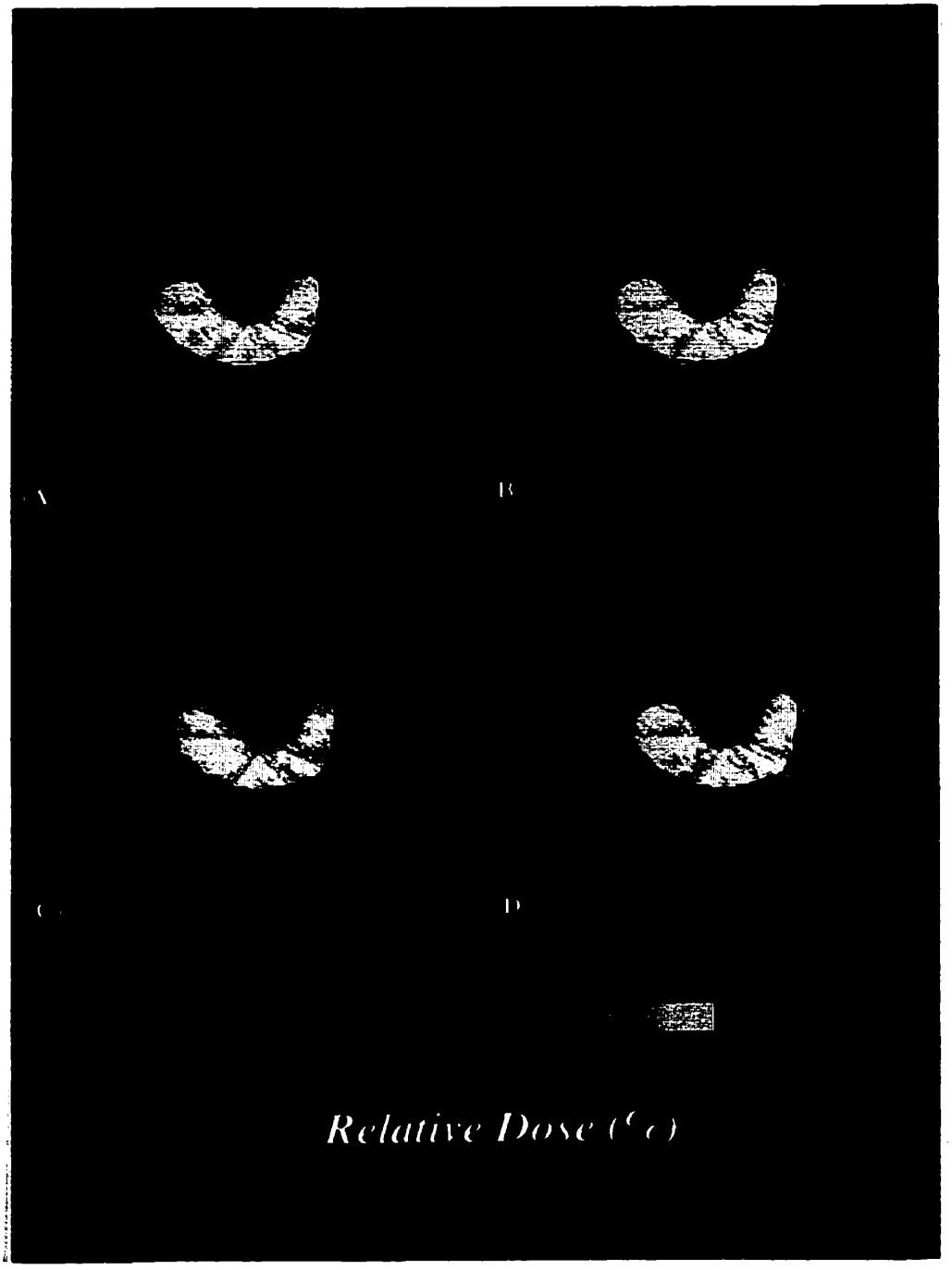

Figure 2.5 *Case A. Dose distributions obtained by the CSD and the AS. Two combinations of weights* $(\beta : m_1)$ *are used. (a) CSD,* $(1 : 1)$*; (b) AS,* $(1 : 1)$*; (c) CSD,* $(1 : 15)$*; (d) AS,* $(1 : 15)$.

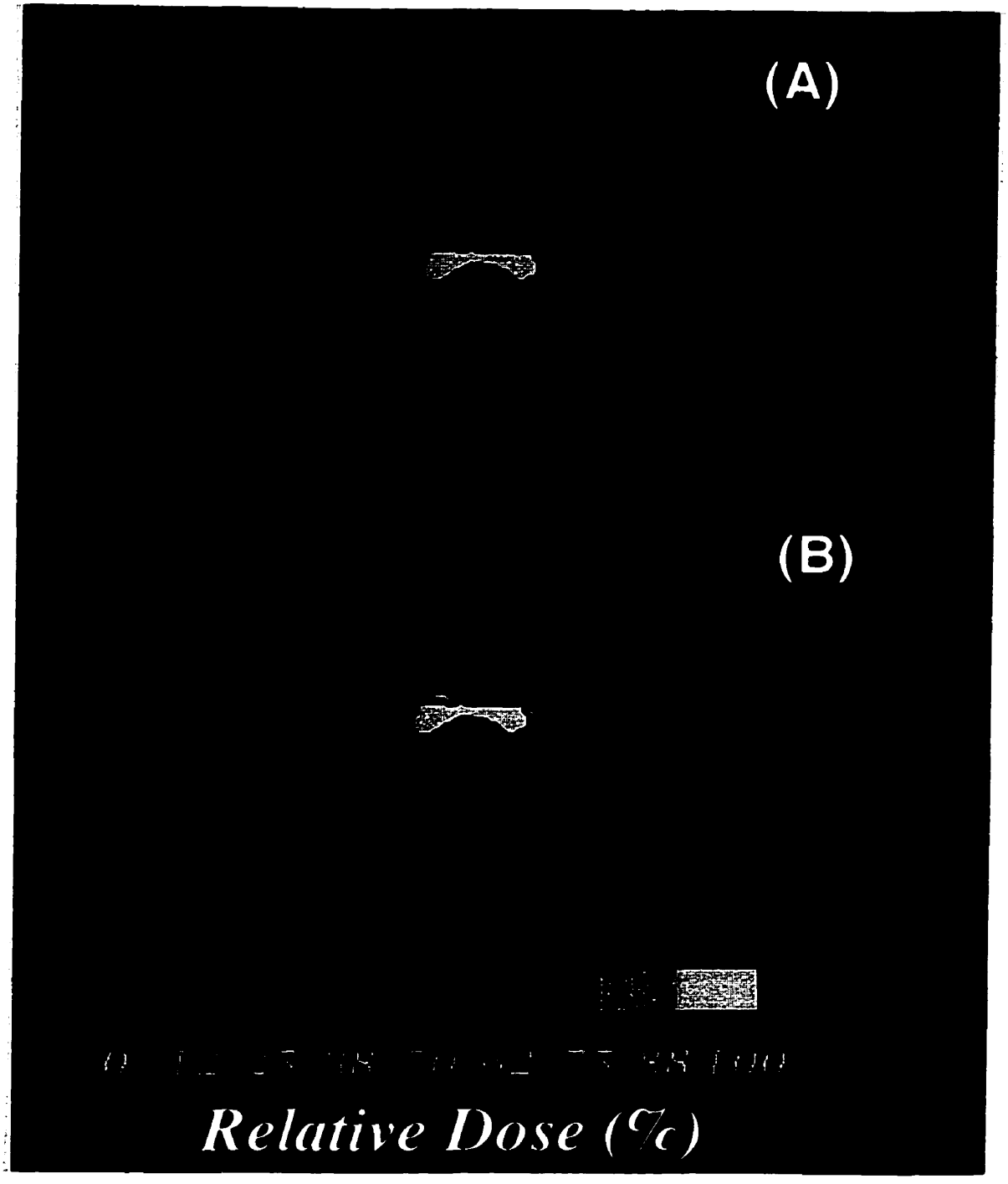

Figure 2.6 *Case B. (a) Dose distribution obtained by the CSD. (b) Dose distribution obtained by the AS. Weights of unity are assigned to the target and the rectum* ($(\beta : m_1) = (1 : 1)$).

Figure 2.7 *Case C. Tumor in the head and neck region. The spinal cord is the organ at risk. For the optimization of the quadratic objective a margin of 1 cm (not shown) was introduced around the target where dose levels were not specified.*

Table 2.2 *Radiobiological and dose parameters used in the optimization. For all cases, homogeneous tumors are assumed. The dose levels for the tumors correspond to those delivered to them with uniform beams and 100 MU delivered by each beam. The dose levels to the other structures are calculated as a percentage of the target dose.*

Cases	Dose (Gy) (%)	D_{50}(Gy)	γ	s
A. Simulated				
Tumor	4.7 (100.0)	...	...	...
1. Organ at risk	1.9 (40.4)	...	...	...
2. Normal tissue	2.8 (59.5)	...	...	...
B. Prostate cancer				
Tumor	5.5 (100.0)	60.0	4.00	...
1. Rectum	3.0 (54.5)	75.0	2.50	0.70
2. Bladder	3.0 (54.5)	80.0	3.00	1.30
3. Normal tissue	3.0 (54.5)	65.0	2.76	1.00
C. Head and neck				
Tumor	2.9 (100.0)	52.0	3.00	...
1. Spinal cord	1.45 (50.0)	60.0	1.78	1.00
2. Normal tissue	1.45 (50.0)	65.0	2.76	1.00

adapted for pencil beam dose calculation (5x5 mm^2 at source–to–isocenter distance of 100 cm). The DSAR model (Appendix A) takes into account scatter but does not consider inhomogeneities. However, by virtue of Eq. (2.3) the optimization procedure is implemented as a model-free routine. It does not explore any particular properties of the matrix $\widehat{H}$ and can be used with any other model for the calculation of the pencil beams. Since scatter is accounted for by the DSAR model, a margin of 1.5 cm is added around the projection of the target in each beam's eye-view. Photon pencil beams (18 MV) are

Table 2.3 *Weights assigned to the target and an organ at risk for the quadratic objective optimization. The other structures considered in the optimization are given weights of unity.*

Cases	Weights ($\beta : m_1$)					
A. Tumour : Organ at risk B. Tumour : Rectum	1:1	1:5	1:10	1:15	5:1	10:1

Table 2.4 *Irradiation parameters. Coplanar beams in the direction of increasing angles with equal angular separations are employed in all cases. Gantry angle is zero when the gantry is straight up. The angles increase for counterclockwise rotation of the gantry as viewed from the isocenter. The isocenter is always placed in the center of the target volume (in center-of-mass sense). Initial angles are : 37° for case A, 0° for case B, 0° for case C.*

Case	Number of beams	Number of pencil beams	Grid size	Voxel size (cm^2) (cm^3)
A	7	1491	124 x 124 x 1	0.25 x 0.25
B	9	1267	104 x 178 x 1	0.25 x 0.25
C	3	1503	32 x 37 x 22	0.5 x 0.5 x 0.5

precalculated before the optimization for the given beam setup and dose calculation grid. The pencil beam weights map onto Monitor Units (beam-on time) after multiplication by a factor of 100 MU/Gy. For the quadratic objective, no additional scaling of the pencil beam weights was done. Given the dose prescription levels, the weights varied between zero and few units (Gy) for both optimization algorithms. When the biological cost function is minimized, an additional common scaling factor is introduced such that

the uniform beams of unit weight deliver uniform dose to the tumor slightly higher than the 50% response dose.[23, 31] A full three-dimensional optimization was performed for the head and neck case, and two-dimensional optimization was performed for the simulated and the prostate cases by forming a slit irradiation.

2.4.3 Optimization Parameters

In the optimization of the quadratic objective, zero initial weights were used as a starting point to avoid unnecessary healthy tissue irradiation. For the case of uncomplicated local tumour control optimization, uniform unit beams were used as an initial guess since previously such a choice has been found satisfactory.[24, 25]

The value of the termination parameter τ was selected as follows. For the case of the quadratic objective optimization, we applied the CSD to few cases, varying τ in order to investigate the sensitivity of the resulting dose distributions with respect to it. Dose volume histograms were employed to evaluate the optimized plans. A dose volume histogram is a cumulative or a differential dose distribution calculated within a preselected volume. In this thesis, however, dose volume histogram stands exclusively for cumulative dose distributions. For a particular organ, a dose volume histogram allows one to evaluate the fractional volume of the organ that is irradiated at and above a certain dose level. (For example, in Fig. 2.8 (c), for $\tau = 0.001$ approximately 80 % of the spinal cord receives 30 % and more of the maximum dose.)

Based on our analysis of the dose-volume histograms (Fig. 2.8), we set τ to 0.001 for cases A and C, and τ to 0.01 for case B. In the radiobiologically based optimization, we set τ equal to 0.001, thus aiming at improvements of a tenth of a percent. Values of the termination parameter τ below 0.001 were found to lead to excessive computational times and therefore impractical.

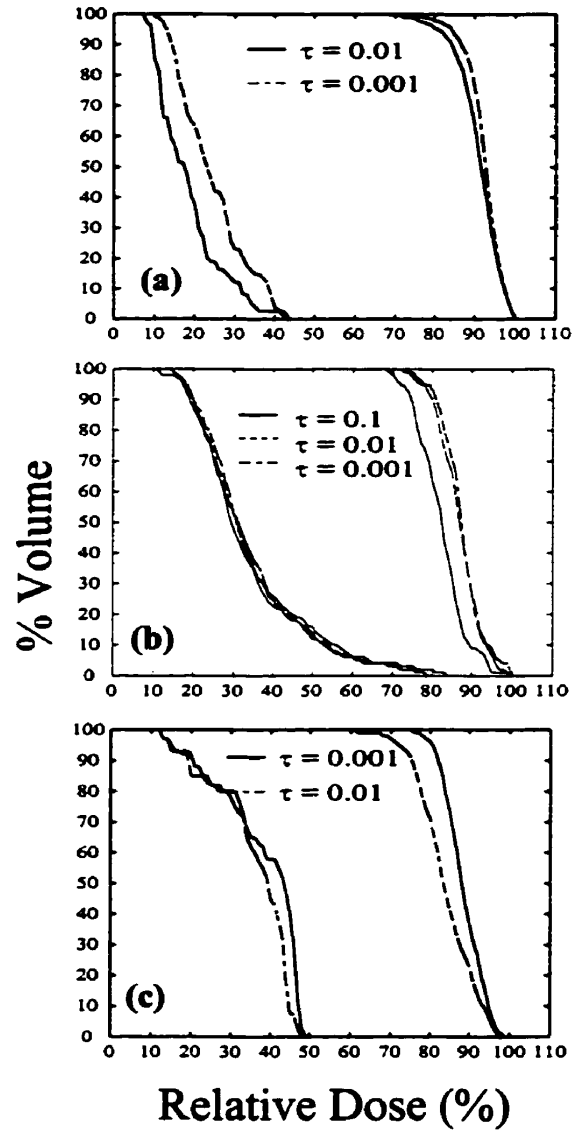

Figure 2.8 *Sensitivity of the inverse treatment planning solutions to the value of the termination criterion τ. All structures are assigned weights of unity for all cases. The dose distributions are normalized to their maxima. The CSD is used for the optimization. (a) Dose-volume histograms for the target and the organ at risk for case A. The resulting solutions are significantly different. (b) Dose-volume histograms for the target and the rectum for case B. The solution for $\tau = 0.01$ is almost identical to that for $\tau = 0.001$. (c) Dose-volume histograms for the target and the spinal cord for case C. The resulting solutions are significantly different.*

In our implementation of the minimization algorithms, an identical safeguarded line minimization routine with parabolic interpolation[39] is used by both the CSD and the AS. This routine combines a guaranteed, reliable *golden section search* in one dimension with a parabolic interpolation to yield an algorithm that converges rapidly if the cost function is well behaved, but is not much less efficient than the guaranteed method in the worst case.

2.5 Results and Discussion

2.5.1 Convergence Properties

Figures 2.9, 2.10 and 2.11 illustrate the convergence properties of the CSD and the AS when applied to the minimization of the square objective. For all cases, the AS exhibits a higher rate of convergence than the CSD. The AS also reaches a lower residual value of the cost function at termination than the CSD does.

Figure 2.12 examines the number of iterations and the number of main iterations required by the AS to achieve the same value of the cost function as the one obtained by the CSD at termination for treatment cases A and B. For case C, the CSD executes 185 iterations and 11 main ones to reach termination. For the same case, the AS requires 107 iterations and 7 main ones to achieve the same value of the cost function. These results and the ones shown in Fig. 2.12 confirm the expected fact that the superiority of the AS to the CSD is mainly due to the lower number of main iterations, which in turn, is due to the improved selection of the descent directions.

Figure 2.13 illustrates the convergence properties of the CSD and the AS when applied to the minimization of the probability $P_{-}(\mathbf{w})$ of obtaining severe complications and/or recurrence for cases B and C. The AS exhibits a higher rate of convergence than the CSD. The AS also reaches a lower residual value of the cost function at termination than does the CSD (Fig. 2.13(a)). The CSD executes 243 iterations with 15 main ones

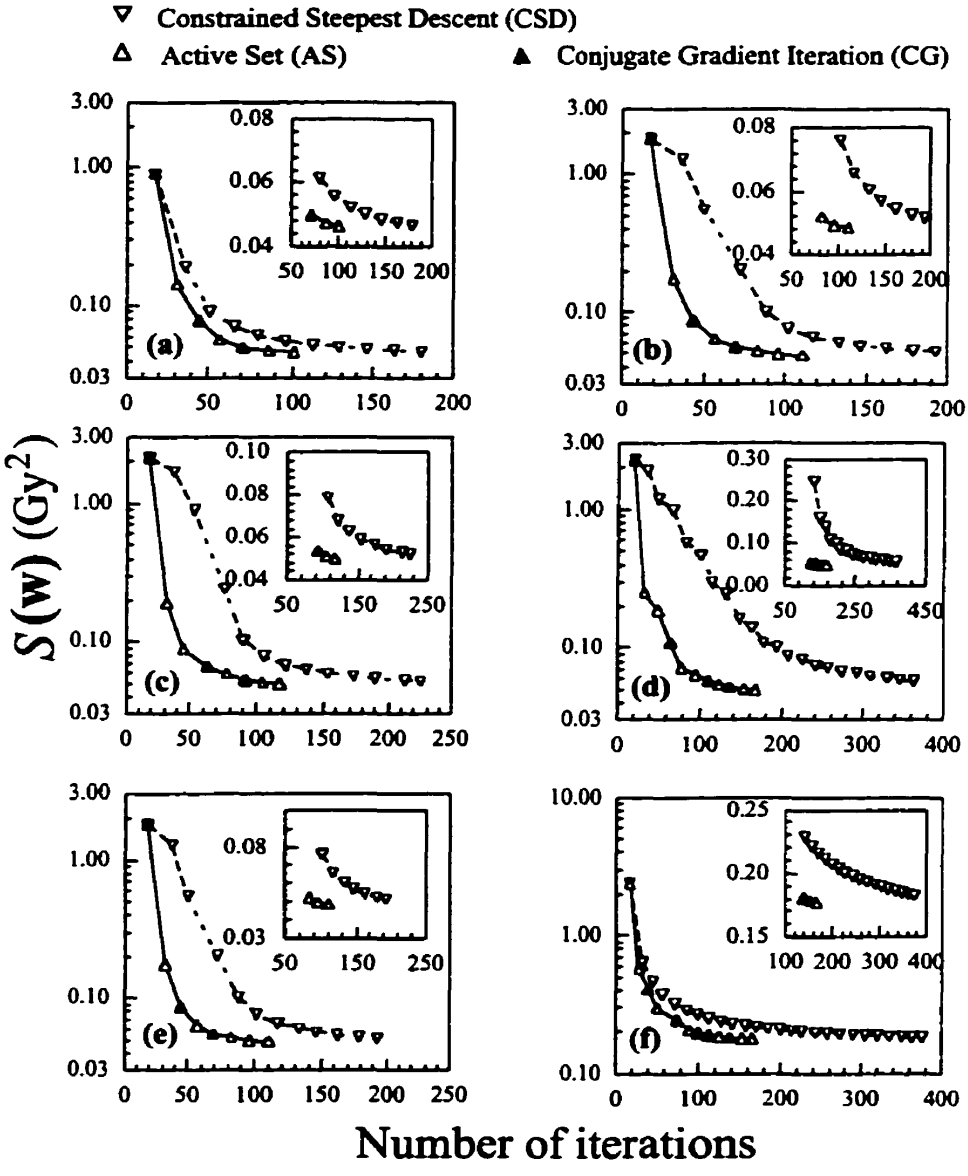

Figure 2.9. *Quadratic objective $S(\mathbf{w})$ versus number of iteration steps for case A. Lines are drawn to show the trend of the cost function values. Plots (a)-(f) correspond to the combinations of weights given in Table 2.3. All plots start with the second main iteration since the first one is common for both the CSD and the AS. The inserts identify the main iteration step at which the AS leads to a value of the cost function lower than the value obtained by the CSD at the termination.*

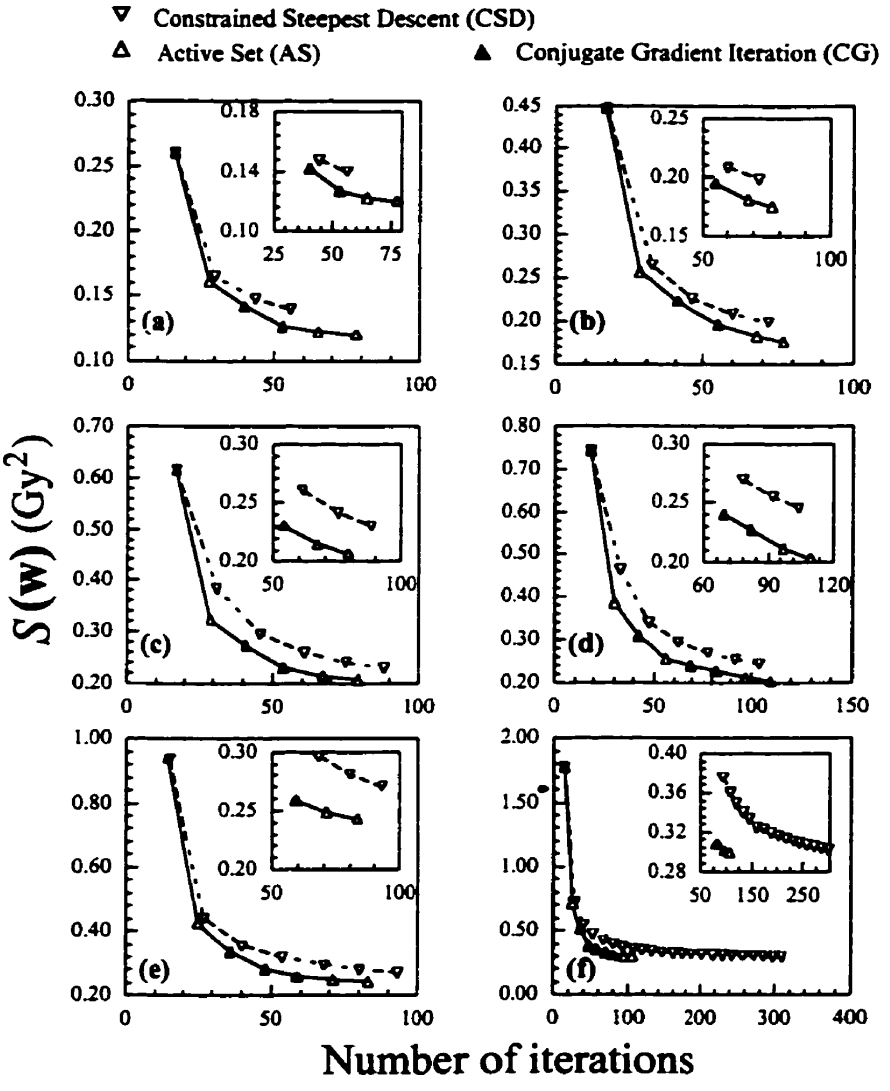

Figure 2.10 *Quadratic objective $S(\mathbf{w})$ versus number of iteration steps for case B. Lines are drawn to show the trend in value of the cost function. Plots (a)-(f) correspond to the combinations of weights given in Table 2.3. All plots start with the second main iteration since the first one is common for both the CSD and the AS. The inserts identify the main iteration step at which the AS leads to a value of the cost function lower than the value obtained by the CSD at the termination. For plot (f) the CSD iterations are terminated as soon as the residual cost is below the value obtained by the AS. For this case the number of iterations required by CSD is as large as three times the one required by the AS.*

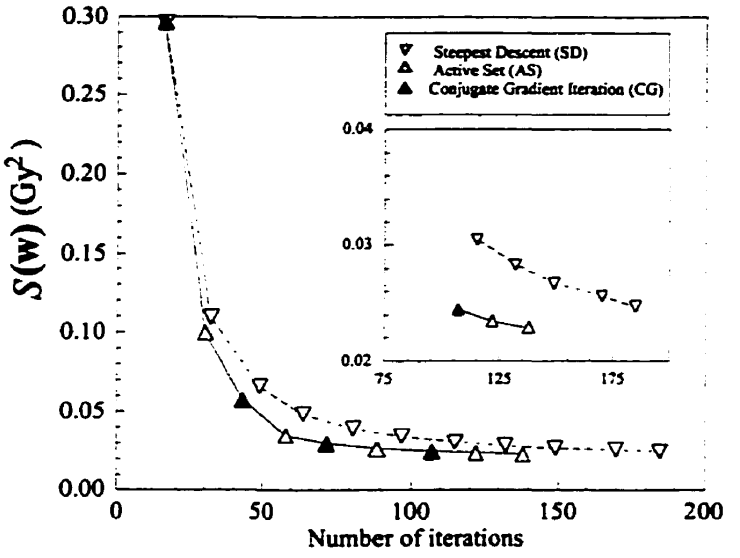

Figure 2.11 *Quadratic objective $S(\mathbf{w})$ versus number of iteration steps for case C. Lines are drawn to show the trend of the cost function values. The plot starts with the second main iteration since the first one is common for both the CSD and the AS. All anatomical structures are given weights of unity. The insert identifies the main iteration step at which the AS leads to a value of the cost function lower than the value obtained by the CSD at the termination.*

to reach termination. For the same case, the AS requires only 100 iterations with 6 main ones to achieve the same value of the cost function.

For the head and neck case the performance of both methods is almost identical, both in terms of their convergence properties and the residual value of the cost function (Fig. 2.13(b)). This suggests that even though a certain algorithm may have the potential of outperforming another, whether or not this occurs is ultimately determined by the cost function, which, for the case of treatment planning optimization of a biological objective, is defined by both the treatment geometry and the radiobiological parameters for the

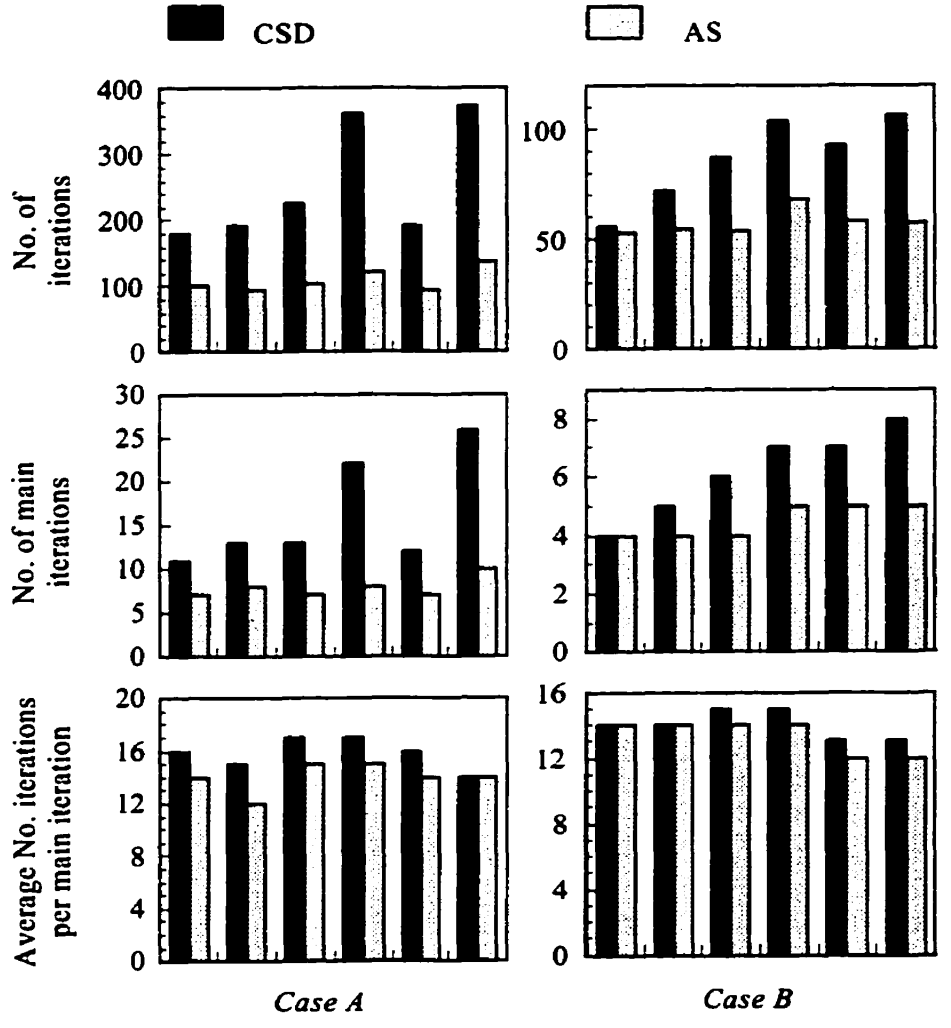

Figure 2.12 *Number of iterations and number of main iterations required by the AS to achieve the same value of the cost function as the one obtained by the CSD at termination for treatment cases A and B. The bar plots from left to right correspond to the combinations of weights given in Table 2.3. The average number of iterations per main iteration rounded to the nearest integer number is also given. In some cases, the last number gives some additional advantage to the AS. However, the superior performance of the AS compared to that of the CSD is mainly due to the lower number of main iterations.*

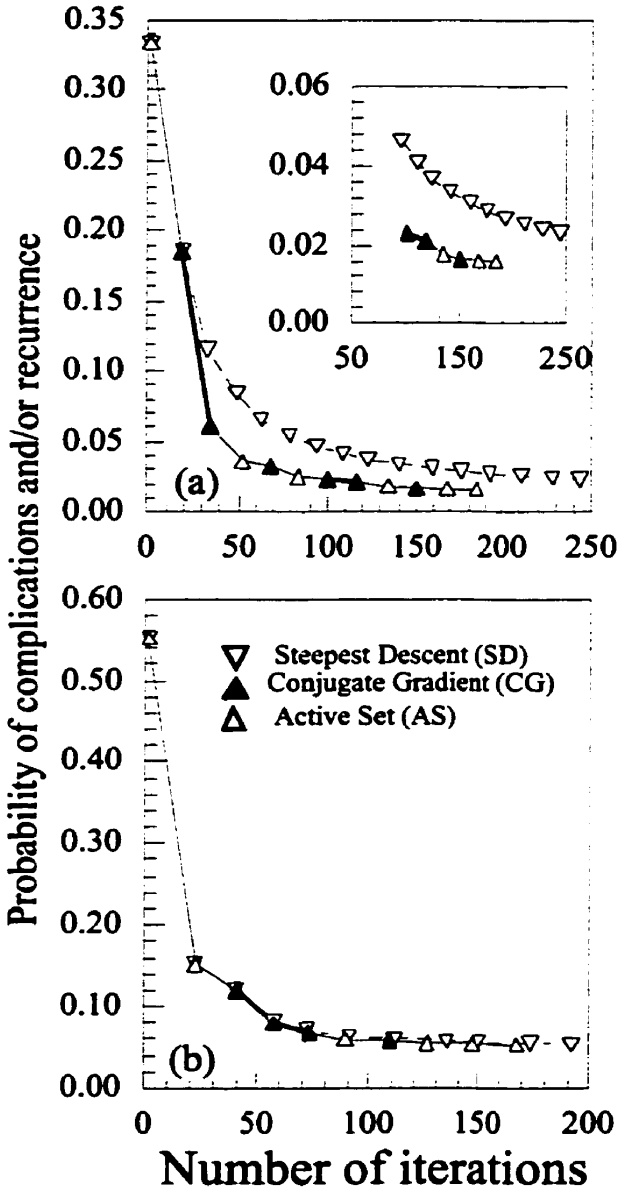

Figure 2.13. *(a) Probability of obtaining severe complications and/or a recurrence* $P_{-}(\mathbf{w})$ *versus number of iteration steps for case B. The insert identifies the main iteration step at which the AS leads to a value of the cost function lower than the value obtained by the CSD at the termination. (b) Probability of obtaining severe complications and/or a recurrence* $P_{-}(\mathbf{w})$ *versus number of iteration steps for case C. Lines are drawn to show the trend in the value of the cost function.*

various organs. Nevertheless, given the statement of the AS (Sec. 2.3.3), one may expect that in the worst scenario, the AS will perform, at least as well as, the CSD with respect to its convergence properties.

Figures 2.9, 2.10, 2.11, 2.13 illustrate that the AS iteration sequences vary significantly with respect to the treatment site. For instance, negative weights are more often encountered for case A (Fig. 2.9) then for case B (Fig. 2.10). For case A, negative weights are usually encountered after the first CG iteration in the current minimization subspace (Fig. 2.9(a-e)) which leads to a frequent update of the active set. However, even a single CG iteration in the current subspace allows significant improvement in the convergence of the iteration sequence. For case B, negative weights are usually encountered once at the beginning of the iteration sequences (Fig. 2.10(a-c, e, f)). After the corresponding active set update several CG iterations (Fig. 2.10(a-c, e, f)) lead to the minimum in the current subspace, which is also the minimum of the constrained problem (according to test **T1**) since the update of the active set and the consequent move down the steepest descent direction do not lead to significant changes in the cost function value.

We emphasize that when negative weights are not encountered the AS automatically becomes the unconstrained CG with the additional calculation of the maximum feasible step $\bar{\alpha}$. Therefore, when there are not any active constraints at the solution the AS is almost as efficient as the CG and produces the same minimum as the CG. When there are active constraints at the solution, the AS attempts to identify them and to proceed as the unconstrained CG in the subspace of the remaining variables. Given its statement, it always converges to a point that will be considered as a minimum by the CSD.[34] On the other hand, a point obtained by the use of the unconstrained CG throughout the minimization and truncation to zero of the negative weights at the end is not guaranteed to be a minimum of the constrained minimization problem. For a particular cost function,

a mathematical analysis of the properties of the objective is needed to evaluate the robustness of such an unconstrained approach.

The additional calculations required by the AS at each main iteration present a negligible computational overhead in comparison with the calculation time necessary for a single dose calculation. For case C, given 1503 pencil beams and 26048 calculation voxels, the AS needs 0.01 s per iteration to evaluate the maximum feasible stepsize and the CG descent direction. A single dose calculation performed as a matrix multiplication (Sec. 2.3.1) takes 15 s. However, in our implementation the pencil beams are stored on disk. Thus the dose calculation time is 75 s due to input-output operations. All the simulations are performed on a Sun SPARC 10 computer.

2.5.2 Dose Distributions

When a quadratic dose objective is optimized the utility of the resulting dose distributions depends not only on the performance of the optimization algorithms but also on the optimal selection of the dose levels and of the weight factors for the various structures. Figure 2.14(a) illustrates this point for treatment case A. The AS leads to a lower residual value at termination than does the CSD (Fig. 2.9(a)), but the AS dose distribution can be considered inferior to that of the CSD (Fig. 2.14(a), Fig. 2.5(a) and Fig. 2.5(b)). However, it cannot be argued that for a particular treatment case, the CSD tends to terminate at a point of a higher cost function value but of better clinical utility than the AS does. Figures 2.14(b), 2.14(c) and 2.5(c), 2.5(d) illustrate that even for the same treatment case and dose levels, depending on the assigned weights, the lower residual value obtained by the AS may reflect dose distributions which are similar or better than those obtained by the CSD.

For treatment case B, the dose distributions obtained by the CSD and the AS were very similar for all combinations of weights given to the anatomical structures despite the

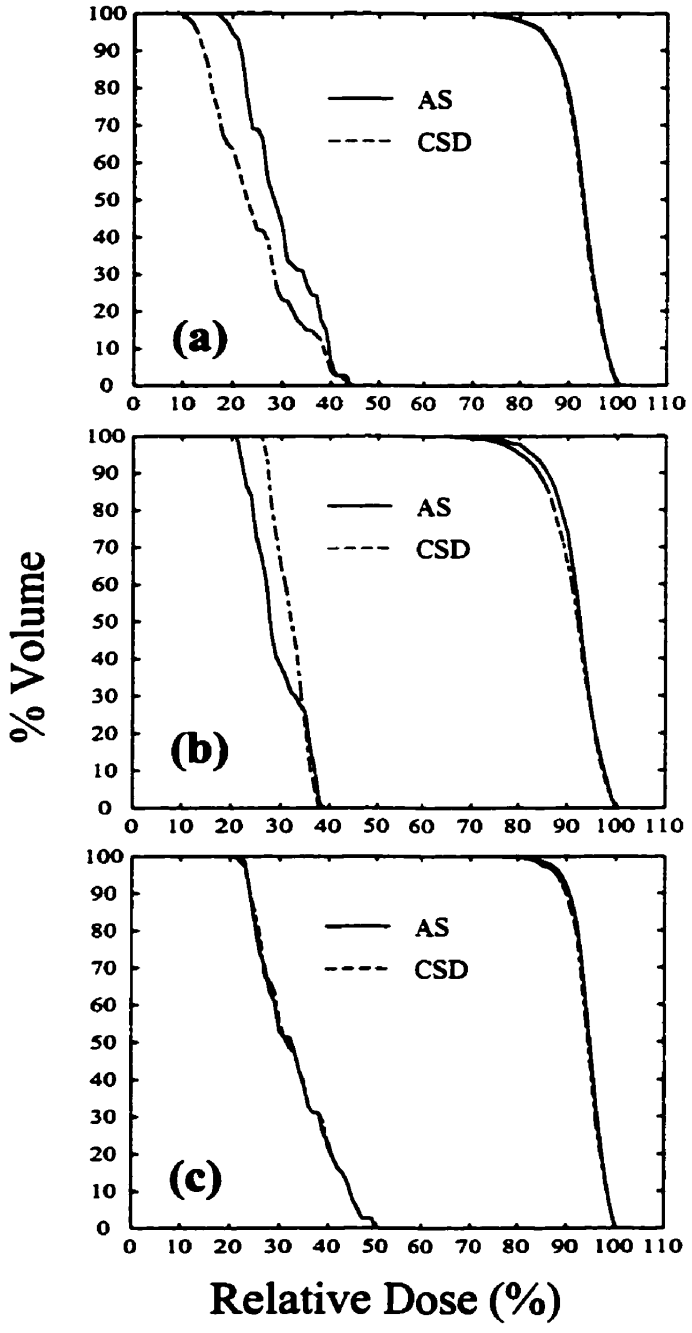

Figure 2.14. *Dose-volume histograms for the target and the organ at risk for case A and for different combinations of weights* $(\beta : m_1)$. *(a)* $(1 : 1)$; *(b)* $(1 : 15)$; *(c)* $(5 : 1)$.

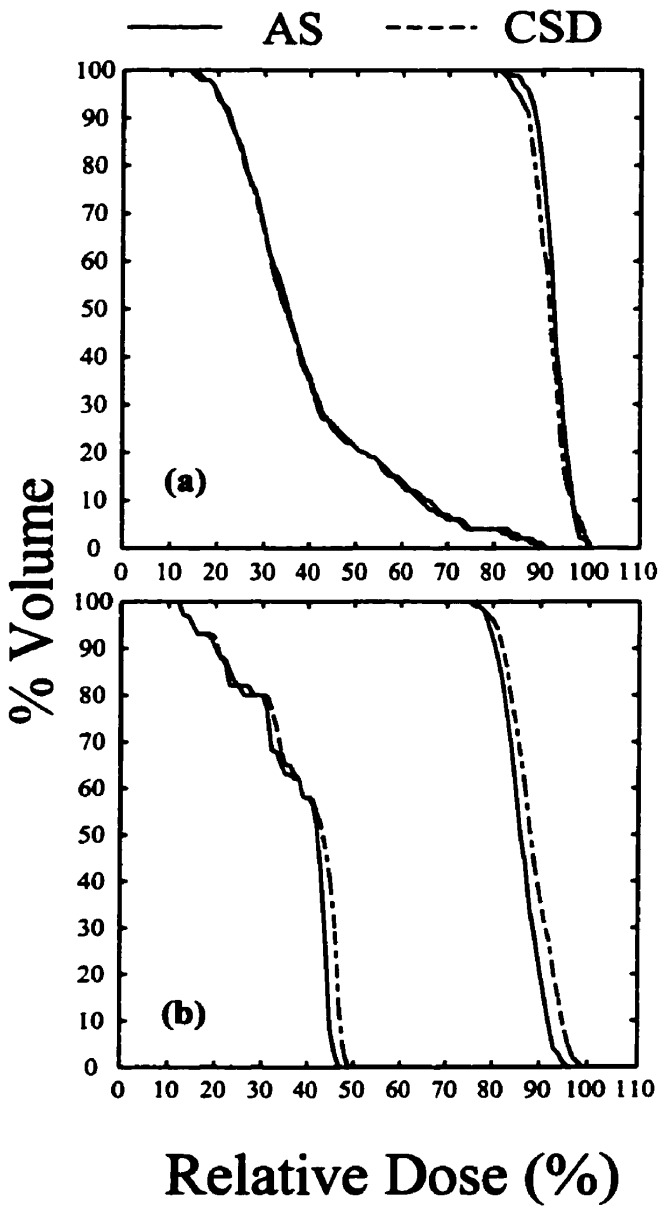

Figure 2.15 *(a) Dose-volume histograms for the target and the rectum for case B. (b) Dose-volume histograms for the target and the spinal cord for case C. Weights of unity are assigned to all anatomical structures.*

consistently lower residual value of the cost function obtained by the AS. Figures 2.15(a) and 2.6 illustrate this fact for a particular selection of the structure weights.

For treatment case C, the lower residual value of the cost function obtained by the AS reflects a better dose distribution both in terms of the sparing of the spinal cord and the tumor dose homogeneity (Fig. 2.15(b)).

To summarize, for a quadratic objective the AS always reaches a lower residual value at termination than the CSD. However, due to the nature of the quadratic objective, the dose distributions obtained by the AS may not always be better than those obtained by the CSD. To ensure a good correlation between a low residual value of the cost function and a useful dose distribution, the dose levels that are specified must be very close to the achievable ones.

2.6 Conclusions

The most time-consuming operation in the optimization of radiotherapy treatment plans by intensity modulated beams is the repeated dose calculation. An active set algorithm for inverse treatment planning is introduced as an alternative tool to the widely used constrained steepest descent method. The algorithm uses a conjugate gradient routine for subspace minimization in order to achieve better convergence at the expense of negligible amount of overhead calculations. By numerical simulations it has been demonstrated that (i) the active set algorithm performs at least as well as the constrained steepest descent in terms of the residual cost at termination for two different type of objective functions; and (ii) due to its better convergence properties, the active set method can decrease significantly the number of iterations necessary to reach a solution of the inverse treatment planning problem.

References

1 A. Brahme, A. J. Roos, and I. Lax. Solution of an integral equation in rotation therapy. *Phys. Med. Biol.*, 27:1221–1229, 1982.

2 A. M. Cormack. A problem in rotation therapy with x-rays. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:623–630, 1987.

3 A. M. Cormack and R. A. Cormack. A problem in rotation therapy with x-rays: Dose distributions with an axis of symmetry. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1921–1925, 1987.

4 N. H. Barth. An inverse problem in radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:425–431, 1990.

5 V. Tulovski, M Ringor, and L Papież. Optimization of rotational radiotherapy treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1205–1214, 1995.

6 Z. Chen, X. Wang, T. Bortfeld, R. Mohan, and L. Reinstein. The influence of scatter on the design of optimized intensity modulations. *Med. Phys.*, 22(11 Pt. 1):1727–1733, 1995.

7 A. Brahme. Optimization of stationary and moving beam therapy techniques. *Radioth. Oncol.*, 12:129–140, 1988.

8 A. L. Boyer, G. E. Desorby, and N. H. Wells. Potential and limitations of invariant kernel conformal therapy. *Med. Phys.*, 18(4):703–712, 1991.

9 T. W. Holmes, T. R. Mackie, and P. J. Reckwerdt. An iterative filtered backprojection inverse treatment planning algorithm for tomotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1215–1225, 1995.

10 T. W. Holmes and T. R. Mackie. A comparison of three inverse treatment planning algorithms. *Phys. Med. Biol.*, 39:91–106, 1994.

11 T. Bortfeld, J. Burkelbash, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied to conformation radiotherapy. *Phys. Med. Biol.*, 35:1423–1434, 1990.

12 L. Xing and G. T. Y. Chen. Iterative methods for inverse treatment planning. *Phys. Med. Biol.*, 41:2107–2133, 1996.

13 L. Xing, C. Pelizzari, F. T. Kuchir, and G. T. Y. Chen. Optimization of relative weights and wedge angles in treatment planning. *Med. Phys.*, 24(2):215–221, 1997.

14 Y. Censor, M. D. Altschuler, and W. D. Powlis. A computational solution of the inverse problem in radiation-therapy treatment planning. *Appl. Math. Comput.*, 25:57–87, 1988.

[15] S. Lee, P. S. Cho, R. Marks, and O. Seho. Conformal radiotherapy computation by the method of alternating projections onto convex sets. *Phys. Med. Biol.*, 42:1065–1086, 1997.

[16] A. Djordjevitch, D. J. Bonham, E. M. A. Hussein, J. W. Andrew, and M. E. Hale. Optimal design of radiation compensators. *Med. Phys.*, 17:397–404, 1990.

[17] S. M. Morrill, , I. I. Rosen, R. G. Lane, and J. A. Belli. The influence of dose constraint point placement in optimizaed radiation therapy treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 19(2):129–142, 1990.

[18] M. Langer, R. Brown, P. Kijewski, and H. Chul. The reliability of optimization under dose-volume limits. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:529–538, 1993.

[19] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[20] N. Metropolis, A. Rosenbluth, M. Rosenbluth, H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.

[21] S. Webb. Optimization by simulated annealing of three-dimensional conformal treatment palnning for radiation fields defined by multileaf collimator. II. Inclusion of two-dimensional modulation of the X-ray intensity. *Phys. Med. Biol.*, 37(8):1689–1704, 1992.

[22] B. K. Lind. Properties of an algorithm for solving the inverse problem in radiation therapy. *Inverse Problems*, 6:415–426, 1990.

[23] P. Källman, B. K. Lind, and A. Brahme. An algorithm for maximizing the probability of complication-free tumour control in radiation therapy. *Phys. Med. Biol.*, 37:871–890, 1992.

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

[25] A. Gustafsson, B. K. Lind, R. Svensson, and A. Brahme. Simultaneous optimization of dynamic multileaf collimation and scanning patterns or compensation filters using a generalized pencil beam algorithm. *Med. Phys.*, 22(7):1141–1156, 1995.

[26] X. Wang, R. Mohan, A Jackson, S. A. Leibel, Z. Fuks, and C. C. Ling. Optimization of intensity-modulated 3d conformal treatment plans based on biological indices. *Radiotherapy and Oncology*, 37:140–152, 1995.

[27] E. Polak. *Computational methods in optimization*. Academic Press, New York, 1971.

[28] D. H. Hristov and B. G. Fallone. Parametrization of the independent variables in treatment planning optimization. *Med. Phys.*, 23(6):1111–1112, 1996. (abstract).

[29] A. Niemierko. Non-linear congugate gradient algorithm for optimization of intensity modulated beams. *Med. Phys.*, 23:1172, 1996. (abstract).

[30] A. Ågren, A. Brahme, and I. Turesson. Optimization of uncomplicated control for head and neck tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 19:1077–85, 1990.

[31] P. Källman, A. Ågren, and A. Brahme. Tumour and normal tissue responses to fractionated non-uniform dose delivery. *Int. J. Radiat. Biol.*, 62(2):249–262, 1992.

[32] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*, page 99. Academic Press, Inc., Orlando, Florida 32887, 1981.

[33] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies in C*, page 426. Cambridge University Press, New York, 1988.

[34] G. P. McCormick. Anti-zig-zagging by bending. *Management Science*, 15:315–319, 1969.

[35] S. Söderstrom, A. Gustafsson, and A. Brahme. The clinical value of different treatment objectives and degrees of freedom in radiation therapy optimization. *Radiotherapy and Oncology*, 29:148–163, 1993.

[36] B. Emami, J. Lyman, A. Brown, L. Coia, M. Goitein, J. E. Munzenrider, B. Shank, L. J. Solin, and A. M. Wesson. Tolerance of normal tissue to therapeutic irradiation. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:109–122, 1991.

[37] P. Okunieff, D. Morgan, A. Niemierko, and H. Suit. Radiation dose-response of human tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1227–1237, 1995.

[38] J. R. Cunningham. Scatter-air ratios. *Phys. Med. Biol.*, 7:45–51, 1972.

[39] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies in C*, pages 299–300. Cambridge University Press, New York, 1988.

CHAPTER 3 A Continuous Penalty Function Method for Inverse Treatment Planning

3.1 Introduction

In Chapter 2 we introduced an active set algorithm as a tool for solving the large-scale inverse problem in conformal therapy when a smooth differentiable objective function of the pencil beam weights is used. The clinical utility, however, of the dose distributions produced by any minimization algorithm depends critically on the proper modelling of the treatment objectives. In this chapter, some of our research effort in this direction is presented.

The goal of inverse treatment planning is to generate treatment plans that conform the high-dose region to the target while minimizing the dose to surrounding healthy tissues by modulating the intensity of a preselected set of fixed beams. In principle, biological objective functions which incorporate tissue architecture and radiobiological response should be used for the mathematical modelling of this objective. However, a clinical placement of a biologically based inverse treatment planning system is unlikely given the current status of both the radiobiological models and their parameters (Chapter 1). Physical objectives, on the other hand, can form the basis for clinical inverse treatment planning since they provide continuity with respect to the existing standard practice by incorporating clinically accepted dose- and dose volume-based criteria for the evaluation of treatment plans.

Among the treatment planning optimization techniques based on physical objective functions the model of partial dose volume constraints[1–5] may be the most clinically relevant one. Within this model various objectives are pursued under certain constraints such as: (i) the desired target dose uniformity; (ii) the acceptable maximum dose to points in healthy structures; (iii) the fractional volumes of critical structures that can receive more than a certain dose, e. g. less than 30% of the lung may receive more than 20 Gy.

Using the above model and the dose to a target point as an objective, Langer *et al.*[1]

solved the optimization problem as a combinatorial linear program. Given k healthy organs, $m(k)$ points in the healthy organ k and $n(k)$ constraint points that have to satisfy the dose limits on the fractional volume, an explicit enumeration of each combination of constraint points in the critical organs leads to a set of $\prod_k \binom{m(k)}{n(k)}$ linear programs, which are solved by the simplex method for linear programming.[6, 7] Linear programs are problems with objectives and constraints which are linear functions of the optimization variables. This approach has been applied to the optimization of the weights of a few beams for a thoracic case, for which one point in the target, one point in the spinal cord and 15 points in the lungs were used.[1]

Such an approach is clearly impractical for large-scale problems. For this reason, in order to find the optimal beam weights for cases that involve up to 36 wedged or open beams and several structures with a few hundred points per structure, Langer *et al.*[2, 4, 5] applied mixed integer linear programming to the partial dose volume model with the dose to a target point as an objective. The integer variables (0/1) indicate whether a particular point exceeds the dose limit on the fractional volume. The algorithm proceeds by solving a sequence of linear programs in which the integer variables with the largest values at the optimum are fixed to unity.

The time to solve a mixed integer linear problem is very sensitive to the number of integer variables (number of points in healthy organs). For a thoracic case, using 450 points in the healthy lung, Langer *et al.*[4] have optimized a few beam weights in average time of 3.8 min on a DEC VAXstation 3520 (Digital Equipment Corp., Maynard, MA). When 800 points were used, the average time increased to 56 min with maximum time of 22 hr. Therefore, even though mixed integer linear programming provides a robust model for the escalation of the target dose under dose volume constraints, the technique is not applicable at the scale of the inverse treatment planning problem, which involves

a few hundred to a few thousand variables and a few thousand to tens of thousands of constraint points.

Morrill *et al.*[3] suggested another approach to solve the partial dose volume model. In their approach dose volume constraints are accounted for by introducing 3-D collars around the target which divide neighboring organs in high- and low- dose regions. A collar of different margin is assigned with respect to each structure so that the ratio of the volume of the high-dose region to the volume of the low-dose region is equal to the specified dose volume constraint for that structure. Points in the target are assigned homogeneity constraints and the integral dose to critical structures is minimized by linear programming. For the case of a pancreatic tumor treated by a combination of beams at every 10°, the approach has been applied for the optimization of up to 216 beam weights. For a few hundred constraint points, calculation times of approximately 35 min have been reported on a DEC VAXstation 3520 (Digital Equipment Corp., Maynard, MA). For routine clinical optimization of 216 variables, Morrill *et al.*[8] estimated that the number of constraint points that can be used with linear programming should be smaller than 1000.[8] This restriction partially results from the need of frequent restart of the optimization process with relaxed constraints since linear programming is unable to produce approximate solutions when a feasible point does not exist in the solution space.

Due to the computational difficulties associated with the application of the partial dose volume model at the scale of the inverse problem, cruder models are used to formulate mathematically the goal of intensity modulation conformal therapy.[9–17] These models aim at matching a prespecified dose distribution by minimizing a physical square objective which penalizes for deviations from the desired dose distributions. The rationale for such an approach is that the solution of a less clinically relevant but manageable problem with a large number of variables (pencil beam weights) can still provide better dose distributions

than a refined model with only a few degrees of freedom. That is clearly the case for targets with concave regions.[9, 13, 18] Furthermore, by exploiting the modulation of the beam intensities, inverse treatment planning may decrease the number of beams necessary to achieve certain conformity.[19] On the other hand, a large number of beams is needed to achieve conformity with unmodulated beams.[1–5]

Apart from the fact that the radiation response properties of the critical structures are not taken into account during the optimization, inverse treatment planning techniques with physical square objectives suffer from other limitations that may render the designed dose distributions clinically unacceptable. In this chapter we develop a new approach to address some of these drawbacks. Section 3.2 reviews briefly the conventional inverse treatment planning technique in order to elucidate some of its limitations. In Sec. 3.3 an alternative constrained minimization problem is proposed. On the premise that further improvement of a 3-D conformal plan with open (or wedged) beams can be achieved by modulation of the intensities of the selected beams, we propose to minimize the dose to healthy tissues while delivering the necessary dose to the target within certain uniformity. A continuous penalty function method is introduced as a numerical technique that finds approximate solutions of the constrained minimization problem. The efficiency of the continuous penalty function method depends on the choice of a numerical integration scheme and a pair of weighting functions, which are investigated in Sec. 3.5. In Sec. 3.6 examples are presented of the application of the technique to the clinical cases described in Sec. 3.4. Based on these examples, in Sec. 3.7 we discuss the use of the technique as a tool for treatment planning optimization in the general context of multiple criteria optimization.

3.2 Conventional Inverse Treatment Planning

3.2.1 Problem Statement

Let D_{S_a} be the maximum tolerance dose of a structure S_a where a is an index that enumerates the various anatomical structures. Let D_T^{min} be the minimum dose that has to be delivered to the target T and let D_T^{max} be the maximum admissible target dose as deduced from dose uniformity considerations. Let P be a voxel of the grid and $D^p(\mathbf{w})$ be the dose given to that voxel. We recall that the vector of dose values is calculated as the weighted sum of the dose contributions of the individual pencil beams as:

$$\mathbf{D} = \hat{\mathbf{H}}\mathbf{w}, \qquad \mathbf{D} = \{D^p(\mathbf{w})\}_{p=1}^{N}, \quad \hat{\mathbf{H}} = \{H_i^p\}, \quad \mathbf{w} = \{w^i : w^i \geq 0\}_{i=1}^{M} \quad . \tag{3.1}$$

The inverse treatment planning is then the process of matching a desired dose distribution by solving the following problem. Find w such that :

$$\begin{aligned} &1. \quad w^i \geq 0 \\ &2. \quad D^p(\mathbf{w}) \leq D_{S_a}, \quad P \in S_a \\ &3. \quad D_T^{max} \geq D^p(\mathbf{w}) \geq D_T^{min}, \quad P \in T \; . \end{aligned} \tag{3.2}$$

Setting $D_T = D_T^{max} = D_T^{min}$ one formulates the above objectives as a non-linear optimization problem with constraints :

$$\begin{aligned} &min\{F(\mathbf{w}) | g_p(\mathbf{w}) \leq 0, w^i \geq 0\} \\ &F(\mathbf{w}) = \frac{1}{n_T} \sum_{P \in T} (D^p(\mathbf{w}) - D_T)^2 \\ &g_p(\mathbf{w}) = D^p(\mathbf{w}) - D_{S_a}, \qquad P \in S_a \end{aligned} \tag{3.3}$$

In the above expression n_T is a normalization factor which here is selected equal to the number of calculation voxels within the target T.

A variant of the exterior penalty function method[20] has been exclusively used for solving the above problem.[9, 10, 12–17] A term proportional to the magnitude of the

constraint violation is added to the objective function to penalize constraint violations. Thus a new objective function is to be minimized :

$$min\{C(\mathbf{w})|w^i \geq 0\}$$

$$C(\mathbf{w}) = \frac{1}{n_T}\sum_{P\in T}(D^p(\mathbf{w}) - D_T)^2 + r\sum_{S_a}\frac{m_a}{n_a}\sum_{P\in S_a}R_p(D^p(\mathbf{w}) - D_{S_a})^2 \tag{3.4}$$

$$R_p = \begin{cases} 1 & if \quad D^p(\mathbf{w}) \geq D_{S_a} \\ 0 & if \quad D^p(\mathbf{w}) < D_{S_a} \end{cases}$$

where r is a penalty parameter and n_a is the number of calculation voxels pertaining to structure S_a. Different structures can be given different weights m_a. Some of the approaches[10–17] require exact matching of the prescribed dose by setting $R_p = 1$. In practice the nonnegativity constraints are imposed by truncation of the negative weights w^i to zero at each iteration step.[9–17, 21]

3.2.2 Limitations

The main limitation of the inverse treatment planning with a physical objective function is that the goal of the radiation treatment as specified by Eq. (3.2) may not be feasible for the selected beam setup. This is likely to be the case for the following two reasons. First, as yet, there is limited experience as to what is achievable with intensity modulated beams. Second, the dose distributions intensity that modulated beams can provide are difficult to conjecture due to the large number of degrees of freedom. When the goal specified by Eq. (3.2) is not feasible, the minimum of the physical objective function [Eq. (3.4)] may result in dose distributions with target underdosage[14, 22–24] that are not clinically acceptable.[25] (The term target underdosage refers to very low-dose regions whose clinical effect cannot be circumvented by acceptable renormalization of the dose distribution.) Even in cases where the desired dose distribution is feasible the question remains whether more stringent requirements could be imposed with respect to the critical structures.

Given that a critical structure is in the immediate neighborhood of a target or a dose value is assigned to each grid voxel,[10–16] the smooth real dose distribution cannot match the desired dose distribution, since the latter has sharp discontinuity at the interfaces between different structures. Even though the match may be the best in least square sense, it again may not be clinically acceptable. Collars can be introduced around the different structures to allow more realistic dose specifications in the transient region.[26] However, this creates the opportunity of assigning collar widths larger than the achievable ones, thus compromising the dose conformation.

An implementation problem arises when gradient-based optimization techniques are employed for the minimization of $C(\mathbf{w})$ in Eq. (3.4).[9–12, 14–16] It is known from the non-linear optimization theory that the problem specified by Eq. (3.3) is to be solved by a sequence of unconstrained minimization problems for increasing values of the penalty r. However, current methods[9, 10, 12, 14–16] minimize $C(\mathbf{w})$ for fixed and somewhat arbitrary value of the penalty parameter r in order to decrease the number of lengthy dose calculations. In this case one does not, in fact, distinguish the constraints from the objective function. The repercussions are that the solution of the optimization problem may not even approximately satisfy the dose constraints.

The above problems are sufficiently suppressed when a large number of fields (>15) is used. However, for conformal treatments with standard facilities (MLC, compensators), within reasonable time limits, only a few beams (5–9) are to be employed and the discussed effects can be significant for certain treatment sites. In what follows we suggest an approach that addresses some of the limitations of the inverse treatment planning technique.

3.3 The Continuous Penalty Function Method (CPFM)

3.3.1 Theory

Most of the limitations of the inverse treatment planning with a physical objective function are related to the feasibility of the specified dose levels. To overcome the difficulties resulting from infeasible dose prescriptions in our approach we aim at minimizing the dose delivered to healthy tissues while delivering the necessary dose to the target. We formulate the above objective as the following constrained minimization problem:

$$
\begin{aligned}
& min\{f(\mathbf{w})|g_j(\mathbf{w}) \leq 0, j = 1, ..., 2n_T + M\} \\
& f(\mathbf{w}) = \sum_{S_a} m_a \frac{1}{n_a} \sum_{p \in S_a} (D^p(\mathbf{w}))^2 = \sum_{S_a} m_a \frac{1}{n_a} \sum_{p \in S_a} \left(H_i^p w^i\right)^2 \\
& g_j(\mathbf{w}) = D^p(\mathbf{w}) - D_T^{max} = H_i^p w^i - D_T^{max}, \quad P \in T, \quad j = 1, ..., n_T \\
& g_j(\mathbf{w}) = D_T^{min} - D^p(\mathbf{w}) = H_i^p w^i - D_T^{min}, \quad P \in T, \quad j = n_T + 1, ..., 2n_T \\
& g_j(\mathbf{w}) = -w^i, \quad j = 2n_T + 1, ..., 2n_T + M, \quad i = 1, ..., M
\end{aligned}
\tag{3.5}
$$

The cost function $f(\mathbf{w})$ is defined as the average of the squared dose delivered to healthy tissues in order to reflect a goal of matching a zero-dose level to healthy tissues in a least-square sense. The average is calculated over the voxels of the various organs and different organs can be assigned different importance by their weights m_a. The number of calculation voxels pertaining to structure S_a is n_a and the number of calculation voxels pertaining to the target T is n_T. The constraints $g_j(\mathbf{w})$ demand that the target dose be within certain limits and that the pencil beam weights be nonnegative.

The limitations of the conventional inverse treatment planning technique are addressed by the statement of the constrained optimization problem given by Eq. (3.5). Concerning the dose prescription, it only requires two dose levels D_T^{min} and D_T^{max}. Based on the clinical experience with uniform beams, a planner can always select feasible values for these quantities. Collars around the target need not be specified, since

a solution of the above constrained minimization problem will autonomously extend the high dose region around the target as much as needed to satisfy the dose constraints. Furthermore, Eq. (3.5) poses a convex programming problem since the objective is a convex function of the pencil beam weights and the constraints are linear and therefore convex functions of the pencil beam weights. A remarkable property of any convex programming problem is that every local minimum is a global minimum.[27] Thus any solution of Eq. (3.5) guarantees the minimum possible value of the cost function without cold spots in the target volume.

To solve the inverse treatment planning problem given by Eq. (3.5) we propose to use the continuous penalty function method (CPFM).[28] It accounts for constraints by introducing a penalty term and varies the penalty coefficient continuously. Introducing the functions $g_j^+(\mathbf{w})$ defined as $g_j^+(\mathbf{w}) \stackrel{def}{=} max[0, g_j(\mathbf{w})]$, we reflect constraint violations by a convex penalty term $U(\mathbf{w})$ written as :

$$U(\mathbf{w}) = \frac{1}{n_T} \sum_{j=1}^{2n_T} \left(g_j^+(\mathbf{w})\right)^2 + \frac{1}{M} \sum_{j=2n_T+1}^{2n_T+M} \left(g_j^+(\mathbf{w})\right)^2. \tag{3.6}$$

Within the frame of the continuous approach functions $P(\mathbf{w}, t)$, $\mu(t)$ and $\tau(t)$ are defined such that :

$$\begin{aligned} &P(\mathbf{w}, t) = \mu(t) f(\mathbf{w}) + \tau(t) U(\mathbf{w}) \\ &\mu(t) > 0, \quad \tau(t) > 0, \quad \int_0^\infty \mu(t) dt = \infty \end{aligned} \tag{3.7}$$

and $\frac{\mu(t)}{\tau(t)} \rightarrow 0$ monotonically as $t \rightarrow \infty$. The limit points ($t \rightarrow \infty$) of the solutions of the following system of ordinary differential equations (ODE) are sought:

$$\frac{d}{dt}\mathbf{w}(t) = -\nabla_{\mathbf{w}} P(\mathbf{w}, t) \quad \mathbf{w}(0) = \mathbf{w}_0 \quad . \tag{3.8}$$

Using the convexity of the objective function $f(\mathbf{w})$ and the penalty term $U(\mathbf{w})$ as well as previous results[28, 29] we have proved (Appendix B) that the limit points of Eq. (3.8)

converge to a solution of the initial optimization problem [Eq. (3.5)] for any starting point $\mathbf{w}(0)$. Furthermore, since any local minimum coincides with a global minimum of Eq. (3.5) the method leads asymptotically to a global minimum. Equation (3.8) can be considered a generalization of methods for unconstrained minimization. Indeed, it turns into the steepest descent method when the time-dependent weighting functions are kept constant and the Euler integration scheme is applied.

The CPFM has several properties that indicate its utility for the large-scale problem of radiotherapy optimization. It has a simple statement and accounts for the constraints explicitly. It does not involve solving a sequence of unconstrained minimization problems, a process that presents considerable computational difficulties for large values of the penalty coefficient. Furthermore, despite the fact that the CPFM is usually used for finding approximate solutions,[28, 30] there is some numerical evidence[31, 32] that optimization routines that solve Eq. (3.8) can perform considerably better than some well-known and successful sequential quadratic programming techniques in terms of reliability, number of function evaluations as well as accuracy.

Rather than accounting for the non-negativity constraints via a penalty term, one can consider truncation to zero of the negative pencil beam weights at each iteration step of the optimization. In this case the term in Eq. (3.6) penalizing for negative pencil beam weights is omitted and the limiting points of the following equation are sought :

$$\frac{d}{dt}\mathbf{w}(t) = -[\nabla_{\mathbf{w}} P(\mathbf{w},t)]^{+} \tag{3.9}$$

where

$$\left([\nabla_{\mathbf{w}} P(\mathbf{w},t)]^{+}\right)^{i} = \begin{cases} 0 & if \quad w^{i} = 0 \quad and \quad (\nabla_{\mathbf{w}} P(\mathbf{w},t))^{i} \geq 0 \\ (\nabla_{\mathbf{w}} P(\mathbf{w},t))^{i} & otherwise. \end{cases} \tag{3.10}$$

Intuitively the above approach should produce the same results as the robust CPFM [Eqs. (3.6), (3.7) and (3.8)]. A comparison of these two approaches is given in Sec. 3.5.

3.3.2 Numerical Methods of ODE Integration

An important observation concerning the numerical solution of the system of ordinary differential equations given by Eq. (3.8) is that some computational effort is likely to be wasted in the process of following an ODE trajectory accurately when all that is really needed is the limit point of this trajectory as $t \rightarrow \infty$. For this reason we investigate several numerical methods to integrate Eq. (3.8) efficiently. All methods have been integrated in our inverse treatment planning system (Sec. 2.4.2).

As a benchmark a Runge-Kutta method[33] with adaptive stepsize control is applied to the integration of Eq. (3.8). A fractional error of five percents is set as a criterion for the routine that evaluates the truncation error as a function of the step size. Thus an accuracy of about five percent is sought for the values of the pencil beam weights at termination. To select an optimal routine to solve Eq. (3.8) we compare the results of several other integration schemes to those calculated by the Runge-Kutta method in terms of computational efficiency and accuracy.

As a first integration scheme we employ Euler's method. At the kth iteration a step from $\mathbf{w}_k$, t_k to $\mathbf{w}_{k+1} = \mathbf{w}_k + \Delta\mathbf{w}$, $t_{k+1} = t_k + h_k$ is done by calculating the correction $\Delta\mathbf{w}$ in the pencil beam weights as:

$$\Delta\mathbf{w} = -h_k \nabla P(\mathbf{w}_k, t_k). \tag{3.11}$$

The step h_k is accepted if $P(\mathbf{w}_k + \Delta\mathbf{w}, t_k) < P(\mathbf{w}_k, t_k)$ and h_{k+1} is set equal to h_k. Otherwise, h_k is reduced and the correction $\Delta\mathbf{w}$ in the pencil beam weights is recalculated from Eq. (3.11). When Eq. (3.9) is integrated with that scheme, the negative components of $\mathbf{w}_{k+1}$ are truncated to zero at each iteration step.

An alternative integration scheme arises after a linearization of Eq. (3.8) about the current iteration point $\mathbf{w}_k$. Using Taylor's expansion of $P(\mathbf{w}, t)$ in the vicinity of $\mathbf{w}_k(t_k)$

we can write an approximation to Eq. (3.8) as

$$\frac{d\mathbf{w}}{dt} = -\nabla P(\mathbf{w}_k, t_k) - \nabla^2 P(\mathbf{w}_k, t_k)(\mathbf{w} - \mathbf{w}_k) \tag{3.12}$$

where $\nabla^2 P(\mathbf{w}_k, t_k)$ is the Hessian of $P(\mathbf{w}, t)$ (a M by M symmetric matrix whose i,jth element is $\frac{\partial^2 P}{\partial w^i \partial w^j}$). We employ a slight modification of the method proposed by Brown and Bartolomew-Biggs[32], which here we refer to as an approximate Hessian method. At the kth iteration a step from $\mathbf{w}_k$, t_k to $\mathbf{w}_{k+1} = \mathbf{w}_k + \Delta\mathbf{w}$, $t_{k+1} = t_k + h_k$ involves the solution of the linear system :

$$\left[I + h_k \nabla^2 P(\mathbf{w}_k, t_k)\right] \Delta\mathbf{w} = -h_k \nabla P(\mathbf{w}_k, t_k), \tag{3.13}$$

which follows from Eq. (3.12). In the equation above I stands for the identity matrix. The step h_k is accepted if $P(\mathbf{w}_k + \boldsymbol{\Delta}\mathbf{w}, t_k) < P(\mathbf{w}_k, t_k)$ and h_{k+1} is set equal to h_k. Otherwise, h_k is reduced and the correction $\Delta\mathbf{w}$ in the pencil beam weights is recalculated from Eq. (3.13). Due to the huge size of the Hessian $\nabla^2 P(\mathbf{w}_k, t_k)$ we approximate it by a diagonal matrix A whose elements are those of the Hessian:

$$\begin{aligned} A_{ii} = \nabla^2_{ii} P(\mathbf{w}_k, t_k) = 2\mu(t_k) \sum_{S_a} \frac{m_a}{n_a} \sum_{P \in S_a} \left(H_i^p\right)^2 \\ + \frac{2\tau(t_k)}{n_T} \sum_{\substack{P \in T \\ DP < D_T^{min} \\ DP > D_T^{max}}} \left(H_i^p\right)^2 + \frac{2\tau(t_k)}{M} \sum_{\substack{i=1 \\ w^i < 0}}^{M} 1 \end{aligned} \tag{3.14}$$

Strictly speaking, the Hessian of $P(\mathbf{w}, t)$ is not a continuous function of the pencil beam weights. However, this fact does not seem to be critical for the numerical implementation of the method. When Eq. (3.9) is integrated with that scheme, the Hessian reads

$$A_{ii} = \nabla^2_{ii} P(\mathbf{w}_k, t_k) = 2\mu(t_k) \sum_{S_a} \frac{m_a}{n_a} \sum_{P \in S_a} \left(H_i^p\right)^2 + \frac{2\tau(t_k)}{n_T} \sum_{\substack{P \in T \\ DP < D_T^{min} \\ DP > D_T^{max}}} \left(H_i^p\right)^2 \tag{3.15}$$

and the negative components of $\mathbf{w}_{k+1}$ are truncated to zero at each iteration step.

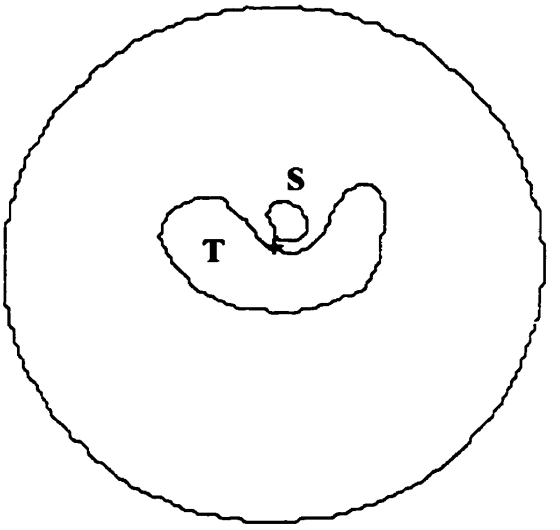

Figure 3.1. *Case A. A concave target and an organ at risk in a cylindrical phantom. The diameter of the phantom is 30 cm. Seven coplanar 18 MV photon beams are employed. The gantry angles are: 37°, 90°, 142°, 193°, 244°, 296°, 346°. Gantry angle is zero when the gantry is straight up. The angles increase for counterclockwise rotation of the gantry as viewed from the isocenter. The isocenter (cross) is placed in the center of the target volume (in center-of-mass sense).*

To adjust the integration step h_k for the Euler and the approximate Hessian method we scale it by a constant positive factor α smaller than unity until the acceptance criteria are met. As a consequence, all algorithms, themselves, establish the proper scale of the integration step after a few iterations.

3.4 Method

3.4.1 Targets, Organs at Risk and Prescription Levels

Three different cases were investigated in this work. Case A is a simulated concave target with an organ at risk in a cylindrical phantom (Fig. 3.1). It was used for the

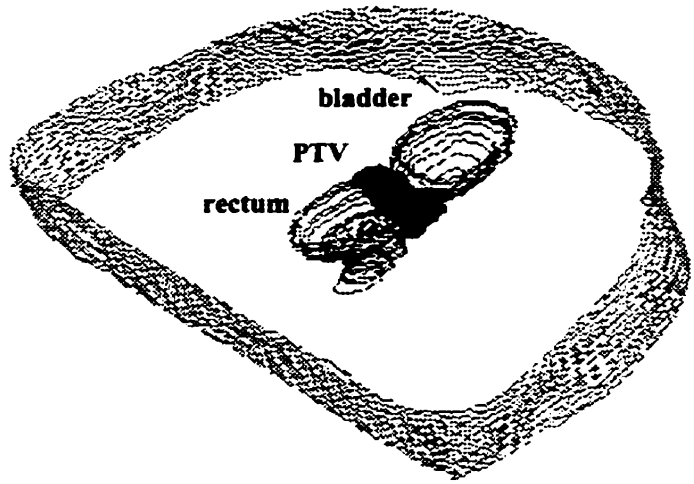

Figure 3.2. *Case B. Cancer of the prostate. Considered organs at risk are the bladder and the rectum. In this particular case the PTV does not overlap with the organs at risk.Nine coplanar 18 MV photon beams are employed. The gantry angles are: 20°, 60°, 100°, 140°, 180°, 220°, 260°, 300°, 340°. The angles increase for counterclockwise rotation of the gantry as viewed from the isocenter. The isocenter is placed in the center of the target volume (in center-of-mass sense).*

evaluation of the optimization algorithm. Case B is a clinical case of a prostate cancer with organs at risk being the bladder and the rectum (Fig. 3.2). Case C represents a target volume located in the head and neck region with organs at risk being the spinal cord and the normal tissue stroma (Fig. 3.3). For both clinical cases the patient is in a supine position with the head towards the gantry. For the clinical cases both the CPFM and a standard inverse treatment planning technique[9] are employed. Since the problem

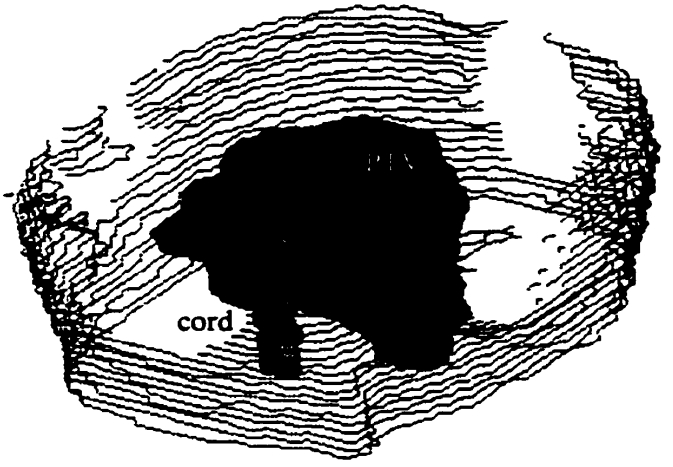

Figure 3.3. *Case C. Target in the head and neck region. The spinal cord is the organ at risk. Five coplanar 18 MV photon beams are employed. The gantry angles are: 0°, 73°, 145°, 218°, 290°. The angles increase for counterclockwise rotation of the gantry as viewed from the isocenter. The isocenter is placed in the center of the target volume (in center-of-mass sense).*

represented by Eq. (3.4) is a convex one, any minimization method will achieve the same optimal solution. Thus particular selection of a minimization algorithm for Eq. (3.4) is not crucial and several algorithms have been used with this model.[9, 14, 17] We applied the scaled gradient projection algorithm[34] to solve the problem given by Eq. (3.4) because it has been widely employed with conventional inverse treatment planning.[9, 19, 26, 35] The parameters used by the two techniques are listed in Table 3.1. For case B, the parameters used by the standard technique have values typical for prostate cases.[19, 26] For case C, the parameters used by the standard technique reflect the goal of delivering 66 Gy to

Table 3.1 *Relative minimum D^{min} and maximum D^{max} prescription dose levels used in the optimization by a standard inverse treatment planning technique and the CPFM. The weights m_a of the various structures are also given.*

Cases	$D^{min}/D^{max}/m_a$	
	Standard technique	CPFM
A. Simulated		
Tumour	...	90/100/1
1. Organ at risk	...	.../.../1
2. Normal tissue stroma	...	.../.../1
B. Prostate cancer		
Tumour	100/100/1	95/100/1
1. Bladder	.../85/1	.../.../1
2. Rectum	.../60/3	.../.../3
3. Normal tissue stroma	...	...
C. Head and neck		
Tumour	100/100/1	90/100/1
1. Spinal cord	.../67/5	.../.../5
2. Normal tissue stroma	.../90/1	... /.../1

the planning target volume (PTV) while keeping the dose to the spinal cord below 44 Gy. The dose levels for the CPFM correspond to those achieved by uniform beams for the same setup.

3.4.2 Pencil Beams

For the examples presented here 18 MV photon pencil beams were precalculated before the optimization for the given beam setup, patient anatomy and dose calculation

grid. The pencil beam weights map onto monitor units after multiplication by a factor of 100 MU/Gy. The minimum and the maximum dose levels for the target volume are specified in Gy and selected as follows. The maximum dose D_T^{max} to be delivered to the target is chosen to be equal to the maximum target dose value delivered by uniform beams of unit weight (100 MU). The minimum dose D_T^{min} to be delivered to the target is then calculated from the maximum dose D_T^{max} given the target dose inhomogeneity that can be tolerated. A two-dimensional optimization was performed for the simulated case and fully three-dimensional optimizations were performed for the other cases. To speed up the optimization, only pencil beams whose integral dose contribution to the target is above a certain level (integral dose contribution to the target over the maximum dose delivered by the pencil beam greater than 10^{-3}) are retained for the optimization. The weights of the remaining pencil beams are set to zero to prevent the irradiation of healthy tissue. For the two-dimensional optimization of case A, slit irradiations were simulated and only pencil beams in the plane of the phantom were retained. For case B the dose-calculation grid used for optimization encompassed only the target and the critical structures. The resulting beam intensities were consequently used for the calculation of the dose distributions within the entire patient volume. The irradiation and pencil beam parameters are given in Table 3.2.

3.5 Optimization Parameters

3.5.1 Integration Scheme Selection

To evaluate the performance of the integration schemes proposed in Sec. 3.3.2, we applied them to perform inverse treatment planning on case A. The dose constraint was to encompass the target by the ninety percent isodose line. The weighting functions were selected as $\mu(t) \equiv 1$ and $\tau(t) \equiv \exp(t)$. The initial pencil beam weights w(0)

Table 3.2 *Pencil beam parameters.*

Case	No. of beams	No. of pencil beams	Pencil beam size (mm^2)	Grid size	Voxel size (cm^2) (cm^3)
A	7	153	5 x 5	124 x 124 x 1	0.25 x 0.25
B	9	572	5 x 10	40 x 69 x 28	0.25 x 0.25 x 0.51
C	5	848	5 x 5	52 x 58 x 21	0.30 x 0.30 x 0.30

have been set to zero for all optimizations reported in this paper to avoid unnecessary irradiation of healthy tissues. This choice is not crucial since any solution of Eq. (3.5) is a global minimum regardless of the starting point of the optimization. For all integration schemes, the iterations were terminated as the value of the penalty term $U(\mathbf{w})$ became smaller than 2.5×10^{-6} Gy^2, which in this particular case indicated close conformity with the constraints (Table 3.3).

The evolution of the cost function $f(\mathbf{w})$ and the penalty term $U(\mathbf{w})$ when the Runge-Kutta method, the Euler method [Eq. (3.11)] and the approximate Hessian method [Eqs. (3.12) and (3.14)] are applied to the integration of Eq. (3.7) is shown in Fig. 3.4. It suggests that the Euler and the Runge-Kutta schemes are equivalent. The cost function trajectory of the approximate Hessian slightly departs from the other two ones, most probably due to the approximation of the Hessian matrix [Eq. (3.14)]. Figure 3.5 suggests the same conclusions when the Runge-Kutta method, the Euler method [Eq. (3.11)] and the approximate Hessian method [Eqs. (3.12) and (3.15)] are applied to the integration of Eq. (3.9). Furthermore, all methods produce similar dose distributions (Fig. 3.6) and Table 3.3) and intensity profiles (Fig. 3.7). Based on these data, we conclude that,

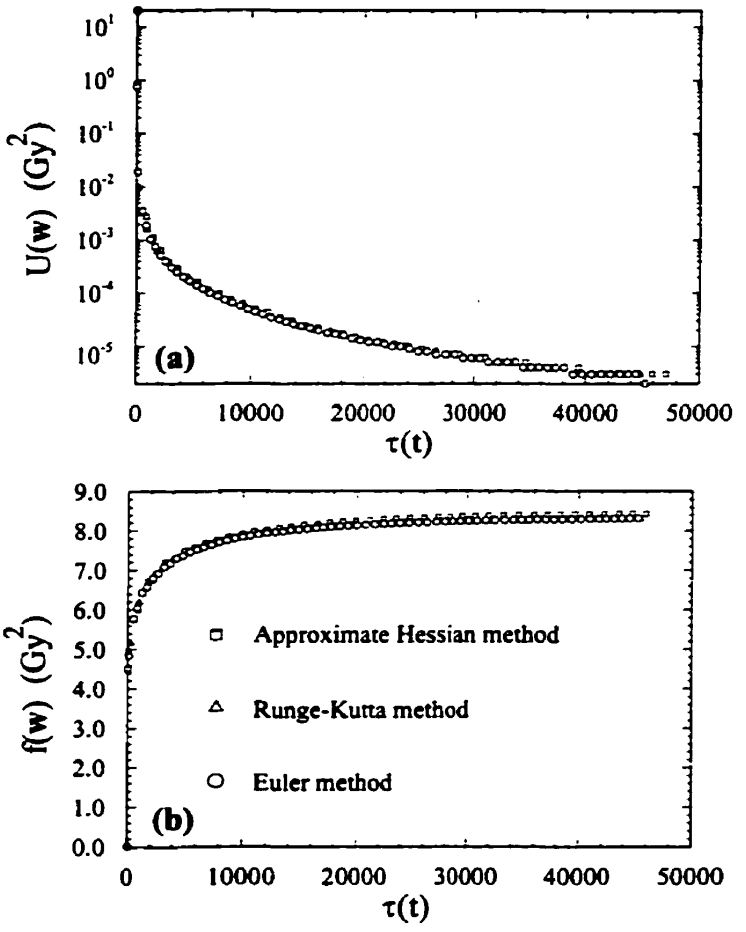

Figure 3.4. *Comparison of the Runge-Kutta method, the Euler method and the approximate Hessian method when the non-negativity constraints are explicitly included in the penalty term* $U(\mathbf{w})$. *(a) Penalty term* $U(\mathbf{w})$ *versus continuous penalty coefficient* $\tau(t)$. *(b) Cost function* $f(\mathbf{w})$ *versus continuous penalty coefficient* $\tau(t)$.

for practical purposes, in terms of their solutions, all methods of numerical integration considered in Sec. 3.3.2 are equivalent to the benchmark Runge-Kutta method.

However, in terms of their computational efficiency the integration schemes under investigation differ significantly. Given N calculation voxels and M pencil beams, a

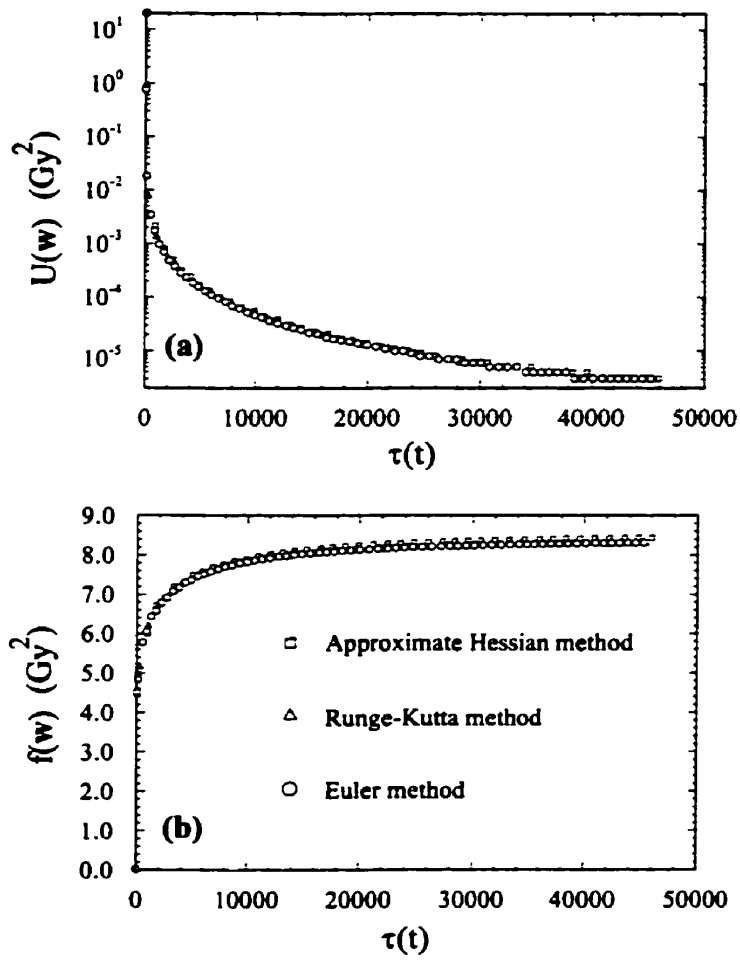

Figure 3.5. *Comparison of the Runge-Kutta method, the Euler method and the approximate Hessian method when the non-negativity constraints are implicitly accounted for by truncation of the negative pencil beam weights at each iteration step. (a) Penalty term $U(\mathbf{w})$ versus continuous penalty coefficient $\tau(t)$. (b) Cost function $f(\mathbf{w})$ versus continuous penalty coefficient $\tau(t)$.*

single dose calculation requires $2MN$ floating point operations according to Eq. (3.1). The number of floating point operations required for the calculations of either the gradient or the approximate Hessian of $P(\mathbf{w})$ is also in order of $2MN$. The evaluation of

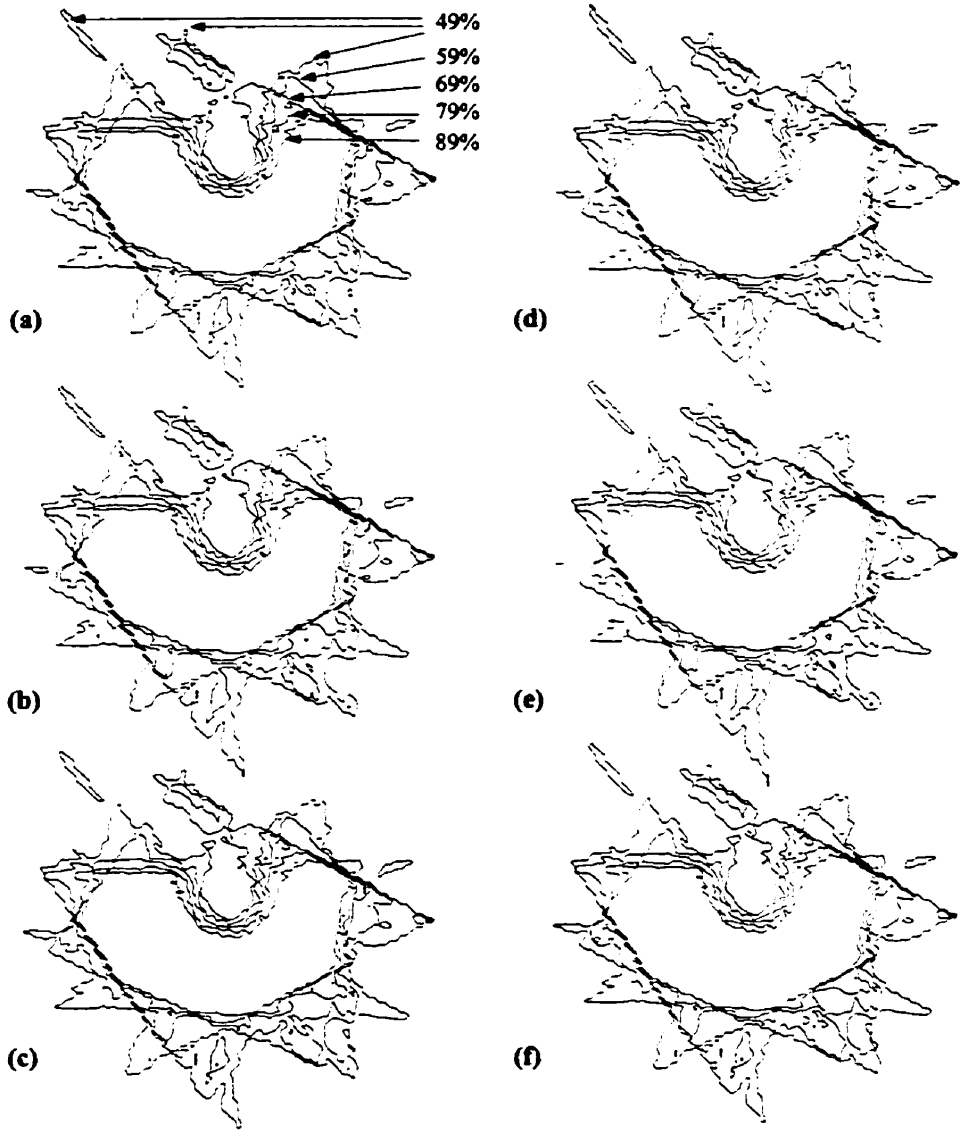

Figure 3.6. *Isodose plots for the plans created by the various integration routines. The dose distributions are normalized to the dose distribution maxima. The 89% isodose line is the one that encompasses the target in all cases. The isodose lines below 59% do not conform to the target, which is to be expected given its size, shape and the relative low number of beams employed. (a)–(c) Runge-Kutta, Euler and approximate Hessian methods when the non-negativity constraints are accounted for by a penalty term. (d)–(f) Runge-Kutta, Euler and approximate Hessian methods when the non-negativity constraints are accounted for by truncation.*

Table 3.3 *Dose statistics (rounded to 0.1%) for the plans created by the various integration routines. The dose distributions are normalized to their maxima. The non-negativity constraints are taken into account by a penalty term (p) and truncation (t). The mean $\overline{D}$, minimum D_{min} and maximum D_{max} doses for each structure are given.*

Method	Tumor			Critical structure			Rest of anatomy		
	$\overline{D}$	D_{min}	D_{max}	$\overline{D}$	D_{min}	D_{max}	$\overline{D}$	D_{min}	D_{max}
Runge-Kutta[p]	94.2	89.5	99.0	38.8	15.6	67.9	31.7	0.0	100.0
Euler[p]	93.9	89.4	98.6	38.7	15.3	68.0	31.7	0.0	100.0
Approx. Hessian[p]	95.0	90.0	100.0	39.5	17.1	68.4	31.9	0.0	99.4
Runge-Kutta[t]	94.1	89.4	99.0	38.3	15.0	67.5	31.8	0.0	100.0
Euler[t]	93.6	89.1	98.2	38.2	14.7	67.6	31.7	0.0	100.0
Approx. Hessian[t]	95.1	90.1	100.0	39.3	16.3	68.8	32.3	0.0	99.5

$P(\mathbf{w})$ itself requires few N floating point operations, an amount at least two orders of magnitude smaller than that necessary for the evaluation of either the dose distribution or the gradient $\nabla P(\mathbf{w})$. Therefore the total number of floating point operations needed for the optimization can be roughly estimated as $2(m_g + m_d)MN$ where m_g is the total number of gradient and approximate Hessian evaluations and m_d is the total number of dose distribution calculations. Based on the above estimates and the results shown in Table 3.4 we conclude that the approximate Hessian method outperforms the other ones in terms of computational efficiency. Furthermore, the time independent part of the first term of the Hessian approximation [Eq. (3.15)] can be calculated only once, which makes the method even more efficient.

The evaluation of the various integration methods strongly suggests the approximate Hessian approach [Eqs. (3.13), (3.15) and (3.9)] with negative pencil beam weights

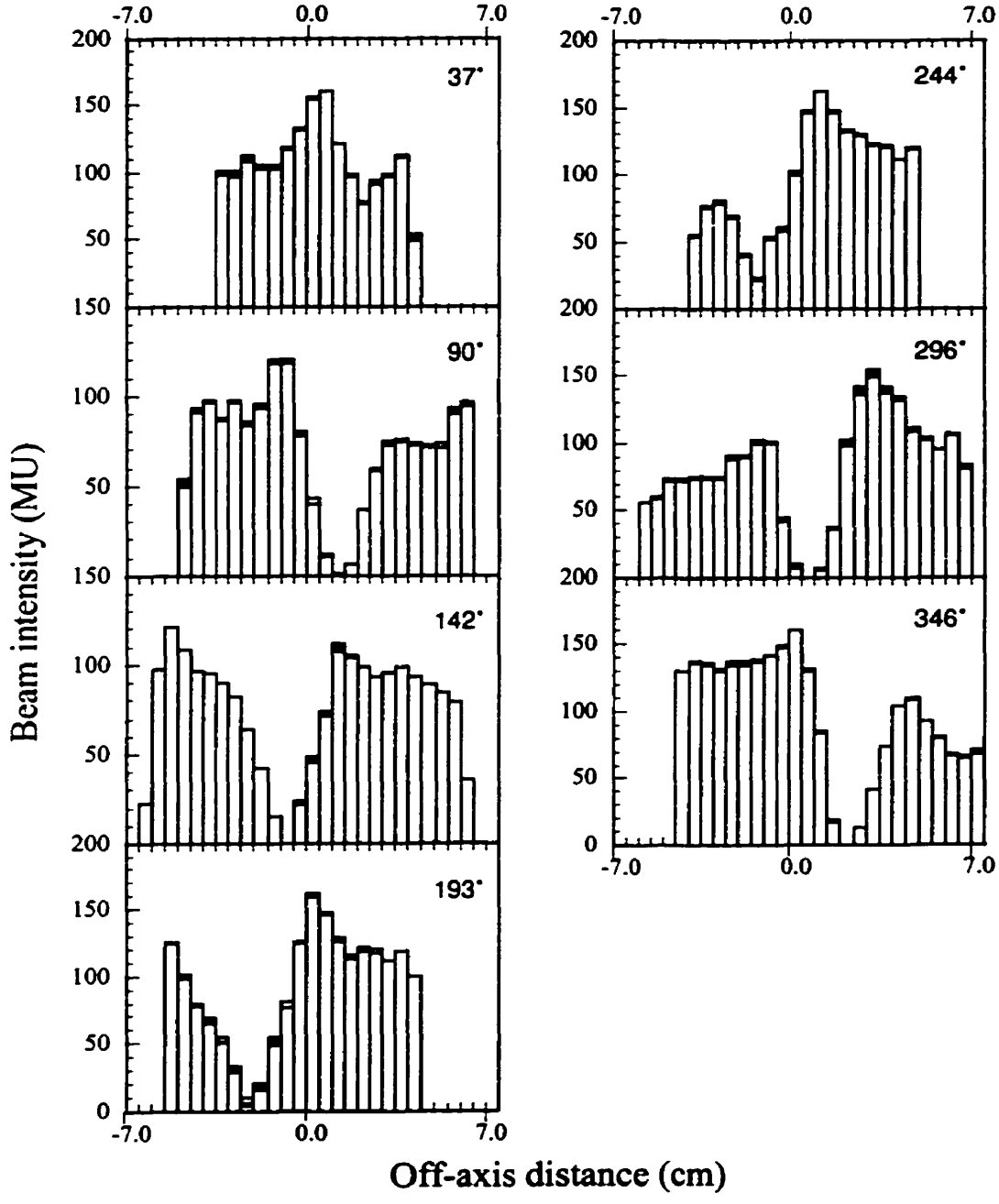

Figure 3.7. *Overlaid intensity profiles (pencil beam weights × 100 MU/Gy) obtained by the considered methods of integration (listed in Table 3.3). For each angle of incidence the profiles obtained by all six methods (Table 3.3) are shown. The gantry angles are given in the upper right corners.*

Table 3.4 *The number of dose, gradient and Hessian evaluations for the integration schemes under consideration. A common initial integration step* $h_0 = 0.2$ *and a scaling factor* $\alpha = 0.8$ *were used for all cases where needed.*

Evaluated quantity	Number of evaluations		
	Runge-Kutta	Euler	Approx. Hessian
	Non-negativity constraints included via penalty		
Dose	4167	760	519
Hessian and/or Gradient	4167	737	1000
	Negative pencil beam weights truncation		
Dose	4403	768	502
Hessian and/or Gradient	4403	745	966

truncation as the method of choice. For this reason we have employed this method for all our further investigations.

3.5.2 Weighting Functions

The requirements with respect to the weighting functions $\mu(t)$, $\tau(t)$ (Sec. 3.3.1) allow considerable freedom in their selection. Theoretically, as long as $\mu(t)$ and $\tau(t)$ possess the desired properties, the limit points $\mathbf{w}(\infty)$ of Eq. (3.8) represent a solution of the inverse treatment planning problem. Therefore, asymptotically, the cost function trajectory should converge to its minimum value. Figure 3.8(a) illustrates this point for a particular selection of the weighting functions when the approximate Hessian method was applied to perform inverse treatment planning on case A and the dose constraint was to cover the target by the 90% isodose line. After reaching a maximum value of 8.53 Gy^2 the cost function starts decreasing at a very low rate. Such a behavior is to

be expected, since, at large values of the penalty coefficient $\tau(t)$, the algorithm corrects the pencil beam weights by small amounts to avoid the large penalty that arises from constraint violation.

In practice, the numerical integration of Eq. (3.8) is to be terminated at some point $t = T$, at which the constraints are judged to be acceptably satisfied. In this case, depending on the choice of the weighting functions, the values of the pencil beam weights $\mathbf{w}(T)$ will be in different proximity of the actual solution $\mathbf{w}(\infty)$. The set of different weighting functions cannot be, of course, explored completely and an *ad hoc* choice is to be made. To the best of our knowledge, the most extensive computational experience with the continuous method for constrained optimization has been reported by Brown and Bartholomew-Biggs.[32] In their parametrization $\mu(t)$ is set to unity and $\tau(t)$ is set to a quadratic function, namely $1 + t + t^2$. Besides this particular selection we also investigated some other ones that have been used for continuous optimization.[30]

Figure 3.8(b) illustrates the cost function trajectories resulting from the various parameterizations under investigation when the approximate Hessian method was applied to perform inverse treatment planning on case A and the dose constraint was to cover the target by the 90% isodose line. As expected, the different parameterizations result in different trajectories. We retained the solutions $\mathbf{w}(T)$ that resulted in penalty values $U(\mathbf{w})$ smaller than 1.5×10^{-6} Gy^2 (2.9×10^{-5} Gy^2 for the case of the quadratic parameterization) in order to evaluate their proximity to the actual solution of the constrained optimization problem [Eq. (3.5)]. However, even for case A, 153 variables and 1900 dose constraints are involved, which would present a formidable task for most of the available non-linear programming codes. For this reason, the active set algorithm (Chapter 2) was applied to the minimization of the function $P_T(\mathbf{w}) = P(\mathbf{w}, T) = f(\mathbf{w}) + \exp(T)U(\mathbf{w})$, which had the largest cost function value for the selected constraint violation (Fig. 3.8). Within the frame of the sequential unconstrained minimization

$\exp(T)$ of the penalty term has a large value (Fig. 3.8). The pencil beam weights at termination $\mathbf{w}(T)$ were used as a starting point of the minimization. The function value was monitored and the minimization was terminated after few thousand iterations as the termination criterion

$$\left| \frac{P_T(\mathbf{w}_l) - P_T(\mathbf{w}_{l-1})}{P_T(\mathbf{w}_l)} \right| < \epsilon \tag{3.16}$$

was satisfied for ϵ equal to 10^{-5}. The comparison of the resulting plan (Fig. 3.9(b)) to the plans produced by the different parameterizations (Fig. 3.9(a), (c) and (d)) suggests that, for large values of the weighting function ratio $\tau(t)/\mu(t)$, the parameterization $\mu(t) = 4$, $\tau(t) = \exp(t)$ produces pencil beam weights $\mathbf{w}(t)$ that are in the closest vicinity of the minimizer $\mathbf{w}^*$. This observation is further confirmed by inspection of the dose statistics (Table 3.5), the cumulative dose histograms (Fig. 3.10) and some of the intensity profiles (Fig. 3.11) for the plans created by the different parameterizations.

The parameterization $\mu(t) = 4$, $\tau(t) = \exp(t)$ does not seem justified with respect to the number of iterations (Table 3.5). However, for clinical cases, acceptable conformity with the constraints is usually achieved when the penalty term is in the order of 10^{-3} $Gy^2 - 10^{-4}$ Gy^2 and a few hundred iterations suffice.

The quadratic parameterization which has been most extensively and successfully used for continuous optimization[32] failed to produce as close conformity with the constraints as the other parameterizations (Table 3.5) did because the approximate Hessian scheme could not extend the integration of Eq. (3.9) for large enough values of the penalty weighting function. Therefore, no definite conclusion can be drawn about the proximity of the pencil beam weights $\mathbf{w}(t)$ to the solution $\mathbf{w}(\infty)$ for large values of $1 + t + t^2$. Nevertheless, the plan produced by the quadratic parameterization (Fig. 3.9(c)) is evaluated against the plan produced by the solution of the constrained optimization problem (Fig. 3.9(b)) in terms of dose statistics (Table 3.5) and cumulative

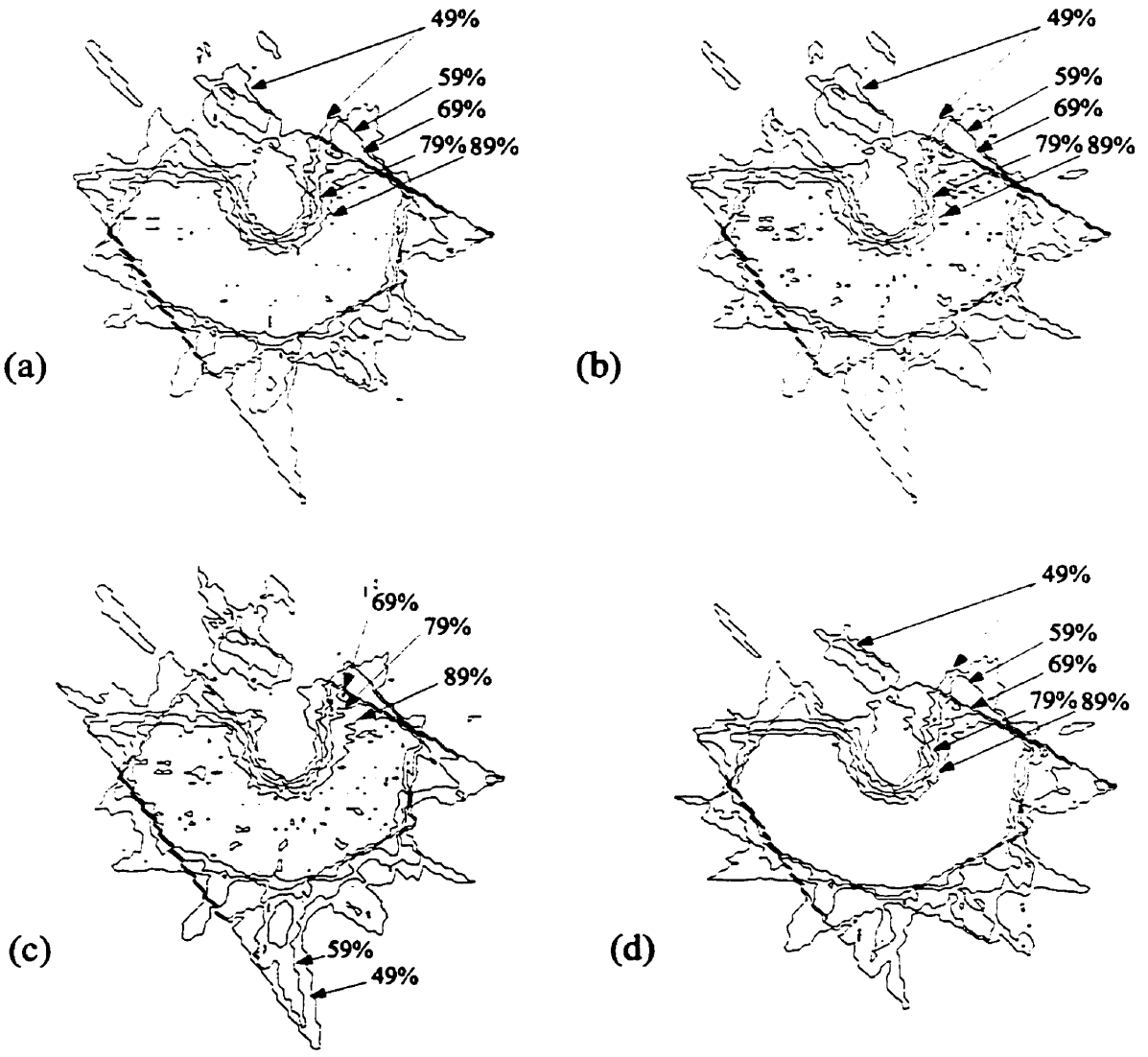

Figure 3.9. *Isodose plots normalized to the dose distribution maximum for the plans created by the approximate Hessian method with different combinations of weighting functions: (a)* $\mu(t) = 4$, $\tau(t) = \exp(t)$; *(b)* $\mu(t) = 1$, $\tau(t) = \exp(t)$ *plus additional minimization; (c)* $\mu(t) = 1$, $\tau(t) = 1 + t + t^2$; *(d)* $\mu(t) = 1$, $\tau(t) = \exp(t)$.

Table 3.5 *Number of iterations, penalty term $U(\mathbf{w})$, cost function $f(\mathbf{w})$, dose statistics (rounded to 0.1%) and the ratios $\tau(T)/\mu(T)$ of the weighting functions for the plans created by the approximate Hessian method for three different combinations of weighting functions. The dose distributions are normalized to their maxima. The mean $\overline{D}$, minimum D_{min} and maximum D_{max} doses for each structure are given.*

Iter.	$U(\mathbf{w})$	$f(\mathbf{w})$	Tumor			Critical structure			Rest of anatomy		
	x10^{-6} [Gy^2]	[Gy^2]	$\overline{D}$	D_{min}	D_{max}	$\overline{D}$	D_{min}	D_{max}	$\overline{D}$	D_{min}	D_{max}
a) $\mu(t) = 4$, $\tau(t) = \exp(t)$, $t_0 = 0$, $\tau(T)/\mu(T) = 62510$											
2800	<1.5	8.08	93.0	88.4	97.4	35.4	10.9	66.6	31.9	0.0	100.0
b) $\mu(t) = 1$, $\tau(t) = \exp(t)$, $t_0 = 0$, $\tau(T)/\mu(T) = 59725$ plus additional minimization											
	1.1	8.04	92.2	87.7	96.8	35.1	11.7	65.7	31.6	0.0	100.0
c) $\mu(t) = 1$, $\tau(t) = 1 + t + t^2$, $t_0 = 0$, $\tau(T)/\mu(T) = 9475$											
3625	29.0	7.47	92.4	86.8	97.4	31.7	8.1	62.5	31.7	0.0	100.0
d) $\mu(t) = 1$, $\tau(t) = \exp(t)$, $t_0 = 0$, $\tau(T)/\mu(T) = 59725$											
599	< 1.5	8.47	95.0	90.5	99.7	39.6	15.8	68.9	32.1	0.0	100.0

dose histograms (Fig. 3.10). These indexes as well as the similarity of the intensity profiles (Fig. 3.11) imply that the quadratic parameterization could have produced a plan close to that of the solution if large enough values of the penalty weighting function had been achieved.

It is to be acknowledged that the utility of the various parameterizations may depend on the treatment site. For this reason, in our applications of the CPFM to clinical cases, we extend the optimization by a few iterations, applying the approximate Hessian method

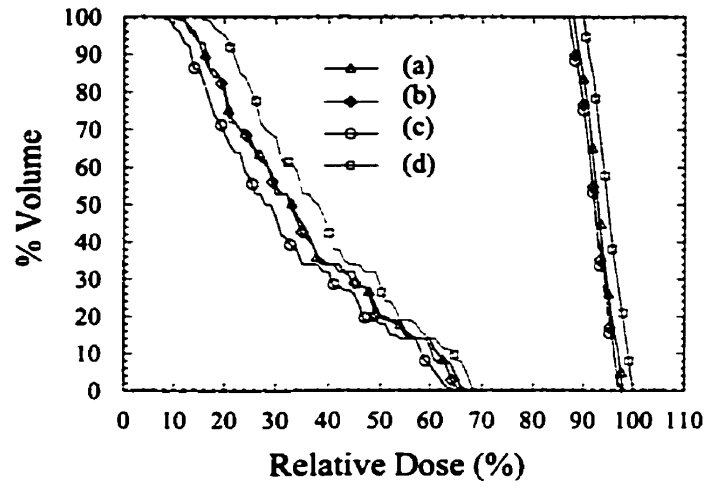

Figure 3.10. *Cumulative dose histograms (dose-volume histograms) for the target and the organ at risk (case A). Plots (a) to (d) correspond to the dose distributions illustrated in Fig. 3.9, (a) to (d). The dose distributions are normalized to their maxima.*

[Eq. (3.13)] to the minimization of the function $P(\mathbf{w}, T) = 4f(\mathbf{w}) + \exp(T)U(\mathbf{w})$ where T is the value of the parameter t at the termination of the integration of Eq. (3.9).

3.6 Examples

In this section we apply the CPFM to two clinical cases and propose a procedure for the optimization of the intensity modulation of the radiation beams when dose constraints with respect to the organs at risk are to be satisfied. Dose-volume histograms (DVH) of the target and the organs at risk (OAR) are used as a primary tool for presenting and comparing dose distributions.

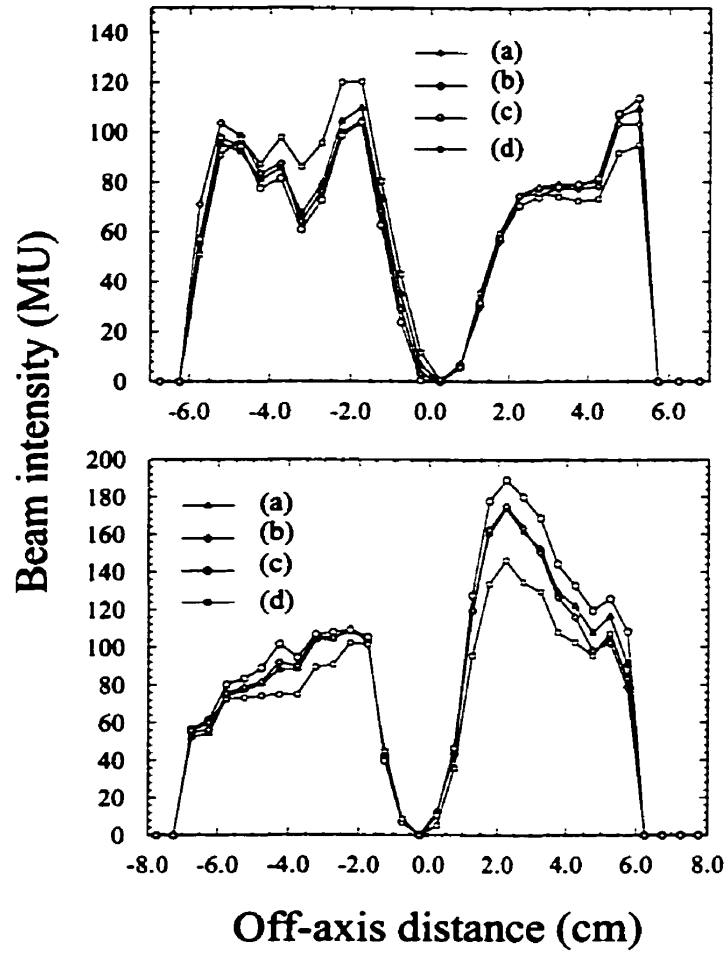

Figure 3.11. *Intensity profiles (pencil beam weights × 100 MU/Gy) of two beams produced by the approximate Hessian method with different combinations of weighting functions. Plots (a) to (d) correspond to the combinations listed in Fig. 3.9, (a) to (d).*

3.6.1 Prostate

Figure 3.12 illustrates DVHs for a family of plans produced by the CPFM for increasing values of the ratio $\tau(t)/\mu(t)$. The target dose uniformity gradually improves at the expense of larger volumes of the rectum being raised to high dose levels. Such

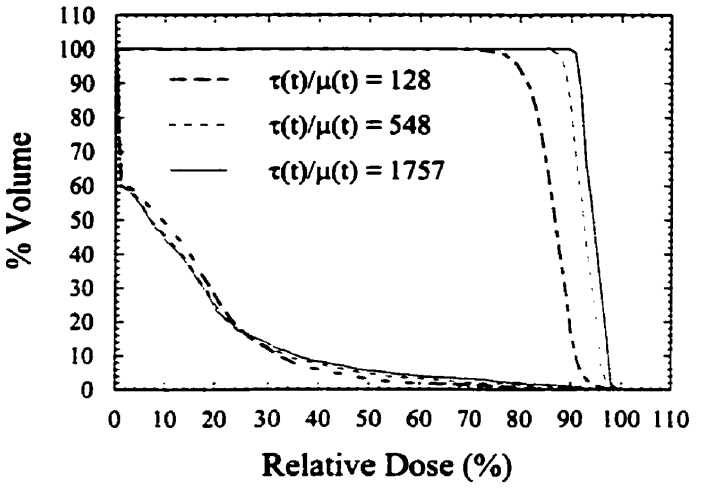

Figure 3.12. *Dose-volume histograms for target and rectum volumes for plans created by the CPFM for increasing values of the ratio $\tau(t)/\mu(t)$ of the weighting functions. The dose distributions are normalized to their respective maxima. The target dose uniformity ((maximum target dose – minimum target dose) vs. minimum target dose) for the last plan is 9% given a prescription of 5%.*

behavior is expected and the use of a large weighting factor for the target as a mean of improving dose homogeneity and conformation has been reported.[37, 23] However, the selection of the target weight has been a process of trial and error, whereas the CPFM autonomously reaches the necessary target weight to achieve the specified target dose and homogeneity.

Figure 3.13 compares a plan (plØ) obtained by the CPFM for a large value of $\tau(t)/\mu(t)$ to the one obtained from plØ after additional minimization. The iterations for the latter minimization were terminated according to our stopping criterion [Eq. (3.16)] with $\epsilon = 10^{-3}$. The marginal improvement supports the utility of the selected parameterization.

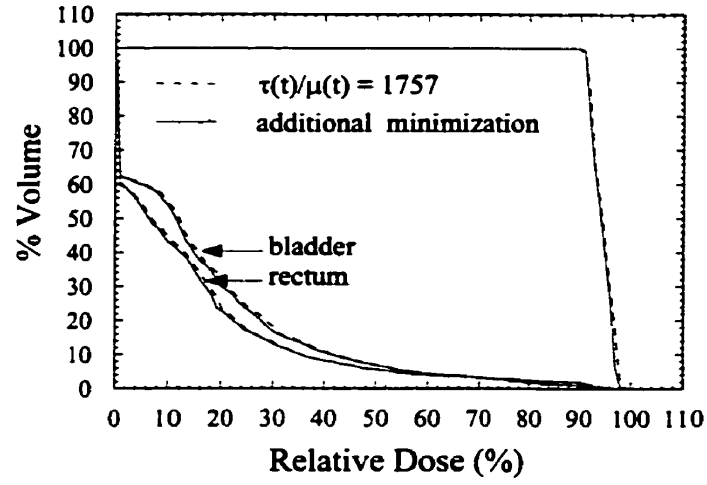

Figure 3.13. *Dose-volume histograms for target, rectum and bladder volumes. The dose distributions are normalized to their respective maxima. Dashed lines correspond to a plan (plØ) created by the CPFM* ($\tau(T)/\mu(T) = 1757$, $f(\mathbf{w}(T)) = 12.17$ *Gy*^{2}, $U(\mathbf{w}(T)) = 7.7 \times 10^{-4}$ *Gy*^{2}, *202 iterations). Solid lines correspond to the plan obtained from plØ after additional minimization (71 iterations, at termination* $f(\mathbf{w}) = 12.10$ *Gy*^{2}, $U(\mathbf{w}) = 3.8 \times 10^{-4}$ *Gy*^{2}*).*

Figure 3.14 illustrates dose distributions in few axial planes representative of a plan obtained by the CPFM at a large value of $\tau(t)/\mu(t)$. The dose distribution is normalized to its maximum. The isodose lines from 90% down to 50% conform closely to the target. In some planes, relatively large 50% hot spots develop near the skin outline (Fig. 3.14(b)). Depending on the prescription dose and the spatial location of other OARs (e. g. femural heads) these hot spots may not be acceptable. In this case, all the relevant organs at risk are to be taken into account by their inclusion in the optimization.

A central question is the application of the CPFM to cases where some dose or dose-volume constraints with respect to the organs at risk are to be satisfied. It is realistic to expect that for certain target-dose homogeneity the solution provided by the CPFM

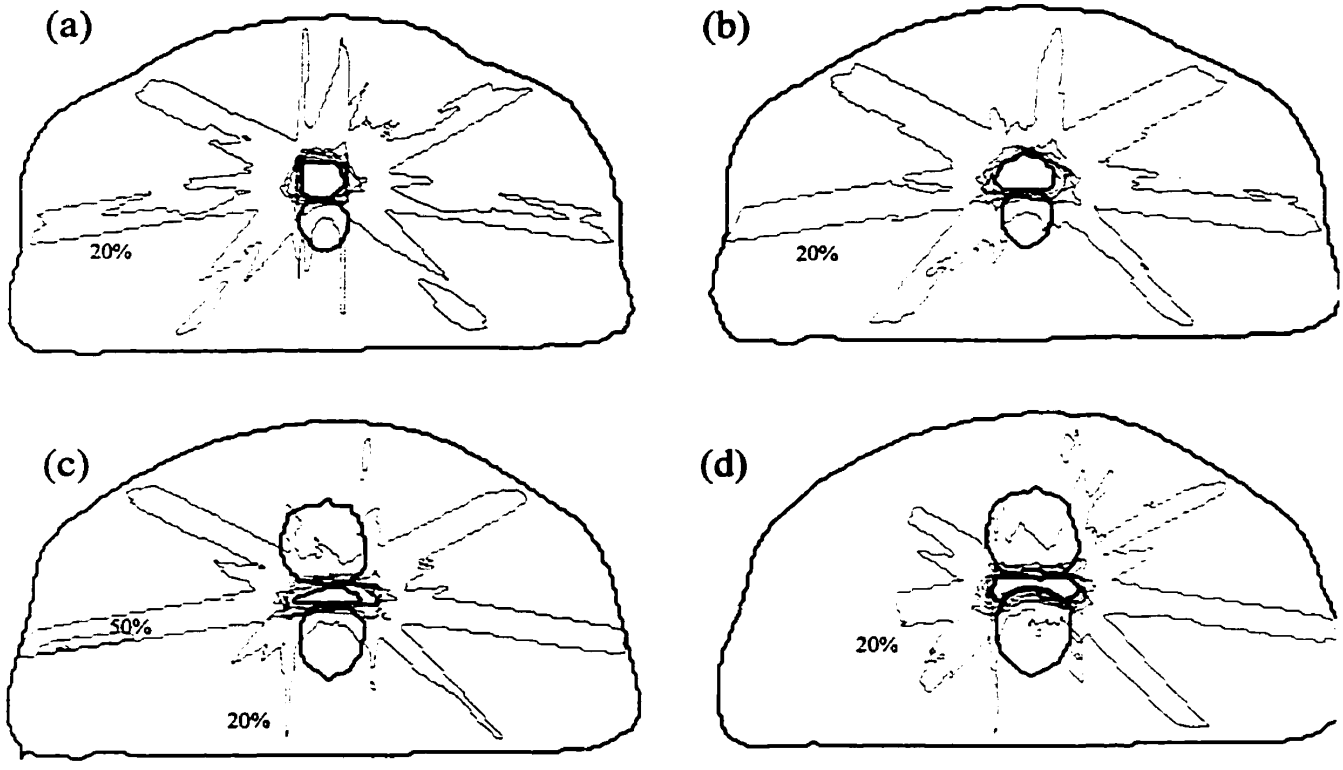

Figure 3.14. *Isodose plots for a plan produced by the CPFM* ($\tau(T)/\mu(T) = 1757$). *The dose distribution is normalized to its maximum. The 90%, 80%, 70%, 60%, 50%, 20% isodose lines are given. a) 20 mm inferior to the isocenter plane. b) 15 mm inferior to the isocenter plane. c) 10 mm superior to the isocenter plane. d) 40 mm superior to the isocenter plane.*

may not be useful due to clinically unacceptable overdosage of healthy tissue. Since one generally expects to achieve a larger level of healthy tissue sparing at the expense of decreasing uniformity of the target dose a straightforward approach is to apply the method with increasingly relaxed requirements for the target dose homogeneity (for instance 5%, 10%, 15%) and to compare the resulting plans.

Another approach is based on the observation that the CPFM creates a family of plans that present different compromises between target coverage and normal tissue sparing (Fig. 3.12). These plans can be stored for a planner's consideration. Furthermore, the conformity with certain dose and dose-volume constraints can be automatically examined at each iteration step. The last plan in the integration sequence that is still acceptable can be kept for a physician's consideration. If the ratio $\tau(T)/\mu(T)$ at termination is not large, further minimization of the function $P(\mathbf{w}, T) = 4f(\mathbf{w}) + \exp(T)U(\mathbf{w})$ can improve the plan since at this point the current estimate of the pencil beam weights $\mathbf{w}(T)$ can be significantly different from the minimizer of the function $P(\mathbf{w}, T)$. The pencil beam weights $\mathbf{w}(T)$ are to be used as a starting point of the minimization.

Figure 3.15 illustrates such a sequential procedure. The plan obtained by the CPFM for a large value of $\tau(t)/\mu(t)$ delivers the target dose within 9% uniformity but a larger volume of the rectum is raised to high dose levels in comparison to the plan obtained by the standard inverse treatment technique (Fig. 3.15(a)). We assume that this plan is clinically acceptable for illustrative purposes only. In fact, even though the parameters for the standard inverse treatment planning technique have values typical for prostate cases[19] this plan may not be optimal for our particular case due to differences in the organ and cost function definitions. From the family of plans created by the CPFM for increasing values of $\tau(t)/\mu(t)$ (Fig. 3.12) the one that produces rectum DVH close to that given by the standard technique is used for additional minimization (Fig. 3.15(b)). The resulting plan (Fig. 3.15(c)) improves both the target dose homogeneity and the sparing of the rectum. Figure 3.15(d) illustrates that a plan similar to that produced by the sequential procedure discussed above can be of course directly obtained by the CPFM under relaxed dose constraints. We emphasize this fact, since the sequential procedure can sometimes degrade either the target dose homogeneity or the sparing of the critical structures. In these cases the CPFM should be reapplied with relaxed constraints on the target dose.

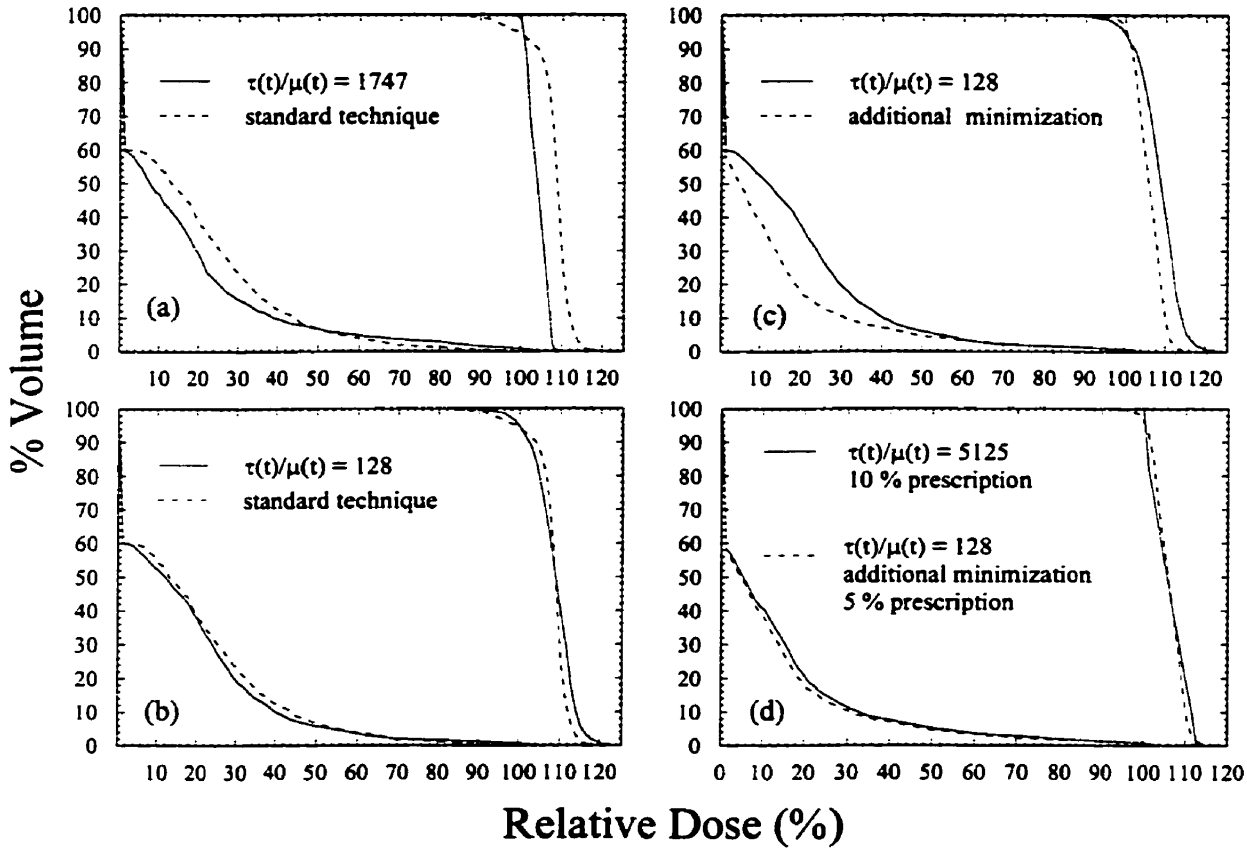

Figure 3.15. *Dose-volume histograms for target and rectum volumes for plans produced by: (a) the standard inverse treatment planning technique ($\epsilon = 10^{-3}$, 89 iterations) and the CPFM for a large ratio $\tau(t)/\mu(t)$; (b) the standard inverse treatment planning technique and the CPFM when the dose-volume histograms for the rectum are almost identical; (c) the CPFM as in (b) and additional optimization ($\epsilon = 5.0 \times 10^{-3}$, 66 iterations) of the plan created by the CPFM; (d) the CPFM with relaxed target-dose constraints ($\tau(T)/\mu(T) = 5124$, $f(\mathbf{w}(T)) = 9.01$ Gy^2, $U(\mathbf{w}(T)) = 2.3 \times 10^{-5}$ Gy^2, 282 iterations) and a sequential procedure as in (c). All plans are normalized to have nominal 100% delivered to 95% of the target volume.*

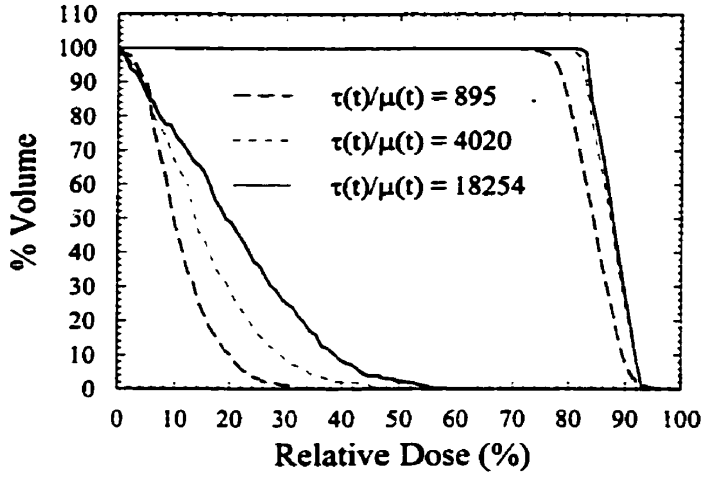

Figure 3.16. *Dose-volume histograms for target and spinal cord volumes for plans created for increasing values of the ratio $\tau(t)/\mu(t)$ of the weighting functions. The dose distributions are normalized to their respective maxima. The target dose uniformity ((maximum target dose – minimum target dose) vs. minimum target dose) for the last plan is 12% given a prescription of 10%.*

3.6.2 Head and Neck

Figure 3.16 illustrates DVHs for a family of plans produced by the CPFM for increasing values of the ratio $\tau(t)/\mu(t)$. As for the prostate case the target dose uniformity gradually improves at the expense of larger volumes of the spinal cord being raised to high dose levels. However, for this case the slight improvements of the target dose homogeneity are associated with significant differences in the irradiation of the organ at risk (Fig. 3.16) contrary to what is observed for the prostate case (Fig. 3.12).

Figure 3.16 further illustrates that for large enough values of $\tau(t)/\mu(t)$ the target is enclosed by the 80% isodose surface but the maximum of the dose distributions is not in the PTV. Therefore, if one prescribes to the 80% isodose surface some Hot Spots[38]

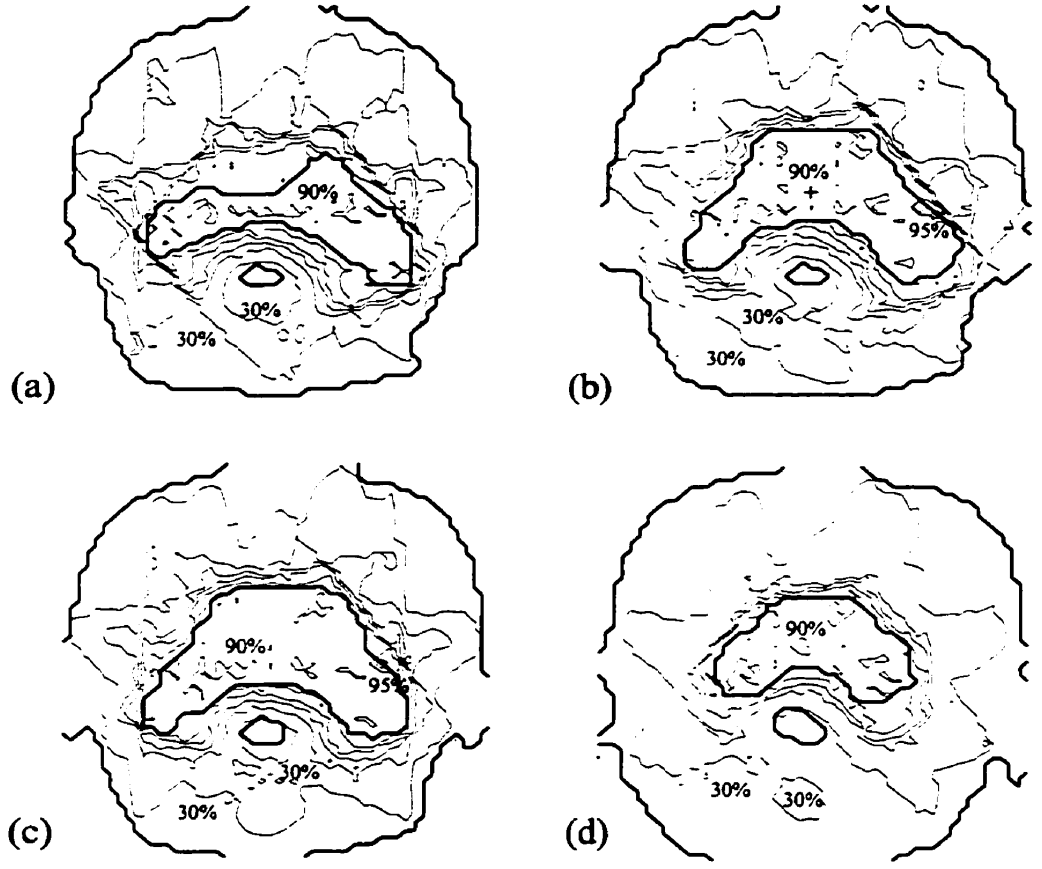

Figure 3.17. *Isodose plots for a plan produced by the CPFM* ($\tau(T)/\mu(T) = 4020$) *in few axial planes. The dose distribution is normalized to its maximum. The 95%, 90, 80%, 70%, 60%, 50%, 30% isodose lines are given. a) 6 mm inferior to the isocenter plane. b) isocenter plane. c) 3 mm superior to the isocenter plane. d) 18 mm superior to the isocenter plane.*

exist outside the PTV. However the Hot Spots above 90% and 95% are found to be very small and in the immediate vicinity of the PTV (Fig. 3.17).

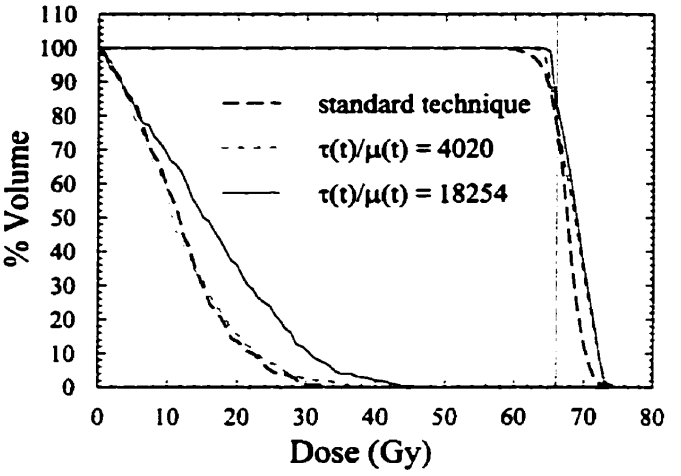

Figure 3.18. *Dose-volume histograms for target and spinal cord volumes for plans created by the CPFM and the standard inverse treatment planning technique. All plans are normalized to deliver 66 Gy to the ICRU Reference Point coinciding with the isocenter. The prescription level is indicated by the vertical line. The CPFM required 86 iterations to achieve a weighting function ratio $\tau(t)/\mu(t)$ of 4020 and 337 iterations to increase that ratio to 18254. The standard inverse treatment planning technique required 290 iterations to achieve termination ($\epsilon = 1.0 \times 10^{-3}$).*

Figure 3.18 compares the plans obtained by the CPFM for two values of the ratio $\tau(t)/\mu(t)$ to the plan obtained by the standard inverse treatment planning technique. As for the prostate case a plan similar to that produced by the standard inverse treatment technique could be selected from the sequence created by the CPFM. A few additional iterations did not improve the plan which was to be expected given the relatively large value of the weighting function ratio ($\tau(t)/\mu(t) = 4020$). The CPFM plan for $\tau(t)/\mu(t) = 18254$ delivers the target dose within -4% to +10% of the prescription dose and keeps the dose to the spinal cord below 44 Gy. However, this plan may not be necessarily the optimal one, since, in comparison with the other two plans, the

improvement in the target dose homogeneity can be considered marginal whereas the increase in the delivered dose to the spinal cord is substantial. The above observation is immediately related to the inherent limitations of the dose-based objectives for treatment planning optimization (Sec. 3.7.2).

The CPU time per iteration was 6 min 30 s for case B and 8 min for case C on a Sun SPARC 4 computer. Thus the overall calculation times were large: 22 hrs for case B and 45 hrs for case C despite the modest number of iterations required by the CPFM for both cases. One can expect a decrease in the computational time of 5 to 10 times if the optimization is performed on a high-end computer along with a more efficient dose-calculation scheme.

3.7 Discussion and Conclusions

3.7.1 The Continuous Penalty Function Method

Several investigators have suggested that any clinically relevant optimization technique should be able to impose some dose, dose-volume and/or TCP/NTCP constraints[39, 40, 19] on the relevant anatomical structures and target volumes. The existing algorithms for treatment planning optimization by beam intensity modulation through objective function minimization account for the dose constraints by the inclusion of a fixed weight penalty term. Thus constraint violations are penalized but close conformity with the constraints is not guaranteed at the conclusion of the optimization.

We introduce the CPFM to treat explicitly the treatment planning optimization as a constrained minimization problem. The method is a simple generalization of gradient-based iterative techniques for inverse treatment planning that use fixed weighting coefficients. The only additional detail is that the CPFM changes the weights from iteration to iteration in a simple, predefined manner. Given that the objective and the

penalty are convex functions of the variables of the optimization,[28, 29] the CPFM finds a solution that approximately minimizes the objective function and closely conforms to the constraints. The method performs this task with the same order of iterations as the conventional method. The latter property is very important if realistic clinical plans are to be optimized in tolerable time limits. Furthermore, numerical experience has shown that the CPFM can perform successfully on other types of objective functions as well.[32, 41] Therefore, the rationale exists for further investigation of the utility of the CPFM for inverse treatment planning when objectives and constraints different from the ones used in the work presented here are used. For instance, the TCP can be specified as an objective function while the constraints require that the NTCPs for the organs at risk be below some specified levels. For this problem the Euler integration scheme is to be used since the calculation of the approximate Hessian of biological functions is a computationally intensive process. Since only few non-trivial constraints are involved in the constrained optimization of biological indexes, such a problem may be also manageable by the available high-accuracy codes for constrained non-linear optimization despite the large number of variables to be optimized. However, the CPFM may be a viable method to tackle the complexity and the scale of the optimization of the beam intensity modulation when both dose and biological constraints and objectives are involved.

3.7.2 Objective Functions and Multiple Objective Optimization

In the work presented here the CPFM is applied to a least-square dose objective of matching zero-dose level to healthy tissues subject to dose constraints with respect to the target volume. The purpose of this statement is twofold. First, we redefine the inverse treatment planning problem based on a least-square objective as a constrained minimization problem to avoid unacceptable underdosage of the target volume, which is often the case for the least-square objective.[22–24] Furthermore, this formulation performs

some form of feasibility study of the minimum dose levels achievable beyond the extent of the target volume since the method attempts to minimize the dose outside the target as much as possible (given the necessary target coverage). Second, the particular definition of the objective and the constraints ensures the applicability of the CPFM to the constrained optimization problem. In strict mathematical terms, similar to other penalty function methods, solution of the constraint optimization problem is achieved by the CPFM at the asymptotic limit. However, a few hundred iterations is sufficient to achieve dose conformity of 2–4% with respect to the prescriptions (Figs. 3.12 and 3.16, respectively). Better conformity can be achieved with greater number of iterations. With the conventional technique, a greater number of iterations will improve the accuracy of the solution. However, if that solution itself does not provide the necessary target coverage, additional number of iterations will still result in unacceptable underdosage of the target volume.[22–24]

Other dose-based objectives can be conceived as well. For instance, the integral dose to healthy tissues can be used as a linear objective under the same dose constraints to the target and the CPFM can be applied with the Euler integration scheme to search for a solution that minimizes the objective. It has to be acknowledged, however, that any inverse treatment planning based solely on dose objectives and constraints is limited since dose-volume and radiobiological effects are not taken into account. In principal, objective functions based on models of TCP and NTCPs are more relevant than physical objectives since these models provide quantitative biophysical measure of dose distributions. However, optimization based solely on models of TCP and NTCP is currently being discouraged because the validity and predictive power of these models has not yet been proven clinically.[42]

Given the current status of the radiobiological models, the radiotherapy treatment

planning is inherently a multiple objective optimization problem.[43] Several objectives expressed in terms of dose statistics, dose-volume histograms, and estimates of TCP and NTCP are to be simultaneously optimized. A clinical decision is based on a score that combines these incompatible objectives. However, currently there is not a unanimously accepted mathematical representation of this score since it is difficult to capture the clinical judgement about the relative importance of each component of the score.[39, 44] Thus the utility function that combines the individual objectives is not yet known and the optimal plan selection requires significant computer-human interaction. Some characteristics of the inverse treatment planning approach presented here facilitate the semi-interactive process of optimal plan selection. First, the technique creates a sequence of plans with a well-defined asymptotic point. These plans present different compromises between target coverage and healthy tissue sparing that can be kept for clinical consideration. Second, the sequential approach autonomously suggests the target importance factor that may relate the minimization of the resulting square objective to a clinically relevant optimization. Third, compared to other techniques based on physical objectives, fewer dose specifications are required. Preferences with respect to various organs can be accounted for by assigning different importance weighting factors. Within the framework of CPFM, such assignments do not result in target underdosage since the penalty term weighting coefficient increases constantly thus asymptotically forcing the dose constraints [Eq. (3.7)].

To summarize, a continuous penalty function method is introduced as a tool to find approximate solutions of the large-scale constrained minimization problems that are encountered in the process of treatment planning optimization. The method is applied to an alternative formulation of the inverse treatment planning problem that obviates the need of dose specifications that may not be feasible. Several features of the resulting technique demonstrated on clinical examples suggest that it can be a viable alternative

References

1 M. Langer and J. Leong. Optimization of beam weights under dose-volume restrictions. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1255–1259, 1987.

2 M. Langer, R. Brown, M. Urie, J. Leong, M. Stracher, and J. Shapiro. Large-scale optimization of beam weights under dose-volume restrictions. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:887–893, 1990.

3 S. M. Morrill, R. G. Lane, J. A. Wong, and I. I. Rosen. Dose-volume considerations with linear programming. *Med. Phys.*, 18(6):1200–1210, 1991.

4 M. Langer, R. Brown, P. Kijewski, and H. Chul. The reliability of optimization under dose-volume limits. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:529–538, 1993.

5 M. Langer, S. Morrill, R. Brown, M. Urie, O. Lee, and R. Lane. A comparison of mixed integer programming and fast simulated annealing for optimizing beam weights in radiation therapy. *Med. Phys.*, 23(6):957–964, 1996.

6 K. G. Murty. *Linear Programming*. Joihn Wiley & Sons, Inc., New York, 1983.

7 I. Rosen, R. G. Lane, S. M. Morrill, and J. A. Belli. Treatment plan optimization using linear programming. *Med. Phys.*, 18(2):141–152, 1991.

8 S. M. Morrill, , I. I. Rosen, R. G. Lane, and J. A. Belli. The influence of dose constraint point placement in optimizaed radiation therapy treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 19(2):129–142, 1990.

9 T. Bortfeld, J. Burkelbash, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied to conformation radiotherapy. *Phys. Med. Biol.*, 35:1423–1434, 1990.

10 B. K. Lind. Properties of an algorithm for solving the inverse problem in radiation therapy. *Inverse Problems*, 6:415–426, 1990.

11 A. L. Boyer, G. E. Desorby, and N. H. Wells. Potential and limitations of invariant kernel conformal therapy. *Med. Phys.*, 18(4):703–712, 1991.

12 T. W. Holmes and T. R. Mackie. A unified approach to the optimization of brachytherapy and external beam therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 20:859–873, 1991.

13 S. Webb. Optimization by simulated annealing of three-dimensional conformal treatment palnning for radiation fields defined by multileaf collimator. II. Inclusion of two-dimensional modulation of the X-ray intensity. *Phys. Med. Biol.*, 37(8):1689–1704, 1992.

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

[15] T. W. Holmes, T. R. Mackie, and P. J. Reckwerdt. An iterative filtered backprojection inverse treatment planning algorithm for tomotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1215–1225, 1995.

[16] A. Gustafsson, B. K. Lind, R. Svensson, and A. Brahme. Simultaneous optimization of dynamic multileaf collimation and scanning patterns or compensation filters using a generalized pencil beam algorithm. *Med. Phys.*, 22(7):1141–1156, 1995.

[17] L. Xing and G. T. Y. Chen. Iterative methods for inverse treatment planning. *Phys. Med. Biol.*, 41:2107–2133, 1996.

[18] A. Brahme. Optimization of stationary and moving beam therapy techniques. *Radioth. Oncol.*, 12:129–140, 1988.

[19] J. Stein, R. Mohan, X.-H. Wang, Q. Wu, K. Preiser, C. Ling, and W. Schlegel. Number and orientations of beams in intensity-modulated radiation treatments. *Med. Phys.*, 24(2):149–160, 1997.

[20] E. Polak. *Computational methods in optimization*. Academic Press, New York, 1971.

[21] C. Raphael. Mathematical modelling of objectives in radiation therapy treatment planning. *Phys. Med. Biol.*, 37:1293–1312, 1992.

[22] S. Söderstrom, A. Gustafsson, and A. Brahme. The clinical value of different treatment objectives and degrees of freedom in radiation therapy optimization. *Radiotherapy and Oncology*, 29:148–163, 1993.

[23] V. Tulovski, M Ringor, and L Papież. Optimization of rotational radiotherapy treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1205–1214, 1995.

[24] A. Niemierko. Selection of the objective function for optimization of intensity modulated beams. *Med. Phys.*, 23:1172, 1996. (abstract).

[25] M. Goitein and A. Niemierko. Intensity modulated therapy and inhomogeneous dose to the tumor: A note of caution. *Int. J. Radiat. Oncol. Biol. Phys.*, 36(2):519–522, 1996.

[26] R. Mohan, X. Wang, T. Bortfeld, A. L. Boyer, G. J. Kutcher, S. A. Leibel, Z. Fuks, and C. C. Ling. The potential and limitations of the inverse radiotherapy technique. *Radiotherapy and Oncology*, 32:232–248, 1994.

[27] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, page 89. John Wiley and Sons, Inc, 1968.

[28] Y. G. Evtushenko. Numerical methods for solving problems of non-linear programming. *Zh. vȳchisl. Mat. mat. Fiz.*, 16(2):307–324, 1976.

[29] Y. G. Evtushenko. *Numerical Optimization Techniques*, page 213. Optimization Software, Inc., Publication Division, New York, 1st edition, 1985.

[30] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, page 126. John Wiley and Sons, Inc, 1968.

[31] A. A. Brown and M. S. Bartholomew-Biggs. Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations. *Journal of Optimization Theory and Applications*, 62(2):211–224, 1989.

[32] A. A. Brown and M. S. Bartholomew-Biggs. ODE versus SQP methods for constrained optimization. *Journal of Optimization Theory and Applications*, 62(3):371–386, 1989.

[33] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies in C*, pages 569–580. Cambridge University Press, New York, 1988.

[34] D. Bertsekas and J. Tsitsiklis. *Parallel and Distributed Processing — Numerical Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1989.

[35] C. C. Ling, C. Burman, C. S. Chui, G. J. Kutcher, S. A. Leibel, T. LoSasso, R. Mohan, T. Bortfeld, L. Reinstein, S. Spirou, X. Wang, Q. Wu, M. Zelefsky, and Z. Fuks. Conformal radiation treatment of prostate cancer using inversely-planned intensity modulated photon beams with dynamic multileaf collimation. *Int. J. Radiat. Oncol. Biol. Phys.*, 35(4):721–730, 1996.

[36] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley and Sons, Inc, 1968.

[37] S. Webb. Optimizing the planning of intensity-modulated radiotherapy. *Phys. Med. Biol.*, 39:2229–2246, 1994.

[38] ICRU. *ICRU Report 50: Prescribing, Recording, and Reporting Photon Beam Therapy*. Bethesda, Maryland: International Commission on Radiation Units and Measurements, 1993.

[39] A. Niemierko. Random search algorithm (RONSC) for optimization of radiation therapy with both physical and biological end points and constraints. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:89–98, 1992.

[40] A. Söderström and A. Brahme. Small is beautiful — and often enough. *Int. J. Radiat. Oncol. Biol. Phys.*, 34:757–758, 1996. (letter to the editor).

[41] Y. Evtushenko and V. Zhadan. Private communication, 1996.

[42] A. Niemierko. Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Med. Phys.*, 24(1):103–110, 1997.

[43] Ralph E. Steuer. *Multiple Criteria Optimization: Theory, Computation, and Application.* Joihn Wiley & Sons, Inc., New York, 1986.

[44] A. Niemierko. Treatment evaluation. In T. R. Mackie and J. R. Palta, editors, *Teletherapy: Present and Future*, page 186, Vancouver, Canada, 1996. AAPM. Proceedings of the 1996 Summer School, Advanced Medical Publishing.

CHAPTER 4 A Gray-Level Image Correlation Algorithm for Anatomy Registration

4.1 Introduction

The geometric accuracy of the patient setup is an important factor for the successful accomplishment of external-beam radiotherapy. The verification process is concerned with two major components: (i) the shape, the size and the orientation of the prescribed radiation fields themselves and (ii) the correct positioning of the patient anatomy with respect to the radiation fields. The former problem has been solved to a large extent by a number of robust algorithms which automatically extract the radiation field mask (or edges) from the portal images[1–9], analyze it and report errors in the field shape and size.[3, 6–9]

Different approaches have been developed for the registration of portal images to simulator (conventional or CT) images in order to quantify the displacement of the anatomy under the radiation field. These include manual point matching[10], semiautomatic point matching[11, 12], curve matching,[13] as well as much more automated techniques based on the extraction of bone edges in the portal images[3, 8, 14–16] and consequent chamfer matching[3] or image correlation.[8] These approaches as well as the emerging techniques for automatic three-dimensional verification of the patient setup[16] rely heavily on the extraction (either interactive or automatic) of anatomical features. This is a difficult task due to the low inherent contrast of portal images and the lack of *a priori knowledge* about the image content due to the inter-patient and treatment site variability (for the case of automatic extraction). Automatic extraction techniques require extensive testing and tuning of various edge detection and morphological operators as well as threshold levels.[14] However, the selection of these parameters is based on a cohort of patients and therefore it does not adapt from patient to patient.

In a different, two-stage approach to the anatomy matching, the portal image from the first treatment is registered interactively to a simulator image, thus producing a *reference*

portal image and the images acquired during subsequent treatments are registered to the new reference image. For this case correlation of subimages from the reference image with a portal image can be done for anatomy matching.[11, 17, 18] The advantages of this approach are that no feature extraction is necessary and the algorithms are not model based. However, the cross-correlation operator is not rotationally invariant and previous methods have used it to identify only translations.[17, 11] The cross-correlation approach is also computationally extensive. The fast Fourier transform (FFT) implementation[17] of the normalized cross-correlation was shown to improve the speed significantly, but not sufficiently to include transformations other than translations. In another approach, Radcliffe *et al.*,[18] applied Monte Carlo techniques to the calculation of the cross-correlation integral, which decreased the computational time and allowed the search space of the geometric transformations between the images to include also rotation and magnification. However, the number of samples required for satisfactory performance may varies with the anatomical site used for matching, thus forcing the use of a high number of samples which in turn impedes the speed of the registration algorithm.

In this chapter we present an approach which, in addition to the quantification of translations, incorporates in-plane rotation search in a correlation based algorithm. Since the reliability of the matching depends on the size, the shape and the contrast of the subimages used for correlation we have adopted a full calculation of the cross-correlation integral rather than its sampling through Monte Carlo techniques. The pursuit for a higher speed is done by what can be called sequential matching and a FFT implementation of the cross-correlation operators (normalized cross correlation and Pearson's linear correlation coefficient). The feasibility of this approach to the automatic registration of portal images and the automatic registration of portal images to digitally reconstructed radiographs (DRRs) is investigated.

4.2 Materials and Methods

4.2.1 Megavoltage DRRs

To test the performance of the algorithm and study the feasibility of portal-to-DRR registration a CT data set of 62 slices (PQ2000; Picker Int., Cleveland, OH) of a humanoid skull phantom (Nuclear Associate, Carle Place, NY, 11514–0349) with placed radiopaque fiducial markers (radiopaque catheter, 1 x 1 mm^2) was acquired. The settings were: 130 kVp, 24 cm field of view (pixel size=0.47 mm), 3 mm slice thickness, 0 gap. A lateral portal image of the phantom was taken with a 10 MV photon beam (Clinac 18, Varian Associates, Palo Alto, CA). For the portal image lead squares (4 x 4 x 1.5 mm^3) were placed on the skull with their inner top corners located at the positions of the radiopaque fiducial markers described above. The source-to-isocenter distance was 100.0 cm and source-to-film distance was 124.9 cm. The image was digitized to 14 bits, 0.43 mm pixels with a Du Pont LINX FD-2000 laser digitizer, and cropped to 512 x 512 pixels and processed with a 3 x 3 smoothing filter to decrease the noise. To ensure that the phantom was aligned properly to avoid out-of-plane rotations we aligned the markers placed on the phantom with the CT laser localizers and then with the laser localizers in the treatment room. Megavoltage DRRs to be used as reference images for registration were simulated by modifying the original 3D CT data set in the following manner:

I. The spectrum of the bremsstrahlung radiation for 10 MV was determined by the EGS4 Monte Carlo code.[19, 20]

II. The linear attenuation coefficients of some tissue substitutes (lung, fat, muscle, bone) of known composition were calculated from the available data for the photon cross sections of the constituents[21] and the radiation spectrum. The CT numbers

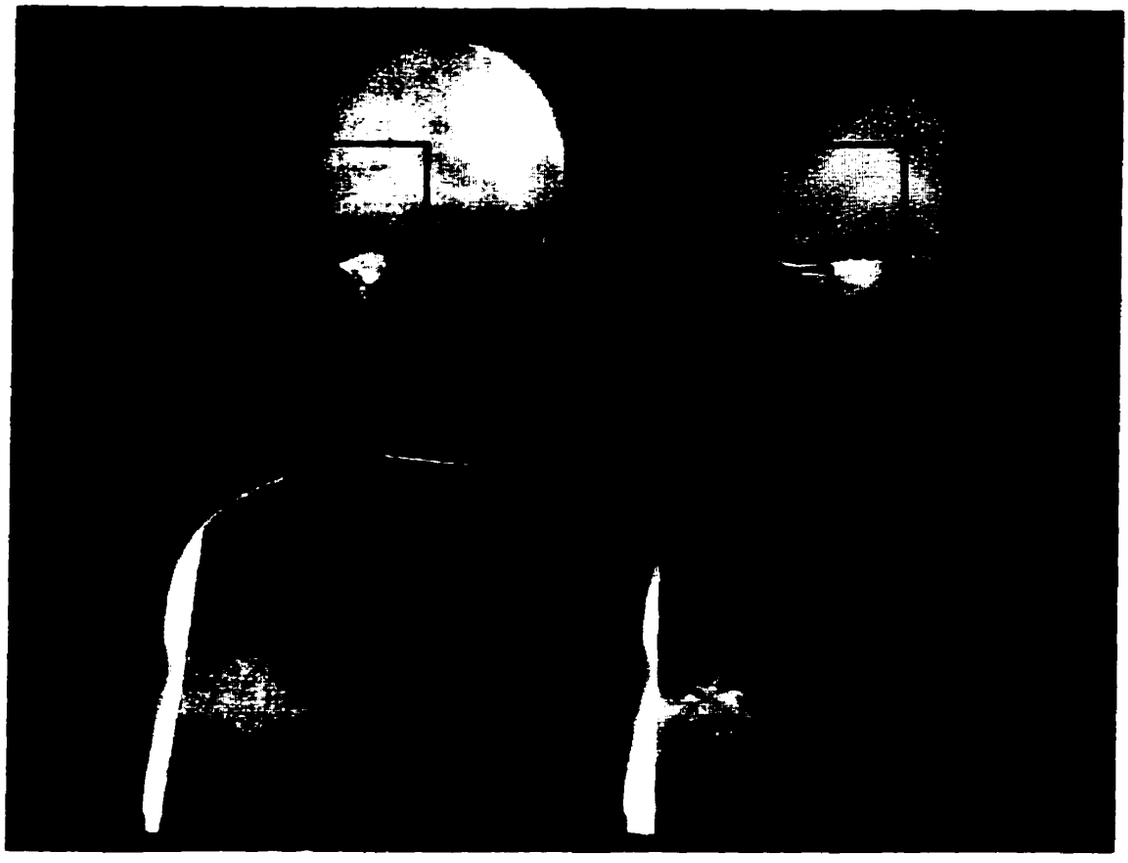

Figure 4.1. *Top right : megavoltage DRR with a selected reference subimage (large window) and the features for matching (small windows). The sphenoid is enclosed by the top small window and the Petrous bone (ear) is enclosed by the small central window. Top left : the portal image of the skull phantom with the anatomy displaced with respect to reference one (10 mm, –6 deg). The search subimage is shown by the large window. Bottom right : the difference image before the registration. The improper cancelation of bony structures is conspicuous. Bottom left : the difference image after the registration. The transformation parameters are :* $a = -25$ *pixels, b=-4 pixels,* $\alpha = 6.2$ *deg,* $\Delta a = 0$ *pixels,* $\Delta b = -2$ *pixels. The images are normalized for visualization.*

of the materials were then measured and a calibration curve of linear attenuation coefficient versus CT number was built.

III. The CT data were then mapped onto linear attenuation coefficients at 10 MV and DRRs simulated by raytracing and trilinear interpolation.[22, 23] The parameters for the raytracing were the ones used for the acquisition of the portal image.

IV. The portal-film dose response, and the calibration of the film digitizing system were finally used to modify the DRRs to simulate the effects introduced on the portal image by the portal film and the digitizing system. The average gray level value in the open portion of the field was used as a reference for the determination of the entrance fluence.

The above procedure considers only the effect of primary radiation on the image formation. In order to account partially for scatter, Dong and Boyer[24, 25] suggested to match the intensity histogram of the reference DRR to that of the one of the portal image. However, such an approach is not very robust for it implicitly relies on the notion of nearly registered images.[26] For this reason in our investigations we used the primary DRRs which are computed on sound physical basis.

The portal image was manually registered to a zero-displacement DRR by point matching (top inner corners of the lead markers (Fig. 4.1, top left) to provide a registration point for quantifying the results of the algorithm. The root-mean square (RMS) difference between the known positions of the fiducial markers and the ones given by the registration algorithm was used for this purpose. The fiducial markers were barely discernible on the megavoltage DRR (see Fig. 4.1, top right). For this reason their positions were determined from a DRR with a diagnostic quality, simulated with the same parameters as the zero-displacement megavoltage one.

4.2.2 Portal Images

Another set of five lateral portal images of the same phantom was acquired with the same treatment machine to test the portal-to-portal registration. The phantom was kept stationary with respect to the treatment coach. Translations were simulated by moving the treatment coach and in-plane rotations by rotating the collimator of the Clinac18. The source-to-film distance in this case was 126.4 cm. Four images were taken with 1 Monitor Unit (1.05 cGy/MU with isocenter setup) and one with 2 MU to investigate the sensitivity of the different cross-correlation operators to scale and shift transformations of the intensity of the images. These were digitized with the same system as above and the radiation fields were registered with the automatic extraction and matching algorithm available at our institution.[6] The images were then cropped to the size of the radiation fields-512 x 540 pixels. After this procedure one is presented with images of shifted and rotated anatomy with respect to the common reference frame established by the registered radiation field masks (Fig. 4.2). We used the results from the field matching as *a gold standard* for the angles of rotation as given by the anatomy matching algorithm.

To get some insight in the utility of our approach for clinical purposes we also applied our algorithm to the registration of a pair of clinical portal images obtained in our institution (Fig. 4.3). These were digitized as above and cropped to 512 x 512 pixels.

The algorithm is developed and tested as a MATLAB script (The Mathworks, Inc., Natick, MA 01760), which reduced the time devoted to technicalities of the implementation. MATLAB was running on a DEC 3000–M300 computer.

4.3 Registration algorithm

4.3.1 Cost Function

To register a pair of portal images we propose to determine the in-plane translations

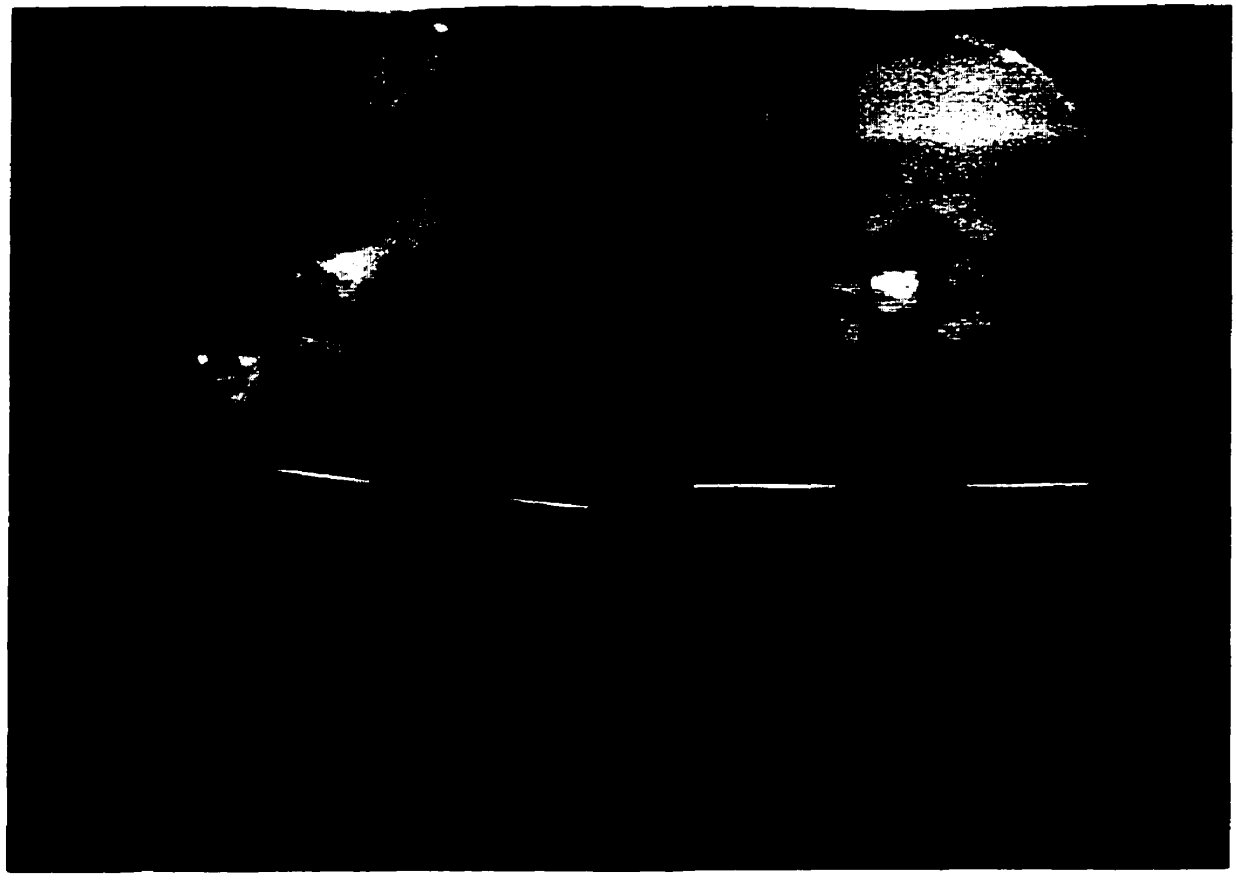

Figure 4.2. *Top right : a portal image taken at 1 MU with a selected reference subimage (large window) and a feature for matching (small window). Top left : the portal image taken at 2 MU. The windows represent search regions of decreasing size. In the different windows the position of the searched feature with respect to the axis of rotation going through the center of the corresponding search window is different. However, the dependence of the maximum correlation coefficient on the angle of rotation remains the same (see text). Bottom right : the difference image after the first (translational) stage of the registration. The improper cancelation of bony structures is conspicuous. Bottom left : the difference image after the second stage of the registration. The images are normalized for visualization.*

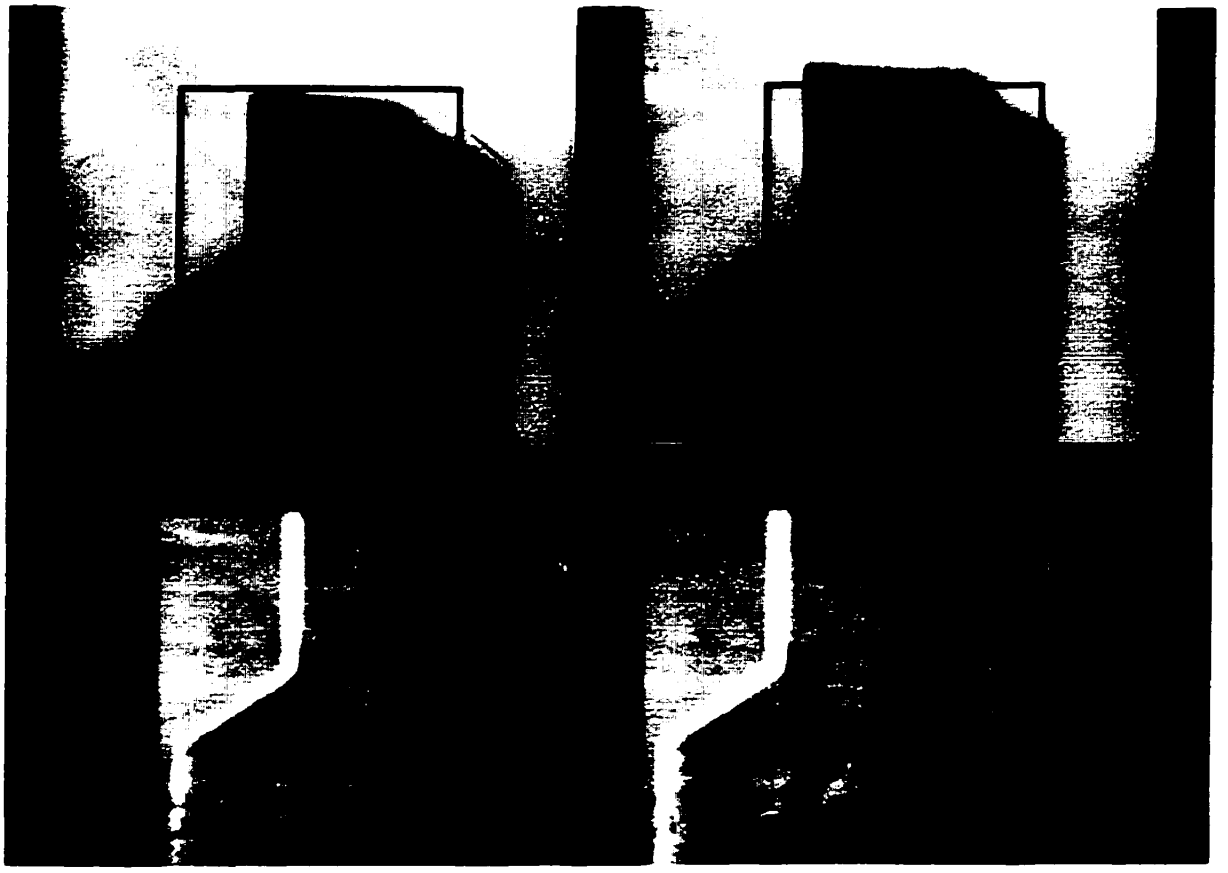

Figure 4.3. *Top left : a lateral image of a patient's head, used as a reference one with a selected feature for matching (small window) within a reference subimage (large window). Top right : a lateral image of the patient's head taken in another treatment session. The window represents the search subimage. Bottom right : the difference image after the first translational and rotational alignment. Bottom left : the difference image after the final fine translational alignment. The improved cancelation of bony structures under the radiation field due to the last fine adjustment is clearly visible. The transformation parameters are :* $a = 44$ *pixels, b=25 pixels,* $\alpha = 1.7$ *deg,* $\Delta a = -3$ *pixels,* $\Delta b = -2$ *pixels. Note that the field adjustment (indicated by the arrow) suggested by the physician is easily identifiable after the anatomy registration (top left, bottom left).*

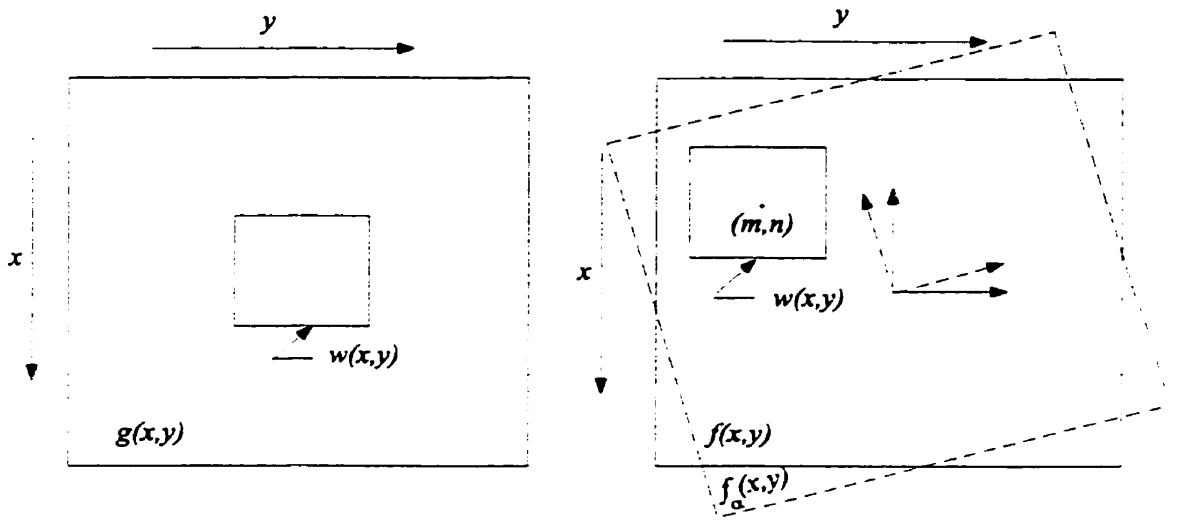

Figure 4.4. *The algorithm employs a search mask $w(x,y)$ taken from the reference image $g(x,y)$ and searches for the location of the minimum of a cost function among the rotated versions $f_\alpha(x,y)$ of the search window $f(x,y)$.*

$T(\vec{r})$ and rotation $R(\alpha)$ which maximize the value of the cross-correlation integral between a subimage within the reference image (a template) and a search image. Since the maximum correlation value within the distribution decreases as the angle of rotation between the template and the search image increases,[27] a cost function can be established to evaluate their alignment.

Let $g(x,y)$ be the *reference image* of the patient in the correct treatment position. Let $w(x,y)$ be a *search mask* of size $J \times K$ within the reference image (Fig. 4.4). The search mask encloses a feature to be matched with the same feature in the *search window* $f(x,y)$ of size $M \times N$ larger than $J \times K$. The search window includes the feature and other anatomical structures which may have been shifted and/or rotated with respect to their correct positions reflected in the reference image $g(x,y)$. Let $f_\alpha(x,y)$ be a *test*

image obtained from $f(x,y)$ after rotation at an angle α around the center of $f(x,y)$, $\bar{w}$ be the average intensity of the mask, and let $\bar{f}_\alpha(x,y)$ be the average value of $f_\alpha(x,y)$ in the region coincident with $w(x,y)$. Then, for each position (m,n) within the test image a cost function measuring the similarity between $w(x,y)$ and the region of $f_\alpha(x,y)$ under the search mask can be calculated. We considered two cost functions: the first one, $L_P(f_\alpha, w)$ being the negative of the Pearson's linear correlation coefficient[28] (PCC),

$$L_P^{m,n}(f_\alpha, w) = -r_\alpha(m,n) = \\ = -\frac{\sum_x \sum_y \left[f_\alpha(x,y) - \bar{f}_\alpha(x,y)\right]\left[w(x-m,y-n) - \bar{w}\right]}{\sqrt{\sum_x \sum_y \left[w(x-m,y-n) - \bar{w}\right]^2}\sqrt{\sum_x \sum_y \left[f_\alpha(x,y) - \bar{f}_\alpha(x,y)\right]^2}} \tag{4.1}$$

and the second one, $L_N(f_\alpha, w)$ being the negative of the normalized cross-correlation[29] (NCC),

$$L_N^{m,n}(f_\alpha, w) = -c_\alpha(m,n) = -\frac{\sum_x \sum_y f_\alpha(x,y) w(x-m,y-n)}{\sqrt{\sum_x \sum_y w^2(x-m,y-n)}\sqrt{\sum_x \sum_y f_\alpha^2(x,y)}} . \tag{4.2}$$

In the above expressions all the quantities are to be calculated for the common region of $w(x,y)$ and $f_\alpha(x,y)$ at each test location (m,n). The NCC [Eq. (4.2)] is invariant under scaling of the image intensities, $g(x,y) \longrightarrow C_1 \times g(x,y)$, whereas the PCC [Eq. (4.1)] is invariant under a more general transformation, $g(x,y) \longrightarrow C_1 \times g(x,y) + C_2$, including both scaling and shift of the image intensities. In the above expressions C_1 and C_2 are constants. Therefore the PCC is expected to be more robust for applications where the images to be registered have undergone some intensity changes due to different dose rate and detector response.

It has been shown that the NCC can be implemented through FFT-based cross-correlations[17] which increases the speed significantly due to the efficiency of the fast Fourier transform algorithms.[28, 30] The FFT-implementation of the NCC we use is that

given by Jones and Boyer.[17] In the Appendix we demonstrate that the PCC can also be implemented by FFT-based cross-correlations, which, to our knowledge, has not been reported previously.

A few more words are to be said concerning the approach to the incorporation of in-plane rotations. Clearly, the test image $f_\alpha(x, y)$ must not be cropped to the size of the initial search image $f(x, y)$. Otherwise there could be cases in which the maximum of the correlation integral is not produced at the correct location, because after the rotation, the feature to be matched has been partially cropped in the test image $f_\alpha(x, y)$, even though it has been fully included in the initial search image $f(x, y)$.

Also in this approach, if the feature under the search mask is not fully included in the search image $f(x, y)$, the correlation value at the correct location will drop because an artificial edge in $f_\alpha(x, y)$ has been matched. For this reason the feature to be matched is usually chosen at the center of the reference image thus maximizing the probability that it will be found within the search area during the registration. In this case a drop of the maximum correlation value below a properly established threshold will indicate that the feature to be located is not entirely included in the search image and a gross deviation in the patient setup has occurred.

To summarize, the registration parameters are given as follows. The angle of rotation α is the one at which the minimum of the cost function(s) [Eq. (4.1) and (4.2)] occurs. The translation vector $\vec{r}$ is determined as $\vec{r} = (m, n)_{maximum} - (m, n)_{center}$, where $(m, n)_{maximum}$ is the location of the correlation peak and $(m, n)_{center}$ is the expected position of the center of the search mask after the rotation. Therefore, to register the images one rotates the search image at an angle α and afterwards translates it by $-\vec{r}$.

4.3.2 Cost Function Minimization

The cost functions [Eq. (4.1) and (4.2)] have to be minimized with respect to (m, n, α) to obtain the transformation parameters, which in the general case may be amenable only to procedures such as Fast Simulated Annealing and Genetic Algorithms[31] which search the parameter space thoroughly in order to guarantee convergence to the absolute minimum. However, the amount of iterations involved makes such an approach impractical. We propose another approach, based on the observed properties of the cost functions.

For the range of rotations which is likely to occur clinically $(-10^o < \alpha < 10^o)$ an investigation of the properties of the NCC[17] has shown that the maximum value within the correlation distribution $max(c(f_\alpha, w))$ [Eq. (4.1)] is a unimodal, convex function of the angle of rotation with a maximum corresponding to the proper rotational alignment of the search image and the search mask. This also holds true for the case of the PCC, illustrated in Fig. 4.5. In fact the above observation holds true for in-plane rotations up to 15 deg. Therefore, the minimization of the cost functions [Eq. (4.1) and (4.2)] can be done efficiently as follows:

I. at a trial angle α the maximum value within the correlation distribution $max(c(f_\alpha, w))$ [Eq. (4.1)] or $max(r(f_\alpha, w))$ [Eq. (4.2)] and its pixel coordinates (m, n) are found by a sorting algorithm.

II. the negative of the maximum is assigned to be the value of a restricted cost function, $L(\alpha)$ at the trial angle α.

III. the cost function is then minimized by Brent's method in one dimension[32], which is in fact *golden section search*, but with improved convergence when the function has continuous second order derivative, which, as shown in Fig .4.5, seems to be the case for $L(\alpha)$. Due to the smoothness of $L(\alpha)$ and the interpolation used by

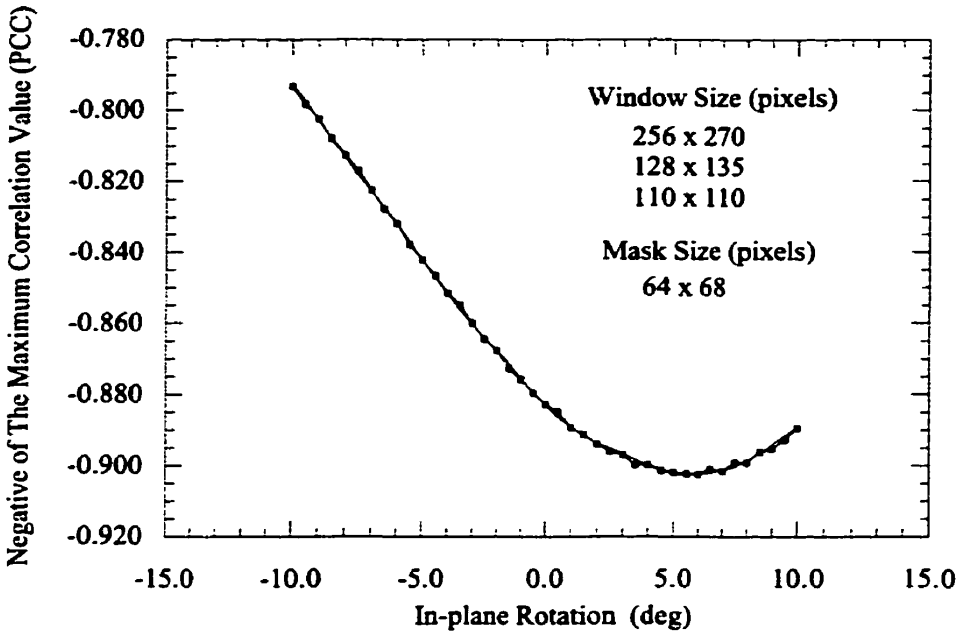

Figure 4.5. *A plot of the negative of the maximum correlation value of the correlation distribution as a function of the angle of rotation of the search windows in Fig. 4.2. The images are rotated with respect to each other at 5.8 deg. Even though the relative position of the mask with respect to the center of rotation is different, the plots are identical since for all the cases the search mask is entirely within the search window.*

Brent's method, the improvement in the accuracy of the search below some value of the user supplied rotational angle tolerance is marginal. We set this tolerance to 1 deg because it resulted in registration of no more than 0.06 deg rotation for identical images, which we considered sufficiently accurate.

4.3.3 Sequential Search

The amount of computations involved in the algorithm outlined above depends on the size of the search mask and the search window as well as the range of the angle of

rotation to be explored. However, saving computational time at the expense of decreased size of the search window and the search mask will decrease the degree of automation of the algorithm and affect the uniqueness of the feature to be matched.

It has been shown[17, 11] that for a certain range of in-plane rotations ($-15^o < \alpha < 15^o$), the maximum of the correlations value does not shift significantly (± 5pixels) from its correct position given by the coordinates of the center of the mask after the rotation. The reason is that for a 64 x 64 pixel search mask the average pixel is displaced 16 pixels from the center, and after a rotation of 10 to 15 deg it is shifted by 3 to 4 pixels. This gives rise to the uncertainty in the position of the correlation maximum discussed above.

Therefore, the approximate position of the search mask within the search window can be determined as a preliminary step by a cross-correlation (NCC or PCC) of the search mask and the search window. Then a smaller search window centered at the approximate position of the search mask is selected and the search algorithm described in Sec. 4.3.2 is initiated. A diagram of the sequential implementation of the algorithm is given in Fig. 4.6. It gives the angle of rotation, and after the rotation, the fine adjustments of the translations. To avoid the problems described in Sec. 4.3.1, the size of the new search window is calculated from the size of the search mask as shown in Fig. 4.7. At this stage the cross-correlaton integral is computed in the spatial domain, which is faster given the size of the search mask (we used a 64 x 64 search mask). The importance of the final adjustment of the translations is clearly demonstrated in Fig. 4.3.

4.4 Results

4.4.1 Algorithm Performance

To evaluate the performance of the algorithm itself we simulated 28 DRRs with

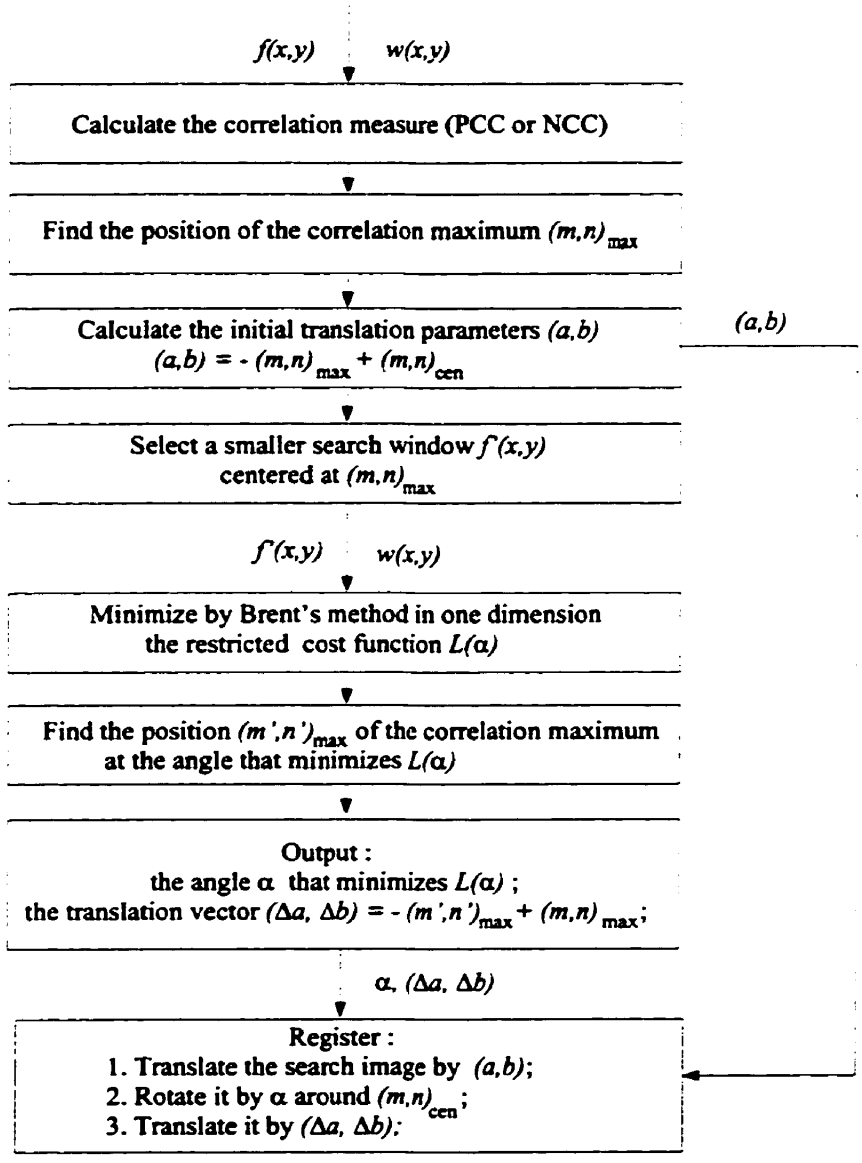

Figure 4.6. *Sequential search of the transformation parameters. The initial translational parameters* (a, b) *are calculated from the position of the maximum of the correlation distribution* $(m, n)_{max}$ *and its expected position in the reference window* $(m, n)_{cen}$. *The expected position coincides with the position of the center of the mask in the reference window. The fine adjustment of the translational parameters* $(\Delta a, \Delta b)$ *is done after the rotation by the angle* α.

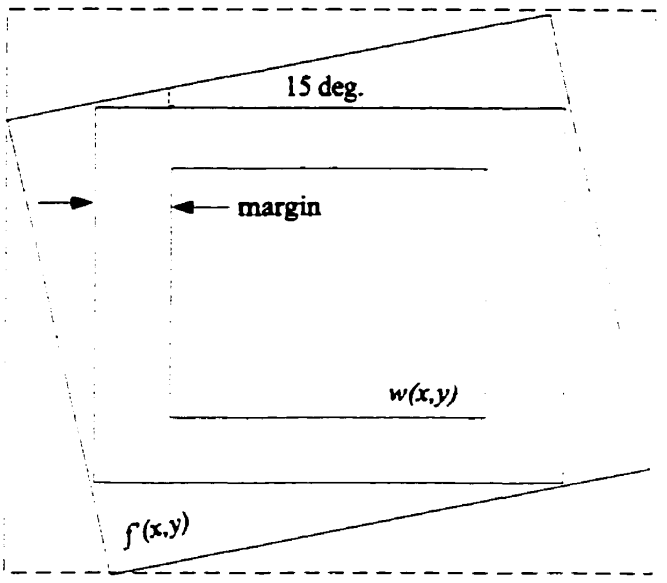

Figure 4.7. *Arrangement for selecting the size of the smaller search window $f'(x,y)$. The size of the margin is selected to be 8 pixels, twice as large as the displacement of the correlation maximum observed at 15 deg.*[17]

the anatomy shifted 0 mm, 5 mm, 10 mm, 15mm and rotated from -9 deg to +9 deg in steps of 3 deg. These images served as reference ones for the registration of the zero-displacement DRR. We used three different 256 x 256 reference/search windows with 64 x 64 search masks always placed at the center of the reference window, thus simulating 84 different cases for the algorithm. The range of the angle of rotation was set to ± 10 deg. For all the subsequent experiments no histogram modifications (normalization, etc.) were done on the images. For all results presented below, the x axis runs horizontally from right to left, the y from top down and the rotational angle increases in clockwise direction.

At this stage there was no significant difference in the performance of the algorithm because of the different cost functions. The images were aligned to 0.3 deg on average,

with maximum misalignment of 0.7 deg. The average RMS difference between the known positions of the fiducial markers and the ones given by the registration algorithm (Sec. 4.2.1) was 2.8 pixels and the maximum was 9.4 pixels, corresponding to 1 mm and 3 mm at the isocenter plane, respectively. Although the different DRRs were simulated by changing the geometry of the raytracing, the maximum correlation values were very high, above 0.994 for the PCC and above 0.9999 for the NCC, indicative of virtually identical images. The maximum translational shift given by the algorithm at the final stage was 4 pixels, which justifies our sequential approach. An interesting observation is the fact that even though the slice thickness was 3 mm along the simulated translations, alignment down to 1 mm was possible, due to the divergent beam geometry.

The average times taken by the algorithm are 1 min with the NCC cost function and 1.25 min with the PCC cost function compared to 15 min (NCC) and 20 min (PCC) with no sequential approach.

4.4.2 Portal-DRR Registration

To test the feasibility of automatic portal-to-DRR registration, we repeated the above experiments, using the same search masks and search windows, but with the portal image of the skull (Sec. 4.2.1) as a search image, instead of the zero-displacement DRR (Fig. 4.1). Since the portal image was registered to the zero-displacement DRR the true transformation parameters were known. The algorithm performed successfully only with the PCC as a cost function and only for two of the selected three features for matching- the sphenoid and Petrous bone (ear). The typical maximum correlation values were 0.86 and 0.83 respectively, reflecting the different degrees of similarity obtained in the simulation of the DRRs. The average error in the rotational angle was 0.9 deg with a maximum misalignment of 2.2 deg (Fig. 4.8). The average RMS difference between the known positions of the fiducial markers and the ones given by the registration algorithm

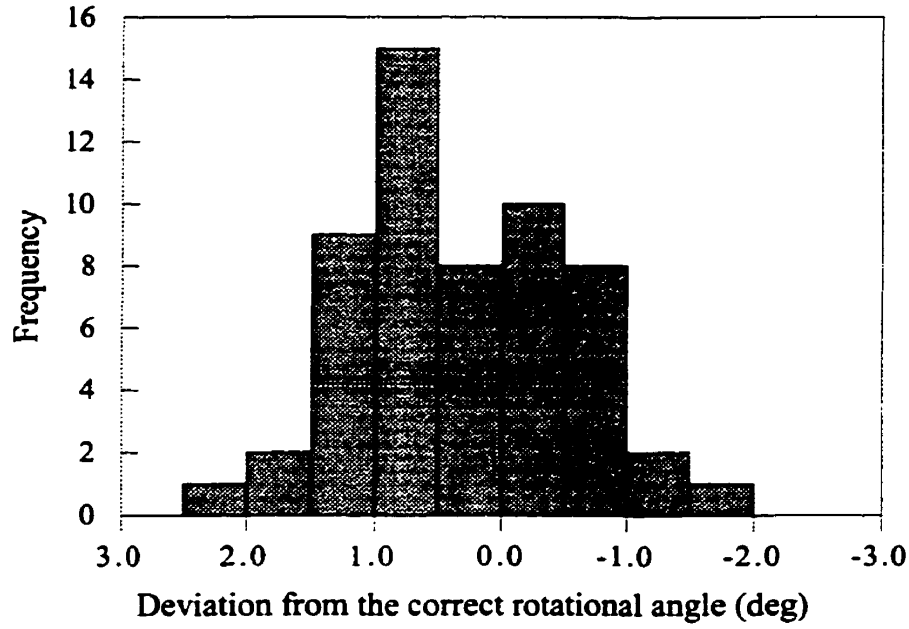

Figure 4.8. *Frequency distribution of the difference between the correct rotational angle and the one given by the registration algorithm, evaluated from 56 cases. They correspond to shifts of 0 mm to 15 mm in step of 5 mm, rotations from -9 deg to 9 deg in step of 3 deg, and 2 different features to be matched.*

(Sec. 4.2.1) was 4.2 pixels and the maximum was 10.3 pixels, corresponding to 1.5 mm and 3.3 mm, respectively, at the isocenter plane (Fig. 4.9). For the third feature, even though its approximate position was located by the cross-correlation operator, the error in the rotation angle was larger than 3 deg. These failures exist because there is not sufficient similarity between this particular feature in the DRRs and its counterpart in the portal image.

In none of the cases did the NCC detect the correct position of the features within the search window in the first stage of the algorithm which resulted in misregistration. The reason is the weaker adaptability of the NCC to the changes in the intensity of the portal

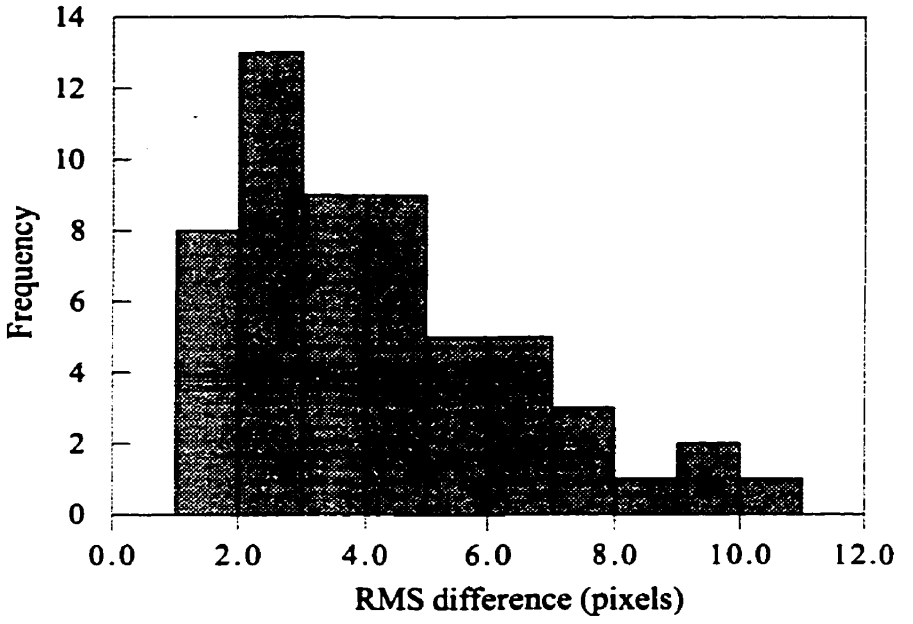

Figure 4.9. *Frequency distribution of the RMS difference between the known positions of the fiducial markers and the ones given by the registration algorithm, evaluated from 56 cases. They correspond to shifts of 0 mm to 15 mm in step of 5 mm, rotations from -9 deg to 9 deg in step of 3 deg, and 2 different features to be matched.*

image due to the in-phantom scatter which was not taken into account in the simulation of the DRRs. However, from our experiments on pairs of portal images (see Section 4.4.2) it follows that both cost functions lead to similar results if the intensities of the images are similar. To validate this we normalized the intensities of the portal image and the DRRs to achieve somewhat better similarity. This improved the performance of the NCC, which produced results close to that of the PCC for the sphenoid bone region. However in this approach there is no objective criterion when the proper histogram modification is done. For this reason we concluded that the NCC was inappropriate for portal-to-DRR registration given the quality of the DRRs we have simulated.

Table 4.1. *Tests of the algorithm on images obtained at different exposures (Fig. 4.2) with the PCC cost function. The angle given by the field registration algorithm is 5.8 deg. Both the reference and the search image have been blurred to a different degree prior to registration. Mask 1,2,3 encompass the same features as the small windows shown from left to right, respectively on Fig. 1 (Top right). Mask 4 encompasses an area between small masks 1 and 2. The maximum correlation value is denoted by M.*

	No processing		3 x 3 smoothing filter		5 x 5 smoothing filter	
mask	Angle (deg)	1-M	Angle (deg)	1-M	Angle (deg)	Maximum correlation value
1	4.9	0.265	4.9	0.109	5.2	0.101
2	4.9	0.235	5.9	0.098	5.4	0.088
3	4.6	0.367	3.6	0.150	3.0	0.138
4	4.5	0.188	4.2	0.057	4.3	0.047

4.4.3 Portal-Portal Registration

One of the portal images from our second set (Sec. 4.2.1, Fig. 4.2) was used as a reference for the registration of portal images. Four different 256 x 256 reference/search windows with 64 x 64 search masks placed at the center of the reference window were selected for the registration, thus simulating 16 different cases for the algorithm. We only investigated the detection of the in-plane rotations by the algorithm for the translations were difficult to control with our experimental setup. The RMS difference between the rotation angles as given by the registration of the anatomy from the ones given by the field registration were calculated for both cost functions, with the acceptance range set to 3 deg. For both cost functions the RMS difference was 1.4 deg, but with different failure : 6% for the PCC and 19% for the NCC. This was due to the failure of the NCC in the registration of images obtained at different exposures (Table 4.1 and Table 4.2).

Table 4.2. *Tests of the algorithm on images obtained at different exposures (Fig. 4.2) with the NCC cost function. The angle given by the field registration algorithm is 5.8 deg. Both the reference and the search image have been blurred to a different degree prior to registration. Mask 1,2,3 encompass the same features as the small windows shown from left to right, respectively on Fig. 1 (Top right). Mask 4 encompasses an area between masks 1 and 2. The maximum correlation value is denoted by M.*

	No processing		3 x 3 smoothing filter		5 x 5 smoothing filter	
mask	Angle (deg)	1-M (10^{-4})	Angle (deg)	1-M (10^{-4})	Angle (deg)	Maximum correlation value (10^{-4})
1	4.9	3.8	4.9	1.4	4.9	1.2
2	-5.8	4.2	-5.9	1.8	-5.2	1.6
3	-7.5	3.8	-1.7	1.2	-2.0	1.1
4	6..7	3.5	0.0	1.2	0.1	1.1

In these cases the minimum of the cost function did not correspond to the true translation parameters at the first stage of the algorithm. Our experiments also showed that for some features (with low spatial frequencies) the cost functions are very robust against change in the resolution of the images, a fact observed previously by Radcliffe et al.[18] This is important for it allows some preprocessing of the images in order to remove noise and increase the significance of the correlation peak as a reflection of the quality of the match. More work is required to correlate the performance of the algorithm to specific types and shapes of anatomical structures, and to different smoothing processes.

4.5 Discussion and conclusions

One of the goals of this work is to demonstrate an approach to the determination of in-plane rotation (down to 1 deg) in a correlation-based algorithm for image registration

in a reasonable time frame. The approach clearly relies on the assumption that the range of the in-plane rotation is ± 15 deg, which is greater than the one expected clinically (around 4–5 deg).[11] When using film-screen systems for imaging, translational errors and rotational errors greater than those above may be introduced because of the film-digitization process. For this reason registration of radiation fields has to be performed prior to the registration of anatomy. The images are aligned first with the parameters obtained from the registration of the fields which establishes a common reference frame. The misregistration of the anatomy in this frame should, in general, be small when patient setup is properly performed. Well calibrated on-line portal systems do not pose such a problem because one can assume that a change of position of a patient can be detected as a change in position of the patient's anatomy relative to the pixel matrix of the image.

We did not attempt to develop any correlation procedure for the determination of the magnification. A previous study[18] has shown that this parameter could not be precisely determined by a correlation based algorithm. The fact that the correlation operator can determine translations and rotation (by a search) indicates that the divergent geometry of the beam do not change the appearance of the projection image significantly. Because the beam is not grossly divergent, the redistribution of the pixels in the resultant images due to the small changes in magnification may not be precisely detected by the correlation operator. In what follows we discuss a possible approach to the incorporation of the magnification in an automatic registration algorithm.

The performance of the algorithm outlined above depends on the ability of the correlation operators to match the search feature within the search window, which itself depends on several factors: (i) the similarity of the images, (ii) the presence of prominent anatomical feature(s), as well as (iii) the quality of the images.

The size and the shape of the anatomical feature as well as its subject contrast

determine how well it can be matched in the search region.[11, 33] We did not attempt any investigation in this direction since we believe that the identification of the most appropriate features for matching and registration for a certain treatment setup can be easily and reliably identified by experience from a few clinical cases. This learning process is also necessary since the choice of the best landmarks also improves the robustness of the algorithm against noise, change in the resolution as well as the existence of structures that can confound the correlation operator. For those selected features the value of the cost functions after the registration can also be used as a criterion for the quality of the match, given that the imaging chain (linear accelerator and portal imaging device) is stable. The use of the PCC as a cost function makes the registration more robust in cases where the images have undergone some intensity modifications due to changes in the dose rate or the acquisition technique (different number of MU).

An important consideration of the utility of the above algorithm to the automatic registration of portal images is the existence of nonrigid transformations of the patient. An example is the difference image in Fig. 4.3. Even though the structure to be matched is registered well as indicated by the uniform gray levels in the corresponding region, some minor mismatch under the radiation field is noticeable. This cannot be managed by the algorithm in its present form. The reason is that it finds the transformation parameters which best register the selected feature locally and assigns those parameters as a global transformation parameters under the assumption of rigid transformations.

This problem as well as the determination of the magnification can be approached at the expense of increased computational time in the following manner. A set of features can be selected and the registration parameters can be found for each one by the algorithm in its present form. Then a model based transformation can be found that is in the best agreement with the parameters calculated before and weighted (that can be done in several

ways) with the values of the cost function. In this way, nonrigid transformations and image magnification can be determined.

Such a procedure will be advantageous even in the case of rigid transformations. In this case the Procrustes algorithm[11, 13, 34] can be used to find the translation, rotation and scale values, which, when applied to the search image, minimize the least square differences between the paired features in the search and the reference image. This approach is similar to the one suggested by Moseley and Munro[11] but has the advantage of (i) incorporating a larger search window and (ii) decreasing the error introduced by the correlation operator in the identification of the match points when the images are rotated.

Even though the proposed automatic algorithm may never be completely successful, it can be implemented as a tool in more complex strategies for registration of portal images. To make these strategies effective, a user-friendly interface that provides a selection of different options with a different degree of automatization is to be created. The best options for each individual setup (automatic, semiautomatic or manual) can then be determined through experience.

The accuracy of the algorithm is of course inferior to what has been achieved when the radiation fields are matched. However, as it can be seen in Fig. 4.3 some deviations in the treatment set up can be easily recognized as a mismatch of the radiation fields above the registered anatomical structures, thus facilitating the process of decision making in the course of radiation therapy.

The fact that registration of portal images and DRRs was possible suggests that some regions of the anatomy are amenable to automatic registration even though the in-phantom scatter is not taken into account. In fact, as a reference image, a DRR is of better utility compared to any portal image since the anatomy position with respect to the radiation field is precisely known. When the first portal image is registered to a conventional simulator

image, the registered portal image becomes the gold standard. However, an error in the registration would be carried on systematically to subsequent measurements. Our work suggests that the megavoltage DRR can be registered with each portal image directly thus removing this systematic error and obviating the need of conventional simulator images for registration purposes. This emphasizes the added importance of CT simulation in modern radiation oncology. Another important consideration is that an anatomical feature for matching not obscured by external structures (for example, the mounting of the shield-supporting trays) can always be selected in the reference image. Further investigations are, however, needed to determine the accuracy of the registration as a function of the parameters of the CT data.

To summarize, we have developed a grey-level image alignment algorithm based on cross-correlation that takes into account both in-plane rotations and translations.‡‡ The search for the in-plane rotation is made possible in a reasonable time frame by using a sequential approach and the FFT-implementation of the Pearson correlation coefficient, which proved to be a better option for matching than normalized cross-correlation. The algorithm can also be used for registration of portal images and DRRs, given that certain modifications are done on the DRRs, thus making the use of a portal image as a reference image unnecessary. The algorithm has the potential to be a tool in fast and automated approaches to images registration and patient setup verification.

‡‡ After the appearance of this work in Medical Physics, similar developments were published by Dong and Boyer.[35]

References

[1] J. Bijhold, K. G. A. Gilhuijs, M. van Herk, and H. Meertens. Radiation field edge detection in portal images. *Phys. Med. Biol.*, 36:1705–1710, 1991.

[2] J. Bijhold, K.G.A. Gilduijs, and M. van Herk. Automatic verification of radiation field shape using digital portal images. *Med. Phys.*, 19:1007–1014, 1992.

[3] K. G. A. Gilhuijs and M. van Herk. An automatic in-line inspection of patient setup in radiation therapy using digital portal images. *Med. Phys.*, 20:667–677, 1993.

[4] I. Crooks and B.G. Fallone. Contrast enhancement of portal images by selective histogram equalization. *Med. Phys.*, 20:199–204, 1993.

[5] H. Wang and B.G. Fallone. A robust morphological algorithm for automatic radiation field-extraction and correlation of portal images. *Med. Phys.*, 21:237–244, 1994.

[6] H. Wang and B.G. Fallone. A mathematical model of radiation field edge localization. *Med. Phys.*, 22:1107–1110, 1995.

[7] S. Zhou and L.J. Verhey. A robust method of multileaf collimator (mlc) leaf-configuration verification. *Phys. Med. Biol.*, 39:1929–1948, 1993.

[8] K. Eilertsen, A. Skretting, and T. L. Tennvassås. Methods for fully automated verification of patient set-up in external beam radiotherapy with polygon shaped fields. *Phys. Med. Biol.*, 39:993–1012, 1994.

[9] K. W. Leszczynski, S. Loose, and P. Dunscombe. Segmented chamfer matching for prescription-treatment image registration in radiotherapy. *Phys. Med. Biol.*, 40:83–94, 1995.

[10] H. Meertens, J. Bijold, and J. Strackee. A method for the measurement of field placement errors in digital portal images. *Phys. Med. Biol.*, 35:299–323, 1990.

[11] J. Moseley and P. Munro. A semiatomatic method for registration of portal images. *Med. Phys.*, 21(4):551–558, 1994.

[12] B. J. McParland. Digital portal registration be sequential anatomical matchpoint and image correlations for real-time continuous field alignment verification. *Med. Phys.*, 22(7):1063–1075, 1995.

[13] J.M. Balter, C.A. Pelizzare, and G.T.Y. Chen. Correlation of projection radiographs in radiation therapy using open curve segments and points. *Med. Phys.*, 19:329–334, 1992.

[14] K. G. A. Gilhuijs, A. Touw, M. van Herk, and E. Vijlbrief. Optimization of automatic portal image analysis. *Med. Phys.*, 22(7):1089–1099, 1995.

[15] D. S. Fritsch, E. L. Chaney, A. Boxwala, M. J. McAuliffe, S. Raghavan, A. Thall, and J. R. Earnhart. Core-based portal image registration for automatic radiotherapy treatment verification. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1287–1230, 1995.

[16] K. G. A. Gilhuijs, P. J. H. van de Hen, and M. van Herk. Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Med. Phys.*, 23(3):389–399, 1996.

[17] S. Jones and A. Boyer. Investigation of an FFT-based correlation technique for verification of radiation treatment setup. *Med. Phys.*, 18:1116–1125, 1991.

[18] T. Radcliffe, R. Rajapakshe, and S. Shalev. Pseudocorrelation: A fast, robust, absolute, grey-level image alignment algorithm. *Med. Phys.*, 21:761–769, 1994.

[19] W. R. Nelson, H. Hirayma, and D. W. O. Rogers. The EGS4 code system. Internal report SLAC 265, Stanford Linear Accelerator Center, 1985.

[20] Corey E. Zankowski. Monte Carlo analysis of the 10 MV x-ray beam from a CLINAC-18 linear accelerator. Master's thesis, Medical Physics Dept., McGill University, Montreal, 1994.

[21] E. Storm and H.I. Israel. Photon cross sections from 1 KeV to 100 MeV for elements Z=1 to Z=100. *Nuclear Data Tables*, A7:565–681, 1970.

[22] G.W. Sherouse, K. Novins, and E.L. Chaney. Computation of digitally reconstructed radiographs for use in radiotherapy treatment design. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:651–658, 1990.

[23] B. Siddon. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med. Phys.*, 12(2):252–255, 1985.

[24] L. Dong and A. L. Boyer. Creating digitally reconstructed radiographs to simulate electronic portal inages for automatic image correlation studies. In *3rd International Workshop on Electronic Portal Imaging*, San Francisco, CA, October 7-8, 1994.

[25] L. Dong and A. L. Boyer. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1053–1060, 1995.

[26] E. L. Hall. *Computer image processing and recognition*, page 181. Academic Press, New York, 1st edition, 1979.

[27] R. J. Schalkoff. *Digital Image Processing and Computer Vision*. John Wiley & Sons, Inc., 1989.

[28] R. C. Gonzales and P. Wintz. *Digital Image Processing*. Addison-Wesley, Reading, MA, 2nd edition, 1987.

[29] W. K. Pratt. *Digital Picture Processing*. John Wiley & Sons, Inc., New York, 1978.

[30] R. N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, 1986.

[31] D. L. G. Hill and D. J. Hawkes. Medical image registration using knowledge of adgacency of anatomical structures. *Image Vision Computing*, 12:173–178, 1994.

[32] Richard P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632, 1973.

[33] L. Lemieux, R. Jagor, D. R. Fish, N. D. Kitchen, and D. G. T. Thomas. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Med. Phys.*, 21:1749–1760, 1994.

[34] A. C. Evans, S. Marett, L. Collins, and T. M. Peters. Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems. *Proc. SPIE Med. Image. III Image Process.*, **1092**:264–270, 1989.

[35] L. Dong and A. L. Boyer. A portal image alignment and patient setup verification procedure using moments and correlation techniques. *Phys. Med. Biol.*, 41(4):697–724, 1996.

CHAPTER 5 Summary and Future Developments

5.1 Inverse Treatment Planning

5.1.1 Summary

In the first part of this thesis, two numerical techniques for inverse treatment planning were presented.

The first technique, the active set algorithm is a hybrid optimization routine applicable to any differentiable function of the pencil beam weights. The minimization algorithm combines the constrained steepest descent and the conjugate gradient methods in order to accelerate the design of intensity modulated beams while accounting for the non-negativity constraints imposed on the pencil beam weights. The algorithm attempts to identify the active set at the solution (pencil beam weights that have value of zero at the solution) so as to proceed with unconstrained minimization in the space of the remaining variables employing the conjugate gradient method. If the prediction of the active set is incorrect, a constraint is encountered and the prediction of the active set is updated. Consequently, the algorithm restarts the conjugate gradient method in the new optimization subspace by resetting the current descent direction to be opposite to the local gradient of the objective function. For two popular objectives, our numerical simulations demonstrated that the active set method outperforms the constrained steepest descent in terms of (i) the residual value of the cost function at termination and (ii) the number of iterations required to achieve an objective value lower or equal to that obtained by the constrained steepest descent at termination. Thus, as a tool for inverse treatment planning, the active set algorithm is a viable alternative to the constrained steepest descent method.

The second algorithm, the continuous penalty function method, is an approximate numerical technique for solving large-scale constrained optimization problems. The method is applied to the least-square dose objective of matching zero dose level to healthy tissues. The constraints require that the target dose be within certain levels.

They are accounted for by the introduction of a penalty term. The continuous penalty function method combines the objective and the penalty term into a single function and proceeds by increasing the weight of the penalty term at each iteration step in order to conform closely the final dose distribution with the constraints. During this iterative process which has a well-defined asymptotic point, the optimization technique creates plans with different compromises between target coverage and healthy tissue sparing. These plans can be kept for a physician's consideration. The selection of a particular plan from the iteration sequence automatically indicates the target importance weight which can relate the minimization of the resulting squared dose objective to a clinically relevant optimization in which many treatment factors are implicitly accounted for by the physician. This approach adds some flexibility to the continuous penalty function method for its application to cases where the desired target coverage results in inadmissible overdosage of critical organs.

Being independent of beam energies and modalities, both the active set algorithm and the continuous penalty function method can also optimize these parameters in addition to the beam profiles. For this purpose, for each direction of the beam setup, multiple beams (different energies, type of radiation) are to be employed. After the optimization, the pencil beams pertaining to more beneficial beams will be assigned greater weights than the ones that pertain to beams with little contribution to the improvement of the dose distribution. The latter beams can then be removed without much degradation of the treatment plan.

The inverse treatment planning methods presented in this thesis are independent of both the dose calculation model and the dose sampling technique. Furthermore, both algorithms are applicable to large classes of cost functions. The active set technique can be employed with any differentiable objective. The continuous penalty function method is

robust for convex objectives and convex penalty terms but it can also provide solutions for other types of objective functions.[1, 2] To benefit fully from inverse treatment planning, one must understand the impact of the dose calculation models, of the dose sampling techniques and of the various objectives on both the beam profiles and the optimized dose distributions. These are some of the future research tracks that need to be explored.

5.1.2 Future Work

Dose calculation model An accurate dose calculation algorithm is essential to assure that the optimized dose distributions obtained by intensity modulated beams can be actually delivered. The DSAR model employed in this thesis (Appendix A) does not consider inhomogeneities in the patient volume. Furthermore, its particular implementation restricts the minimum beamlet size to 5 × 5 mm. A better model based on first principle convolution/superposition dose calculations[3–5] must be implemented so that the sensitivity of the optimized dose distributions with respect to lateral scatter, pencil beam resolution as well as inhomogeneities can be investigated. Some insight in the role of these parameters in the design of intensity modulated beams may significantly accelerate the process of inverse treatment planning and radiation delivery for certain treatment sites and objectives.

Dose sampling techniques The dose distributions produced by intensity modulated beams are characterized with steep high-gradient regions and small but very low- or very high-dose domains. Therefore, fine sampling of the dose distributions is necessary for consistent and reliable inverse treatment planning. The adequate density of dose calculation points for evaluation of dose distributions generally depends on the treatment site, the sampling method and the properties of the dose distribution.[6–8] Therefore, further investigations in this direction are warranted.

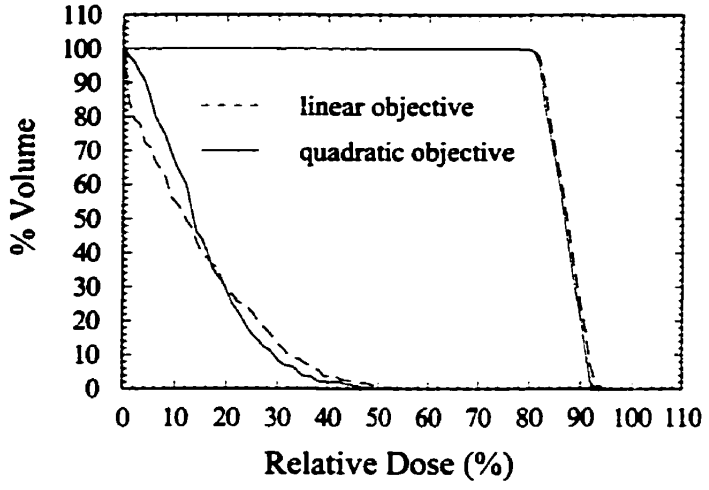

Figure 5.1. *Dose-volume histograms for target and spinal cord volumes for plans created by the continuous penalty function method with linear and quadratic objectives for the same constraint violations as measured by the penalty term. The dose distributions are normalized to their respective maxima. The same importance weights are given to the critical organs for both objectives. For the quadratic objective* $\tau(t)/\mu(t) = 4020$ *at termination. Euler's integration scheme was used for the linear objective.*

Objective functions The least-square objective of matching zero-dose level was used in conjunction with the continuous penalty function to provide continuity with the conventional inverse treatment planning technique and to ensure the applicability of the minimization algorithm. However, other physical objectives can be envisaged and should be explored in order to feed the process of inverse treatment planning with some *a priori* knowledge. For example, a linear objective equal to the weighted sum of the average doses to critical structures penalizes equally for equal dose increases in the high- and the low-dose regions of the critical structures, whereas the quadratic objective penalizes preferentially for dose increase in the high-dose regions of the irradiated organs (Fig. 5.1 and Fig. 5.2). These features of the objective functions guide the optimization process

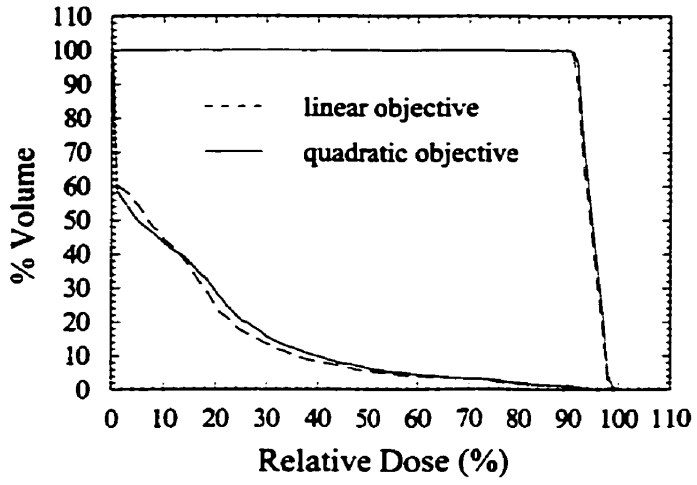

Figure 5.2. *Dose-volume histograms for target and rectal volumes for plans created by the continuous penalty function method with linear and quadratic objectives for the same constraint violations as measured by the penalty term. The dose distributions are normalized to their respective maxima. The same importance weights are given to the critical organs for both objectives. For the quadratic objective* $\tau(t)/\mu(t) = 1757$ *at termination. Euler's integration scheme was used for the linear objective.*

to produce different dose distributions and different sparing of the critical organs for the same constraint violation (Fig. 5.1 and Fig. 5.2). Therefore, one may consider the use of a quadratic objective for nearly serial organs that can tolerate high doses only in very small, insignificant volumes. Typical examples are the spinal cord[9] and the rectum.[10] On the other hand, a linear objective may be better suited for organs with large volume effects such as the lungs[10] which can be kept functional by sacrificing a considerable part of their volume while keeping the rest to very low dose levels. These considerations underline once again that the continuous improvement in the modelling of the treatment objectives is an important research direction to be explored.

5.2 Anatomy Registration for Treatment Setup Verification

5.2.1 Summary

The gray-level image correlation algorithm, presented in the second part of this thesis (Chapter 4), was designed with the goal of automatic quantification of the two-dimensional displacements of the patient anatomy with respect to the radiation field. Under the assumption of rigid transformations and image similarity (up to a linear transformation of the intensities), the algorithm searches for the transformation parameters which produce the highest correlation value between the reference image and the registered one. Certain properties of the correlation operators were explored and consequently exploited to accelerate the image registration. First, a frequency domain representation of the Pearson correlation coefficient was derived which allowed its fast Fourier transform implementation. Second, the value of the maximum of the correlation distribution was shown to be a unimodal function of the angle of rotation between the pair of images to be registered, with a maximum corresponding to the proper image alignment. These properties, along with the property of the correlation maximum to indicate the approximate position of the image feature under small rotations ($\pm 15^\circ$) were used to design the registration algorithm as a two-stage sequential procedure. At the first stage, the approximate value of the displacement between the two images is obtained from the detected and the expected positions of the maximum of the correlation distribution which is calculated with the fast Fourier transform implementation. At the second stage of the algorithm, a smaller search window in the vicinity of the previously located maximum is selected and a search for both the rotational and the translational parameters is conducted. The accuracy of the registration procedure for our phantom study was in the order 1 mm and 1° with typical execution times in the order of a minute on a DEC 3000–M300 computer.

5.2.2 Future Work

Imaging parameters and treatment sites The results from the application of the correlation algorithm to the registration of portal images and megavoltage DRRs indicates that certain regions of the anatomy were amenable to automatic matching even though the in-phantom scattered radiation was not modelled. The success of the Pearson correlation coefficient, which is invariant under a shift and scaling of the image intensities, implies that for our particular setup the scatter contribution across the image is a slowly varying (almost constant) function of the image coordinates, which is added to the primary image signal. This speculation is consistent with the failure of the normalized cross-correlation, which is invariant only under scaling of the image intensities. Therefore, the performance of the algorithm for the cases of portal-to-DRR registration should improve when scatter from the patient is reduced by acquiring the portal image at a large patient-to-detector distance. On the other hand, given the fixed size of the portal imaging detectors (electronic portal imaging devices or film-screen detectors) the patient-to-detector distance is often determined by the size of patient anatomy that has to be imaged. Therefore, for each intended treatment site, phantom studies at the maximum possible patient-portal imager separation should be conducted to investigate the performance of the registration algorithm for the various anatomical features that are envisaged as correlation templates. The identification of the anatomical landmarks, however, needs to be done only once. Furthermore, their selection on the reference image of the patient to be treated can be performed very fast simply by placing (on a computer screen) predefined square drawings at the proper locations. This procedure needs to be executed only once at the beginning of the treatment sessions.

Extensions and applications The registration algorithm relies on the matching of unique, asymmetric and prominent anatomical landmarks. Its robustness for anatomy

alignment can be further improved by combining the information obtained from the registration of several anatomical features. The transformation parameters for each landmark can be found by the technique in its present form and a global transformation based on the properly weighted (e. g. with the values of the correlation maxima) local transformations can be designed to register the portal image and the reference one.

A challenging problem is the extension of the present technique for 3D verification of the patient setup, especially when out of plane rotations are presented. Lemieux *et al.*[11] reported the application of a correlation technique to the 3D registration of CT data to high-quality diagnostic radiographs. However, the reported computational times ($\sim$ 30 min) were impractical for on-line portal imaging. Furthermore, given the portal image quality, the feasibility of their correlation approach[11] in the context of portal imaging needs to be investigated.

However, when only translations are present, our correlation algorithm can align the patient anatomy (represented by the 3D CT data) in the desired treatment position by registering a pair of orthogonal portal images to the corresponding DRRs. This approach may also evaluate accurately the actual 3D displacement when small out-of-plane rotations in the order of $\sim 2^{o}$ are presented.[12] Such small out-of-plane rotations of the bony anatomy are typical for the conformal treatments of patients with prostate cancer after immobilization with custom thermoplastic body cast.[13] Furthermore, given the small volume of the target volume, a correction of the small out-of-plane rotations may not be necessary.[13] Therefore, the 3D anatomy verification of such treatments is a potential application of our registration tool. Furthermore, the algorithm can be used for detecting and quantifying setup errors in retrospective studies on setup uncertainties and target motion. The subsequent findings should then be incorporated in the treatment planning process[14, 15] and properly accounted for in the design of conformal treatments.

References

1 A. A. Brown and M. S. Bartholomew-Biggs. ODE versus SQP methods for constrained optimization. *Journal of Optimization Theory and Applications*, 62(3):371–386, 1989.

2 Y. Evtushenko and V. Zhadan. Private communication, 1996.

3 T. R. Mackie, J. W. Scrimger, and J. J. Battista. A convolution method of calculating dose for 15-MV x-rays. *Med. Phys.*, 12:188–196, 1985.

4 A. L. Boyer and E. C. Mok. A photon dose distribution model employing convolution calculations. *Med. Phys.*, 12:169–177, 1985.

5 R. Mohan, C. Chui, and L. Lidofsky. Differential pencil beam dose computation model for photons. *Med. Phys.*, 13:64–73, 1986.

6 A. Jackson, R. Mohan, and B. Baldwin. Comments on "Sampling techniques for the evaluation of treatment plans". *Med. Phys.*, (5):1375–1376, 1993.

7 X.-Q. Lu and L. M. Chin. Further discussions on sampling methods-a response to Letters to the Editor [Med. Phys. 20, 1375-1376 (1993)] and [Med. Phys. 20, 1377-1380 (1993)]. *Med. Phys.*, (5):1375–1376, 1993.

8 M. Langer, R. Brown, P. Kijewski, and H. Chul. The reliability of optimization under dose-volume limits. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:529–538, 1993.

9 S. Söderstrom, A. Gustafsson, and A. Brahme. The clinical value of different treatment objectives and degrees of freedom in radiation therapy optimization. *Radiotherapy and Oncology*, 29:148–163, 1993.

10 R. Mohan, X. Wang, T. Bortfeld, A. L. Boyer, G. J. Kutcher, S. A. Leibel, Z. Fuks, and C. C. Ling. The potential and limitations of the inverse radiotherapy technique. *Radiotherapy and Oncology*, 32:232–248, 1994.

11 L. Lemieux, R. Jagor, D. R. Fish, N. D. Kitchen, and D. G. T. Thomas. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Med. Phys.*, 21:1749–1760, 1994.

12 S. Jones and A. Boyer. Investigation of an FFT-based correlation technique for verification of radiation treatment setup. *Med. Phys.*, 18:1116–1125, 1991.

[13] J. Hanley, M. A. Lumley, G. S. Mageras, J. Sun, J. Zelefsky, S. A. Leibel, Z. Fuks, and G. J. Kutcher. Measurement of patient positioning errors in three-dimensinal conformal radiotherapy of the prostate. *Int. J. Radiat. Biol. Oncol. Phys.*, 27(2):435–444, 1997.

[14] E. Fontenla, C. A. Pelizzari, and G. T. Chen. Implications of 3-dimensional target shape and motion in aperture design. *Med. Phys.*, 23(8):1431–1441, 1996.

[15] A. Bel, M. van Herk, and J. V. Lebesque. Target margins for random geometrical treatment uncertainties in conformal radiotherapy. *Med. Phys.*, 23(9):1537–1545, 1996.

APPENDIX A DSAR Dose Calculation Model for Inverse Planning

Inverse treatment planning requires dose calculation algorithms that model the effect of beam intensity modulation on the dose distributions. A short description of the dose model used throughout this work is presented below.

The dose $D^p(x, y, d, F_p)$ delivered to a calculation voxel P lying at a physical depth d (Fig. A.1), off-axis position (x, y) in the beam's eye-view plane and distance F_p from the source is[1]:

$$D(x,y,d,F_p) = \frac{D^n(d_n, F_n; W_X, W_Y)}{TAR(d_n, W_X, W_Y)} \left(\frac{F_n}{F_p}\right)^2 \frac{1}{w_{norm}} [f(x,y,d,F_p) TAR(d,0,0) + \\ + \sum_{i=1}^{I} \sum_{j=1}^{J} f\left(x_i, y_j, d'_{ij}, F_p\right) \frac{\Delta^2 SAR\left(x'_i - x, y'_j - y, d'_{ij}\right)}{\Delta x \Delta y} \Delta x_i \Delta y_j] \,. \tag{A.1}$$

In the above equation, the first term in the brackets accounts for the primary dose and the second one for the scattered dose from the different volume elements. The differential scatter-air ratio $\Delta^2 SAR/\Delta x \Delta y$ describes the amount of scatter from pencil beams such as the one shown in Fig. A.1. The scattered dose contributed to a voxel by a pencil beam is determined by the distance $r_{ij} = \sqrt{(x_i - x)^2 + (y_j - y)^2}$ between the calculation voxel and the pencil beam and the depth d'_{ij} of the pencil beam which is determined by the curvature of the entrance surface (Fig. A.1). Values for the DSAR function can be obtained either by differentiating tissue-air ratios[2] (TARs) for rectangular fields or by differentiating scatter-air ratios[2] (SAR) with respect to radius :

$$\frac{\Delta^2 SAR\left(x'_i - x, y'_j - y, d\right)}{\Delta x \Delta y} = \\ = \frac{1}{2\pi r_{ij} \Delta r} \left[SAR\left(r_{ij} + \frac{\Delta r}{2}, d\right) - SAR\left(r_{ij} - \frac{\Delta r}{2}, d\right) \right] \tag{A.2}$$

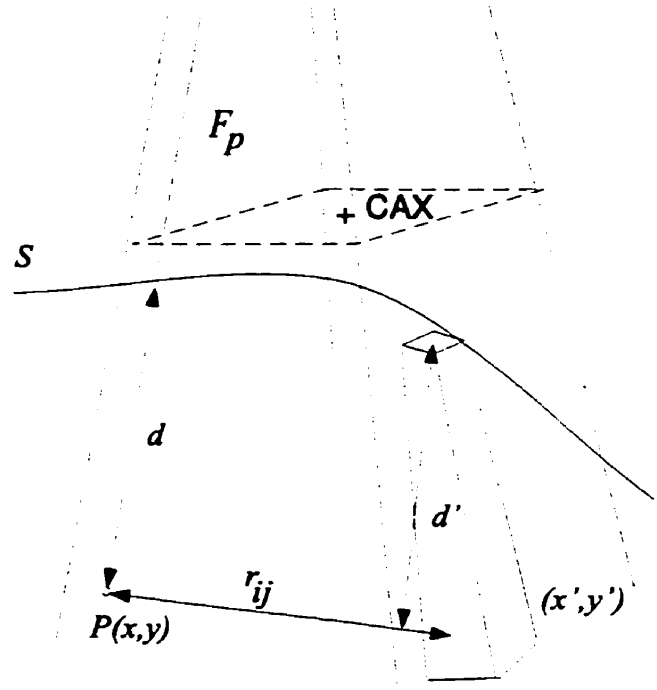

Figure A.1. *The dose to a voxel $P(x,y)$ is calculated by summing the weighted scattered dose contributions from all columns of tissue such as the one shown. The dose scattered by such a column is a function of its distance r_{ij} to the voxel of interest and the depth d'.*

given that the increments $\Delta x, \Delta y$ are small. The function $f(x,y,d,F_p)$ describes the fluence distribution. It can be represented as $f(x,y,d,F_p) = g(x,y,d,F_p)w(x,y)$ where $g(x,y,d,F_p)$ describes the radiation field at each voxel in the patient accounting for the intensity distribution of the radiation source and penumbra effects. The transmission of a filter is given by $w(x,y)$ and the attenuation of the beam along the central axis due to the filter is accounted by $1/w_{norm}$. The dose to the normalization point lying on the central axis of the beam at a depth d_n at a source-to-normalization point distance F_n is $D^n(d_n, F_n, W_X, W_Y)$. The tissue-air ratio for the depth of normalization is given

by $TAR(d_n, W_X, W_Y)$ where the field size at the normalization point defined by the collimator jaws is $W_X \times W_Y$. When the normalization and the calibration points coincide (a typical case being $F_n = SAD$, $d_n = d_{max}$) the normalization dose is given as :

$$D^n(d_n, F_n, W_X, W_Y) = MU \times OF_{med}(W_X, W_Y) \tag{A.3}$$

where $OF_{med}(W_X, W_Y)$ is the output factor at the depth of dose maximum d_{max} for the given field size and MU is the number of the delivered monitor units (beam on time).

Certain preliminary calculations are to be done that account for the fact that tissue-air ratios are not readily available for megavoltage linear accelerators. Instead, the beam data are obtained by measurements in phantom and normalized dose functions as tissue-phantom ratio (TPR) or tissue-maximum ratio (TMR) are evaluated.[3] TARs can be calculated from the corresponding TMRs by :

$$TAR(d, W_X, W_Y) = \frac{TMR(d, W_X, W_Y)}{TAR(d_{max}, W_X, W_Y)} = \frac{TMR(d, W_X, W_Y)}{BSF(W_X, W_Y)} \tag{A.4}$$

where $BSF(W_X, W_Y)$ is the back scatter factor for the corresponding field size. Once the TARs are calculated, the corresponding SARs are given as :

$$SAR(r, d) = TAR(r, d) - TAR(0, d) \; . \tag{A.5}$$

Within the DSAR model, inverse treatment planning can be performed by introducing pencil beam weights $w(x, y)$ in monitor units (MU) and setting the normalization (wedge) factor $1/w_{norm}$ to one. The resulting beam profiles determine the beam-on time for each portion of the beam. However, when a sequence of static fields defined by a multileaf collimator is used for the delivery of the profiles, the change of the output with the field size is to be accounted for.

References

[1] J. R. Cunningham and J. J. Battista. Calcualtion of dose ditributions for x-ray therapy. *Physics in Canada*, 51:190–195, 1995.

[2] J. R. Cunningham. Scatter-air ratios. *Phys. Med. Biol.*, 7:45–51, 1972.

[3] H. E. Johns and J. R. Cunningham. *The physics of radiology*. Charles C. Thomas, Springfield, Illinois, 1983.

APPENDIX B

Applicability of the CPFM to the Inverse Treatment Planning Problem

B.1 Problem Statement

A non-linear constrained optimization problem is written as:

$$\min_{x \in X} f(x), \quad X = \{x \in E_n : g(x) \geq 0\}. \tag{B.1}$$

Here, E_i is the i-dimensional Euclidian space, $x = \left[x^1, ..., x^n\right]^T \in E_n$, $f(x)$ is the function to be minimized, X is the set of admissible values for x determined by the constraints $g(x)$. A local solution x^* is a point in a neighborhood of which, there is no other point satisfying the constraints that gives a smaller value of the objective function. A global solution can be defined as any local solution that yields the smallest objective function value. We denote by X^* the set of all solutions x^* of the minimization problem (solution set) and we assume that X^* is not empty.

A quadratic optimization problem with linear inequality constraints is given by:

$$\min_{x \in X} \|Ax - b\|^2, \quad X = \{x \in E_n : g(x) = Bx - c \leq 0\}. \tag{B.2}$$

Here $b \in E_m$, $c \in E_l$, A is an $m \times n$ matrix and B is an $l \times n$ matrix. In Eq. (B.2), $\|Ax - b\|$ is the usual Euclidian norm of the vector $Ax - b$ and $g(x) = \left[g^1(x), ..., g^l(x)\right]^T$ is a linear vector function, that is for any $\alpha, \beta \in E_1$, $x, y \in E_n$, $g(\alpha x + \beta y) = \alpha g(x) + \beta g(y)$.

Our purpose is: (i) to reassert that Eq. (B.2) poses a convex programming problem and (ii) to prove that the continuous penalty function method solves that problem for the particular case of inverse treatment planning, where $b = 0$, c has finite components and the elements of the matrices A and B are real positive numbers. To this end we first cite some definitions and results from the theory of mathematical programming.

B.2 Convexity — Definitions and Properties

Definition 1. A function $f(x) : E_n \rightarrow E_1$ is a convex function of x in E_n if for every two points $x_1, x_2 \in E_n$ and every λ, where $0 \leq \lambda \leq 1$

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)\,. \tag{B.3}$$

If $f(x)$ is a convex function of x in E_n we simply say that $f(x)$ is a convex function of x.

Definition 2. A function $g(x) : E_n \rightarrow E_1$ is a concave function of x in E_n if $-g(x)$ is a convex function.

As a corollary of the above definitions linear functions are both convex and concave.

Lemma 1. If $f_1, ..., f_p$ are convex functions then $f(x) = \sum_{i=1}^{p} f_i(x)$ is a convex function.

The statement of the lemma follows immediately from the definition of a convex function.

Definition 3. A symmetric matrix A is said to be positive semidefinite (nonnegative definite) if for every vector y, $y^T A y \geq 0$.

Let $\nabla^2 f(x)$ denotes a square symmetric matrix of the order n whose $(i,j)^{\text{th}}$ element is $\frac{\partial^2}{\partial x^i \partial x^j} f(x)$. The matrix $\nabla^2 f(x)$ is also called the *Hessian matrix* of $f(x)$. The following theorem holds true.[1]

Theorem 1. If the function $f(x)$ is twice differentiable in E_n, then $f(x)$ is convex if $\nabla^2 f(x)$ is semidefinite everywhere in E_n.

Definition 4. The convex programming problem is written as

$$\min_{x \in X} f(x), \quad X = \{x \in E_n : g(x) \geq 0\} \tag{B.4}$$

where $f(x)$ is a convex function and each component $g^i(x)$ of the vector function $g(x)$ is a concave function.

For convex programming problems the following theorem holds true.[2]

Theorem 2. [Local-Global Convexity Property] Every local minimum x^* of the convex programming problem is a global minimum.

B.3 Continuous Penalty Function Method

Within the frame of the continuous approach one solves the constrained optimization problem (Eq. B.1) by introducing a penalty term $S(x)$ such that $S(x) = 0$ if $x \in X$ and $S(x) > 0$ otherwise. A new function $P(x,t)$ is defined as

$$P(x,t) = \mu(t)f(x) + \tau(t)S(x) \tag{B.5}$$

where $\mu(t)$ and $\tau(t)$ are continuous functions of the scalar argument t such that

$$\mu(t) > 0, \quad \tau(t) > 0, \quad \int_0^\infty \mu(t)dt = \infty. \tag{B.6}$$

A set $G(t)$ is defined such that

$$G(t) = \{x \in E_n : P(x,t) \leq \mu(t)f(x^*)\}. \tag{B.7}$$

The following theorem holds true:[3, 4]

Theorem 3. Let $f(x)$ and $S(x)$ be convex continuously differentiable functions everywhere in E_n, let the set $G(0)$be non-empty and bounded, let the continuous functions $\mu(t)$ and $\tau(t)$ satisfy conditions (B.6), and let the ratio $\frac{\mu(t)}{\tau(t)} \to 0$ monotonically as $t \to \infty$. Then the set of limit points (as $t \to \infty$) of the solution $x(x_0, t)$ of the Cauchy problem

$$\frac{dx(t)}{dt} = -\nabla_x P(x,t), \quad x(0) = x_0 \tag{B.8}$$

is non-empty and all the limit points belong to X_* whatever the initial condition x_0.

B.4 Local-Global Convexity Property of the Inverse Treatment Planning Problem

A remarkable property of the problem stated by Eq. (B.2) is that any local solution is also a global solution. Indeed, introducing a vector function $h(x)$ such that $h^i(x) = -g^i(x)$ the constrained optimization problem [Eq. (B.2)] can be rewritten as:

$$\min_{x \in X} \|Ax - b\|^2, \quad X = \{x \in E_n : h(x) \geq 0\} \tag{B.9}$$

The Hessian $A^T A$ of the twice differentiable objective function $\|Ax - b\|^2$ is a symmetric positive semidefinite matrix, since for every vector y, $y^T A^T A y = (Ay)^T (Ay) = \|Ay\|^2 \geq 0$ (Definition 3). Therefore, according to Theorem 1 the objective function is convex. The components of the vector function $h(x)$ are linear and therefore concave functions of x. Thus the problem posed by Eq. (B.9) is a convex programming problem (Definition 4) and therefore every local minimum is a global minimum (Theorem 2).

The inverse treatment planning problem is given by Eq. (B.9) and a particular choice of the parameters b, c, A and B. Therefore every local minimum of the inverse treatment planning problem is a global minimum. If one considers the case of zero weights to the critical structures and relaxed requirements with respect to the target one can demonstrate that several global minima of different clinical utility are possible.

B.5 The Inverse Treatment Planning Problem

We have shown in the previous section that the objective $f(x)$ is in accordance with the requirements of Theorem 3 since it is continuously differentiable and convex.

In order to use the continuous penalty function method[3] we form a convex penalty term

$$S(x) = \sum_{i=1}^{l} \left(g_i^+(x)\right)^2 \tag{B.10}$$

to account for the constraints. The functions $g_i^+(x)$ are defined as $g_i^+(x) \stackrel{def}{=} max[0, g_i(x)]$. We prove that $\left(g_i^+(x)\right)^2$ is a convex function of x and therefore $S(x)$, being a sum of convex functions, is also convex (Lemma 1). Omitting the index i, we verify that $\left(g^+(x)\right)^2$ complies with the requirements of Definition 1, that is

$$\left(g^+(\lambda x_1 + (1-\lambda)x_2)\right)^2 \leq \lambda\left(g^+(x_1)\right)^2 + (1-\lambda)\left(g^+(x_2)\right)^2 \tag{B.11}$$

for every λ, $0 \leq \lambda \leq 1$ and any x_1, x_2. All possibilities are considered below:

I. [$\lambda = 0$ **or** $\lambda = 1$, **any** x_1, x_2]

Let us consider the case $\lambda = 0$. Equation (B.11) reads:

$$\begin{aligned}\left(g^+(\lambda x_1 + (1-\lambda)x_2)\right)^2 &= \left(g^+(x_2)\right)^2 \\ &= \lambda\left(g^+(x_1)\right)^2 + (1-\lambda)\left(g^+(x_2)\right)^2\end{aligned} \tag{B.12}$$

since for positive λs

$$g^+(\lambda x) = max[0, g(\lambda x)] = max[0, \lambda g(x)] = \lambda g^+(x). \tag{B.13}$$

Similarly it can be shown that the inequality in (Eq. B.11) is satisfied for the case $\lambda = 1$.

II. $[0 < \lambda < 1,\ x_1, x_2 : g(x_1) \geq 0,\ g(x_2) \geq 0]$

For this case

$$g^+(\lambda x_1 + (1-\lambda)x_2) = \lambda g(x_1) + (1-\lambda)g(x_2). \tag{B.14}$$

The left side of Eq. (B.11) reads:

$$\begin{aligned}\left(g^+(\lambda x_1 + (1-\lambda)x_2)\right)^2 &= (\lambda g(x_1) + (1-\lambda)g(x_2))^2 = \\ &= \lambda^2 g^2(x_1) + 2\lambda(1-\lambda)g(x_1)g(x_2) + (1-\lambda)^2 g^2(x_2).\end{aligned} \tag{B.15}$$

After rearranging all terms of Eq. (B.11) on the left, one has to verify:

$$\lambda(\lambda-1)g^2(x_1) + 2\lambda(1-\lambda)g(x_1)g(x_2) - (1-\lambda)\lambda g^2(x_2) \stackrel{?}{\leq} 0 \tag{B.16}$$

After division by $\lambda(1-\lambda)$, $(\lambda(1-\lambda) > 0)$ we have

$$-g^2(x_1) + g(x_1)g(x_2) - g^2(x_2) = -(g(x_1) - g(x_2))^2 \leq 0 \tag{B.17}$$

and therefore Eq. (B.11).

III. $[0 < \lambda < 1,\ x_1, x_2 : g(x_1) < 0,\ g(x_2) < 0]$

For this case Eq. (B.11) turns into the trivial equality $0 = 0$ and therefore holds true.

IV. $[0 < \lambda < 1,\ x_1, x_2 : \lambda g(x_1) + (1-\lambda)g(x_2) \leq 0,\ g(x_1) \leq 0, g(x_2) \geq 0$ **or** $g(x_1) \geq 0, g(x_2) \leq 0]$

For this case Eq. (B.11) reads: $0 \leq (1-\lambda)g^2(x_2)$ or $0 \leq \lambda g^2(x_1)$ and therefore holds true.

V. a. $[0 < \lambda < 1,\, x_1, x_2 : \lambda g(x_1) + (1-\lambda)g(x_2) \geq 0,\, g(x_1) \leq 0, g(x_2) \geq 0]$

For this case Eq. (B.11) reads:

$$\begin{aligned}\left(g^+(\lambda x_1 + (1-\lambda)x_2)\right)^2 &= \left(\lambda g(x_1) + (1-\lambda)g(x_2)\right)^2 \\ &= \lambda^2 g^2(x_1) + 2\lambda(1-\lambda)g(x_1)g(x_2) + (1-\lambda)^2 g^2(x_2) \\ &\overset{?}{\leq} (1-\lambda)g^2(x_2)\end{aligned} \tag{B.18}$$

Since $(1-\lambda)^2 g^2(x_2) \leq (1-\lambda)g^2(x_2)$ it suffices to show that

$$\lambda^2 g^2(x_1) + 2\lambda(1-\lambda)g(x_1)g(x_2) \overset{?}{\leq} 0. \tag{B.19}$$

Dividing by $\lambda g(x_1)$ ($\lambda g(x_1) < 0$) and rearranging the resulting terms one confirms that

$$\underbrace{\lambda g(x_1) + (1-\lambda)g(x_2)}_{\geq 0} + \underbrace{(1-\lambda)g(x_2)}_{\geq 0} \geq 0. \tag{B.20}$$

and therefore Eq. (B.11) holds true for this case.

VI. a. $[0 < \lambda < 1,\, x_1, x_2 : \lambda g(x_1) + (1-\lambda)g(x_2) \geq 0,\, g(x_1) \geq 0, g(x_2) \leq 0]$

For this case Eq. (B.11) reads:

$$\begin{aligned}\left(g^+(\lambda x_1 + (1-\lambda)x_2)\right)^2 &= \left(\lambda g(x_1) + (1-\lambda)g(x_2)\right)^2 \\ &= \lambda^2 g^2(x_1) + 2\lambda(1-\lambda)g(x_1)g(x_2) + (1-\lambda)^2 g^2(x_2) \\ &\overset{?}{\leq} \lambda g^2(x_1)\end{aligned} \tag{B.21}$$

Since $\lambda^2 g^2(x_1) \leq \lambda g^2(x_1)$ it suffices to show that

$$2\lambda(1-\lambda)g(x_1)g(x_2) + (1-\lambda)^2 g^2(x_2) \overset{?}{\leq} 0. \tag{B.22}$$

Dividing by $(1-\lambda)g(x_2)$ ($(1-\lambda)g(x_2)$) and rearranging the resulting terms one confirms that

$$\underbrace{\lambda g(x_1) + (1-\lambda)g(x_2)}_{\geq 0} + \underbrace{\lambda g(x_1)}_{\geq 0} \geq 0. \tag{B.23}$$

and therefore Eq. (B.11) holds true for this case.

The penalty term $S(x)$ can be written as

$$S(x) = \sum_{i=1}^{l} \varphi(g_i(x)) \tag{B.24}$$

where

$$\varphi(g) = \begin{cases} g^2, & g \geq 0 \\ 0, & g < 0 \end{cases}. \tag{B.25}$$

The function $\varphi(g)$ is continuously differentiable with respect to g since

$$\varphi'(g) = \begin{cases} 2g, & g > 0 \\ 0, & g = 0 \\ 0, & g < 0 \end{cases}. \tag{B.26}$$

The linear functions $g_i(x)$ are continuously differentiable with respect to the components of the vector x and therefore the composite penalty term $S(x)$ is a continuously differentiable with respect to x.

The weighting functions used in this study are of the form

$$\mu(t) = const > 0, \quad \tau(t) = exp(t),\ 1 + t + t^2, ... \tag{B.27}$$

and therefore satisfy the requirements of Theorem 3.

Let consider the requirement of Theorem 3 that concerns the set $G(0)$. First we show that if a solution x^* exists (that is the case if feasible target dose levels are specified), the set $G(0)$ is not empty since $x^* \in G(0)$. Indeed, for the functions used in this thesis,

$$P(x, 0) = \mu(0)f(x) + \tau(0)S(x) = const \times f(x) + S(x) \tag{B.28}$$

and therefore $G(0)$ is given by

$$G(0) = \{x \in E_n : S(x) \leq const \times (f(x^*) - f(x))\}. \tag{B.29}$$

For the solution x^* the equality holds true since

$$S(x^*) = const \times (f(x^*) - f(x^*)) = 0 \tag{B.30}$$

and therefore $x^* \in G(0)$.

The requirement that $G(0)$ be bound implies that there exists a large but finite number N such that $\|x\| < N$ for any vector $x \in G(0)$. Therefore, one requires that there be no vectors x with infinite components that pertain to $G(0)$, that is there be no vectors x with infinite components such that

$$S(x) \leq const \times (f(x^*) - f(x)). \tag{B.31}$$

The functions $S(x)$, $f(x)$ are non-negative. Therefore, Eq. (B.31) can hold true only if

$$0 \leq S(x) \leq const \times f(x^*). \tag{B.32}$$

However, for a vector with one or several infinite components x^i, Eq. (B.32) cannot hold true since the penalty term always tends to infinity. Only positive or only negative infinite components would cause $S(x) \to \infty$ due to the violation of the dose constraints. Infinitely large positive and negative components could lead to conformity with the dose constraints but the penalty term again tends to infinity due to the violation of the non-negativity constraints, since

$$S(x) \propto (x^i)^2 \underset{x^i \to -\infty}{\to} \infty. \tag{B.33}$$

The above argument demonstrates that for the particular case of inverse treatment planning the set $G(0)$ is bound and non-empty if feasible dose levels for the target are specified.

To summarize, for the particular formulation of the inverse treatment planning proposed in this work, the objective function $f(x)$ and the penalty term $S(x)$ are convex continuously differentiable functions of x, the weighting functions $\mu(t)$, $\tau(t)$ satisfy the

requirements of Theorem 3 and the set $G(0)$ is bound and non-empty. Therefore, the continuous penalty function method proposed in Theorem 3 is applicable to the inverse treatment planning problem and the method can be used for finding approximate solutions.

References

1 Y. G. Evtushenko. *Numerical Optimization Techniques*, page 25. Optimization Software, Inc., Publication Division, New York, 1st edition, 1985.

2 Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, pages 89–90. John Wiley and Sons, Inc, 1968.

3 Y. G. Evtushenko. Numerical methods for solving problems of non-linear programming. *Zh. výchisl. Mat. mat. Fiz.*, 16(2):307–324, 1976.

4 Y. G. Evtushenko. *Numerical Optimization Techniques*, page 213. Optimization Software, Inc., Publication Division, New York, 1st edition, 1985.

APPENDIX C

FFT-implementation of the Pearson's correlation coefficient

Given a digital image $f(x,y)$ of size $M \times N$ (search window) and a search mask $w(x,y)$ of size $J \times K$, where $J < M$ and $K < N$ the correlation coefficient is defined as

$$r(m,n) = \frac{\sum_x \sum_y [f(x,y) - \bar{f}(x,y)][w(x-m,y-n) - \bar{w}]}{\sqrt{\sum_x \sum_y [w(x-m,y-n) - \bar{w}]^2} \sqrt{\sum_x \sum_y [f(x,y) - \bar{f}(x,y)]^2}} \tag{C.1}$$

where $m = 0, 1, 2, ..., M+J-1$, $n = 0, 1, 2. ..., N+K-1$, $\bar{w}$ is the average intensity of the mask, $\bar{f}(x,y)$ is the average value of $f(x,y)$ in the region coincident with $w(x,y)$, and the summations are carried over the coordinates common to both f and w. The way to interpret the correlation operator given above is the following: the search mask $w(x,y)$ is placed at some point (m,n) of the search window $f(x,y)$ and the value of the correlation at this point is calculated by Eq. (C.1) where the averages and the sums are calculated over the points of overlap between $w(x,y)$ and $f(x,y)$.

Our approach consists of expanding Eq. (C.1) and implementing the resulting terms by FFT-based correlations. According to the correlation theorem if $F(s,t) = \mathcal{F}[f(x,y)]$ is the Fourier transform of f and $W(s,t) = \mathcal{F}[w(x,y)]$ is the Fourier transform of w, then the cross-correlation matrix

$$[f \circledast w]_{m,n} = \sum_x \sum_y f(x,y) w(x-m, y-n) \tag{C.2}$$

can be written through the Fourier transforms of the functions as

$$f \circledast w = \mathcal{F}^{-1}[F(s,t) W^*(s,t)] \tag{C.3}$$

where "*" denotes the complex conjugate and $\mathcal{F}^{-1}$ is the inverse Fourier transform. Let us define two additional functions (binary masks) as

- $p(x,y) = \begin{cases} 1, & \text{if } (x,y) \text{ is within the search mask,} \\ 0, & \text{otherwise} \end{cases}$
- $q(x,y) = 1,\ x = 1..M,\ y = 1..N$

and their Fourier transforms as:

- $P(s,t) = \mathcal{F}[p(x,y)]$ and $Q(s,t) = \mathcal{F}[q(x,y)]$.

The expansion of the numerator of Eq. C.1 leads to the following terms :

$$\sum_x \sum_y f(x,y)w(x-m,y-n) = [f \circledast w]_{m,n} \tag{C.4}$$

and

$$-\sum_x \sum_y f(x,y)\bar{w} = -\bar{w}(m,n)\left[\sum_x \sum_y f(x,y)\right]_{m,n} . \tag{C.5}$$

The average value of the search mask over the common area of f and w is

$$\bar{w}_{m,n} = \bar{w}(m,n) = \frac{1}{S(m,n)}\left[\sum_x \sum_y w(x,y)\right]_{m,n} \tag{C.6}$$

where the area of overlap is

$$S_{m,n} = S(m,n) = [q \circledast p]_{m,n} = \left[\mathcal{F}^{-1}[Q(s,t)P^*(s,t)]\right]_{m,n} \tag{C.7}$$

and the sums over the common area of f and w can be written as

$$\begin{aligned} \left[\sum_x \sum_y w(x,y)\right]_{m,n} &= [q \circledast w]_{m,n} \\ \left[\sum_x \sum_y f(x,y)\right]_{m,n} &= [f \circledast p]_{m,n} . \end{aligned} \tag{C.8}$$

The other two terms in the expansion of the numerator cancel each other and we obtain for the numerator

$$\left[f \circledast w - \frac{(q \circledast w)(f \circledast p)}{q \circledast p}\right]_{m,n} . \tag{C.9}$$

In the above expression as well as in all the subsequent ones the arithmetic operations on the images are done point by point. For the denominator we have:

$$\sqrt{\sum_x \sum_y [w(x-m,y-n)-\bar{w}]^2} =$$
$$= \sqrt{\sum_x \sum_y w^2(x-m,y-n) - 2\bar{w}\sum_x \sum_y w(x-m,y-n) + \sum_x \sum_y \bar{w}^2(m,n)}$$
$$= \sqrt{\sum_x \sum_y w^2(x-m,y-n) - S(m,n)\bar{w}^2(m,n)} = \left[\sqrt{q \circledast w^2 - \frac{(q \circledast w)^2}{q \circledast p}}\right]_{m,n} \quad \text{(C.10)}$$

Similarly,

$$\sqrt{\sum_x \sum_y [f(x,y)-\bar{f}]^2} =$$
$$= \sqrt{\sum_x \sum_y f^2(x,y) - 2\bar{f}(m,n)\sum_x \sum_y f(x,y) + \sum_x \sum_y \bar{f}(m,n)\bar{f}(m,n)} \quad \text{(C.11)}$$
$$= \left[\sqrt{f^2 \circledast p - \frac{(f \circledast p)^2}{q \circledast p}}\right]_{m,n}$$

Thus the FFT-based implementation of the linear correlation coefficient reads:

$$r(m,n) = \left[\frac{(f \circledast w)(q \circledast p) - (q \circledast w)(f \circledast p)}{\sqrt{(q \circledast w^2)(q \circledast p) - (q \circledast w)^2}\sqrt{(f^2 \circledast p)(q \circledast p) - (f \circledast p)^2}}\right]_{m,n} . \quad \text{(C.12)}$$

The above expression gives the value of the correlation coefficient at any location of the search mask over the search window. Clearly the boundary regions (in order of few pixels) are going to produce high value of the correlational coefficient. Therefore, cropping of the correlation matrix r is desirable to avoid a false maximum of the correlation distribution. In the case where one wants to calculate the correlation coefficient only for the cases when the whole search mask is within the search window Eq. (C.12)

can be simplified even further. In this case:

$$\begin{aligned} q \circledast p &= JK\widehat{I} \\ q \circledast w^2 &= \widehat{I}\sum_{x=1}^{J}\sum_{y=1}^{K} w^2(x,y) \\ q \circledast w &= \widehat{I}\sum_{x=1}^{J}\sum_{y=1}^{K} w(x,y) \end{aligned} \tag{C.13}$$

where $\widehat{I}$ is a unitary matrix of size $M-J+1$, $N-K+1$ and the other correlation matrices in Eq. (C.12) are to be cropped to this size. In this case some computational time can be saved by substituting some of the correlations with the expressions in Eq. (C.13).

Bibliography

[1] ICRU Report No. 33: Radiation Quantities and Units. Technical report, Interanational Commision on Radition Units and Measurements, Washington, DC, 1980. **p. 1–2.**

[2] Cancer Trends. In *1987 annual cancer statistics review*, pages II.1–II.203. US Department of Health and Human Services, 1988.

[3] H. Amols and T. Lowinger. An inexpensive microcomputer based port film image enhancement system. *Med. Phys. (abstract)*, 14:483, 1987. **p. 1-23.**

[4] G. K. Bahr, J. G. Kereiakes, H. Horwitz, R. Finney, J. Galvin, and K. Goode. The method of linear programming applied to radiation treatment planning. *Radiology*, 91:686–693, 1968. **pp. 1–14, 1–15.**

[5] J. M. Balter, C. A. Pelizzare, and G. T. Y. Chen. Correlation of projection radiographs in radiation therapy using open curve segments and points. *Med. Phys.*, 19:329–334, 1992. **pp. 1-23, 4–2, 4–25.**

[6] N. H. Barth. An inverse problem in radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:425–431, 1990. **pp. 1–11, 1–16, 2–3.**

[7] A. Bel, M. van Herk, and J. V. Lebesque. Target margins for random geometrical treatment uncertainties in conformal radiotherapy. *Med. Phys.*, 23(9):1537–1545, 1996. **pp. 1–11, 5–9.**

[8] D. Bertsekas and J. Tsitsiklis. *Parallel and Distributed Processing — Numerical Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1989. **p. 3–17.**

[9] J. Bijhold, K.G.A. Gilduijs, and M. van Herk. Automatic verification of radiation field shape using digital portal images. *Med. Phys.*, 19:1007–1014, 1992. **pp. 1-23, 1–42.**

[10] J. Bijhold, K. G. A. Gilhuijs, M. van Herk, and H. Meertens. Radiation field edge detection in portal images. *Phys. Med. Biol.*, 36:1705–1710, 1991. **pp. 1-23, 4–2.**

[11] R. Boellaard, M. van Herk, and B. J. Mijnheer. A convolution model to convert transmission dose images to exit dose distributions. *Med. Phys.*, 24(2):189–199, 1997. **p. 1–11.**

[12] T. Bortfeld, J. Burkelbash, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied to conformation radiotherapy. *Phys. Med. Biol.*, 35:1423–1434, 1990. **pp. 1–11, 1–15, 1–16, 2–4, 3–5, 3–6, 3–7, 3–8, 3–9, 3–16, 3–17.**

[13] T. Bortfeld, D. L. Kahler, T. J. Waldron, and A. L. Boyer. X-ray field compensation with multileaf collimators. *Int. J. radiat. Oncol. Biol. Phys.*, 28:723–730, 1994. **pp. 1–11, 1–16.**

[14] A. L. Boyer, L. Antonuk, A. Fenster, M. van Herk, H. Meertens, P. Munro, L. E. Reinstein, and J. Wong. A review of electronic portal imaging devices (EPID). *Med. Phys.*, 19(1):1–16, 1992. **pp. 1–11, 1–20.**

[15] A. L. Boyer, G. E. Desorby, and N. H. Wells. Potential and limitations of invariant kernel conformal therapy. *Med. Phys.*, 18(4):703–712, 1991. **pp. 1–15, 2–3, 3–5, 3–8, 3–9.**

[16] A. L. Boyer and E. C. Mok. A photon dose distribution model employing convolution calculations. *Med. Phys.*, 12:169–177, 1985. **pp. 1–16, 5–4.**

[17] J. A. Brace, T. J. Davy, D. B. L. Skeggs, and H. S. Williams. Conformation therapy at the royal free hospital: a progress report on the tracking cobalt project. *Br. J. Radiol.*, 54:1068–1074, 1981. **pp. 1–10, 1–12.**

[18] R. N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, 1986. **p. 4–11.**

[19] A. Brahme. Dosimetric precision requirements in radiation therapy. *Acta Radiol. Oncol.*, 23:379–391, 1984. **p. 1–18.**

[20] A. Brahme. Design principles and clinical possibilities for a new generation of radiation therapy equipment. *Acta Oncol.*, 26:403–412, 1987. **pp. 1–11, 1–16.**

[21] A. Brahme. Optimization of stationary and moving beam therapy techniques. *Radioth. Oncol.*, 12:129–140, 1988. **pp. 1–15, 2–3, 3–6.**

[22] A. Brahme. Treatment optimization using physical and radiobiological objective functions. In A. Smith, editor, *Radiation Therapy Physics*, pages 209–246. Springer, Berlin, 1995. **p. 1–15.**

[23] A. Brahme, T Kraepelien, and H. Svensson. Electron and photon beams from a 50 MeV racetrack microtron. *Acta Radiol. Oncol.*, 19:305–319, 1980. **p. 1–11.**

[24] A. Brahme and D. Reisstad. Microtrons for electron and photon radiotherapy. *IEEE Trans. Nucl. Sci.*, NS-28:1880–1883, 1981. **p. 1–11.**

[25] A. Brahme and A. K. Ågren. Optimal dose distributions for eradication of heterogeneous tumors. *Acta Oncol.*, 26:377–385, 1987. **p. 1–2.**

[26] A. Brahme, A. J. Roos, and I. Lax. Solution of an integral equation in rotation therapy. *Phys. Med. Biol.*, 27:1221–1229, 1982. **pp. 1–11, 1–16, 2–3.**

[27] Richard P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632, 1973. **p. 4–13.**

[28] A. A. Brown and M. S. Bartholomew-Biggs. ODE versus SQP methods for constrained optimization. *Journal of Optimization Theory and Applications*, 62(3):371–386, 1989. **pp. 3–12, 3–14, 3–27, 3–29, 3–41, 5–4.**

[29] A. A. Brown and M. S. Bartholomew-Biggs. Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations. *Journal of Optimization Theory and Applications*, 62(2):211–224, 1989. **p. 3–12.**

[30] R. W. Byhardt, J. D. Cox, A. Hornburg, and G. Liermann. Weekly localization films and detection of field placement errors. *Int. J. Radiat. Oncol. Biol. Phys.*, 4:881–887, 1978. **p. 1–20.**

[31] J. Cai, S.-Q. Zhou, J. Hopkinson, V. A. Saxena, and R. S. Cox. Alignment of multi-segmented anatomical features from radiation therapy images by using least square fitting. *Med. Phys.*, 23(12):2069–2075, 1996. **p. 1–23.**

[32] Statistics Canada. Mortality – Summary List of Causes, 1994. Toronto, 1994. Catalogue no. 84-209-XPB, **p. 1–2.**

[33] M. P. Carol. Beam modulation conformal therapy. In *International Symposium on 3D Radiation Treatment Planning and Conformal Therapy*, pages 435–445, St. Louis, Missouri, April 21–23, 1994. Proceedings, **p. 1–11.**

[34] Y. Censor, M. D. Altschuler, and W. D. Powlis. A computational solution of the inverse problem in radiation-therapy treatment planning. *Appl. Math. Comput.*, 25:57–87, 1988. **p. 2–4.**

[35] Z. Chen, X. Wang, T. Bortfeld, R. Mohan, and L. Reinstein. The influence of scatter on the design of optimized intensity modulations. *Med. Phys.*, 22(11 Pt. 1):1727–1733, 1995. **pp. 1–16, 2–3.**

[36] C. K. Chung, J. A. Styker, M. O'Neil, and W. E. DeMuth. Evaluation of adjuvant postoperative radiotherapy for lung cancer. *Int. J. Rad. Oncol. Biol. Phys.*, 8:1877–1880, 1982. **p. 1–8.**

[37] A. M. Cormack. A problem in rotation therapy with x-rays. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:623–630, 1987. **p. 2–3.**

[38] A. M. Cormack and R. A. Cormack. A problem in rotation therapy with x-rays: Dose distributions with an axis of symmetry. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1921–1925, 1987. **p. 2–3.**

[39] I. Crooks and B.G. Fallone. Contrast enhancement of portal images by selective histogram equalization. *Med. Phys.*, 20:199–204, 1993. **pp. 1-23, 4–2.**

[40] I. Crooks and B.G. Fallone. A novel algorithm for the edge-detection and edge-enhancement of medical images. *Med. Phys.*, 20:993–998, 1993. **p. 1-23.**

[41] J. R. Cunningham. Scatter-air ratios. *Phys. Med. Biol.*, 7:45–51, 1972. **pp. 2–21, A–1.**

[42] J. R. Cunningham and J. J. Battista. Calcualtion of dose ditributions for x-ray therapy. *Physics in Canada*, 51:190–195, 1995. **p. A–1**.

[43] I. Daftari, P. L. Petti, J. M. Collier, J. R. Castro, and P. Samuel. The effect of patient motion on dose uncertainty in charged particle irradiation for lesions encircling the brain stem or spinal cord. *Med. Phys.*, 18:1105–1115, 1991. **p. 1–18**.

[44] W. De Neve. An overview of correction strategies for setup errors. In *4th International Workshop on Electronic Portal Imaging*, Amsterdam, The Netherlands, 1996. Proceedings, **p. 1–21**.

[45] W. de Neve, F. van den Heuvel, M. de Beukeeler, M. Coghe, L. Thon, W. de Roover, M. van Lancher, and G. Storme. Routine clinical on-line portal imaging followed by immediate field adjustment using a telecontrolled patient coach. *Radiother. Oncol.*, 24:45–54, 1992. **p. 1–11**.

[46] V. T. DeVita, M. Lippman, S. A. Hubbard, D. C. Idhe, and S. A. Rosenberg. The effect of combined modality therapy on local control and survival. *Int. J. Rad. Oncol. Biol. Phys.*, 12:487–501, 1986. **p. 1–7**.

[47] V.T. DeVita. Progress in cancer management. *Cancer*, 51:2401–2409, 1983. **p. 1–7**.

[48] G. X. Ding, S. Shalev, and G. Gluchev. A ρ-θ technique for treatment verification in radiotherapy and its clinical applications. *Med. Phys.*, 20:1135–1143, 1993. **p. 1–23**.

[49] S. Dische. Radiotherapy in the nineties. *Acta Oncol.*, 31(5):501–511, 1992. **p. 1–7**.

[50] A. Djordjevitch, D. J. Bonham, E. M. A. Hussein, J. W. Andrew, and M. E. Hale. Optimal design of radiation compensators. *Med. Phys.*, 17:397–404, 1990. **pp. 1–15, 2–4**.

[51] J. S. Donald. Computed tomography applied to radiotherapy treatment palnning: techniques and results. *Radiology*, 133:477–482, 1979. **p. 1–10**.

[52] L. Dong and A. L. Boyer. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1053–1060, 1995. **pp. 1–23, 4–6**.

[53] L. Dong and A. L. Boyer. A portal image alignment and patient setup verification procedure using moments and correlation techniques. *Phys. Med. Biol.*, 41(4):697–724, 1996. **p. 4–26**.

[54] L. Dong and A. L. Boyer. Creating digitally reconstructed radiographs to simulate electronic portal inages for automatic image correlation studies. In *3rd International Workshop on Electronic Portal Imaging*, San Francisco, CA, October 7-8, 1994. **p. 4–6**.

[55] R. E. Drzymala, R. Mohan, L. Brewster, J. Chu, M. Goitein, W. Harms, and M. Urie. Dose volume histograms. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:71–78, 1991. **p. 1–10**.

[56] P. Dunscombe. The future of portal verification. In *Joint Meeting of Canadian Association of Radiologists, Canadian Organization of Medical Physicists and the Canadian College of Physicists in Medicine*, page 3, Montreal, Quebec, June 4–7, 1995. CCPM Symposium: Future Trends in Medical Imaging, **p. 1–23**.

[57] P. B. Dunscombe and K. Fox. The precision of determining compliance with prescribed fields from conventional portal films. *Br. J. Radiol.*, 62:48–52, 1989. **pp. 1–20, 1–22, 1–23**.

[58] P. B. Dunscombe and K. Fox. Portal radiography practice in Canada. *Int. J. Radiat. Oncol. Biol. Phys.*, 24:177–179, 1992. **pp. 1–20, 1–22, 1–23**.

[59] A. Dutreix. When and how can we improve precision in radiotherapy. *Acta Radiol. Oncol.*, 2:275–292, 1984. **p. 1–18**.

[60] A. Dutriex. Prescription, precision and decision in treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1291–1296, 1987. **p. 1–14**.

[61] K. Eilertsen, A. Skretting, and T. L. Tennvassås. Methods for fully automated verification of patient set-up in external beam radiotherapy with polygon shaped fields. *Phys. Med. Biol.*, 39:993–1012, 1994. **pp. 1–23, 4–2**.

[62] M. M. Elkind. DNA damage and cell killing: cause and effect? *Cancer*, 56:2351–2363, 1985. **p. 1–2**.

[63] B. Emami, J. Lyman, A. Brown, L. Coia, M. Goitein, J. E. Munzenrider, B. Shank, L. J. Solin, and A. M. Wesson. Tolerance of normal tissue to therapeutic irradiation. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:109–122, 1991. **p. 2–21**.

[64] A. C. Evans, S. Marett, L. Collins, and T. M. Peters. Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems. *Proc. SPIE Med. Image. III Image Process.*, **1092**:264–270, 1989. **p. 4–25**.

[65] P. M. Evans, J. Q. Gildersleve, C. Rawlings, and W. Swindell. Technical note: the implementation of patient position correction using a megavoltage imaging device on a linear accelerator. *Br. J. Radiol.*, 66(789):833–838, 1993. **p. 1–11**.

[66] Y. Evtushenko and V. Zhadan. Private communication, 1996. **pp. 3–43, 5–4**.

[67] Y. G. Evtushenko. Numerical methods for solving problems of non-linear programming. *Zh. výchisl. Mat. mat. Fiz.*, 16(2):307–324, 1976. **pp. 3–11, 3–12, 3–43, B–3, B–5**.

[68] Y. G. Evtushenko. *Numerical Optimization Techniques*. Optimization Software, Inc., Publication Division, New York, 1st edition, 1985. **pp. 3–11, 3–43, B–1**.

[69] Y. G. Evtushenko. *Numerical Optimization Techniques*, page 213. Optimization Software, Inc., Publication Division, New York, 1st edition, 1985. **p. B–3.**

[70] B. G. Fallone, D. H. Hristov, E. B. Podgorsak, and L. Souhami. Digital portal verification of stereotactic radiosurgery. Montreal, Quebec, Sept. 13–17, 1995. Annual Meeting of The Royal College of Physicians and Surgeons of Canada. **p. 1–24.**

[71] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley and Sons, Inc, 1968. **p. 3–28.**

[72] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, page 89. John Wiley and Sons, Inc, 1968. **p. 3–11.**

[73] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, page 126. John Wiley and Sons, Inc, 1968. **pp. 3–12, 3–27.**

[74] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, pages 89–90. John Wiley and Sons, Inc, 1968. **p. B–2.**

[75] J. K. Folkman. Fighting cancer by attacking its blood suply. *Sci. Am.*, 275(3):150–154, 1996. **p. 1–2.**

[76] E. Fontenla, C. A. Pelizzari, and G. T. Chen. Implications of 3-dimensional target shape and motion in aperture design. *Med. Phys.*, 23(8):1431–1441, 1996. **pp. 1–11, 5–6.**

[77] B. A. Fraass. The development of conformal radiation therapy. *Med. Phys.*, 22(11 Pt.2):1911–1921, 1995. **p. 1–12.**

[78] B. A. Fraass, L. Marsh, M. K. Martel, K. Forster, D. L. McShann, and K. TenHanken. Multileaf collimator-based intensity modulation for conformal therapy. *Med. Phys.*, 21:1008, 1994. Abstract, **p. 1–11.**

[79] D. S. Fritsch, E. L. Chaney, A. Boxwala, M. J. McAuliffe, S. Raghavan, A. Thall, and J. R. Earnhart. Core-based portal image registration for automatic radiotherapy treatment verification. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1287–1230, 1995. **pp. 1–23, 4–2.**

[80] Z. Fuks, S. A. Leibel, K. E. Wallner, C. B. Begg, W. R. Fair, L. L. Anderson, B. S. Hilaris, and W. F. Whitmore. The effect of local control on metastatic dissemination in carcinoma of the prostate: Long-term results in patients treated with ^{125}I implantation. *Int. J. Rad. Oncol. Biol. Phys.*, 21(3):537–47, 1991. **pp. 1–8, 1–9.**

[81] J. G. Galvin, A. T. Turrisi, and Cheng. Treatment simulation using a CT unit. In *Radiation Oncology Physics, AAPM Monograph No. 15*. American Institute of Physics, New York, 1986. **p. 1–5.**

82 J. M. Galvin, X.-G. Chen, and R. M. Smith. Combining multileaf fields to modulate fluence distributions. *Int. J. Radiat. Oncol. Biol. Phys.*, 27:697–705, 1993. **p. 1–11.**

83 J. M. Galvin, A. R. Smith, and B. Lally. Characterization of a multileaf collimator system. *Int. J. Radiat. Oncol. Biol. Phys.*, 25:181–192, 1993. **p. 1–11.**

84 J. M. Galvin, A. R. Smith, R. D. Moeller, R. L. Goodman, W. D. Powlis, J. Rubenstein, L. Solin, B. Michael, M. Needham, C. J. Huntzinger, and Kligermann. Evaluation of multileaf collimator design for a photon beam. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:789–801, 1993. **p. 1–11.**

85 J. P. Gerard. High dose high-precision in radiation oncology. *Radiother. Oncol.*, 29:167–168, 1993. **p. 1–17.**

86 K. G. A. Gilhuijs, A. Touw, M. van Herk, and E. Vijlbrief. Optimization of automatic portal image analysis. *Med. Phys.*, 22(7):1089–1099, 1995. **p. 4–2.**

87 K. G. A. Gilhuijs, P. J. H. van de Hen, and M. van Herk. Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Med. Phys.*, 23(3):389–399, 1996. **pp. 1–23, 4–2.**

88 K. G. A. Gilhuijs and M. van Herk. An automatic in-line inspection of patient setup in radiation therapy using digital portal images. *Med. Phys.*, 20:667–677, 1993. **pp. 1-23, 4–2.**

89 P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, Inc., Orlando, Florida 32887, 1981. **p. 2–12.**

90 P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*, page 99. Academic Press, Inc., Orlando, Florida 32887, 1981. **p. 1–2.**

91 P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*, page p. 270. Academic Press, Inc., Orlando, Florida 32887, 1981. **p. 1–2.**

92 D. J. Gladstone, M. van Herk, and L. M. Chin. Verification of lung attenuator positioning before total body irradiation using an electronic portal imaging device. *Int. J. Radiat. Biol. Oncol. Phys.*, 27(2):449–454, 1993. **p. 1–11.**

93 M. Goitein. The utility of computed tomography in radiation therapy: an estimate of outcome. *Int. J. Radiat. Oncol. Biol. Phys.*, 5:1799–1807, 1979. **p. 1–10.**

94 M. Goitein. The comparison of treament plans. *Semin. Radiat. Oncol.*, 2:246–256, 1992. **p. 1–14.**

95 M. Goitein. The potential of real time portal imaging. In *3rd International Workshop on Electronic Portal Imaging*, San Francisco, U.S.A., 1994. **pp. 1–22, 1–23.**

[96] M. Goitein and M. Abrahms. Multi-dimensional treatment planning: I. Delineation of anatomy. *Int. J. Radiat. Oncol. Biol. Phys.*, 8:777–787, 1983. **p. 1–10**.

[97] M. Goitein, M. Abrahms, D. Rowell, H. Pollari, and J. Wiles. Multi-dimensional treatment planning: II. Beam's eye-view back projection, and projection through CT sections. *Int. J. Radiat. Oncol. Biol. Phys.*, 9:780–797, 1983. **p. 1–10**.

[98] M. Goitein and J. Busse. Immobilization error: some theoretical considerations. *Radiology*, **117**:407–412, 1975. **p. 1–18**.

[99] M. Goitein and A. Niemierko. Intensity modulated therapy and inhomogeneous dose to the tumor: A note of caution. *Int. J. Radiat. Oncol. Biol. Phys.*, 36(2):519–522, 1996. **p. 3–8**.

[100] M. Goitein and T. Schultheiss. Strategies for treating possible tumor extensions: some theoretical considerations. *Int. J. Radiat. Oncol. Biol. Phys.*, **11**:1519–1528, 1985. **p. 1–18**.

[101] R. C. Gonzales and P. Wintz. *Digital Image Processing*. Addison-Wesley, Reading, MA, 2nd edition, 1987. **p. 4–11**.

[102] Photon Treatment Planning Collaborative Working Group. State-of-the-art of external photon beam radiation treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:9–23, 1991. **p. 1–10**.

[103] Photon Treatment Planning Collaborative Working Group. Three-dimensional dose calculations for radiation treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 21:25–36, 1991. **p. 1–10**.

[104] A. Gustafsson, B. K. Lind, and A. Brahme. A generalized pencil beam algorithm for optimization of radiation therapy. *Med. Phys.*, 21(3):343–356, 1994. **pp. 1–14, 1–15, 2–6, 2–7, 2–14, 2–21, 2–27, 3–5, 3–7, 3–8, 3–9, 3–17**.

[105] A. Gustafsson, B. K. Lind, R. Svensson, and A. Brahme. Simultaneous optimization of dynamic multileaf collimation and scanning patterns or compensation filters using a generalized pencil beam algorithm. *Med. Phys.*, 22(7):1141–1156, 1995. **pp. 1–14, 2–6, 2–7, 2–14, 2–21, 2–27, 3–5, 3–7, 3–8, 3–9**.

[106] G. H. Hahn. *Hyperthermia and Cancer*. Plenum, New York, 1982. **p. 1–2**.

[107] P. S. Hahn, S. Shalev, D. Viggars, and P. Therrien. Treatment planning for protocol-based radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:937–939, 1990. **p. 1–14**.

[108] E. L. Hall. *Computer image processing and recognition*, page 181. Academic Press, New York, 1st edition, 1979. **p. 4–6**.

[109] J. Hanley, M. A. Lumley, G. S. Mageras, J. Sun, J. Zelefsky, S. A. Leibel, Z. Fuks, and G. J. Kutcher. Measurement of patient positioning errors in three-dimensinal conformal radiotherapy of the prostate. *Int. J. Radiat. Biol. Oncol. Phys.*, 27(2):435–444, 1997. **p. 5–9**.

[110] V. N. Hansen, P. M. Evans, and W. Swindell. The application of transit dosimetry to precision radiotherapy. *Med. Phys.*, 23(5):713–721, 1996. **p. 1–11**.

[111] J. Hayward and M. Galeffi. The significance of local control in the primary treatment of breast cancer. *Ann. Surg*, 122:1244–1247, 1987. **p. 1–8**.

[112] M. G. Herman, R. A. Abrams, and R. Mayer. Clinical use of on-line portal imaging for daily patient treatment verification. *Int. J. Radiat. Oncol. Biol. Phys.*, 24:216, 1992. Suppl. 1, **pp. 1–22, 1–23**.

[113] D. L. G. Hill and D. J. Hawkes. Medical image registration using knowledge of adgacency of anatomical structures. *Image Vision Computing*. 12:173–178, 1994. **p. 4–13**.

[114] L. Hodes. Semiautomatic optimization of external beam radiation treatment planning. *Radiology*, 110:191–196, 1974. **pp. 1–14, 1–15**.

[115] T. W. Holmes and T. R. Mackie. A unified approach to the optimization of brachytherapy and external beam therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 20:859–873, 1991. **pp. 1–15, 3–5, 3–7, 3–8, 3–9**.

[116] T. W. Holmes and T. R. Mackie. A comparison of three inverse treatment planning algorithms. *Phys. Med. Biol.*, 39:91–106, 1994. **p. 2–3**.

[117] T. W. Holmes, T. R. Mackie, and P. J. Reckwerdt. An iterative filtered backprojection inverse treatment planning algorithm for tomotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1215–1225, 1995. **pp. 1–15, 2–3, 3–5, 3–7, 3–8, 3–9**.

[118] D. H. Hristov, T. Falco, and B. G. Fallone. Automatic image correlation of megavoltage digitally reconstructed radiographs and portal images. Montreal, Quebec, June 4–7, 1995. Joint Meeting of Canadian Association of Radiologists, Canadian Organization of Medical Physicists and the Canadian College of Physicists in Medicine. **p. 1–24**.

[119] D. H. Hristov, T. Falco, and B. G. Fallone. Automatic image correlation of megavoltage digitally reconstructed radiographs and portal images. In *Joint Meeting of Canadian Association of Radiologists, Canadian Organization of Medical Physicists and the Canadian College of Physicists in Medicine*, pages 17–18, Montreal, Quebec, June 4–7, 1995. Proceedings, **p. 1–24**.

[120] D. H. Hristov and B. G. Fallone. Optimization and applications of an inverse treatment planning algorithm. In *43nd Annual General Meeting of The Canadian Organization of Medical Physicists*,

[133] D. H. Hristov and B. G. Fallone. An active set algorithm for treatment planning optimization. In *39th Annual General Meeting of The American Association of Medical Physicists*, Milwakee, Wisconsin, 27-31 July, 1997. **p. 1–24.**

[134] D. H. Hristov and B. G. Fallone. Automatic image correlation of megavoltage digitally reconstructed radiographs and portal images. In *4th International Workshop on Electronic Portal Imaging*, pages 17–18, Amsterdam, The Netherlands, June 10–12, 1996. Proceedings, **p. 1–24.**

[135] M. A. Hunt, T. E. Schultheiss, G. E. Desobry, M. Hakki, and G. E. Hanks. An evaluation of setup uncertainties for patients treated to pelvic sites. *Int. J. Rad. Oncol. Biol. Phys.*, 32(1):227–233, 1995. **p. 1–11.**

[136] S. Hyödynmaa, A. Gustafsson, and A. Brahme. Optimization of conformal electron beam therapy using energy- and fluence-modulated beams. *Med. Phys.*, 23(5):659–666, 1996. **p. 1–14.**

[137] ICRU. *ICRU Report 50: Prescribing, Recording, and Reporting Photon Beam Therapy*. Bethesda, Maryland: International Commission on Radiation Units and Measurements, 1993. **pp. 1–4, 3–39.**

[138] A. Jackson, R. Mohan, and B. Baldwin. Comments on "Sampling techniques for the evaluation of treatment plans". *Med. Phys.*, (5):1375–1376, 1993. **p. 1–16.**

[139] A. C. Jackson, G. J. Kutcher, and E. D. Yorke. Probability of radiation-induced complications for normal tisues with parallel architecture subject to non-uniform irradiation. *Med. Phys.*, 20:613–625, 1993. **pp. 1–14, 5–4.**

[140] N. L. Jain, M. G. Kahn, R. E. Drzymala, B. E. Emami, and J. A. Purdy. Objective evaluation of 3-d radiation treatment plans:A decision-analytic tool incorporating treatment preferences of radiation oncologists. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:321–333, 1993. **p. 1–14.**

[141] H. E. Johns and J. R. Cunningham. *The physics of radiology*, pages 682–683. Charles C. Thomas, Springfield, Illinois, 1983. **p. 1–2.**

[142] H. E. Johns and J. R. Cunningham. *The physics of radiology*. Charles C. Thomas, Springfield, Illinois, 1983. **p. A–3.**

[143] S. Jones and A. Boyer. Investigation of an FFT-based correlation technique for verification of radiation treatment setup. *Med. Phys.*, 18:1116–1125, 1991. **pp. 1-23, 4–3, 4–11, 4–12, 4–13, 4–15, 5–9.**

[144] T. J. Jordan and P. C. Williams. The design and performance characteristics of a multileaf collimator. *Phys. Med. Biol.*, 39:231–251, 1994. **p. 1–11.**

[145] P. Källman, B. K. Lind, and A. Brahme. An algorithm for maximizing the probability of complication-free tumour control in radiation therapy. *Phys. Med. Biol.*, 37:871–890, 1992. **pp. 1–14, 2–6, 2–11, 2–12, 2–14, 2–21, 2–27.**

[146] P. Källman, B. K. Lind, A. Eklof, and A. Brahme. Shaping of arbitrary dose distributions by dynamic multileaf collimation. *Phys. Med. Biol.*, 33:1291–1300, 1988. **p. 1–11.**

[147] P. Källman, A. Ågren, and A. Brahme. Tumour and normal tissue responses to fractionated non-uniform dose delivery. *Int. J. Radiat. Biol.*, 62(2):249–262, 1992. **pp. 1–14, 2–11, 2–27.**

[148] M. L. Kessler, R. K. TelHaken, B. A. Fraass, and D. L. McShan. Expanding the use and effectiveness of dose-volume histograms for 3-D treatment planning I: Integration of 3-D dose display. *Int. J. Radiat. Oncol. Biol. Phys.*, 29:1125–1131, 1994. **p. 1–10.**

[149] B. Kihlen and B. I. Ruden. Reproducibility of field alignment in radiation therapy. *Acta Oncol.*, 13:1583–1588, 1989. **p. 1–20.**

[150] J. J. Kinzie, G. E. Hanks, C. J. Maclean, and S. Kramer. Patterns of care study: Hodgkin's disease relapse rates and adequacy of portals. *Cancer*, 52:2223–2226, 1983. **pp. 1–18, 1–20.**

[151] M. C. Kirby and P. C. Williams. Portal imaging for the verification of breast treatments. *Radiother. Oncol.*, 22:314–316, 1991. **p. 1–11.**

[152] M. C. Kirby and P. C. Williams. Measurement possibilities using an electronic portal imaging device. *Radiother. Oncol.*, 29(2):237–243, 1993. **p. 1–11.**

[153] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983. **p. 2–5.**

[154] A. H. Knowlton. General principles of radiation therapy. In D. S. Fishter and J. C. Marsh, editors, *Cancer Therapy*, pages 169–175. G. K. Hall Medical Publishers, Boston, 1982. **p. 1–2.**

[155] D. A. Kuban, A. M. El-Mahdi, and P. F. Schellhammer. Effect of local tumor control on distant metastasis and survival in prostatic adenocarcinoma. *Urology*, 30(5):420–426, 1987. **p. 1–8.**

[156] G. J. Kutcher and C. Burman. Calculation of complication probability factors for nonuniform normal tissue irradiation: the effective volume method. *Int. J. Radiat. Oncol. Biol. Phys.*, 16:1623–1630, 1989. **p. 1–14.**

[157] G. J. Kutcher, R. Mohan, S. A. Leibel, Z. Fuks, and C. C. Ling. Computer-controlled 3D conformal therapy. In *Radiation Therapy Physics*. Springer-Verlag, Berlin, Germany, 1992. **p. 1–12.**

[158] R. G. Lane, M. D. Loyd, C. H. Chow, E. Eklelunu, and I. I. Rosen. Custom beam profiles in computer-controlled radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 22:167–174, 1991. **p. 1–11.**

[159] M. Langer, R. Brown, P. Kijewski, and H. Chul. The reliability of optimization under dose-volume limits. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:529–538, 1993. **pp. 1–16, 2–5, 3–3, 3–4, 3–6, 5–4.**

[160] M. Langer, R. Brown, S. Morrill, R. Lane, and O. Lee. A genetic algorithm for generating beam weights. *Med. Phys.*, 23(6):965–971, 1991. **pp. 1–14, 1–15.**

[161] M. Langer, R. Brown, M. Urie, J. Leong, M. Stracher, and J. Shapiro. Large-scale optimization of beam weights under dose-volume restrictions. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:887–893, 1990. **pp. 1–14, 1–15, 3–3, 3–4, 3–6.**

[162] M. Langer and J. Leong. Optimization of beam weights under dose-volume restrictions. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1255–1259, 1987. **pp. 1–14, 3–3, 3–4, 3–6.**

[163] M. Langer, S. Morrill, R. Brown, M. Urie, O. Lee, and R. Lane. A comparison of mixed integer programming and fast simulated annealing for optimizing beam weights in radiation therapy. *Med. Phys.*, 23(6):957–964, 1996. **pp. 1–14, 1–15, 3–3, 3–4, 3–6.**

[164] J. V. Lebesque, A. Bel, J. Bijhold, and A. A. M. Hart. Detection of systematic patient errors by portal film analysis. *Radiother. Oncol.*, 23:198, 1992. **p. 1–20.**

[165] S. Lee, P. S. Cho, R. Marks, and O. Seho. Conformal radiotherapy computation by the method of alternating projections onto convex sets. *Phys. Med. Biol.*, 42:1065–1086, 1997. **p. 2–4.**

[166] S. A. Leibel, R. Heimann, G. J. Kutcher, M. Zelefsky, C. M. Burman, E. Mellian, Orazem. J., S. Kelson, R. Mohan, T. LoSasso, Y.-C. Lo, J. Wiseberg, D. S. Chapman, C. C. Ling, and Z. Fuks. Three-dimensional conformal radiation therapy in localized carcinoma of the prostate: preliminary results of a phase I dose-escalation study. *Int. J. Radiat. Oncol. Biol. Phys.*, 28(1):55–65, 1994. **pp. 1–6, 1–8.**

[167] S. A. Leibel, G. J. Kutcher, R. Mohan, L. B. Harrison, J. G. Armstrong, M. J. Zelefsky, T. J. LoSasso, C. M. Burman, G. S. Mageras, C. S. Chui, L. J. Brewster, M. E. Masterson, Y. C. Lo, C. C. Ling, and Z. Fuks. Three-dimensional conformal radiation therapy at the Memorial Sloan-Kettering Cancer Centre. *Semin. Radiat. Oncol.*, 2:274–289, 1992. **pp. 1–8, 1–14.**

[168] S. A. Leibel, C. C. Ling, G. J. Kutcher, R. Mohan, C. Cordon-Cordo, and Z. Fuks. The biological basis for conformal three-dimensional radiation therapy. *Int. J. Rad. Oncol. Biol. Phys.*, 21:805–811, 1991. **pp. 1–6, 1–8.**

[169] S. A. Leibel, C. B. Scott, M. Mohiuddin, V. A. Marcial, L. R. Coia, L. W. Davis, and Z. Fuks. The effect of local-regional control on distant metastatic dissemination in carcinoma of the head and neck:

Results of an analysis from the RTOG head and neck database. *Int. J. Radiat. Oncol. Biol. Phys*, 21(3):549–556, 1991. **pp. 1–6, 1–8.**

[170] S. A. Leibel, M. Zelefsky, G. J. Kutcher, C. M. Burman, S. Kelson, and Z. Fuks. Three-dimensional conformal radiation therapy in localized carcinoma of the prostate: interim report of a phase I dose-escalation study. *J. Urol.*, 152:1792–1798, 1994. **p. 1–9.**

[171] L. Lemieux, R. Jagor, D. R. Fish, N. D. Kitchen, and D. G. T. Thomas. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Med. Phys.*, 21:1749–1760, 1994. **pp. 4-24, 5–9.**

[172] J. Leong. A digital image processing system for high energy x-ray portal images. *Phys. Med. Biol.*, 29:1527–1535, 1984. **p. 1-11.**

[173] J. Leong and M. A. Stracher. Visualisation of internal motion within a treatment portal during a radiation treatment. *Radiother. Oncol.*, 9:153–156, 1987. **p. 1–11.**

[174] K. W. Leszczynski, S. Loose, and P. Dunscombe. Segmented chamfer matching for prescription-treatment image registration in radiotherapy. *Phys. Med. Biol.*, 40:83–94, 1995. **pp. 1–23, 4–2.**

[175] K.W. Leszczynski, S. Shalev, and N.S. Cosby. The enhancement of radiotherapy verification images by an automated edge detection technique. *Med. Phys.*, 19:611–621, 1992. **p. 1-23.**

[176] M. B. Levene, P. K. Kijevski, L. M. Chin, S. M Bengt, B. E. Bjarngard, and S. Hellman. Computer-controlled radiation therapy. *Radiology*, 129:769–775, 1978. **p. 1–10.**

[177] A. S. Lichter. Three-dimensional conformal radiation therapy: a testable hypothesis. *Int. J. Rad. Oncol. Biol. Phys.*, 21:853–855, 1991. **pp. 1–7, 1–8.**

[178] Allen S. Lichter and Theodore S. Lawrence. Recent advances in radiation oncology. *The New England Journal of Medicine*, 332(6):370–372, 1995. **pp. 1–2, 1–7.**

[179] B. K. Lind. Properties of an algorithm for solving the inverse problem in radiation therapy. *Inverse Problems*, 6:415–426, 1990. **pp. 1–15, 2–6, 2–7, 2–14, 3–5, 3–7, 3–8, 3–9.**

[180] C. C. Ling, C. Burman, C. S. Chui, A. Jackson, G. J. Kutcher, S. A. Leibel, T. LoSasso, G. Mageras, R. Mohan, E. Yorke, and Z. Fuks. Perspectives of multidimensional conformal radiation treatment. *Radiotherapy and Oncology*, 29:129–139, 1993. **pp. 1–9, 1–12.**

[181] C. C. Ling, C. Burman, C. S. Chui, G. J. Kutcher, S. A. Leibel, T. LoSasso, R. Mohan, T. Bortfeld, L. Reinstein, S. Spirou, X. Wang, Q. Wu, M. Zelefsky, and Z. Fuks. Conformal radiation treatment of prostate cancer using inversely-planned intensity modulated photon beams with dynamic multileaf collimation. *Int. J. Radiat. Oncol. Biol. Phys.*, 35(4):721–730, 1996. **pp. 1–11, 1–17, 3–17.**

[182] X.-Q. Lu and L. M. Chin. Further discussions on sampling methods-a response to Letters to the Editor [Med. Phys. 20, 1375-1376 (1993)] and [Med. Phys. 20, 1377-1380 (1993)]. *Med. Phys.*, (5):1375–1376, 1993. **pp. 1–16, 5–4.**

[183] J. T. Lyman. Complication probabilities as assessed from dose-volume histograms. *Radiat. Res.*, 104:S13–S19, 1985. **p. 1–14.**

[184] T. R. Mackie, J. W. Scrimger, and J. J. Battista. A convolution method of calculating dose for 15-MV x-rays. *Med. Phys.*, 12:188–196, 1985. **pp. 1–16, 5–4.**

[185] G. S. Mageras and R. Mohan. Application of fast simulated annealing to optimization of conformal treatments. *Med. Phys.*, 20(3):639–647, 1993. **p. 1–14.**

[186] J. Mantel, H. Perry, and J. J. Weinkam. Automatic variation of field size and dose rate in rotation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 2:697–704, 1974. **pp. 1–10, 1–14, 1–15.**

[187] J. E. Marks and A. G. Haus. The effect of immobilization on localization error in the radiotherapy of head and neck cancer. *Clin. Radiol.*, 27:175–177, 1976. **p. 1–20.**

[188] J. E. Marks, A. G. Haus, H. G. Sutton, and M. L. Griem. Localization error in the radiotherapy of Hodgkin's disease and malignant lymphoma with extended mantle fields. *Cancer*, 34:83–90, 1974. **p. 1–20.**

[189] J. E. Marks, A. G. Haus, H. G. Sutton, and M. L. Griem. The value of frequent treatment verification films in reducing localization error in the irradiation of complex fields. *Cancer*, 37:2755–2761, 1976. **p. 1–20.**

[190] G. P. McCormick. Anti-zig-zagging by bending. *Management Science*, 15:315–319, 1969. **p. 2–14.**

[191] J. E. McGary and A. L. Boyer. An interactive, parallel, three-dimensional fast Fourier transform convolution dose calcualtion using a supercomputer. *Med. Phys.*, 24(4):519–522, 1997. **p. 1–16.**

[192] T. R. McNutt, T. R. Mackie, P. Reckwerdt, and B. R. Paliwal. Modeling dose distributions from portal dose images using the convolution/superposition method. *Med. Phys.*, 23(8):1381–1392, 1996. **p. 1–11.**

[193] T. R. McNutt, T. R. Mackie, P. Reckwerdt, and B. R. Paliwal. Modeling dose distributions from portal dose images using the convolution/superposition method. *Med. Phys.*, 23(8):1381–1392, 1996. **p. 1–11.**

[194] B. J. McParland. Digital portal registration be sequential anatomical matchpoint and image correlations for real-time continuous field alignment verification. *Med. Phys.*, 22(7):1063–1075, 1995. **p. 4–2.**

[195] D. L. McShan, B. A. Fraass, and A. S. Lichter. Full integration of the Beam's Eye View concept into clinical treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:1485–1494, 1990. **p. 1–10.**

[196] H. Meertens. Digital processing of high energy photon beam images. *Med. Phys.*, 12:111–113, 1985. **p. 1-23.**

[197] H. Meertens, J. Bijold, and J. Strackee. A method for the measurement of field placement errors in digital portal images. *Phys. Med. Biol.*, 35:299–323, 1990. **p. 1–23.**

[198] H. R. Menick, L. Garfinkel, and G. D. Dodd. Preliminary report of the National Cancer Data Base. *CA-A Cancer Journal for Clinicians*, 41:7–18, 1991. **p. 1–6.**

[199] N. Metropolis, A. Rosenbluth, M. Rosenbluth, H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953. **p. 2–5.**

[200] R. Mohan, C. Chui, and L. Lidofsky. Differential pencil beam dose computation model for photons. *Med. Phys.*, 13:64–73, 1986. **pp. 1–16, 5–4.**

[201] R. Mohan, G. S. Mageras, B. Baldwin, L. J. Brewster, G. J. Kutcher, S. Leibel, C. M. Burman, C. C. Ling, and Z. Fuks. Clinically relevant optimization of 3-d conformal treatments. *Med. Phys.*, 19(4):933–944, 1993. **p. 1–14.**

[202] R. Mohan, X. Wang, T. Bortfeld, A. L. Boyer, G. J. Kutcher, S. A. Leibel, Z. Fuks, and C. C. Ling. The potential and limitations of the inverse radiotherapy technique. *Radiotherapy and Oncology*, 32:232–248, 1994. **pp. 1–14, 1–16, 3–9, 3–17, 5–6.**

[203] R. Mohan, Q. Wu, X.—H. Wang, and J. Stein. Intensity modulation optimization, lateral transport of radiation and margins. *Med. Phys.*, 23(12):2011–2021, 1996. **p. 3–9.**

[204] S. M. Morrill, , I. I. Rosen, R. G. Lane, and J. A. Belli. The influence of dose constraint point placement in optimizaed radiation therapy treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 19(2):129–142, 1990. **pp. 1–16, 2–5, 3–5.**

[205] S. M. Morrill, R. G. Lane, J. A. Wong, and I. I. Rosen. Dose-volume considerations with linear programming. *Med. Phys.*, 18(6):1200–1210, 1991. **pp. 1–14, 1–15, 3–3, 3–5, 3–6.**

[206] J. Moseley, D.A. Jaffray, J-P. Bissonnette, A. Fenster, and P. Munro. Image analysis techniques for portal imaging. *Med. Phys.*, 19:822, 1992 (abstract). **pp. 4-2, 4–3.**

[207] J. Moseley and P. Munro. A semiatomatic method for registration of portal images. *Med. Phys.*, 21(4):551–558, 1994. **pp. 1–23, 4–2, 4–15, 4–23, 4–24, 4–25.**

[208] T. R. Munro and C. W. Gilbert. The relation between tumor lethal dose and the radiosensitivity of tumor cells. *Br. J. Radiol.*, 34:246–251, 1961. **p. 1–2.**

[209] J. E. Munzenrider, M. Pilepich, J. B. Rene-Ferero, I. Tchakarova, and B. L. Carter. Use of body scanner in radiotherapy treatment planning. *Cancer*, 40:170–179, 1979. **p. 1–10.**

[210] K. G. Murty. *Linear Programming*. Joihn Wiley & Sons, Inc., New York, 1983. **p. 3–4.**

[211] M. H. Myers and B. F. Hankey. Cancer patient survival experience in the United States. Publication No. (NIH) 80–2148, DHHS, U. S. Goverment Printing Office, Washington, D.C., 1980. **p. 1–6.**

[212] W. R. Nelson, H. Hirayma, and D. W. O. Rogers. The EGS4 code system. Internal report SLAC 265, Stanford Linear Accelerator Center, 1985. **p. 4–4.**

[213] A. Niemierko. Random search algorithm (RONSC) for optimization of radiation therapy with both physical and biological end points and constraints. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:89–98, 1992. **pp. 1–14, 1–15, 3–42, 3–45.**

[214] A. Niemierko. Non-linear congugate gradient algorithm for optimization of intensity modulated beams. *Med. Phys.*, 23:1172, 1996. (abstract) **p. 2–7.**

[215] A. Niemierko. Selection of the objective function for optimization of intensity modulated beams. *Med. Phys.*, 23:1172, 1996. (abstract) **pp. 3–8, 3–43, 3–44.**

[216] A. Niemierko. Treatment evaluation. In T. R. Mackie and J. R. Palta, editors, *Teletherapy: Present and Future*, page 186, Vancouver, Canada, 1996. AAPM. Proceedings of the 1996 Summer School, Advanced Medical Publishing. **p. 3–45.**

[217] A. Niemierko. Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Med. Phys.*, 24(1):103–110, 1997. **pp. 1–14, 3–44.**

[218] A. Niemierko and M. Goitein. Modelling of normal tissue response to radiation: the critical volume-model. *Int. J. Radiat. Oncol. Biol. Phys.*, 25:135–145, 1992. **p. 1–14.**

[219] A. Niemierko, M. Urie, and M. Goitein. Optimization of 3D radiation therapy with both physical and biological end points and constraints. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:99–108, 1992. **p. 1–15.**

[220] T. Nishidai, Y. Nagata, M. Takahashi, M. Abe, N. Yamaoka, H. Ishihara, Y. Kubo, H. Ohta, and C. Kazusa. CT simulator: a new 3-D planning and simulating system for radiotherapy. *Int. J. Rad. Oncol. Biol. Phys.*, 18:499–504, 505–513, 1990. **p. 1–5.**

[221] National Cancer Institute of Canada. Canadian cancer statistics 1997. Toronto, 1997. **p. 1–2.**

[222] P. Okunieff, D. Morgan, A. Niemierko, and H. Suit. Radiation dose-response of human tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1227–1237, 1995.

[223] L. J. Old. Immunotherapy for cancer. *Sci. Am.*, 275(3):136–143, 1996. **p. 1–2.**

[224] A. Oliff, J. B. Gibbs, and F. McCormick. New molecular targets for cancer therapy. *Sci. Am.*, 275(3):144–149, 1996. **p. 1–2.**

[225] C. A. Perez, R. R. Kuske, H. M. Camel, A. E. Galakatos, M. A. Hederman, M. S. Kao, and B. J. Walz. Analysis of pelvic tumor control and impact on survival in carcinoma of the uterine cervix treated with radiation therapy alone. *Int. J. Rad. Oncol. Biol. Phys*, 14(4), 1988. **p. 1–8.**

[226] C.A. Perez, M. V. Pilepich, and F. Zivnuska. Tumor control in definitive irradiation of localized carcinoma of the prostate. *Int. J. Rad. Oncol. Biol. Phys*, 12:523–531, 1986. **p. 1–8.**

[227] E. Polak. *Computational methods in optimization.* Academic Press, New York, 1971. **pp. 2–7, 2–17, 3–7.**

[228] A. Pollack, G. Zagars, G. Starkshall, C. H. Childress, S. Kopplin, A. L. Boyer, and I. I. Rosen. Conventional vs. conformal radiotherapy for prostate cancer: preliminary results of dosimetry and acute toxicity. *Int. J. Radiat. Oncol. Biol. Phys.*, 34(3):555–564, 1994. **p. 1–9.**

[229] E. H. Porter. The statistics of dose/cure relationships for irradiated tumors. *Br. J. Radiol.*, 53:210–227, 336–345, 1980. **p. 1–2.**

[230] W. Powlis, I. Brikman, S. Seshardri, and P. Block. Portal radiographs: Digital enhancement of contrast. *Radiology*, 169:839–841, 1988. **p. 1-23.**

[231] W. K. Pratt. *Digital Picture Processing.* John Wiley & Sons, Inc., New York, 1978. **p. 4–11.**

[232] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies in C.* Cambridge University Press, New York, 1988. **pp. 3–13, 2–13.**

[233] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies in C*, pages 569–580. Cambridge University Press, New York, 1988. **p. 3–13.**

[234] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies in C*, page 426. Cambridge University Press, New York, 1988. **p. 2–13.**

[235] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies in C*, pages 299–300. Cambridge University Press, New York, 1988. **p. 2–29.**

[236] B. S. Proimos. Synchronous field shaping in rotational megavolt therapy. *Radiology*, 74:753–757, 1960. **p. 1–10.**

[237] I. Rabinowitz, J. Broomberg, M. Goitein, K. McCarthy, and J. Leong. Accuracy of radiation field alignment in clinical practice. *Int. J. Radiat. Oncol. Biol. Phys.*, 11:1857–1867, 1985. **pp. 1–19, 1–20.**

[238] T. Radcliffe, R. Rajapakshe, and S. Shalev. Pseudocorrelation: A fast, robust, absolute, grey-level image alignment algorithm. *Med. Phys.*, 21:761–769, 1994. **pp. 1-23, 4–3, 4–22, 4–23.**

[239] C. Raphael. Mathematical modelling of objectives in radiation therapy treatment planning. *Phys. Med. Biol.*, 37:1293–1312, 1992. **pp. 1–14, 3–8.**

[240] C. Rawlings, S. Griffits, H. Probst, and N. Howson. Clinical experience with megavoltage portal imaging. In *4th International Workshop on Electronic Portal Imaging*, Amsterdam, The Netherlands, 1996. Proceedings, **pp. 1–22, 1–23.**

[241] A. T. Redpath, B. L. Vickery, and D. H. Wright. A new technique for radiotherapy using quadratic programming. *Phys. Med. Biol.*, 21:781–791, 1974. **pp. 1–14, 1–15.**

[242] L. E. Reinstein, H. I. Amols, P. J. Biggs, R. T. Droege, A. B. Filimonov, W. R. Lutz, and S. Shalev. Radiotherapy portal imaging quality. Report #24, AAPM, New York, American Institute of Physics, 1988. **p. 1–20.**

[243] A. Ågren, A. Brahme, and I. Turesson. Optimization of uncomplicated control for head and neck tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 19:1077–85, 1990. **pp. 1–14, 2–11, 2–12.**

[244] I. Rosen, R. G. Lane, S. M. Morrill, and J. A. Belli. Treatment plan optimization using linear programming. *Med. Phys.*, 18(2):141–152, 1991. **pp. 1–15, 3–4.**

[245] S. A. Rosenthal, J. M. Galvin, J. W. Goldwein, A. R. Smith, and P. H. Blitzer. Improved methods for determination of variability in patient positioning for radiation therapy using simulation and serial portal film measurements. *Int. J. Radiat. Oncol. Biol. Phys.*, 23:621–625, 1992. **p. 1–20.**

[246] P. Rubin and W. D. Siemann. Principles of radiation oncology and cancer radiotherapy. In P. Rubin, editor, *Clinical Oncology*. W.B. Saunders Company, The Curtis Center, Independence Square West, Philadelphia, PA 19106, 7th edition, 1993. **p. 1–2.**

[247] V. Rudat, M. Flentje, D. Oetzel, M. Menke, W. Schlegel, and M. Wannenmacher. Influence of the positioning error on 3D conformal dose distributions during fractionated radiotherapy. *Radiother. Oncol.*, 33:157–166, 1994. **pp. 1–17, 1–18.**

[248] R. J. Schalkoff. *Digital Image Processing and Computer Vision*. John Wiley & Sons, Inc., 1989. **p. 4–10.**

[249] S. E. Schild, J. A. Jr. Martenson, L. L. Gunderson, D. M. Ilstrup, K. K. Berg, M. J. O'Connell, and L. H. Weiland. Postoperative adjuvant therapy of rectal cancer: An analysis of disease control, survival, and prognostic factors. *Int. J. Rad. Oncol. Biol. Phys.*, 17(1):55–62, 1989. **p. 1–8.**

[250] G. Sherouse, J. Rosenman, H. McMurr, S. Pizer, and E. Chaney. Automatic digital contrast enhancement of radiation therapy portal films. *Int. J. Radiatian Oncology Biol. Phys.*, 13:801–806, 1987. **p. 1-23.**

[251] G. W. Sherouse and E. L. Chaney. The portable virtual simulator. *Int. J. Rad. Oncol. Biol. Phys.*, 21:475–482, 1991. **pp. 1–5, 1–10.**

[252] G.W. Sherouse, K. Novins, and E.L. Chaney. Computation of digitally reconstructed radiographs for use in radiotherapy treatment design. *Int. J. Radiat. Oncol. Biol. Phys.*, 18:651–658, 1990. **p. 4–6.**

[253] T. E. Shultheiss and C. G. Orton. Models in radiotherapy: Definitions of decision criteria. *Med. Phys.*, 12:183–187, 1985. **p. 1–14.**

[254] B. Siddon. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med. Phys.*, 12(2):252–255, 1985. **p. 4–6.**

[255] A. Söderström and A. Brahme. Small is beautiful — and often enough. *Int. J. Radiat. Oncol. Biol. Phys.*, 34:757–758, 1996. (letter to the editor) **p. 3–42.**

[256] S. Söderstrom and A. Brahme. Optimization of the dose delivery in a few field techniques using radiobiological objective functions. *Med. Phys.*, 20:1201–1210, 1993. **p. 3–42.**

[257] S. Söderstrom, A. Gustafsson, and A. Brahme. The clinical value of different treatment objectives and degrees of freedom in radiation therapy optimization. *Radiotherapy and Oncology*, 29:148–163, 1993. **p. 3–42.**

[258] D. Sonderman and P. G. Abrahamson. Radiotherapy treatment design using mathematical programming models. *Oper. Res.*, 33:705–725, 1985. **pp. 1–14, 1–15.**

[259] S. V. Spirou and C. S. Chui. Generation of arbitrary intensity profiles by dynamic jaws or multileaf collimators. *Med. Phys.*, 21(1031-1041), 1994. **p. 1–14.**

[260] G. Starkshall. A constrained least-square optimization method for external beam radiation therapy treatment planning. *Med. Phys.*, 11:659–665, 1984. **p. 1–15.**

[261] J. Stein, R. Mohan, X.-H. Wang, Q. Wu, K. Preiser, C. Ling, and W. Schlegel. Number and orientations of beams in intensity-modulated radiation treatments. *Med. Phys.*, 24(2):149–160, 1997. **pp. 1–15, 3–6, 3–17, 3–37, 3–42.**

[262] Ralph E. Steuer. *Multiple Criteria Optimization: Theory, Computation, and Application.* Joihn Wiley & Sons, Inc., New York, 1986. **p. 3–45.**

[263] E. Storm and H.I. Israel. Photon cross sections from 1 KeV to 100 MeV for elements Z=1 to Z=100. *Nuclear Data Tables*, A7:565–681, 1970. **p. 4–4.**

[264] H. D. Suit. Potential for improving survival rates for the cancer patient by increasing the efficacy of the treatment of the primary tumor. *Cancer*, 50:1227–1234, 1982. **pp. 1–7, 1–9.**

[265] H. D. Suit. Local control and patient survival. *Int. J. Radiat. Oncol. Biol. Phys.*, 23(3):653–660, 1992. **pp. 1–7, 1–14.**

[266] H. D. Suit, H. J. Mankin, W. C. Wood, M. C. Gebhardt, D. C. Harmon, A. Rosenberg, J. E. Tepper, and D. Rosenthal. Treatment of the patient with stage M0 soft tissue sarcoma. *Journal of Clinical Oncology*, 6(5):854–62, 1988. **p. 1–8.**

[267] S. Takahashi. Conformation radiotherapy: rotation techniques as applied to radiography and radiotherapy of cancer. *Acta Radiol. Suppl.*, (242):1–141, 1965. **pp. 1–8, 1–10.**

[268] A. Tinger, J. M. Michalski, W. R. Bosch, R. K. Valicenti, D. A. Low, and R. J. Myerson. An analysis of intratreatment and intertreatment displacements in pelvic radiotherapy using electronic portal imaging. *Int. J. Rad. Oncol. Biol. Phys.*, 34(3):683–690, 1996. **p. 1–11.**

[269] V. Tulovski, M Ringor, and L Papież. Optimization of rotational radiotherapy treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, 32(4):1205–1214, 1995. **pp. 2–3, 3–8, 3–34, 3–43, 3–44.**

[270] A. J. van der Kogel. Radiobiological and clinical data on dose/volume effect relationships in normal tissues. *Radiother. Oncol.*, 29:127–128, 1993. **p. 1–14.**

[271] M van Herk, K Gilhuijs, P Williams, P. Cinquin, L. Cionini, and W. Swindell. Role of electronic imaging in high dose/high rate precision therapy. *Radioth. Oncol.*, 29:269–270, 1993. **p. 1–11.**

[272] A. Vigliotti, T. A. Rich, M. M. Romsdahl, H. R. Withers, and M. J. Oswald. Postoperative adjuvant radiotherapy for adenocarcinoma of the rectum and rectosigmoid. *Int. J. Rad. Oncol. Biol. Phys*, 13(7):999–1006, 1987. **p. 1–8.**

[273] E. Vigneault, J. Pouliot, J. Laverdiere, J. Roy, and M. Dorion. Electronic portal imaging device detection of radioopaque markers for the evaluation of prostate position during megavoltage irradiation: clinical study. *Int. J. Radiat. Biol. Oncol. Phys.*, 37(1):205–212, 1997. **p. 1–11.**

[274] H. Wang. Private communication, **p. 1–22.**

[275] H. Wang and B.G. Fallone. A robust morphological algorithm for automatic radiation field-extraction and correlation of portal images. *Med. Phys.*, 21:237–244, 1994. **pp. 1-23, 4–2.**

[276] H. Wang and B.G. Fallone. A mathematical model of radiation field edge localization. *Med. Phys.*, 22:1107–1110, 1995. **pp. 1–23, 4–2, 4–7.**

[277] X. Wang, R. Mohan, A Jackson, S. A. Leibel, Z. Fuks, and C. C. Ling. Optimization of intensity-modulated 3d conformal treatment plans based on biological indices. *Radiotherapy and Oncology*, 37:140–152, 1995. **pp. 1–14, 2–6, 2–7, 2–14.**

[278] S. Webb. Optimization of conformal radiotherapy dose distributions by simulated annealing. *Phys. Med. Biol.*, 34:1349–1370, 1989. **p. 1–15.**

[279] S. Webb. Optimization by simulated annealing of three-dimensional, conformal treatment planning for radiation fields defined by a multileaf collimator. *Phys. Med. Biol.*, 36:1201–1226, 1991. **p. 1–15.**

[280] S. Webb. Optimization by simulated annealing of three-dimensional conformal treatment palnning for radiation fields defined by multileaf collimator. II. Inclusion of two-dimensional modulation of the X-ray intensity. *Phys. Med. Biol.*, 37(8):1689–1704, 1992. **pp. 1–14, 2–6, 3–5, 3–6, 3–7, 3–8, 3–9.**

[281] S. Webb. *The physics of three-dimensional radiation therapy*. Institute of Physics Publishing, Bristol, 1993. **pp. 1–8, 1–15.**

[282] S. Webb. Optimizing the planning of intensity-modulated radiotherapy. *Phys. Med. Biol.*, 39:2229–2246, 1994. **pp. 1–11, 1–16, 3–34.**

[283] K. J. Weeks, L. Marks, S. K. Ray, D. P. Spencer, D. A. Turner, and A. H. Friedman. 3-dimensional optimization of multiple arcs for stereotactic radiosurgery. *Int. J. Radiat. Oncol. Biol. Phys.*, 26:147–154, 1993. **p. 1–15.**

[284] K. J. Weeks and M. R. Sontag. 3-D dose-volume compensation using nonlinear least-squares regression technique. *Med. Phys.*, 18:474–480, 1991. **p. 1–15.**

[285] J. E. White, T. Chen, J. McCracken, P. Kennedy, H. G. Seydel, G. Hartman, J. Mira, M. Khan, F. Y. Durrance, and O. Skinner. The influnce of radiation therapy quality control on survival, response and sites of relapse in oat cell carcinoma of the lung. *Cancer*, 50:1084–1090, 1982. **p. 1–19.**

[286] M. V. Williams, J. Denekamp, and J. F. Fowler. Dose-response relationships for human tumors: Implications for clinical trials of dose-modifying agents. *Int. J. Rad. Oncol. Biol. Phys*, 20:1703–1707, 1984. **p. 1–8.**

[287] H. R. Withers, J. M. Taylor, and B. Maciejewski. Treatment volume and tissue tolerance. *Int. J. Radiat. Oncol. Biol. Phys.*, 14:751–759, 1988. **p. 1–14.**

[288] A. B. Wolbarst, L. M. Chin, and G. K. Svensson. Optimization of radiation therapy: integral-response of a model biological system. *Int. J. Radiat. Oncol. Biol. Phys.*, 8:1761–1769, 1982. **p. 1–14.**

[289] L. Xing and G. T. Y. Chen. Iterative methods for inverse treatment planning. *Phys. Med. Biol.*, 41:2107–2133, 1996. **pp. 1–15, 2–4, 3–5, 3–7, 3–8, 3–17.**

[290] L. Xing, C. Pelizzari, F. T. Kuchir, and G. T. Y. Chen. Optimization of relative weights and wedge angles in treatment planning. *Med. Phys.*, 24(2):215–221, 1997. **pp. 1–15, 2–4.**

[291] F.-F. Yin, P. Rubin, M. SChell, R. Wynn, B. West, G. Uschold, and D. Nelson. Clinical implementation of epid at the university of rochester cancer center. In *3rd International Workshop on Electronic Portal Imaging*, San Francisco, U.S.A., 1994. Proceedings, **pp. 1–22, 1–23**.

[292] E. D. Yorke, Z. Fuks, L. Norton, W. Whitmore, and C. C. Ling. Modelling the development of metastases from primary and locally recurrent tumors: comparison with a clinical data base of prostate cancer. *Cancer Res.*, 53:2987–2993, 1993. **p. 1–9**.

[293] E. D. Yorke, G. J. Kutcher, A. C. Jackson, and C. C. Ling. Probability of radiation-induced complications in normal tisues with parallel architecture under conditions of uniform whole or partial organ irradiation. *Radiother. Oncol.*, 26:226–237, 1993. **pp. 1–9, 1–14**.

[294] A. S. Zacarias, R. G. Lane, and I. I. Rosen. Assesment of a linear accelerator for segmented conformal radiation therapy. *Med. Phys.*, 20:193–198, 1993. **p. 1–11**.

[295] Corey E. Zankowski. Monte Carlo analysis of the 10 MV x-ray beam from a CLINAC-18 linear accelerator. Master's thesis, Medical Physics Dept., McGill University, Montreal, 1994. **p. 4–4**.

[296] S. Zhou and L.J. Verhey. A robust method of multileaf collimator (mlc) leaf-configuration verification. *Phys. Med. Biol.*, 39:1929–1948, 1993. **pp. 1–23, 4–2**.