

Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy

Ali Khamene ^{a,*}, Peter Bloch ^b, Wolfgang Wein ^{a,1}, Michelle Svatos ^c, Frank Sauer ^a

^a *Imaging and Visualization Department, Siemens Corporate Research, Inc., 755 College Road East, Princeton, NJ 08540, USA*

^b *Department of Radiation Oncology, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104, USA*

^c *Oncology Care Systems, Siemens Medical Solutions, Inc., 4040 Nelson Avenue, Concord, CA 94520, USA*

Received 23 February 2004; received in revised form 12 August 2004; accepted 10 June 2005

Available online 8 September 2005

Abstract

The efficacy of radiation therapy treatment depends on the patient setup accuracy at each daily fraction. A significant problem is reproducing the patient position during treatment planning for every fraction of the treatment process. We propose and evaluate an intensity based automatic registration method using multiple portal images and the pre-treatment CT volume. We perform both geometric and radiometric calibrations to generate high quality digitally reconstructed radiographs (DRRs) that can be compared against portal images acquired right before treatment dose delivery. We use a graphics processing unit (GPU) to generate the DRRs in order to gain computational efficiency. We also perform a comparative study on various similarity measures and optimization procedures. Simple similarity measure such as local normalized correlation (LNC) performs best as long as the radiometric calibration is carefully done. Using the proposed method, we achieved better than 1 mm average error in repositioning accuracy for a series of phantom studies using two open field (i.e., 41 cm²) portal images with 90° vergence angle.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Patient positioning; Radiation therapy; 2D–3D registration; DRR generation

1. Introduction

External beam radiation therapy is an effective method of treating cancer, in which the lesion is irradiated with high-energy beams produced by a linear accelerator (LINAC). Treatment techniques such as 3D conformal radiation therapy and intensity modulated radiation therapy (IMRT) provide very accurate radiation to the lesion, while sparing healthy tissues. Radiation treatment always starts with the acquisition of planning computerized tomographic (CT) images. CT

data play a central role in radiotherapy treatment planning (RTP) (Herman et al., 2001).

Efficacy of the RTP depends on the patient setup at each daily fraction. The problem is to reproduce the patient position at the time of acquiring the planning CT scans (i.e., used for RTP) for each fraction of the treatment process. Discrepancies between the planned and delivered treatment positions significantly degrade the therapeutic ratio (White et al., 1982). One of the conventional methods to reduce patient setup error includes matching the skin tattoo marks generated at the planning phase with the treatment room laser lights. Immobilization devices such as polyurethane foam casts or thermoplastic masks are also used to assist in reproducing the daily treatment position (Bentel, 1999). These approaches can fail to adequately ensure the location reproducibility, since the positioning is done based on

* Corresponding author. Tel.: +1 609 734 6553; fax: +1 609 734 6565.

E-mail address: akhamene@scr.siemens.com (A. Khamene).

¹ Present address: Technical University of Munich, Boltzmannstr. 3, 85748 Garching, Germany.

the outer surface of the patient's body (Nutting, 2000). Adjustment of the patient setup can also be achieved using internal anatomical landmarks identified on portal images acquired with the patient in the treatment position. Portal images can be acquired using an electronic portal imaging device. The process of patient position verification using the EPID is explained in detail in Herman et al. (2001) and is summarized as follows:

- At the planning stage a set of DRRs is generated from the CT scan of the patient. This set of images is used as the *Reference*, representing the ideal patient position at the time of treatment.
- At the treatment stage a set of *Localization* portal images is taken at the same gantry angles as the *Reference* images.
- *Localization* and *Reference* images are visually compared and corresponding anatomical landmarks are identified in both sets.
- The setup error is estimated and the patient position is corrected to reduce the discrepancies between the *Localization* and *Reference* images.

The whole process of manual position verification using portal images is rather subjective and error prone. Furthermore, it is difficult to correct for all degrees of freedom (i.e., translations and rotations) based on such a procedure (Gilhuijs et al., 1996a; Hanley et al., 1995; Lujan et al., 1998).

In this paper, we propose an automatic method for estimating the patient pose (i.e., position and orientation – fully six-dimensional) with respect to the pre-treatment CT volume. This method is based on registration of the intra-treatment (i.e., portal images) and pre-treatment (i.e., planning CT) data. The couch position offsets can be computed based on the sought after registration parameters, and the iso-center location in the pre-treatment CT data. In order to perform this registration accurately, the geometry of the treatment machine and portal imaging device and the process of portal image formation have to be known. In other words, both radiometric and geometric properties of the process have to be calibrated. We show that although the portal images are of low contrast, they contain enough information to yield accurate registration. Furthermore, we perform a comparative study for various options (e.g., choice of similarity measure, optimizer, setup parameters) that arise throughout the registration process.

2. Background

The first step in setup verification involves determining the position of the patient on the treatment couch. From an algorithmic perspective, this means extracting a rigid transformation between the pre-treatment plan-

ning CT coordinate frame and that of the LINAC, which is represented by one or several portal images. Thus, in order to perform patient positioning using portal images, we need to solve a medical image registration problem (Maintz and Viergever, 1998). Image registration algorithms fall into two categories: *feature-based* and *intensity-based*. Feature-based methods usually have a segmentation step, during which a number of features, either artificial (e.g., implanted radiopaque markers) or natural (i.e., usually bony anatomy) landmarks, have to be extracted both from the planning CT and the portal images. Intensity-based methods, on the other hand, involve the computation of DRRs. DRRs are two-dimensional projection images, which are computed from the planning CT volume, by knowing the geometry of the LINAC/EPID system. Matching of the DRRs and portal images may either be 2D (i.e., three degrees of freedom; two in-plane translations and one in-plane rotation), or 3D (i.e., six degrees of freedom; three translations and three rotations). Once the rigid transformation is found, we can compute the couch position offsets and move the couch (depending on the degrees of freedom that is provided) in order to line up the planned iso-center in the pre-treatment CT with the actual LINAC iso-center. If the treatment couch provides additional degrees of freedom for rotation, it can be used to also adjust the orientation of one or two planes passing through the iso-center (Hurkmans et al., 2001).

There are a few feature-based methods described in the literature (e.g., Bijhold et al., 1991; Bijhold, 1993; Gilhuijs et al., 1996a; Michalaski et al., 1993; Remeijer et al., 2000). These methods are based on matching of the delineated/segmented bony (visible) structures or some feature points from the portal images and the reference DRR images. In (Bansal et al., 1999), authors proposed an iterative automatic segmentation and registration framework, where pose estimation and labeling of the portal images are done simultaneously. Implanted radio-opaque markers have also been used to recover patient setup error in 3D, which includes out-of-plane rotations (Lam et al., 1993). In Lavalée and Szeliski (1995), the authors describe a method for recovering the pose based on matching projected contours of some 3D shapes. Feature-based methods are in general appealing mostly because the computational complexity is low relative to their intensity-based counterparts. The main reason is that the number of features and in turn the problem dimensions are smaller. However, the solution critically depends on the robustness of the extracted features. In other words, the burden has been shifted to the segmentation part, which is a challenging problem by itself, especially in the case of portal images containing little tissue contrast.

On the other hand, intensity-based methods generally use portal images and DRRs to recover the patient setup error directly. These methods do not require

pre-processing steps, which include extracting features and/or segmentation. There are some 2D methods reported that work based on registering portal images and DRRs using correlation coefficients (Dong and Boyer, 1995; Hristov and Fallone, 1996). There are also a variety of techniques for the general 2D/3D registration application (LaRose, 2001; Penney et al., 1998, 2001; Tomazevic et al., 2003; Weese et al., 1997). Most of these methods have targeted a pre-operative CT and intra-operative X-ray fluoroscopy image registration application, which is quite similar to patient positioning and setup verification in radiation therapy. In most of these methods, it is required to generate DRRs in an iterative optimization loop. Computation of DRRs is expensive by nature. In order to tackle this problem, a variety of approaches have been investigated. Some researchers have proposed using only partial DRRs (Zöllei et al., 2001). In LaRose (2001), the author used a graphics accelerated method for fast DRR computation. In LaRose et al. (2000), Rohlfing et al. (2002), the authors proposed a pre-computation step during which a large set of line integrals with various starting positions and orientations are computed. In the on-line mode, the DRRs are generated based on pre-computed line integrals. If the desired DRR contains line integrals that have not been pre-computed, interpolated ones among close ones in the database is used instead. There are drawbacks to this

approach. First, in order to get high quality DRRs, one has to pre-compute the line integrals with very fine sampling, which will make the size of the pre-computed database prohibitively large. This requirement will be more evident if the field of view of the DRR, or in other words the divergence of the rays, is rather large. Second, the spatial interpolation between pre-computed line integrals results in severe artifacts in the estimated DRRs. Again, the degree of artifacts is more noticeable for DRRs with a large field of view. These effects collectively reduce the robustness of the registration process, which heavily relies on the quality of the generated DRR to estimate/refine the pose. In Gilhuijs et al. (1996b), authors proposed an interactive method that is based on fast generation of DRRs. This method requires good operator training especially for non-orthogonal fields.

For intensity-based algorithms, the choice of similarity measures and the optimization methods also play a role in the practicality of the approach (Penney et al., 1998; Rohlfing et al., 2002; Sarrut and Clippe, 2000). Several papers have been published proposing and/or investigating various registration methods using portal images for patient positioning applications. However, there are some issues, which need to be further investigated.

For example, geometric calibration of LINAC source (see Fig. 1) and EPID is often not thoroughly addressed.

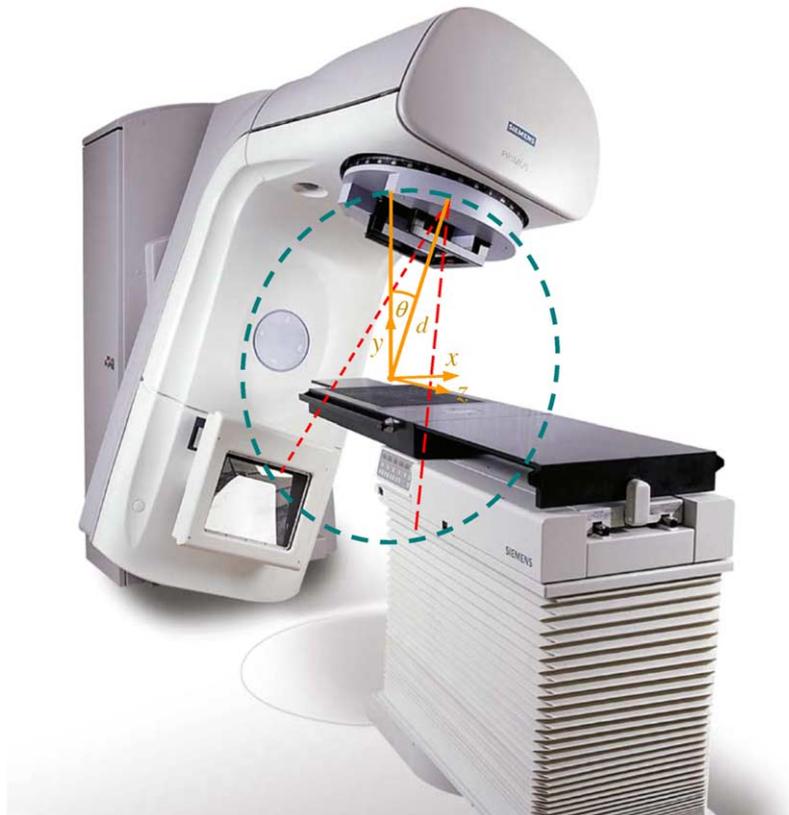


Fig. 1. Linear accelerator with overlaid iso-centric coordinate system.

Geometric calibration parameters consist of two sets of variables defining first the iso-centric beam geometry and second the iso-centric FPD (Flat Panel Detector) position. The field edge matching technique, which is often used as a means to estimate the parameters of projection geometry of the system (for example [Dong and Boyer, 1995](#); [Kim et al., 2001](#)), can only deal accurately with planar motion of the FPD, usually referred to as detector sag. Precise calibration of the beam geometry is required more importantly for the scenarios, where multiple portal images are used for registration. A fully three-dimensional calibration procedure similar to the one used for cone-beam reconstruction has to be employed ([Cho et al., 2004](#)). Another issue is that the portal images are acquired with quantum radiation energy in Mega electron-Volt (MeV) range orders of magnitude higher than the kilo electron-Volt (keV) range used for CT projections. This energy discrepancy alters the type of anatomy that is visible in the portal images ([Bansal et al., 1999](#)). In the intensity-based registration approach, it is crucial to take these energy variations into account within the DRR generation algorithm. This consequently affects the utilization and robustness of the similarity measure. Finally, there has not been a systematic and comprehensive evaluation of various setup parameters such as number of portal images, the angles between the images, and the image field size, on the overall registration procedure. In this paper, we outline a registration procedure in which these relevant practical factors are considered. Furthermore, we design and carry out experiments to investigate the issues, which are mentioned earlier.

3. 2D/3D Registration of portal images and CT volume

Using intensity-based methods for registration of DRRs and portal images, we determine the rigid trans-

formation relating the iso-centric coordinate system of the treatment machine to that of pre-treatment CT. We use one to several portal images taken at various gantry angles to achieve more robustness. Let us divide the whole process into three major phases, *Calibration*, *Planning*, and *Treatment positioning* (see [Fig. 2](#)). In the calibration step, the focus is to estimate the beam geometry at each gantry angle, while isolating any detector sag during rotation. Discarding the detector sag is done through a rectification process, which will be explained later. In the planning stage, CT scan data are used to define the planned target volume (PTV) and to devise the treatment plan. In the treatment positioning phase, the portal images are taken and rectified in order to discard the gantry sag. This is because the internal camera calibration parameters estimated in the calibration phase are only valid for rectified images. Furthermore, an iterative optimization loop maximizes the similarity between the computed DRRs and portal images by operating on the pose parameters. Once the transformation between CT and LINAC is established, the treatment couch offsets (both translational and rotational) are computed and applied to realize the treatment plan. The procedure to obtain the transformation is outlined in this section.

3.1. Problem formulation

Let us denote the CT volume by V , and let P^θ represent the portal image taken at the gantry angle θ . The transformation operator $\mathcal{T}(\cdot)$ maps the CT volume to the LINAC coordinate system. This mapping is done based on the homogeneous matrix \mathbf{T}_{CT}^{LINAC} (we use \mathbf{T} from now on). Furthermore, the projection operator $\mathcal{P}^\theta(\cdot)$, generates a DRR from the transformed CT data set. Finally, the similarity measure computation for two images is done by the operator $\mathcal{S}(\cdot, \cdot)$. The combination of the two operators $\mathcal{T}(\cdot)$ and $\mathcal{P}^\theta(\cdot)$, once

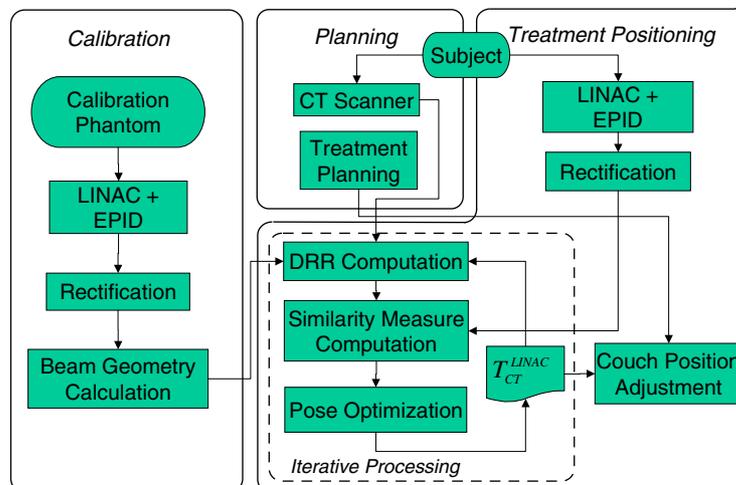


Fig. 2. Flowchart of the process.

applied on the CT volume V , provides the floating DRR \tilde{P}^θ (defined up to \mathbf{T}), which should be compared against the reference portal image P^θ . Therefore, we have:

$$\tilde{P}^\theta = \mathcal{P}^\theta(\mathcal{F}(V)). \quad (1)$$

The goal of the registration is to compute the CT to LINAC transformation \mathbf{T} , maximizing the similarity measure, as follows:

$$\tilde{\mathbf{T}} = \arg \max_{\mathbf{T}} \mathcal{S}(\tilde{P}^\theta, P^\theta). \quad (2)$$

In the case of multiple reference portal images and multiple floating DRRs, an additional issue is how to combine the similarity measures. Our approach is to concatenate the images (either row-wise or column-wise) and compute one similarity measure on the larger images. Some similarity measure algorithms, which use neighborhood information, e.g., 2D gradients, have to be adapted to omit the borderline between the images in order not to produce incorrect results.

3.2. Geometry

In this section, we describe how to characterize the operator \mathcal{P}^θ . As depicted in Fig. 1, we establish a coordinate frame at the iso-center of the LINAC. Assuming the beam geometry does not change for different gantry angles, the following equation can be used to derive the 3×4 projection matrix at given angle θ :

$$\mathbf{P}^\theta = \mathbf{\Pi} \mathbf{H}_\theta, \quad (3)$$

where $\mathbf{\Pi}$ is a 3×4 matrix containing the internal camera parameters of the LINAC, and \mathbf{H}_θ is the 4×4 homogeneous transformation matrix converting point coordinates from an iso-centric system to the one established at the LINAC source. The transformation matrix \mathbf{H}_θ can simply be specified as follows:

$$\mathbf{H}_\theta = \mathbf{R}_x(90^\circ) \mathbf{T}_y(d) \mathbf{R}_z(-\theta), \quad (4)$$

where d is the distance between source and iso-center, \mathbf{R}_x and \mathbf{R}_z are homogeneous matrices containing rotations about the x and z axes, and \mathbf{T}_y only contains the translation along the y axis.

3.3. Calibration

In the calibration step, we estimate the internal parameters of the LINAC source, packed in the matrix $\mathbf{\Pi}$, and the parameters needed for acquiring the transformation matrix \mathbf{H}_θ in Eq. (4). In order to do so, we use portal images of a cylindrical phantom with radio-opaque tungsten markers (see Figs. 3 and 4). The positions of the markers are known within a pre-defined coordinate system located on the phantom axis. We place the phantom on the treatment couch and take six images at multiples of 60° gantry angle. Let us assume Q_i is the coordinate of the i th radio-opaque marker in the phantom coordinate frame, and q_i^θ is its projection on the portal image, taken at gantry angle θ . The calibration involves recovering eight parameters (i.e., six for the transformation from the phantom to iso-centric coordinate frame, field of view f , and source to iso-center distance d) by minimizing the re-projection error of all the points, as follows:

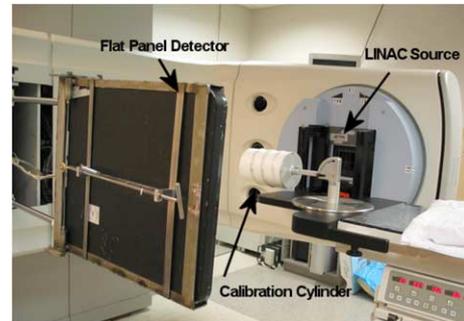


Fig. 3. LINAC calibration procedure.

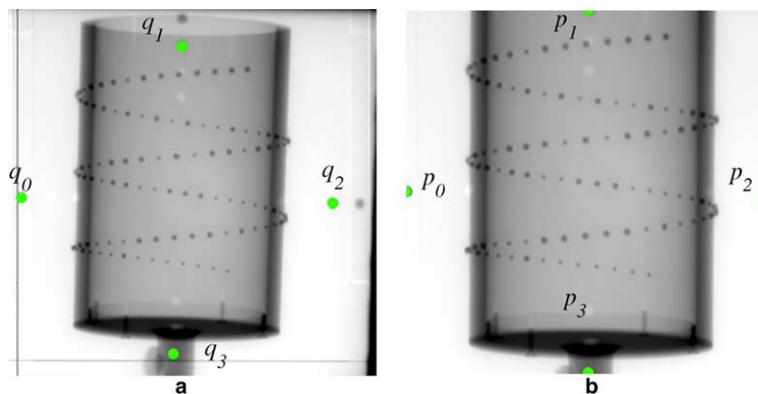


Fig. 4. (a) Depicts a portal image before rectification. The accessory tray markers are labeled as q_0 to q_3 . (b) Shows the rectified image, in which the accessory markers, labeled as p_0 to p_3 , are mapped to the center of the image sides.

$$\{\tilde{\mathbf{T}}, \tilde{f}, \tilde{d}\} = \arg \min_{\mathbf{T}, f, d} \sum_n \times \sum_i \|\mathbf{P}^{60n}(f, d)\mathbf{T}Q_i - \mathbf{M}q_i^{60n}\|^2, \quad (5)$$

where n is an integer in $[0, 5]$, and \mathbf{M} represents a planar motion (i.e., homography) matrix accounting for detector movement, which is also referred to as the rectifying homography (see Appendix B for details). Finally, $\mathbf{P}^0(f, d)$ depicts the projection matrix dependency to the focal length and source to iso-center distance. We use the Levenberg–Marquardt optimization method to minimize the nonlinear cost function in Eq. (5). To ensure success of the optimization procedure, we use a rough closed form solution for \mathbf{T} , f , and d , as the starting values (see Appendix C for details).

3.4. DRR generation

A radiograph can be reconstructed digitally from volumetric CT data, by computing the integral of the attenuation function along a ray passing through the volume at a given energy spectrum. The knowledge of the ray geometry (i.e., the field of view and flat panel detector size) is required in order to generate the DRRs. In our case, this geometrical information is encoded in the projection matrix Π .

We assume that the scattered radiation component at the level of the EPID is small and can be neglected. For example, it has been shown in McDonough et al. (2002) that for 6 and 15 MV photon treatment beams, the scattered to primary ratio with a 60 cm gap between the beam iso-center and the EPID was less than 0.06 and 0.02, respectively (for an irradiated volume of size $20 \times 20 \times 18 \text{ cm}^3$ at the iso-center). Beam hardening is also relatively small for high energy photons. For example, for a 6 MV photon beam, we determined from measured attenuation data (in 30 cm of water) typically less than 6% difference in beam attenuation associated with beam hardening. The difference is even smaller for shorter path lengths and higher photon energies. By overlooking the scattering and beam hardening effects, the intensity values of the radiograph image pixels can be acquired from the following equation:

$$\tilde{P}_{ij} = \int S(E) \left(\exp \left(- \int_{Q \in l_{ij}} \mu(Q, E) dl_{ij} \right) \right) dE, \quad (6)$$

where $S(E)$ is the source spectrum, and $\mu(Q, E)$ is the attenuation function at point Q for various energies E . Furthermore, dl_{ij} constitutes the line integral element corresponding to the point (i, j) on the radiograph \tilde{P} . The direction of the line l_{ij} is determined by connecting (i, j) on the radiograph to the center of projection (i.e., source focal point) that is specified by the matrix Π .

There are some obstacles for reconstructing near perfect portal images like radiographs. The first is that CT

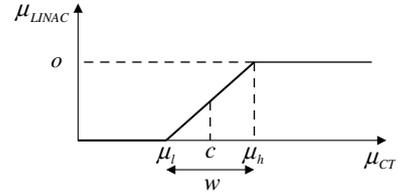


Fig. 5. The conversion function $\mathcal{C}(\cdot)$.

reconstructs $\mu(Q)$ at an effective energy, and does not provide energy dependent attenuation. Second, the $S(E)$ for LINAC is in MeV range, but $\mu(Q)$ has been reconstructed from CT using the source in keV range. We have to make some assumptions to be able to address these problems. First, we assume that CT attenuation $\mu(Q)$ is reconstructed using CT projections with an effective beam energy E_{CT}^0 . Second, the treatment beam effective energy is E_{LINAC}^0 . Moreover, there exists a functional dependency $\mathcal{C}(\cdot)$ between $\mu(E_{CT}^0)$ and $\mu(E_{LINAC}^0)$. Therefore, Eq. (6) can be simplified as follows:

$$\tilde{P}_{ij} = \exp \left(- \int_{Q \in l_{ij}} \mathcal{C}(\mu(Q, E_{CT}^0)) dl_{ij} \right). \quad (7)$$

We characterize the conversion operator $\mathcal{C}(\cdot)$ as a linear function within a limited support of $[\mu_l, \mu_h]$, as shown in Fig. 5 (see Hubbell, 1982; Milickovic et al., 2000 for details). Three parameters characterizing $\mathcal{C}(\cdot)$ are the width of the support w , the center of the support c and the saturation value o . We estimate the parameters of the conversion function $\mathcal{C}(\cdot)$ by minimizing the difference metric between a sample portal image and the corresponding DRR. This so-called radiometric calibration has to be done only once for each pair of LINAC and EPID. The difference metric can also be defined in a way that represents only global brightness and contrast differences. This relaxes the requirement of having one pair of registered DRR and portal image. The radiometric calibration procedure can simply be formulated as follows:

$$\{\tilde{w}, \tilde{c}, \tilde{o}\} = \arg \min_{w, c, o} \|\tilde{P}(w, c, o) - P\|^2. \quad (8)$$

We take advantage of 3D texture-based volume rendering techniques to generate DRRs (Cabral et al., 1994). The method, which utilizes the graphics processing unit (GPU), yields better computational efficiency. In this approach, a 12 or 16 bit CT volume is loaded into graphics memory and two dimensional cross-sectional slices are computed perpendicular to the dynamic viewing direction and integrated from back to front with a certain transparency value, implementing Eq. (7). The attenuation conversion operator is implemented using dependent texture mapping capabilities, now available in most graphics cards. A dependent texture constitutes the mapping that converts 12 or 16 bit attenuation

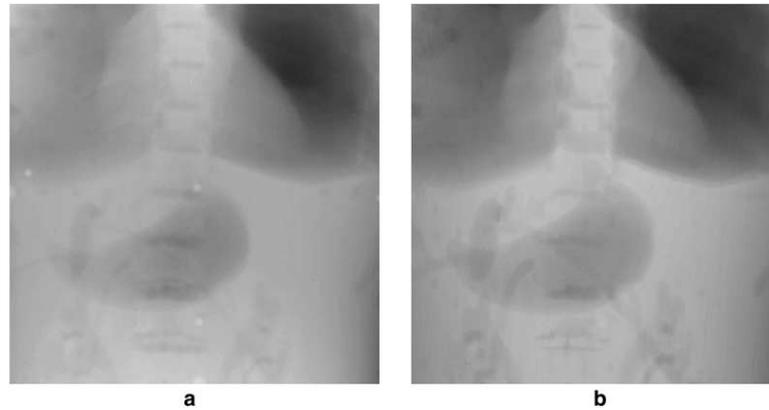


Fig. 6. (a) Portal image taken from the Rando[®] body phantom, (b) DRR generated with 3D texture mapping/dependent texture, corrected for energy discrepancies.

values of a 3D volume to new values. The rendering then happens using the mapped attenuation coefficients. The dependent texture mapping procedure is performed without performance penalty in the graphics cards, which support such an operation (e.g., a GeForce FX 5600 from NVidia Corporation,² Santa Clara, CA).

Using this technique, we reconstruct high quality DRRs very efficiently. It takes about 60 ms to generate 256×256 DRRs from a $256 \times 256 \times 216$ volume with a NVidia FX 5600 card with 256 megabytes of texture memory. The computation time reduces to 35 ms once the volume size is reduced to $128 \times 128 \times 108$. Fig. 6 depicts the quality of the DRRs generated using this method. Radiometric calibration plays the central role in making the DRRs look very similar to the corresponding portal image.

3.5. Similarity measures

Various similarity measures have been proposed and tested for medical image registration algorithms. There is a large body of literature on this issue, and readers may refer to them for detailed descriptions of each similarity measure and its formulation (Hipwell et al., 2003; Penney et al., 1998; Roche et al., 1998; Sarrut and Clippe, 2000; Wells et al., 1996).

As shown in Eq. (2), a similarity measure can be modeled by an operator, which takes two images and returns a scalar value representing the level of resemblance between them. An ideal similarity measure peaks once the two images are geometrically registered, regardless of the radiometric discrepancies between them. Furthermore, it smoothly decreases as one of the images is geometrically changed. In other words, a perfect similarity measure should be able to discriminate between geometric and radiometric changes in the

images. Various similarity measures can tolerate and disregard radiometric variations.

The simplest image dissimilarity measure is in fact the sum of square differences (SSD) of the pixels of the two images. In order for this measure to perform, the images have to be radiometrically identical. In the correlation coefficient (CC) computation, the assumption is that there exists a linear dependency between the intensity profile of the two images. In other words, only global contrast and brightness variations in the images can be disregarded. Local normalized correlation (LNC) alleviates the latter assumption by subdividing the images into blocks and computing the correlation coefficient for each block separately, averaging the results to get a scalar value. Variance weighted correlation (VWC) is similar to LNC, the only difference is that corresponding blocks correlation coefficients are weighted by the block's energy before averaging. Gradient difference (GD) computes the sum of square differences (or some other error function) over the gradient images, assuming the intensity gradients are more robust than the original pixel intensities. On the other hand, gradient correlation (GC) measures the correlation coefficients of the gradient images, enforcing a less restrictive assumption of having the gradient images linearly correlated. Pattern intensity (PI) directly assesses the patterns in the difference image. The measure reaches its maximum once there is no "pattern" present in the difference image (Penney et al., 1998).

Besides the deterministic (dis)similarity measures, we have implemented a family of stochastic similarity measures. The correlation ratio (CR) assumption is that the intensity probability distributions of the two images are functionally dependent (Roche et al., 1998). The measure estimates the strength of the functional dependency. Entropy estimators gauge the information content of random variables. Overlapping information content of two variables can be computed once one knows how to estimate the entropy of a random variable. Mutual

² www.nvidia.com

information (MI) is the overlapping information content computed based on Shannon's entropy (Wells et al., 1996). MI as a similarity measure is appealing since it does not have a restrictive assumption over the image intensities.

The similarity measure between a pair of floating and static images can be thought of as the cost function defined over the transformation parameters of the floating image. The general properties of cost functions (i.e., convexity, smoothness) dictate the robustness of the optimization algorithm in finding the extremum. Cost function properties have a direct link to the type of the similarity measure and the nature of the data. One of our goals in this paper is to investigate the robustness of the various similarity measures for the portal image to CT registration problem.

3.6. Optimization

Optimization procedures involve finding a set of independent variables, in the domain of a cost function, which take the function to a maximum or minimum. In our application, the cost function is the similarity measure between the portal image and the floating DRR as shown in Eq. (2). The floating DRR is dependent on the transformation between pre-treatment CT and the LINAC coordinate systems. Consequently, the similarity measure is the function of the same six independent variables mentioned earlier. The optimization constitutes a search algorithm through the parameter space with a specific iterative strategy. Optimization methods can be divided into two major categories based on using the derivatives of the cost function. We implemented and used a gradient based (e.g., *Gradient Descent*; Griewank, 1981) and various non-gradient based (e.g., *Best Neighbor Search*; Törn and Žilinskas, 1989), and *Powell–Brent* (Brent, 1973; Powell, 1964) methods for the application at hand. One problem in global optimization techniques is the possibility of being trapped in a local optimum. This could be alleviated in our application by designing/choosing a similarity measure, which yields a smoother cost function. There are also some techniques to address this problem by taking advantage of redundant optimization passes with randomized starting points or some randomized steps, which has the trade-off of increasing the overall computation time (Gorse et al., 1994). Furthermore, having a large discrepancy between the reference and floating images also increases the chance of being trapped in local optima and may eventually lead to failure of the optimization process as a whole. In order to tackle this problem, we consider two approaches. In the first approach, the user has the option to maneuver the pre-treatment CT volume in six-dimensional space, and observe the generated DRR interactively in order to find a reasonably good starting point for the automatic registration procedure

[a similar approach has been reported in Gilhuijs et al. (1996b)]. The second approach is to perform multi-optimization passes on a pyramid of images with dyadic resolution starting from the lowest. Lower resolution images produce smoother cost functions, decreasing the chance of having local optima problems. Once the optimization is converged on one level, the result will be used as the starting point of the next higher resolution level. We choose to have three levels of dyadic resolution starting from the highest level at 256×256 resolution.

Another important consideration is the way the steps for searching/probing the parameter space are defined. The search steps at each dimension of the parameter space have to be ideally arranged in a way that the cost function changes are homogeneously sensible. The latter directly translates to the appearance change (consequently, changes in the similarity measure value) in the generated DRRs as we change each parameter of the search space. It is widely known that once the focal point of the source is far from the object, lateral translations are virtually indistinguishable from out-of-plane rotations (Young and Chellappa, 1992). This causes the cost function not to have required observability to arrive at a robust solution. To overcome this, we always compute the search steps within the coordinate system placed at the center of the object. The scaling factor for rotational steps is also set in a way that the total movements of the bounding box vertices of the object are identical to the translational steps. In this way, we make sure that the cost function changes with regard to each parameter variation are at least in the same order of magnitude. Moreover, we avoid overoptimization of a subset of parameters (Törn and Žilinskas, 1989).

3.7. Couch position adjustments

As it is explained in Section 1, the treatment begins with specifying the targeted volume in the planning CT data set. The position of the beam iso-center is also specified in the planned target volume. The transformation T specified through the optimization process in Section 3.6, along with the specified iso-center in the PTV (let us call it \bar{C}) can be used to compute the treatment couch adjustments. The nature of the adjustment also depends on the number of movements (i.e., degrees of freedom), which is provided by the couch top mechanics. Most of the modern couches support at least three translations (i.e., $\Delta t_x, \Delta t_y, \Delta t_z$), and one rotation ($\Delta \phi_y$). In addition, couch tops are becoming available that permit roll and tilt ($\Delta \phi_z$ and $\Delta \phi_x$, respectively) (Litzenberg et al., 1999).

First, let us assume that the axial plane in the CT planning coincides with the plane passing through the LINAC iso-center and trajectory of the LINAC for various gantry angles (i.e., the LINAC *axial plane*). In

this case no rotational adjustment is necessary. The couch translation that maps the CT planned iso-center to the LINAC iso-center can be calculated as follows:

$$\Delta \mathbf{t} = -\mathbf{T}\mathbf{C}, \quad (9)$$

where \mathbf{C} is the homogeneous coordinates of \bar{C} , and $\Delta \mathbf{t} = [\Delta t_x, \Delta t_y, \Delta t_z, 1]$. Let us also consider a case, where two perpendicular planes with the normals of \mathbf{n}_x and \mathbf{n}_z are specified in the PTV, and have to be coincided with the LINAC's sagittal and axial planes, respectively. In order to realize this setup, we need to have additional angular corrections. The angular corrections can be derived as follows:

$$\Delta \phi = \mathcal{A} \left((\mathbf{R}_T [\mathbf{n}_x | \mathbf{n}_x \times \mathbf{n}_z | \mathbf{n}_z])^{-1} \right), \quad (10)$$

where $\Delta \phi = [\Delta \phi_x, \Delta \phi_y, \Delta \phi_z]$, \mathbf{R}_T is 3×3 rotation part of the homogeneous matrix \mathbf{T} , and \mathcal{A} is the operator that decomposes rotation matrices into three consecutive rotation angles about x , y , and z , based on the Euler–Cardan formula. In cases, where couch tops do not support all three angles, the ability to map any two planes from PTV to the corresponding LINAC ones will also be restricted.

4. Experimental results

A series of experiments was performed to test and validate (1) the implementation, (2) merit of various similarity measures and (3) the efficiency of the preferred algorithm for various setup parameters. The performance measure in all cases was calculated based on the mean and standard deviation of distance in millimeters between the computed couch position adjustment vector from the intensity based algorithm (i.e., $-\mathbf{T}\mathbf{C}$) and that from the “gold standard” pose (i.e., $-\mathbf{T}_0\mathbf{C}$, where \mathbf{T}_0 is the gold standard transformation). In other words, the error measure is defined to be the target registration error (TRE) of the planned iso-center C :

$$e = \text{TRE}(C) = \|\mathbf{T}_0\mathbf{C} - \mathbf{T}\mathbf{C}\|. \quad (11)$$

We use the error measure in Eq. (11) throughout the paper. The planned iso-center for these calculations was considered to be the point at the center of the CT volume. For all tests, the initial poses consisted of 100 randomly selected poses with uniform independently distributed translations in millimeters within $[-10 \ 10]$ and rotations in degrees within $[-10 \ 10]$ around the gold standard pose.

In the first set of experiments, synthetic data were used. Tests were performed on the CT volume of a Rando[®] head phantom (The Phantom Laboratory, Inc.,³ Salem, NY) with voxel size $256 \times 256 \times 216$. We gener-

ated one or multiple DRRs from the CT volume. We used the DRR(s) as reference portal image(s). Ground truth pose in this case was the pose, at which the DRRs were generated. We also set the initial pose for the registration process by random variation of the ground truth pose. Multiple runs of the registration process using the simulated portal images from various gantry angles yield normal-like distribution of the TRE with a mean of 0.1 mm. These results were more or less regardless of the choice of similarity measure and optimization procedure. They validated the implementation of the algorithms, permitting further tests on real portal images of a phantom and patients.

4.1. Experiments with phantom data

We used portal images of a Rando[®] body phantom to validate the registration procedure. For this set of experiments, the beam was generated by a Siemens Primus LINAC (Siemens Medical Solutions Inc., Concord CA, USA). The portal imaging device was an amorphous silicon detector from BioScan (Bio Scan Inc., Washington, DC, USA). A CT Volume of the Rando[®] body phantom was acquired using a Siemens Somatom CT scanner (Siemens Medical Solutions Inc., Forchheim, Germany).

We attached four radio-opaque markers on a plate, which was then placed into the accessory tray of the LINAC to be able to account for flat panel detector movements (see Section 3.3). The LINAC system was then calibrated by taking six portal images of the calibration phantom as it was explained in Section 3.3.

In order to establish a ground truth (i.e., gold standard pose), 16 radio-opaque tungsten markers were positioned on the Rando[®] body phantom. The phantom was then CT scanned. The CT volume original size was $512 \times 512 \times 216$ with 0.9 mm slice resolution and 3.0 mm slice thickness. For the registration procedure, we down-sampled the images to reduce the total volume size to $256 \times 256 \times 216$. The phantom was then placed on the treatment couch and 60 portal images were acquired from various gantry angles. After the image rectification procedure, we segmented out the radio-opaque markers from both the CT volume and portal images. We then computed the gold standard pose by matching up the 2D location of the radio-opaque markers and their corresponding 3D locations in the CT volume through the minimization shown in Eq. (5). The only difference here was that the two parameters f and d had already been specified through the calibration procedure. The LINAC plus EPID pair was also radiometrically calibrated using the method outlined in Section 3.4.

We performed a series of registrations using one or two portal images with 6° , 12° , and 90° vergence angles. The initial pose for all these tests was generated by randomly changing the “ground truth” pose using the

³ <http://www.phantomlab.com>

procedure previously discussed. The registration results with random initial poses provide the basis for the error analysis (see Fig. 7).

The error analysis (mean and standard deviation of TRE, as explained in Eq. (11)) for various similarity measures is depicted in Fig. 8. The registrations were performed using a pair of portal images with 90° vergence angle. The non-stochastic similarity measures LNC, GC, and PI performed better than the information theoretic ones such as MI and CR. The simple explanation is that the high quality and the degree of resemblance of DRRs to the portal images causes the measures with more restrictive assumptions to perform more robust. We also examined the effect of using various optimization techniques, namely the Best Neighbor Search, Gradient Descent, and the Powell–Brent method. The mean and standard deviation of the TRE for various registration experiments using different optimization techniques revealed no significant difference. In other words, the type of optimization technique did not seem to be relevant.

Fig. 9 shows the error analysis results for the two measures LNC and MI for pairs with various vergence angles. It is also intuitive to observe that as we use two portal images instead of one and increase the vergence angle, both mean and standard deviation of the TRE decrease. The mean TRE doubles once the vergence angle is decreased from 90° to 12°, however, it still remains less than 3 mm. The 68.3% ellipsoids represent-

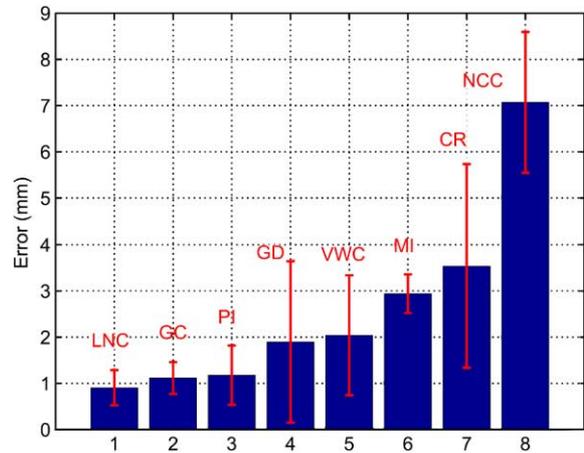


Fig. 8. TRE mean and standard deviation for registration using 0–90° portal images for various similarity measures.

ing the spatial distribution of the TRE is also depicted in Fig. 11. This figure shows that the error magnitude is higher in the viewing direction compared to the other directions and the error distribution tends to become more isotropic as the vergence angle increases. The lateral in-plane error distribution is limited to -1.5 to 1.5 mm, which might be sufficient for some applications.

Fig. 10 (a) shows the error analysis results based on the registration using two portal images taken at 0° and 90° with various field sizes. It is always desired to reduce the

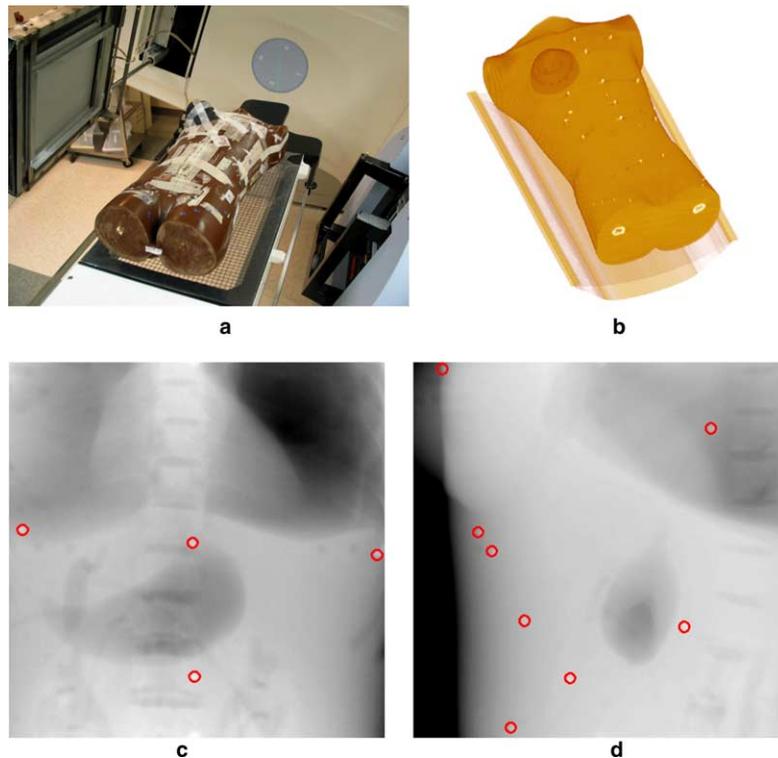


Fig. 7. (a) Rando® phantom on treatment couch, (b) CT volume of the Rando® phantom, radio-opaque markers appear as bright spots, (c) and (d) portal images taken at gantry angle 0° and 90°, in which the radio-opaque markers also show up as bright spots.

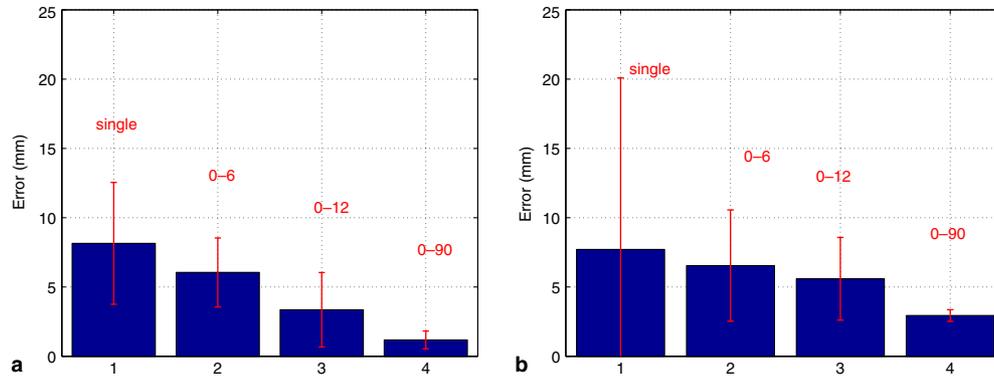


Fig. 9. TRE mean and standard deviation for portal image registration using (a) LNC and (b) MI similarity measures for various gantry angles.

amount of unnecessary radiation to the patient by decreasing the field size. These results show that by reducing the field size from 41 to 17 cm² the average error increases from less than 1 to about 3 mm. Having a field size of 25 cm² seems to be a good compromise for this specific data set as it yields a mean TRE better than 1.5 mm.

Processing times for the total registration procedure have been graphed in Fig. 12. The average number of iterations was 88 for the stereo case with 90° vergence angle, and 101 for the single portal image case. The smaller number of iterations in the stereo case shows that the cost function was more well-defined and stable compared to the single portal image case. Finally, Figs. 13–15 depict the sample DRR images at start and end of the optimization process overlaid by outlines of the portal images for comparison. These figures depict that using this algorithm, we can recover translational and rotational discrepancy in range of –10 to 10 mm and –10° to 10°, respectively, within a reasonable amount of time.

4.2. Experiments with patient data

In absence of radio-opaque markers, which can provide a gold standard registration pose, our patient

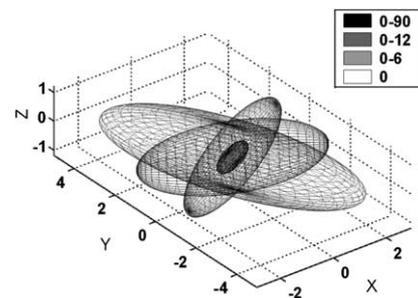


Fig. 11. Spatial distribution of the iso-center TRE for registration using single or stereo portal pairs with various vergence angles.

experimental results can only be verified qualitatively. Figs. 16 and 17 show the registration results for two different patients. Fig. 16 (a) shows the left lateral (LL) portal image of a patient’s brain, (b) depicts the DRR generated at the estimated pose, and (c) shows the estimated position of the patient planning CT volume in the LINAC coordinate system. Fig. 17 shows the same sequence for another patient’s abdomen, which was imaged in anterior–posterior (AP) direction.

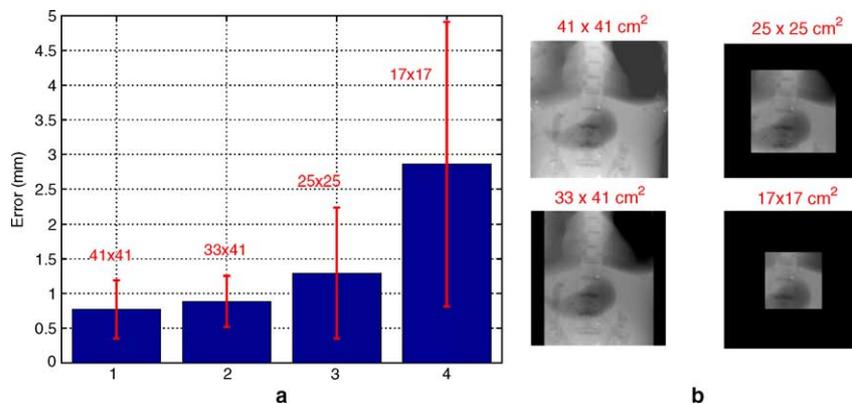


Fig. 10. (a) TRE mean and standard deviation for portal image registration using LNC similarity measures for various field sizes. (b) Sample portal images of various field sizes.

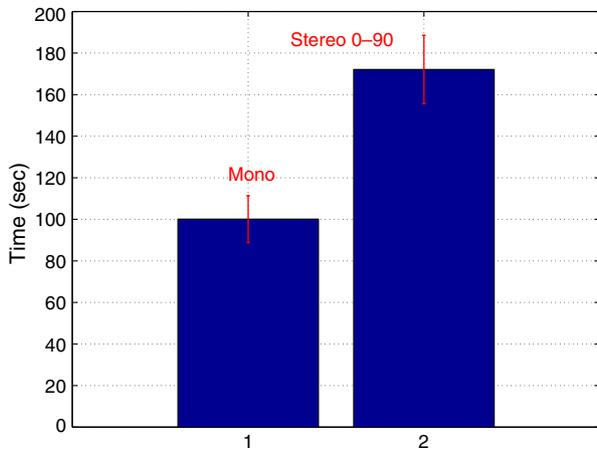


Fig. 12. Mean and standard deviation of total processing time.

5. Summary and conclusion

In this paper, we presented and evaluated a method for automatic patient positioning in radiation therapy. The centerpiece of the approach is geometric and radio-

metric calibration of the LINAC. The detector sag problem has been addressed in the geometric calibration process. The radiometric calibration parameters are then exploited in the DRR generation procedure, which yields high quality images. We also performed the DRR generation through the graphics pipeline, which increases the computational efficiency while preserving the quality. We have done a comparative analysis of various similarity measures and optimization methods. The conclusion was that simple similarity measures such as LNC yield more accurate results as long as the radiometric calibration is carefully performed. Furthermore, various optimization schemes did not change the final error measures. Error analysis in a phantom study showed that the patient positioning can be achieved with less than 1.0 mm mean TRE for large field (i.e., 41 cm²) portal images with 90° vergence angle. The mean error increased to about 1.5 mm as the field size was reduced to 25 cm². Further error analysis on patient data has to be performed in the future. More importantly, the efficacy of the automatic registration algorithm on the overall treatment procedure has to be evaluated.

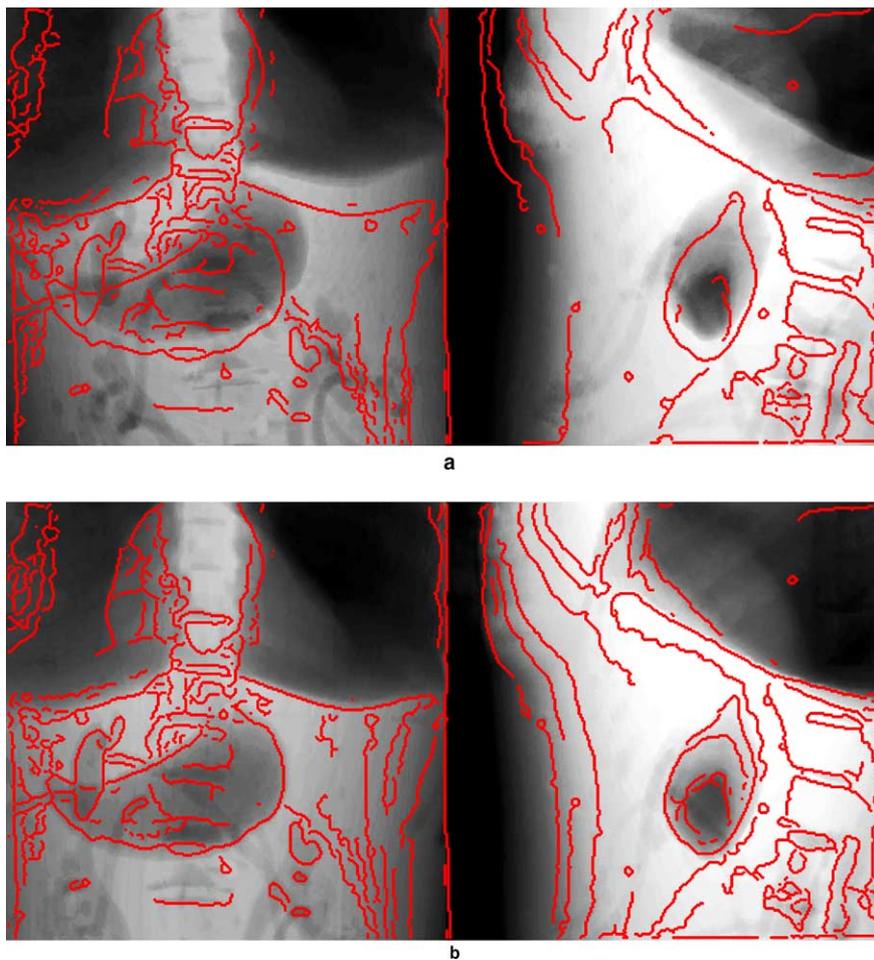


Fig. 13. Sample DRR pair taken at 0° and 90° gantry angles, overlaid with outlines of corresponding portal images (a) before, and (b) after registration.

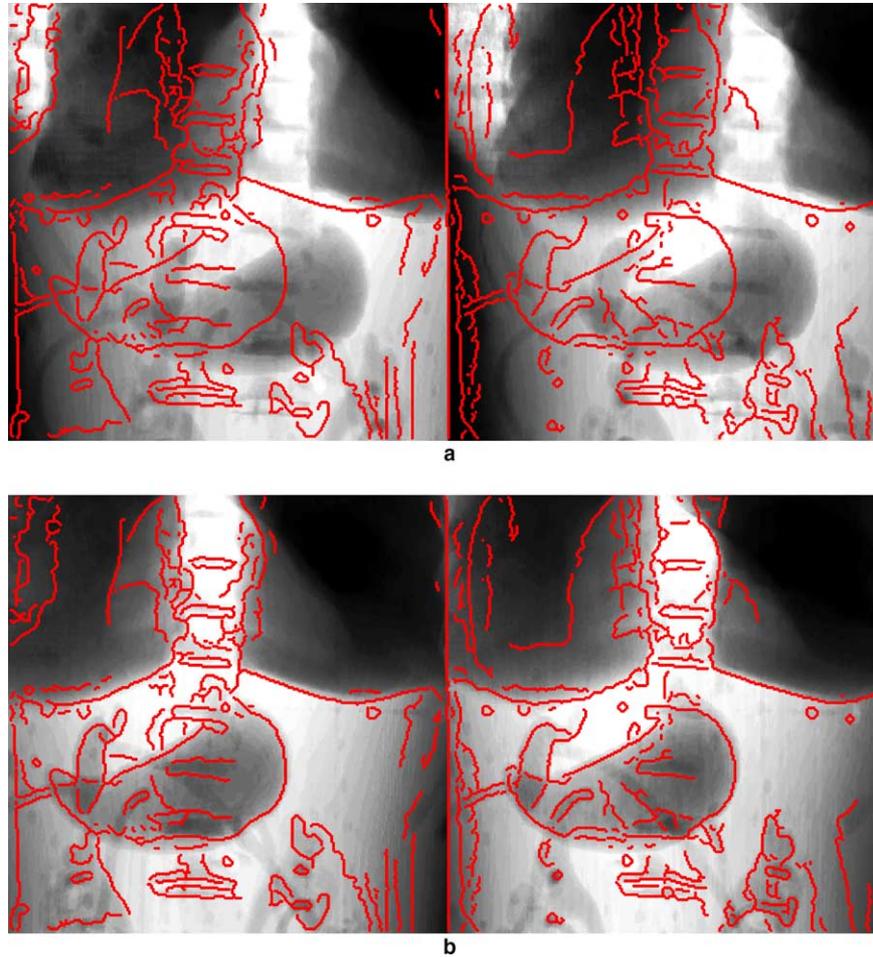


Fig. 14. Sample DRR pair taken at 0° and 12° gantry angles, overlaid with outlines of corresponding portal images (a) before, and (b) after registration.

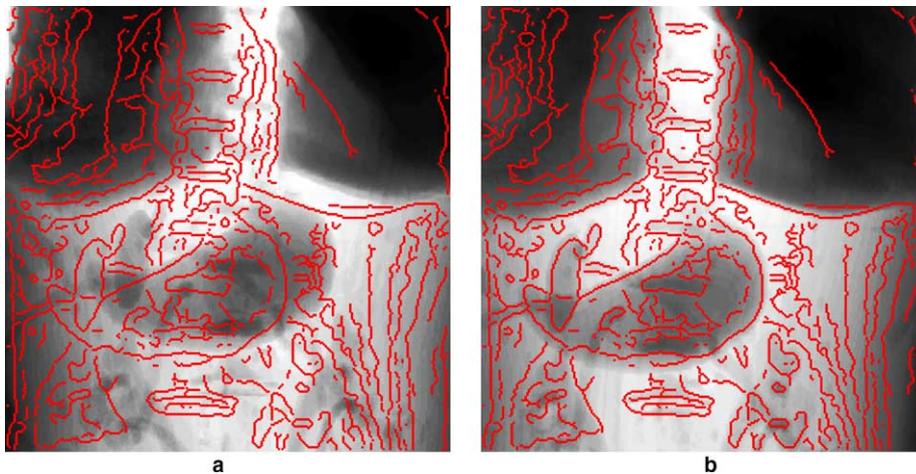


Fig. 15. Sample DRR image taken at 0° gantry angle, overlaid with outlines of corresponding portal image (a) before, and (b) after registration.

Appendix A. Projection matrix computation

Let us assume that a 3D point \bar{Q}_i is projected under a 3×4 projection matrix \mathbf{P} onto a 2D image point \bar{q}_i .

Therefore, we will have $q_i \equiv \mathbf{P}Q_i$, where q_i and Q_i are homogeneous coordinates of points \bar{q}_i and \bar{Q}_i and \equiv depicts an up-to-scale equality. The previous equation can be re-written as follows:

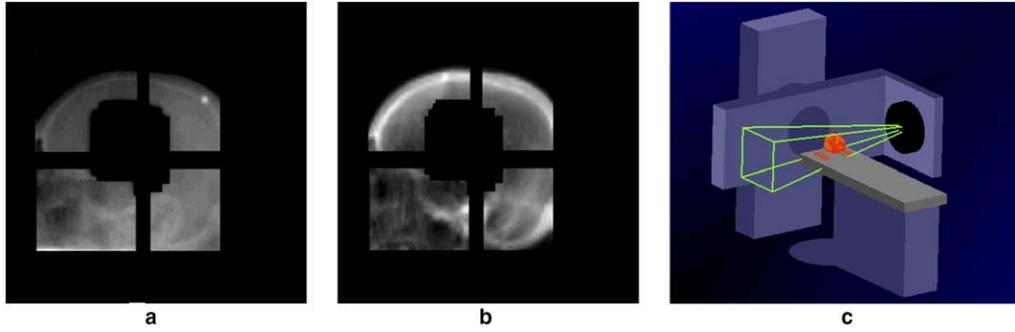


Fig. 16. (a) Shows portal image of the brain, (b) depicts the matching DRR image, and (c) shows the estimated patient's position on treatment couch.

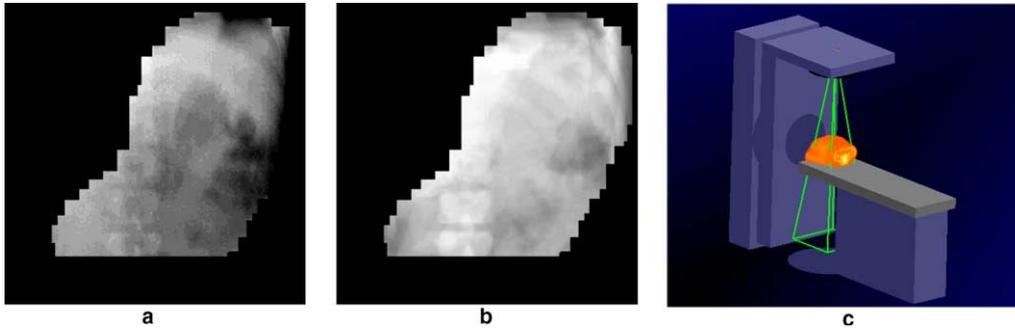


Fig. 17. (a) Shows portal image of the abdomen, (b) depicts the matching DRR image, and (c) shows the estimated patient's position on treatment couch.

$$\begin{bmatrix} Q_{ix} & Q_{iy} & Q_{iz} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & Q_{ix} & Q_{iy} & Q_{iz} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Q_{ix} & Q_{iy} & Q_{iz} & 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{P}_{11} \\ \mathbf{P}_{12} \\ \vdots \\ \mathbf{P}_{3,4} \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} q_{ix} \\ q_{iy} \\ 1 \end{bmatrix}, \quad (\text{A.1})$$

where λ is an unknown scaling factor. In Eq. (A.1), if we call the 3×12 matrix \mathbf{Q}_i , the 3×1 vector q_i , and assign P to be the row-wise unwrapped version of \mathbf{P} , we will have $\mathbf{Q}_i P = \frac{1}{\lambda} q_i$. If the number of corresponding pairs is n , we can arrange a combined matrix equation like $\mathbf{Q}P = \frac{1}{\lambda} \mathbf{q}$, where \mathbf{Q} and \mathbf{q} can be acquired by stacking \mathbf{Q}_i and q_i for $i \in [0, n - 1]$ on top of one another. If $n \geq 12$, we can compute an up-to-scale ambiguous solution for the projection matrix elements as follows:

$$P = \frac{1}{\lambda} (\mathbf{Q}^T \mathbf{Q})^{(-1)} (\mathbf{Q}^T \mathbf{q}), \quad (\text{A.2})$$

where superscript T depicts the transpose operation. Finally, the projection matrix \mathbf{P} can be constructed by re-arranging the elements of vector P .

Appendix B. Image rectification

A homography matrix models planar motion of points projected onto a pair of images. If we have two

images with n pairs of corresponding points, such as \bar{p}_i and \bar{q}_i for $i \in [0, n - 1]$, and these points are lying on a plane, then we will have $p_i \cong \mathbf{M}q_i$, where p_i and q_i are the homogeneous 2D coordinates of \bar{p}_i and \bar{q}_i , respectively, \mathbf{M} is the 3×3 homography matrix, and \cong depicts up-to-scale equality. Similar to Eq. (A.1), the planar motion equation can be written as follows:

$$\begin{bmatrix} q_{ix} & q_{iy} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{ix} & q_{iy} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & q_{ix} & q_{iy} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{11} \\ \mathbf{M}_{12} \\ \vdots \\ \mathbf{M}_{33} \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} p_{ix} \\ p_{iy} \\ 1 \end{bmatrix}, \quad (\text{B.1})$$

where λ is an ambiguous scale factor. Eq. (B.1) can be re-written as $Q_i M = \frac{1}{\lambda} p_i$. By stacking Eq. (B.1) for different corresponding points (i.e., $i \in [0, n - 1]$) on top of one another, we will have $\mathbf{Q}M = \frac{1}{\lambda} \mathbf{p}$, where $\mathbf{Q} = [Q_0 | Q_1 | \dots | Q_{n-1}]^T$, and $\mathbf{p} = [p_0 | p_1 | \dots | p_{n-1}]^T$. Finally, by having more than four corresponding points, one is able to compute the elements of the homography matrix up to a scaling factor, as follows:

$$M = \frac{1}{\lambda} (\mathbf{Q}^T \mathbf{Q})^{(-1)} (\mathbf{Q}^T \mathbf{p}). \quad (\text{B.2})$$

The homography matrix \mathbf{M} can be computed by re-arranging the elements of M into a 3×3 matrix. In order to account for flat panel detector sag, which can be modeled as planar motion, we use a plate (placed in the accessory tray close to the LINAC source) with four markers

that appear at the periphery of the portal image (see Fig. 4(a)). The locations of these points p_i for $i \in [0, 3]$ change as the gantry rotates. This is the result of the flat panel detector movement with respect to the LINAC source. We assume a pre-defined ideal location for these points q_i , $i \in [0, 3]$, that have been located at the center of the image sides (i.e., $[0.0, 511.5]$, $[511.5, 1023.0]$, $[1023.0, 511.5]$, and $[511.5, 0.0]$ for q_0 , q_1 , q_2 , and q_3 , respectively, assuming the portal image size is 1024×1024). The process of rectification includes: (1) locating the p_i for $i \in [0, 3]$ in the portal image, (2) calculating the homography matrix from Eq. (B.2) using the ideal position of the points (i.e. q_i for $i \in [0, 3]$), (3) applying the homography matrix \mathbf{M} to all points within the portal image. Sample images are shown in Fig. 4. The rectified images can be thought to have been acquired from a static position of the flat panel detector (i.e., a so-called *virtual detector*) with respect to the LINAC source. All further processing, including the geometric calibration and registration is done on the rectified images.

Appendix C. LINAC closed form calibration

We first have to assume that the flat panel detector and the LINAC source are rigidly attached or as it is explained in Appendix B, the portal images are corrected for any movement of the flat panel detector. The process of calibration then involves characterization of the projection matrices, which can be used to project any point from the iso-centric coordinate frame of the LINAC to the image point on the flat panel detector for various gantry angles. The projection matrix for the gantry angle θ can be derived from Eq. (3), and its elements are as follows:

$$\mathbf{P}^\theta = \begin{bmatrix} f \cos(\theta) + o_x \sin(\theta) & -f \sin(\theta) + o_y \cos(\theta) & 0 & o_x d \\ o_x \sin(\theta) & o_y \cos(\theta) & -f & o_y d \\ \sin(\theta) & \cos(\theta) & 0 & d \end{bmatrix}, \quad (\text{C.1})$$

where o_x and o_y are the x and y coordinates of the image center. In our case, where we use the rectified images with size 1024×1024 , we have $o_x = o_y = o = 511.5$. Furthermore, f is the focal length in pixel units, which can also be expressed in terms of the FOV and image size (i.e., $f = 511.5 / \tan(\text{FOV})$) in the case of rectified images. In order to have an estimate of the parameters of this matrix (i.e., f and d), we use a calibration phantom as shown in Fig. 3. We take several portal images of the calibration phantom (six at multiples of 60° angle starting from 0°), and after rectification and finding the radio-opaque markers in the portal images (depicted in Fig. 7), we establish 2D–3D correspondences and find the projection

matrices $\hat{\mathbf{P}}_i$ for $i \in [0, 5]$ (note that $\theta = 60i$) using the procedure outlined in Appendix A. It should be noted that $\hat{\mathbf{P}}_i$ are the projection matrices, which map a point from the calibration phantom coordinate frame onto the portal images and *not* from the iso-centric coordinate frame of the LINAC. Therefore, we first have to find the iso-center and establish the iso-centric coordinate frame as shown in Fig. 1.

It can easily be shown that the LINAC source location S_θ (at the gantry angle θ) in the calibration phantom coordinate frame can be computed using $-\mathbf{C}_i^{-1} \mathbf{C}_i$, where \mathbf{C}_i is a 3×3 matrix consisting of the first three columns of $\hat{\mathbf{P}}_i$, and \mathbf{C} is the last column of $\hat{\mathbf{P}}_i$. All the six source points (i.e., S_θ for various θ) should ideally lie on a plane. The LINAC iso-center in the phantom coordinate frame O^{ph} can approximately be computed using the average of S_θ for $\theta = 60i$, where integer $i \in [0, 5]$. More accurate estimation of the iso-center involves fitting a circle to the six points and computing the center of that circle. The normal of the plane, which is fitted to these points, V_z indicates the z direction of the iso-centric coordinate frame within that of the calibration phantom. The unit vector in y direction can be computed by normalizing the vector connecting S_0 to O^{ph} as follows, $V_y = (S_0 - O^{\text{ph}}) / (|S_0 - O^{\text{ph}}|)$. Finally, the unit vector in x direction is computed using the following equation: $V_x = V_y \times V_z$. We then have to arrange V_x , V_y , V_z and O^{ph} into a homogeneous matrix $\mathbf{H}_{\text{iso}}^{\text{ph}}$, which transforms the points from the LINAC iso-centric coordinate frame to that of the calibration phantom, as shown below:

$$\mathbf{H}_{\text{iso}}^{\text{ph}} = \begin{bmatrix} V_x & V_y & V_z & O^{\text{ph}} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (\text{C.2})$$

It can easily be shown that the estimates of \mathbf{P}^θ can be computed as follows:

$$\tilde{\mathbf{P}}^\theta = \hat{\mathbf{P}}^\theta \mathbf{H}_{\text{iso}}^{\text{ph}}. \quad (\text{C.3})$$

As it is shown in Eq. (C.1), the elements of $\tilde{\mathbf{P}}^\theta$ depend on the parameters f and d . Following equations can be used to recover d and f given the elements of the matrix $\tilde{\mathbf{P}}^\theta$:

$$\tilde{f} = \frac{1}{12} \sum_{i=0}^5 \frac{-\tilde{\mathbf{P}}^{60i}(2, 3) + \sqrt{\tilde{\mathbf{P}}^{60i}(1, 1)^2 + \tilde{\mathbf{P}}^{60i}(1, 2)^2 - o^2}}{\sqrt{\tilde{\mathbf{P}}^{60i}(3, 1)^2 + \tilde{\mathbf{P}}^{60i}(3, 2)^2}}, \quad (\text{C.4})$$

$$\tilde{d} = \frac{1}{18} \sum_{i=0}^5 \frac{\tilde{\mathbf{P}}^{60i}(3, 4) + (\tilde{\mathbf{P}}^{60i}(1, 4) + \tilde{\mathbf{P}}^{60i}(2, 4))/o}{\sqrt{\tilde{\mathbf{P}}^{60i}(3, 1)^2 + \tilde{\mathbf{P}}^{60i}(3, 2)^2}}, \quad (\text{C.5})$$

where the denominators are used only to recover scale factor ambiguity, which exists in estimation of the projection matrices. These estimates provide the initial values needed for non-linear optimization with the cost function shown in Eq. (5).

References

- Bansal, R., Staib, L., Chen, Z., Rangarajan, A., Knisely, J., Nath, R., Duncan, J., 1999. Entropy-based multiple portal to 3DCT registration for prostate radiotherapy using iteratively estimated segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI'99)*. LNCS, vol. 1679, pp. 567–578.
- Bentel, C.G., 1999. *Patient Positioning and Immobilization in Radiation Oncology*. McGraw-Hill, New York.
- Bijhold, J., 1993. Three-dimensional verification of patient placement during radiotherapy using portal images. *Med. Phys.* 20, 347–356.
- Bijhold, J., van Herk, M., Vijlbrief, R., Lebesque, J., 1991. Fast evaluation of patient setup during radiotherapy by aligning features in portal and simulator images. *Phys. Med. Biol.* 36 (12), 1665–1679.
- Brent, R., 1973. *Algorithms for Minimization without Derivatives*. Prentice-Hall, Englewood-Cliffs, NJ.
- Cabral, B., Cam, N., Foran, J., 1994. Accelerated volume rendering and topographic reconstruction using texture mapping hardware. In: *ACM Symp. on Vol. Vis.*, pp. 91–98.
- Cho, Y., Moseley, D.J., Siewerdsen, J.H., Jaffray, D.A., 2004. Geometric calibration of a cone-beam computed tomography system and medical linear accelerator. In: *ICCR 2004*.
- Dong, L., Boyer, A.L., 1995. An image correction procedure for digitally reconstructed radiographs and electronic portal images. *Int. J. Radiat. Oncol. Biol. Phys.* 33 (5), 1053–1060.
- Gilhuijs, K., Drukker, K., Touw, A., van de Ven, P.H., van Herk, M., 1996a. Automatic three-dimensional inspection of patient setup in radiation therapy using port images, simulator images, and computed tomography data. *Med. Phys.* 23 (3), 389–399.
- Gilhuijs, K., Drukker, K., Touw, A., van de Ven, P.H., van Herk, M., 1996b. Interactive three dimensional inspection of patient setup in radiation therapy using digital portal images and computed tomography data. *Int. J. Radiat. Oncol. Biol. Phys.* 34 (4), 873–885.
- Gorse, D., Shepherd, A., Taylor, J.G., 1994. A classical algorithm for avoiding local minima. In: *Proc. of WCNN'94*, San Diego, CA, pp. III 364–369.
- Griewank, A., 1981. Generalized descent for global optimization. *J. Optim. Theory Appl.* 34 (1), 11–39.
- Hanley, J., Mageras, G.S., Sun, J., Kutcher, G.J., 1995. The effects of out-of-plane rotations on two dimensional portal image registration in conformal radiotherapy of the prostate. *Int. J. Radiat. Oncol. Biol. Phys.* 5, 1331–1343.
- Herman, M.G., Balter, J.M., Jaffray, D.A., McGee, K.P., Munro, P., Shalev, S., Van Herk, M., Wong, J.W., 2001. Clinical use of electronic portal imaging: report of AAPM radiation therapy committee task group 58. *Med. Phys.* 28 (5), 712–737.
- Hipwell, J.H., Penney, G.P., McLaughlin, R.A., Rhode, K., Summers, P., Cox, T.C., Byrne, J., Noble, J., Hawkes, D., 2003. Intensity-based 2D–3D registration of cerebral angiograms. *IEEE Trans. Med. Imaging* 22 (11), 1417–1426.
- Hristov, D., Fallone, B., 1996. A gray-level image alignment algorithm for registration of portal images and digitally reconstructed radiographs. *Med. Phys.* 23 (1), 227–233.
- Hubbell, J.H., 1982. Photon mass attenuation and energy-absorption coefficients from 1 keV to 20 MeV. *Int. J. Appl. Radiat. Isot.* 33, 1269–1290.
- Hurkmans, C., Remeijer, P., Lebesque, J., Mijnheer, B., 2001. Set-up verification using portal imaging; review of current clinical practice. *Radiother. Oncol.* 56, 105–120.
- Kim, J., Fessler, A., Lam, K.L., Balter, J.M., Haken, R.K.T., 2001. A feasibility study of mutual information based set-up error estimator for radiotherapy. *Med. Phys.* 28 (12), 2391–2395.
- Lam, K., Ten Haken, R., McShan, D., Thornton, A., 1993. Automated determination of patient setup errors in radiation therapy using spherical radio-opaque markers. *Med. Phys.* 20 (4), 1145–1152.
- LaRose, D.A., 2001. *Iterative X-ray/CT registration using accelerated volume rendering*. Ph.D. Thesis, Carnegie Mellon University.
- LaRose, D.A., Bayouth, J., Kanade, T., 2000. Transgraph: interactive intensity-based 2D/3D registration of X-ray and CT data. In: *SPIE Medical Imaging 2000: Image Processing*, pp. 385–396.
- Lavallee, S., Szeliski, R., 1995. Recovering the position and orientation of free-form objects from image contours using 3D distance maps. *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (4), 378–390.
- Litzenberg, D., Lam, K.L., Balter, J.M., Hornick, D.C., Ten Haken, R.K., 1999. A mathematical model for correcting patient setup errors using a tilt and roll device. *Med. Phys.* 26, 2586–2588.
- Lujan, A.E., Balter, J.M., Ten Haken, R.K., 1998. Determination of rotations in three dimensions using two-dimensional portal image registration. *Med. Phys.* 25 (5), 703–708.
- Maintz, J.B.A., Viergever, M.A., 1998. A survey of medical image registration. *Med. Image Anal.* 2 (1), 1–37.
- McDonough, J., Rusu, I., Wang, S., Chapuy, M., Dimocovski, M., Dimocovski, Z., Bloch, P., 2002. Dosimetric evaluation of an amorphous silicon portal imaging device. In: *7th International Workshop on Electronic Portal Imaging (EPI2K2)*, pp. 130–132.
- Michalaski, J., Wong, J., Bosch, W., 1993. A evaluation of two methods of anatomical alignment of radiotherapy portal images. *Int. J. Radiat. Oncol. Biol. Phys.* 27 (5), 1199–1206.
- Milickovic, N., Baltas, D., Giannouli, S., Lahanas, M., Zamboglou, N., 2000. CT imaging based digitally reconstructed radiographs and their application in brachytherapy. *Phys. Med. Biol.* 45 (10), 2787–2800.
- Nutting, C., Khoo, V., 2000. A randomized study of the use of a customized immobilization system in the treatment of prostate cancer with conformal radiotherapy. *Radiother. Oncol.* 54 (1), 1–9.
- Penney, G.P., Weese, J., Little, J.A., Desmedt, P., Hill, D.L.G., Hawkes, D.J., 1998. A comparison of similarity measure for use in 2D–3D medical image registration. *IEEE Trans. Med. Imaging* 17 (4), 586–595.
- Penney, G.P., Weese, J., Little, J.A., Desmedt, P., Hill, D.L.G., Hawkes, D.J., 2001. Validation of a two- to three-dimensional registration algorithm for aligning preoperative CT images and intra-operative fluoroscopy images. *Med. Phys.* 28, 1024–1032.
- Powell, M., 1964. An efficient method of finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* 7, 155–162.
- Remeijer, P., Geerlof, E., Ploeger, L., Gilhuijs, K., van Herk, M., Lebesque, J.V., 2000. 3D portal image analysis in clinical practice: an evaluation of 2D and 3D analysis techniques as applied to 30 prostate cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.* 45 (5), 1281–1290.
- Roche, A., Malandain, G., Pennec, X., Ayache, N., 1998. The correlation ration as a new similarity measure for multimodal image registration. In: *Lecture Notes in Computer Science MICCAI'98*, pp. 1115–1124.
- Rohlfing, T., Russakoff, D.B., Murphy, M.J., Maurer, C.R., 2002. An intensity-based registration algorithm for probabilistic images and its application for 2D to 3D image registration. In: *SPIE Medical Imaging 2003: Image Processing*, pp. 581–591.
- Sarrut, D., Clippe, S., 2000. Patient positioning in radiotherapy by registration of 2D portal to 3D CT images by a content-based research with similarity measures. In: *Computer Assisted Radiology and Surgery*, pp. 707–712.
- Tomazevic, D., Likar, B., Slivnik, T., Pernus, F., 2003. 3D/2D registration of CT and MR to X-ray images. *IEEE Trans. Med. Imaging* 22 (11), 1407–1416.
- Törn, A., Zilinskas, A., 1989. *Global Optimization*. Springer, Berlin.
- Weese, J., Penny, G.P., Buzug, T.M., Hill, D.L.G., Hawkes, D.J., 1997. Voxel-based 2D/3D registration of fluoroscopy images and

- CT scans for image-guided surgery. *IEEE Trans. Inf. Technol. Biomed.* 1, 284–293.
- Wells, W., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R., 1996. Multi-modal volume registration by maximization of mutual information. *Med. Image. Anal.* 1, 35–51.
- White, J., Chen, T., McCracken, J., et al., 1982. The influence of radiation therapy quality control on survival, response and sites of relapse in oat cell carcinoma of the lung. *Cancer* 50, 1084–1090.
- Young, G., Chellappa, R., 1992. Statistical analysis of inherent ambiguities in recovering 3-D motion from a noisy flow field. *IEEE Trans. Pattern Anal. Mach. Intell.* 14, 995–1013.
- Zöllei, L., Grimson, W.E.L., Norbash, A., Wells III, W.M., 2001. 2D–3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. In: *IEEE Computer Vision and Pattern Recognition*, pp. 696–703.