A Feasibility Study on Mutual Information based Set-up Error Estimator for Radiotherapy

Jeongtae Kim*, Jeffrey A. Fessler*, Kwok L. Lam†, James M. Balter†, and Randall K. Ten Haken†

October 8, 2001

Abstract

We have investigated a fully automatic set-up error estimation method that aligns DRRs (Digitally Reconstructed Radiographs) from a 3D planning CT image onto 2D radiographs that are acquired in a treatment room. We have chosen a MI (Mutual Information)-based image registration method, hoping for robustness to intensity differences between the DRRs and the radiographs. The MI-based estimator is fully automatic since it is based on the image intensity values without segmentation. Using 10 repeated scans of an anthropomorphic chest phantom in one position and two single scans in two different positions, we evaluated the performance of the proposed method and a correlation-based method against the set-up error determined by fiducial markers-based method. The mean differences between the proposed method and the fiducial marker-based method were smaller than 1mm for translational parameters and 0.8 degree for rotational parameters. The standard deviations of estimates from the proposed method due to detector noise were smaller than 0.3 mm and 0.07 degree for the translational parameters and rotational parameters, respectively.

Key Words: Radiotherapy, Set up error, Image Registration, Parameter Estimation, Mutual Information.

*Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109
†Dept. of Radiation Oncology, University of Michigan, Ann Arbor, MI 48109

1 Introduction

The goal of external beam radiotherapy is to irradiate a tumor to a high x-ray dose while sparing normal tissues as much as possible. To achieve this goal, the x-ray beam arrangement and the dose distribution are carefully planned based on the target tumor position within an individualized patient model. Each 3D patient model is constructed using a 3D planning x-ray CT (Computed Tomography) scan or 2D x-ray scan that is acquired several days before treatment.

Accurate radiotherapy requires the patient to be set up for treatment in a coordinate system that is consistent with the treatment plan. In practice, however, patient set-up errors occur despite the use of laser alignment. Such errors continue to be a concern in radiotherapy not only due to the unnecessary irradiation of normal tissues but also due to the sub-optimal irradiation of the target tumor.

Many studies have been conducted to quantify the statistics of the patient set-up error, to consider the effect of the set-up error in planning, to reduce the set-up error and/or to compensate for the set-up error by adjusting the x-ray beam position or the patient position [1–10]. Such considerations require an accurate and automatic set-up error estimation method. If the patient set-up error can be correctly estimated after a treatment, one can retrospectively calculate the actual x-ray dose accumulation from the treatment and review the patient set-up procedure. Moreover, if the position estimate can be completed quickly, then one can compensate for the set-up error by adjusting either the radiotherapy table or the x-ray beam position prior to treatment delivery.

Technically, the patient setup error estimation problem is the estimation of patient pose difference in the treatment room from the planning CT. Similar application arises in the area of image-guided surgery that requires the mapping of positions in pre-operative data to a coordinate system defined in the operating room [10] [11].

The patient setup is usually estimated by an image registration method that registers 2D radiographs after patient set-up to 2D simulated images or 3D planning CT image at the desired position. In most cases, the 2D simulator images or the 3D planning CT image is geometrically transformed to achieve the registration, and the patient set-up is estimated as the geometrical transformation that accomplishes the registration.

Although approaches that use 2D simulator images have the advantage of fast computation time, those are not very accurate since the 3D transformation is estimated by registering 2D images in a 2D plane [5]. It has been

reported that these methods have problems in estimating rotations in planes non-parallel to the 2D radiographs planes [12]. Approaches that register 3D planning CT image onto the 2D radiographs overcome this limitation at the expense of greater computation.

Several groups have investigated 3D/2D registration methods [5] [11–16]. Those methods can be classified into feature-based methods [12] [13] [15] and intensity-based methods using DRRs (Digitally Reconstructed Radiographs) [11] [14] [16]. Feature-based methods use anatomical or artificial landmarks segmented from the images to perform the registration. Because of the segmentation, the feature based methods have several drawbacks. For example, segmentation procedures often require skilled human interaction, thereby making the procedure difficult to automate. Automatic procedures for segmentation have been studied [13]. Even if one designs an automatic segmentation algorithm, segmentation errors may occur and cause setup estimation errors. However, since feature-based methods have the advantage of fast computation, those are widely investigated for setup error estimation [12] [13] [15].

Unlike feature-based methods, intensity-based methods using DRRs do not require segmentation. The DRRs, which are computed 2D projection images of the planning CT volume, are computed at several angles and compared to the radiographs from the same angles. The registration is achieved by maximizing a similarity measure based on the intensity values of such DRRs and the radiographs. Since intensity-based methods do not require segmentation, those can be easily automated and avoid the segmentation error. However, these methods require much more computation than feature-based methods, thus posing a significant challenges for clinical application. In intensity-based methods, one must design an effective similarity measure for registration. The correlation between images and other application-specific similarity measures have been investigated for 3D/2D registration [11] [14] [16].

We investigate an approach for estimating the set-up error based on the 3D/2D, intensity-based registration method using DRRs. We have chosen a MI (Mutual Information) registration criterion since it is robust to the intensity differences between two images. In image registration for radiotherapy, the radiographs and the DRRs have different intensities since the radiographs and the CT are generated by x-ray sources with different spectra. Moreover, other effects such as different scalings between two imaging devices, beam hardening, scattering and

the radiotherapy table also cause intensity differences. The MI-based image registration method has been successful in 3D-3D multi-modality image registration [17] [18] [19]. By adopting the MI-based image registration technique, we have tried to design a fully automatic and accurate estimator.

Despite the computation time issues, we have chosen the intensity-based method using DRRs since we want to design an automatic and accurate estimation method. Moreover, we believe that the computation time will be eventually reduced by faster computer hardware, storing pre-computed DRRs with appropriate interpolation, etc.

To evaluate the performance of the MI-based estimator, we conducted an experiment with an anthropomorphic chest phantom. We placed 11 radiopaque markers on the phantom, enabling determination of the "ground truth" set-up error by registering the positions of the markers in the DRRs and the radiographs. We evaluated the performance of the MI-based estimator by comparisons with the marker-based method since we expected the latter to be more accurate in this experiment.

The structure of this paper is as follows. In section 2, we formulate the technical problem of the set-up error estimation. We also report the materials that were used for the experiment. In section 3, we present experimental results of the proposed MI-based method in comparison with a well-known correlation-based method [16]. In section 4, we discuss the proposed method and suggest future research directions.

2 Materials and Methods

2.1 Experimental Methods

We attached eleven 1mm diameter lead markers to the exterior surface of an anthropomorphic phantom to help establish the "ground truth" set-up error. By placing markers on locations that would be imaged around the boundaries of the projection views, we could apply the MI-based method using just the center portions of the projection images excluding the markers.

A $512\times512\times398$ voxel planning CT image with $0.9375\times0.9375\times1$ mm spacing was acquired on a GE CT/i scanner with a 140 kV x-ray source. Tattoos were drawn on the phantom where three alignment laser planes crossed the phantom to facilitate consistent set-up in a treatment room.

Next, the phantom was moved to the treatment room and it was set up at the isocenter by manually aligning tattoos to three laser planes in the treatment room. Four radiographs were obtained from different angles by rotating the x-ray source and Varian Portal Vision amorphous silicon active matrix flat panel image detector in 30^o increments around the Z-axis as in Figure 1. For each of the 0^o and 90^o views, we acquired 10 repeated radiographs without realignment for evaluating the effect of noise on the estimator. The x-ray source voltage was 6 MV and the detector size was 512×384 pixels with $0.78\text{mm}\times0.78\text{mm}$ spacing. We used only two radiographs from 0^o, 90^o (i.e, AP and lateral images) for the MI-based and the correlation-based method. However, to enhance the accuracy of the "ground truth", we used all four radiographs for the fiducial marker-based method.

We acquired additional radiographs in the treatment room after manually moving the phantom to each of two other positions for testing the robustness of the proposed method to different set-up errors. The experiments for the three different phantom positions are called Experiment A, B and C, respectively in subsequent sections.

We applied three methods for estimating the set-up error: the MI-based method, the correlation-based method, and the marker-based method. For the MI-based method and the correlation-based method, the planning CT image was down-sampled by four along each axis to reduce computation time and memory usage. Since our planning CT had finer axial sampling than typical CTs in clinical use, the down sampling yields more representative CT quality in axial sampling. However, it also caused slightly coarser in-plane sampling.

The down-sampling was implemented by averaging the nearest $4\times4\times4$ voxel values. As a result, the planning CT that was used to compute DRRs had $3.750\times3.750\times4$ mm spatial spacing. However to preserve the accuracy of the "ground truth", we did not down-sample the CT for the marker based method.

For the MI-based method and the correlation-based method, we used only the central 400×300 sub-image of each of the DRRs and the radiographs to avoid the effect of the markers which are not usually used in clinical practice.

We have established the geometry of the EPID imaging systems by determining radiation field edges using simple thresholding method [20]. We assumed that the distance from source to detector known during calibration is correct.

For numerical search, we used the *Nelder-Mead simplex* method for all three methods [21]. We started the

search from nominal unrotated and untranslated position. Termination criteria for the simplex method were that both the sum of the six parameter changes was less than 0.1, 0.1, 10^{-12} and the maximum difference of the objective function within the simplex was less than 0.0001, 0.0001, 10^{-12} for the MI-based, the correlation-based, and the marker-based method, respectively. Units for the rotation parameters were degrees and the translational parameters were mm.

2.2 Technical Problem formulation

For the phantom experiments, we assume that the set-up errors are generated by rigid body motion and there is no change in the attenuation coefficients from the time of the planning CT scan. With those assumptions, we can model each voxel intensity value U_k of the planning CT with a mono-energetic source approximation as follows:

$$U_k = \alpha\mu(\vec{x}, \varepsilon_{\mathrm{kV}}), \qquad k = 1, 2, \ldots, M, \tag{2.1}$$

where, $\mu(\vec{x}, \varepsilon)$ is 3D attenuation coefficients distribution, $\vec{x}$ is 3D spatial index, ε is the photon energy, $\varepsilon_{\mathrm{kV}}$ is the effective energy of the x-ray source for the planning CT, α is the constant scaling of the x-ray CT, and M is the number of CT voxels.

The radiographs acquired in the treatment room correspond to the projection images of the rotated and translated attenuation coefficients based on MV spectrum source since the MV treatment beam is also used for the imaging. Making a mono-energetic source approximation and neglecting scattered radiation, we model the measured pixel intensity values of the radiographs as follows:

$$\tilde{Y_k} = I_o e^{-\int_{L_k} \mu(T_{\tilde{\theta}}(\vec{x}), \varepsilon_{\mathrm{MV}}) d\vec{l}} + N_k, \qquad k = 1, 2, \ldots, N \tag{2.2}$$

$$\tilde{\theta} = [\phi_x\, \phi_y\, \phi_z\, t_x\, t_y\, t_z] \tag{2.3}$$

$$T_{\tilde{\theta}}(\vec{x}) = R(\phi_x)R(\phi_y)R(\phi_z)\vec{x} + [t_x\, t_y\, t_z]', \tag{2.4}$$

where $\tilde{Y}_k$ is the measured intensity value of kth detector pixel, I_o is the MV x-ray source intensity, L_k is the x-ray path from source to kth detector pixel, $T_{\tilde{\theta}}$ is the translate-rotation transform with parameters $\tilde{\theta}$, $\varepsilon_{\mathrm{MV}}$ is the effective source energy for the radiographs, N_k is measurement noise, N is the number of detector pixels, $R(\cdot)$ is a 3×3 rotation matrix, ϕ_x, ϕ_y, ϕ_z are the unknown rotations around X, Y, Z axes, and t_x, t_y, t_z are the unknown translations along X, Y, Z axes from the planning CT scan to treatment.

To estimate $\tilde{\theta}$, we compute the DRRs of the planning CT transformed according to any given guess for the parameter θ using the following formula:

$$X_k(\theta) = \int_{L_k} \mu(T_\theta(\vec{x}), \varepsilon_{\mathrm{kV}}) d\vec{l}, \qquad k = 1, 2, \ldots, N, \tag{2.5}$$

where X_k is the value of kth pixel in DRR.

In practice, the DRRs are computed by approximate summations for the line integral using the values U_k. We have implemented the line integral by computing the ray crossing lengths within each voxel and summing over the lengths multiplied by voxel values [22]. The performance of the MI-based method and the correlation-based method are unaffected by global constant scaling differences.

The technical problem of set-up error estimation is to estimate the six elements of the registration parameters $\tilde{\theta}$ by using the intensity values of the radiographs and the planning CT image. By defining $Y_k = -\log(I_o/\tilde{Y}_k)$ and ignoring the measurement noise, the problem becomes estimating $\tilde{\theta}$ by using Y_k, which is the measured line integrals of the attenuation coefficient based on mega-voltage spectra, and X_k, which is the computed line integrals of the attenuation coefficients with kilo-voltage source spectra, as defined in (2.5) and the following:

$$Y_k = \int_{L_k} \mu(T_{\tilde{\theta}}(\vec{x}), \varepsilon_{\mathrm{MV}}) d\vec{l}, \qquad k = 1, 2, \ldots, N. \tag{2.6}$$

If there were a functional relationship between Y_k and X_k, we could use the MLE (Maximum Likelihood

Estimator) which has many desirable properties [23]. Moreover, if there were a linear relationship, we could use successfully the well-known correlation-based method [16]. However, since the DRRs and the radiographs do not have a exact functional relationship, we would like to use a similarity measure that is robust to the intensity differences between the two images.

2.3 MI-based Method

Like many other intensity-based image registration methods using statistical analysis, the MI-based method conceptually considers the values of pixels in an image as samples of some random variable. Likewise, we can define a joint pdf (probability density function) of two random variables based on the corresponding pixel values in two images.

The idea of the MI-based method is that the two random variables are less jointly random if two images are more registered. When registered, pixels that have the same intensity value in one image will correspond to a more clustered distribution of the intensity values in the other image. This observation need not to be limited to the case that the intensities of the corresponding pixels are clustered around single value. There can be two or more clusters around different values. If we estimate the joint pdf from corresponding pixels in two images, the conditional pdf becomes more clustered as the two images become better registered.

Based upon this observation, the MI-based method achieves registration by minimizing the estimated conditional entropy given the random variable which represents the transformed image or equivalently maximizing the estimated mutual information between two images [17] [18] [19].

We treat the pixel intensity values of the DRRs $X_k(\theta)$ in (2.5) from AP image as samples of a random variable X_{AP} and those from lateral image as samples of random variable X_{LA}. We also consider the pixel values of the radiographs Y_k from AP and lateral images as samples of a random variable Y_{AP} and Y_{LA}.

To estimate the mutual information, we first estimate the joint pdf of (X_{AP},Y_{AP}) and (X_{LA},Y_{LA}) by using the samples X_k and Y_k. We used a 128$\times$128 bin joint histogram to estimate the joint pdf. The estimated mutual information is then computed from the estimated joint pdf by the definitions of the entropy and the mutual information [18] [19] [24]. The set-up error is estimated by seeking θ which maximizes the sum of the two

estimated mutual information as follows,

$$\hat{\theta}_{\mathrm{MI}} = \arg\max_{\theta} \hat{I}_{\theta}(X_{AP}, Y_{AP}) + \hat{I}_{\theta}(X_{LA}, Y_{LA}) = \arg\min_{\theta} \hat{H}_{\theta}(Y_{AP}|X_{AP}) + \hat{H}_{\theta}(Y_{LA}|X_{LA}) \tag{2.7}$$

where $\hat{I}_{\theta}(X_{AP}, Y_{AP})$, $\hat{I}_{\theta}(X_{LA}, Y_{LA})$ are the estimated mutual information between random variable (X_{AP}, Y_{AP}), (X_{LA}, Y_{LA}), and $\hat{H}_{\theta}(Y_{AP}|X_{AP})$, $\hat{H}_{\theta}(Y_{LA}|X_{LA})$ are the estimated conditional entropy of Y_{AP} given X_{AP}, Y_{LA} given X_{LA}.

To solve this nonlinear optimization problem numerically, we applied the *Nelder-Mead simplex method*.

2.4 Correlation-based Method

For comparison purposes, we also applied a correlation-based method. Although the intensity scales of the MV and the kV X-ray images are not exactly linearly related in theory, we may expect this method to work well since if MV attenuation is large, so is kV attenuation.

Moreover, one may try to improve the performance of the correlation-based method by computing more MV-like DRR. The MV attenuation coefficients may be computed from kV attenuation coefficient by classifying voxels into different compounds and referencing typical MV attenuation coefficients of those compounds [16]. However, these efforts require additional procedure and periodic calibration since the calculations depend on the characteristics of the CT scanner. Moreover, the performance improvement may be limited by other factors such as beam hardening effects, scattering effects, presence of the unmodeled radiotherapy table, etc.

We compute the correlation between a DRR and a radiograph of an AP image as follows:

$$\hat{\rho}_{\theta}(X_{AP}, Y_{AP}) = \frac{\sum_{k=1}^{N}(X_k(\theta) - \bar{X}_k(\theta))(Y_k - \bar{Y}_k)}{[(\sum_{k=1}^{N}(X_k(\theta) - \bar{X}_k(\theta))^2)(\sum_{k=1}^{N}(Y_k - \bar{Y}_k)^2)]^{1/2}}, \tag{2.8}$$

where $X_k(\theta)$ is the k_{th} pixel value of AP DRR, $\bar{X}_k(\theta)$ is the mean of $X_k(\theta)$, $Y_k(\theta)$ is the value of k_{th} pixel in

AP radiograph, $\bar{Y}_k$ is the mean of Y_k, and N is the number of pixels.

Similarly, we can compute the correlation between a DRR and a radiograph from lateral images. The set-up error is estimated by maximizing the sum of the estimated correlation coefficients from AP and lateral images as follows,

$$\hat{\theta}_{\mathrm{co}} = \arg\max_{\theta} \hat{\rho}_{\theta}(X_{AP}, Y_{AP}) + \hat{\rho}_{\theta}(X_{LA}, Y_{LA}) \tag{2.9}$$

where $\hat{\rho}_{\theta}(X_{AP}, Y_{AP})$ is the estimated correlation between a DRR and a radiograph from the AP direction and $\hat{\rho}_{\theta}(X_{LA}, Y_{LA})$ is that from lateral images.
We also applied the *Nelder-Mead simplex method* to solve this optimization problem.

2.5 Marker-based Method

To establish the "ground truth", we estimated the set-up error using the positions of the radiopaque markers. The results from the marker-based method were assumed to be the true phantom position. As a result, the performance of the MI-based estimator and the correlation-based method were evaluated by referencing the position estimated by the marker based method.

The procedures for estimating the set-up error by the marker-based method were as follows. First, we identified the slices of the CT which contain the markers. Figure 2 shows example two slices among slices that contain markers. We manually identified $8\times8\times4$ voxel region around each marker and estimated each center position using the centroid method [25]. We also manually identified 7×7 pixels around each marker in the four different radiographs and identified each center position as p_i using the centroid method. Next, while transforming the coordinates of the center position of each marker in 3D space, we projected marker center positions onto 2D planes which are the same planes as the radiograph planes, and identified the 2D indices of the projected center positions as $d_i(\theta)$.

We estimated the set-up error by minimizing the mean square Euclidean distance between the p_i and $d_i(\theta)$ as

follows:

$$\hat{\theta} = \arg\min_{\theta} \sum_{i} \|p_i - d_i(\theta)\|^2. \tag{2.10}$$

There were 22, 19, and 21 clearly identifiable markers in the four radiographs for Experiment A, B, and C respectively. Although the projection of 11 markers on the phantom to four different projection views resulted in 44 markers in the radiographic planes, some markers were not visible since those markers projected beyond the bounds of the detector.

3 Results

3.1 Radiographs

Figure 3 (a) shows the estimated MI from radiograph/DRR 1 and Figure 3 (b) shows the average of two estimated MI from radiograph/DRR 1,4 as the planning CT is translated along the X, Y and Z axes from the registered position respectively while other five parameters were kept at the registered position. The registered position was defined as the mean of the marker-based estimated positions from 10 radiographs.

For the radiograph/DRR 1 case, the estimated MI changes only slightly with respect to the translation along the Y-axis, and the point of the maximal estimated MI is far from the registered point. This large error could be due to the fact that the movement along the Y-axis does not cause much change in DRR 1. Thus 3D/2D registration based on a single-view similarity measure would be sensitive to noise. The insensitivity of the single-view estimated MI with respect to changes in t_y is clear from Figure 1, because the Y-axis translation causes only a small change of magnification in the projection image.

Similar phenomena occur if only radiograph/DRR 4 is used. Thus, using only one projection view may cause a significant error in estimating the translation that is orthogonal to the detector plane.

This problem is alleviated by using two radiograph/DRR images. Figure 3 (b) shows the estimated MI using radiograph/DRR 1 and 4. The maximal mutual information position is close to the true position for all

six parameters. Based on this observation, we used radiograph/DRR 1,4 for evaluating the proposed MI-based estimator and the correlation-based estimator.

3.2 Estimated joint pdf

Figure 4 shows an example of non-registered and registered radiograph/DRR 1 and 4. The non-registered DRR was generated by translating the planning CT 12.5mm along the Z-direction from the registered position. For non-registered images, the bottom rows of DRR 1 contain bright pixels, unlike the corresponding pixels in radiograph 1. Also the top left parts of radiograph 4 are imaged dark, since those areas are air, while some of the corresponding pixels in DRR 4 have bright pixels. These phenomena diffuse the joint histogram and decrease the estimated mutual information.

Figure 5 shows the joint histograms that are estimated from the registered and the non-registered DRR/Radiograph 1 and 4. The joint histogram from the registered images shows interesting characteristics. It was expected that the joint histogram from the registered images would show more clustered shape along increasing functional curve, since the MV attenuation coefficients tend to be high if the kV attenuation coefficients are high. However, a range of DRR intensity values corresponded to high radiograph intensity values in Figure 5 (b), the estimated joint histogram from lateral images. This can be explained as the effect of the radiotherapy table. The radiotherapy table is slightly visible in the rightmost parts of radiograph 4 in Figure 1 (e). Although most of the radiotherapy table parts were not used for the MI-based and the correlation-based method, it still made the upper right parts of the radiograph 4 brighter than the DRR 4 since the corresponding pixels in the DRR 4 only represents air as we can see in Figure 4. For increased intensity values in the upper right parts due to the radiotherapy table in Figure 4 (b), even if the DRR and the radiographs are registered, corresponding pixels of DRR in Figure 4 (f) are darker than radiograph 4.

3.3 Position Estimation Results

In Experiment A, we used laser alignment to set up the phantom without error, while in Experiment B and C, we generated set-up errors deliberately to test the robustness of the proposed method to different set-up errors.

In Experiment A, radiographs 1 and 4 were acquired 10 times. Table 1 summarizes the empirical means and the standard deviations of each method in Experiment A. The estimated set-up errors by the proposed MI-based method were close to the set-up errors determined by the marker-based method. The differences between the means of the MI-based method and the marker-based method did not exceed 1.0 mm for translation parameters and 0.8° for rotational parameters. Considering that the voxel spacing of the sub-sampled planning CT that was used for computing DRR was $3.75 \times 3.75 \times 4$ mm, estimation errors for every parameter were sub-voxel.

The sample STD (Standard Deviation) of the proposed estimator was very small. This was because the EPID has very low noise and we used an automatic method without human interaction.

The results of the correlation-based method were also fairly good. This indicates that the the relationship between DRR and radiograph is approximately linear. Compared with the proposed method, the result of the correlation-based method shows relatively larger variance to noise. However, since the noise level was quite low, the noise-induced variability was insignificant compared to the mean errors.

Table 2 shows the results from Experiment B. In this experiment, we tried to generate patient set-up error of t_x=12mm, t_y=-8 mm t_z=-7mm and $\phi_x = \phi_y = \phi_z = 0^{\circ}$. Compared with the results from Experiment A, the correlation-based method performed worse while the MI-based method still worked well.

Table 3 shows the results from Experiment C. For this experiment, we tried to generate rotational set-up error. The planned set-up error was ϕ_x=0°, ϕ_y=2.2°, ϕ_z=1.2°, $t_x = t_y = t_z$ =0mm.

In summary, for three different phantom positions, we estimated the set-up error 12 times including 10 repeated estimates using 10 different acquisitions. The means of the estimated set-up error differences between the marker-based method and the MI-based estimator did not exceed 1mm for translation parameters and did not exceed 0.8° for rotation parameters.

4 Discussion

By using two orthogonal radiograph/DRR pairs, we achieved average accuracies of better than 1mm for translational movement parameter and better than 0.8° for rotational movement parameters in estimating the set-up error.

We established the "ground truth" positions using the marker-based method. The accuracy of the marker-based method was expected to be the best among the three methods that were tested because the higher resolution CT was used to identify the marker position in the CT and four radiographs/DRRs were used in the marker-based method while only two radiographs/DRRs were used for the other methods. Although this method also contains error due to segmentation errors and noise, we used the marker-based method as "ground truth" since it is expected to be more accurate than the other methods.

There was approximately 3mm set-up variation from the positioning of the phantom at its proper reference location. Factors that contribute to this value include the limits of human operators in positioning relative to laser marks, as well as differences in laser calibration between the CT scanner and the treatment room. A retrospective review indicated a 1mm offset of the CT lasers from the center of the image matrix. Such errors in transferring a phantom or patient from one system to another have been previously reported, and are unlikely to be dramatically reduced in routine radiotherapy quality assurance.

For rigid body set-up error estimation using the chest phantom, two radiographs/DRRs were adequate for sub-voxel accuracy. We suspect that the performance of the estimator would be improved only modestly if more DRRs and radiographs from different angles were added. Practically, using fewer radiographs/DRRs is strongly preferable because of acquisition time and computation time.

The correlation-based method also achieved sub-voxel accuracy. Even though many factors could cause the intensity relation between DRR and radiograph to be nonlinear, we found the relationship is approximately linear so the correlation-based method worked fairly well.

We have found that the standard deviation of the correlation-based method was larger than that of the MI-based method in Experiment A. This implies that the correlation-based method may be more sensitive to the noise, although the effects of noise were not significant in the experiment since the noise level was very low. We are not sure if this phenomena is just for our experiments or general. One may investigate this problem by theoretically approximating the variances of the estimator [26] as well as experimentally.

We also tried to compute the MV DRRs for the correlation-based method. However, results were no better but more complicated. We think that was because we did not have CT data from test phantoms for voxel classification

[16] nor information about scattering, etc. Since we were unable to implement the algorithm thoroughly, we have excluded the results.

One may try to compute better MV DRRs by following thorough procedures, however we suspect that the result will not be improved dramatically since unmodeled effects such as the presence of the radiotherapy table, difficulty of correctly compensating for scatter and beam hardening may limit the performance.

Compared with methods that require *a priori* model, we think that the MI-based method may perform better in the presence of unexpected objects. As presented in the experimental results section, even though there exist non-modeled effects of the radiotherapy table, the results of the MI-based estimator show good performance. This robustness of the MI-based estimator to non-modeled effects partly supports the advantages of the method over *a priori* model-based methods such as MLE [23]. Therefore, the MI-based method may be useful in applications in which non-modeled objects may present, such as image-guided surgery.

For estimating patient set-up error, around 225 evaluations of the estimated mutual information were required. Each evaluation of the estimated mutual information requires computation of two DRRs. The joint histogram and the mutual information also must to be computed for each mutual information evaluation. The most time consuming part was generating DRRs; it took around 16 sec to compute one 400×300 DRR from the $128\times128\times85$ planning CT on Pentium II 600 MHz machine. As a result, estimating one patient set-up error using two orthogonal radiographs/DRRs required about 2 hours.

This long computation time is not due to the MI criteria but because of DRR computation. Compared with other intensity-based method using DRR, our computation time was longer since we implemented a more accurate line integral instead of a trilinear interpolation approximation [14]. One might use trilinear interpolation approximations for computing DRR with proposed MI criteria for faster estimation.

Due to long computation time, it is difficult to estimate the patient set-up error in real time with the proposed method. As a result, adjusting the position of the x-ray source or the position of the radiotherapy table to compensate for the patient set-up error is presently impractical. Instead, as a first application, we expect our method to be applied to review the patient set-up procedure. However, in near future, we expect that the proposed method can be used for estimating the set-up error in real time using pre-computed DRRs stored in large memory

with appropriate interpolation and perhaps parallel processing. Furthermore, methods such as multi-resolution optimization technique may further reduce computing time [27].

Our investigation can be discussed in comparison of other investigations. Lemieux *et al.* applied the correlation-based similarity measure to skull phantom experiments [14]. They used kV X-ray source for both CT and radiographs so that the correlation measure may work better.

Dong *et al.* also investigated the correlation-based method by testing the method using a head phantom [16]. Although they used MV radiographs and kV CT, they generated MV DRRs so that the DRRs and the radiographs have linear intensity relationship. MV DRRs were generated by classifying voxels into several categories such as bone, muscles, etc. based on CT numbers and seeking typical MV attenuation coefficients of those.

There also have been many studies of set-up estimation [5] [11–16]. Most investigations reported set-up estimation methods of less than a few mm estimation error for translation parameters and less than a few degree for rotation parameters. It is hard to compare the performance of each method directly since different types of CT images and radiographs were used. We believe that the performance of an estimator can depend greatly on the image characteristics. For example, the presence of high contrast objects can improve performance. Moreover, different image resolution and different optimization stopping criteria may also affect the performance of the estimators.

Penney *et al.* investigated the performance of several similarity measures by applying those similarity measures to the image registration of a 3D CT to a fluoroscopy image [28]. However, since only one fluoroscopy image was used for estimating six parameters, the results may not be directly applicable to our study that used two radiographs for registration.

The MI-based method was investigated for the set-up error estimation problem by Hadley *et al.* [29]. In their investigation, the MI-based method worked well for simulated images but did not show good performance for clinical images. Since the investigation was for 2D/2D image registration and there was no comparison to other similarity measures for clinical images, it does not necessarily imply that the MI-based method perform worse in 3D/2D registration problems than other similarity measures.

Our future work on set-up error estimation includes the application of the MI-based estimator to clinical

data. Although we have achieved good registration results in a phantom experiment, we believe that much more verification with clinical data is crucial for potential application of the MI-based method to the clinical practice.

We plan to investigate non-rigid body motion set-up error. Since the human body is not rigid and radiotherapy may change the size and shape of the tumor, an accurate non-rigid body motion set-up error estimation is required for practical application.

We also plan to investigate other intensity-based similarity measures. For example, Rényi entropy may be a more flexible and computationally efficient similarity measure to accomplish image registration [30] [31] since it is more generalized definition of entropy. Investigating better similarity measures such as Rényi entropy may improve both speed and performance.

5 Conclusion

We have investigated a set-up error estimation method using 3D/2D, intensity-based method. To achieve 3D to 2D image registration, the radiographs and the DRRs of the planning CT were registered by maximizing the MI between DRRs and radiographs. In the experiment with an anthropomorphic chest phantom, we achieved accuracies better than 1mm for estimating the translational parameters and 0.8 degree for estimating the rotational parameters using two orthogonal pairs of the MV radiographs and DRRs. The true set-up error was established by the fiducial marker-based method. Based upon theoretical background and the experimental results, we believe that MI has significant potential as an effective similarity measure for 3D/2D intensity-based registration.

Acknowledgments

This work was supported by NIH Grant CA60711, P01-CA59827 and R01-CA81161. Dr. Balter is supported as a Kimmer scholar.

References

[1] G J Kutcher, G S Mageras, and S A Leibel, "Control, correction and modeling of setup errors and organ motion," *Seminars in Radiation Oncology*, vol. 5, no. 2, pp. 134–145, 1995.

[2] D A Jaffray, D Yan, and J W Wong, "Managing geometric uncertainty in conformal intensity-modulated radiation therapy," *Seminars in Radiation Oncology*, vol. 9, no. 1, pp. 4–19, 1999.

[3] S Shalev, "Treatment verification using digital imaging," in *Radiation therapy physics*, A R Smith, Ed., pp. 155–173. Springer-Verlag, New York, 1995.

[4] M van Herk, A Bel, KGA Gilhuijs, JV Lebesque, A van Dalen, P van der Ven, and R Vijlbrief, "Electronic portal imaging," *Bull Cancer*, vol. Suppl 5, no. 82, pp. 601s–605s, 1995.

[5] K G A Gilhuijs, P J H van de Ven, and M van Herk, "Automatic three-dimensional inspection of patient setup in radiation therapy using port images, simulator images, and computed tomography data," *Med. Phys.*, vol. 23, no. 3, pp. 389–399, 1996.

[6] J E Schewe, J M Balter, K L Lam, and R K Ten Haken, "Measurement of patient setup errors using port films and a computer-aided graphical alignment tool," *Med. Dosim.*, vol. 21, no. 2, pp. 97–104, 1996.

[7] L J Verhey, "Patient immobilization for 3-D RTP and virtual simulation," in *A practical guide to 3-D planning and conformal radiation therapy*, J A Purdy and G Starkschall, Eds., pp. 151–174. Advanced Medical Publishing, Inc., WI, 1999.

[8] R K Ten Haken and B A Fraass, "Components of a 3-D CRT quality assurance program," in *A practical guide to 3-D planning and conformal radiation therapy*, J A Purdy and G Starkschall, Eds., pp. 309–322. Advanced Medical Publishing, Inc., WI, 1999.

[9] J M Balter, R K Ten Haken, and K L Lam, "Treatment setup verification," in *Teletherapy:present and future*, T R Mackie and J R Palta, Eds., pp. 471–493. Advanced Medical Publishing, Inc., WI, 1996.

[10] M J Murphy, "An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotactic radiosurgery," *Med. Phys.*, vol. 24, no. 6, pp. 857, June 1997.

[11] J Weese, G P Penney, P Desmedt, T M Buzug, D L G Hill, and D J Hawkes, "Voxel-based 2-d/3-d registration of fluoroscopy images and CT scans for image-guided surgery," *IEEE Trans. on Info. Tech. in Biomedicine*, vol. 1, no. 4, pp. 284–293, 1997.

[12] K G A Gilhuijs, K Drukker, A Touw, P J H van de Ven, , and M van Herk, "Interactive three dimensional inspection of patient setup in radiation therapy using digital portal images and computed tomography data," *Int. J. Radiation Oncology Biol. Phys.*, vol. 34, no. 4, pp. 873–885, 1996.

[13] K G A Gilhuijs and M van Herk, "Automatic on-line inspection of patient setup in radiation therapy using digital portal images," *Med. Phys.*, vol. 20, no. 3, pp. 667–677, 1993.

[14] L Lemieux, R Jagoe, 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.*, vol. 21, no. 11, pp. 1749–1760, 1994.

[15] J Bijhold, "Three-dimensional verification of patient placement during radiotherapy using portal images," *Med. Phys.*, vol. 20, no. 2, pp. 347–356, 1993.

[16] L Dong and A L Boyer, "An image correlation procedure for digitally reconstructed radiographs and electronic portal images," *Int J. Radiation Oncology Biol. Phys.*, vol. 33, no. 5, pp. 1053–1060, 1995.

[17] W M Wells, P Viola, H Atsumi, S Nakajima, and R Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Im. Anal.*, vol. 1, no. 1, pp. 35–51, Mar. 1996.

[18] F Maes, A Collignon, D Vandermeulen, G Marchal, and P Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Tr. Med. Im.*, vol. 16, no. 2, pp. 187–98, Apr. 1997.

[19] C R Meyer, J L Boes, B Kim, P H Bland, et al., "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin plate spline warped geometric deformations," *Med. Im. Anal.*, vol. 1, no. 3, pp. 195–206, 1997.

[20] L Dong and A L Boyer, "A portal image alignment and patient setup verification procedure using moments and correlation techniques," *Phys. Med. Biol.*, vol. 41, pp. 697–723, 1996.

[21] W H Press, B P Flannery, S A Teukolsky, and W T Vetterling, *Numerical recipes in C*, Cambridge Univ. Press, 1988.

[22] R L Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Med. Phys.*, vol. 12, no. 2, pp. 252–255, Mar. 1985.

[23] M Cramer, *Mathematical methods of statistics*, Princeton University Press, 1946.

[24] T M Cover and J A Thomas, *Elements of information theory*, John Wiley and Sons, New York, 1991.

[25] J M Fitzpatrick, D L G Hill, and C R Maurer, "Image registration," in *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*, M. Sonka J. Michael Fitzpatrick, Ed., pp. 477–513. SPIE, Bellingham, 2000.

[26] J A Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Tr. Im. Proc.*, vol. 5, no. 3, pp. 493–506, Mar. 1996.

[27] F Maes, D Vandermeulen, and P Suetens, "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information," *Med. Im. Anal.*, vol. 3, no. 4, pp. 373–86, 1999.

[28] G P Penney, J Weese, J A Little, P Desmedt, D L G Hill, and D J Hawkes, "A comparison of similarity measures for use in 2-d-3-d medical image registration," *IEEE Trans. Medical Imaging*, vol. 17, no. 4, pp. 284–293, August 1998.

[29] S W Hadley, C A Pelizzari, L S Johnson, and G T Y Chen, "Automated registration of portal and simulation films by mutual information," in *Proceedings of XIIth ICCR*, May 1997, pp. 243–244.

[30] A Rényi, "On measures of entropy and information," in *Proc. of the Fourth Berkeley Symposium on Math. Stat. and Prob.*, 1961, vol. 1, pp. 547–61.

[31] A O Hero and O Michel, "Asymptotic theory of greedy approximations to minimal k-point random graphs," *IEEE Tr. Info. Theory*, vol. 45, pp. 1921–1939, Sept. 1999.

Table 1. Estimated set-up error for Experiment A

	ϕ_x	ϕ_y	ϕ_z	t_x	t_y	t_z
Marker (mean)	0.601	-0.114	-1.892	-1.968	-3.568	-2.995
MI error (mean)	-0.064	0.035	-0.729	-0.635	0.692	-0.151
Corr. error (mean)	0.336	0.143	-1.308	-0.939	-0.597	0.102
Marker STD	0.004	0.002	0.002	0.012	0.009	0.004
MI STD	0.045	0.008	0.071	0.088	0.060	0.071
Corr. STD	0.318	0.080	0.158	0.254	0.114	0.248

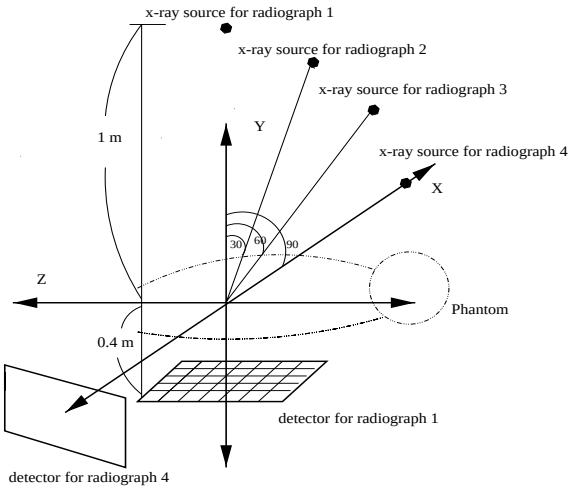
(*In each table, the units for rotational parameters are degrees and for translational parameters are mm.)

Table 2. Estimated set-up error for Experiment B

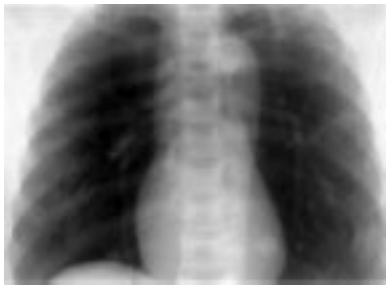
	ϕ_x	ϕ_y	ϕ_z	t_x	t_y	t_z
Marker	0.519	0.005	-1.879	9.703	-11.32	-9.58
MI error	-0.153	-0.044	-0.594	-0.597	1.030	0.668
Corr. error	0.286	-0.001	-2.534	-0.924	-3.092	1.451

Table 3. Estimated set-up error for Experiment C

	ϕ_x	ϕ_y	ϕ_z	t_x	t_y	t_z
Marker	0.755	1.864	-0.698	-1.515	-3.201	-2.947
MI error	0.067	-0.121	-0.667	-1.043	0.992	-0.453
Corr. error	0.72	0.069	-1.172	-1.050	-0.391	-0.010



(a) Radiographs and coordinate system



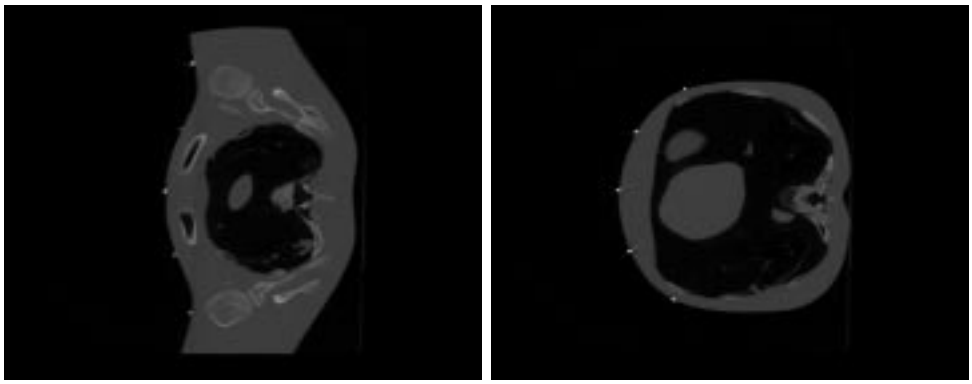
(b) Radiograph 1: angle 0

(c) Radiograph 2: angle 30

(d) Radiograph 3: angle 60

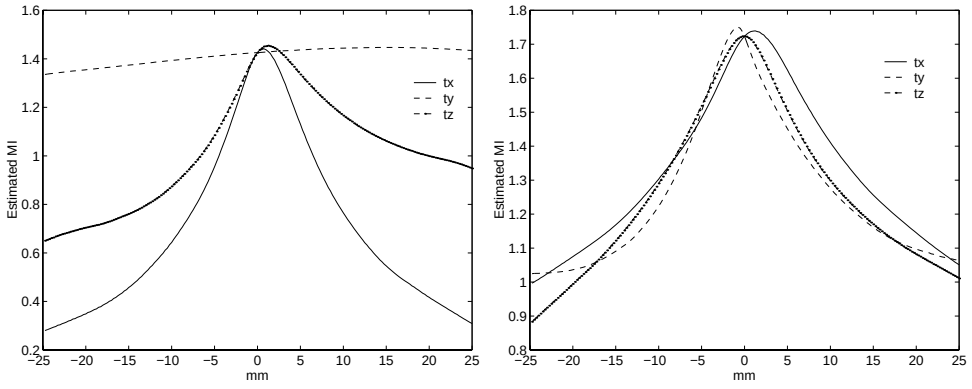
(e) Radiograph 4: angle 90

Figure 1: Radiographs



(a) 86th slice (b) 251th slice

Figure 2: Slices with markers

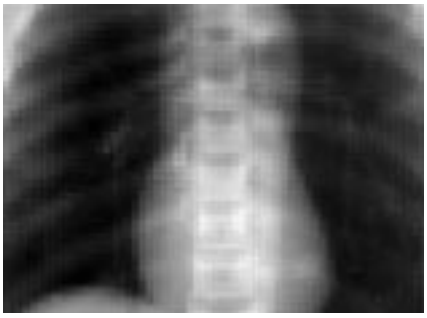


(a) Radiograph/DRR 1 (b) Radiograph/DRR 1, 4

Figure 3: Estimated mutual information with respect to the translation errors

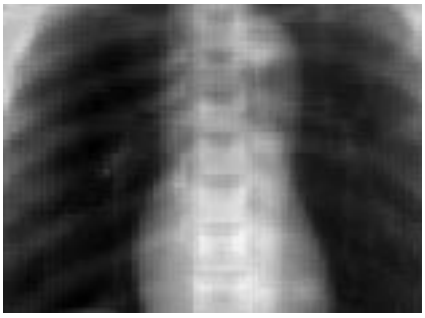
(a) Radiograph 1

(b) Radiograph 4

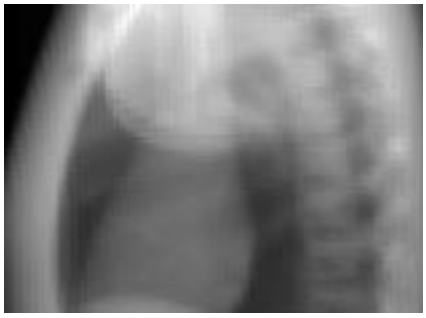


(c) Non-registered DRR 1

(d) Non-registered DRR 4

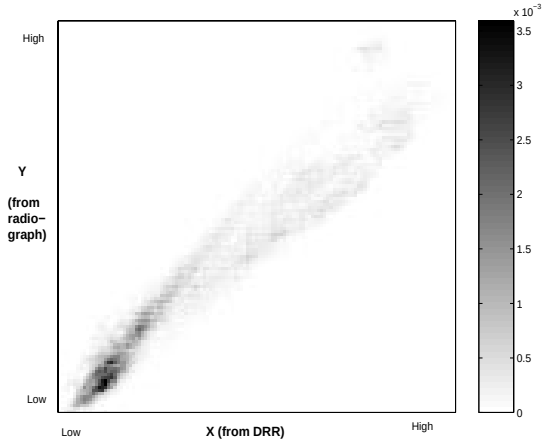


(e) Registered DRR 1

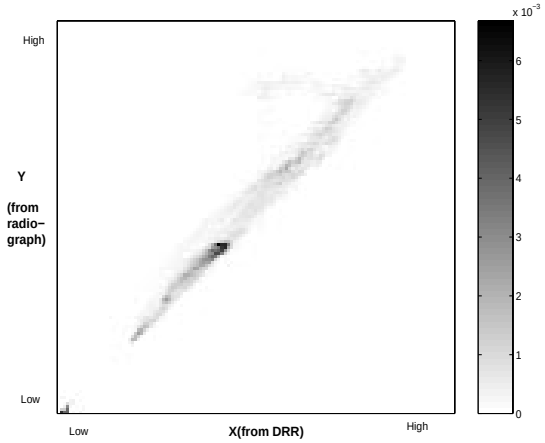


(f) Registered DRR 4

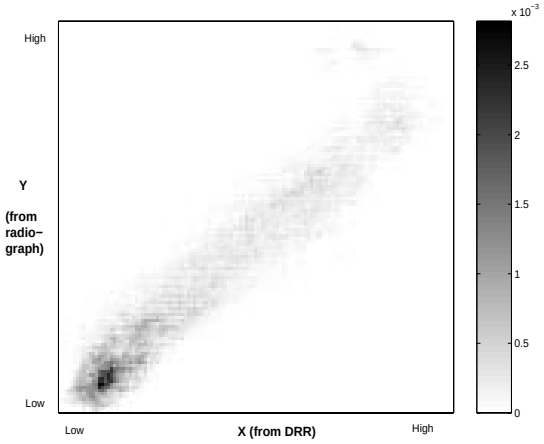
Figure 4: Non-registered and registered radiographs and DRRs



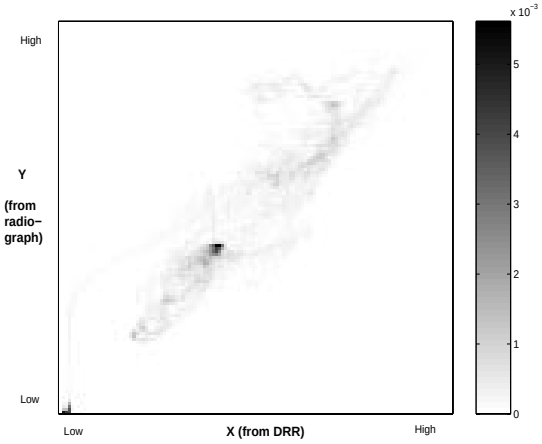
(a) Registered (DRR/radiograph 1)



(b) Registered (DRR/radiograph 4)



(c) Non-registered (DRR/radiograph 1)



(d) Non-registered (DRR/radiograph 4)

Figure 5: Joint histograms from registered and non-registered images