Towards Automatic, Model-Driven Determination of 3D Patient Setup Errors in Conformal Radiotherapy

Paul Yushkevich, Daniel Fritsch, Stephen Pizer, and Edward Chaney

UNC Medical Image Display & Analysis Group

The University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3175

Abstract

The accurate and quantitative determination of three-dimensional patient setup errors in conformal radiotherapy, including setup errors due to out-of-plane rotations, requires methods for registering pre-treatment, three-dimensional planning CT images with intra-treatment, two-dimensional portal images. We have developed a method for performing such a registration based on structural models that emphasize medial aspects of shape. Such models (1) provide an ability to pre-select those structures in a reference image which are known to be reliable fiducials for registration, (2) allow for the stable recognition of the same structures in treatment portal images, and (3) can be combined with images in such a way as to yield a measure of agreement between the model and the features in the image. We describe the means for creating a model in a reference image generated from the planning CT, for deforming the model to identify corresponding structures in a treatment portal image, and for optimizing an objective function based on combining the deformed model with a collection of digitally reconstructed radiographs generated from CT at tentative poses. The optimum of the objective function yields the three-dimensional pose of the patient relative to the planning pose, thereby indicating the three-dimensional setup error. Pilot results using simulated images with known patient positioning errors have shown that such an objective function obtains an optimum very near to truth.

Correspondence to:
Paul Yushkevich
Sitterson Hall, CB# 3175
The University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3175
email: pauly@cs.unc.edu

1 INTRODUCTION

1.1 Accuracy in radiotherapy treatment planning and delivery

The recent trend in radiotherapy toward conformal treatments — where smaller, tumor-conforming fields are used in conjunction with escalated radiation doses — has placed greater demands for accuracy in both treatment planning and treatment delivery. For treatment planning, it is necessary to have accurate knowledge of the three-dimensional (3D) extent of the tumor and surrounding, healthy tissues so that conformal fields may be prescribed which both adequately encompass the tumor and which also minimize the dose imparted to healthy tissues. The widespread adoption of 3D radiotherapy treatment planning (RTP) systems has had much to do with the increased interest in conformal therapy, as such 3D RTP systems permit more complex treatment fields to be prescribed as shown in Figure 1 [Sherouse, 1991]. The basic assumption of 3D RTP systems is that the user has accurately delineated (i.e., segmented) the 3D structures that will be impacted by the treatment such that dose-volume metrics computed on such structures from the conforming fields are reliable. For the remainder of this paper, we will assume that the result of 3D RTP is an optimal treatment plan.

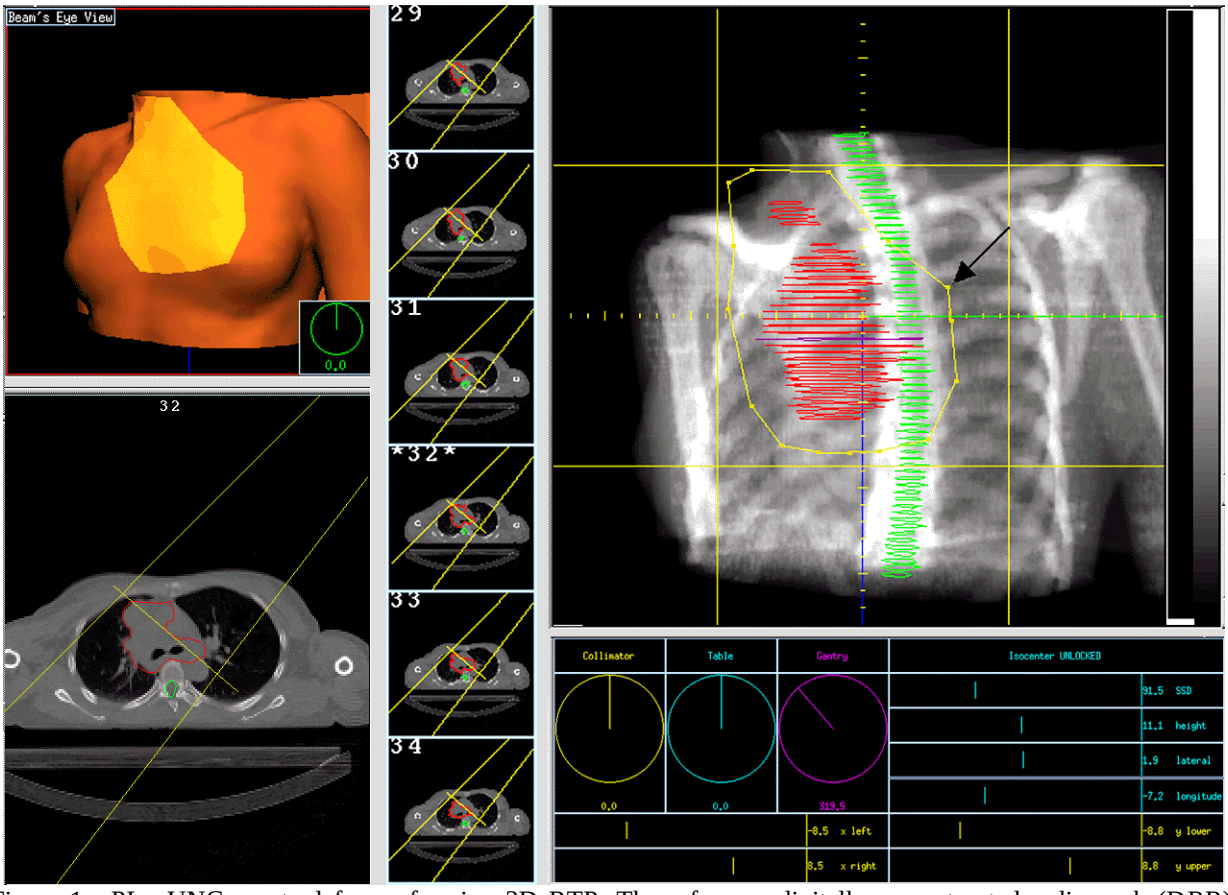

Figure 1. PLanUNC – a tool for performing 3D RTP. The reference digitally reconstructed radiograph (DRR), along with the projection of the tumor and spinal cord, is shown in the upper right panel.

1.2 Verifying the accuracy of patient setup

The main subject of this paper has to do with the second demand for accuracy in conformal radiotherapy; namely, that of accuracy in the delivery of treatment.[1] For conformal therapy to be successful, i.e., for the dose distribution to match the prescription, the patient must receive the prescribed treatment plan. In turn, for the patient to receive the prescribed plan he or she must be positioned accurately with respect to the treatment machine over the entire course of therapy. While skin markers and immobilization devices may help to improve positioning accuracy, the success of conformal therapy depends so highly on accurate patient positioning that more direct methods of setup verification are needed.

Traditionally, portal images acquired during treatment with the mega-voltage treatment beam have been used to verify patient setup. Such treatment images are compared with reference images obtained either (1) using a simulator with the same geometric configuration as the treatment machine but with a diagnostic-energy beam or (2) via synthetic reprojection of the planning CT volumetric image in the case where 3D RTP has been performed. The latter type of reference image is often called a digitally reconstructed radiograph (DRR) and is typically printed to radiographic film [Chaney, 1995]. In standard clinical practice, the radiation oncologist performs verification of patient setup by viewing the reference and portal images side-by-side on a film-viewing light box. The radiation oncologist attempts to quantify setup errors, using simple measuring devices like rulers and protractors, and then records any corrections that are to be made for that particular field and which are then applied by the radiotherapist prior to the next scheduled treatment session. In essence, during this film-to-film comparison the radiation oncologist is performing an image registration, albeit a crude one, based on structures that appear in both images and which are known to be relatively rigid with respect to the tumor.

Unfortunately, for several reasons the current methodology for verifying treatment setup is inadequate, at least in the conformal therapy situation. First, visualization of structures on which to base the image comparison is often very difficult because of the poor quality of portal images, which arises from the imaging physics of the high-energy radiation beam. Second, most setup errors are corrected retrospectively allowing for the possibility that, especially under the escalated doses delivered during conformal treatment, the tumor may fail to receive a sufficient dose and healthy tissues may be irreparably damaged. Third, radiation oncologists are often confounded by compound patient setup errors that involve more than simple translations or in-plane rotations.[2] In fact, recent observer studies have shown that in-plane rotations are often confused with translational errors [Boxwala, 1997]. In the presence of out-of-plane rotations the interpretation and quantification of setup errors are even more confounding.

The recent introduction of electronic portal imaging devices (EPIDs) has created much interest in the radiotherapy community as such devices permit portal images to be acquired in near real time using a small fraction of the dose delivered during an entire treatment session (see Boyer [1992] for a thorough review of EPIDs). Because these images are digital, contrast enhancement algorithms may be applied which can improve the radiation oncologists ability to identify structures on which to base an image comparison [Rosenman, 1993]. Moreover, because of the

[1] We define accuracy in the RMS sense. That is, it includes both systematic and random error components.
[2] In-plane rotation refers to rotations about the central beam axis.

immediacy of such electronic portal images (EPIs), it is possible to check patient setup at the onset of a treatment session, rather than retrospectively as is current practice, allowing for the opportunity to detect and correct setup errors online. Perhaps most importantly, EPIs have fostered a number of computer-based image registration algorithms, which may allow for more accurate determination of patient setup errors [Boyer, 1992].

1.3 Determining three-dimensional patient setup errors

To date, most portal image registration algorithms in the literature have considered only the dimensions of 3D translation and rotation only in the plane perpendicular to the beam axis. Hence, the assumption is that out-of-plane rotations do not exist or at least are negligible. While actual data on the frequency and magnitude of out-of-plane rotational errors is scarce, it is generally held that such errors do occur and that they may have significant dosimetric consequences [Hanley, 1995, 1997]. Moreover, out-of-plane rotational errors may cause registration algorithms that assume such errors do not exist to report erroneous values for patient repositioning.

There have been several reports of attempts to quantify 3D patient positioning errors in radiotherapy using planning CT data and portal images, most coming from a group in the Netherlands Cancer Institute. Bijhold [1993] describes a method based on correspondence of anatomical landmarks in the planning CT and the portal image. The difficulty with this method is that it is often not possible to accurately identify such point fiducials in both 3D and in projection in the portal image. Gilhuijs [1995] describes an interactive approach using visual comparison of portal images with DRRs. This approach has been shown to be effective but is limited to retrospective analysis due to the time requirements. A completely automatic method has also been developed by Gilhuijs [1996] wherein ridge-like features extracted in portal images are matched against projections through the planning CT data. This approach has been shown to work well but requires portal images acquired at two different angles, and thus cannot be used in online mode for patient repositioning.

We argue that an algorithm for determining 3D patient positioning should

1. be unobtrusive such that it requires no modifications to the patient (e.g., implanted or external fiducial markers) nor the treatment machine (e.g., externally mounted x-ray or camera devices),
2. perform well over a population of patients and a number of treatment sites,
3. be insensitive to image disturbances such as noise, blurring, and to non-fiducial structures which vary in shape and position relative to the tumor,
4. be relatively unaffected by image resolution both in the CT and the portal image,
5. involve minimal user interaction,
6. be computationally efficient enough to allow for online patient repositioning.

The method described in the following section is designed to satisfy all of the above criteria.

2 METHODS AND MATERIALS

We define patient pose as the 3D position and orientation of the patient relative to the treatment machine. The objective of 3D patient setup verification is to determine the pose of the patient at the time of treatment relative to the pose of the patient as specified in the treatment plan. In radiotherapy, patient pose may be adjusted by changing six parameters of the treatment machine: the gantry, collimator and table angles and the *x*, *y* and *z* position of the table on which the patient lies.

The basis of our method (shown schematically in Figure 2) is to compare the *treatment pose* as indicated in the portal image with the *planning pose* as indicated in the planning DRR. Since it is not possible to derive the 3D setup error from these two images alone, we perform error measurement by systematically generating candidate poses, which are modifications to the planning pose, computing candidate DRRs at those poses, and finding the candidate pose that “best matches” the portal image. The remainder of this section describes the means by which we can measure this match.

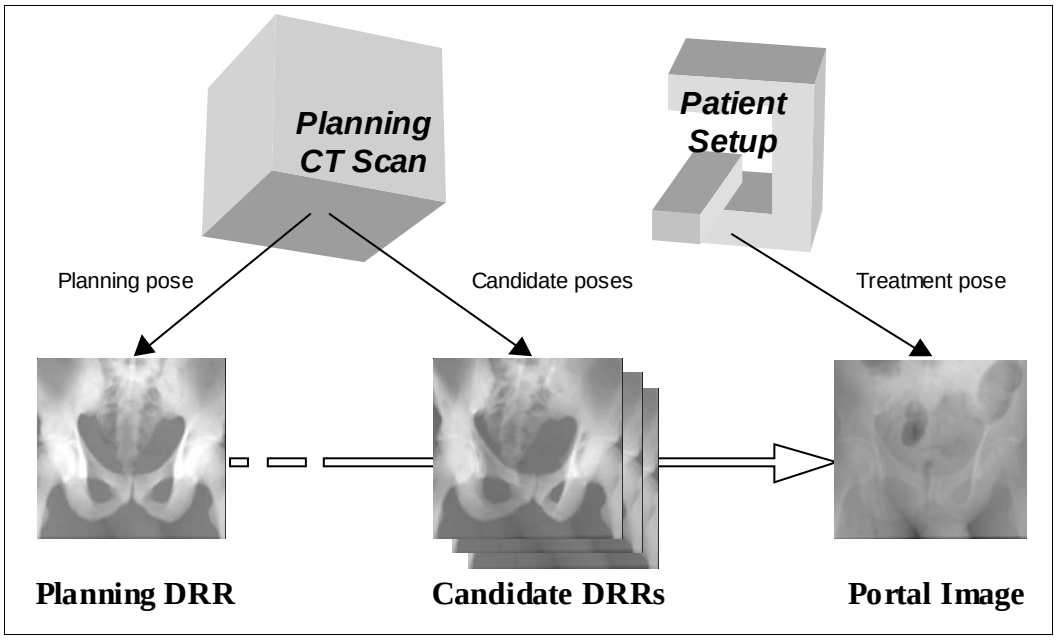

Figure 2. The basis for finding treatment pose from planning pose.

2.1 Overview of the registration method

As illustrated in Figure 2, we should be able to determine the difference in pose of the patient at treatment time and the pose in the treatment plan if we can find that particular candidate pose in which certain structures exhibit the same shape in projection in the portal image and the candidate DRR. We base this shape comparison on fiducial structures that are projections of bones that are known to be relatively rigid with respect to the tumor volume since soft tissue contrast is almost non-existent in the high-energy portal image.

To find that optimal candidate pose requires a means for (1) choosing and representing structures on which the match is to be made from the planning DRR, (2) extracting those same structures in

the portal image, and (3) optimizing the measure of agreement between the extracted structures in the portal image and the same structures in the candidate DRRs. These steps are summarized in Figure 3 and are described in the subsections to follow.

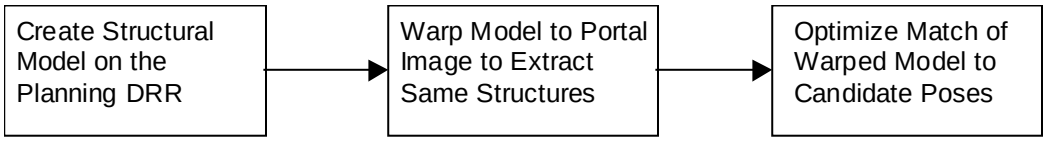

Figure 3. Overview of the registration method.

2.2 Defining reference structures on the Planning DRR

Our method for choosing and representing those structures in the reference image that are to be used in the registration process is based on a theory of figural models of shape [Pizer, 1996; Pizer, 1998]. A figural model can be extracted from one image and deformed to best match structures in another image [Fritsch 1997a]. Figural models consist of collections of medial primitives located along the medial axis, or skeleton, of an object. Medial primitives encode their positions on the medial axis, the widths of the object at each position, and specify an aperture (a spatial weighting function) with which image measurements at the primitive are to be made. We will describe these measurements shortly. The primitive may also encode first order and higher order information, such as orientation of medial axis at the primitive, or normal directions to the boundary at boundary sites corresponding to the primitive. The links between adjacent medial primitives in the figural model encode important shape relationships between the primitives. Such shape relationships are used to limit the deformation of a figural model to those configurations that are reasonable when the figural model is applied to extract like objects in other images.

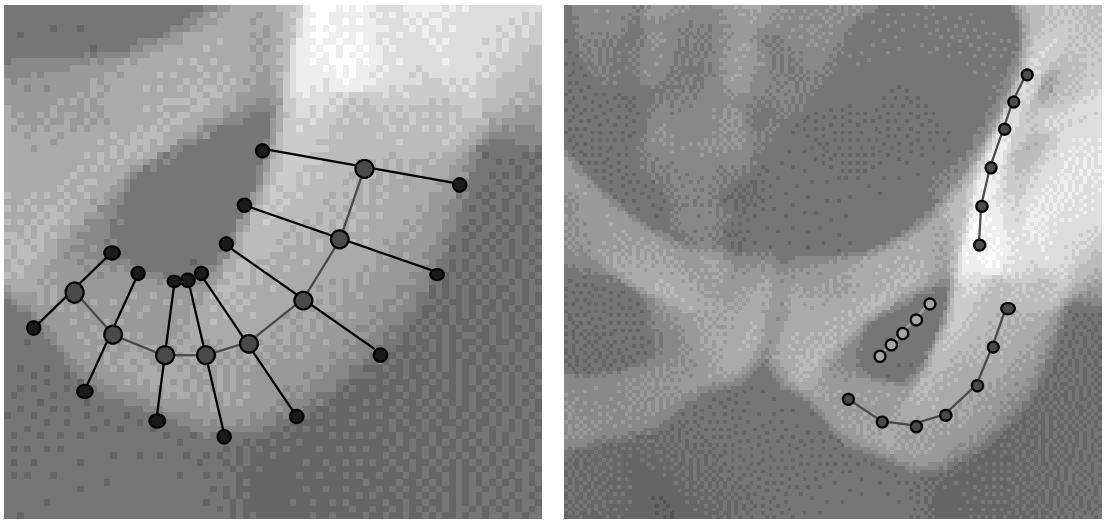

Figure 4. Structural models of the projection of bones and their medial primitives.

Figure 4 (left) shows a medial model corresponding to a bone structure. As discussed in Pizer [1998], such figural models are especially attractive for use in image analysis because (1) they provide shape measures that are invariant under the operations of translation, rotation, and zoom, (2) they provide access to intuitive shape properties such as object width and the widening or narrowing rate of an object and the direction and curvature of the skeleton of the object, (3) they permit efficiency in shape representation due to the coarse sampling of medial primitives, and (4) they can be combined with unclassified image data to extract like objects in a way that is insensitive to noise, blurring, and other image disturbances [Fritsch, 1997a].

We use figural models to analyze and quantify shape of anatomical structures in radiograph projections. Pre-treatment, we extract figural models from the planning DRR using either manual drawing or using an automatic method called core extraction [Pizer 1998]. We include only those structures whose position and shape in the body change little with respect to the tumor between planning and treatment times. Some examples of such structures in projection include pelvic bones, the overlaps of two or more bones, and the gaps between bones. Figure 4 (right) shows models corresponding to each one of these types of structures.

To obtain a figural model from a reference image and to apply it to another image we make measurements at each medial primitive. These measurements involve the application of one or more aperture functions, each centered at the spatial location of the medial primitive and each specific to the type of medial primitive used (i.e., dependent on the type of information encoded in the primitive). The size of the aperture (that is, its spatial "footprint") is proportional to the object width recorded in the medial primitive.

In results to follow, we use a medial primitive that encodes just its medial position $\boldsymbol{x}$ and the object width w. We use a Laplacian of Gaussian aperture function centered at the medial primitive position x, $\pm\sigma^2[G_{xx}(\boldsymbol{x},\sigma)+G_{yy}(\boldsymbol{x},\sigma)]$, where σ is the standard deviation of the Gaussian and is proportional to the medial primitive's width w. The positive signed aperture is used to extract structures whose intensities are darker than the surround while the negative signed aperture is used to extract structures whose intensities are lighter than the surround.

2.3 Warping a structural model to the portal image

We assume that at treatment all machine parameters have been correctly set and the only unknown is the actual pose of the patient at treatment time. In the first instants of treatment we obtain the portal image, a low-contrast projection of the treatment area taken with the high-energy treatment beam. We locate fiducial bone structures in the portal image automatically by deforming structural models that were extracted from the planning DRR. The deformation method is described fully in Fritsch [1997a] and involves optimizing an objective function that is the sum of an image match term and a penalty term that involves a measure of shape change as the model warps. If the portal image is not grossly different from the planned DRR, a reasonable assumption given that the patient should be in a pose rather near to the planned pose, models of bones will deform to match corresponding bones in the portal image. Note that we could instead have chosen to extract medial models from the portal image itself. We choose to perform deformation because it is automatic, while model extraction is an interactive process that should be performed before treatment.

2.4 Measuring the match between a warped model and a candidate DRR

To determine the patient pose, we evaluate image match between the warped structural model from the portal image and a series of DRRs computed from candidate poses. Candidate poses are specified as deviations from the planned treatment pose. A model's image match to a candidate DRR is a sum of image matches at each medial primitive, where each medial primitive applies its own aperture function to measure a property called *medialness*.

Medialness is a measure of how skeletal a medial primitive is in reference to an object in an image. The sum of medialness over all points in a structural model, called integrated medialness, is a measure of match between the model and an image. This measurement is analytically equivalent to applying an aperture function associated with the structural model to the image. Shown in figure 5, such an aperture function is constructed as a sum of individual aperture functions of each medial primitive contained in the structural model.

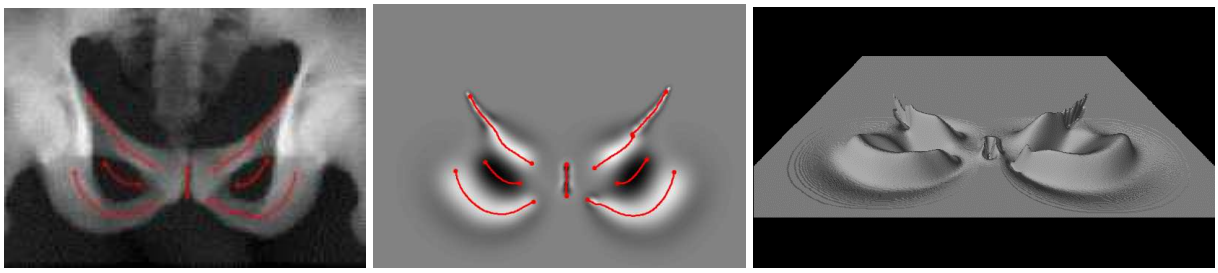

Figure 5. Left : A structural model superimposed on a planning DRR. Middle: The aperture function associated with the structural model. Right: The aperture function shown as a graph.

Since for every candidate pose we generate a corresponding DRR, we can speak of image match computed by integrated medialness as a function of pose, and we call this our objective function. To determine the patient pose, we find the pose that optimizes this objective function. As long as this function is smooth over all feasible poses, a good optimization algorithm can determine the patient pose correctly. One iteration in such optimization algorithm would involve evaluating a pose, which is equivalent to obtaining a DRR projection and computing image match. The issue here is the evaluation of this objective function. If it is to be useful (a) it should have an optimum at a pose with a small error from the actual pose, and (b) it should be efficiently evaluated. In our implementation, it takes under a minute to evaluate one pose on a Pentium®-class machine. The next section describes how we have evaluated the method using simulated setup errors where truth is known exactly.

3 EVALUATION

It is difficult to evaluate the accuracy of registration methods in a clinical situation because geometry of patient setup cannot be known exactly. However, in a computer simulation the truth about treatment geometry and patient pose are known exactly. In such a situation we can examine accuracy of our method relative to known truth. We use the NLM Visible Human frozen CT, $2mm^3$ voxels, and the PLanUNC 3D RTP software to generate planning and candidate DRRs as well as the digitally reconstructed portal radiographs (DRPRs) used to simulate portal images. Methodology for creating realistic DRPRs is described in Fritsch [1997b].

For our method to determine patient pose accurately, it is critical that the image match function has its maximum when the candidate pose matches the treatment pose. Our ability to find this maximum depends on smoothness of the image match function over test poses near the treatment pose. Our pilot results show that the image match function is smooth and reaches maximum at candidate poses very near the treatment pose.

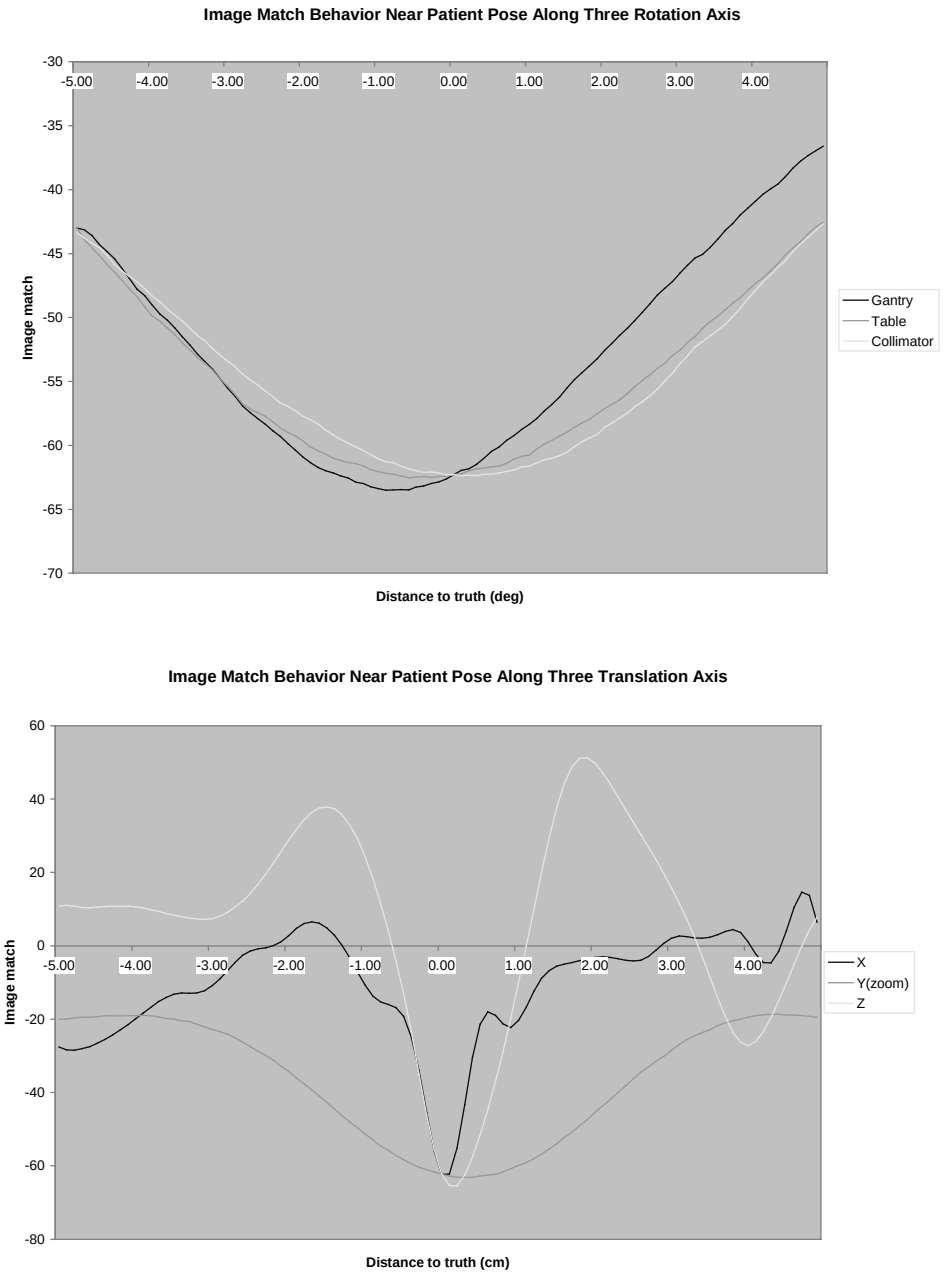

Figure 6. Image match function behavior near the treatment pose, projected on six cardinal axis.

We randomly generate a feasible treatment pose $\boldsymbol{v_0}$ and produce a digitally reconstructed portal radiograph (DPRR) from that pose. In the future, we may add intensity noise or shape noise to this DPRR to more accurately simulate a real portal image of a patient whose anatomy may have

changed from the time the CT scan was obtained. We extract structural models from the simulated portal image. We generate candidate DRRs over a range of poses $\boldsymbol{v} \in [\boldsymbol{v}_0 - k\boldsymbol{u}_i, \boldsymbol{v}_0 + k\boldsymbol{u}_i]$, where $\boldsymbol{u}_i$ is a unit vector in one of six cardinal directions in pose domain. At each pose, we compute image match of the model and DRR. Figure 6 plots the image match function over six cardinal directions about the truth for a single simulated treatment pose. The optimal value of image match is reached within a degree of rotation and several millimeters of translation from the truth. Table 1 shows statistics for 8 such experiments.

	Mean	*Standard Deviation*
Gantry	< 0.31°	< 0.29°
Table	< 0.24°	< 0.30°
Collimator	< 0.15°	< 0.15°

	Mean	*Standard Deviation*
X	< 0.06 cm	< 0.06 cm
Y	< 0.45 cm	< 0.27 cm
Z	< 0.09 cm	< 0.13 cm

Table 1. Distribution of difference of optimum from the truth, in degrees of rotation or cm of translation, over 8 simulated patient poses.

Since model extraction is a manual task, we must address the effect that human bias in extraction has on the results. In our pilot study we found that results obtained in an experiment performed by a second model builder were similar to, and fell within one standard deviation of, the first builder's results which are shown in Table 1.

4 DISCUSSION

Our results were obtained from a single CT scan of one pelvic region. Results from this region are encouraging; however further results on different areas of the body and for more patients are needed to thoroughly evaluate our method. Work in underway to produce such results.

The results show that the pose at which image match is greatest does not coincide exactly with the pose at which structural models are extracted. To understand reasons behind this slight inaccuracy, and to correct it, we must examine structural models and the image match function in close detail. The aperture function used to perform measurements at each medial primitive has a size proportional to the width encoded in the primitive and when evaluated to four standard deviations covers a significant portion of the image. The large size of the aperture has a positive effect on our method because the primitive is "attracted" to structures that are away from the model. This attraction makes the image match function smooth. The disadvantage of a large aperture is in its attraction to other objects in images, such as gas bubbles in portal images and other bones and bone overlaps in the DRR. A structural model is a set of medial primitives that best match a given structure in the image from which they were extracted. However, it is possible that another structure in another image (e.g., a contrast-enhanced image) may have an even greater image match with the same model. In our application we may extract a model at a pose, slightly perturb the pose and find that the DRR at perturbed pose matches the model better than the DRR at the true pose. This is due to large size of the aperture function and dependency on contrast in computation of image match. Since this inaccuracy only occurs near the true pose, we can refine our method once the optimal pose has been found.

Improvements to our method can be made by introducing an alternative type of medial primitives and associated measurement of image match to our method. Such an variation can be introduced either as a replacement of the current type of primitives or as a refinement at a late stage of optimization. A refinement would involve using medial primitives with measurement of image match more sensitive to local image structure, less sensitive to distant areas of the image, and unaffected by changes in contrast between portal and photoelectric images. The matching of grayscale profiles found in work by Cootes et al. [1995] is such a measurement. An oriented medial primitive with an aperture function that is a pair of first derivatives of Gaussians located at boundary sites and elongated along the medial track of the model is another candidate [Pizer 1998].

Models extracted from different structures respond differently to small changes in pose. Vertical structures respond strongly to horizontal shifts and rotations, but do not respond to as well to vertical ones. A way to minimize image match error is to choose the set of structures in such a way that at least one structure responds strongly to any small change in pose. This is achieved by extracting models from projections of bones that are well spread, and differently oriented in the body.

5 CONCLUSION

Our method for registering 3D CT with 2D portal treatment images appears to provide a robust algorithm for determining treatment setup error. Our method bases registration on those anatomical structures that do not vary in shape and position relative to the tumor. We locate such structures in portal treatment images and extract figural models from them. We find the treatment pose by optimizing image match of structural models with DRR projection of the CT made at various, systematically generated guess poses. Our method allows us to detect treatment setup errors in translation, in-plane rotation as well as out of plane rotation. Our research is underway with building a robust optimization strategy for finding the treatment pose, introducing finer image match measurements to improve accuracy at near-optimal poses, building a system integrating model-warping step and optimizations step, and improving performance in order to make our method a desirable alternative to current methods of error detection.

6 ACKNOWLEDGEMENTS

The research reported in this paper has been conducted under the support of NCI Grant 1 RO1 CA67183, Fundamental Approaches for Portal Image Processing. The authors thank other members of UNC Computer Science and Radiation Oncology Departments, in particular Tim Cullip, Alan Liu and Gregg Tracton for their assistance and advice.

7 REFERENCES

Bijhold J, Three-dimensional verification of patient placement during radiotherapy using portal images. Med. Phys. 20 (2):347-356, 1993.

Boxwala AA, Friedman CP, Fritsch DS, Rosenman JG, Chaney EL: Acceptability and usage patterns of an image analysis workstation. In Ed: D. Masys. Proc AMIA Ann Fall Symp., Philadelphia: Hanley and Belfus. 444-448 (1997)

Boyer AL, Antonuk L, Fenster A, van Herk M, Meertens H, Munro P, Reinstein LE, Wong J, A review of electronic portal imaging devices (EPIDs). Med. Phys. 19:1-16, 1992.

Chaney EL, Thorn JS, Tracton G, Cullip T, Rosenman JG, Tepper JE, A portable software tool for computing digitally reconstructed radiographs. Int. J. Radiat. Oncol. Biol. Phys. 31(2):491-497, 1995.

Cootes TF, Cooper D, Taylor CJ, Graham J, Active shape models - their training and application. Computer Vision and Image Understanding. Vol. 61(1):38-59, 1995.

Fritsch DS, Raghavan S, Boxwala A, Earnhart J, Tracton G, Cullip T, Chaney EL, Benchmark test cases for evaluation of computer-based methods for detection of setup errors: realistic digitally reconstructed electronic portal images with known setup errors. Int. J. Radiat. Oncol. Biol. Phys. 37(1):199-204, 1997b.

Fritsch DS, Pizer SM, Yu L, Johnson V, Chaney EL, Segmentation of medical image objects using deformable shape loci. Proc. Information Processing in Medical Imaging (IPMI97), J Duncan and G Gindi, eds., Springer-Verlag, Lecture Notes in Computer Science 1230:127-140, 1997a.

Guilhuijs KGA, Drukker K, Touw A, van de Ven, PJH, van Herk, M, 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, 1995.

Guilhuijs KGA, van de Ven, PJH, van Herk, M. Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. Med. Phys. 23(3):389-399, 1996.

Hanley J, Lumley MA, Mageras GS, Sun J, Zelefsky MJ, Leibel SA, Fuks Z, Kutcher GJ, Measurement of patient positioning errors in three-dimensional conformal radiotherapy of the prostate. Int. J. Radiat. Oncol. Biol. Phys. 37:435-444, 1997.

Hanley J, Mageras GS, Sun J, Kutcher GJ, 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. 33:1331-1343, 1995.

Pizer SM, Fritsch DS, Johnson V, Chaney EL, Segmentation, registration, and measurement of shape variation via image object shape, 1996. University of North Carolina Computer Science Department Technical Report TR96-020. Tutorial notes, *Visualization in Biomedical Computing* ’96.

Pizer SM, Fritsch DS, Low K, Furst JD, 2D & 3D Figural models of anatomic objects from medical images. To appear in Proc. Int. Soc. Math. Morph. 98, Kluwer series in Computational Imaging and Vision.

Rosenman JG, Roe CA, Cromartie R, Muller KE, Pizer SM, Portal film enhancement: Technique and clinical utility. Int. J. Radiat. Oncol. Biol. Phys. 25:33-338, 1993.

Sherouse GW, Chaney EL, The portable virtual simulator. Int. J. Radiat. Oncol. Biol. Phys. 21:475-482, 1991.