A learning-based patient repositioning method from limited-angle projections

Chen-Rui Chou[1], C. Brandon Frederick[2], Sha Chang[3], and Stephen M. Pizer[1,2,3]

[1] Department of Computer Science, UNC-Chapel Hill, USA,
[2] Department of Biomedical Engineering, UNC-Chapel Hill, USA,
[3] Department of Radiation Oncology, UNC-Chapel Hill, USA.

Abstract. This paper presents a novel patient repositioning method from limited-angle tomographic projections. It uses a machine learning strategy. Given a single planning CT image (3D) of a patient, one applies patient-specific training. Using the training results, the planning CT image, and the raw image projections collected at the treatment time, our method yields the difference between the patient's treatment-time position and orientation and the planning-time position and orientation. In the training, one simulates credible treatment-time movements for the patient, and by regression it formulates a multiscale model that expresses the relationship giving the patient's movements as a function of the corresponding changes in the tomographic projections. When the patient's real-time projection images are acquired at treatment time, their differences from corresponding projections of the planning-time CT followed by applications of the calculated model allows the patient's movements to be estimated. Using that estimation, the treatment-time 3D image can be estimated by transforming the planning CT image with the estimated movements, and from this, changes in the tomographic projections between those computed from the transformed CT and the real-time projection images can be calculated. The iterative, multiscale application of these steps converges to the repositioning movements. By this means, this method can overcome the deficiencies in limited-angle tomosynthesis and thus assist the clinician performing an image-guided treatment. We demonstrate the method's success in capturing patients' rigid motions with sub-voxel accuracy with noise-added projection images of head and neck CTs.

1 Introduction

Approximately 40% of people with cancer have radiotherapy as part of their treatment. While damaging cancer cells, radiation can also affect surrounding healthy cells if it is not directed with a high degree of accuracy. In order to attain high accuracy, a real-time imaging device can be employed to guide the radiation treatment. This approach is known as image-guided radiotherapy (IGRT). To

support IGRT, traditional imaging approaches, such as computed tomography (CT), are not very appropriate. Despite CT's fine 3D spatial resolution in imaging tissue, the expense of having a CT scanner in each treatment room, the long image acquisition time and the intense dose of ionizing radiation make its direct application in real-time imaging impractical. As described in section 2, recent tomographic imaging advances, like nanotube stationary tomosynthesis (NST) [1] and cone-beam CT (CBCT), are designed to decrease the number of imaging sources or limit the range of projection orientations to achieve higher image acquisition speed, lowered dose and lowered cost. However, with the deficiency in geometry, traditional tomographic reconstruction methods, like filtered-back projection (FBP) or the simultaneous algebraic reconstruction technique (SART) [2], cannot reconstruct a 3D image with full volumetric information but rather induce many streaking artifacts. (See Fig.1)

(a) **(b)** **(c)**
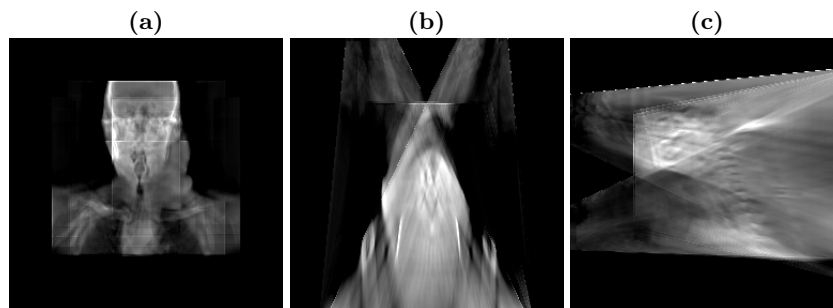
Fig. 1. (a) Coronal, **(b)** axial and **(c)** sagittal views of the head and neck image reconstructed from NST's 52 projections. The reconstruction method is SART with 20 iterations.

One way to overcome this geometric deficiency is to use the patient's planning CT image as prior information. With this prior information, one can utilize image registration techniques to register the prior image to either the 2D raw projections [3] or the reconstructed 3D image [4,5,6,7]. In [5,6,7], mutual information was used to register the planning CBCT to a 3D image reconstructed from a set of limited-angle projections (80 projections in 45° DTS [5,6,7].) Their registration quality varied with the landmark objects used for MI calculations: with bony structures, the mean shift error is about 2 mm [5]; while with soft tissues, there are up to 1.7 mm "one-dimensional" shift errors [7]. Apparently, even when high contrast bony structures are present, their methods require many projections to detect patient's movements. In [8], they register the prior CT (3D) images to several X-ray projection (2D) images by maximizing the similarity between the X-ray projection images and those in the CT projection space. However, almost half of their test cases failed to converge to the correct positions when the initial shifts are more than 1 cm.

Our method, on the other hand, can capture patient's shifts up to 3 cm. Moreover, our method requires very few projections (e.g., 2 for both 5° CBCT and NST) but yields patient's treatment-time 3D motions with sub-voxel accuracy in simulated oblique-angle NST and simulated limited-angle CBCT for the head and neck.

Our method incorporates the image-residuals-to-shape-parameters learning strategy in Active Appearance Models (AAMs) [9]. It involves two stages: training and treatment. In the training stage we sample from a range of patient movements, and for each such movement we generate 2D projections by transforming and reprojecting the patient's planning CT. We compute a linear regression between the patient movements and the difference between the projections of the moved CT and those of the CT in the planning position. In the treatment stage, the learned regression model is applied iteratively to the successive residues between the real-time projections and those of the moving CT transformed by the previously predicted parameters. This iteration yields the predicted transformations with high accuracy.

This paper is organized as follows: In section 2, we sketch the imaging arrangements of NST and of limited-angle CBCT. In section 3, we explicitly describe our method with three parts: training, treatment and an implementation design. In section 4, we demonstrate the numerical results and analyze upon them. Finally, we conclude our paper, make some interesting observations and set future work in section 5.

2 Imaging geometries

While neither CBCT and NST are not presently available for our experiments, we simulate their IGRT projection images by reprojecting the planning-time CT based on the corresponding IGRT geometries. Therefore we describe the imaging geometries as below.

2.1 Nanotube Stationary Tomosynthesis (NST)

NST [1] is a stationary imaging system that can perform imaging without interfering with treatment delivery. As illustrated in Fig. 2 (a), it consists of an arrangement of radiation sources arrayed around the treatment portal, together with a planar detector. Firing the sources in sequence produces a different 2D projection image per source. Each projection image requires ∼200 ms. The imaging dose is very low relative to CT.

2.2 Limited-angle Cone-Beam CT (CBCT)

CBCT is a rotational imaging system with a single radiation source and a planar detector. This pair rotates by a limited angle during IGRT, taking projection

images during traversal (Fig. 2 (b)). The limited-angle scan provides the shortened imaging time and the lowered imaging dose, e.g., for 5° rotation, it takes less than one second.
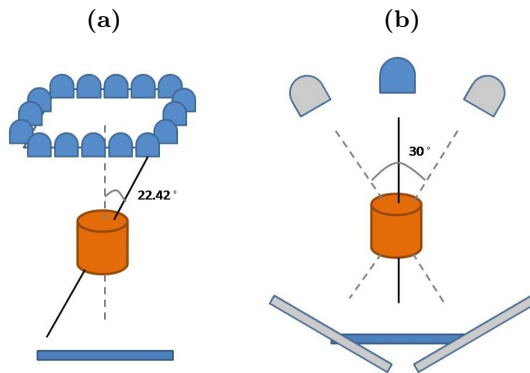
Fig. 2. **(a)** The NST geometry: stationary sources arrays with max tomographic angle=22.42°**(b)** The 30° CBCT geometry: rotational imaging system with tomographic angle=15°.

3   Method

Our method is made up of two stages: training (sec. 3.1) and treatment (sec. 3.2). A particular form of training, which operates hierarchically by scale, has turned out to yield a good implementation (sec. 3.3).

3.1   The training stage

The training stage requires a patient's planning 3D image with full volumetric information, e.g., a CT image: CT_{plan}, prior to the radiotherapy. The planning image, CT_{plan}, is used as a reference to generate N possible treatment-time images, $\mathbf{CT}_{trans}$, by a collection $\mathbf{C}$ of credible rigid transformations C_i. The rigid transformations are sampled from a space Ω represented by the six rigid-motion parameters t_x, t_y, t_z for translation; r_x, r_y, r_z for rotation (Euler angles):

$$\mathbf{CT}_{transf} = \{CT^i_{transf} = Transform(CT_{plan},\ C_i)\mid \bigcup_{i=1}^{N} C_i \in \mathbf{C}\} \qquad (1)$$

$$\mathbf{C} = span(t_x,\ t_y,\ t_z,\ r_x,\ r_y,\ r_z) \in \Omega \qquad (2)$$

At training time our method computes the residuals between 2D projections of CT_{plan} and the corresponding projections of $\mathbf{CT}_{trans}$. The tomographic projections are generated by reprojecting the 3D images, with operator $\mathbf{P}$, based on the

geometry of the target machine, e.g. NST or limited-angle CBCT. The residuals array $\mathbf{R}$ is a concatenation, over the projections of the differences between corresponding projection pairs. In the work described here the difference between two projections is simply the pixel-by-pixel tuple of intensity differences.

$$\mathbf{R}_{i,j} = P_j(CT^i_{transf}) - P_j(CT_{plan}), \tag{3}$$

where P_j is the j-th projection in $\mathbf{P}$ and $1 \leqslant j \leqslant M$.

$$\mathbf{R} = \mathbf{P}(\mathbf{CT}_{transf}) - \mathbf{P}(CT_{plan}) \tag{4}$$

$$= \begin{bmatrix} P_1(CT^1_{transf},\ C_1) - P_1(CT_{plan}) & \cdots & P_M(CT^1_{transf},\ C_1) - P_M(CT_{plan}) \\ P_1(CT^2_{transf},\ C_2) - P_1(CT_{plan}) & \cdots & P_M(CT^2_{transf},\ C_2) - P_M(CT_{plan}) \\ \vdots & \ddots & \vdots \\ P_1(CT^N_{transf},\ C_N) - P_1(CT_{plan}) & \cdots & P_M(CT^N_{transf},\ C_N) - P_M(CT_{plan}) \end{bmatrix}$$

We calculate the matrix β that when multiplied by the residuals array $\mathbf{R}$ minimizes squared errors in the prediction of the transformation parameters $\mathbf{C}$ (5). That is, our estimate $\hat{\beta}$ is calculated by a deterministic linear regression (6).

$$\mathbf{C} = \mathbf{R} \cdot \beta + \epsilon \tag{5}$$

$$\hat{\beta} = (\mathbf{R^T R})^{-1} \mathbf{R^T C} \tag{6}$$

Fig. 3 (a) is the flow chart of the training procedure.

3.2 The treatment stage

At treatment time the objective of each iteration i is to produce a geometric transformation C_i to be applied to the present best estimate of the treatment CT, as well as a new estimate of the treatment CT. Once the patient's real-time projections $\mathbf{P}_{real-time}$ are collected, we apply an iterative procedure, (7)-(9), to transform and reproject the planning 3D image such that its tomographic projections, $\mathbf{P}(CT_{moving})$, are consistent with the real-time ones. The transformation C_i is estimated by the projection residuals $\mathbf{R}$ and the weightings $\hat{\beta}$ calculated in the training stage. The reprojection process is implemented on GPUs to make it run within 10 ms.

$$\mathbf{R} = \mathbf{P}_{real-time} - \mathbf{P}(CT_{moving}) \tag{7}$$

$$C_i = \mathbf{R} \cdot \hat{\beta} \tag{8}$$

$$CT_{moving} = Transform(CT_{moving}, C_i) \tag{9}$$

Because the composite changes of multiple rigid-motion parameters may be linearly dependent with other composite changes, instead of using the full transformation computed by (8), at each iteration of (7)-(9) we transform the moving image by the one rigid-motion parameter in C_i that decreases the magnitude of the residual $||\mathbf{R}||_2$ most. The method is taken to have converged if the magnitude of the residual is below the threshold. Fig. 3 (b) is the flow chart of the treatment procedure.
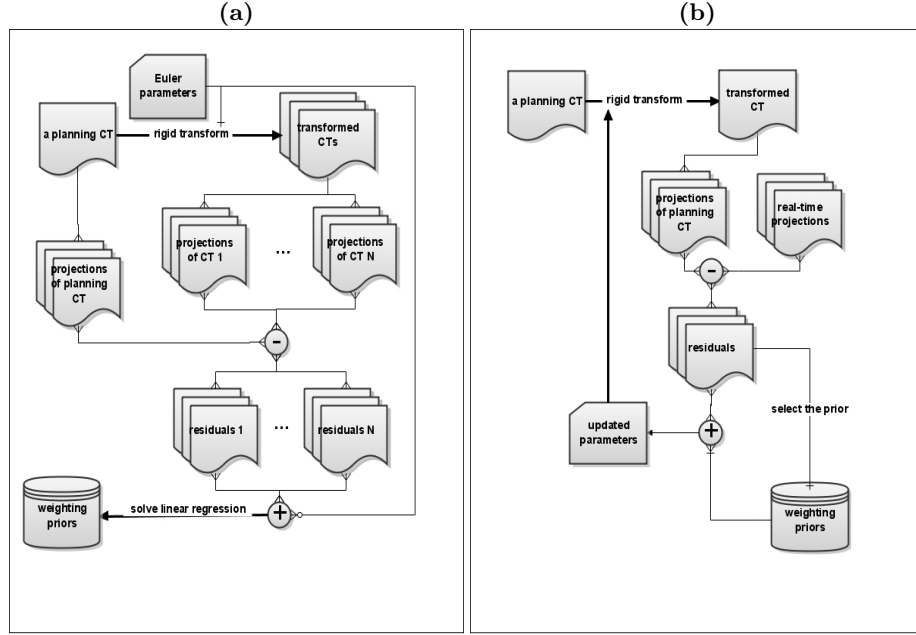
Fig. 3. Flow chart of **(a)** the training procedure and **(b)** the treatment procedure.

3.3 Hierarchical trainings

It is unrealizable to enumerate many combinations in **C**. Therefore, we used a piecewise approximation scheme called "hierarchical training":

In the training stage, K hierarchical weighting matrices $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_K$ are generated from large to small scales of training. At the k-th level of training, the transformation parameters are collected from the combinations of translations with $\pm\mu_k, 0$ (cm) and rotations with $\pm\nu_k, 0$ (degrees). In order to have accurate estimations in the whole training domain, the selection of μ_k and ν_k depends on the interpolation accuracy of $\hat{\beta}_k$ calculated in (6) such that the union of each level's confidence intervals, $\tau_{translation}^k$ and $\tau_{rotation}^k$, covers the whole training domain. (See 1D example in Fig. 4)

$$\overset{K}{\underset{k=1}{\cup}} \tau_{translation}^k \supseteq [-Tr_{limit},\ Tr_{limit}], \quad \overset{K}{\underset{k=1}{\cup}} \tau_{rotation}^k \supseteq [-Ro_{limit},\ Ro_{limit}] \quad (10)$$

where $\pm Tr_{limit}$ and $\pm Ro_{limit}$ are the training limits for translations and rotations, respectively. For each level of training, the training values in **C** are formed by all combinations, over the six rigid motion parameters, of $\pm Tr_{limit}$ and $\pm Ro_{limit}$, respectively.

In the treatment stage the calculated hierarchical weighting matrices are applied sequentially, from $\hat{\beta}_1$ to $\hat{\beta}_K$, to give new advection forces when our optimizer fails to decrease the magnitude of the residual by all of the six rigid-motion

parameters. In our current implementation, we applied a four-level hierarchy ($K = 4$) with $\mu_k = (K - k + 1) \cdot \frac{Tr_{limit}}{K}$, $\nu_k = (K - k + 1) \cdot \frac{Ro_{limit}}{K}$. As a result, after applying the hierarchical training, our optimizer can almost surely converge to the global minimum in the residual space.
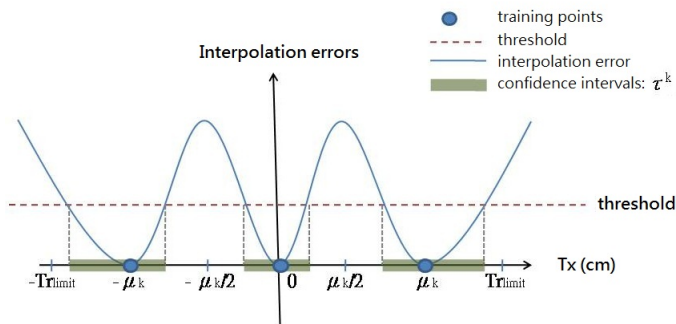

Fig. 4. An illustration of the confidence intervals of the translations along x-axis in the k-th level of training. $\pm\mu_k$, 0 are the training points. Confidence intervals τ^k are calculated by a threshold that outputs the limits of the tolerable wrong estimations. $\pm Tr_{limit}$ are the training limits of translations.

4 Results

4.1 Test Environments

We tested our method using three patients' (P5202, P5203, P5204) head and neck CTs with a voxel size of 0.12 cm lateral × 0.12 cm anterior-posterior × 0.30 cm superior-inferior. The simulated treatment-time projections were generated by transforming and reprojecting the patient's planning CT and then adding Gaussian-distributed noise to each pixel. A total of 180 testing moves, 90 in NST and 90 in CBCT, were combinatorial in both translations and rotations and were randomly picked within -2 to 2 cm and -5 to 5 degrees, respectively. In CBCT, imaging at a variety of arc angles have been tested; the performance with the smallest arc angle with good performance, 5 degrees, is reported. In both NST and CBCT experiments the number of imaging positions was varied to find the minimum number with sub-voxel accuracy.

Zero mean, constant standard deviation Gaussian noise was added to the simulated IGRT projection images. The standard deviation σ of the noise=0.2·(mean bony intensity−mean soft tissue intensity). This noise level is far higher than that produced in either NST or CBCT. Example projection images are shown in Fig. 5.
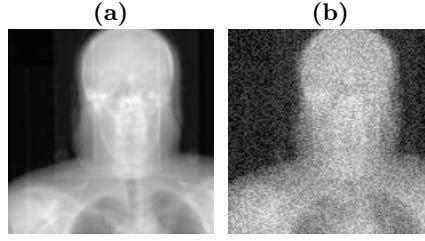
Fig. 5. Sample simulated CBCT projection images: **(a)** before adding noise **(b)** after adding Gaussian noise.

The errors are measured by mean absolute error (MAE) and mean target registration error (mTRE). The MAE in any dimension is the mean, over the test cases, of the error magnitude in the corresponding dimension of $\mathbf{C}$. mTRE for a case is the mean displacement error, over all voxels in a $16 \times 16 \times 16$ cm^3 bounding box (probable tumor region) centered at the center of the pharynx in the planning image (11).

$$mTRE = \frac{1}{N} \sum_{i=1}^{N} \|T_{true} \cdot \mathbf{x}_i - T_{est} \cdot \mathbf{x}_i\|_2 \qquad (11)$$

where N is the number of pixels in the probable tumor region, $\mathbf{x}_i$ is the tuple of the i-th pixel position, and T_{true}, T_{est} are the true and the estimated transformation matrices, respectively.

4.2 Tests on oblique-angle NST and limited-angle CBCT
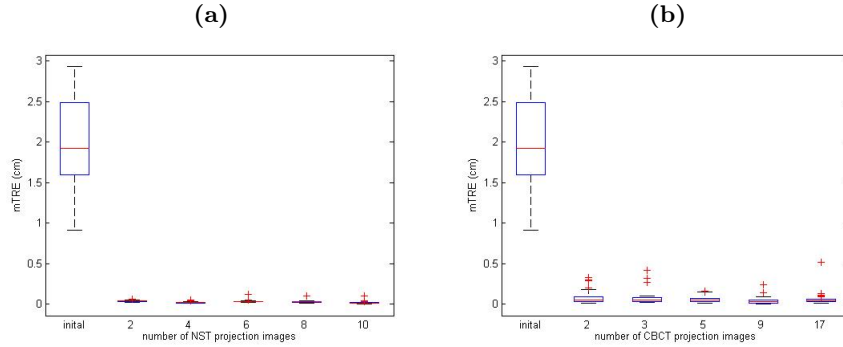
Fig. 6. mTRE statistics on a set of 30 random moves of patient P5202 with **(a)** NST, 4 level of training and the varied numbers of projection images **(b)** CBCT, 4 level of training and the varied numbers of projection images.

We first studied how many projection images are needed to obtain sub-voxel accuracy. The results on patient P5202, displayed in Fig. 6, show that two projection images are enough for both NST and 5° CBCT to have sub-voxel accuracy. Fig. 7 shows the geometries for the two x-ray sources in NST and CBCT, respectively.
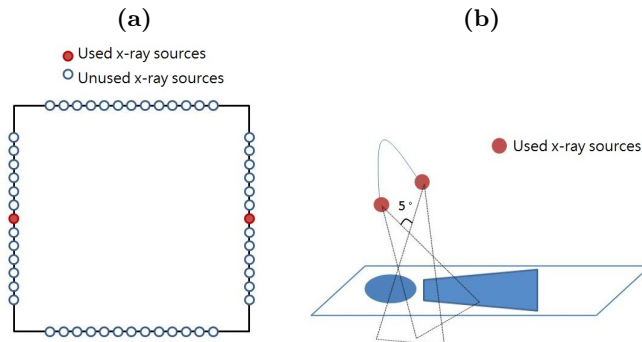
Fig. 7. (a) The geometry of the 2 x-ray sources used in generating the simulated NST projections. **(b)** The geometry of the 2 x-ray sources used in generating the simulated 5° CBCT projections.

Next we studied the accuracy of our method with the geometries shown in Fig. 7. The results, over all three patients are displayed in Tables 1 and 2, are that both the mean absolute errors (MAEs) and their standard deviations (SDs) are less than 0.09 cm in each dimension of translation and less than 0.3 degree in each dimension of rotation. Moreover, in Fig. 8, 97.8% (176 out of 180) of mTREs in our test cases are less than one voxel size, 0.34 cm. Due to the larger tomographic angle of NST (about 20°) rather than CBCT's (2.5°), the NST registration errors are smaller than the registration errors of CBCT. In addition, for both NST and CBCT, the registration errors in Ry are smaller because major projection images lie in the patient's plane which is parallel to the rotation plane of Ry.

Table 1. Position and Orientation Errors - NST with 2 projections

	Tx(cm)	Ty(cm)	Tz(cm)	Rx(deg)	Ry(deg)	Rz(deg)	mTRE(cm)
MAE	0.0094	0.0302	0.0262	0.1489	0.0248	0.1540	0.0524
SD	0.0085	0.0211	0.0715	0.1093	0.0174	0.2824	0.0728

**Table 2.** Position and Orientation Errors - 5° CBCT with 2 projections

	Tx(cm)	Ty(cm)	Tz(cm)	Rx(deg)	Ry(deg)	Rz(deg)	mTRE(cm)
MAE	0.0165	0.0904	0.0353	0.0877	0.0302	0.1334	0.1094
SD	0.0141	0.2007	0.0817	0.0845	0.0217	0.2998	0.2130

deg: degree; Axis directions– x: RL; y: AP; z: SI

(a)   **(b)**
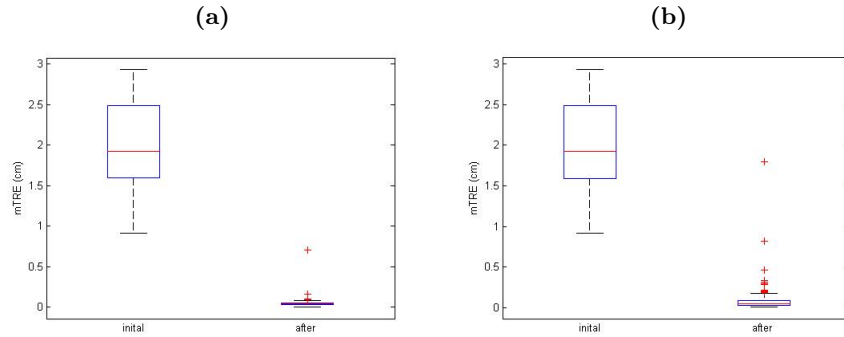
Fig. 8. The mTRE in **(a)** 90 NST and **(b)** 90 CBCT test cases across three patients. The mTREs between the target and the planning images are marked as "initial," and the mTREs between the estimated and the target images are marked as "after." The means and the standard deviations of mTRE after repositioning are 0.0524 ± 0.0728 cm and 0.1094 ± 0.2130 cm for NST and CBCT test cases, respectively.

Finally we studied the effect of how many levels of training were used. Fig. 9 (a)(b) show that increasing the level of training significantly reduces the registration errors in both NST and CBCT.
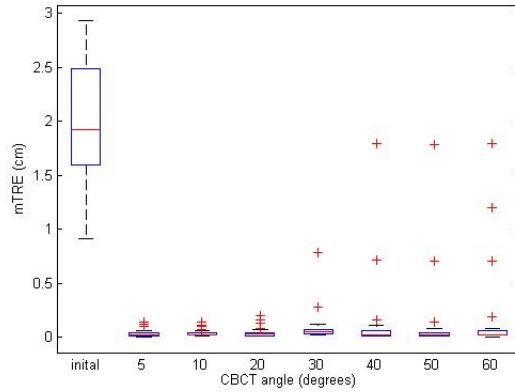
(a)   **(b)**
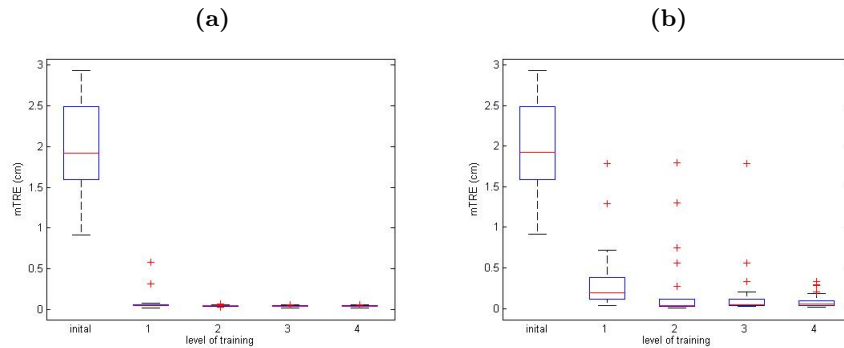
Fig. 9. mTRE statistics on a set of 30 random moves of patient P5202 with **(a)** NST, 2 projection images and the varied levels of training **(b)** CBCT, 2 projection images, 5 degrees and the varied levels of training.

5 Discussion and Conclusion

Remarkably, our novel patient repositioning method requires only two projections and as little as $2.5°$ tomographic angle but yields patient's treatment-time position and orientation for the head and neck with sub-voxel accuracy as well as

a fully-volumetric estimated treatment-time 3D image. Our method is a general patient positioning scheme for IGRT across limited-angle imaging modalities. The imaging time required is about 400 ms for NST and about 1 second for CBCT with 2.5° tomographic angle. After image acquisition, the average computation time of our method is 3 seconds (can be extended to half second with multithreading). Thus the full method is fast enough for clinical use.

Fig. 10. mTRE statistics on a set of 30 random moves of patient P5202 with 2 CBCT projection images, 4 level of training and the varied angles.

Fig. 10 studies the repositioning errors as a function of CBCT projection angle. In each case two projections are used. Surprisingly, the number of repositioning values with more than sub-voxel errors increases with the CBCT projection angle. We must study this behavior further but suggest it shows a sensitivity of our method to the initial values of the six rigid-motion parameters or it implies that the confidence interval of our 4-level training doesn't completely cover the training domain.

Our method has been tested only with head and neck images simulated from patient CTs. It needs to be tested on real IGRT images. Also, it needs to be tested on CT images with lower contrast, such as those of soft tissue. Initial testing on simulated rigid transformations of soft tissue images, namely those of the breast, have shown good accuracy.

A future extension of our method is to capture not only the rigid transformations but also those in combination with treatment-time organ/tissue warps. The approach will use a shape space of credible warps.

6 Acknowledgments

References

1. Maltz, J.S., Sprenger, F., Fuerst, J., Paidi, A., Fadler, F., Bani-Hashemi, A.R.: Fixed gantry tomosynthesis system for radiation therapy image guidance based on a multiple source x-ray tube with carbon nanotube cathodes. Medical Physics **36**(5) (2009) 1624–1636

2. Andersen, A.H., Kak, A.C.: Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. Ultrasonic imaging **6**(1) (1984) 81–94

3. Sadowsky, O., Ramamurthi, K., Ellingsen, L.M., Chintalapani, G., Prince, J.L., Taylor, R.H.: Atlas-assisted tomography: registration of a deformable atlas to compensate for limited-angle cone-beam trajectory. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006. (2006) 1244–1247

4. Thilmann, C., Nill, S., Tucking, T., Hoss, A., Hesse, B., Dietrich, L., Bendl, R., Rhein, B., Haring, P., Thieke, C.: Correction of patient positioning errors based on in-line cone beam CTs: clinical implementation and first experiences. Radiation Oncology **1**(1) (2006) 16

5. Wu, Q.J., Godfrey, D.J., Wang, Z., Zhang, J., Zhou, S., Yoo, S., Brizel, D.M., Yin, F.F.: On-board patient positioning for head-and-neck IMRT: comparing digital tomosynthesis to kilovoltage radiography and cone-beam computed tomography. International Journal of Radiation Oncology, Biology, Physics **69**(2) (2007) 598–606

6. Yoo, S., Wu, Q.J., Godfrey, D., Yan, H., Ren, L., Das, S., Lee, W.R., Yin, F.F.: Clinical evaluation of positioning verification using digital tomosynthesis (DTS) based on bony anatomy and soft tissues for prostate image-guided radiation therapy (IGRT). International journal of radiation oncology, biology, physics **73**(1) (2009) 296

7. Zhang, J., Wu, Q.J., Godfrey, D.J., Fatunase, T., Marks, L.B., Yin, F.F.: Comparing digital tomosynthesis to Cone-Beam CT for position verification in patients undergoing partial breast irradiation. International Journal of Radiation Oncology, Biology, Physics **73**(3) (2009) 952–957

8. Russakoff, D., Rohlfing, T., Maurer, C.: Fast intensity-based 2D-3D image registration of clinical data using light. In: Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. (2003) 416–422 vol.1

9. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(6) (2001) 681–685