2013

Time dependent cone-beam CT reconstruction via a motion model optimized with forward iterative projection matching

David Staub
Virginia Commonwealth University

Time dependent cone-beam CT reconstruction via a motion model optimized with forward iterative projection matching

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

By

David Staub
B.S., Duke University, May 2008

Director: Martin J. Murphy, Ph.D., Department of Radiation Oncology

Virginia Commonwealth University
Richmond, Virginia
April, 2013

Approved 5/8/2013

Acknowledgement

My completion of this dissertation was made possible with the support of a large number of people, both inside and outside the department. Within the faculty ranks I would first and foremost like to thank my advisor Dr. Martin Murphy. He provided willing guidance when I needed it, and his clear gaze on the big picture helped steady me when I got bogged down in the little details. Dr. Murphy's obvious enthusiasm for his work and rigorous approach to problem solving has served as a shining example to apply to my own (hopefully) budding scientific career. I would also like to thank our program director Dr. Jeff Siebers for the unwavering support he has shown to all the students in VCU's Medical Physics program, myself included. Knowing that the guy in charge is in your corner has been a huge asset during this process. In addition, I'd like to thank Dr. Geoff Hugo for sharing patient data-sets, giving me advice, and allowing me to pester him with numerous questions. Dr. Hugo's door is always open and he reigns supreme as the most easygoing faculty member. I would also like to recognize Dr. Scott Brock, a post-doc in our group, and one of the nicest people I've met during my time at VCU. The FIPM-bspline algorithm that became an integral part of my thesis is his work, and he cheerfully spent hours of his time helping me get it up and running and troubleshooting problems. Likewise I would like to thank Andrew Sampson who helped me set up and run the PTRAN Monte Carlo code used for scatter calculations, a harrowing process to which he cheerfully donated hours. I would also like to thank Shufei Chen for providing reconstructed 3D marker trajectories which aided us greatly in the final validation of our method. In addition I would like to thank Deanna Pace, our program administrator, who does a ton of work behind the scenes and tenaciously bugged me to get things done on time. Beyond the department I would like to thank my girlfriend Mackenzie, who has made my life in Virginia so much richer and keeps me grounded by regularly reminding me that I'm a huge nerd. Last but not least I would like to thank my mom Sandra, my dad Jonathan, and my sister Judy for gamely trying to read my papers and supporting and being proud of me despite my terrible explanations of what it is that I've been doing up here.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

4DCBCT	Four dimensional cone-beam computed tomography
4DCT	Four dimensional computed tomography
4DFBCT	Four dimensional fan-beam computed tomography
AAPC	Auto adaptive phase correlation
AP	Anterior-posterior
ASD-POCCS	Adaptive steepest descent projection onto complex sets
BFGS	Broyden-Fletcher-Goldfarb-Shano
CBCT	Cone-beam computed tomography
CR	Computed radiography
CT	Computed tomography
CTV	Clinical target volume
DIR	Deformable image registration
DRR	Digitally reconstructed radiograph
DVF	Displacement vector field
FBCT	Fan-beam computed tomography
FBP	Filtered backprojection
FIPM	Forward iterative projection matching

GPU	Graphics processing unit
GSL	GNU scientific library
GTV	Gross tumor volume
i.u.	Intensity units
LAC	Linear attenuation coefficient
LR	Left-right
MKB	McKinnon-Bates
MR	Magnetic resonance
PC	Phase correlated
PCA	Principal component analysis
PDF	Probability density function
PICCS	Prior image constrained compressed sensing
PSF	Point spread function
PTV	Planning treatment volume
RMS	Root mean square
RMSE	Root mean square error
RPM	Real-time Position Management
SI	Superior-inferior
SNR	Signal-to-noise ratio
SSD	Sum of squared differences

Abstract

TIME DEPENDENT CONE-BEAM CT RECONSTRUCTION VIA A MOTION MODEL OPTIMIZED WITH FORWARD ITERATIVE PROJECTION MATCHING

By David Staub, B.S.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2013.

Director: Martin J. Murphy, Ph.D., Department of Radiation Oncology

The purpose of this work is to present the development and validation of a novel method for reconstructing time-dependent, or 4D, cone-beam CT (4DCBCT) images. 4DCBCT can have a variety of applications in the radiotherapy of moving targets, such as lung tumors, including treatment planning, dose verification, and real time treatment adaptation. However, in its current incarnation it suffers from poor reconstruction quality and limited temporal resolution that may restrict its efficacy. Our algorithm remedies these issues by deforming a previously acquired high quality reference fan-beam CT (FBCT) to match the projection data in the 4DCBCT

data-set, essentially creating a 3D animation of the moving patient anatomy. This approach combines the high image quality of the FBCT with the fine temporal resolution of the raw 4DCBCT projection data-set. Deformation of the reference CT is accomplished via a patient specific motion model. The motion model is constrained spatially using eigenvectors generated by a principal component analysis (PCA) of patient motion data, and is regularized in time using parametric functions of a patient breathing surrogate recorded simultaneously with 4DCBCT acquisition. The parametric motion model is constrained using forward iterative projection matching (FIPM), a scheme which iteratively alters model parameters until digitally reconstructed radiographs (DRRs) cast through the deforming CT optimally match the projections in the raw 4DCBCT data-set. We term our method FIPM-PCA 4DCBCT. In developing our algorithm we proceed through three stages of development. In the first, we establish the mathematical groundwork for the algorithm and perform proof of concept testing on simulated data. In the second, we tune the algorithm for real world use; specifically we improve our DRR algorithm to achieve maximal realism by incorporating physical principles of image formation combined with empirical measurements of system properties. In the third stage we test our algorithm on actual patient data and evaluate its performance against gold standard and ground truth data-sets. In this phase we use our method to track the motion of an implanted fiducial marker and observe agreement with our gold standard data that is typically within a millimeter.

1 Introduction

1.1 Purpose

Radiotherapy of moving targets, in particular tumors of the lungs, presents two complications over radiotherapy of static targets. The first issue is that computed tomography (CT) images acquired for planning purposes are often plagued by motion artifacts. This can make accurate delineation of the tumor volume difficult, which can lead to a plan that treats an inaccurately shaped and/or positioned target volume. The second complication is that motion compensation must be included in the treatment plan. Typically this is done by extending a margin, termed the planning treatment volume (PTV), around the delineated tumor volume that will be treated as if it too is a part of the tumor. Thus the tumor can move around during treatment and still receive full dose coverage as long as it stays within the margin. This approach comes at the expense of exposing normal tissue to high radiation doses, which can lead to unpleasant and dangerous complications and also increase the risk of secondary cancers. These problems can be reduced by explicitly including the changes of anatomy over time in the imaging, planning, and treatment processes. Collectively these strategies are known as 4D radiotherapy.[1] Each phase of 4D radiotherapy depends on the quality of the previous phase. Success at the planning stage depends on the quality of the 4D image set used as input, and treatment can only be as accurate as the plan it implements.

4D imaging for the purposes of radiotherapy planning and treatment is generally accomplished via time correlated CT imaging, otherwise known as 4DCT imaging. Instead of the single 3D image produced by regular CT imaging, 4DCT generates a set of 3D images that each correspond to a specific anatomic breathing pose of the object. Within 4DCT there exist two sub-modalities, 4D fan-beam CT (4DFBCT) and 4D cone-beam CT (4DCBCT).

Current 4D planning approaches can generally be divided into three groups: plans that incorporate 4D anatomic information to reduce the PTV, plans that incorporate 4D information directly into the optimization process, and plans that require real-time adaptation of the treatment beam and/or patient position to target the tumor at its current position. Harsolia et al. compared several different 4D based planning techniques against each other and a non-4D technique.[2] In one of the 4D techniques the gross tumor volume (GTV) was constructed as the union of the tumor volume outlined on six different phases of a 4DCT, and an isotropic PTV margin was added solely to account for setup error. Another technique convolved a probability density function (PDF) characterizing the probability of the patient's tumor being at a certain location, determined from fluoroscopy, with a static dose distribution planned on a phase of the 4DCT at which the tumor was at approximately its mean position. The dose to the clinical target volume (CTV) from this moving dose distribution was calculated, and the beam aperture was enlarged until adequate coverage was achieved. In comparison with the 3D technique both of the 4D techniques provided adequate coverage of the CTV using a smaller PTV, and thus had better normal tissue sparing, with the PDF technique performing the best.[2] Dietrich et al. also describe several approaches to 4D treatment planning.[3] The first convolves a pencil beam kernel with a probability density function characterizing the 3D tumor trajectory, resulting in a smeared out dose distribution that is less sensitive to tumor motion. Another is the class of inverse planning techniques that incorporate real time tumor-tracking to minimize dose to organs at risk while maintaining prescription dose to the tumor. Finally they cite as most common the co-registration of dose distributions calculated on individual CTs from a 4D set in order to accumulate the time weighted distributions on a reference CT. All of these techniques either require or greatly benefit from a 4DCT used as input to the planning process. Thus they depend on the accuracy of the planning 4DCT for their success.

Dietrich lists several drawbacks to 4DCT which might hamper 4D radiotherapy, including errors due to data sorting, image registration, and image artifacts, as well as a very limited time sampling of the patient's anatomy.[3] Yamamoto et al. investigated the type, frequency, and magnitude of image artifacts in phase sorted 4DCTs of 50 cancer patients. They found that most patient 4DCTs had at least one artifact and that the median number was two.[4] Xing et al. outlined several outstanding issues facing 4DCT as a means for accurate treatment planning. They cite irregular patient breathing as a source of binning complication and subsequent image artifacts. In addition, they note that each piece of internal anatomy may have its own phase and amplitude relationship with the recorded external fiducial. Also, they point out that smoothly interpolating between points in the respiratory cycle can miss non-rigid organ deformations. [5]

From this picture we can describe the characteristics of an ideal 4DCT reconstruction algorithm for the purposes of the purposes of 4D radiotherapy:

- High image resolution and quality.
- Minimal image artifact.
- High temporal resolution that accurately depicts the deformation undergone by the moving patient anatomy.

The goal of this work is to describe a 4DCT algorithm that exceeds methods found in the literature with regards to all of these parameters for the purpose of improving radiotherapy of lung tumors.

1.2 Current approaches

In general, 4DCT reconstruction requires an oversampled CT data-set acquired concurrently with a surrogate measure of patient breathing motion. The motion surrogate can be measured from the CT data-set itself, e.g. 1D diaphragm position measured from chest CBCT projections, or from a separate device, e.g. the Real-time Position Management™ (RPM) system (Varian Medical Systems, Palo Alto, CA). Current approaches to 4DCT reconstruction can be divided into two categories: data binning and motion model based algorithms. The defining characteristic of data binning approaches is the sorting of raw CT into bins according to the breathing phase or amplitude at which they were acquired. The bin into which a particular piece of CT data is sorted is determined by the value of the motion surrogate at the time at which it was acquired. It should be noted that the data being sorted is different for 4DFBCT and

4DCBCT. For 4DFBCT reconstructed 3D axial slices captured at different couch positions are sorted, while for 4DCBCT raw 2D projections taken at a single couch position are sorted. From this point each phase bin is reconstructed as if it were a normal 3DCT. For FBCT the reconstructed slices are simply stitched together, while for CBCT the raw data are reconstructed using an image reconstruction algorithm, the most common of which is filtered backprojection (FBP). The result is a series of images covering one time-averaged cycle of patient breathing, typically consisting of ~10 images. The reconstructed sequence is time averaged in the sense that data from multiple breathing cycles are combined to form each individual image. The class of motion model based algorithms attempts to sidestep the requirement of data binning. These algorithms seek to construct an un-averaged portrait of patient anatomy during breathing, typically with a temporal resolution equivalent to that of the raw 4DCT data. Much work has been done with regards to both reconstruction approaches, and the following two sections explore some of these methods.

1.2.1 Binning based 4DCT reconstruction

One of the first 4DCT protocols was presented by Vedam et al. as early as 2003. The procedure required an oversampled spiral FBCT which it retrospectively sorted into phase bins based on a respiratory breathing trace acquired with an external marker.[6] Also in 2003 Taguchi published a paper comparing the performance of candidate backprojection algorithms for the reconstruction of 4DCBCTs.[7] Since then many groups have sought to improve various aspects of the binning based 4DCT reconstruction process.

One topic of research has been the binning of CT data according to motion surrogate phase, amplitude, or both. Most investigators seem to have come to the consensus that amplitude based binning produces CTs with less motion artifact than phase-based techniques, though it can aggravate bin undersampling issues.[8-10] Langner et al. investigated a prospective sorting technique that binned images based on both displacement and velocity criteria.[11] Their objective was to reduce patient dose by avoiding acquisition of redundant data as well as to reduce residual CT motion artifact by using their more stringent binning requirements. Compared to retrospective displacement sorting they found that the spread in displacements within the bins was at least as good using the prospective method and that patient dose was reduced by 20-50%. However prospective binning came at the price of increased acquisition time.[11]

Another variable is the method used to acquire the breathing motion surrogate. Fitzpatrick et al. used optical tracking of an external chest marker (RPM system).[8] Lu et al. used the combination of a spirometer and pressure sensitive belt wrapped around the chest (pneumo-bellows).[10] Sonke et al. determined the position of the diaphragm directly from the projections in their 4DCBCT data-set to form their respiratory trace.[12]

Other work has examined the tradeoff between number of projections and dose per projection. In a direct attempt to eliminate undersampling Li et al. investigated the effect of reducing mAs in order to acquire a greater number of projections with the same patient dose. They found that this approach helped to alleviate view aliasing while maintaining a satisfactory signal-to-noise ratio (SNR).[13]

Conventional binning based 4DCBCT methods typically reconstruct individual bins via filtered backprojection.[6, 8, 12-14] With FBP image quality is often impaired by streak artifacts resulting from projection undersampling due to binning.[15-18] To assess the potential of more advanced methods to improve image quality Bergner et al. investigated five 4DCBCT reconstruction algorithms and analyzed the results based on image noise, streak artifacts, temporal resolution, and computational effort. They examined two filtered backprojection algorithms: conventional phase correlated (PC) reconstruction in which projections are binned according to phase and backprojected, and autoadaptive phase correlation (AAPC) reconstruction in which motion unaffected areas of each projection are backprojected in every bin regardless of that projection's phase. They examined two forward iterative projection approaches, prior image constraint compressed sensing (PICCS) and adaptive steepest descent projection onto complex sets (ASD-POCS), which seek to minimize both the integral of the absolute value of the gradient of an image as well as the mismatch between forward projections taken through the reconstructed images and raw data images of the selected phase. Additionally, they looked at a hybrid approach, the McKinnon-Bates algorithm (MKB), combining a deterministic backprojection reconstruction with a single forward projection step. Their results failed to identify any clear cut best choice. The backprojection algorithms were quicker and produced more dependable results, while the iterative solutions took far longer and were dependent on initial data and starting points but produced lower noise and artifact images. They mentioned that the MKB algorithm fell between the two groups and might be an acceptable compromise. [19]

Other investigators have observed that the information contained in different phases of a 4DCT is highly correlated, which can be exploited to increase image quality and/or reduce patient dose. In the simplest example Schreibmann et al., whose goal was to reduce patient dose and equipment workload, only acquired images at the end inspiration and expiration phases of breathing. They used deformable image registration (DIR) to produce a displacement vector field (DVF) linking the two images, and then interpolated the intermediate phase images by deforming one of the starting CTs with successive fractions of this DVF. They compared their interpolated results to a full 4DCT and determined the two sets to match reasonably well.[20] Li et al. assessed several ways of enhancing 4DCBCT to address undersampling concerns. Their strategy was to bin projections into possibly undersampled phase bins and then backproject. These view-aliased CT volumes were then all deformably registered to a desired phase, and the deformed CTs were added together to form a non-aliased image at that particular phase. Starting from the reconstructed 4DCBCT with view aliasing they investigated three different ways of achieving this result, the first of which was to simply deformably register the view aliased phases to a particular phase and then add the deformed volumes to that phase. The second approach consisted of registering each phase to the artifact free planning CT acquired prior to the 4DCBCT. With this approach only one of the volumes in the registration had view aliasing artifacts, and thus it was expected that the deformable registrations would be more accurate. The third and most novel approach involved deformably registering the planning CT to unreconstructed bins of projections via 2D to 3D registration, and then using the displacement vector fields (DVFs) from this process to deform the reconstructed phases to a desired phase. Its advantage over the first two methods lay in the fact that it did not require the registration of artifact affected images. Their results showed that the third method produced the most artifact free reconstructions while the first two methods performed approximately equally.[15] Gau et al. investigated an iterative reconstruction seeking to minimize a data fidelity term and a regularization term penalizing large intensity gradients. The regularization term operated in both the spatial and temporal domains and thus was able to combine information from all phases to improve the reconstruction of individual phases.[21]

1.2.2 Motion model based 4DCT reconstruction

The above methods differ in many respects, but they all require the binning of image data in order to reconstruct a single time-averaged breathing cycle. A second class of algorithm attempts to circumvent this limitation using modeling of patient breathing motion. For 4DFBCT there already exist an un-averaged time series of reconstructed 3D images at each couch position. Standard 3D to 3D deformable image registration can be used to create a motion model tied to a patient breathing trace for each couch position, and then the motion models for all couch positions can be combined. McClelland et al. designed a 4DFBCT reconstruction algorithm that created separate 4D motion models for each slab or couch position via deformable registration between phases within a slab. B-splines were used to register each individual time series of slabs, and the resulting time series of control points were fit to a respiratory trace to model their temporal evolution. Finally, the motion models for separate slabs were combined to form a motion model covering the entire imaging volume. In this manner the motion model was able to propagate the FBCT to unique poses based on the patient respiratory trace.[22] A similar approach was employed by Yang et al. who used their ‘5D motion model’ to fit the motion described by deformable registration of 4DFBCT slabs.[23] Zhang et al. reconstructed and registered the phases of a 4DFBCT and performed principal component analysis (PCA) on the resultant DVFs to linearly correlate the DVF motion with a small number of variables on a voxel by voxel scale. They chose the two eigenmodes with the largest variance as their variables and modeled the temporal evolution of the two eigenmodes using the patient breathing trace and the patient breathing trace time-shifted by three data points.[24]

Similar motion model based reconstruction of 4DCBCT images via motion modeling is more challenging since the relevant data are raw 2D projection images rather than reconstructed 3D slices. However, the process is simpler in that there is only a single couch position to consider due to the large area of CBCT detectors. Motion model based 4DCBCT reconstruction methods combine three elements: a parametric time dependent deformation model, a static CT image, and an optimization framework based on minimizing the difference between raw CBCT projections and forward projections of the current image sequence. A single pass through the algorithm proceeds by animating the static CT with the deformation model to create a sequence of images, then uses the known CBCT system geometry to simulate forward projections, or digitally reconstructed radiographs (DRRs), cast through the animated CT image sequence, and then computes the mismatch between the DRRs and the measured CBCT projection data. The deformation model is iteratively updated until designated stopping criteria, such as data mismatch falling below a threshold, are met. We term these algorithms time dependent forward iterative projection matching (FIPM) approaches. Several investigators have explored this method. Zeng et al. used such a projection matching approach to constrain a deformation motion model consisting of the tensor product of 3 spatial and 1 temporal B-splines.[25] Docef et al. used a model consisting of the tensor product of 3 spatial B-splines to represent spatial deformation and modulated the control points in time using a parameterized function of the patient breathing trace.[26] Li et al. reconstructed a 4DFBCT via backprojection, deformably registered the sequential phases to a reference phase to obtain DVFs, and then calculated the eigenmodes of the DVFs to form a PCA model similar to that of Zhang et al.[24, 27, 28] Rather than build an explicit time dependence into their deformation model they optimized the model separately for each individual CBCT projection.[27] Their principal aim was to reconstruct the trajectory of a lung tumor. More recently they have explored the mathematical underpinnings of the PCA model and compared it to the 5D lung motion model of Low et al.[29, 30]

The principal challenge in this type of approach is achieving a physically realistic deformation model with a tractable number of parameters that is easily generalized to any imaging subject. The 4D DVF B-spline model of Zeng et al., while completely general, would require many thousands of control points to provide a reasonable time resolution for the duration of a typical CBCT exam. [25] The method of Docef et al. greatly reduces the number of control point parameters but requires assumptions about the relative phase and amplitude of the individual displacement vectors. [26] The method of Li et al. requires few parameters and makes no assumptions, but it does not explicitly model time dependence, instead requiring re-optimization of the model for each time point. [27]

1.3 Our approach

Our algorithm follows the time dependent FIPM approach. Its salient feature is the deformation model employed, which is based on the principal motion eigenmodes of breathing, obtained via PCA. The principal components are trained on a conventionally reconstructed 4DCBCT of the patient, whereby they “learn” the relative amplitudes and phases of motion among the moving image elements without recourse to a specific physical model. Specifically they are trained on DVFs linking a high quality static prior CT of the patient to phases of the conventionally reconstructed 4DCBCT. Such a static CT might be a breath hold FBCT or an artifact free phase of a 4DFBCT captured of the patient previously. The time dependence of the model is accounted for in the principal coefficients, which are modeled as parametric functions of the patient breathing trace. We term our method FIPM-PCA 4DCBCT.

With regards to other time dependent FIPM reconstruction algorithms our approach has several key advantages. Compared to the 4D and 3D B-spline models of Zeng et al. and Docef et al. our deformation model has far fewer free parameters, making its optimization both more robust and potentially fast enough for clinical implementation. As compared to Li et al. we fit the motion model to the entire set of CBCT projection radiographs simultaneously, which has the benefit of explicitly modeling the temporal behavior of the deformation. An additional difference between our implementation and that of Li et al. is our use of the 4DCBCT data set to generate the training DVFs as opposed to a previously acquired 4DFBCT. Though we do not test for this specifically we expect, and they comment, that a motion model based on the most current training DVFs, i.e. those obtained from the raw 4DCBCT data-set itself, will be more accurate than a model based on training DVFs from a 4DFBCT that might be days or weeks old.[28]

The advantage of our 4DCBCT algorithm over motion model based 4DFBCT lies in the advent of linac mounted, or “onboard”, CBCT imaging devices within radiotherapy clinics. These systems allow for online imaging of the patient on the same treatment couch and in the exact pose in which radiotherapy will be delivered. As only CBCT systems are suitable for such integration CBCT has an intrinsic advantage that FBCT cannot match. Workarounds such as couches that translate between a FBCT and linac are costly and may still induce changes in the patient pose between imaging and treatment delivery. We also take note of the TomoTherapy treatment system (Accuray, Sunnyvale, CA), which is essentially a mega-voltage helical CT device that can be used to image the patient on the same couch and in the same pose in which treatment will be performed. However, the soft tissue contrast in mega-voltage CT images is inferior to that obtained using kilo-voltage CBCT systems.

Compared to binning based 4DCT algorithms the benefits of our algorithm are numerous. By deforming a static high quality FBCT to produce our reconstructed image sequence our method preserves the high image quality of the initial image. Binning based 4DCBCT in particular suffers from streak artifacts due to undersampling that can severely degrade image quality, while binning based 4DFBCT can exhibit discontinuities at couch position boundaries. Furthermore, the non-uniform scatter collected across the detector surface makes CT numbers in FBP reconstructed CBCT unreliable. Since our reconstructed images are simply deformed versions of a superior quality FBCT, the high CT number fidelity of the reference FBCT is preserved in our reconstructed 4DCBCTs. In addition, our method provides a non-time averaged, high temporal resolution portrait of moving patient anatomy. Whereas binning based 4DCT reconstruction may provide ~10 images for an acquisition time covering a minute or more, our method can produce a reconstructed image at the rate of raw data acquisition, which may be on the order of 10 Hz or more. Finally, our method intrinsically produces DVFs linking reconstructed images to a reference phase. Such DVFs are often an integral part of 4D planning methods and are normally produced by deformably registering reconstructed images in a separate step. Our algorithm simplifies this workflow and may avoid errors associated with a separate DIR procedure.

Our time dependent FIPM 4DCBCT reconstruction algorithm relies on three key factors: the assumption that the correlation between our measured breathing motion surrogate and the motion of internal anatomy remains relatively constant during image acquisition (as noted in Section 1.2 either internal or external surrogates may be used), the quality of the DVFs used to train our PCA deformation model, and the accuracy with which our DRR module simulates the geometry and physics underlying the formation of our raw projection data. A large portion of our work explorers the effect played by these factors.

1.4 Layout of this manuscript

The rest of this manuscript is laid out in four chapters. Chapter 2 details the specifics of our algorithm and describes its proof of concept testing using simulated data. Chapter 3 describes adjustments made to the algorithm to account for real world considerations. Specifically we discuss the optimizations made to our DRR algorithm to generate physically realistic forward projections. Chapter 4 describes the application of our algorithm to actual patient data. Each of these chapters begins with an introduction that describes the problem to be addressed, followed by a description of the methods used to solve the problem, and is concluded with results and discussion. Chapter 5 serves as a reflection on the process as a whole, including a summary of our progress and thoughts on future directions.

2 Algorithm proof of concept

2.1 Introduction

This chapter begins with a detailed description of our FIPM-PCA 4DCBCT algorithm including required input data, mathematical framework, and output data. It proceeds by describing our method of numerically generating synthetic data-sets to enable testing with a known gold standard. We then list various aspects of the algorithm for which different approaches are tested and describe the matrix of different tests that these variations create. Next we present the results of testing and verification of our algorithm on our synthetic data-sets, and conclude with a discussion of the results.

2.2 Algorithm details

2.2.1 Mathematical framework

Our reconstructed 4DCBCT can be formulated as the solution to the following minimization problem:

$$\mathbf{X} = \arg\min_{\mathbf{X}} \left[\sum_{t=1}^{N_T} \left\| \mathbf{A}\mathbf{X}_t - \mathbf{Y}_t \right\|_2^2 \right] \quad \textbf{(1)}$$

where $\mathbf{X}_t$ is a reconstructed 3D image of the patient anatomy at time *t*, $\mathbf{Y}_t$ is a 2D projection from the raw 4DCBCT data-set captured at time *t*, $\mathbf{A}$ is the forward projection operator, N_T is the total number of time points at which projections were captured in the raw 4DCBCT data-set, and $\mathbf{X} = \left\{ \mathbf{X}_t \right\}_{t=1:N_T}$ is the complete set of N_T reconstructed 3D images, i.e. the reconstructed 4DCBCT. We further define $\mathbf{X}_t$ as the following:

$$\mathbf{X}_t\left(\vec{r}\right) = \mathbf{X}_0\left(\vec{r} + \mathbf{u}_t\left(\vec{r}\right)\right) \quad \textbf{(2)}$$

where $\mathbf{X}_t\left(\vec{r}\right)$ is the intensity value of image $\mathbf{X}_t$ at location $\vec{r}$, $\mathbf{X}_0$ is a static prior 3D image of the patient anatomy (the reference image), and $\mathbf{u}_t\left(\vec{r}\right)$ is the value of time dependent DVF deformation model $\mathbf{u} = \left\{ \mathbf{u}_t \right\}_{t=1:N_T}$ at time *t* and location $\vec{r}$ (a 3 component vector). Thus $\mathbf{X}_t$ is a deformed version of static image $\mathbf{X}_0$. In general, the smaller the time gap between the acquisition of the reference image $\mathbf{X}_0$ and the CBCT projection data-set $\mathbf{Y}$ the better the expected reconstruction quality. Gross anatomic changes, such as the shrinkage of a tumor or the collapse or reinflation of a portion of the lung, occurring in the interval between FBCT and CBCT acquisition will be difficult for our algorithm to recover. Minimizing this acquisition interval to ideally a day or less will minimize the risk of such changes occurring. $\mathbf{u}_t$ can be further expanded as:

$$\mathbf{u}_t = \mathbf{u}_{avg} + \sum_{i=1}^{i=M} w_i(t)\mathbf{e}_i \tag{3}$$

where $\mathbf{u}_{avg}$ is a mean DVF, $\mathbf{e}_i$ is an orthonormal basis vector of the same dimensionality as $\mathbf{u}_t$, $w_i(t)$ is a scalar time dependent weighting coefficient, and M is the total number of basis vectors. Finally, $w_i(t)$ can be expanded as:

$$w_i(t) = A_i b(t+\tau_i) - C_i \tag{4}$$

where $b(t)$ is a time dependent patient breathing motion surrogate and A_i, τ_i, and C_i are free parameters that modify the amplitude, time offset, and amplitude offset of the breathing motion surrogate respectively. For convenience we define a generic free parameter vector that stores all $3*M$ free parameters, $\mathbf{p} = \{\mathbf{p}_i\}_{i=1:M} = \{A_i, \tau_i, C_i\}_{i=1:M}$.

2.2.2 Algorithm workflow

The goal of our algorithm is to find the set of A_i, τ_i, and C_i such that Eqn. (1) is satisfied. This goal is achieved iteratively. As an example, the workflow for iteration *j* of the algorithm is outlined below:

1. Plug the free parameter values, $\mathbf{p}^j = \{\mathbf{p}_i\}_{i=1:M}^j = \{A_i, \tau_i, C_i\}_{i=1:M}^j$, into the DVF deformation model to get the current guess at the time dependent deformation model, $\mathbf{u}^j = \{\mathbf{u}_t\}_{t=1:N_T}^j$.
2. Deform the static prior 3D image, $\mathbf{X}_0$, with the deformation model to generate an initial guess at the reconstructed 4DCBCT, $\mathbf{X}^j = \{\mathbf{X}_t\}_{t=1:N_T}^j$.
3. Compute the value of the objective function, formulated as the sum of squared differences (SSD) between forward projections cast through reconstructed 4DCBCT and the known projections of the raw 4DCBCT data-set, $SSD^j = \sum_{t=1}^{N_T} \left\| \mathbf{AX}_t^j - \mathbf{Y}_t \right\|_2^2$.
4. If SSD^j falls below a threshold ε terminate the algorithm. If not, proceed to step 5.
5. Use the optimizer to generate a new set of free parameters, $\mathbf{p}^{j+1} = \{\mathbf{p}_i\}_{i=1:M}^{j+1} = \{A_i, \tau_i, C_i\}_{i=1:M}^{j+1}$.
6. Advance the iteration number $j = j+1$ and return to step 1.

For the first iteration this workflow is preceded by an additional step 0:

0. Generate an initial guess at the free parameters of the DVF deformation model, $\mathbf{p}^0 = \{\mathbf{p}_i\}_{i=1:M}^0 = \{A_i, \tau_i, C_i\}_{i=1:M}^0$.

This workflow is presented diagrammatically in Figure 1.

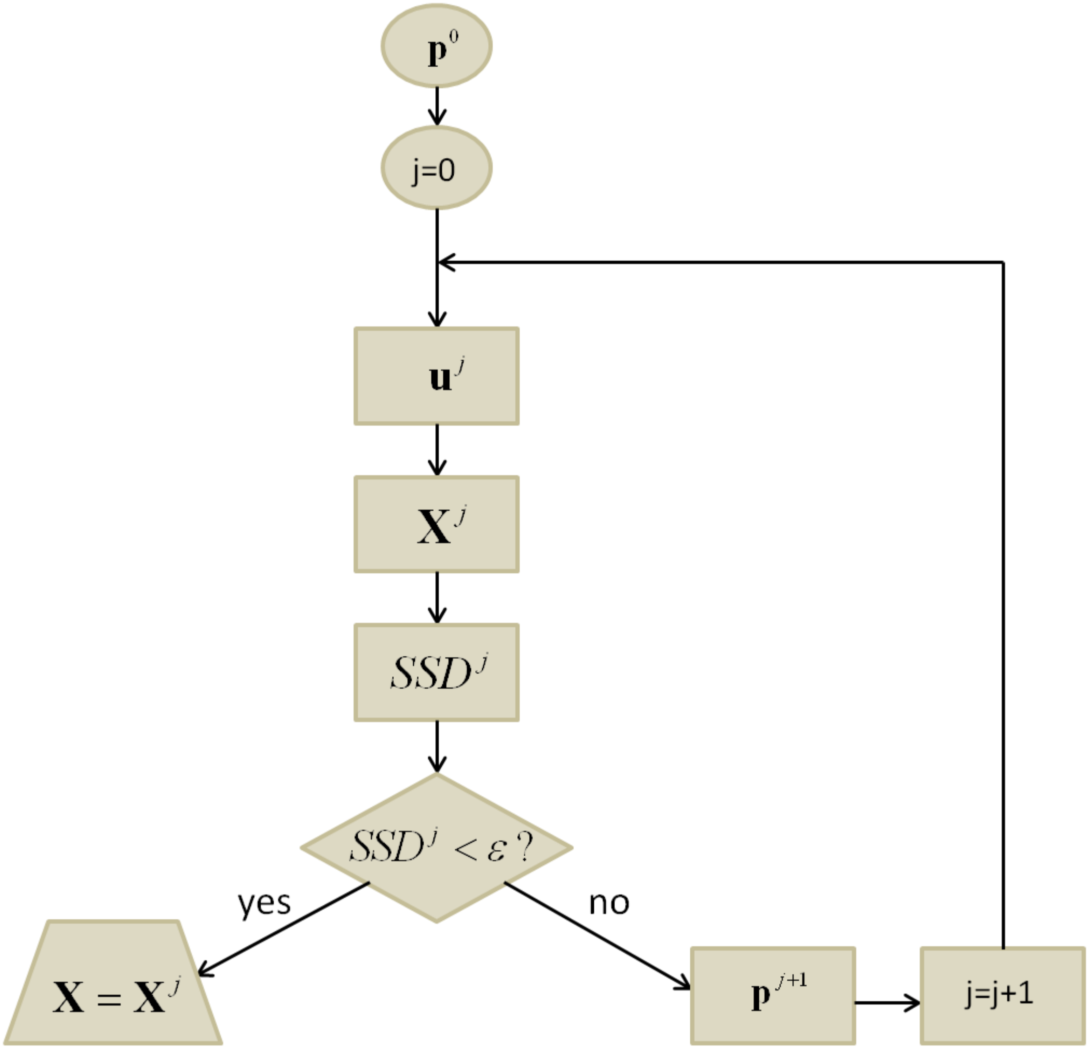


Figure 1: 4DCBCT reconstruction algorithm workflow diagram.

2.2.3 DVF deformation model

Our deformation model is based on the fundamental assumption that the time-changing displacements of all the anatomical elements of the patient can be completely described by their relative amplitudes and phases, and that these relative amplitude and phase relationships remain approximately constant for the duration of the CBCT scan. Under this assumption, we expect that the motion of each voxel will track with the motion of a single reference point, to within their relative amplitude and phase shift factors. If the motion of the reference point is represented by a measured breathing trace $b(t)$, then we assume that the motion of the other voxels will track with $b(t)$. This forms the basis for our breathing model.

As described in Eqn. (3) our deformation model is constructed as a weighted sum of orthonormal basis vectors. The most desirable basis is one in which the variance of the motion data is maximally concentrated along the first few basis vectors and becomes zero or negligible for all other basis vectors. In addition, we desire that the motion data vary independently along each basis vector in relation to all other basis vectors. Then we must only determine weights for the limited set of important basis vectors and need not worry about any interdependence between them. Such a basis is represented by the principal components of the breathing motion. Although the principal component basis vectors are not necessarily stochastically independent they are maximally de-correlated, which serves as a reasonable approximation of independence. We determine this optimal basis by performing PCA on a training set comprised of a temporal sequence of DVFs derived from the breathing motion of the imaging subject.

The training set must be of sufficient quality to characterize the phase and amplitude relationships present between voxels during patient breathing. Unlike Li et al. we produce our training set directly from the CBCT data to be reconstructed. [27] Methods 3 and 4 in Section 2.4.1 detail the two realistic approaches that we test to generate such a group of DVFs. For example, in method 3 from Section 2.4.1 we first group our raw CBCT projections into bins based on breathing phase using the breathing motion surrogate signal. We then perform a 2D/3D registration x between each bin of projections and the reference FBCT using a novel DIR algorithm developed within our group. [31] Each registration produces one DVF for the training set. Once the training set is acquired its covariance matrix is calculated and the eigenvectors of the covariance matrix are determined using PCA. These eigenvectors, or principal components, are the orthonormal basis vectors $\mathbf{e}_i$ in Eqn. (3). They capture the relative amplitude and temporal phase relationships of all of the individual displacement vectors. As in Eqn. (4) the weights $w_i(t)$, or principal coefficients, are modeled in time using a patient breathing motion surrogate or, breathing trace, $b(t)$ as a template so that, per our assumption above, the displacement vectors will approximately track with the observed movement of the breathing reference point. The idea of creating a complete spatial motion model by driving a set of principal component vectors by the movement of a single reference point has been independently investigated by Li et al, who also find that it is a reasonable approach.[30]

We relate each principal coefficient to the breathing trace through three adjustable parameters: A_i, τ_i, and C_i that modify the amplitude, time offset, and amplitude offset of the breathing trace respectively. The rationale for this simple temporal parameterization of the principal coefficients (and our underlying assumption about the relative movement of voxels) is found in the behavior of actual 4DCT principal coefficients, as shown for example in Vaman et al., where the first two principal coefficients track the breathing cycle to within an amplitude, phase, and offset factor as in Eqn. (4).[32] This parameterization has been observed to provide a reasonable representation of the temporal behavior of principal coefficients derived directly from the principal component analysis of a set of empirical breathing DVFs, as discussed below.

The model utilizes M different eigenvector-principal coefficient pairs. In practice it has been found that the first two or three PCA eigenvectors are usually sufficient to accurately describe the breathing motion in the training set.[24, 27, 28, 32] In the tests of our fitting algorithm a motion model utilizing 5, 7, or 10 eigenvectors produced no improvement in fitting accuracy over the 3 eigenvector model.

2.3 Generating a synthetic 4DCBCT data-set

Rather than use actual CBCT data for our proof of concept testing, we used numerically simulated raw 4DCBCT data-sets. This afforded us control over the motion complexity of the moving 3D object underlying the data and provided us a gold standard with which to evaluate our results at every step of the reconstruction process. A numerical 4DCBCT data-set was produced as follows:

1. We constructed a numerical CT phantom, also referred to as the reference CT, in the form of a cubic volume cropped from a patient pelvis CT.
2. An initial DVF was created with 3 spatial components, with the amplitudes varying in space according to a 3 dimensional Gaussian function with its maximum at the image center.
3. A time series of DVFs was produced by scaling the vectors in the initial DVF by the amplitude of a breathing trace sampled at successive time points.
4. The time series of DVFs was applied to the reference CT (our static prior image), producing a time series of deformed CTs.
5. DRRs at different simulated gantry angles were cast through the time series of deformed CTs to produce a rotating set of projections taken through moving anatomy. This became our simulated 4DCBCT data-set. In our CBCT geometry the CTs were

composed of a 64x64x64 matrix of 1x1x1 mm voxels, and a flat panel detector was simulated as a 256x256 matrix of 1x1 mm pixels.

Generating our 4DCBCT in this manner not only produced the requisite set of rotating projections of our moving object, but also gave us the true CTs underlying each projection and the true DVFs linking those CTs to the reference CT. These are used in our analysis of the factors contributing to reconstruction accuracy.

2.4 Algorithm test matrix

For the purposes of testing and validating our algorithm we performed 4DCBCT reconstructions while varying three different environment variables. The different unique combinations of environment variables, 48 in all, yield different settings in which to test our algorithm, ranging from most ideal to most realistic. The three environment variables and the different values which they can take are described in the following sections.

2.4.1 DVF training set

The DVF training set serves as the input to the PCA that produces the basis of eigenvectors, which is the heart of the deformation model. Each DVF in the training set serves as a snapshot of the phase and amplitude relationships between voxel displacements at a particular time. For the motion model to perform well it is important that the training DVFs be reasonably free of artifacts and errors and sufficiently numerous to capture the essential features of the patient’s breathing. We investigated four different methods of producing the DVF training set, two ideal methods to serve as benchmarks, and two realistic methods. The ideal methods take advantage of the fact that our numerically simulated data-sets contain the known CTs underlying each projection image and the known DVFs linking those CTs back to the reference CT. The realistic methods utilize only the projection data, as would be possible in real life.

1. **Method 1:** Populate the training set with 16 DVFs sampled directly from the full time series of known DVFs.
   - **a.** The sampled DVFs are guaranteed to be artifact free.
   - **b.** This is our most ideal method; it produces the gold standard DVF training set.
2. **Method 2:** Deformably register the reference CT to 16 deformed CTs sampled from the full time series of known deformed CTs to produce 16 DVFs.
   - **a.** The sampled deformed CTs are guaranteed to be artifact free.
   - **b.** The DVFs may contain artifacts from the 3D/3D DIR process.
   - **c.** This method is an idealization but is expected to produce a lower quality training set than method 1.
3. **Method 3:** Cast 100 projections through each of 16 deformed CTs sampled from the full time series of known deformed CTs to simulate 16 phase bins of raw 4DCBCT data. Deformably register the reference CT to each bin of projection images.
   - **a.** The projection images within each bin are guaranteed to be artifact free.
   - **b.** The DVFs may contain artifacts from the 2D/3D registration process. The registrations are performed using a novel time independent FIPM DIR algorithm

developed within our group. It is termed FIPM-bspline registration due to its use of a bspline DVF model.[31]

   - **c.** This method simulates a procedure that will be possible in real life.

4. **Method 4:** Cast 100 projections through each of 16 deformed CTs sampled from the full time series of known deformed CTs to simulate 16 phase bins of raw 4DCBCT data. Use filtered backprojection to reconstruct each phase bin. Deformably register the reference CT to each reconstructed phase bin to produce the DVF training set.
   - **a.** The reconstructed phase bins may contain artifacts due to the FBP reconstruction.
   - **b.** The DVFs may contain errors due the 3D/3D registration process, which will be exacerbated by artifacts from the FBP reconstruction.
   - **c.** This method simulates a procedure that will be possible in real life.

It should be noted that we have implemented an idealized projection binning procedure for the FIPM and backprojection based methods. By doing this we sidestepped the complications of intra-bin motion that can be expected to occur in real life. Our goal was to test the two methods under identical conditions and discover the best possible performance we could expect from each when those conditions are favorable. Thus we were more interested in identifying which method offered the best performance relative to the other rather than precisely quantifying that performance.

For purposes of comparison Figure 2 displays an example slice of the reference CT, a known (i.e., ground truth) deformed CT phase, the same CT phase reconstructed by the 2D/3D registration algorithm outlined in method 3, and the CT phase reconstructed via conventional FBP 4DCBCT as in method 4.

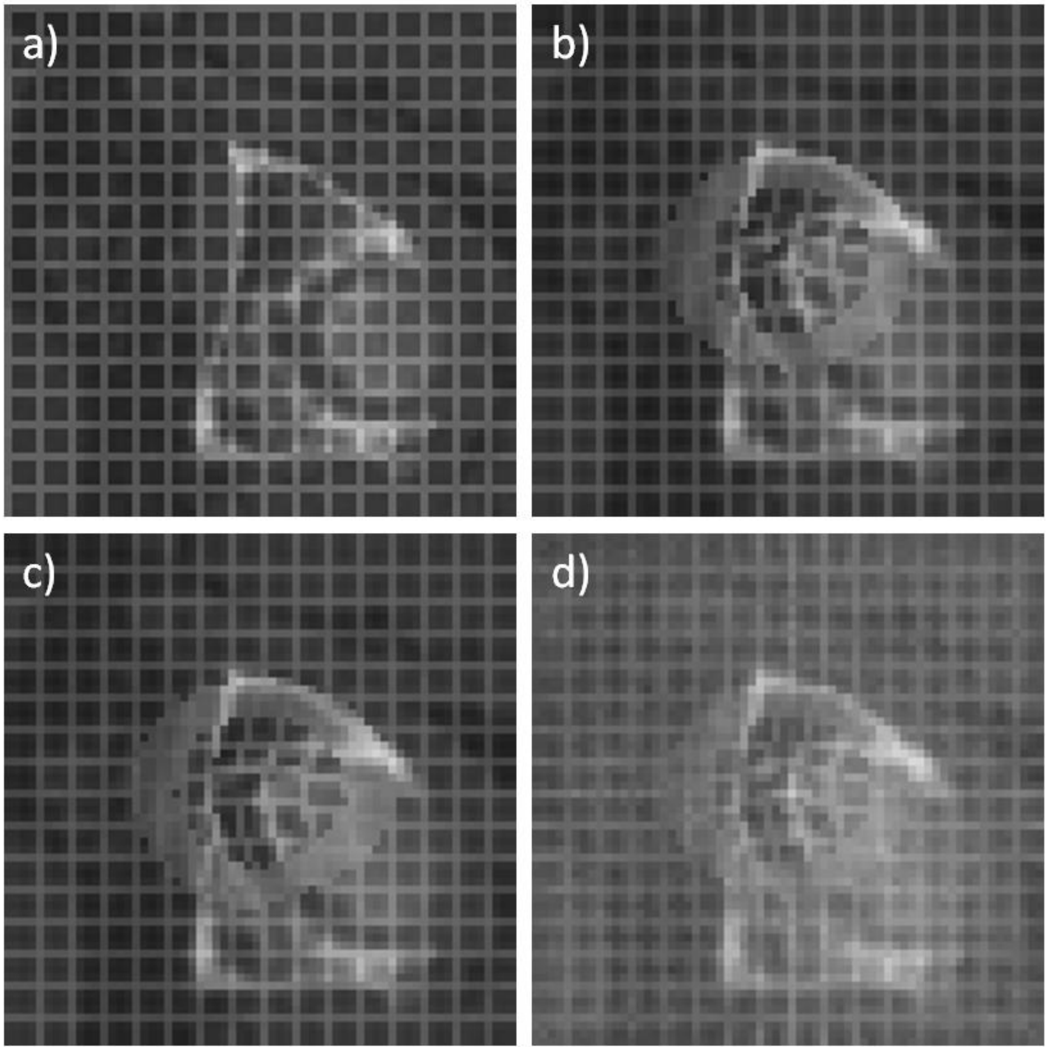


Figure 2: A slice through the source CT, b) a slice through a known deformed CT, c) a slice through the deformed CT reconstructed using the FIPM method, d) a slice through the same deformed CT reconstructed with filtered backprojection. The simulated motion type for this example uses a sine waveform with no phase shifts. The grid overlaying the images was added to enhance the visibility of the deformation.

2.4.2 Simulated breathing motion complexity

As described in Section 2.3 the time series of DVFs that determines our simulated motion was generated by scaling each vector component of an initial DVF with sequential amplitude values sampled from a breathing trace. We altered the complexity of this simulated motion in two different ways: by introducing phase shifts in the breathing trace used to scale different DVF vectors and by using breathing traces of varying complexity. For the phase shift case we tested two different scenarios: all DVF vectors moving in phase and vectors moving with three relative phase shifts. Phase shifts were distributed by scanning across rows of vectors and applying the different shifts in sequence, repeating the pattern every three vectors. For the case of breathing trace complexity we also tested two scenarios: a synthetic breathing trace sampled from a sinusoid and an actual breathing trace acquired from a patient. In combination this setup produced four different simulated motion complexities, one simple, two intermediate, and one complex. They are as follows:

1. No phase shifts, sinusoidal breathing trace; the simplest motion.
2. No phase shifts, actual patient breathing trace; intermediate complexity motion.
3. Three phase shifts, sinusoidal breathing trace; intermediate complexity motion.
4. Three phase shifts, actual patient breathing trace; the most complex motion.

In real life patient breathing will likely span a wide spectrum of complexity. With these different tests our goal was to simulate different motion complexities in a highly controlled manner to identify their effect on the results of our reconstruction algorithm.

2.4.3 Fitting target

The fitting target is the known data to which the algorithm compares its output. In real life the known data will always be the raw 4DCBCT projections. However, our numerically generated data-sets also contained the true CTs and DVFs underlying the projections. Therefore we could modify our algorithm to test against either of these two fitting targets to determine the effect on reconstruction quality. For the three possible fitting targets the objective function from step 3 in the algorithm workflow from section 2.2.2 is formulated as:

1. **DVF fitting target:** $SSD^j = \sum_{t=1}^{N_T} \left\| \mathbf{u}_t^j - \mathbf{u}_t^{known} \right\|_2^2$, where $\mathbf{u}^{known}$ is the known set of DVFs generated as part of the simulated 4DCBCT data-set.
2. **CT fitting target:** $SSD^j = \sum_{t=1}^{N_T} \left\| \mathbf{X}_t^j - \mathbf{X}_t^{known} \right\|_2^2$, where $\mathbf{X}^{known}$ is the known set of CTs generated as part of the simulated 4DCBCT data-set.
3. **Projection fitting target:** $SSD^j = \sum_{t=1}^{N_T} \left\| \mathbf{A}\mathbf{X}_t^j - \mathbf{Y}_t \right\|_2^2$, identical to section 2.2.2.

The DVFs are a direct representation of our deformation model, therefore the set of known DVFs was expected to provide the most sensitive fitting target to optimize the model with. The CTs are further removed from the deformation model, so their use as the fitting target was expected to produce worse results. The projections are yet another step removed from the deformation model. In addition, they compress the 3D data of the CTs into a 2D form. Therefore we expected them to provide the least sensitive fitting target. In practice, the effect of using successively less sensitive fitting targets is to challenge to optimizer to greater and greater degrees in finding the true global minimum, which represents the most correct deformation model. By performing these tests we were able to separate these optimization related effects from those introduced by the previously described tests in sections 2.4.2 and 2.4.1.

In addition, this test suite enabled the evaluation of several different optimization routines. We tested the Levenberg-Marquardt, Polak-Ribiere conjugate gradient, BFGS, and Nelder-Mead downhill simplex algorithms. All optimization algorithms were implemented using the GNU Scientific Library (GSL), a C library containing various routines for scientific computing.[33] Of these options, the Nelder-Mead simplex algorithm was the only one to provide robust convergence for the projection fitting target, therefore it was used for all reported results.

2.5 Results

The product of our reconstruction algorithm is two sets of data, a time series of CTs and a time series of DVFs linking those CTs to the reference CT. From our numerically generated data-sets we knew both the ground truth CT time series and the ground truth DVF time series. Therefore we could evaluate the results of our algorithm using two metrics, the DVF root mean square error (RMSE) and the CT RMSE. The DVF and CT RMSEs were formulated as

$$RMSE_{DVF} = \sqrt{\frac{1}{N_T N_X N_Y N_Z N_V} \sum_{t=1}^{N_T} \sum_{x=1}^{N_X} \sum_{y=1}^{N_Y} \sum_{z=1}^{N_Z} \sum_{v=1}^{N_V} \left(\mathbf{u}_{txyzv} - \mathbf{u}_{txyzv}^{known} \right)^2} \tag{5}$$

and

$$RMSE_{CT} = \sqrt{\frac{1}{N_T N_X N_Y N_Z} \sum_{t=1}^{N_T} \sum_{x=1}^{N_X} \sum_{y=1}^{N_Y} \sum_{z=1}^{N_Z} \left(\mathbf{X}_{txyz} - \mathbf{X}_{txyz}^{known} \right)^2} \quad (6)$$

respectively, where N_X is the number of voxels along the CT X (left-right) direction, N_Y is the number of voxels along the CT Y (anterior-posterior) direction, N_Z is the number of voxels along the CT Z (superior-inferior) direction, N_V is the number of directional components composing a single DVF vector (always 3), $\mathbf{u}_{txyzv}$ is the value in mm of the *v*th component of the DVF vector located at position x*,y,z* at time *t*, and $\mathbf{X}_{txyz}$ is the value in CT number intensity units (i.u.) of the reconstructed 4DCBCT at position *x,y,z* at time *t*. Vector components in the known DVFs ranged from about 0 to 6 mm (CT voxels are 1x1x1 mm), while CT number intensity units in the known CT time series ranged from 0 to about 3000 i.u. For our testing on simulated data N_T=15 and N_X= N_Y= N_Z=64. Our test matrix included four different methods of generating the DVF training set, four different types of simulated motion, and three different fitting targets for a total of 4*4*3=48 unique test scenarios. We calculated the $RMSE_{\text{DVF}}$ and the $RMSE_{\text{CT}}$ for each of these cases and plotted the results. To display this data we took a particular combination of two environment variables and created a trend by holding them constant while the third variable was changed. An overlay of all trends for the changing variable composed a finished graph. Thus we were able to plot the 48 data-points for $RMSE_{\text{DVF}}$ and $RMSE_{\text{CT}}$ in three different ways to focus on each environment variables. In Figure 3 we varied the DVF training set generation method for different combinations of motion complexity and fitting target, in Figure 4 we varied the motion complexity for different combinations of DVF training set generation method and fitting target, and in Figure 5 we varied the fitting target for different combination of motion complexity and DVF training set generation method.

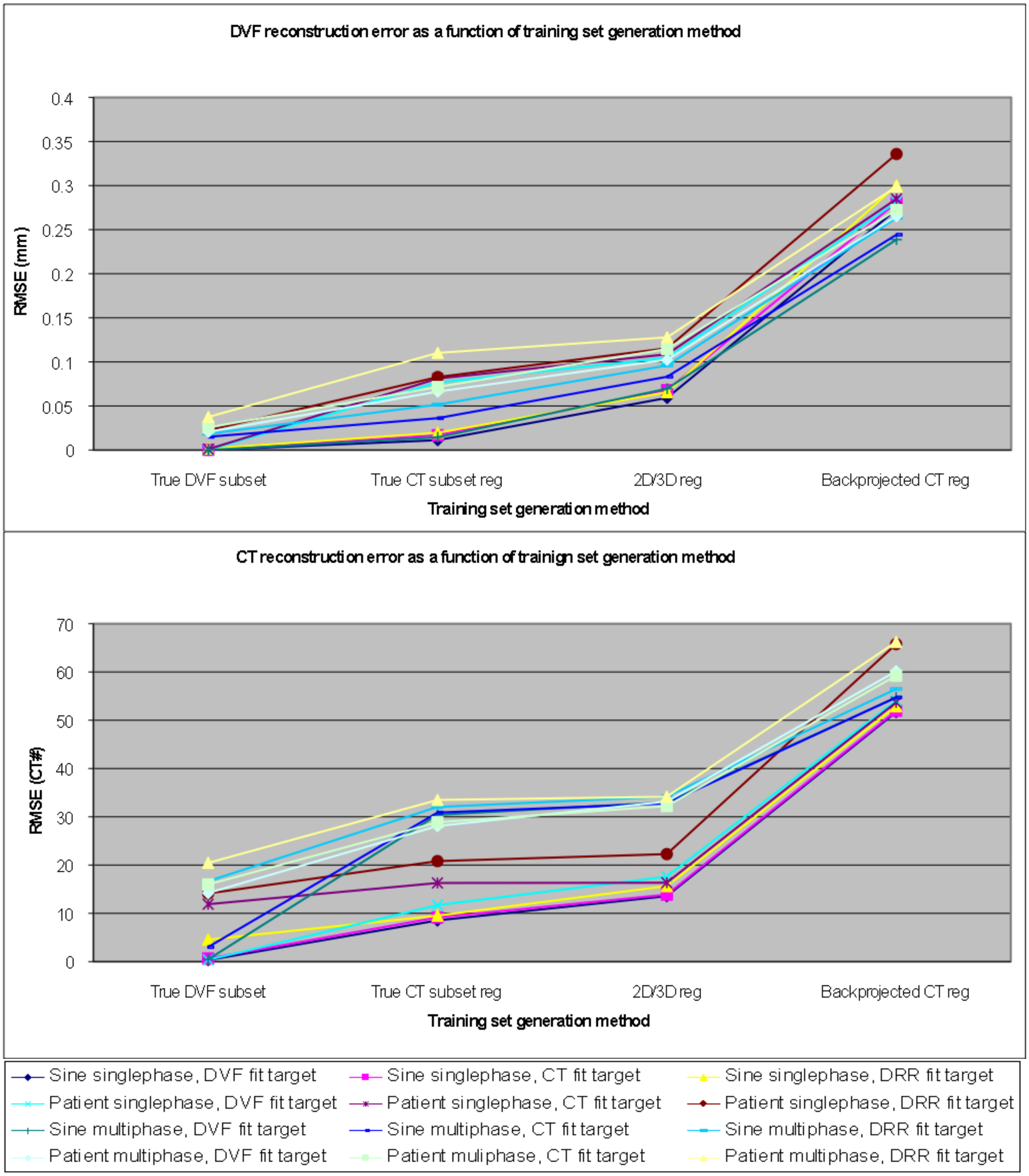


Figure 3: Variation of DVF (top) and CT (bottom) RMSE with training set generation method for different combinations of simulated motion complexity and fit target.

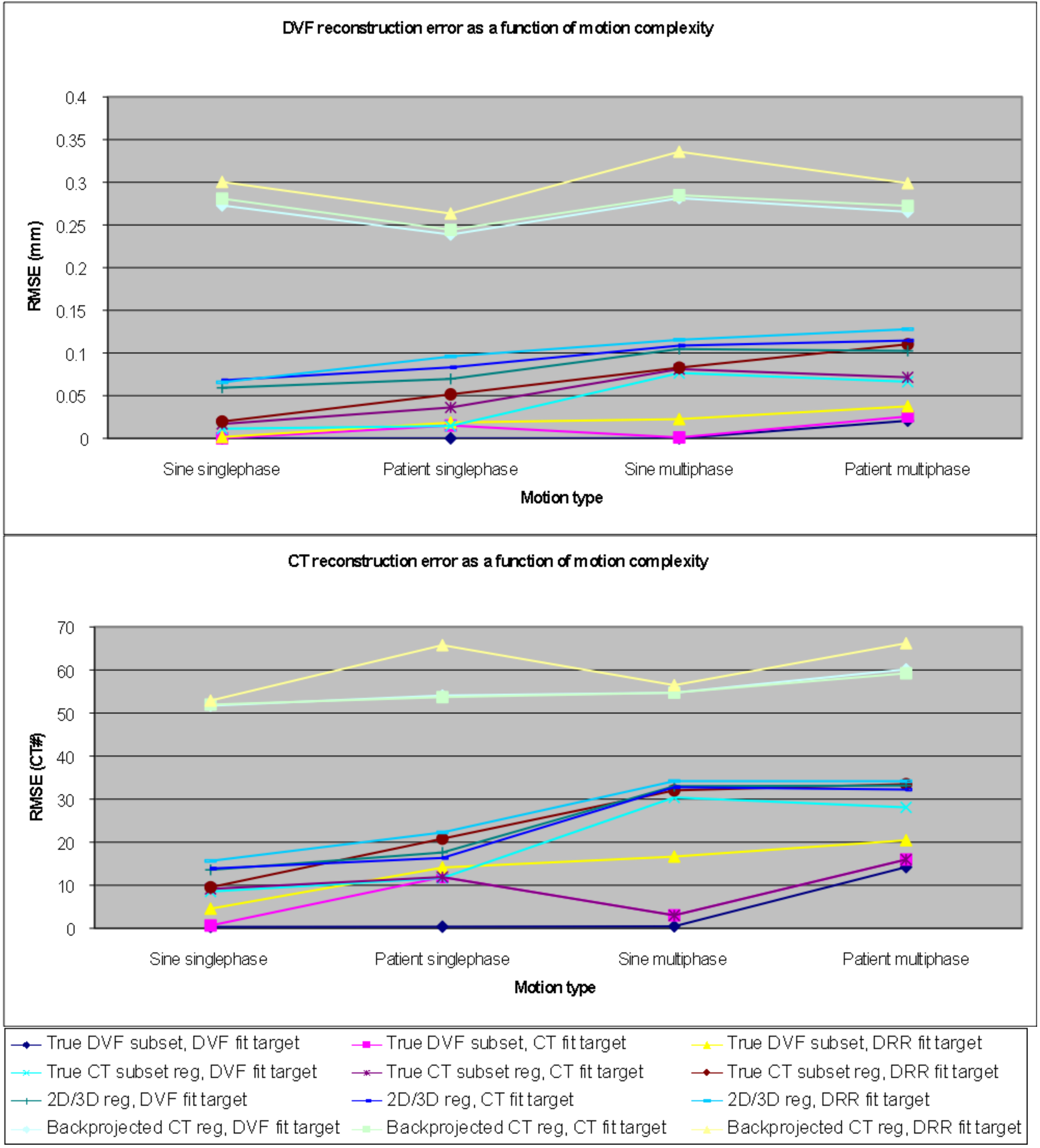


Figure 4: Variation of DVF (top) and CT (bottom) RMSE with simulated motion complexity for different combinations of training set generation method and fit target.

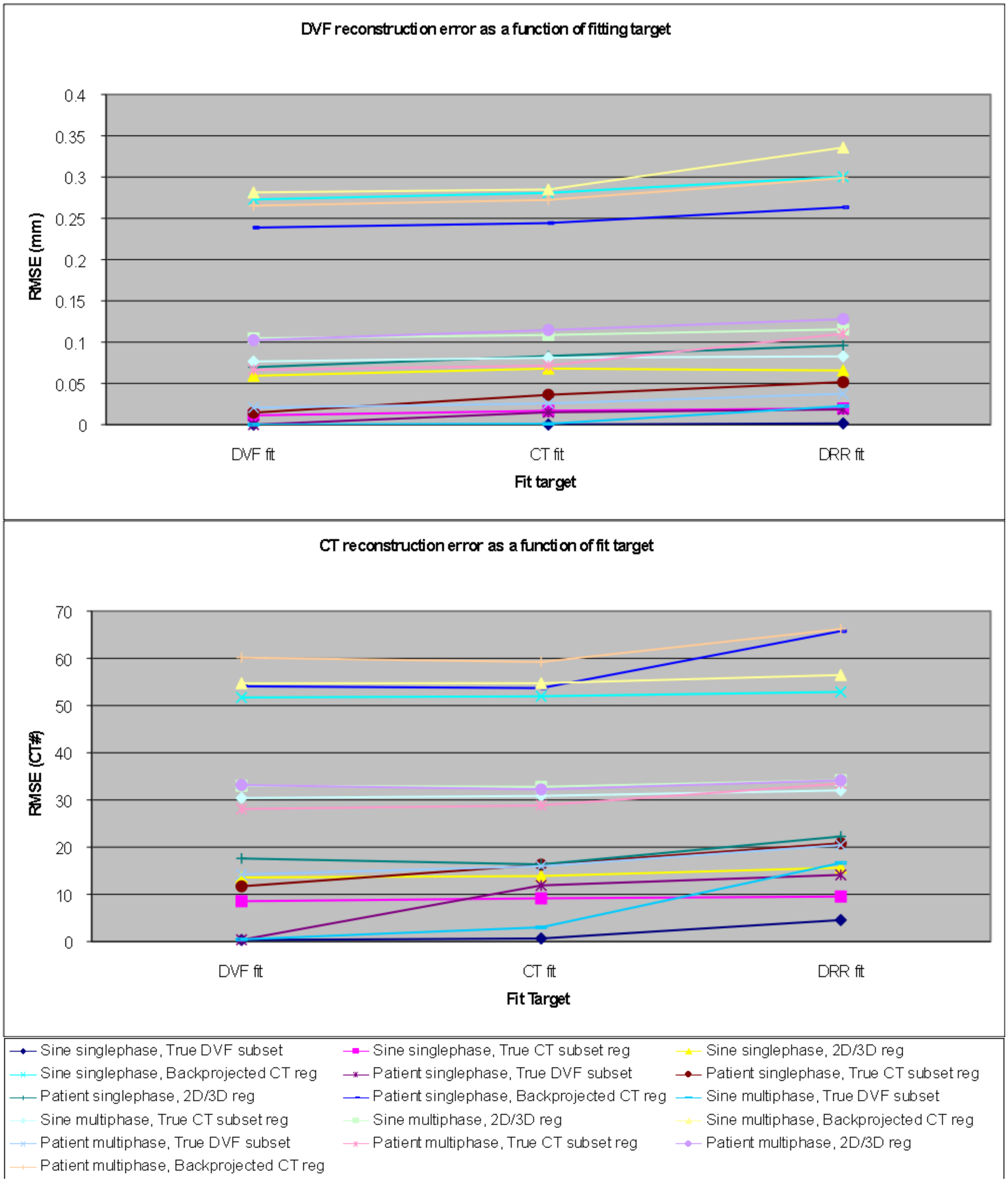


Figure 5: Variation of DVF (top) and CT (bottom) RMSE with fit target for different combinations of simulated motion complexity and training set generation method.

From these figures we can discern the overall trend correlated with each environment variable and we can also focus on individual trends to make useful observations. We quantified these figures by computing the mean of each column of data points, which are listed in Table 1, Table 2, and Table 3 for DVF training set generation method, motion complexity, and fitting target respectively.

Table 1: Average reconstructed DVF and CT RMSE for different DVF training set generation methods.

RMSE variation with DVF training set generation method				
	True DVF subset	True CT subset reg	2D/3D reg	FBP 3D/3D reg
DVF RMSE (mm)	0.012	0.053	0.093	0.278
CT RMSE (i.u.)	8.56	21.65	24.91	57.3

Table 2: Average reconstructed DVF and CT RMSE for different simulated motion complexities.

RMSE variation with simulated motion complexity				
	sine, single-phase	patient, single-phase	sine, multi-phase	patient, multi-phase
DVF RMSE (mm)	0.091	0.094	0.124	0.126
CT RMSE (i.u.)	19.37	25.41	31.61	35.52

Table 3: Average reconstructed DVF and CT RMSE for different fitting targets.

RMSE variation with fitting target			
	DVF fit target	CT fit target	DRR fit target
DVF RMSE (mm)	0.099	0.107	0.122
CT RMSE (i.u.)	25.76	26.97	31.2

The data plotted in Figure 3, Figure 4, and Figure 5 was obtained by fitting 15 time steps spanning about two breathing cycles. This provided a close up look at the algorithm's performance over a short, closely sampled time interval. For 4DCBCT data-sets generated using the sine breathing trace each breathing period is identical so the number of breathing cycles considered in the fit is not important. However, to ensure that the performance of our algorithm did not degrade when we considered more extended periods of patient breathing we also fit data at 100 time points covering about 10 patient breathing cycles for the most realistic fitting scenario, i.e., motion driven by the patient breathing trace with multiple phase shifts and projections as the fitting target. For all training set generation methods, including more time points spanning a larger number of breathing cycles improved the accuracy of the fit. For training set methods 1-4 the reductions in reconstruction error of the extended fit over the limited fit in terms of percent DVF RMSE/CT RMSE were 19 % / 15 %, 12 % / 8 %, 9 % / 7 %, and 5 % / 5 % respectively.

2.6 Discussion

Conventional backprojection, algebraic, and statistical reconstruction methods for CBCT must recover all of the Hounsfield numbers from the projection data. This places maximal demands on the amount of sampled projection data and the accuracy of the physical processes incorporated in the reconstruction process, resulting in well-known CBCT deficiencies. Phase-binned reconstruction for 4DCBCT only makes the problems worse.   We have presented here a new reconstruction approach that sidesteps these problems. Our strategy of deforming a prior FBCT image to match CBCT projection images not only relaxes the demands on projection data sampling through the use of prior information, it obviates the need to recover Hounsfield numbers directly from the CBCT projections. Furthermore, the use of a motion model to animate the prior image allows us to obtain a temporal portrait of anatomical motion that is limited only by the time resolution and the duration of the CBCT scan, rather than by the practical considerations of conventional 4DCT phase bins.   The gain in image quality when moving from a typical FBP reconstruction algorithm to a FIPM based reconstruction algorithm incorporating prior information is clearly seen in Figure 2.

Our 4DCBCT reconstruction concept was subjected to tests designed to address three questions:

1. What is the best way to generate the DVF training set for PCA?
2. How is the algorithm’s performance influenced by the complexity of the breathing motion?
3. How does each proxy connecting the DVF deformation model to the output DRRs affect the accuracy with which the optimal model parameters can be determined?

2.6.1 DVF training set generation method

We investigated two practical means of obtaining a DVF training set -- the 2D/3D FIPM-bspline forward projection matching approach (method 3 from section 2.4.1), and the conventional backprojected 3D/3D registration approach (method 4 from section 2.4.1) -- and compared them to the idealized methods 1 and 2. The two idealized methods represent the gold standards. Figure 3 clearly shows that the FIPM-bspline 2D/3D approach produces better results than the backprojected 3D/3D method. Furthermore, results from the 2D/3D method don’t differ substantially from those obtained by directly registering a sub-sample of the known CTs to the reference CT (method 2 from section 2.4.1). As expected, avoiding registration altogether by using a sub-sample of the known DVFs (method 1 from section 2.4.1) provides the best result. Looking at Table 1 we see that method 4 produced a roughly threefold increase in

RMSE over method 3, while method 3 introduced a less substantial increase in error over method 2. One point of interest is that while the average DVF RMSE of method 3 was nearly twice that of method 2, the CT RMSE was only slightly higher. This is due to the fact that large differences in the DVF often cause negligible changes in the resulting deformed CT, especially when the CT contains objects that are more or less homogeneous. Exceptions can occur at object boundaries, where small DVF changes may produce large intensity changes in edge voxels of the deformed image. Finally, we note that even areas of the numerical phantom that appear homogenous tend to have a voxel-to-voxel standard deviation of around 8-10 i.u., which is typical of a fan-beam CT. Average CT RMSEs from the first 3 methods fall within 2 or 3 factors of this number, while the RMSE of method 4 is greater by a factor of 5 or 6.

In our comparison of the DVF training set generation methods we chose to implement an idealized image binning procedure for methods 3 and 4. Our purpose was to eliminate any residual motion from the bins and ensure that each contained an adequate projection count in order to compare both approaches under ideal circumstances. In the real world, projections will be binned according to the patient breathing trace, which will result in intra-bin residual motion and irregular numbers of projections per bin. We expect that under these conditions the forward iterative projection matching method will further distance itself from the backprojection method. The FIPM-bspline method has been shown to produce accurate results with far fewer projections than are required for a FBP reconstruction of decent quality.[31] This allows for narrower bin widths which will in turn reduce residual motion within the bin. In addition, if bins are narrower, then more bins can be used without redundancy, allowing for a more densely populated DVF training set that should lead to more accurate PCA eigenvectors.

2.6.2 Simulated motion complexity

Our second concern was the performance of the algorithm when faced with different degrees of simulated motion complexity. We have broken this down into two sub-areas of inquiry: 1) How many eigenmodes are sufficient to create an accurate motion model, and 2) can our breathing trace parameterization accurately represent known principal coefficient traces as motion complexity increases?

Our expectation based on our previous experience[32] as well as that of others[24, 27, 28] was that three eigenmodes would be sufficient for an accurate model. Figure 6 illustrates the eigenvalue spectra for the four different complexities of simulated motion when PCA is computed on DVFs sampled directly from the known DVF time series.

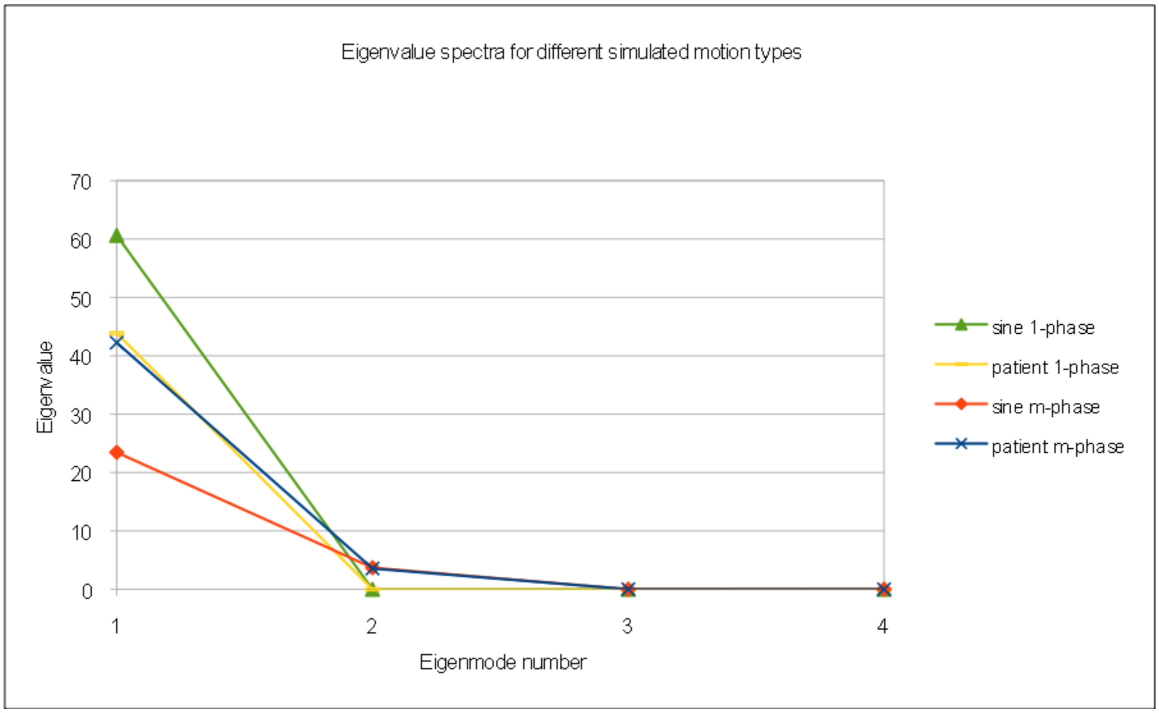


Figure 6: Eigenvalue spectra for PCA performed on DVFs generated with different simulated motion complexity. The DVFs are directly sampled from the sets of known DVFs.

The two single-phase motion types each generate only one non-zero eigenmode, the sine multi-phase type generates two non-zero modes, and the patient multi-phase type generates 3 non-negligible modes. This is what one would expect from the principal components analysis of time-varying DVFs. For a single-phase, sinusoidal motion, the covariance matrix of the DVFs has rank one, and thus only one non-zero eigenvector. For a sinusoidal motion with different voxel phases the covariance matrix has rank two and there are only two non-zero eigenvectors. For more complex motions, the covariance matrix has higher rank (cf reference [34] for the proof). For these idealized cases, three eigenmodes will be sufficient to fully model any of these motion types. However, when imperfect DVF training data are used from methods 2-4 a tail appears in the eigenvalue spectra extending non-zero values all the way to mode 15 (there are 16 DVFs per training set, which allows for a maximum of 15 non-zero eigenmodes). Even so, we have observed that in terms of DVF or CT error there is no benefit gained by modeling more than 3 modes. In terms of convergence of the optimization routine, including more eigenmodes does not appear to have any effect. In the cases tested the algorithm reduced the objective function to the same value when fitting 3 modes as it did when fitting 4 or more modes.

To examine the effectiveness of our breathing trace parameterization it is useful to look at the bottom-most trend of Figure 4. This trend plots the RMSE for fits to motion data of varying complexity under the most ideal conditions, i.e. the known DVFs are used as the fitting target and method 1 is used to generate the DVF training set. The first three data-points are more or less flat, with an increase in RMSE occurring only for patient multi-phase motion complexity. This is due to the fact that pseudo-periodic motion with multiple phases cannot be perfectly modeled by the parameterized function of the breathing trace, while the simpler three motion types can. Figure 7 displays the parameterized breathing trace of Eqn. (4) fit directly to the known principal coefficient traces for the first and second breathing modes of patient multi-phase type data. Such principal coefficient traces can be constructed by projecting the known DVFs from the simulated 4DCBCT data-set into error free eigenvectors generated using method 1 from section 2.4.1. Figure 7 shows that the breathing trace fit is close to the known principal coefficient traces but cannot represent them perfectly. This limitation becomes a greater issue when multiple eigenmodes contribute equally to the overall variance of the motion data. However, we have previously observed only one dominant mode with one or two much weaker modes in a data-set tracking patient breathing[32], and have designed our simulated motion to match. Thus the increase in fitting error is small when going from the sine multi-phase motion type to the patient multi-phase motion type, as illustrated by Figure 4 and quantified by Table 2. Overall Figure 4 and Table 2 indicate a very gradual increase in reconstruction error as the complexity of our simulated motion increases.

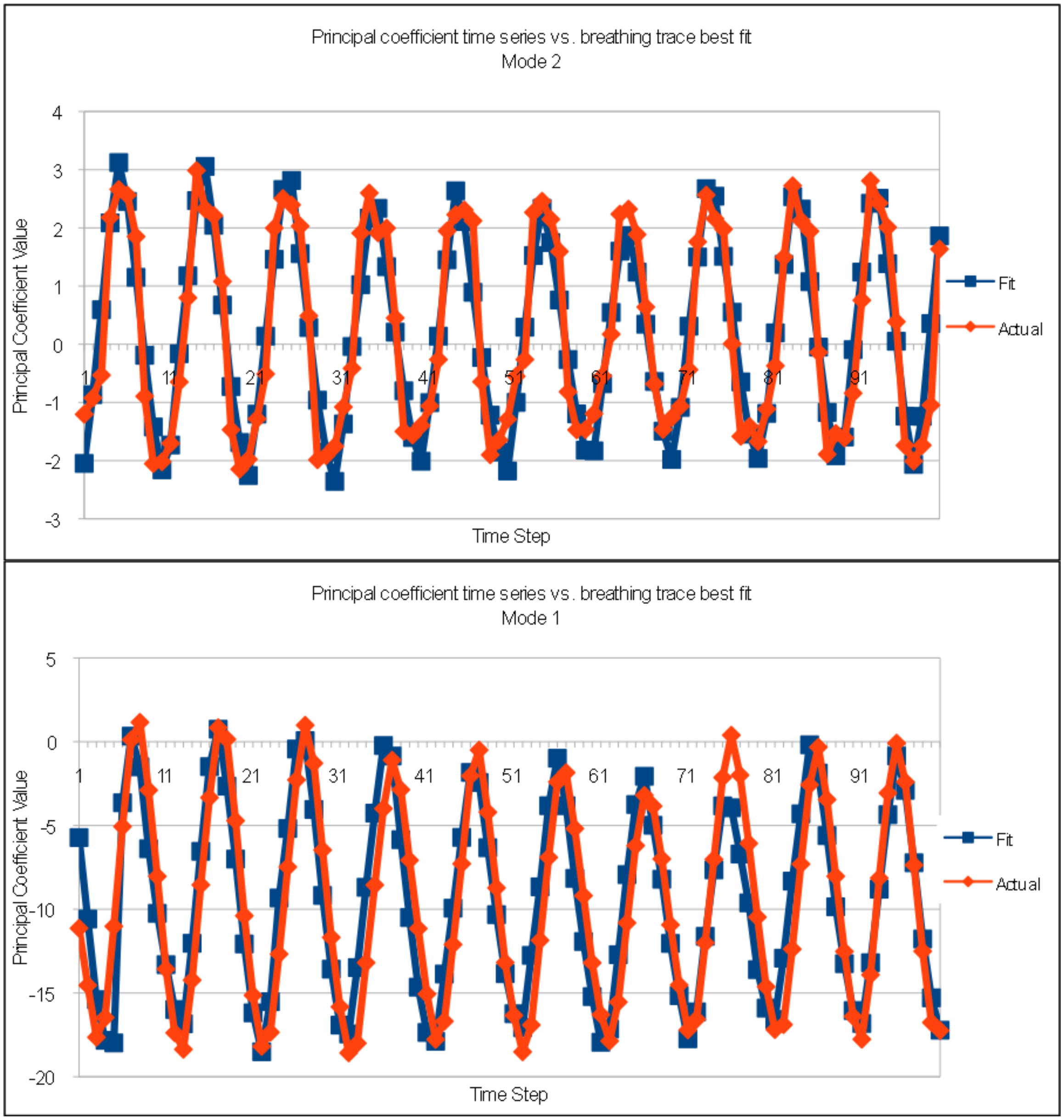


Figure 7: Parameterized breathing trace fit to known principal coefficient traces for the first (top) and second (bottom) eigenmodes. The known principal coefficient traces were produced by taking DVF time series generated with the patient multi-phase motion type and projecting them into the first and second eigenvectors. Eigenvectors were trained on a set of DVFs sampled from the known DVF time series.

2.6.3 Fitting target

Our final area of inquiry concerned the compression of DVF data down to the DRR level at which matching to known data is performed. This gives the optimization algorithm less information to constrain the model parameters. When starting from parameter values initialized to zero, markedly worse results can occur when using DRRs as the target data instead of DVFs or CTs. However, this problem can be avoided by using a good starting guess for the fit parameters, as evidenced by the results in Figure 5. One can get a good starting set of model parameters by projecting the DVF training set into the eigenvectors to get estimated principal coefficient traces and then fitting them directly with an average phase of the parameterized breathing trace function. As with our test of motion complexity the overall trend depicted by Figure 5 and Table 3 is a very gradual increase in reconstruction error as the fitting target becomes less optimal.

In addition, the effects of target data compression were a deciding factor in our choice of optimization algorithm. Our initial pick was the Levenberg-Marquardt algorithm, which offered fast convergence in a small number of iterations for the DVF fitting case, but failed at the CT and DRR matching levels. At the DVF matching level the optimization landscape is relatively smooth and contains a steep and well pronounced global minimum. However, when CT and DRR matching are introduced the sharp global minimum becomes a more gently sloping basin and shallow local minima appear as ripples in the texture of the optimization landscape. Examples of the different optimization landscapes encountered when using the DVFs or the DRRs as the matching targets are presented in Figure 8. Both the Polak-Ribiere conjugate gradient method and the BFGS algorithm provided at least partial success, but the Nelder-Mead simplex algorithm, our only non-gradient method, proved to be the most dependable. Our motion model requires the optimization of 3*M free parameters, where M is the number of eigenmodes incorporated. For our simulations M was equal to 3, giving us 9 free parameters. For comparison, the B-spline models of Docef et al,[26] and Zeng et al.[25] both contain thousands of free parameters while the PCA model of Li et al.[27, 28] requires only M free parameters, where they set M equal to 3. However, unlike our model and that of Docef et al. and Zeng et al. their model contains no temporal dependence and therefore the fitting process must be repeated at every time step. In terms of algorithm efficiency, when fitting data simulated with the most realistic conditions around 200 iterations are generally required for convergence of the Nelder-Mead simplex algorithm. On an Intel core i7 CPU running at 2.8 GHz with 12 GB of memory and considering 15 time steps in the fit this process generally took between 5-10 minutes. On our initial GPU implementation of the code (GeForce GTX480) the same trial finished in about 1-1.5 minutes.

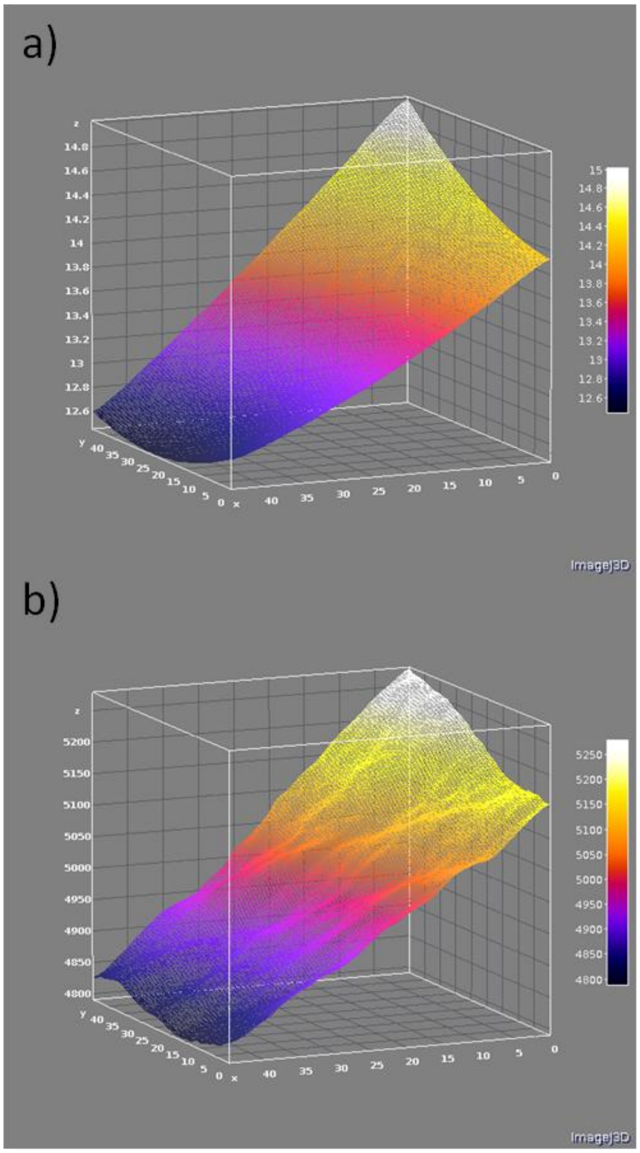


Figure 8: a) Slice taken through the optimization landscape with DVFs used as the matching target. b) A slice through the same section of parameter space but with DRRs used as the matching target. The data for this fit was generated using sine motion with no phase shifts and DVFs sampled from the known set of DVFs for the training set.

2.7 Conclusions

The contents of this chapter are designed to establish the mathematical framework upon which our algorithm is based and put it through proof of concept testing designed to validate its worthiness for further pursuit. Starting from conditions in which it is essentially guaranteed to perform well we have seen that reconstruction performance degrades in a gradual and orderly manner when more realistic test scenarios are considered. The one exception to this rule was the use of FBP 4DCBCT reconstruction coupled with 3D/3D DIR to generate the DVF training set (method 4 from section 2.4.1). In this case both DVF and CT RMSE experienced a marked jump compared to the gold standard methods as well as the FIPM based 2D/3D registration method (method 3 from section 2.4.1). Therefore the FIPM based method is the clear choice moving forward.

3 Tuning for the real world

3.1 Introduction

For the most realistic tests described in Chapter 2 we endeavored to simulate complex breathing motion, used DVF training set generation methods that will be available for actual patient data, and used projections as the fitting target data. However, those projections were generated using the same process as that used to create our DRRs. In the real world our DRR module will not perfectly simulate the projection formation process, and the degree to which the two disagree may have a large effect on the success of our FIPM-PCA 4DCBCT reconstruction algorithm. The influence of such projection/DRR mismatch was explored in the testing and validation of the 2D/3D FIPM-bspline DIR algorithm used by method 3 in Section 2.4.1 to generate the DVF training set. In that work contrast mismatch was simulated by adding values sampled from a non-linear gamma function to the target projections, with peak gamma function amplitudes equal to either 0.25 or 0.5 % of the max pixel value. Results indicated that the normalized root mean squared error (nRMSE) of the deformed CT when using a high contrast numerical phantom increased from a baseline of about 1 % to 3 and 5 % for the two contrast mismatch amplitudes. For a more realistic pelvic CT image deformed CT nRMSEs increased from 1 % to 12 % and 23 %.[31] This chapter is devoted to our efforts to devise a DRR algorithm that achieves the optimal combination of realism and computational efficiency in order to minimize the considerable degradation possible due to this mismatch effect.

Our group has developed and investigated several FIPM based 2D/3D registration methods, of which the topic of this work is one.[35-38] In general, this class of algorithms requires the calculation of DRRs from the FBCT for comparison to actual cone-beam radiographic projections. The anatomical features that must be registered, as well as the quantitative manner in which the DRRs and the projections are compared (i.e., the matching metric, or merit function), determine the accuracy with which the two images must match.

For example, rigid 2D/3D registration for radiotherapy applications typically focuses on high-contrast bony features, while a deformable registration algorithm must be sensitive to low contrast soft tissue features. Among other applications of DRR-to-projection matching we take notice of simultaneous algebraic reconstruction techniques and prior-image-constrained remote sensing. [18, 39] Each of these FIPM applications has its own requirements for the DRR quality.

The problem of rigid 2D/3D registration via projection matching to DRRs has a long history of study, during which many different image matching metrics have been investigated.[40-42] Though our interest is in deformable 2D/3D registration, insights from rigid registration can provide us with useful guiding principles with regards to such optimization parameters. The lengthy experience documented in the literature demonstrates that each metric has merits and weaknesses. The sum-of-squared intensity differences metric has the advantage of being robust with respect to noise and is easy to formulate for a gradient-driven iterative search process, but requires a close correspondence of the absolute image intensities in the DRRs and the projections.[38, 43, 44] The many formulations of mutual information allow much greater variability in the comparative image intensities, but most studies of its effectiveness have involved 3D images. [45] Mutual information has been shown to be less robust for 2D image matching, owing to the relatively small number of available data samples compared to 3D/3D registration.[40, 45] This weakness is compounded when there is noise. It also has a smaller and more sharply delimited capture range, which makes it less robust in a gradient-driven iterative search. Formulating gradients of the iterative cost function for mutual information is also more difficult than for SSD.[44]

In our development of deformable FIPM we have preferred a gradient search and therefore have opted to use the sum-of-squared differences as our image matching metric. We have already shown via simulations that this formulation is sensitive to contrast mismatch between the DRRs and the projection images, but is relatively insensitive to noise.[38] This has motivated the present work. Our goal here was to calculate DRRs solely from fundamental physical principles underlying image formation to accurately replicate the absolute image intensities in a CBCT projection; i.e. to calculate DRRs from first principles. We avoid recourse to intensity normalization or any other artificial numerical intensity adjustment, which could bias the registration towards an inaccurate deformation. It is not our purpose to maximize the contrast in the DRRs, but to reproduce what is present in the projections. The effects of noise on FIPM are minor and relatively easy to incorporate, and therefore were not considered here.

Many investigators have described algorithms covering various aspects of the DRR generation process. In particular, much attention has been paid to the x-ray volume rendering component, with regards to geometric realism, spatial resolution, and computational efficiency. Siddon introduced a widely used algorithm for efficiently calculating the exact radiological path through a CT volume. [46] Russakoff et al. developed the concept of “attenuation fields” for fast DRR rendering.[47] Moore et al. used pencil beams rather than rays as projectors to increase DRR realism and avoid aliasing effects.[48] Birkfellner et al. used a “wobbled splatting” technique for fast DRR generation.[49] Other techniques including shear-warp volume rendering, frequency domain volume rendering, and Monte-Carlo integration based methods have also been investigated.[50-52] We note that much of this prior work is tailored to rigid 2D/3D registration for patient setup, which emphasizes precise geometrical reproduction of the anatomical pose and benefits from the high contrast of skeletal features. In addition, many authors have used a graphics processing unit (GPU) to speed up DRR computation.[53-57] Our algorithm uses a standard ray-casting approach with GPU acceleration.[57]

The basic ray-tracing component of a DRR algorithm models only the attenuation of primary photons, whereas projection image formation is influenced by additional physical effects. One such effect is scatter, which can have a large presence in CBCT projections.[58] Various methods of estimating CBCT scatter exist, including beam stop arrays, scatter kernel methods, and Monte Carlo simulation. Ruhrnschopf et al. give a comprehensive summary of scatter estimation and correction techniques.[59] These solutions are generally framed in the context of scatter mitigation in CBCT projections, rather than scatter modeling for DRRs. Moore et al. applied a scatter correction specifically aimed at improving DRRs using the beam stop array method.[48] Our approach is to use an in-house Monte Carlo algorithm simulating the CBCT system to obtain scatter estimates.[60] In contrast to the beam stop array method, this approach does not require any special equipment and more importantly avoids any extra scans, which deliver extra dose to a patient.

A second effect influencing projection image formation is beam hardening. Again, there are many papers describing beam hardening and beam hardening corrections, but nearly all are applied to improving CT reconstruction. Two popular methods assume that the subject is composed entirely of water or bone.[61, 62] Yan et al. describe a method for correcting single and dual energy CT using an iterative method by representing each voxel in the CT as a sum of basis materials.[63] In their DRR algorithm Moore et al. explicitly modeled a poly-energetic beam rather than applying a correction. For our DRR algorithm we evaluated both an analytical and measurement approach to a beam hardening correction implemented in post-processing.

A final effect is veiling glare, or leakage of signal from a detector pixel to its neighbors.[64-66] A group at our institution has measured the effect of veiling glare on the flat panel detector used in this study.

We maximize the realism in our DRRs by using an efficient ray-tracing routine combined with scatter, beam hardening, and veiling glare corrections. In this chapter we provide a step by step account of the measurements and calculations needed by our algorithm. Though many descriptions of these effects have been described separately, few works exist that consolidate them into one algorithm for the purpose of computing realistic DRRs.

3.2 Projection definitions

We begin by defining a CBCT projection through an object:

$$P_{CBCT}(\mathbf{r}_i) = A_1 \Phi_{CBCT}(\mathbf{r}_i) \alpha_{CBCT}(\mathbf{r}_i) R_{CBCT}(\mathbf{r}_i) + S_{CBCT}(\mathbf{r}_i) \qquad (7)$$

where $\mathbf{r}_i$ is the location of detector pixel *i* (all coordinates are in the world coordinate frame), $P_{CBCT}(\mathbf{r}_i)$ is the signal recorded by detector pixel *i*, A_1 is a tube output factor corresponding to mAs, $\Phi_{CBCT}(\mathbf{r}_i)$ is the source fluence distribution aimed at the detector pixel, $\alpha_{CBCT}(\mathbf{r}_i)$ is the transmission, or the fraction of primary particles that remain when the beam reaches the detector pixel, $R_{CBCT}(\mathbf{r}_i)$ is a pixel dependent detector response function, and $S_{CBCT}(\mathbf{r}_i)$ is the contribution of scattered photons to the detector signal. We further define the transmission as:

$$\alpha_{CBCT}(\mathbf{r}_i) = \exp[-\mathrm{M}(\mathbf{r}_i)] = \exp\left(-\int_{\mathbf{r}_0}^{\mathbf{r}_i} \mu(\mathbf{r}) dr \right) \qquad (8)$$

where $\mu(\mathbf{r})$ is the linear attenuation coefficient (LAC) at location $\mathbf{r}$ within the object being imaged, $\mathbf{r}_0$ is the location of the source, and $\mathrm{M}(\mathbf{r}_i)$ is the line integral of linear attenuation coefficients (cumulative LAC) along the ray connecting $\mathbf{r}_0$ to $\mathbf{r}_i$. We can expand the scatter term as:

$$S_{CBCT}(\mathbf{r}_i) = A_1 \sigma_{CBCT}(\mathbf{r}_i) R_{CBCT}(\mathbf{r}_i) \tag{9}$$

where $\sigma_{CBCT}(\mathbf{r}_i)$ is the object scatter distribution. In general, the scatter distribution is a complex function of the object being imaged, the source fluence distribution, and the detector pixel location.

Finally, it is useful to define a special case of Eq. (7) termed a flood projection. A flood projection is taken with no object in the beam, thus transmission is 1 and scatter is 0 across the detector surface. This leads to the following simplified mathematical description:

$$F_{CBCT}(\mathbf{r}_i) = A_0 \Phi_{CBCT}(\mathbf{r}_i) R_{CBCT}(\mathbf{r}_i) \tag{10}$$

where $F_{CBCT}(\mathbf{r}_i)$ is the flood projection signal recorded by detector pixel *i*. The flood projection empirically captures the source fluence distribution and detector response function.

3.3 DRR challenges

A CT image represents the x-ray attenuation coefficients of materials in the imaged object, relative to the attenuation coefficient of water, for the particular energy spectrum of the incident x-rays. The goal of our algorithm is to use a CT image of an object to simulate an x-ray radiograph (projection) of the object as defined in Eq. (7). We can divide the physical processes that produce a projection into primary and secondary effects.

The primary process responsible for image formation is the attenuation of primary photons by the object. This process is governed by the object's LAC distribution, $\mu(\mathbf{r})$, in the following manner:

$$I_i = I_0 \exp\left(-\int_{\mathbf{r}_0}^{\mathbf{r}_i} {}_C\, \mu(\mathbf{r}) d\mathbf{r} \right) \tag{11}$$

where I_0 is the initial intensity of the (mono-energetic) photon beam, I_i is the intensity of primary photons remaining after the beam passes through the object, and the integral runs along the line connecting $\mathbf{r}_0$ to $\mathbf{r}_i$. We can use Eq. (11) to calculate the transmission factor $\alpha(\mathbf{r}_i)$ in Eq. (7). The line integral in Eq. (11) can be estimated using a ray-tracing routine; however the CT through which we make projections has dimensionless units of CT number (attenuation relative to water), not absolute LAC. Therefore, our first requirement is to convert CT number to absolute LAC in order to render the CT in physically meaningful units. This is complicated by the fact that the absolute LAC depends on the incident x-ray energy spectrum. Therefore one must measure the x-ray attenuation through objects of known composition using the precise x-ray spectrum of the CBCT system whose projections one wishes to emulate.

The secondary processes that we need to model are scatter, beam hardening, and veiling glare. Scatter refers to particles that have been deflected by the object but are still absorbed by the detector. Beam hardening refers to the effect whereby the energy distribution of a poly-energetic beam is shifted towards higher energies as it passes through an attenuating medium. Veiling glare denotes the leakage of signal from the detector pixel in which it is generated to its neighbors. These effects must be accounted for in two ways: they must be removed from the measurements used to determine the CT number to LAC conversion function, and they must be folded into the raw DRR produced by the ray-tracing routine. This separation ensures that scatter, beam hardening, and veiling glare are explicitly and precisely accounted for in each individual DRR rather than entangled within the CT number conversion function.

3.4 Secondary physics effects

3.4.1 Beam hardening

As a poly-energetic beam passes through a material the lower energy photons are preferentially attenuated, causing the average beam energy to increase. This effect is known as beam hardening. The LAC is an energy dependent quantity, so for a given material with a uniform CT number, beam hardening causes the beam to "see" a progressively lower effective LAC as the beam traverses the material. In order to produce physically realistic DRRs we must account for this effect. To this end we must first acquire raw data that characterizes the beam hardening effect in material similar to that composing the human body. Then we must process the raw data into correction functions that can either impart or remove the beam hardening effect. We explored two different approaches for both the data collection and data processing steps. With regards to the collection of raw data we tested both measurement and calculation based methods, and for data processing we explored methods that produced an aggregate correction function applied to cumulative LAC and an “in-place” correction function applied to single LAC values.

3.4.1.1 Beam hardening data

The raw data for the measurement approach is a set of 32 CBCT projections taken through solid water blocks ranging in thickness from 2 mm to 180 mm. The solid water slabs are positioned so that the couch does not interfere with the measurement, and a minimal volume of the slabs is exposed to the beam to minimize scatter. In addition, we acquire a flood projection at the same CT technique. We begin by choosing a position on the detector that lies beneath the shadow of the solid water slabs as our measurement point and calculate the transmission for each slab thickness. Assuming scatter to be negligible, Eqs. (7) and (10) can be used to calculate transmission as follows:

$$\alpha_{CBCT}(\mathbf{r}_i) = \frac{A_0}{A_1}\frac{P_{CBCT}(\mathbf{r}_i)}{F_{CBCT}(\mathbf{r}_i)} \qquad (12)$$

In general $\frac{A_0}{A_1}$ can be estimated as the ratio of the mAs of the flood projection to that of the object projection, or calculated as the ratio of the flood projection to the object projection in an area where there is no attenuation in the object projection, i.e. $\alpha_{CBCT}(\mathbf{r}_i) = 1$. The transmission for each thickness l of solid water can be converted to the hardened cumulative LAC using Eq. (8):

$$\mathrm{M}_{bh1}(l) = -\ln[\alpha_{CBCT}(l)] \tag{13}$$

where the subscript “bh1” refers to the presence of beam hardening. After making these measurements for each thickness we estimate the corresponding set of unhardened cumulative LAC values as follows:

$$\mathrm{M}_{bh0}(l) = \mu_0 l \tag{14}$$

where μ_0 is the average LAC of the material in the unhardened beam spectrum, l is the thickness of the solid water slab, and the subscript “bh0” refers to the absence of beam hardening. While we know l for each slab we must calculate μ_0. To do this we start by fitting an 8th order polynomial function to our previously determined hardened $\mathrm{M}_{bh1}(l)$ data plotted as a function of l, generating a function $y(l)$ such that:

$$\mathrm{M}_{bh1}(l) \approx y(l) = \mathrm{a}l + \mathrm{b}l^2 + \ldots \mathrm{h}l^8 \tag{15}$$

As thickness l approaches zero we know that $\mathrm{M}_{bh1}(l)$ and $\mathrm{M}_{bh0}(l)$ will converge:

$$\lim_{l \to 0} \mathrm{M}_{bh1}(l) = \mathrm{M}_{bh0}(l) = \mu_0 l \tag{16}$$

Thus we can estimate μ_0 as the linear coefficient of our polynomial function from Eq. (15):

$$\mu_0 \approx \mathrm{a} \tag{17}$$

From this we can calculate $\mathrm{M}_{bh0}(\mathbf{r}_i)$ for each slab thickness using Eq. (14).

Our second approach is to analytically calculate the attenuation process through slabs of a test material to obtain the same data-set as above. This requires three items of input: the beam spectrum of our CBCT source, $W_0(E)$, the LAC distribution of the bowtie filter material (aluminum), $\mu_{Al}(E)$, and the LAC distribution of a test material, $\mu_{mat}(E)$, through which we calculate the passage of the beam. We use liquid water as our test material, though any material whose attenuation coefficient is known is a possible candidate. We begin by propagating the beam through the middle of the bowtie filter, where it is at its thinnest:

$$W_{BT,\min}(E) = W_0(E)\exp\left[-\mu_{Al}(E)l_{BT,\min}\right] \tag{18}$$

For all practical purposes this is our true unhardened spectrum, as images are never acquired without the bowtie filter and this is the area of the beam used to calibrate our CT number to LAC conversion function, described in section 3.5. We can now calculate the μ_0 value for our test material using an average of $\mu_{mat}(E)$ weighted by $W_{BT,\min}(E)$:

$$\mu_0 = \frac{\int W_{BT,\min}(E)\mu_{mat}(E)EdE}{\int W_{BT,\min}(E)EdE} \tag{19}$$

We then proceed on a pixel by pixel basis, propagating the spectrum through the thickness of bowtie filter that each pixel “sees”. For detector pixel *i* this new spectrum is calculated as:

$$W_{BT,i}(E) = W_0(E)\exp\left[-\mu_{Al}(E)l_{BT,i}\right] \tag{20}$$

Next, the spectrum for each pixel is filtered through thicknesses of the test material ranging from 0 mm to 150 mm in 5 mm increments. For a given thickness of test material $l_{mat,j}$ this newly filtered spectrum is:

$$W_{mat,i,j}(E) = W_{BT,i}(E)\exp\left[-\mu_{mat}(E)l_{mat,j}\right] \tag{21}$$

From here the cumulative hardened LAC corresponding to thickness $l_{mat,j}$ of material for detector pixel *i* is calculated as:

$$\mathrm{M}_{bh1}(\mathbf{r}_i) = -\ln\left(\frac{\int W_{mat,i,j}(E)EdE}{\int W_{BT,\min}(E)EdE}\right) \tag{22}$$

and the corresponding unhardened cumulative LAC is calculated as:

$$\mathrm{M}_{bh0}\left(\mathbf{r}_i\right) = \mu_0 l_{mat,j} \tag{23}$$

In practice the unhardened cumulative LAC values need only be calculated once since they have no detector pixel dependence.

3.4.1.2 Beam hardening data processing

Having obtained our data-set of M_{bh0} and corresponding M_{bh1} values, either by measurement or calculation, we are prepared to calculate our beam hardening and unhardening functions. The procedure that generates a function correcting cumulative LAC is the most straightforward of our two methods. It generates a function, $h\left(\mathrm{M}_{bh1}\left(\mathbf{r}_i\right)\right)$, that will convert hardened cumulative LAC from measured attenuation data to unhardened cumulative LAC in order to eliminate beam hardening. It also produces a function $g\left(\mathrm{M}_{bh0}\left(\mathbf{r}_i\right)\right)$ that does the reverse.

The beam hardening function $g\left(\mathrm{M}_{bh0}\left(\mathbf{r}_i\right)\right)$ is obtained by fitting an 8th order polynomial to the M_{bh1} values plotted as a function of M_{bh0}. The input to the function is an unhardened cumulative LAC value and the output is a hardening cumulative LAC value, as follows:

$$\mathrm{M}_{bh1}\left(\mathbf{r}_i\right) = g\left(\mathrm{M}_{bh0}\left(\mathbf{r}_i\right)\right) = a\mathrm{M}_{bh0}\left(\mathbf{r}_i\right) + b\mathrm{M}_{bh0}\left(\mathbf{r}_i\right)^2 + \ldots + h\mathrm{M}_{bh0}\left(\mathbf{r}_i\right)^8 \tag{24}$$

Likewise, the beam unhardening function $h\left(\mathrm{M}_{bh1}\left(\mathbf{r}_i\right)\right)$ is determined by fitting an 8th order polynomial to the M_{bh0} values plotted as a function of M_{bh1}. The input to this function is a hardened cumulative LAC value and the output is an unhardened cumulative LAC value:

$$\mathrm{M}_{bh0}\left(\mathbf{r}_i\right) = h\left(\mathrm{M}_{bh1}\left(\mathbf{r}_i\right)\right) = i\mathrm{M}_{bh1}\left(\mathbf{r}_i\right) + j\mathrm{M}_{bh1}\left(\mathbf{r}_i\right)^2 + \ldots + p\mathrm{M}_{bh1}\left(\mathbf{r}_i\right)^8 \quad (25)$$

These functions constitute the aggregate beam hardening correction.

The “in-place” beam hardening correction is more complex to produce and apply than the aggregate method, but its structure is motivated by a more natural simulation of the physical mechanism of beam hardening. The goal is a function that takes as input the total unhardened cumulative LAC the beam has encountered up to its current point and returns a scaling factor to be applied to the unhardened LAC at that point to correct for beam hardening. Thus the magnitude of the effect evolves as the beam passes through more material. To begin we extract the following relationship from Eq. (8):

$$\mathrm{M}_{bh1}\left(\mathbf{r}\right) = \int_{\mathbf{r_0}}^{\mathbf{r}} \mu_{bh1}\left(\mathbf{r}'\right) d\mathbf{r}' \quad (26)$$

Taking the derivative of both sides of Eq. (26) yields:

$$\frac{\partial}{\partial \mathbf{r}} \mathrm{M}_{bh1}\left(\mathbf{r}\right) = \mu_{bh1}\left(\mathbf{r}\right) \quad (27)$$

Or in terms of unhardened cumulative LAC using Eq. (24):

$$\frac{\partial}{\partial \mathbf{r}} g\left(\mathrm{M}_{bh0}\left(\mathbf{r}\right)\right)=\mu_{bh1}\left(\mathrm{M}_{bh0}\left(\mathbf{r}\right)\right) \tag{28}$$

Using Eq. (28) we define our in-place beam hardening correction function as:

$$k\left(\mathrm{M}_{bh0}\left(\mathbf{r}\right)\right)=\frac{\frac{\partial}{\partial \mathbf{r}} g\left(\mathrm{M}_{bh0}\left(\mathbf{r}\right)\right)}{\mu_0}=\frac{\mu_{bh1}\left(\mathrm{M}_{bh0}\left(\mathbf{r}\right)\right)}{\mu_0} \tag{29}$$

This function produces a value between 0 and 1 that scales an unhardened LAC value to impart beam hardening effects based on the quantity of material (denoted by $\mathrm{M}_{bh0}\left(\mathbf{r}\right)$) that the beam has passed through up to that point. It is applied as follows:

$$\mu_{bh1}\left(\mathbf{r}\right)=k\left(\mathrm{M}_{bh0}\left(\mathbf{r}\right)\right)\mu_{bh0}\left(\mathbf{r}\right) \tag{30}$$

A function to convert hardened LAC to unhardened LAC can be similarly computed:

$$j\left(\mathrm{M}_{bh1}\left(\mathbf{r}\right)\right)=\frac{\frac{\partial}{\partial \mathbf{r}} h\left(\mathrm{M}_{bh1}\left(\mathbf{r}\right)\right)}{\mu_{bh1}\left(\mathbf{r}\right)}=\frac{\mu_{bh0}\left(\mathrm{M}_{bh1}\left(\mathbf{r}\right)\right)}{\mu_{bh1}\left(\mathbf{r}\right)} \tag{31}$$

And applied as:

$$\mu_{bh0}(\mathbf{r}) = j\left(\mathrm{M}_{bh1}(\mathbf{r})\right)\mu_{bh1}(\mathbf{r}) \tag{32}$$

For both aggregate and in-place corrections if the raw beam hardening data are obtained by the measurement method only one of each type of function will be produced. If the data are obtained via calculation unique versions of both functions will be produced for each pixel.

3.4.1.3 Equivalence of aggregate and in-place beam hardening corrections

In general Eq. (32) is not particularly useful since it requires knowledge of the object's hardened LAC distribution, which is dependent not only on initial beam energy but also the precise geometrical setup of the object with respect to the beam. However, Eq. (30) can be practically applied to include beam hardening effects in the calculation of the transmission term described in Eq. (8). From the standpoint of computational efficiency the aggregate correction method is clearly superior as it is applied once after ray-tracing has finished rather than many times as the ray progresses through the object. However, the in-place method is a more faithful simulation of the actual beam hardening process. Therefore, it is useful to investigate whether or not the two corrections produce different results with regards to hardened cumulative LAC in order to choose one or the other. To this end we compute cumulative hardened LAC values through two objects with analytically defined unhardened LAC distributions using both methods and compare the results.

We begin by re-defining our beam hardening corrections. From Eq. (24), which we truncate to a third order polynomial for simplicity, the aggregate correction is:

$$\mathrm{M}_{bh1}(\mathbf{r}) = g\left(\mathrm{M}_{bh0}(\mathbf{r})\right) = a\mathrm{M}_{bh0}(\mathbf{r}) + b\mathrm{M}_{bh0}(\mathbf{r})^2 + c\mathrm{M}_{bh0}(\mathbf{r})^3 \tag{33}$$

Plugging Eq. (33) into Eq. (29) yields the form of the in-place correction:

$$\begin{aligned} k\left(\mathrm{M}_{bh0}(\mathbf{r})\right) &= \frac{\frac{\partial}{\partial \mathbf{r}} g\left(\mathrm{M}_{bh0}(\mathbf{r})\right)}{\mu_0} \\ &= \frac{a\frac{\partial}{\partial \mathbf{r}}\mathrm{M}_{bh0}(\mathbf{r}) + 2b\mathrm{M}_{bh0}(\mathbf{r})\frac{\partial}{\partial \mathbf{r}}\mathrm{M}_{bh0}(\mathbf{r}) + 3c\mathrm{M}_{bh0}(\mathbf{r})^2\frac{\partial}{\partial \mathbf{r}}\mathrm{M}_{bh0}(\mathbf{r})}{\mu_0} \\ &= \frac{\frac{\partial}{\partial \mathbf{r}}\mathrm{M}_{bh0}(\mathbf{r})\left(a + 2b\mathrm{M}_{bh0}(\mathbf{r}) + 3c\mathrm{M}_{bh0}(\mathbf{r})^2\right)}{\mu_0} \\ &= a + 2b\mathrm{M}_{bh0}(\mathbf{r}) + 3c\mathrm{M}_{bh0}(\mathbf{r})^2 \end{aligned} \tag{34}$$

where the final equality is obtained by using Eq. (14) to note that $\mathrm{M}_{bh0}(\mathbf{r}) = \mu_0 |\mathbf{r}|$ and subsequently $\frac{\partial}{\partial \mathbf{r}}\mathrm{M}_{bh0}(\mathbf{r}) = \mu_0$.

Next we define two simple objects to be used as test cases. Objects are defined as spatial distributions of unhardened LAC as follows:

$$\mu_{bh0}(\mathbf{r}) = |\mathbf{r}| \tag{35}$$

and

$$\mu_{bh0}\left(\mathbf{r}\right)=\cos\left(\left|\mathbf{r}\right|\right) \tag{36}$$

To begin, we find the cumulative hardened LAC for the first object using the aggregate method. From Eq. (8) we first calculate the cumulative unhardened LAC for the object:

$$\mathrm{M}_{bh0}\left(\mathbf{r}\right)=\int_{\mathbf{r_0}}^{\mathbf{r}}\mu_{bh0}\left(\mathbf{r}'\right)d\mathbf{r}'=\int_{\mathbf{0}}^{\mathbf{r}}\left|\mathbf{r}'\right|d\mathbf{r}'=\frac{\left|\mathbf{r}\right|^2}{2} \tag{37}$$

where we have assumed that $\mathbf{r_0}=\mathbf{0}$ for convenience. Plugging this result into Eq. (33) yields the desired expression for cumulative hardened LAC:

$$\mathrm{M}_{bh1}\left(\mathbf{r}\right)=g\left(\frac{\left|\mathbf{r}\right|^2}{2}\right)=a\frac{\left|\mathbf{r}\right|^2}{2}+b\frac{\left|\mathbf{r}\right|^4}{4}+c\frac{\left|\mathbf{r}\right|^6}{8} \tag{38}$$

Next we test the in-place method for the same object. Using Eq. (26), Eq. (34), and Eq. (37) the hardened cumulative LAC is calculated:

$$
\begin{aligned}
\mathrm{M}_{bh1}(\mathbf{r}) &= \int_{\mathbf{0}}^{\mathbf{r}} \mu_{bh1}(\mathbf{r}')d\mathbf{r}' \\
&= \int_{\mathbf{0}}^{\mathbf{r}} k\left(\mathrm{M}_{bh0}(\mathbf{r}')\right)\mu_{bh0}(\mathbf{r}')d\mathbf{r}' \\
&= \int_{\mathbf{0}}^{\mathbf{r}} \left(a + 2b\frac{|\mathbf{r}'|^2}{2} + 3c\frac{|\mathbf{r}'|^4}{4} \right)\left(|\mathbf{r}'|\right)d\mathbf{r}' \\
&= \int_{\mathbf{0}}^{\mathbf{r}} \left(a|\mathbf{r}'| + b|\mathbf{r}'|^3 + 3c\frac{|\mathbf{r}'|^5}{4} \right)d\mathbf{r}'
\end{aligned}
\tag{39}
$$

Evaluating the final integral yields:

$$
\begin{aligned}
&\int_{\mathbf{0}}^{\mathbf{r}} \left(a|\mathbf{r}'| + b|\mathbf{r}'|^3 + 3c\frac{|\mathbf{r}'|^5}{4} \right)d\mathbf{r}' \\
&= a\frac{|\mathbf{r}'|^2}{2} + b\frac{|\mathbf{r}'|^4}{4} + 3c\frac{|\mathbf{r}'|^6}{24}\Bigg|_{\mathbf{0}}^{\mathbf{r}} \\
&= a\frac{|\mathbf{r}|^2}{2} + b\frac{|\mathbf{r}|^4}{4} + c\frac{|\mathbf{r}|^6}{8}
\end{aligned}
\tag{40}
$$

Comparing Eq. (40) and Eq. (38) we see that the two are identical. Next we used object two to test a more complex non-linear LAC distribution. As before, we began by evaluating hardened cumulative LAC using the aggregate beam hardening correction. To start we computed the form of the unhardened cumulative LAC for object two:

$$
\mathrm{M}_{bh0}(\mathbf{r}) = \int_{\mathbf{r_0}}^{\mathbf{r}} \mu_{bh0}(\mathbf{r}')d\mathbf{r}' = \int_{\mathbf{0}}^{\mathbf{r}} \ d\mathbf{r}' = \sin\left(|\mathbf{r}|\right)
\tag{41}
$$

Plugging this result into Eq. (33) yields:

$$\mathrm{M}_{bh1}(\mathbf{r}) = g\left(\sin\left(|\mathbf{r}|\right)\right) = a\sin\left(|\mathbf{r}|\right) + b\sin^2\left(|\mathbf{r}|\right) + c\sin^3\left(|\mathbf{r}|\right) \tag{42}$$

Next we tested the in-place method for the second object, using the same procedure as before:

$$\begin{aligned}
&\mathrm{M}_{bh1}(\mathbf{r}) = \int_{\mathbf{0}}^{\mathbf{r}} \mu_{bh1}(\mathbf{r}')d\mathbf{r}' \\
&= \int_{\mathbf{0}}^{\mathbf{r}} k\left(\mathrm{M}_{bh0}(\mathbf{r}')\right)\mu_{bh0}(\mathbf{r}')d\mathbf{r}' \\
&= \int_{\mathbf{0}}^{\mathbf{r}} \left(a + 2b\sin\left(|\mathbf{r}'|\right) + 3c\sin^2\left(|\mathbf{r}'|\right)\right)\left(\cos\left(|\mathbf{r}'|\right)\right)d\mathbf{r}' \\
&= \int_{\mathbf{0}}^{\mathbf{r}} \left(a\cos\left(|\mathbf{r}'|\right) + 2b\sin\left(|\mathbf{r}'|\right)\cos\left(|\mathbf{r}'|\right) + 3c\sin^2\left(|\mathbf{r}'|\right)\cos\left(|\mathbf{r}'|\right)\right)d\mathbf{r}'
\end{aligned} \tag{43}$$

Evaluating the final integral yields:

$$\begin{aligned}
&\int_{\mathbf{0}}^{\mathbf{r}} \left(a\cos\left(|\mathbf{r}'|\right) + 2b\sin\left(|\mathbf{r}'|\right)\cos\left(|\mathbf{r}'|\right) + 3c\sin^2\left(|\mathbf{r}'|\right)\cos\left(|\mathbf{r}'|\right)\right)d\mathbf{r}' \\
&= a\sin\left(|\mathbf{r}|\right) - b\cos^2\left(|\mathbf{r}|\right) + c\sin^3\left(|\mathbf{r}|\right)\Big|_{\mathbf{0}}^{\mathbf{r}} \\
&= a\sin\left(|\mathbf{r}|\right) - b\cos^2\left(|\mathbf{r}|\right) + c\sin^3\left(|\mathbf{r}|\right) + b \\
&= a\sin\left(|\mathbf{r}|\right) + b\left(1 - \cos^2\left(|\mathbf{r}|\right)\right) + c\sin^3\left(|\mathbf{r}|\right) \\
&= a\sin\left(|\mathbf{r}|\right) + b\sin^2\left(|\mathbf{r}|\right) + c\sin^3\left(|\mathbf{r}|\right)
\end{aligned} \tag{44}$$

Once again the result from the in-place correction is identical to that of the aggregate correction.

Though this exercise does not represent a proof that the two methods are equal it does provide convincing evidence to that effect. In addition, empirical testing confirmed that DRRs corrected by the two methods were identical to within margins of round-off error. This being the case we chose the aggregate method as our standard of practice due to its superior computational efficiency over the in-place correction.

3.4.2 Scatter

Scatter refers to those photons reaching the detector that have been previously deflected by the object. For our purposes we must be able to remove and add the effects of scatter. To do this we used an in-house Monte Carlo algorithm capable of simulating scatter-only projections of an object.[60, 67] A Monte Carlo scatter projection can be represented as:

$$S_{MC}\left(\mathbf{r}_i\right) = A_2\sigma_{MC}\left(\mathbf{r}_i\right)R_{MC}\left(\mathbf{r}_i\right) \tag{45}$$

The Monte Carlo algorithm can also simulate a flood projection:

$$F_{MC}\left(\mathbf{r}_i\right) = A_2\Phi_{MC}\left(\mathbf{r}_i\right)R_{MC}\left(\mathbf{r}_i\right) \tag{46}$$

Dividing Eq. (45) by Eq. (46) produces:

$$\frac{S_{MC}\left(\mathbf{r}_i\right)}{F_{MC}\left(\mathbf{r}_i\right)}=\frac{\sigma_{MC}\left(\mathbf{r}_i\right)}{\Phi_{MC}\left(\mathbf{r}_i\right)} \tag{47}$$

Making use of a previously acquired CBCT flood projection we multiply Eq. (47) by Eq. (10) to yield:

$$F_{CBCT}\left(\mathbf{r}_i\right)\frac{S_{MC}\left(\mathbf{r}_i\right)}{F_{MC}\left(\mathbf{r}_i\right)}=A_0\Phi_{CBCT}\left(\mathbf{r}_i\right)R_{CBCT}\left(\mathbf{r}_i\right)\frac{\sigma_{MC}\left(\mathbf{r}_i\right)}{\Phi_{MC}\left(\mathbf{r}_i\right)} \tag{48}$$

Finally, multiplying Eq. by the inverse of the output factor from Eq. (12), $\frac{A_1}{A_0}$, and assuming that $\Phi_{MC}\left(\mathbf{r}_i\right)$ is approximately equal to $\Phi_{CBCT}\left(\mathbf{r}_i\right)$, we obtain an estimate of the scatter described in Eq. (9):

$$\tilde{S}_{CBCT}\left(\mathbf{r}_i\right)=A_1\sigma_{MC}\left(\mathbf{r}_i\right)R_{CBCT}\left(\mathbf{r}_i\right) \tag{49}$$

This scatter projection could be added or subtracted from another projection to add or remove scatter effects.

3.4.3 Veiling glare

Veiling glare refers to the leakage of signal from its detector pixel of origin to surrounding detector pixels. We modeled this effect using a point spread function (PSF) that characterized this signal spread over an area up to 10 cm distant from the signal origin. The veiling glare PSF was modeled by fitting a parametric function to the edge response function obtained from half-blocked projection images, which were obtained by acquiring projections with a lead plate positioned over half of the detector surface. Measurements and fits were performed by a group at our institution using the same hardware as used in our DRR experiments. The parametric function was a weighted sum of Gaussian and exponential terms used to describe the short and long range effects of veiling glare, respectively. The veiling glare effect was introduced to a projection by convolving it with the PSF:

$$P_{vg1} = P_{vg0} \otimes PSF \tag{50}$$

where the subscript "vg1" indicates that veiling glare is present, "vg0" indicates that veiling glare is absent, and $\otimes$ is the convolution operator. Likewise veiling glare was removed by deconvolving the PSF:

$$P_{vg0} = P_{vg1} \otimes^{-1} PSF \tag{51}$$

where $\otimes^{-1}$ denotes the deconvolution operator.

3.5 CT number to LAC conversion function

In order to evaluate the integral in Eq. (11) we must convert the CT numbers to units of linear attenuation coefficient. This conversion depends on the energy spectrum (and thus on the kVp) at which one wishes to emulate a CBCT projection. Our method is the same general procedure as that used by Moore et al.[48] We began by collecting cylindrical plugs composed of different materials. The plugs were inserts of known dimensions taken from a Gammex electron density calibration phantom. We first acquired a FBCT of the phantom with plugs in place to determine the CT number of each plug. Next we separately captured CBCT projections at a given technique for each plug. After removing beam hardening, scatter, and veiling glare from these projections, LAC values for each plug material were calculated using Eqs. (8) and (14). Once this was accomplished, a 2nd order polynomial fit to the LACs plotted as a function of CT number yielded our desired CT number to LAC conversion function, $f_{LAC}\left[CT\left(\mathbf{r}\right)\right]$, where $CT\left(\mathbf{r}\right)$ is the CT number at location $\mathbf{r}$. First, veiling glare was removed from the plug projections:

$$P_{CBCT,plug,vg0} = P_{CBCT,plug} \otimes^{-1} PSF \qquad (52)$$

Next scatter was subtracted:

$$P_{CBCT,plug,vg0sc0}\left(\mathbf{r}_i\right)=P_{CBCT,plug,vg0}\left(\mathbf{r}_i\right)-\tilde{S}_{CBCT,plug}\left(\mathbf{r}_i\right) \tag{53}$$

Then the projection was reduced to cumulative LAC:

$$\mathrm{M}_{CBCT,plug,vg0sc0}\left(\mathbf{r}_i\right)=-\ln\left(\frac{A_0}{A_1}\frac{P_{CBCT}\left(\mathbf{r}_i\right)}{F_{CBCT}\left(\mathbf{r}_i\right)}\right) \tag{54}$$

and a correction was applied to remove beam hardening:

$$\mathrm{M}_{CBCT,plug,vg0sc0bh0}\left(\mathbf{r}_i\right)=h(\mathrm{M}_{CBCT,plug,vg0sc0}\left(\mathbf{r}_i\right)) \tag{55}$$

Finally, LAC was calculated from the cumulative unhardened LAC:

$$\mu_{CBCT,plug,vg0sc0bh0}=\frac{\mathrm{M}_{CBCT,plug,vg0sc0bh0}\left(\mathbf{r}_c\right)}{t} \tag{56}$$

where t is the thickness of the plugs, 7 cm, and $\mathbf{r}_c$ is the coordinate of the detector center over which the plugs were positioned during imaging. Plotting $\mu_{CBCT,plug,vg0sc0bh0}$ as a function of CT number and fitting with a polynomial yielded our conversion function, $f_{LAC}\left[CT\left(\mathbf{r}\right)\right]$. (As previously mentioned, this conversion function is specific to the x-ray tube technique at which the CBCT projections were made.) To illustrate the effect of the scatter, beam hardening, and veiling glare corrections on the final conversion function, the data points used to build the function and their polynomial fits are plotted in Figure 9 for both uncorrected and fully corrected data.

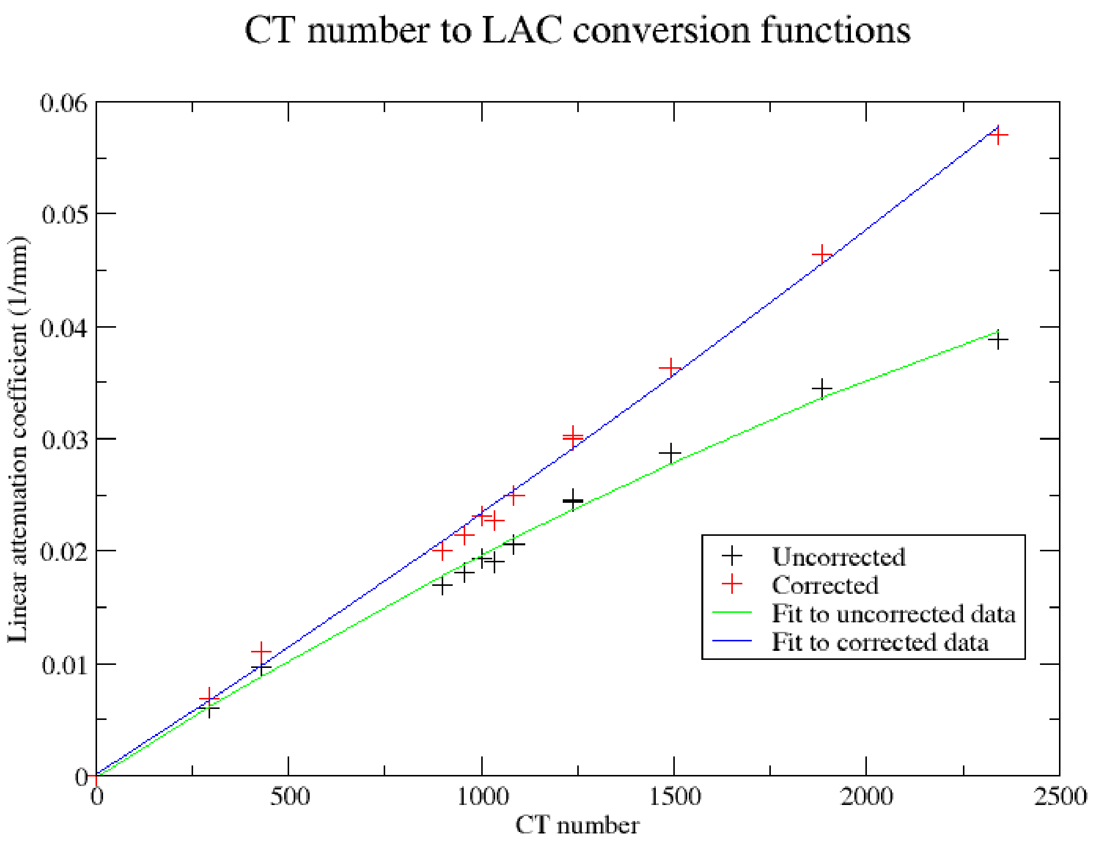


Figure 9: Plot of LAC vs. CT number data points and 2nd order polynomial fits generated from attenuation measurements of different plug materials. Attenuation measurements are either uncorrected or fully corrected for scatter, beam hardening, and veiling glare effects.

3.6 DRR workflow

The workflow for producing a fully corrected DRR was similar to the reverse of that for processing the plug projections. First, a raw DRR was produced using an x-ray volume rendering routine. We used a ray-tracing based approach that approximated the line integral in Eq. (11) as the following:

$$\int_{\mathbf{r}_0}^{\mathbf{r}_i} \mu_{bh0}\left(\mathbf{r}\right) d\mathbf{r} \approx \mathrm{M}_{DRR,bh0}\left(\mathbf{r}_i\right) = \sum_{k=0}^{K} f_{LAC}[CT(\mathbf{r}_0 + k\mathbf{\Delta r})]\Delta r \tag{57}$$

Where K is the number of steps that fit between $\mathbf{r}_0$ and $\mathbf{r}_i$, $\mathbf{\Delta r}$ is the ray-step vector, and Δr is the ray-step length. The CT is sampled at location $\mathbf{r}_0 + k\mathbf{\Delta r}$ using tri-linear interpolation. This algorithm is fully supported by the hardware implementation of ray tracing and interpolation on a GPU platform.[57] After the raw DRR was generated, beam hardening effects were introduced:

$$\mathrm{M}_{DRR,bh1}\left(\mathbf{r}_i\right) = g(\mathrm{M}_{DRR,bh0}\left(\mathbf{r}_i\right)) \tag{58}$$

Using this result and Eqs. (7) and (8) a beam hardening corrected projection image was produced:

$$P_{DRR,bh1}\left(\mathbf{r}_i\right) = \frac{A_1}{A_0} F_{CBCT}\left(\mathbf{r}_i\right) \exp[-\mathrm{M}_{DRR,bh1}\left(\mathbf{r}_i\right)] \tag{59}$$

The flood projection allowed us to reproduce the CBCT source fluence distribution in the DRR, and the output factor allowed us to recover the beam output of the target projection. This image was then corrected for scatter:

$$P_{DRR,bh1sc1}\left(\mathbf{r}_i\right) = P_{DRR,bh1}\left(\mathbf{r}_i\right) + \tilde{S}_{CBCT,obj}\left(\mathbf{r}_i\right) \tag{60}$$

where the scatter projection was calculated using the method of section 3.4.2. Lastly, the veiling glare effect was introduced and the DRR was fully corrected:

$$P_{DRR,bh1sc1vg1} = P_{DRR,bh1sc1} \otimes PSF \tag{61}$$

3.7 Projection acquisition

For verification, DRRs were computed through object FBCTs and compared to actual CBCT projections of the same object. Two different objects were tested in this study, a Catphan™ CT phantom (The Phantom Laboratory, Salem NY), and an actual patient. In each case projections were acquired at 20 ms, 20 mA, and 125 kVp. Catphan™ projections were acquired in full-fan mode while patient projections were captured in half-fan mode, each using their respective aluminum bow-tie filters. All images were acquired on the OBI™ CBCT system (Varian Medical Systems, Palo Alto, CA) which utilizes a G242 kilovoltage x-ray source and a PAXSCAN 4030CB flat panel detector mounted to the CLINAC. In addition a 10:1 antiscatter grid is positioned on top of the detector system, which is protected by a 0.5 mm aluminum cover.

3.8 Results

We tested the accuracy of the DRR algorithm using comparisons between our DRRs and the actual CBCT projections they were intended to match. We calculated two different sets of DRRs – one set using just the measured linear attenuation through the medium (uncorrected) and a second set using the full algorithm with corrections for beam hardening, scatter, and veiling glare in both the LAC measurements and the DRR projections. This displays the relative importance of the corrections as a whole. We do not provide a complete breakdown of the effect of each individual effect, however we do provide a general quantitative analysis as discussed in the latter stages of Section 3.9.1. We used two different objects as test cases - a Catphan™ CT phantom, and an actual patient. For both sets of DRRs, a CBCT flood projection was used to measure the output and source fluence distribution of the CBCT system when the raw DRRs were post-processed. We emphasize that these DRRs are absolute calculations of detected signal intensity, without any normalization factor or other numerical adjustment to match to the CBCT projections. For qualitative validation we focus on visual comparison, comparison of pixel intensity profiles, and comparison of image intensity histograms. For quantitative validation we correlate the DRR pixel intensities with the actual projection pixel intensities.

3.8.1 Visual comparison

Figure 10 displays projections cast through the CT phantom, and Figure 11 and Figure 12 show projections cast through an actual patient at different angles. The images in each figure are set to the same window and level for comparison. In Figure 10 the bottom of each DRR was contaminated by metal artifacts in the Catphan™ FBCT. This has no bearing on the efficacy of our method and is subsequently ignored.

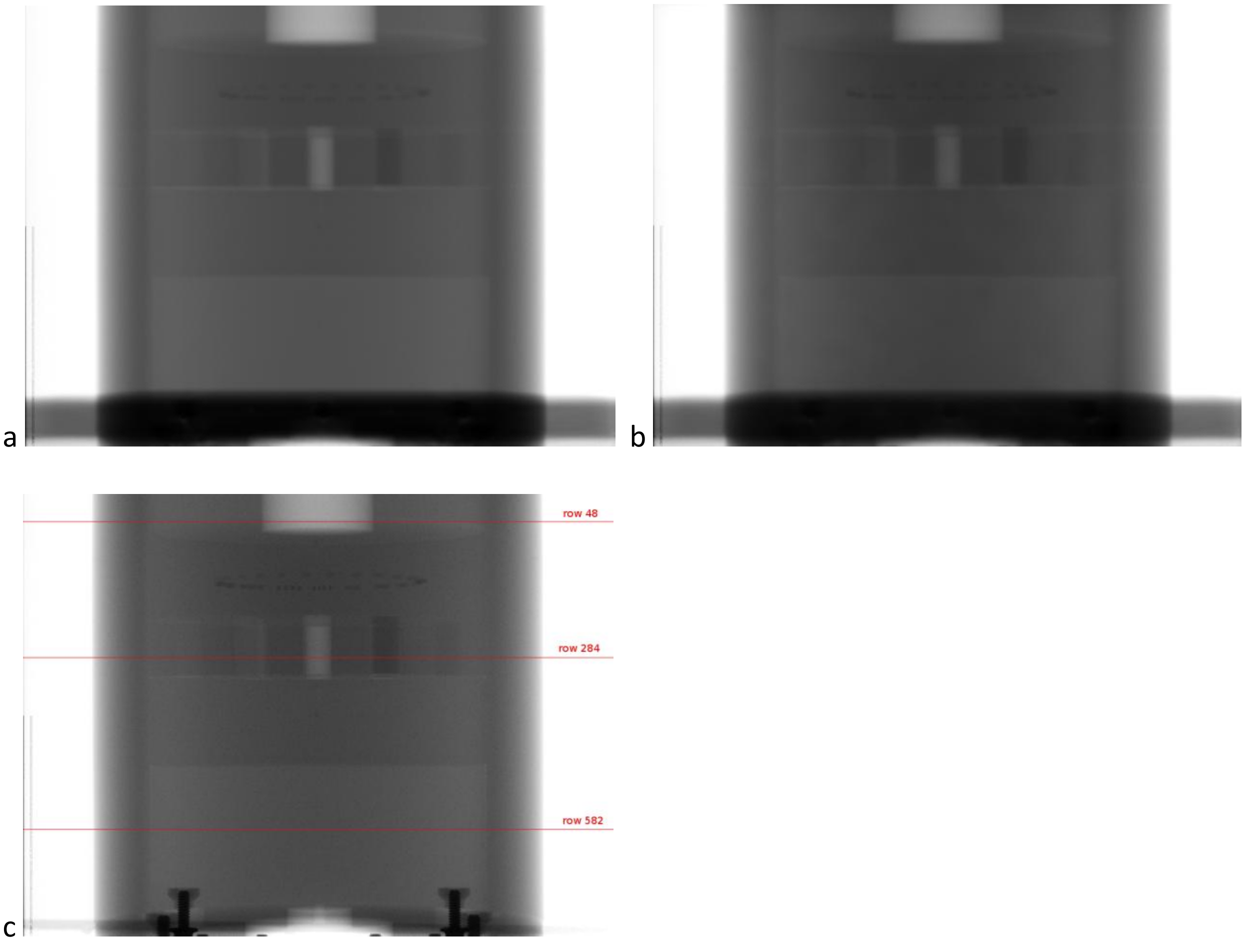


Figure 10: Projections through a Catphan™ CT phantom at a projection angle of 0 degrees. a) uncorrected DRR, b) fully corrected DRR, c) actual CBCT projection. Red lines in c) indicate locations where line profiles are sampled.

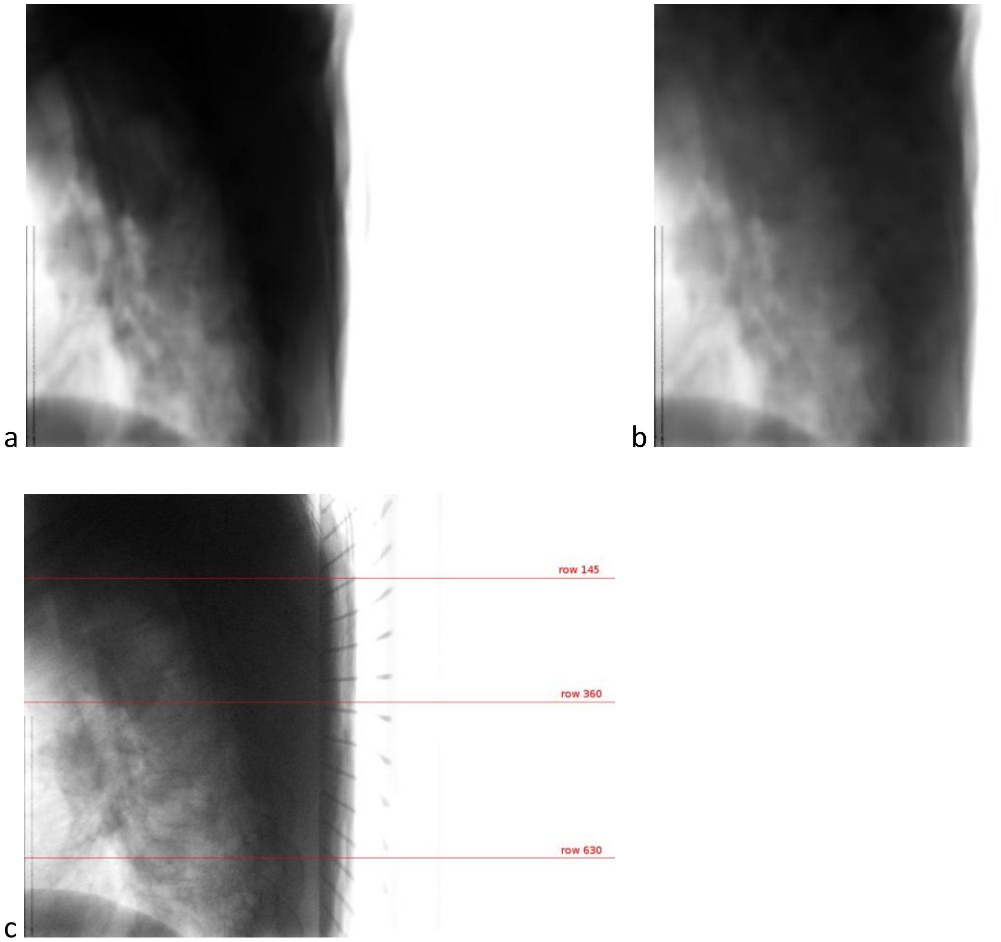


Figure 11: Projections through a patient at a projection angle of 92.67 degrees. a) uncorrected DRR, b) fully corrected DRR, c) actual CBCT projection. Red lines in c) indicate locations where line profiles are sampled.

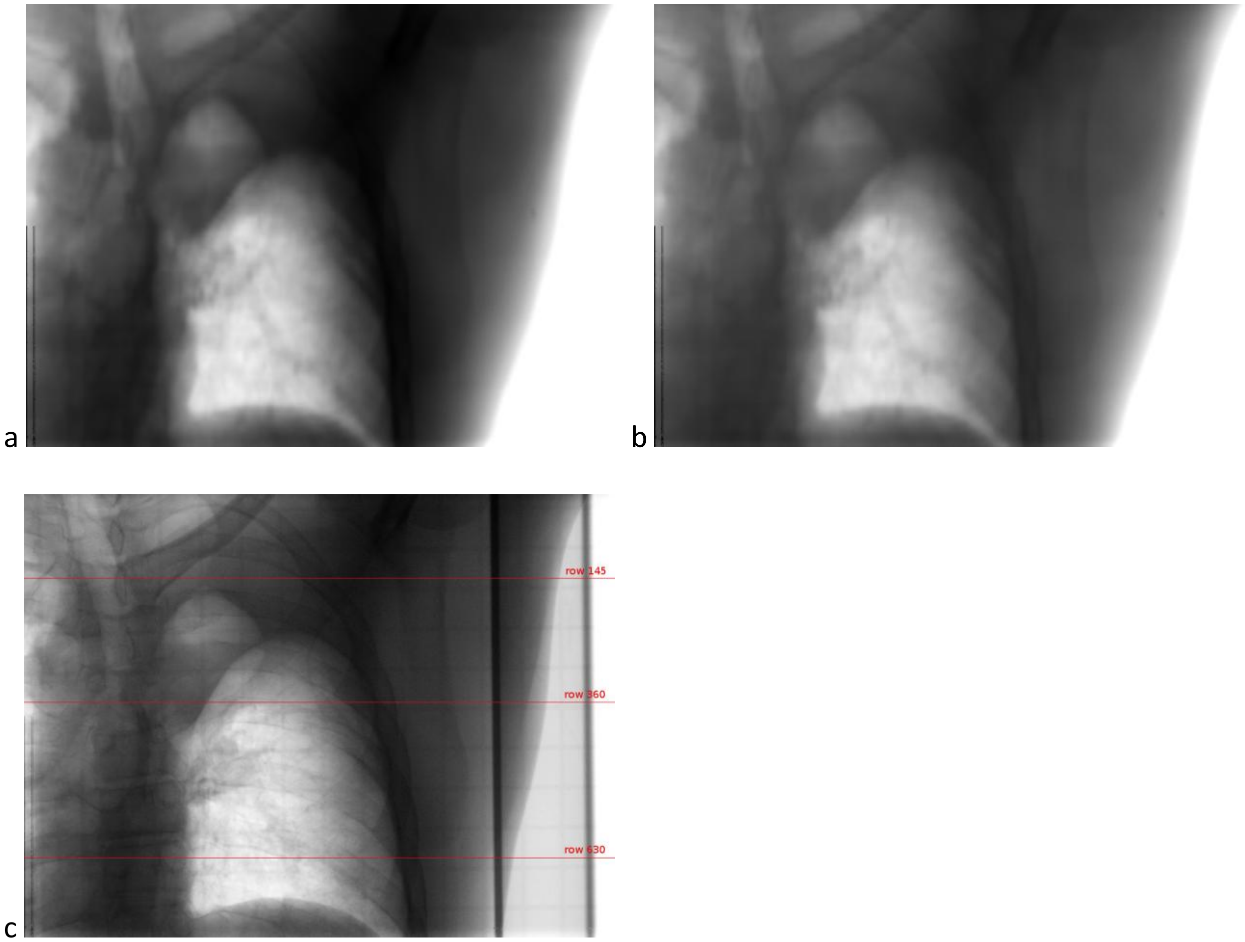


Figure 12: Projections through a patient at a projection angle of 178.93 degrees. a) uncorrected DRR, b) fully corrected DRR, c) actual CBCT projection. Red lines in c) indicate locations where line profiles are sampled.

The goal of this evaluation is not to give an expert analysis of DRR image quality, as might be provided by a radiologist, but to make some basic visual observations on the output of our algorithm. We begin by noting that both uncorrected and corrected DRRs bear a close resemblance to the CBCT projections, with most obvious features clearly reproduced. In general the uncorrected and corrected DRRs look similar; both look less crisp than the CBCT projections due to the coarse resolution of the FBCT relative to the size of the detector pixels. This effect causes the loss of some detail in both corrected and uncorrected DRRs, such as the spine and vertebral disks in Figure 11.

However, in the patient projections in particular the corrected DRRs show more realistic contrast than the uncorrected DRRs with respect to the actual CBCT projections. In Figure 11a, the uncorrected DRR, the upper right side of the patient is essentially saturated in black while the actual projection, Figure 11c shows considerable contrast. In the corrected DRR, Figure 11b, some of this lost contrast has been visibly recovered. Similarly, Figure 12 shows areas, particularly at the top of the image above the crest of the lung, where the uncorrected DRR is saturated in black, while both the corrected DRR and actual CBCT projection maintain visible contrast.

3.8.2 Pixel intensity profiles

Figure 13, Figure 14, and Figure 15 show pixel intensity profiles sampled along the horizontal direction from each set of uncorrected, fully corrected, and actual projections. The locations at which these profiles were sampled are marked on the actual CBCT projections in Figure 10, Figure 11, and Figure 12. For the most part, both the curves for the fully corrected and uncorrected DRRs track with the actual CBCT projection, although the uncorrected curves tend to consistently underestimate the pixel intensity. This is particularly visible in Figure 15. Figure 14 and Figure 15 compare intensity profiles for real patient data. The FBCT through which the DRRs were cast was taken on a different day than the CBCT projections to which the DRRs are compared. For this reason it is impossible to perfectly match patient CBCT projections and DRRs, as the patient setup will invariably differ and some internal anatomic deformation will be unavoidable. In addition, the CBCT and FBCT systems have different couches, and in

Figure 11 and Figure 12 couch structures that are plainly visible in the actual CBCT projections are absent from the DRRs. Areas in the line profiles of Figure 14 and Figure 15 where the DRRs and CBCT projections radically disagree reflect these differences.

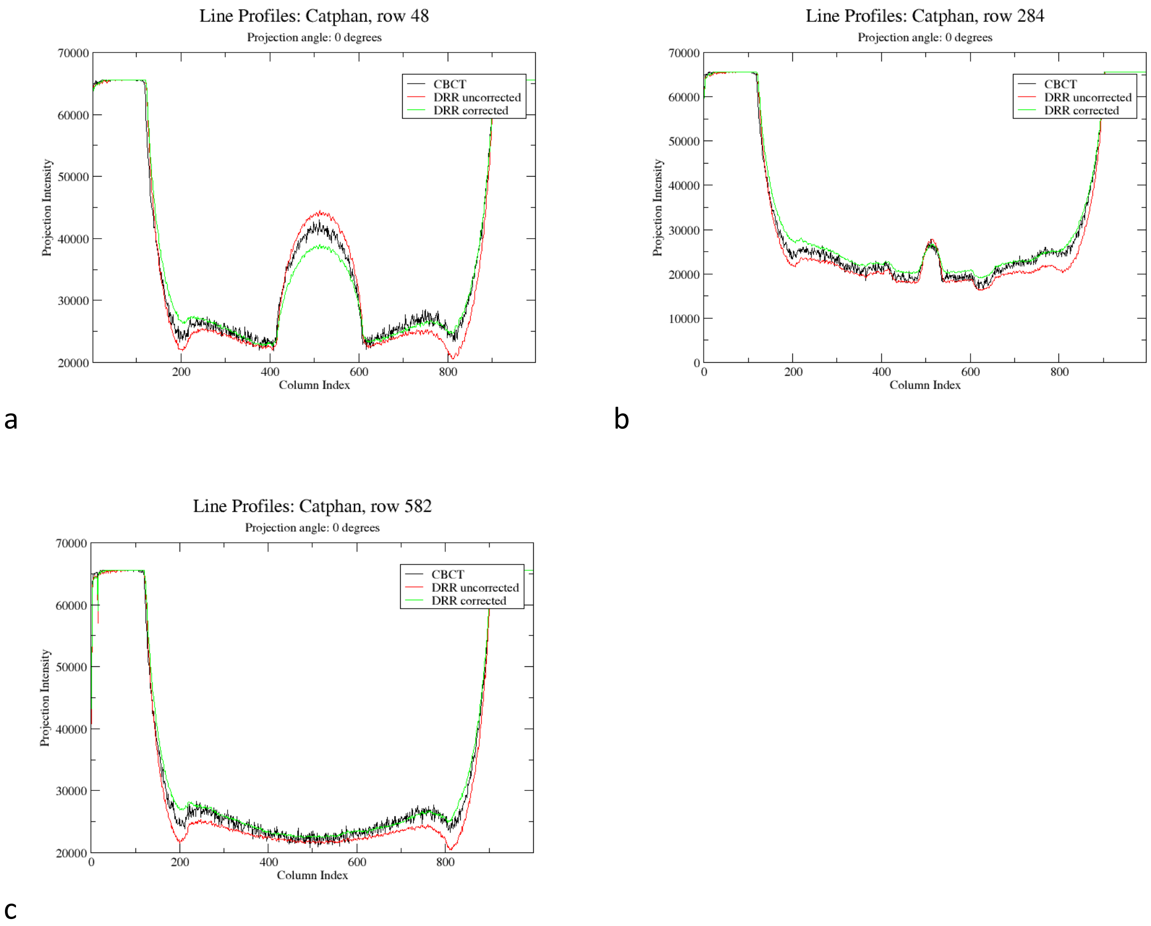


Figure 13: Comparison of line profiles sampled from an actual CBCT projection, a DRR uncorrected for beam hardening, scatter, or veiling glare, and a fully corrected DRR. Projections are taken through a Catphan™ CT phantom at a projection angle of 0 degrees. Line profiles are sampled at a) row 48, b) row 284, and c) row 582.

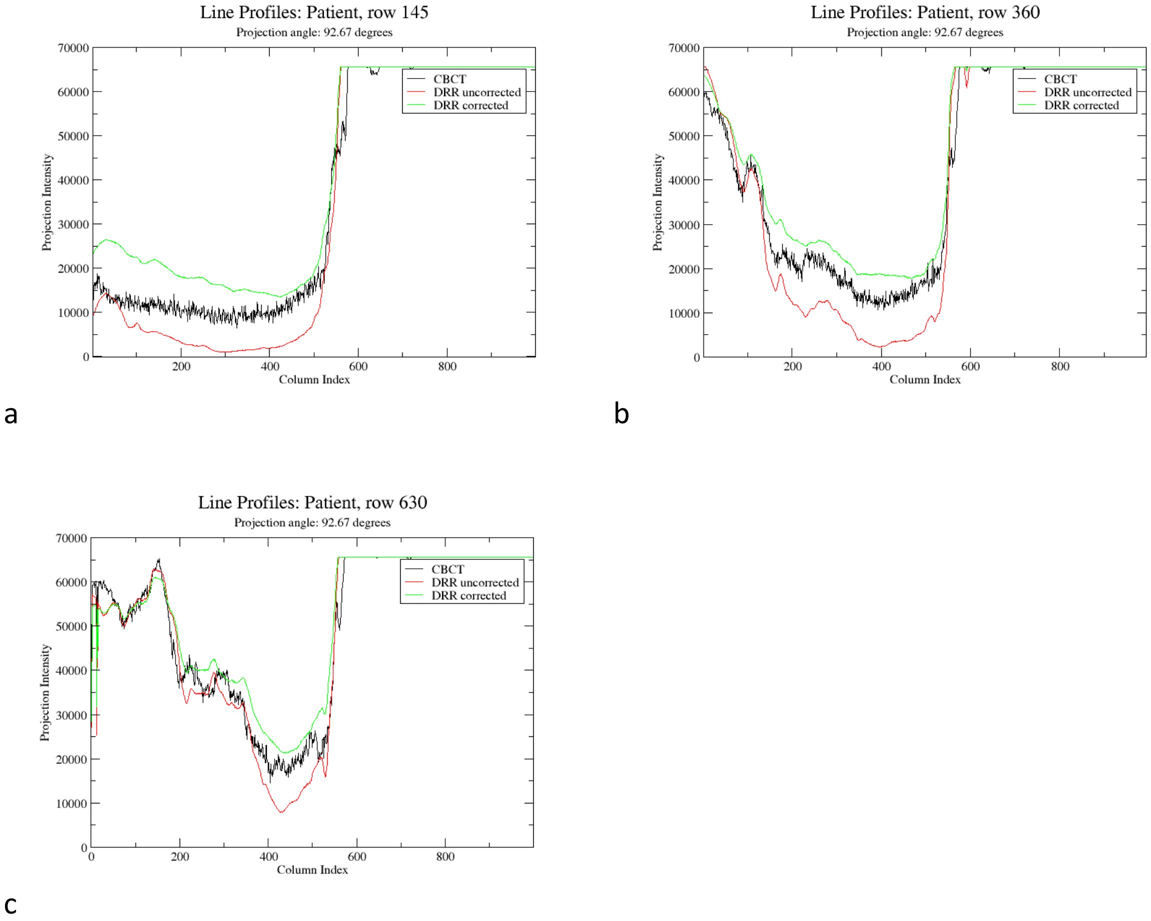


Figure 14: Comparison of line profiles sampled from an actual CBCT projection, a DRR uncorrected for beam hardening, scatter, or veiling glare, and a fully corrected DRR. Projections are taken through an actual patient at a projection angle of 92.67 degrees. Line profiles are sampled at a) row 145, b) 360, and c) row 630.

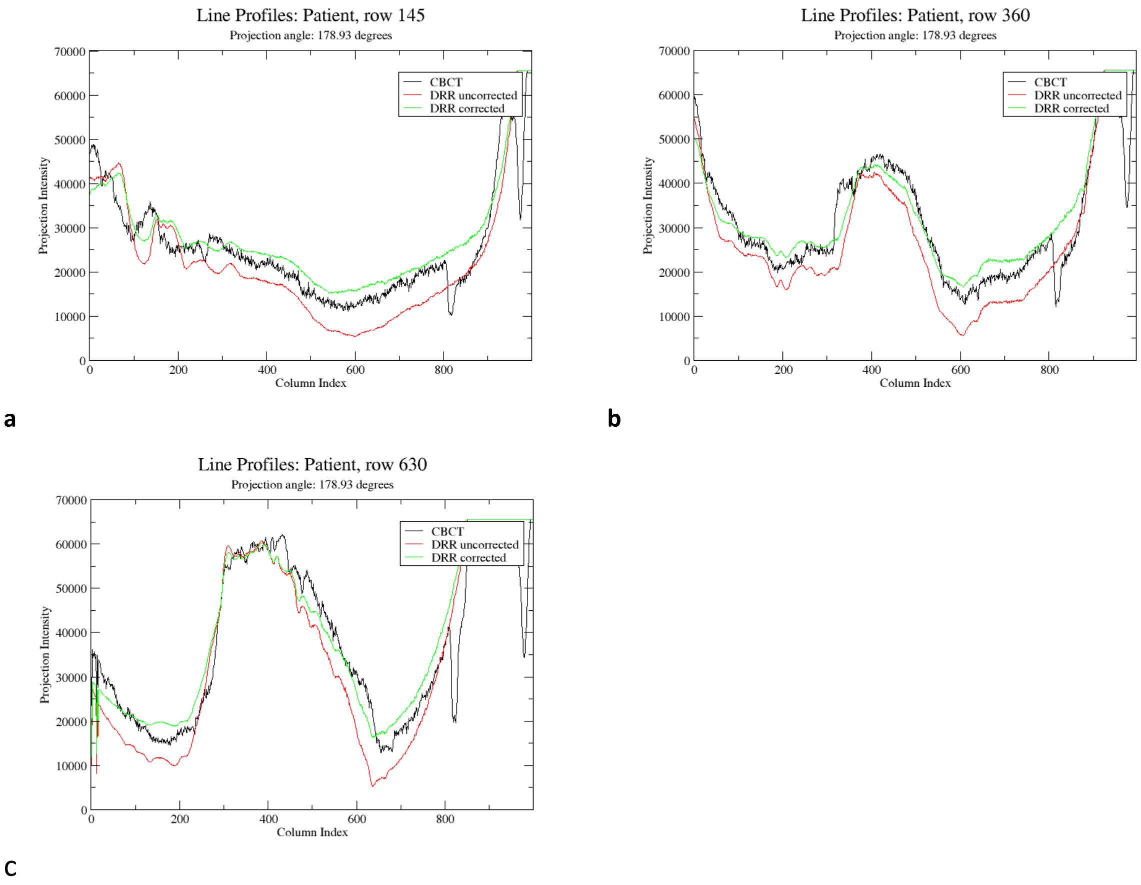


Figure 15: Comparison of line profiles sampled from an actual CBCT projection, a DRR uncorrected for beam hardening, scatter, or veiling glare, and a fully corrected DRR. Projections are taken through an actual patient at a projection angle of 178.93 degrees. Line profiles are sampled at a) row 145, b) 360, and c) row 630.

3.8.3 Intensity histograms

Figure 16 plots intensity histograms of the actual CBCT projections alongside histograms of uncorrected and fully corrected DRRs. Areas contaminated by the treatment couch, in the case of the patient projections, or metal artifact, in the case of the Catphan™ projections, were excluded from the histograms. For all cases the fully corrected DRR histograms exhibit a closer match to the CBCT projections than the uncorrected DRRs. Computing the Dice coefficient between CBCT and DRR histograms yields a measure of the overlap between the two distributions. The Dice coefficient ranges from 0 (no overlap) to 1 (perfect overlap) and is defined as:

$$D = \frac{2(H_{CBCT} \cap H_{DRR})}{H_{CBCT} + H_{DRR}} \tag{62}$$

For the Catphan™ histograms in Figure 16c the Dice coefficients for the fully corrected and uncorrected DRR histograms when measured against the CBCT histogram were 0.87 and 0.76 respectively. For the patient histograms in Figure 16b corrected and uncorrected values were 0.84 and 0.80, and for the patient histograms in Figure 16a corrected and uncorrected values were 0.73 and 0.59.

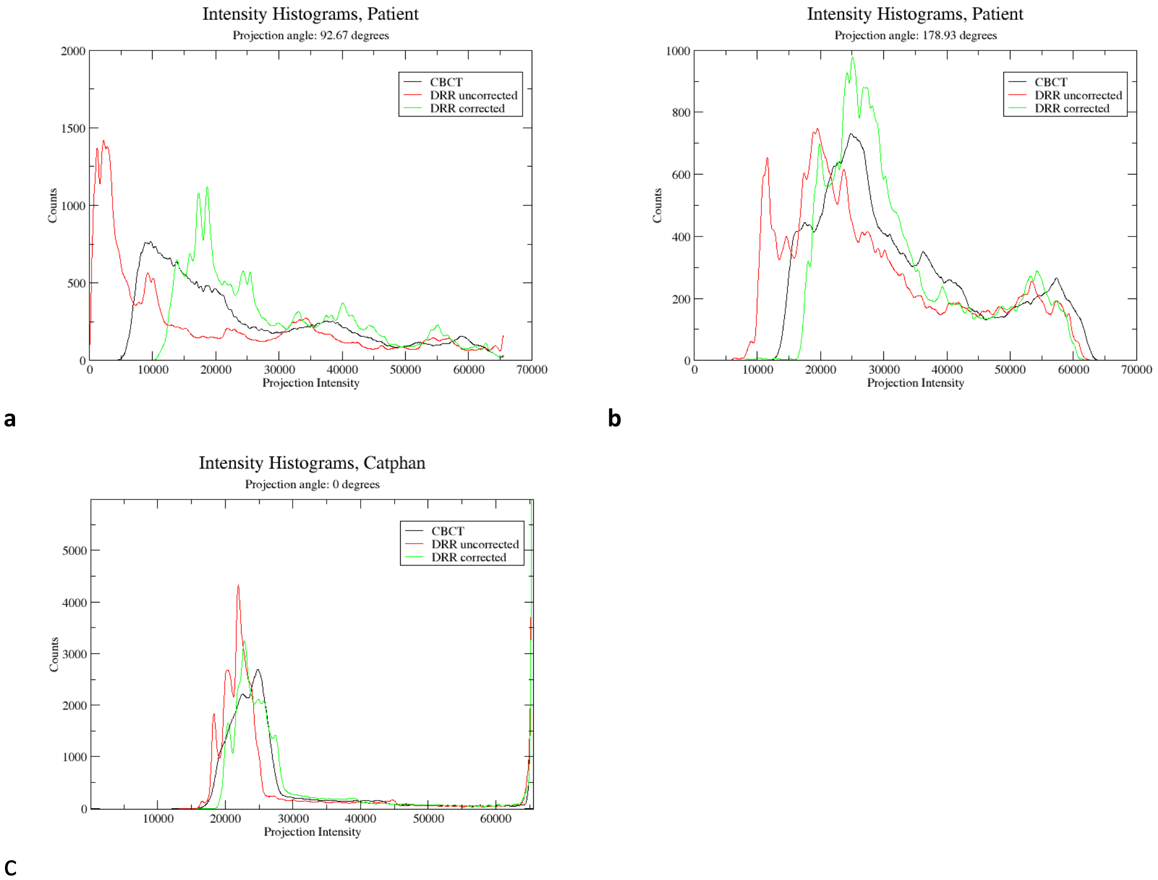
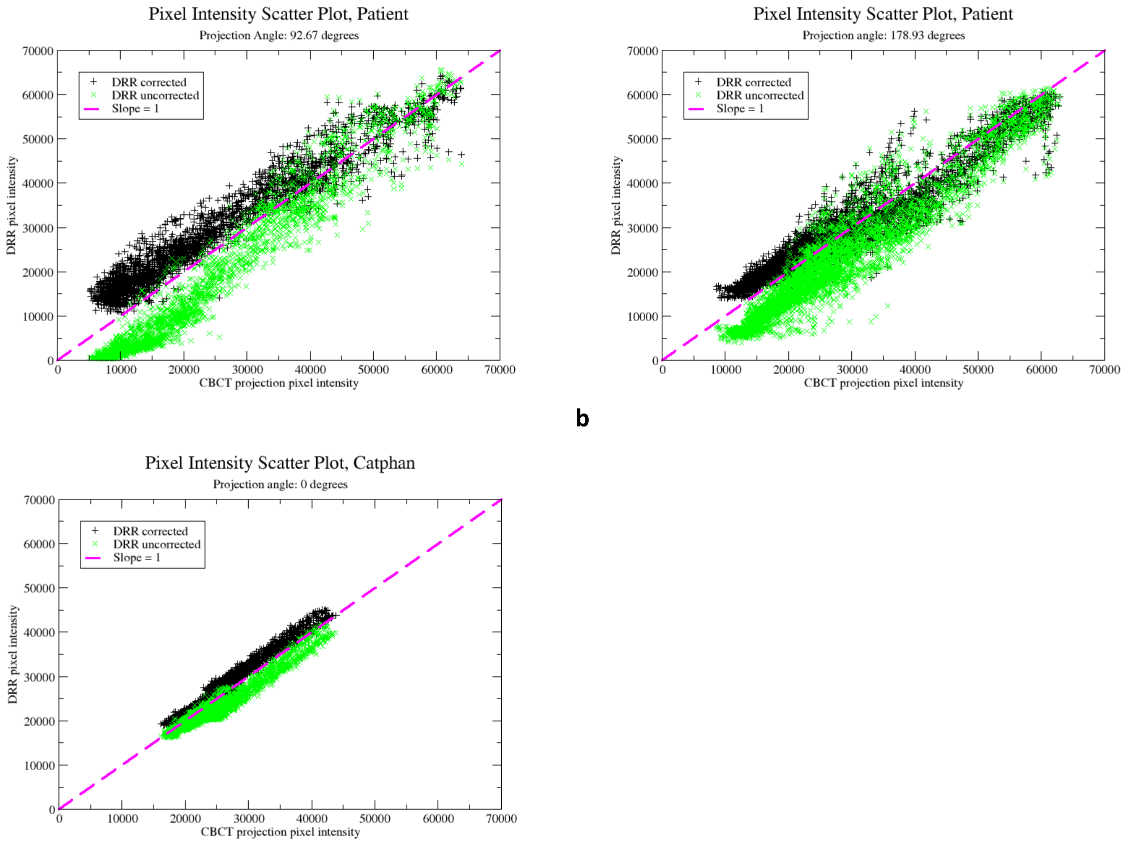


Figure 16: Histograms comparing the intensity distributions of actual CBCT projections, DRRs uncorrected for beam hardening, scatter, or veiling glare, and fully corrected DRRs. Comparisons are for projections cast through a) a patient at a projection angle of 92.67 degrees, b) a patient at a projection angle of 178.93 degrees, and c) a Catphan™ CT phantom at a projection angle of 0 degrees. In a) and b) areas of the projections contaminated by the treatment couch are omitted from the distributions. In c) areas of the projections contaminated by metal artifacts in the FBCT image are omitted from the distributions.

3.8.4 DRR-to-projection intensity correlation

The scatter plots in Figure 17 display the correspondence between intensity values recorded by actual detector pixels and their counterparts in both corrected and uncorrected DRRs. A perfect match of DRR to projection is represented by the (pink) line of slope one. Intensity mismatch is represented by the scatter around this line as well as any systematic offset in slope or intercept.

We obtain a quantitative measure of these correlations by computing the mean percent disagreement between CBCT projection and DRR pixel intensities averaged over all pixels, along with the standard deviation about this mean. The mean value describes the extent to which a scatter distribution is systematically offset from the line of perfect agreement, while the standard deviation describes how tightly grouped the distribution is. A positive mean difference indicates that on average the DRR overestimated the true pixel intensity, while a negative difference signifies that on average the DRR underestimated the true pixel intensity. For the patient data plotted in Figure 17a the percent disagreements for the corrected and uncorrected DRRs were 34.22 ± 33.88 % and -46.61 ± 41.19 % respectively. For the patient data plotted in Figure 17b percent disagreements for corrected and uncorrected DRRs were 6.52 ± 6.46 % and -22.09 ± 21.87 % respectively. For the Catphan™ data plotted in Figure 17c percent disagreements for corrected and uncorrected DRRs were 4.01 ± 3.97 % and -6.58 ± 6.52 % respectively.



a

b

c

Figure 17: Scatter plots showing the correspondence between pixel intensities in actual CBCT projections and those of both corrected (black) and uncorrected (green) DRRs. The pink line indicates the track along which perfect correspondence is achieved. Scatter plots are plotted for projections cast through a) a patient at a projection angle of 92.67 degrees, b) a patient at a projection angle of 178.93 degrees, and c) a Catphan™ CT phantom at a projection angle of 0 degrees.

3.8.5 Beam hardening comparison

Figure 18 displays intensity profiles comparing the measured and calculated beam hardening corrections. For reference, the actual CBCT projection intensity profiles are also included. In the center of the profiles the two beam hardening methods exhibit the best agreement, while the two curves begin to diverge closer to the edges. This is expected, as the data for the measured correction is collected at the center of the detector. At the sides of the detector the calculated method, which accounts for the heterogeneous bowtie filter thickness, increasingly deviates from the measurement based method, which does not.

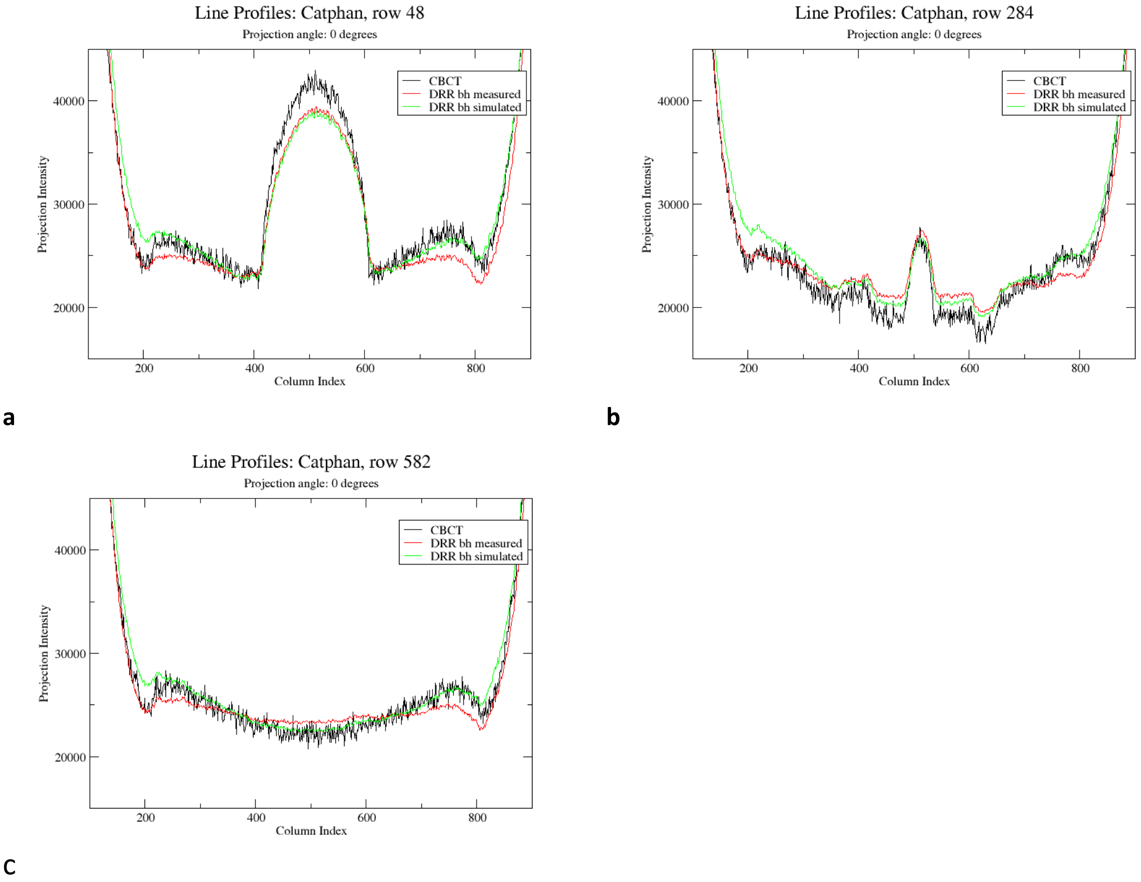


Figure 18: Line profiles sampled from projections cast through Catphan CT phantom for comparison of two different beam hardening correction methods. Line profiles are sampled at a) row 48, b) row 284, and c) row 582.

3.9 Discussion

Before discussing our results it is important to reiterate our motivation for developing a DRR algorithm from first principles for the purpose of FIPM based deformable 2D/3D registration within the context of our FIPM-PCA 4DCBCT algorithm. As our goal is to minimize the discrepancy between DRRs and CBCT projections in absolute units we must have a DRR algorithm that produces realistic image intensities. However, our DRR algorithm must also be unbiased, i.e. it must not conceal or misrepresent differences between DRRs and CBCT projections that correspond to actual deformation features. Any scheme that artificially scales or adjusts the DRRs to the actual CBCT projections runs the risk of compromising such features. By grounding our algorithm in the fundamental physics of image formation our method meets the dual goals of closely reproducing CBCT projection intensities while avoiding bias that could affect the integrity of our image registration process.

3.9.1 Qualitative and quantitative algorithm assessments

In testing our algorithm we calculated DRRs both with and without physics corrections and compared the results to actual CBCT projections. This allowed us to compare our DRRs to a "ground truth" in order to determine the absolute accuracy achieved by our algorithm, as well as to examine the relative effect of our physics based corrections. In our assessment of absolute DRR and CBCT projection agreement there were two potentially confounding factors. The DRRs were manually aligned to the CBCT system geometry, with small misalignments possibly resulting in large discrepancies between pixel intensities. Additionally, for patient data some deformation inevitably occurred in the time between FBCT and CBCT projection acquisition. This likely caused differences between DRRs and CBCT projections unrelated to the DRR algorithm. Due to these factors our quantitative assessments likely indicate an upper bound on the accuracy we can expect from our algorithm.

With respect to our visual comparison of DRRs to the actual CBCT projections, both corrected and uncorrected methods appear to produce images that look similar to each other and to the actual projection. Differences mainly appear in low intensity areas where the uncorrected DRR became saturated in black while the corrected DRR and actual projection maintained some visible contrast. The line plots in Figure 13, Figure 14, and Figure 15 show that in general both corrected and uncorrected DRRs follow the general shape of the actual CBCT projection. It is observed that the uncorrected DRRs tend to underestimate the true CBCT projection intensity, while the corrected DRRs slightly overestimate the CBCT projection intensity. A similar effect is observed in the intensity histogram plots of Figure 16, with histograms of corrected DRRs shifted to the right and those of uncorrected DRRs shifted to the left. In each of the three cases shown the Dice coefficient for the corrected DRRs was larger than that of the uncorrected DRRs, indicating that the physics corrections consistently provide a quantifiable improvement in DRR quality. Assessment of the scatter plots in Figure 17 supports this conclusion. Scatter distributions from corrected DRRs all fall somewhat above the line of perfect agreement indicating overestimation of pixel intensity, while distributions for uncorrected DRRs consistently fall below the line, indicating underestimation. In all cases the mean error for the corrected DRRs is less than the uncorrected DRRs. The same goes for the standard deviations of the errors, indicating a tighter distribution for the corrected DRRs. The smallest error and error standard deviation of $4.01 \pm 3.97\ \%$ occurs for the corrected DRR cast through the Catphan™. We expect the Catphan™ to give us the smallest errors since it is a rigid object and will thus avoid the deformations that can occur in patient images. For this reason it is our best estimate of the true error of our algorithm. We choose to quantitatively evaluate our algorithm using percent mean error of the absolute DRR intensities since this is the most relevant quantity with regards to the SSD merit function. This measure is a broad indicator, in the sense that a small percent error in absolute intensity will usually be indicative of a small error in relative metrics such as contrast, while the reverse may not be true.

In a previous study we quantitatively examined the effect of intensity mismatch on registration quality by introducing an artificial, non-linear variation between projection and DRR intensities.[38] The magnitude of disagreement required to induce appreciable degradation of the registration results was in the neighborhood of 1 %, which is below our present best value of 4.01 %. However, looking at Figure 17c, the variations in our Catphan™ DRR appear to follow a straight line, and therefore may have less of an effect on the FIPM algorithm than those previously considered. We are now in a position to repeat our test of the sensitivity of 2D/3D FIPM to DRR accuracy using actual rather than contrived contrast mismatch.

The corrected DRRs consistently overestimate true projection intensities, while our uncorrected DRRs consistently underestimate them. In the uncorrected DRRs the effects of scatter, beam hardening, and veiling glare are included in the CT number to LAC conversion function to the degree that they are present in the plug projections. If the sum of those effects is greater in the plug projections than in the actual object projections the DRR will overestimate the true CBCT projection intensity. If the reverse is true the DRR will underestimate the true CBCT projection intensity. Averaged over all 12 plug materials, the absence of the scatter, beam hardening, and veiling glare effects reduces the image intensity within the plug shadows by 9, 5, and 19 % respectively when each is individually removed from the plug projections. Likewise, when scatter, beam hardening, and veiling glare effects are not accounted for in DRR post-processing the average DRR pixel intensities are decreased by 32, 19, and 1 % respectively, using the Catphan™ DRRs as an example. From this we see that the veiling glare effect is overrepresented in the plug projections as compared to the object projections, while the beam hardening and scatter effects are underrepresented. This suggests that the two large underestimations with regards to scatter and beam hardening outweigh the overestimation of veiling glare, leading to a net underestimation of the total effect, which in turn produces the observed underestimation of CBCT projection intensity in the uncorrected DRRs. Analyzing the source of the consistent overestimation of CBCT pixel intensity in the corrected DRRs is somewhat trickier, since our whole effort has been to accurately mimic the physics behind image formation. The most likely source is an overly aggressive beam hardening correction, which assumes that the object is composed of a single material (solid water in the case of the measured correction, liquid water in the case of the calculated correction). Using our calculation based framework we can test many different materials and examine the beam hardening functions they produce. For example, using polystyrene or adipose as the test material produces a less aggressive beam hardening function than liquid water, using muscle or lung produces a nearly identical correction, and using bone produces a more aggressive correction.

In addition to the systematic intensity offset present in the DRRs, the scatter distributions in Figure 17 are also spread about the mean. This is likely a result of geometric misalignment of the FBCT scan, deformation of the object between FBCT and CBCT projection acquisition (if applicable), and the volume averaging effect caused by the relatively coarse resolution of the FBCT with respect to the size of the detector pixels. Indeed, the Catphan™ scatter plot, where the second of these factors is absent, shows the tightest distribution.

Of additional interest is our comparison of the calculated and measured beam hardening corrections. While the measurement approach has the benefit of using the actual x-ray technique and detector used to form the CBCT projections, the calculation method intrinsically avoids scatter and veiling glare, is uncontaminated by noise, and can be calculated separately for each detector pixel to account for the non-uniform thickness of the bowtie filter. In addition, any desired test material can be used to simulate the beam hardening effect, opening up the possibility of a material-sensitive correction. The two corrections tend to agree well in the center of projection images, where we expect the measured approach to perform well since the beam hardening data are recorded there. They diverge towards the edges of the projections, with the calculation method providing a better match to the actual CBCT projections as increased bowtie filter thickness becomes a factor. For this reason we chose the calculation based correction as our method of choice, and all corrected DRRs other than those used for Figure 18 were post-processed using calculated beam hardening.

3.9.2 Comparison to the literature

The x-ray volume rendering approach that we have employed has various benefits over other methods for the purpose of FIPM. Splatting algorithms are fast but produce inferior quality DRRs compared to ray-tracing methods, especially with respect to image sharpness. [68, 69] Shear-warp methods produce similarly blurry images.[69] Fourier volume rendering does not support perspective projection and thus is of limited use.[52] Exact path length ray-tracing methods are known to suffer from artifacts due to aliasing.[48, 69] Our method avoids such artifacts because the tri-linear interpolation used to sample the CT serves as an anti-aliasing (smoothing) filter. Moore et al. modified an exact path length algorithm to use a pencil beam rather than a ray as the projector to avoid artifacts.[48] This approach provides a similar smoothing effect but is not as easily ported to run on the GPU as our method, which takes advantage of hardware-implemented tri-linear interpolation. [57] Light field methods provide a means of interpolating between pre-computed DRRs to generate DRRs from arbitrary angles.[47] However, FIPM algorithms for deformable registration must calculate projections through continuously changing CT volumes, thus any pre-computed projection data become useless as the algorithm iterates. Likewise, Monte Carlo methods rely on a pre-computed sample of points within the CT space, limiting their utility for our purposes.[52]

While many different approaches to the x-ray volume rendering problem exist in the literature, references that construct realistic DRRs based on complete image formation physics are difficult to find. Moore et al. provide one such work, with the goal of creating a tool to aid in the optimization of computed radiography (CR) imaging parameters for chest radiographs that avoids additional patient exposures. Their algorithm contains many of the same elements as ours, though implemented in different ways. As noted, their ray-tracing routine uses a volume based projector to sample the CT, whereas we use tri-linear interpolation. Both algorithms are likely to produce similar results, but our approach is attractive since it is easily ported to run on the GPU. Beam hardening is another area of differentiation. Moore et al. explicitly account for beam hardening by casting many projections through the CT at different simulated energies, and then computing an average projection weighted by the measured x-ray spectrum of their machine. In contrast, our approach computes a single DRR at the average beam energy and then applies a beam hardening correction as a post-processing step. While their approach may provide a realistic simulation they pay a massive price in speed, with reported times of 45-90 minutes per DRR. Our correction incurs minimal additional computational overhead, a necessity for an application like FIPM. In addition, Moore et al. use a beam stop array method to measure scatter, as opposed to our Monte Carlo approach. Measurement based approaches have the benefit of intrinsically including the characteristics of the x-ray source/detector system. However, in order to avoid giving patients extra dose Moore et al. were forced to adapt scatter distributions measured from a RANDO phantom to their patient geometries, limiting them to patients with approximately the same proportions as the phantom. Using Monte Carlo, independent scatter distributions can be calculated for any patient, which is preferable for a registration algorithm meant to be generally applicable. Finally, Moore et al. do not account for the veiling glare effect. As discussed in section 3.9.1, this effect has the greatest impact in the plug projections used to make our CT number to LAC conversion function. Moore et al. create a similar conversion function; however instead of directly measuring the LAC of each plug insert they calculate it from reported density and elemental makeup values. This approach has the drawback of depending on manufacturer-reported values for a critical aspect of the overall DRR calibration process, but renders the absence of a veiling glare correction more or less harmless. Overall, Moore et al. report CBCT and DRR intensity histograms that look similar, as well as comparable signal-to-noise ratios and tissue-to-rib ratios (their measure of contrast) for different tissues within their real and calculated projections. The differences between their algorithm and our own do not necessarily reflect better or worse ways of producing DRRs, but instead relate to the different applications for which the DRRs are meant to be applied, particularly with regards to speed.

3.9.3 Computational efficiency

Computational efficiency is an area of concern for the application of FIPM. Our test scenario used a 1024x768 detector matrix, a 512x512x281 CT with voxels of approximately 1x1x1 mm, and a ray step-size of 0.5 mm. The workstation on which we calculated projections ran Debian Linux on an Intel Core i7 CPU clocked at 2.8 GHz with 12 GB of memory. Execution time scales linearly with number of detector pixels and ray step-size, and approximately linearly with the length of the projected rays. Under these conditions, uncorrected DRRs were calculated in about 0.35 seconds on average using a GeForce GTX 480 GPU. Over 95% of the computation time was spent performing tri-linear interpolation on the CT image, which was carried out by the GPU hardware using special texture memory. Once the GPU finished ray-tracing, it handed the raw DRR back to the CPU where the scatter, beam hardening, and veiling glare post-processing corrections were applied. The scatter and beam hardening corrections were performed in a matter of milliseconds, however the veiling glare correction, which requires a convolution between a kernel sampled on a 600x600 grid and the 1024x768 pixel projection, generally took around 15 seconds. Of the three post-processing corrections the veiling glare correction is by far the least important. As noted in section 3.9.1, the veiling glare effect produces only a small change in image intensities of large object projections, such as

Catphan™ images. Therefore in practice it will likely be acceptable to omit this step for the sake of efficiency. The veiling glare effect was mainly apparent when a small highly attenuating object was imaged in an otherwise unobstructed x-ray fluence. In this scenario a large proportion of the signal that appeared in the object's shadow could be attributed to veiling glare. As demonstrated in section 3.9.1 the plug projections used for our CT number calibration curve fall into this category, thus it remains important to correct for veiling glare in the attenuation measurement data.

Finally, we note that the post-processing corrections applied to our DRRs can just as easily be applied in reverse to "clean up" actual CBCT projections. This is likely the preferable route in the context of our full FIPM algorithm, as it can be done once, offline, before the iterative reconstruction process is initiated. Even though this type of backwards correction is done offline, the veiling glare effect may still be too costly to include, as in reverse it requires a deconvolution taking 15-20 minutes using an implementation of the Richardson-Lucy algorithm. [70] However, it is likely that a large speedup of the deconvolution algorithm can be obtained using a GPU.

3.10 Conclusions

The goal of this chapter was to devise an algorithm for producing physically-realistic DRRs that match actual CBCT projections with no artificial scaling, renormalization, or other adjustment. We analyzed the effects of beam hardening, scatter, and veiling glare and have found that they play a small but non-negligible role in the agreement between DRR and CBCT projection. In addition, we compared beam hardening corrections generated from measured and simulated data and determined simulation to be the superior method. Finally, we benchmarked our algorithm for speed and found it to be sufficiently fast for application to our FIPM based 4DCBCT reconstruction method.

We were motivated in these efforts by the observation that simulated contrast mismatch in a previous study of an FIPM based algorithm induced a jump in CT nRMSE of over 20 %.[31] While it is not yet clear whether our improvements are sufficient to mitigate this effect, we now have the tools to perform a simulation study using realistic projection/DRR mismatch.

4 Real data testing and verification

4.1 Introduction

In this chapter we document our experience reconstructing actual patient lung movement data using the full implementation of our algorithm as described in Chapters 2 and 3. As in Chapter 2 we designed a matrix of tests of escalating complexity and realism to assess the performance of our method. In each test the accuracy of our reconstruction was evaluated in comparison to a gold standard derived from an independent reconstruction of the 3D movement of fiducial landmarks in the lung. We present both quantitative and qualitative reconstruction results and analyze their implications.

4.2 Patient data-sets

At our institution we have patient data-sets containing multiple 4DFBCTs and 4DCBCTs acquired over the course of lung cancer radiotherapy treatment. The 4DCBCT data-sets contain the raw 4DCBCT projections and their associated meta-data, patient breathing traces captured simultaneously with image acquisition using the RPM system, and final 4DCBCTs reconstructed via phase-binning and backprojection. In addition, patient tumors were implanted with radio-opaque fiducial markers prior to imaging and we have 3D marker positions in the imaging world coordinate system for each projection in the raw 4DCBCT data-set in which the marker was visible. Together these positions form a 3D motion trajectory of the marker as captured during imaging. These 3D marker trajectories were estimated independently using the method of Poulsen et al. as described in section 4.6.1, and were treated as a gold standard of internal anatomic patient motion to which we compared our results. [71]

4.3 Data pre-processing

Before reconstructing a 4DCBCT with our algorithm we first pre-processed the input FBCT and CBCT data by masking it so that the implanted marker was the only visible structure. For the FBCT data we isolated the 3D marker and set all other voxels to zero. In addition, we cropped the full FBCT volume down to a smaller size to enable faster DRR computation. For the CBCT data, we segmented the 2D marker shadow within each raw 4DCBCT projection so that it was the only visible structure. After segmentation we then blurred the marker using a Gaussian function. The blurring step was enacted to make the registration process more robust and is not required for unsegmented data-sets showing the full patient anatomy. This process is visualized in Figure 19.

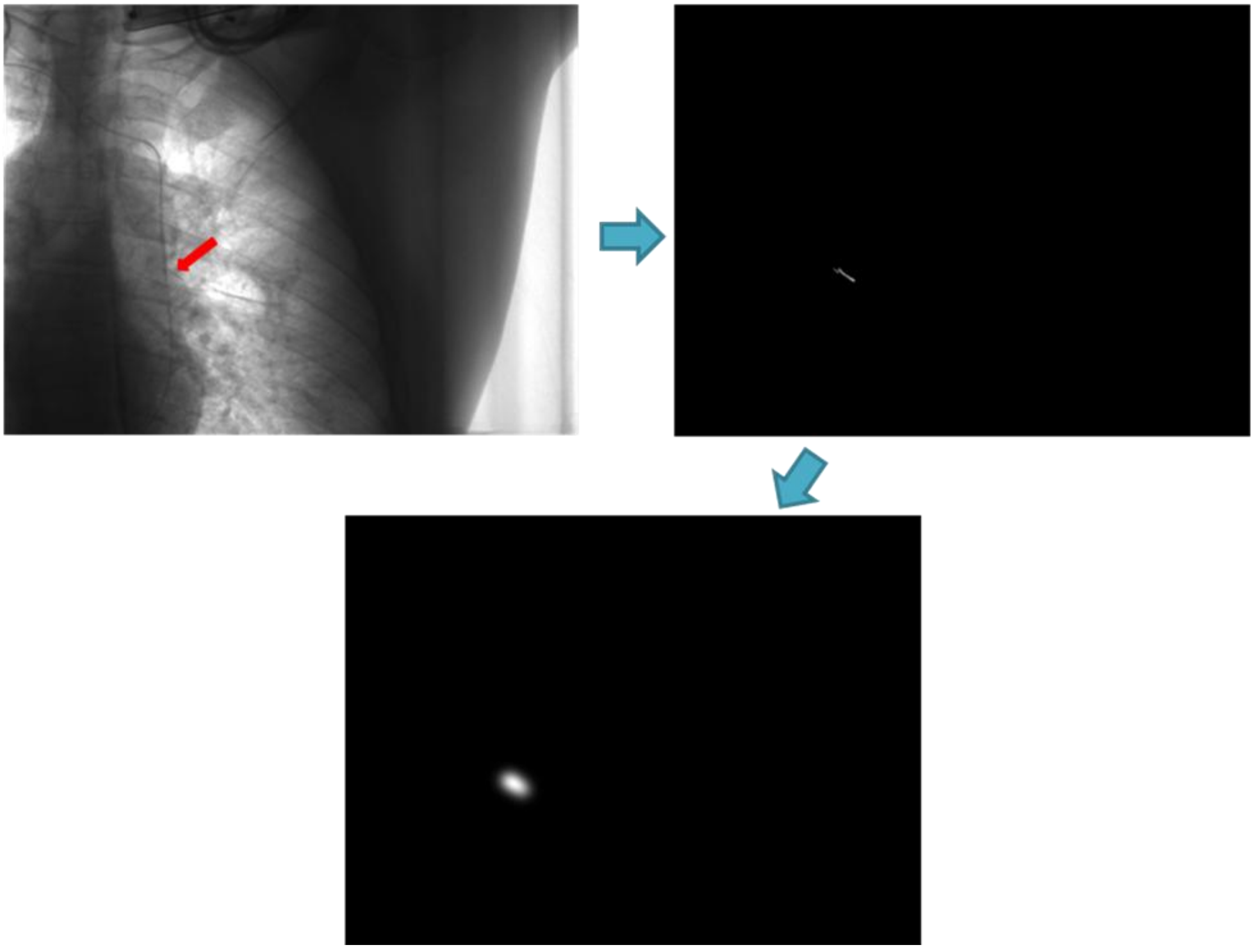


Figure 19: Illustration of the raw 4DCBCT projection pre-processing chain. The initial image (top-left) is segmented so that only the implanted marker is visible (top-right). Then the segmented marker is blurred using a Gaussian function to help mask segmentation irregularities and make it more visible to the optimization algorithm (bottom).

4.4 Test matrix

To test the performance of our algorithm with real patient data we investigated it in scenarios of varying realism, with tests ranging from highly idealized to fully realistic. These tests examined the performance of the algorithm with respect to three fundamental elements: the method used to produce the eigenvector training data, the formulation of the motion model itself, and the fitting target data to which the motion model was optimized. For instance, since the gold standard 3D marker trajectory was one of our validation data-sets we tested how accurately it could be reproduced by our motion model in the absence of any other complications. To do this we trained the motion model eigenvectors on points sampled from the gold standard marker trace itself. We then fit our parametric motion model directly to the full gold standard marker trace. This was analogous to the tests using the DVF fitting target described in Section 2.4.3. At the other end of the spectrum we used our complete 2D/3D FIPM-bspline registration method to produce training DVFs in a fully realistic manner and then fit the resulting motion model to actual patient 4DCBCT projections.

Each fundamental component of our algorithm was configured in multiple ways, as listed below. The components were then combined in the sequence of tests described in order of increasing complexity in Table 4. This set of tests can be decomposed in 3 main sub-groups. Tests 1-8 all use the 3D gold standard maker trajectory as both the optimization fitting target and the source of the eigenvector training data. The purpose of this group is to demonstrate the accuracy with which a motion model based on our patient breathing trace can reproduce the 3D gold standard marker trajectory. It provides a baseline level of 3D marker trajectory disagreements as a reference point for later tests. Tests 9-18 use actual projection data as the fitting target and source of training data. These tests constitute our investigation of actual algorithm performance. Tests 19-25 also use actual patient data as the fitting target and training data source. However, they are devised specifically in order to test the temporal durability of the optimized motion model.

Table 4: Matrix of patient data test conditions.

Test number	Training data	Motion model	Optimization target
Test 1	• None.	• Single parameterized function of breathing trace.	• Individual components of gold standard marker trajectory.
Test 2	• Full set of points composing gold standard marker trajectory, subtracted from reference position to convert to displacement vectors. • 700 points total.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Gold standard marker trajectory.
Test 3	• Subset of points composing gold standard marker trajectory, subtracted from reference position to convert to displacement vectors. • 10 point total.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Gold standard marker trajectory.
Test 4	• Manually tracked marker locations in each phase of backprojection reconstructed 4DCBCT. • 10 points total.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Gold standard marker trajectory.
Test 5	• FIPM-bspline algorithm used to register 3D marker mask to amplitude binned 2D marker projection masks. DVFs from registration used to deform 3D mask. Calculated centroid of each deformed 3D mask. • 10 points total.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Gold standard marker trajectory.
Test 6	• Full set of points composing gold	• Multiple parameterized	• Geometric projections of gold

	standard marker trajectory, subtracted from reference position to convert to displacement vectors. • 700 points total.	functions of breathing trace coupled to eigenvector basis. • 3 modes.	standard marker trajectory positions.
Test 7	• Manually tracked marker locations in each phase of FIPM-bspline reconstructed 4DCBCT, subtracted from reference position to convert to displacement vectors. • 10 points total.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes	• Geometric projections of gold standard marker trajectory positions.
Test 8	• Manually tracked marker locations in each phase of backprojection reconstructed 4DCBCT, subtracted from reference position to convert to displacement vectors. • 10 points total.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Geometric projections of gold standard marker trajectory positions.
Test 9	• Subset of points composing gold standard marker trajectory, subtracted from reference position to convert to displacement vectors. Copies of single vectors used to fill out full DVFs covering full CT ROI. • 10 DVFs total.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections from first 700 time points considered.
Test 10	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • “Bad” phase binning. • 10 bins. • Projections per bin	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections from first 700 time points

	(min, max, avg): 215, 255, 235		considered.
Test 11	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning. • 10 bins. • Projections per bin (min, max, avg): 164, 478, 235	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections from first 700 time points considered.
Test 12	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning. • 20 bins. • Projections per bin (min, max, avg): 71, 291, 117	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections from first 700 time points considered.
Test 13	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning. • 30 bins. • Projections per bin (min, max, avg): 49, 220, 117	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections from first 700 time points considered.
Test 14	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Hybrid binning. • 60 bins. • Projections per bin (min, max, avg): 15, 67, 32	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections from first 700 time points considered.
Test 15	• DVFs from FIPM-bspline	• Multiple	• Raw 4DCBCT

	registration of 3D marker mask to binned 2D marker projection masks. - Hybrid binning. - 60 bins. - Projections per bin (min, max, avg): 15, 67, 32	parameterized functions of breathing trace coupled to eigenvector basis. - 1 mode.	projections; masked so that only marker is visible. - 60 projections included. Only projections from first 700 time points considered.
Test 16	- DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. - Hybrid binning. - 60 bins. - Projections per bin (min, max, avg): 15, 67, 32	- Multiple parameterized functions of breathing trace coupled to eigenvector basis. - 2 modes.	- Raw 4DCBCT projections; masked so that only marker is visible. - 60 projections included. Only projections from first 700 time points considered.
Test 17	- DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. - Hybrid binning. - 60 bins. - Projections per bin (min, max, avg): 15, 67, 32	- Multiple parameterized functions of breathing trace coupled to eigenvector basis. - 4 modes.	- Raw 4DCBCT projections; masked so that only marker is visible. - 60 projections included. Only projections from first 700 time points considered.
Test 18	- DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. - Hybrid binning. - 60 bins. - Projections per bin (min, max, avg): 15, 67, 32	- Multiple parameterized functions of breathing trace coupled to eigenvector basis. - 5 modes.	- Raw 4DCBCT projections; masked so that only marker is visible. - 60 projections included. Only projections from first 700 time points considered.
Test 19	- DVFs from FIPM-bspline registration of 3D marker mask to binned	- Multiple parameterized functions of breathing	- Raw 4DCBCT projections; masked so that only marker

	2D marker projection masks. • Amplitude binning. • 20 bins. • Projections for training set sampled from all valid time points. • Projections per bin (min, max, avg): 71, 291, 117	trace coupled to eigenvector basis. • 3 modes.	is visible. • 60 projections included. Only projections from first 700 time points considered.
Test 20	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning. • 20 bins. • Projections for training set sampled from all valid time points. • Projections per bin (min, max, avg): 71, 291, 117	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections acquired at or after time point 1612 considered.
Test 21	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning. • 20 bins. • Projections for training set sampled from all valid time points. • Projections per bin (min, max, avg): 71, 291, 117	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 120 projections included. Both projections acquired at or before time point 700 and at or after time point 1612 considered.
Test 22	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning.	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only

	• 10 bins. • Projections for training set sampled only from projections acquired at or before time point 700. • Projections per bin (min, max, avg): 32, 172, 70		projections from first 700 time points considered.
Test 23	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning. • 10 bins. • Projections for training set sampled only from projections acquired at or before time point 700. • Projections per bin (min, max, avg): 32, 172, 70	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections acquired at or after time point 1612 considered.
Test 24	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection masks. • Amplitude binning. • 10 bins. • Projections for training set sampled only from projections acquired at or after time point 1612. • Projections per bin (min, max, avg): 42, 133, 73	• Multiple parameterized functions of breathing trace coupled to eigenvector basis. • 3 modes.	• Raw 4DCBCT projections; masked so that only marker is visible. • 60 projections included. Only projections acquired at or after time point 1612 considered.
Test 25	• DVFs from FIPM-bspline registration of 3D marker mask to binned 2D marker projection	• Multiple parameterized functions of breathing trace coupled to	• Raw 4DCBCT projections; masked so that only marker is visible.

	masks. • Amplitude binning. • 10 bins. • Projections for training set sampled only from projections acquired at or after time point 1612. • Projections per bin (min, max, avg): 42, 133, 73	eigenvector basis. • 3 modes.	• 60 projections included. Only projections from first 700 time points considered.

4.4.1 Training data for the PCA eigenvectors

1. **None:** No training data required.
2. **Marker trace sample**: A set of points selected from the gold standard 3D marker trajectory were subtracted from a reference position to create displacement vectors. Each displacement vector served as a DVF feature vector.
    a. Anywhere from 700 to 10 points used in the training set.
3. **Manually tracked marker in reconstructed 4DCBCT**: Markers were manually identified in each of 10 phases of the conventionally reconstructed 4DCBCT. Marker centroids were subtracted from a reference position to create displacement vectors. Each displacement vector served as a DVF feature vector.
    a. Backprojection reconstruction.
    b. FIPM-bspline reconstruction.
4. **Upscaled marker trace sample**: A small number of points from the gold standard 3D marker trajectory were subtracted from a reference position to create displacement

vectors. Each displacement vector was copied to all positions of a DVF with the dimensions of the reference FBCT. Each DVF created in this manner served as a feature vector.

5. **FIPM-bspline registration:** The FIPM-bspline 2D/3D registration method was used to register the FBCT marker mask to the masked 4DCBCT projections grouped into breathing bins. Binning was performed based on breathing trace phase as determined by the RPM system. The DVFs from each registration served as feature vectors.
    a. Binning based on breathing phase, amplitude, or phase/amplitude hybrid.
    b. 10, 20, 30, or 60 bins created.

4.4.2 Motion model formulation

1. **Single parameterized function of breathing trace:** No eigenvectors needed. This formulation was designed to model an individual spatial component of a single displacement vector. It was valid only with training data option 1.
2. **Parameterized functions of breathing trace coupled to eigenvectors:** The full model as defined by Eq. (3) in section 2.2.1.
    a. 1, 2, 3, 4 or 5 modes.

4.4.3 Fitting target

1. **Gold standard 3D marker trajectory, individual components:** The motion model was fit directly to individual spatial components of the gold standard marker trajectory. This target data type was valid only in conjunction with training data type 1 and motion model type 1.
2. **Gold standard 3D marker trajectory:** The motion model was fit directly to the full gold standard marker trajectory. This target data type was valid in conjunction with training data types 2 and 3 and motion model type 2.
3. **Geometric marker projections:** Each point in the gold standard marker trajectory was geometrically projected into the 2D detector plane. Projected points were blurred with a Gaussian function to give them a finite spatial extent. The motion model was used to displace a reference point for which like geometric projections were calculated. The motion model was optimized to minimize intensity mismatch between the two sets of projections. This target data type was valid in conjunction with training data types 2 and 3 and motion model type 2.
4. **Masked 4DCBCT projections:** The motion model was used to deform the reference FBCT marker mask and DRRs were cast through the deformed CT at chosen time points. The motion model was optimized to minimize the intensity mismatch between the DRRs and pre-processed raw 4DCBCT projections. This target data type was valid in conjunction with training data types 4 and 5 and motion model type 2.

4.5 Areas of investigation

The real data testing phase allowed us to evaluate the best implementations and practices for various components of our algorithm. Some of these, such as our investigation of the optimal number of modes for inclusion in the motion model, were previously examined in the proof of concept testing detailed in Chapter 2. Others, like our examination of different projection binning techniques, are new areas of exploration. All proposed refinements based on our real data experience are listed below.

4.5.1 Motion model

The first component that we sought to optimize was the eigenvector basis underlying the DVF motion model. We hypothesized that the eigenvectors would be sensitive to the phase-to-phase DVFs used for our training data, which in turn would be sensitive to the binning method used to obtain the CT time bins. We therefore looked for the projection binning strategy for which the FIPM-bspline 2D/3D registration method produced the highest quality training DVFs. This included an investigation of the number of breathing bins into which the projections were divided (and thus the density of the training set), as well as the overall binning strategy.

The second area we explored was the number of modes included in the motion model. We studied the performance of our algorithm when the motion model incorporated anywhere from 1-5 principal modes.

4.5.1.1 Projection binning methods

4.5.1.1.1 Phase binning

Our projection binning algorithm based on breathing phase was lifted from the output log of the RPM system. This proprietary method labeled each breathing data point with a phase tag ranging from 1-10. Projections with identical phases were sorted into the same breathing bins.

4.5.1.1.2 Amplitude binning

Our projection binning algorithm based on breathing amplitude used the RPM breathing trace to group projections with similar breathing trace amplitude. The total set of bins consisted of two sub-sets of bins with identical boundaries, with one set reserved for breathing points with positive breathing trace gradient, and the other set for breathing points with negative gradient. All bins had nominally identical widths, however projections occurring at breathing points above the maximum bound of the highest bins were included in those bins and projections occurring at breathing points below the minimum bound of the lowest bins were included in those bins. The first step in determining bin boundaries was to label the amplitude of each peak and trough in the RPM breathing trace and calculating average peak and trough amplitudes. The uniform breathing bin width was then calculated as the difference between the average peak and trough heights divided by half the desired number of breathing bins, N/2, as specified by the user (only even numbers of total bins, N, were allowed).

$$A_{width} = \frac{\left(\bar{A}_{peak} - \bar{A}_{trough}\right)}{N/2} \tag{63}$$

Finally, the minimum and maximum amplitude bounds for each bin were calculated.

$$A_{\min,i} = \bar{A}_{trough} + \left(i-1\right)A_{width} \quad , \quad A_{\max,i} = \bar{A}_{trough} + iA_{width} \quad \Big\} i = 1:N/2 \tag{64}$$

4.5.1.1.3 Hybrid binning

We devised a novel hybrid projection binning algorithm that grouped projections based on both their breathing trace amplitude and gradient. The goal of this method was to collect only those projections that were acquired within a jointly-defined breathing phase and amplitude range into time bins, discarding those projections for which either the amplitude or phase identified it as a potential outlier. The expectation was that this approach would weed out projections acquired at irregular points in the breathing history. Bins in this method were two dimensional, having a width in the amplitude direction and a height in the gradient (phase) dimension. The algorithm proceeded in two steps, bin center identification and bin boundary optimization.

The purpose of the bin center identification phase was to identify the amplitude and gradient values around which each breathing bin was to be centered. Step 1 of this phase was to add each breathing point at which a projection was acquired to a finely sampled 2D histogram, where the X direction of the histogram corresponded to breathing trace amplitude and the Y direction corresponded to breathing trace gradient. This step is visualized in part A of Figure 20. Next the histogram was smoothed by convolving it with a 2D Gaussian function. Finally, all peaks in the smoothed histogram were identified and any peak exceeding a user specified threshold was accepted as a bin center. This step is visualized in part B of Figure 20.

In the bin boundary optimization phase a bounding box was dropped around each previously identified bin center and the width and height of the box were adjusted until the breathing points and corresponding projections contained within it were determined to be optimal. The optimality of each box was scored based on four criteria:

1. **Breathing trace amplitude spread:** This criterion was quantified as the standard deviation of the breathing trace amplitude of points falling within the bin. A smaller value indicated a better bin.
2. **Breathing trace gradient spread:** This criterion was quantified as the standard deviation of the breathing trace gradient of points falling within the bin. A smaller value indicated a better bin.
3. **Projection angular distribution:** The criterion was also quantified as a standard deviation, and was calculated in a multi-step process. First the gantry angles of projections within the bin were organized into a list sorted by ascending value. Angular intervals were calculated by computing the difference between adjacent gantry angles.

Next the self-weighted mean angular interval was computed, in which the weight for each value was the value itself, $\Delta\overline{\theta}_w = \frac{\sum_{i=1}^{N_p} \Delta\theta_i w_i}{\sum_{i=1}^{N_p} w_i}$, $w_i = \Delta\theta_i$, where N_p was the number of projections contained within the bin. Finally the sample standard deviation of the angular intervals about the self-weighted mean was calculated, $s = \sqrt{\frac{\sum_{i=1}^{N_p} \left(\Delta\theta_i - \Delta\overline{\theta}_w\right)^2}{N_p - 1}}$. A lower standard deviation indicated a more uniform angular distribution and thus a better bin.

4. **Projection count:** The total number of projections contained within the bin. A larger number indicated a better bin.

The values of each score were normalized to 1 and a weighted average computed to determine the final bin score (weighting factors were user supplied and set to 1 for the results reported here). The highest scoring dimensions for each bin were taken as the final bin parameters. This step is visualized in parts C and D of Figure 20.

It should be noted that parameters for each bin were determined independently, allowing multiple bins to share the same projection while some projections were not included in any bin.

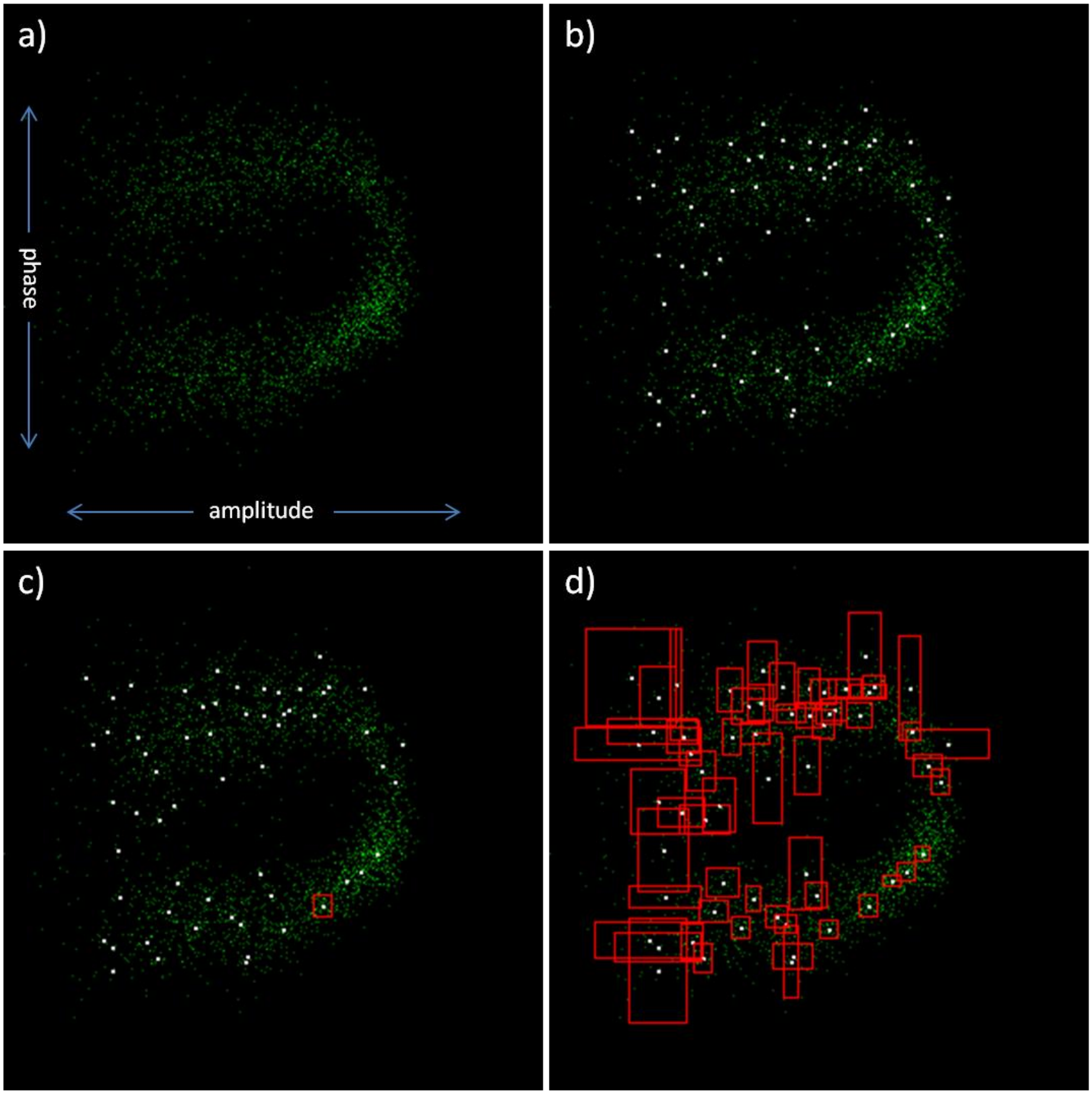


Figure 20: Illustration of steps taken in hybrid binning algorithm. A) Each breathing trace point is added to a cell in a 2D histogram based on its amplitude and phase (amplitude gradient) to build the distribution of (green) points. B) Peaks in the histogram are identified and chosen as bin centers if they exceed a threshold value. Accepted bin centers are labeled as white dots. C) Boxes of different shapes and sizes in the amplitude/phase grid are drawn around each bin center, and the characteristics of the breathing points and corresponding projections within each box are given a score. The highest scoring box defines the boundaries in amplitude and phase for that bin (red box). D) Bounding boxes for all bin centers. Bounding boxes are calculated independently for each bin center and can overlap. In addition, some breathing trace points, and thus their corresponding projections, are not included in any bin.

4.5.1.2 Training set density

Another projection binning related parameter was the number of DVF feature vectors included in the training set. Sorting the projections into more bins resulted in a corresponding increase in the number of feature vectors. With all other circumstances being the same, we expect a more densely sampled training set to more fully capture the imaging target motion and thus generate better eigenvectors. However, changing the binning parameters can affect the quality of the individual 2D/3D registrations and introduce noise or artifact into the eigenvectors.

For the purposes of testing we created eigenvectors from registrations of projections grouped into 10, 20, and 30 amplitude bins as well as 60 hybrid bins. While such large numbers of bins and the corresponding reduction in number of projections per bin are not practical for a FBP based method, the FIPM-bspline algorithm has been shown to be effective with as few as 8 projections.[31]

4.5.1.3 Number of eigenmodes

Our final area of inquiry into the motion model regarded the optimal number of modes to include. Using more modes helps to ensure that no motion information is omitted from the model, but it can also make the optimization process less robust by increasing the dimensionality of the search space, as well as result in the inclusion of modes dominated by noise.

To find the appropriate number we performed reconstructions using anywhere from 1-5 modes under otherwise identical situations. In all cases the eigenvectors were generated from the DVF training set using 60 hybrid bins and the fitting target used was the pre-processed raw 4DCBCT projections.

4.5.2 Optimization

Our second avenue of algorithm development was to make our numerical optimization scheme more robust, i.e. improve its convergence to the objective function global minimum. We used three strategies to achieve this goal, the first and foremost of which was to design a process for initializing the motion model free parameters. The other two were pre-processing the patient breathing trace in order to decouple the motion model amplitude and baseline offset free parameters, and implementing a free parameter scaling vector to rescale parameter magnitudes to a uniform range of values.

4.5.2.1 Initial free parameter guess

Our solution to this problem was achieved in several steps, beginning with the acquisition of average breathing and principal coefficient traces. Average breathing traces were obtained by calculating the average breathing trace amplitude of the projections contained within each breathing bin used to generate the DVF training set. The full set of these average amplitude values represented a single average breathing trace cycle. Average principal coefficient traces were obtained for each mode by projecting each feature vector from the training set into that mode's eigenvector.

The next step was to determine the amplitude, time shift, and baseline offset terms (see Eq. (4) in Section 2.2.1) necessary to make the average breathing trace emulate each average principal coefficient trace as closely as possible. The magnitude of the amplitude term was determined by taking the ratio of the principal coefficient and breathing trace amplitudes. The phase shift term was determined by shifting the average breathing trace in time and calculating the correlation coefficient between the shifted breathing trace and the principal coefficient trace. The shift generating the largest magnitude correlation coefficient was set as the time shift term initial guess, and the sign of the correlation coefficient was assigned as the sign of the amplitude term. The baseline offset term was determined as the difference between the average breathing trace mean and the average principal coefficient trace mean.

Plots of the these average traces are displayed for a training set generated using FIPM-bspline registration to 10 phase bins in Figure 21 and a training set generated from an upscaled marker trace sub-sample in Figure 22.

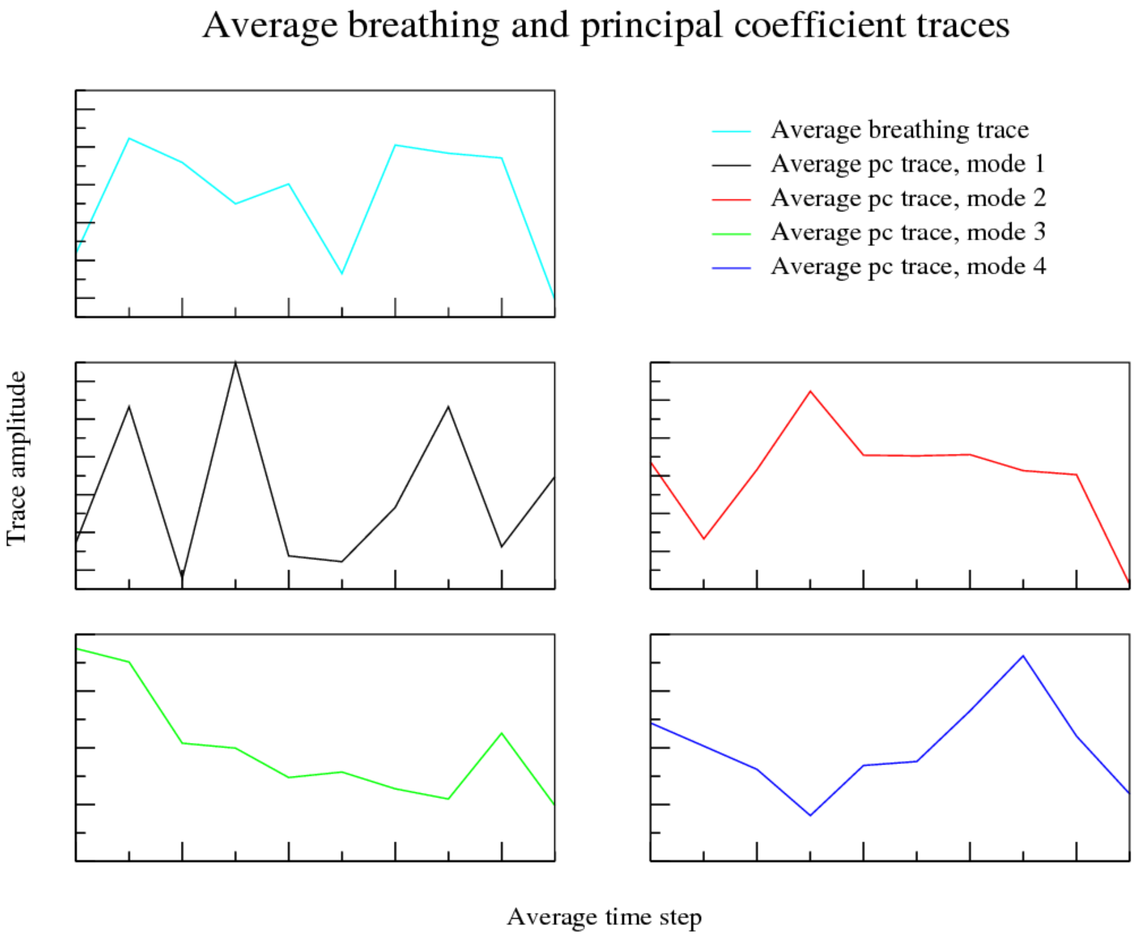


Figure 21: Average breathing and principal coefficient traces. The training set is generated using FIPM-bspline registration to 10 phase bins.

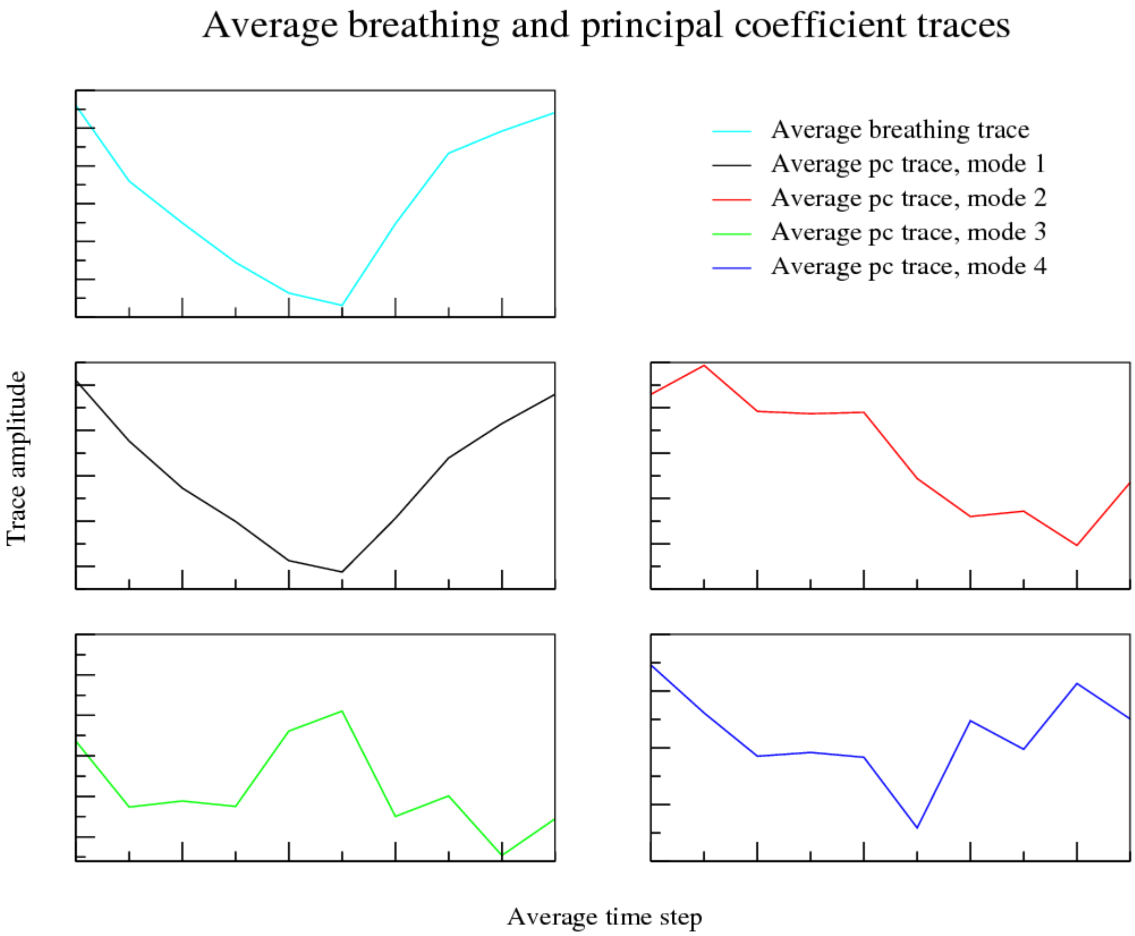


Figure 22: Average breathing and principal coefficient traces. The training set is generated from 10 upscaled gold standard marker displacement points.

4.5.2.2 Breathing trace pre-processing

Before initiating our optimization routine we first pre-processed the breathing trace by subtracting its mean. The rationale for performing this step is described by the following. The patient breathing trace can be considered as a function with the following form:

$$b(t) = Dx(t) + E \tag{65}$$

a normalized, zero mean pseudo-periodic signal, $x(t)$, scaled by an amplitude term, D, and riding on top of a baseline offset E. Likewise, each principal coefficient trace is assumed to be a function with similar form:

$$pc(t') = Fx(t') + G \tag{66}$$

As defined in Eq. (4) the goal of our algorithm is to represent each principal coefficient trace as a parameterized function of the breathing trace:

$$pc(t') = Ab(t+\tau) + C \tag{67}$$

Substituting Eq. (65) into Eq. (67) yields:

$$pc(t') = A\big(Dx(t+\tau)+E\big)+C = ADx(t+\tau)+AE+C \tag{68}$$

Comparing Eqs. (66) and (68) we see that

$$AD = F \rightarrow A = {}^{F}\!\big/_{D} \tag{69}$$

$$t+\tau = t' \rightarrow \tau = t'-t \tag{70}$$

$$AE + C = G \rightarrow C = G - AE \tag{71}$$

Examining Eq. (71) we see that the amplitude and baseline offset terms of our model are coupled rather than independently varying. This dependence could effectively cause our optimization algorithm to change the value of one term when only intending to adjust the value of the other. However, if the term E in Eq. (71) is zero this dependence disappears. Since E is the baseline offset of the breathing trace we can ensure that its value be zero by pre-processing the breathing trace to have zero mean and avoid this possible complication.

4.5.2.3 Parameter scaling vector

Our optimization algorithm attempts to find the optimal set of free parameters by stepping through parameter space along its best estimate of the direction leading towards the objective function minimum. If the magnitudes of our free parameters have greatly different scales, then a small step along one axis of the parameter space may be equivalent to a huge leap along another axis. If this is the case then our optimization algorithm could be forced to move at a lethargic pace to accommodate the small-magnitude parameter, or experience wild objective function fluctuations when taking steps appropriate for the large-magnitude parameter. In our application the amplitude and baseline offset free parameters could be two orders of magnitude greater than the time shift term, so this was a potential issue of concern.

Our solution was to assign a fixed scaling term for each free parameter such that the free parameter magnitudes become approximately equal from the optimization algorithm's point of view. These scaling terms were stored in a data structure referred to as a scaling vector. The amplitude scaling terms were the inverses of the amplitude term initial guesses calculated in section 4.5.2.1. The time shift scaling term was set as the inverse of the average breathing trace period. The baseline offset scaling term was set to a value equal to half the amplitude scaling term. Empirical observations showed that the baseline offset initial guess could be a poor indicator of parameter magnitude, while a number about half the amplitude parameter initial guess was roughly appropriate.

4.5.3 Motion model temporal stability

A final avenue of investigation was the temporal stability of our optimized motion model. Essentially we wanted to know if a motion model trained and optimized on data from an earlier time could accurately model motion recorded at a later time. To explore this question we divided our raw 4DCBCT projection data-set into two groups each composed of approximately 700 projections. The first group contained the first 700 projections from the full 4DCBCT data-set, while the second group contained the last 700 projections from the full data-set. About 900 projections in the middle were excluded, equating to a temporal separation of about 225 seconds based on the projection acquisition rate of approximately 4 Hz. We then created DVF training sets based on the data contained in each group and generated eigenvector bases from each training set. Finally, we optimized a motion model incorporating each eigenvector basis against target projection data sampled from each group. For instance, in one test we optimized a motion model trained on data from the first group against projections from the second group. We also optimized a motion model trained on data from both groups against target projection data from either the first, second, or both groups. Specific descriptions and results for data runs investigating motion model temporal stability are detailed in tests 19-25 in Table 4, Table 5, and Table 6.

4.6 Validation data-sets

As described in Chapter 2, reconstruction using our FIPM-PCA 4DCBCT algorithm produces 3 data-sets when optimization is complete: a time series of DVFs, a time series of CTs, and a time series of DRRs, all of which track the moving anatomy present during imaging. For our pre-processed patient data, the CT time series tracks the 3D motion of the marker in the patient or world reference frame (the gold standard), and the DRR time series tracks the 2D marker motion in the detector reference frame (the ground truth). For purposes of validation we transformed our reconstructed CT and DRR data to 3D and 2D marker motion traces and compared to gold standard and ground truth data respectively.

4.6.1 3D marker trajectory gold standard data

The question of whether or not the 3D position of the marker has been accurately reconstructed is the greatest area of relevance for validation of our algorithm. However, we have no ground truth 3D marker data with which to compare our results. Therefore, our strategy was to use an independent method to reconstruct the 3D marker trajectory as a gold standard.

To perform this task we used 3D trajectories reconstructed from our patient data sets by a group at our institution via the method of Poulsen et al. [71, 72] In this method the 2D position of the marker was first determined in the detector reference frame for each projection using an auto-segmentation algorithm. Next, each 2D position was backprojected in 3D space along the line connecting it to the CBCT x-ray source using the known gantry angle and CBCT projection geometry. A 3D Gaussian function was then fit to the backprojection image to create a 3D PDF characterizing the likelihood of the marker residing at a particular point in space. Finally, the line connecting each 2D marker position to the CBCT x-ray source was revisited. The marker was known to lie somewhere along this line, however its position along the line was unresolved. This final position was estimated by extracting the 1D PDF described by the intersection of the backprojection line and the 3D Gaussian PDF and calculating its expectation value. In this manner the 3D position of the marker for each frame in the raw 4DCBCT projection data-set was estimated.

When the target data to which the motion model was optimized was the gold standard trajectory itself (tests 1-8 in Table 4) then the output of our algorithm was a directly comparable 3D marker motion trace. However, if the fitting target data used was the actual CBCT projections then our reconstructed data was in the form of a series of deformed CT volumes. To transform our data into a series of 3D positions we calculated the centroid of each deformed CT image, obtaining the approximate 3D location of the marker at each time point. Comparison of the two 3D marker trajectories was complicated by the fact that the CBCT imaging geometries to which they were referenced were formulated in different ways.

Misalignment of the two could result in artificial offset differences not indicative of true disagreement. This issue is discussed further in Section 4.8.

4.6.2 2D marker position ground truth data

Our second set of validation data was the 2D position of the marker in the raw 4DCBCT projections themselves. Since these data came directly from known patient measurements with no interpolation or approximation they served as a ground truth. We compared positions in this data-set with the projected position of the marker in DRRs cast through the reference CT deformed by the optimized motion model.

Since the 2D projections and the 3D reference CT were pre-processed to contain only the marker, determining 2D marker positions was a simple procedure. The centroid of each image was calculated for both DRRs and projections, yielding the marker centroid coordinates within the detector reference frame. This validation method was only applicable to test runs using the actual 4DCBCT projection images as the fitting target (tests 9-25 in Table 4).

4.7 Results

4.7.1 Marker trajectory plots

The marker provides a fiducial with which we can detect the breathing motion simulated by our motion model and used to animate the reference CT. After reconstructed and gold standard 3D marker trajectory data were acquired via the methods of Section 4.6.1 they were plotted together for visual comparison. Figure 23, Figure 24, Figure 25, Figure 26, Figure 27, Figure 28, Figure 29, Figure 30, Figure 31, Figure 32, Figure 33, Figure 34, and Figure 35 plot reconstructed and gold standard 3D marker trajectory results for tests 2, 7, 8, 10, 11, 13, 14, 15, 18, 22, 23, 24, and 25 respectively. In each figure motion in the X, or left-right (LR), Y, or anterior-posterior (AP), and Z, or superior-inferior (SI) direction was plotted separately in mm units. Figure 23, Figure 24, Figure 25, Figure 26, Figure 27, Figure 28, Figure 29, Figure 30, and Figure 31 plot results for the motion model and optimization tests described in Section 4.5.1 and 4.5.2. These plots cover the first 700 time points contained in the raw 4DCBCT data-set. Figure 32, Figure 33, Figure 34, and Figure 35 plot results for the temporal stability tests described in Section 4.5.3. They display results for all 2345 time points contained in the raw 4DCBCT data-set. For approximately the middle 900 time points in these figures the gold standard data is missing as the marker was outside of the detector field of view during this period.

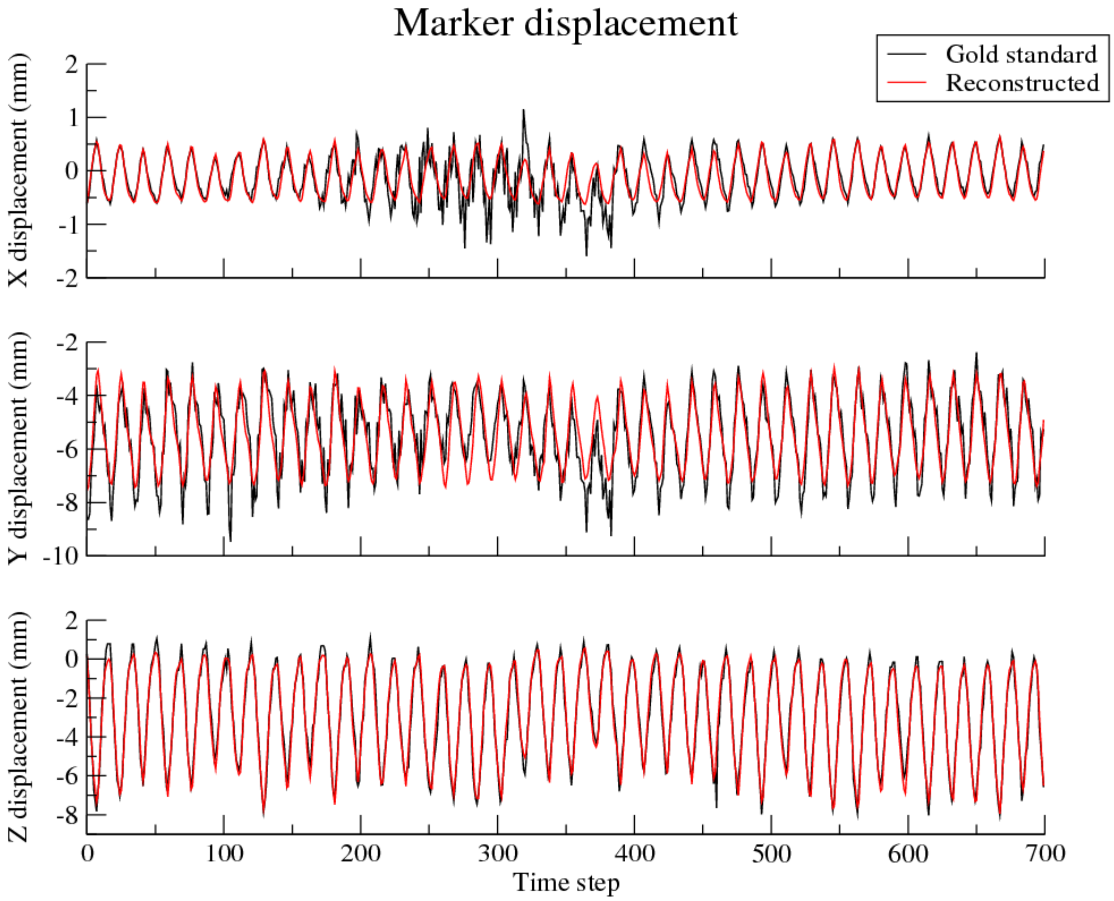


Figure 23: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use the marker trace full-sample training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and the gold standard marker trajectory as the fitting target (Test 2).

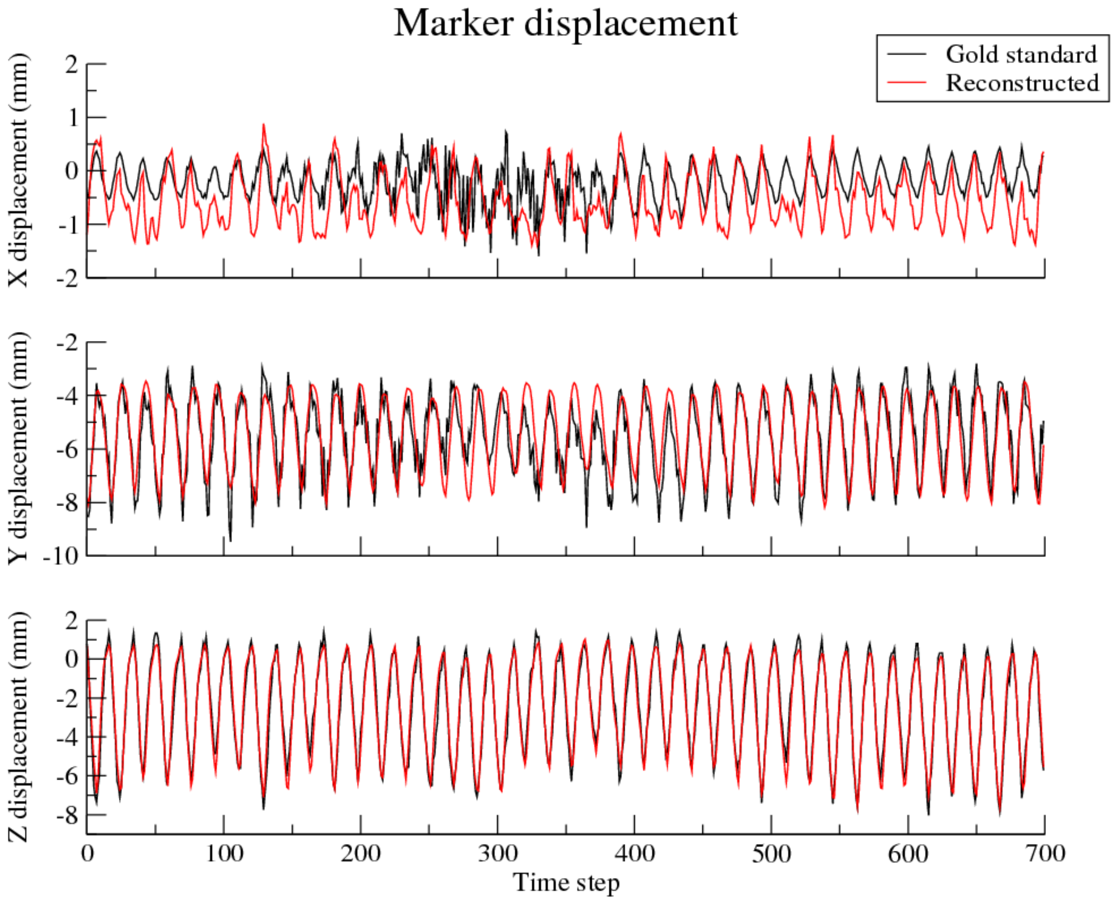


Figure 24: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use manual marker tracking in the FIPM-bspline reconstructed 4DCBCT to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and geometric marker projections as the fitting target (Test 7).

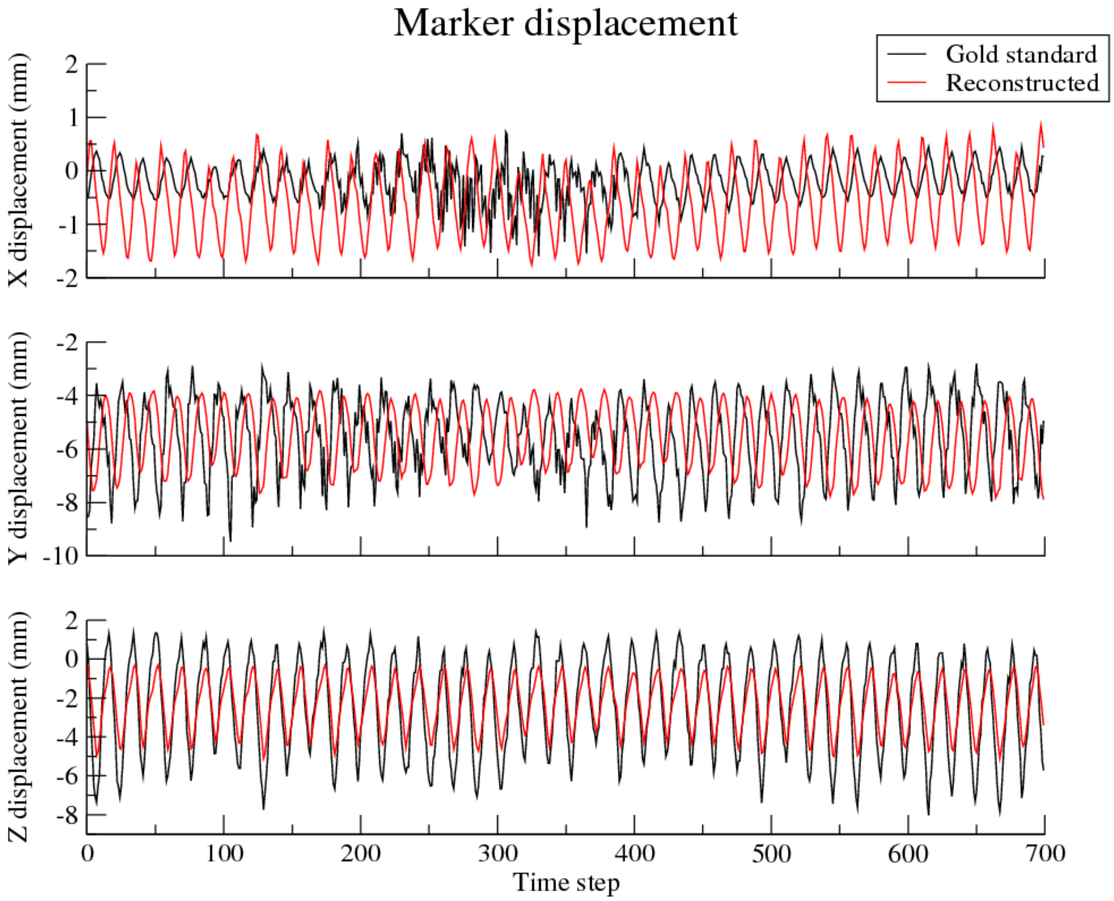


Figure 25: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use manual marker tracking in the backprojection reconstructed 4DCBCT to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and geometric marker projections as the fitting target (Test 8).

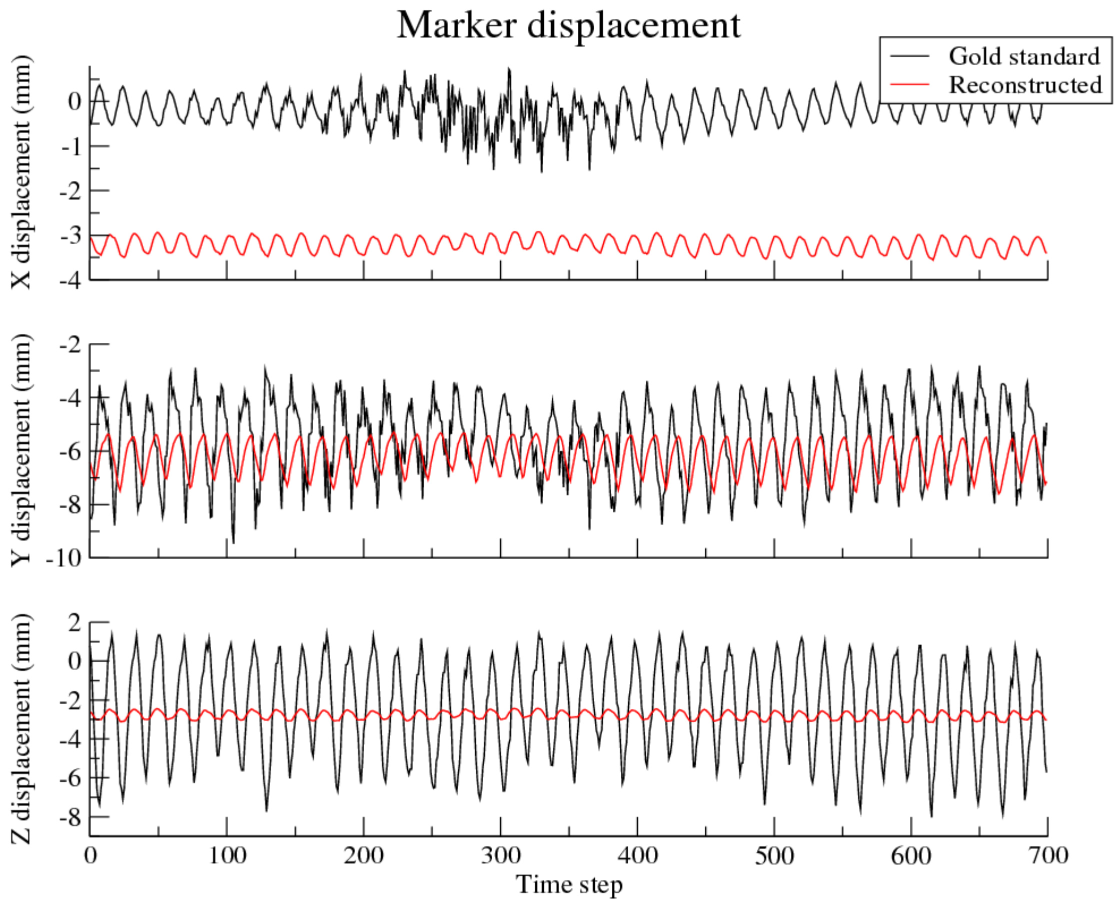


Figure 26: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use FIPM-bspline registration with 10 phase bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 10).

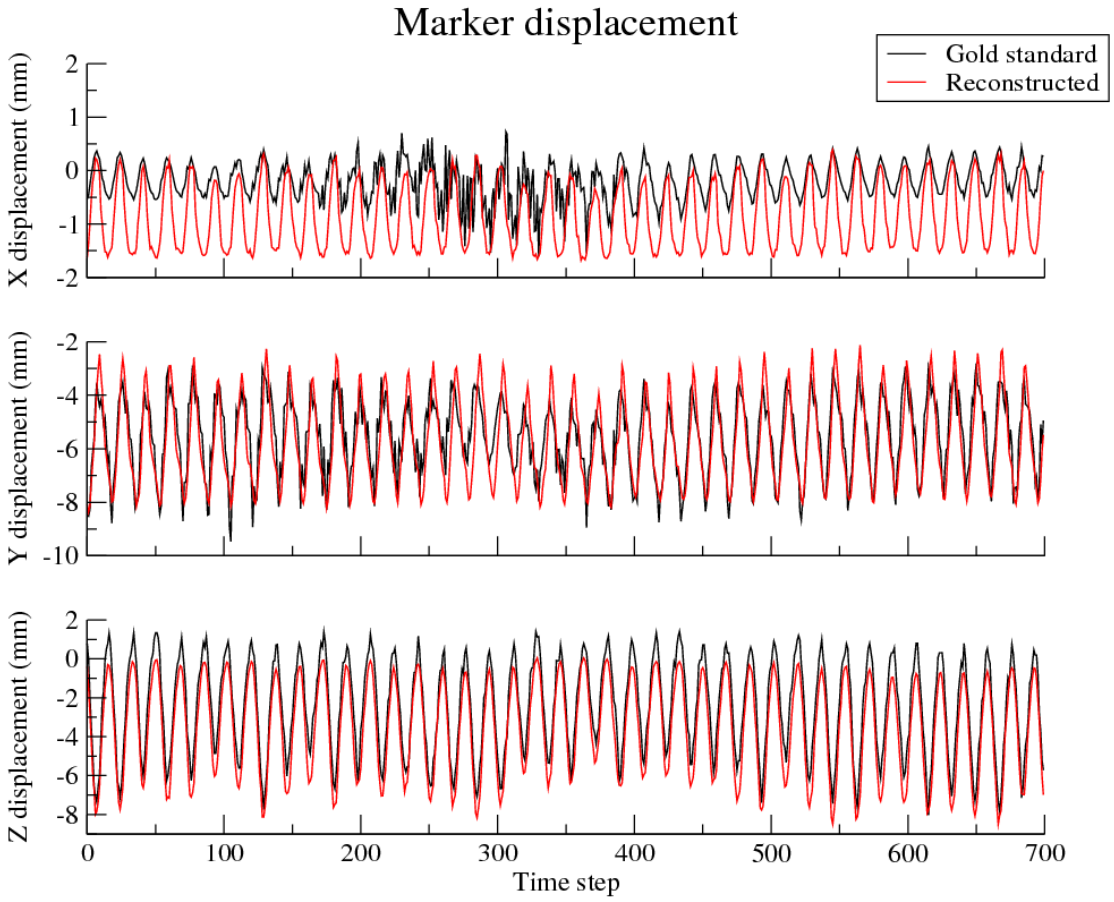


Figure 27: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 11).

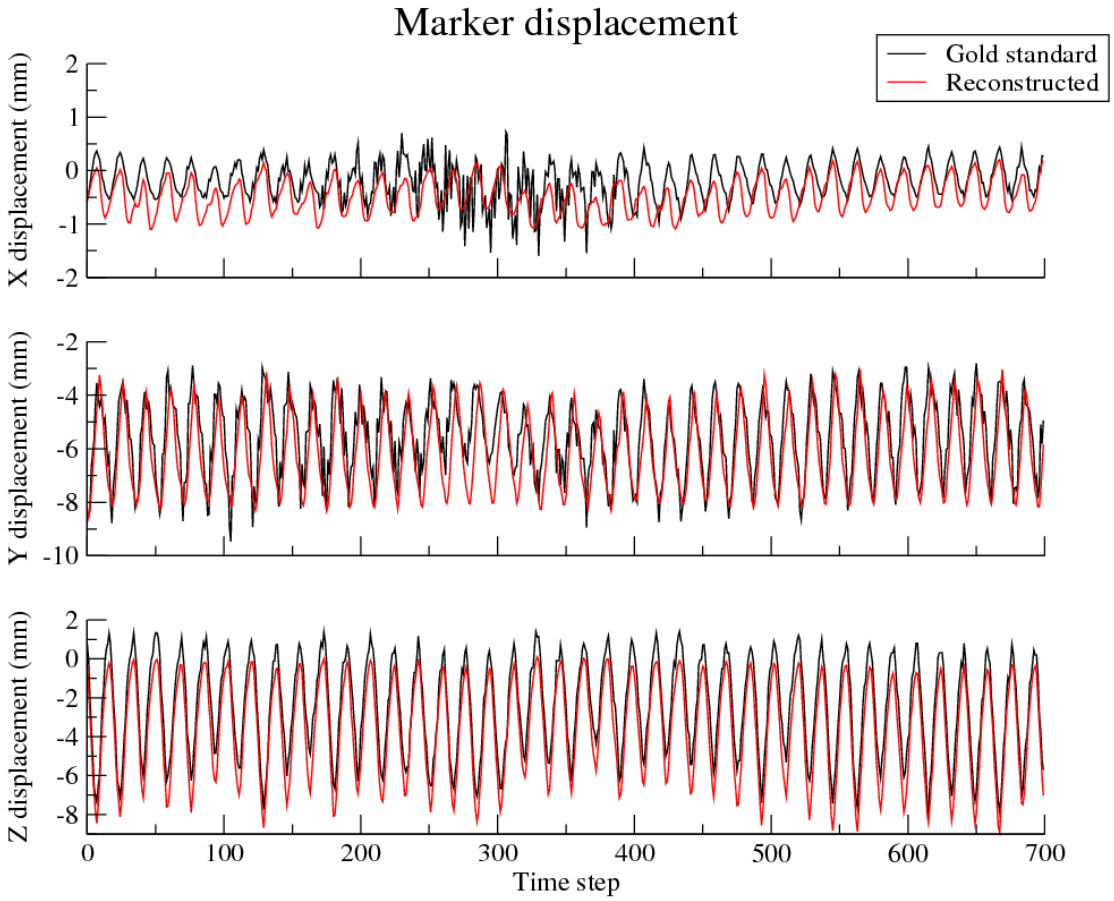


Figure 28: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use FIPM-bspline registration with 30 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 13).

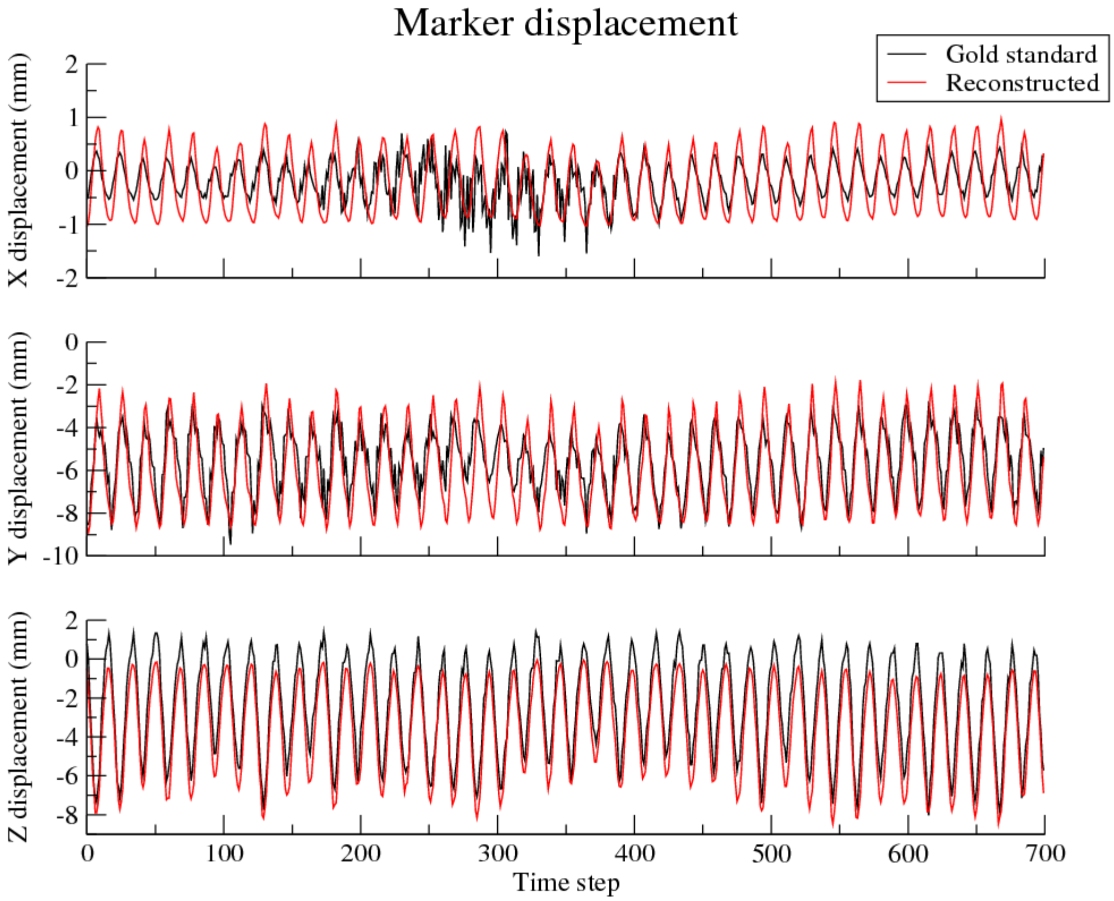


Figure 29: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use FIPM-bspline registration with 60 hybrid bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 14).

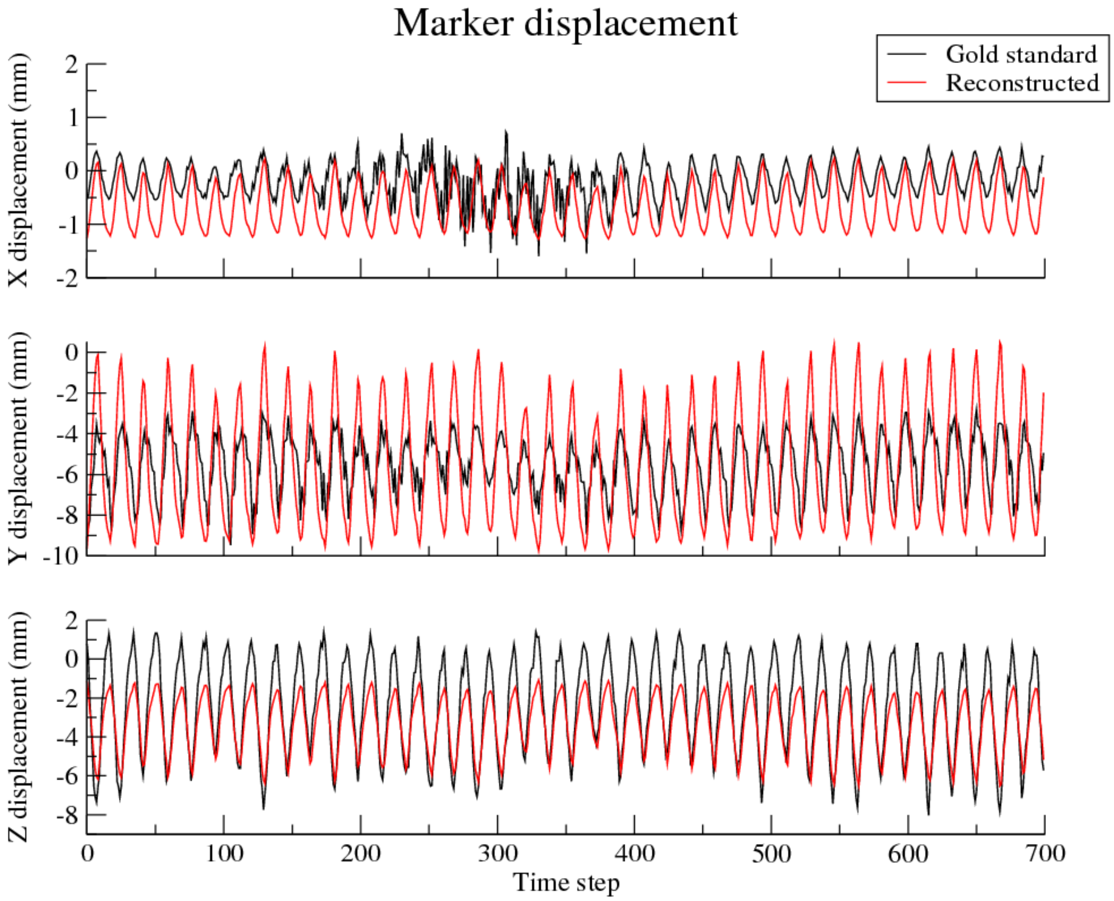


Figure 30: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use FIPM-bspline registration with 60 hybrid bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 1 mode for the motion model, and masked 4DCBCT projections as the fitting target (Test 15).

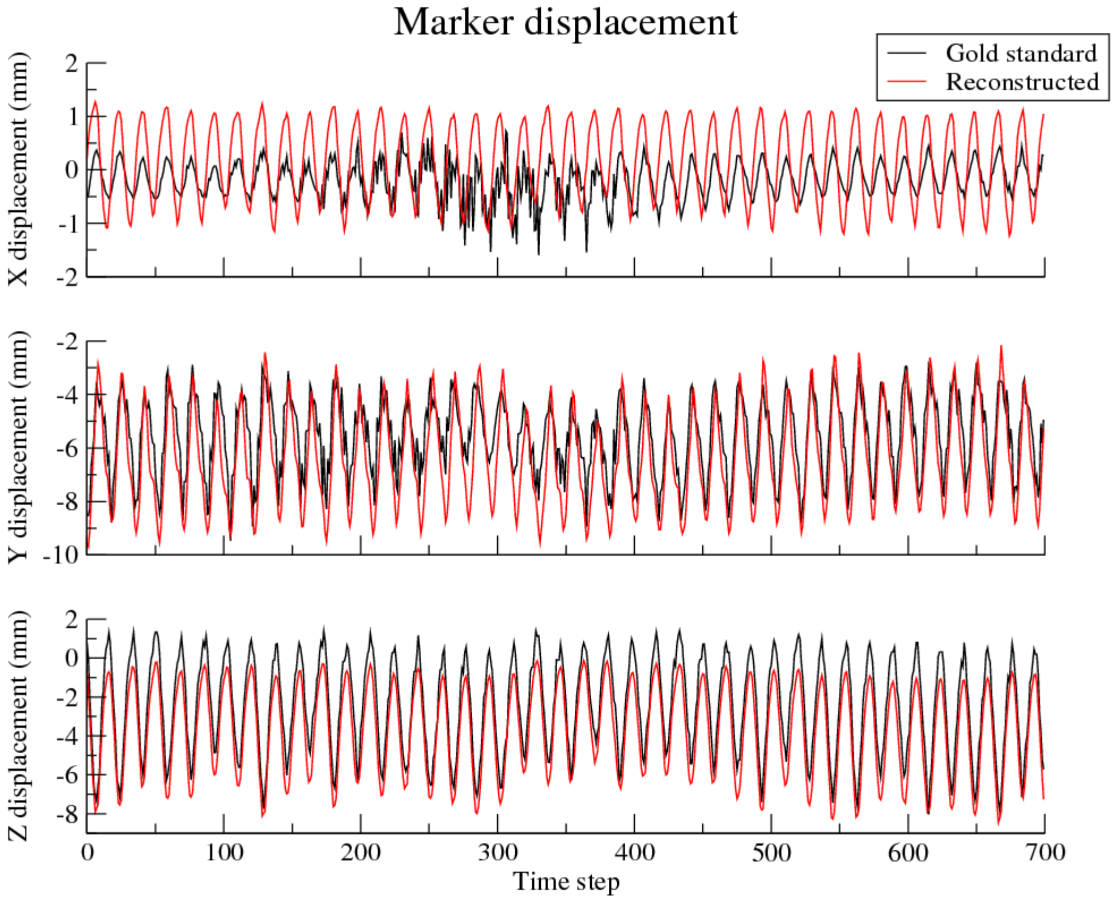


Figure 31: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories use FIPM-bspline registration with 60 hybrid bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 5 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 18).

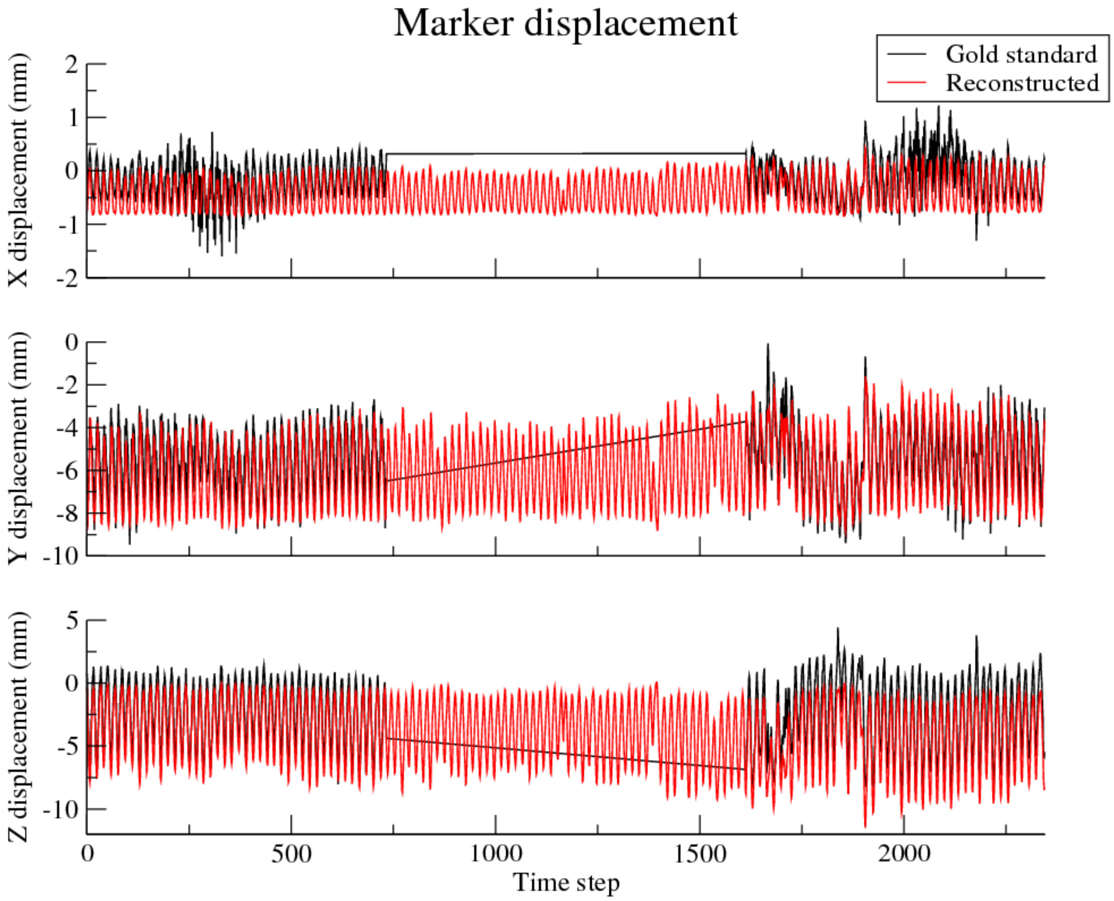


Figure 32: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories used FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the first 700 projections in the raw 4DCBCT data-set, and the fitting target projections were also sampled from the first 700 projections in the raw 4DCBCT data-set (Test 22).

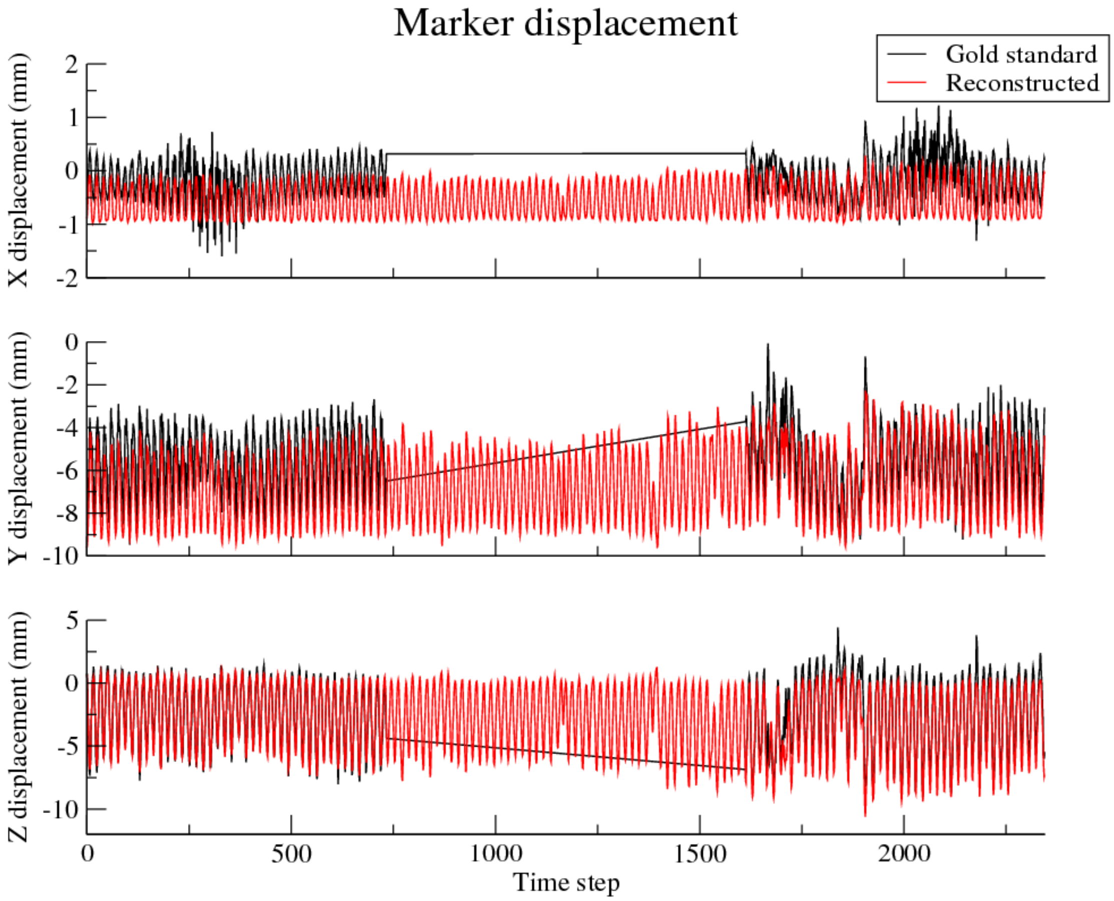


Figure 33: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories used FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the first 700 projections in the raw 4DCBCT data-set, and the fitting target projections were sampled from the last 700 projections in the raw 4DCBCT data-set (Test 23).

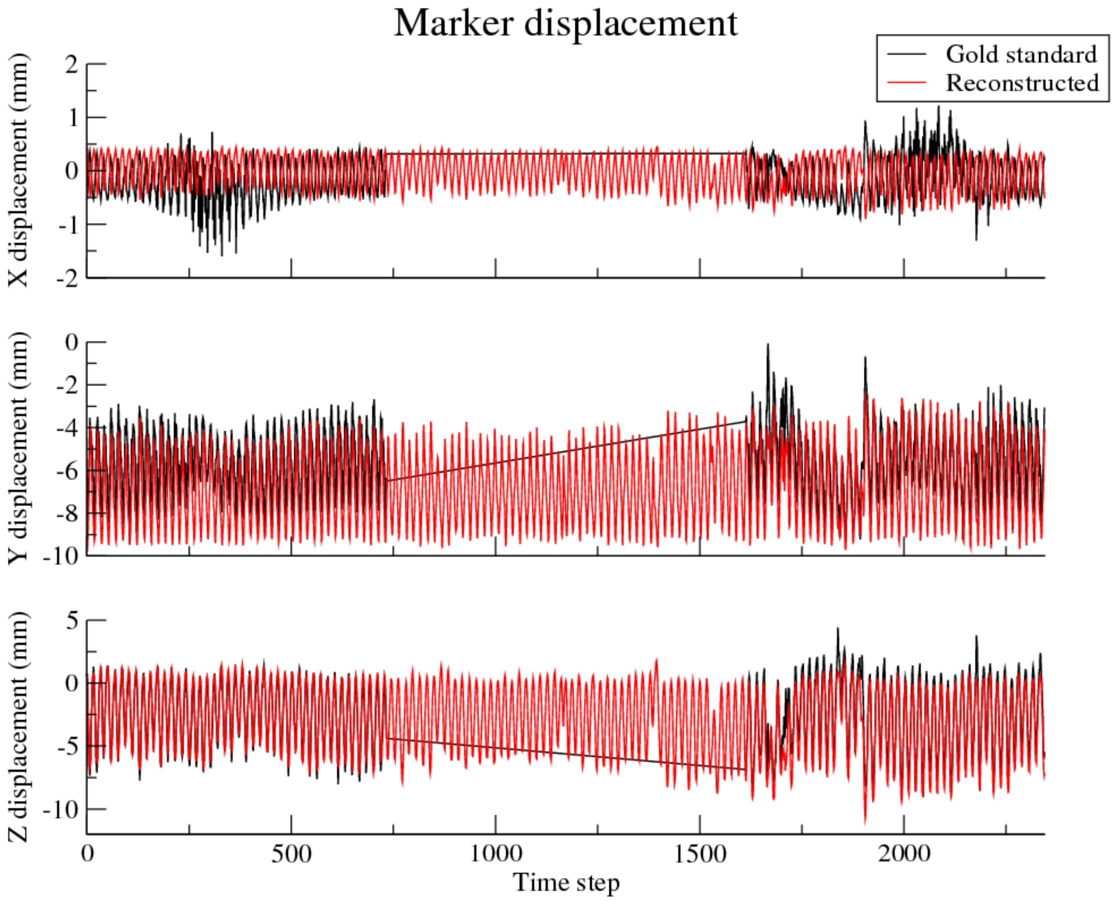


Figure 34: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories used FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the last 700 projections in the raw 4DCBCT data-set, and the fitting target projections were also sampled from the last 700 projections in the raw 4DCBCT data-set (Test 24).

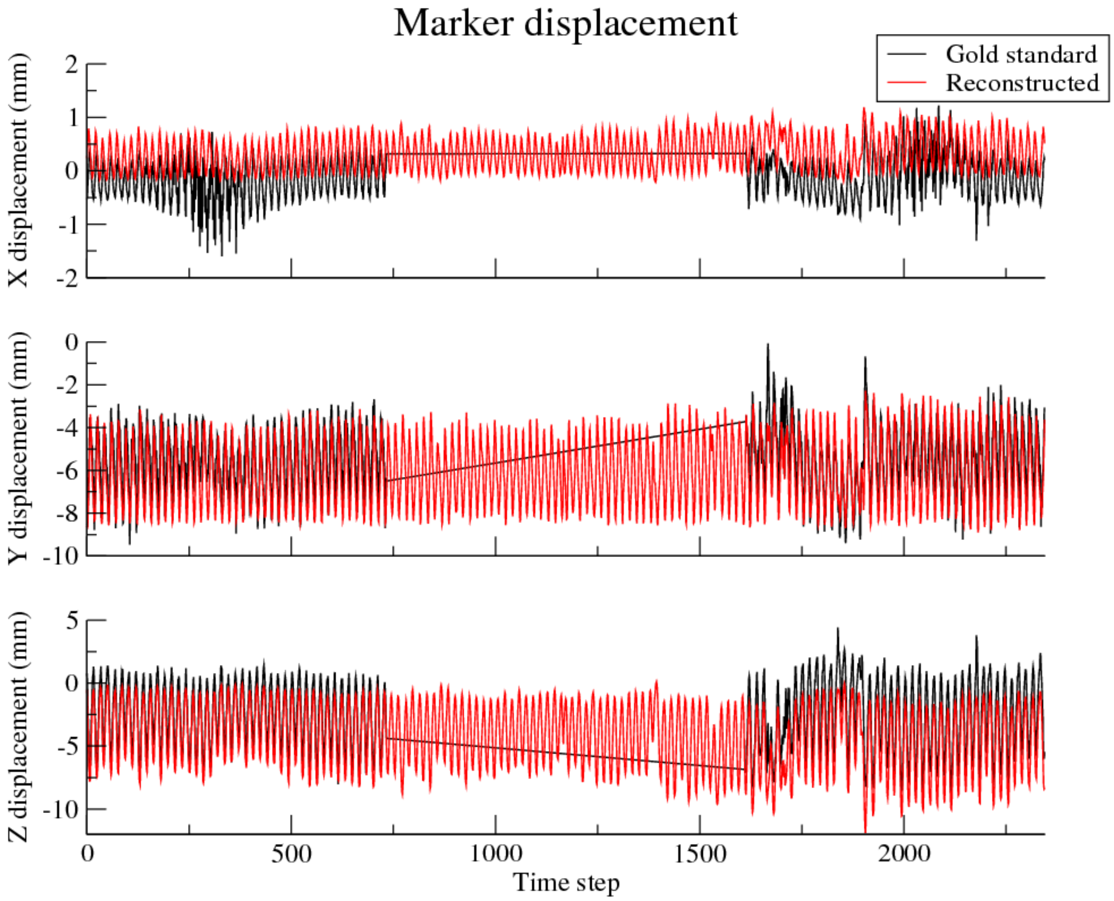


Figure 35: Comparison of gold standard and FIPM-PCA 4DCBCT reconstructed marker trajectories for each spatial component. Reconstructed trajectories used FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the last 700 projections in the raw 4DCBCT data-set, and the fitting target projections were sampled from the first 700 projections in the raw 4DCBCT data-set (Test 25).

Similarly, after reconstructed and ground truth 2D projected marker centroid positions were acquired using the method of Section 4.6.2 they were plotted together for comparison. Figure 36, Figure 37, Figure 38, Figure 39, Figure 40, Figure 41, Figure 42, Figure 43, Figure 44, and Figure 45 plot reconstructed and ground truth 2D projected marker position results for tests 10, 11, 13, 14, 15, 18, 22, 23, 24, and 25 respectively. In each figure results were plotted in the detector U direction (along the long axis of the detector, tangential to the circle swept out by the rotating gantry) and the detector V direction (along the short axis of the detector, pointing along the SI direction). They were left in their native units of detector pixels. For reference, each detector pixel was 0.388 x 0.388 mm, and, assuming the 3D object to be located at the machine isocenter, its projection would have been magnified by approximately a factor of 1.5. Thus a rough conversion from detector pixels to mm was 1 detector pixel ≈ 0.388/1.5 = 0.259 mm. As before Figure 36, Figure 37, Figure 38, Figure 39, Figure 40, and Figure 41 cover the first 700 time points of the 4DCBCT data-set while Figure 42, Figure 43, Figure 44, and Figure 45 span the full 2345 time points.

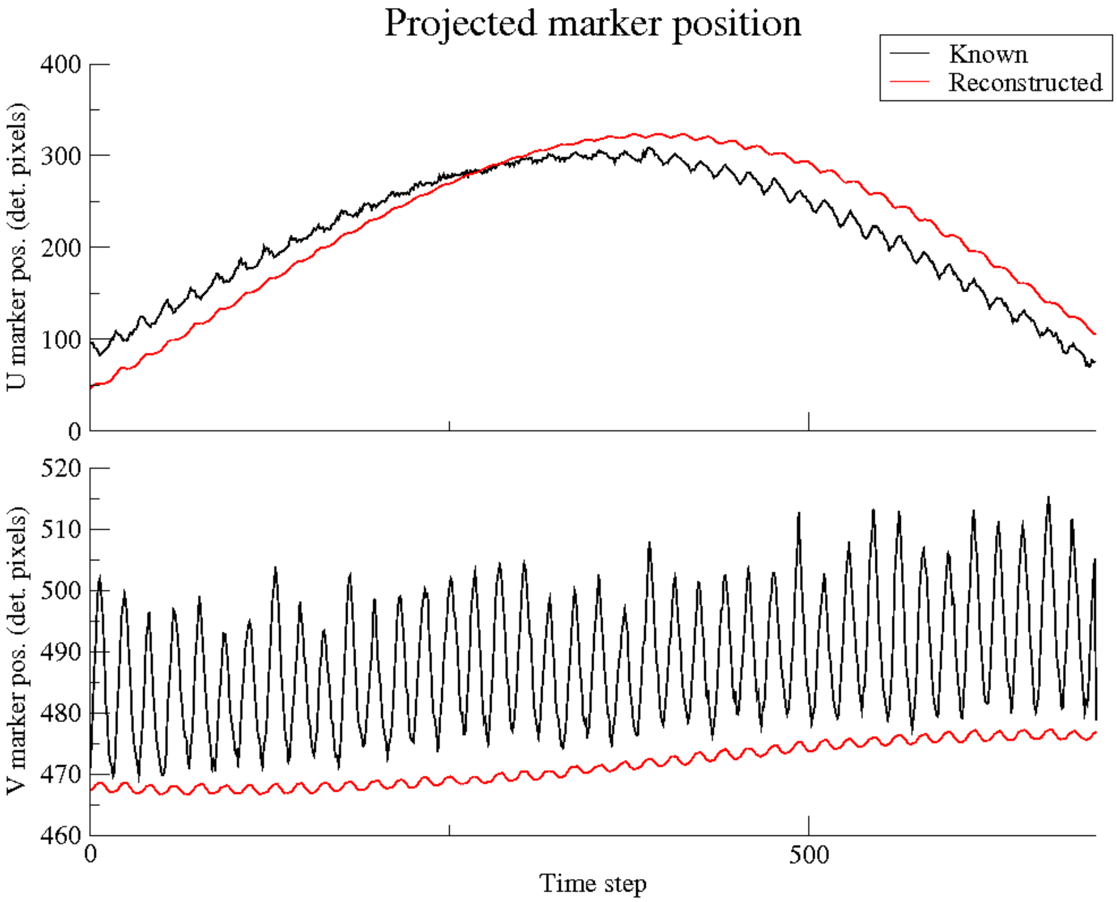


Figure 36: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 10 phase bins to generate the training set, parameterized functions of the breathing trace

coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 10).

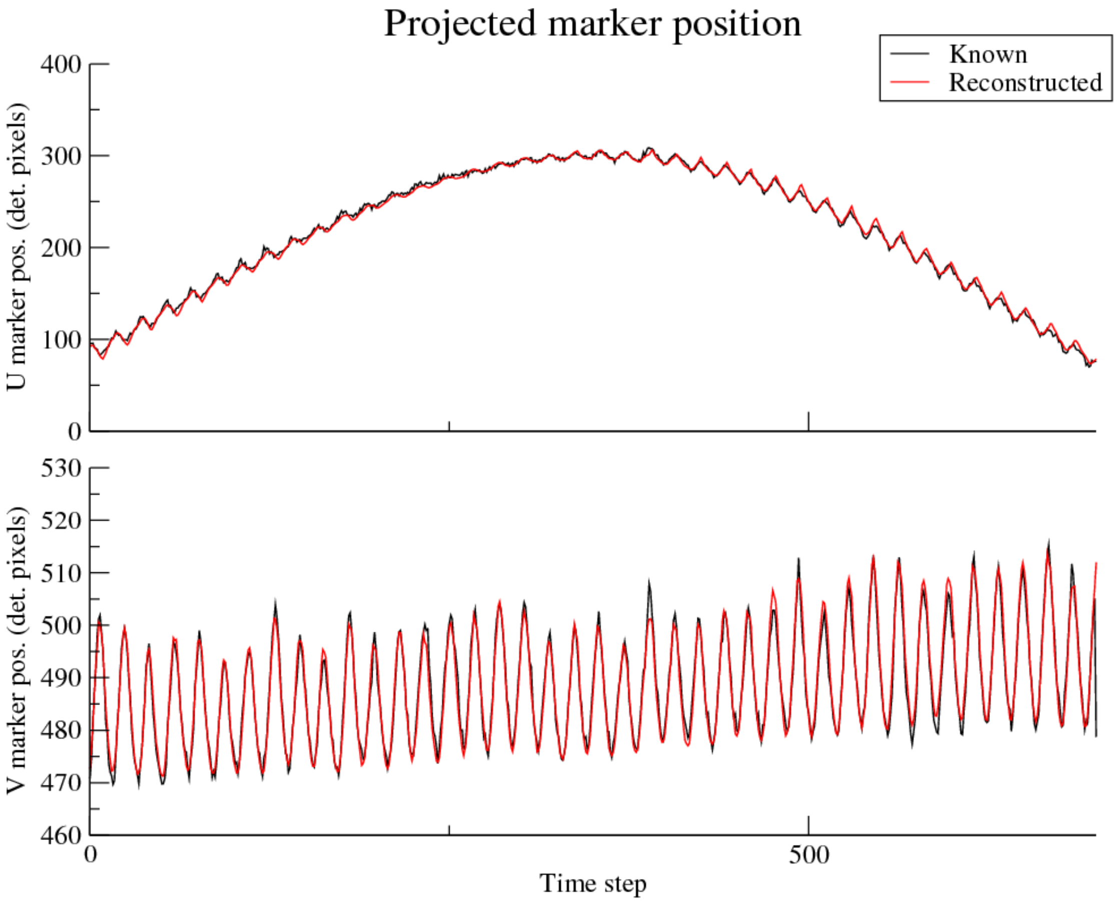


Figure 37: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 11).

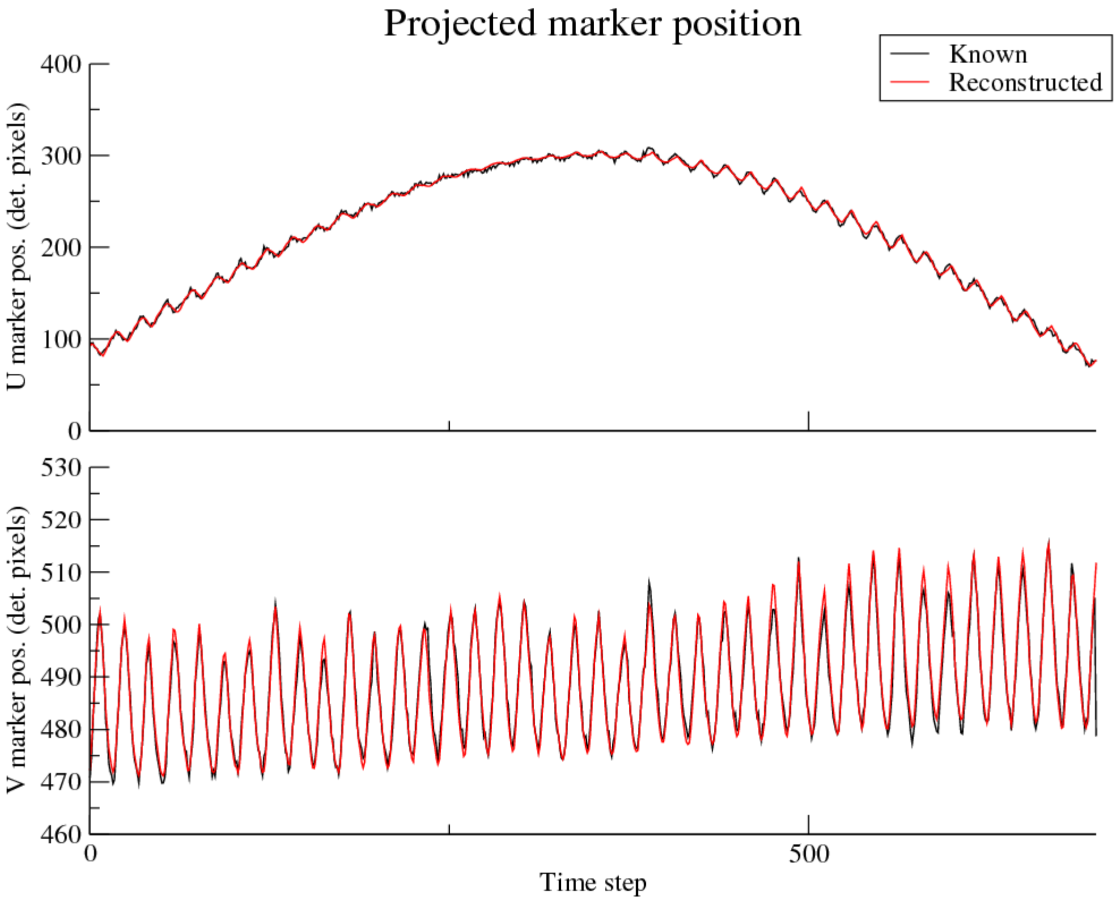


Figure 38: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 30 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 13).

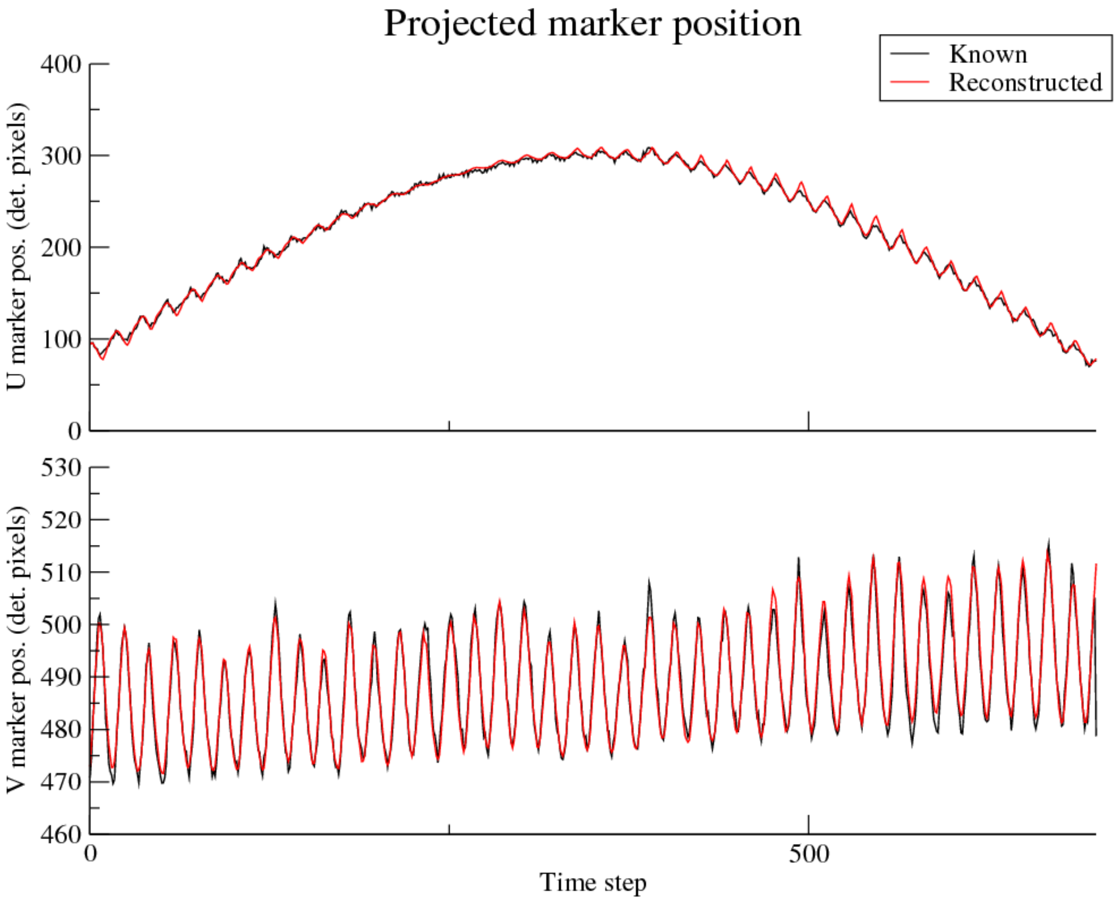


Figure 39: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 60 hybrid bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 14).

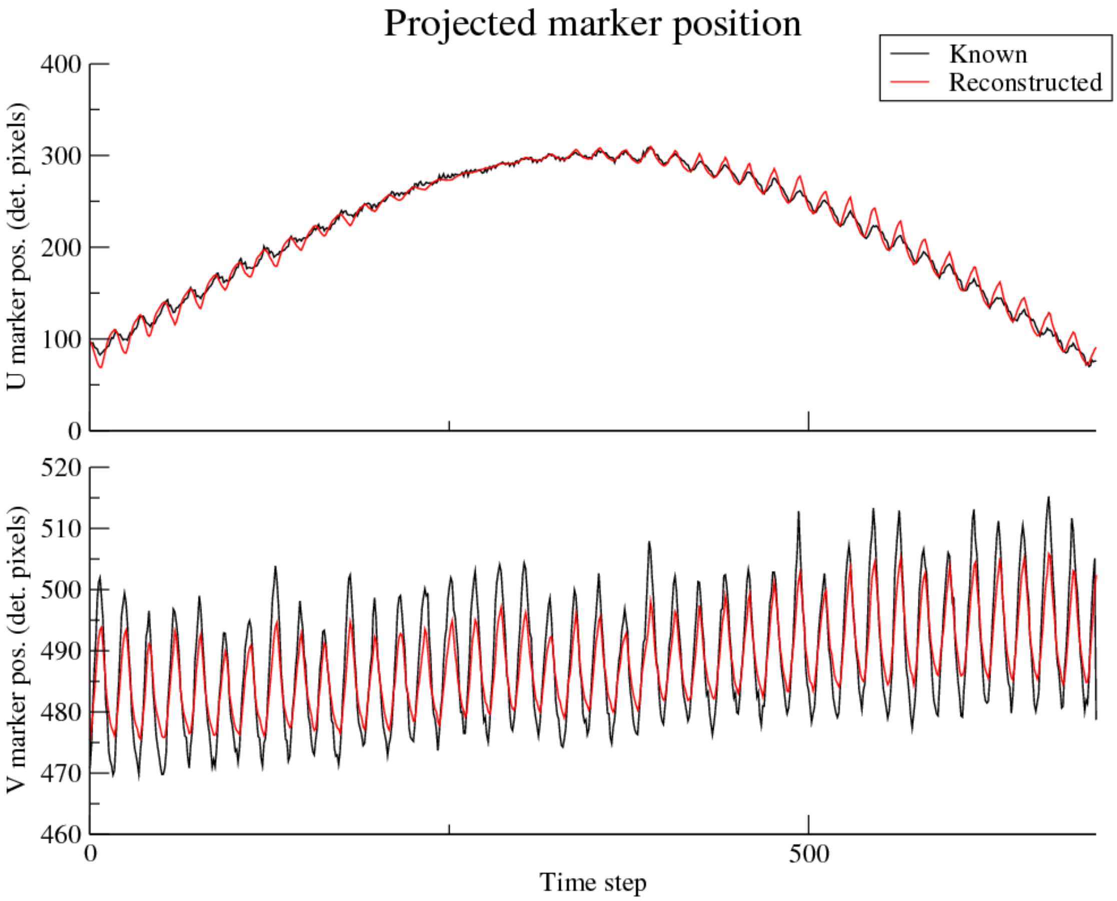


Figure 40: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 60 hybrid bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 1 mode for the motion model, and masked 4DCBCT projections as the fitting target (Test 15).

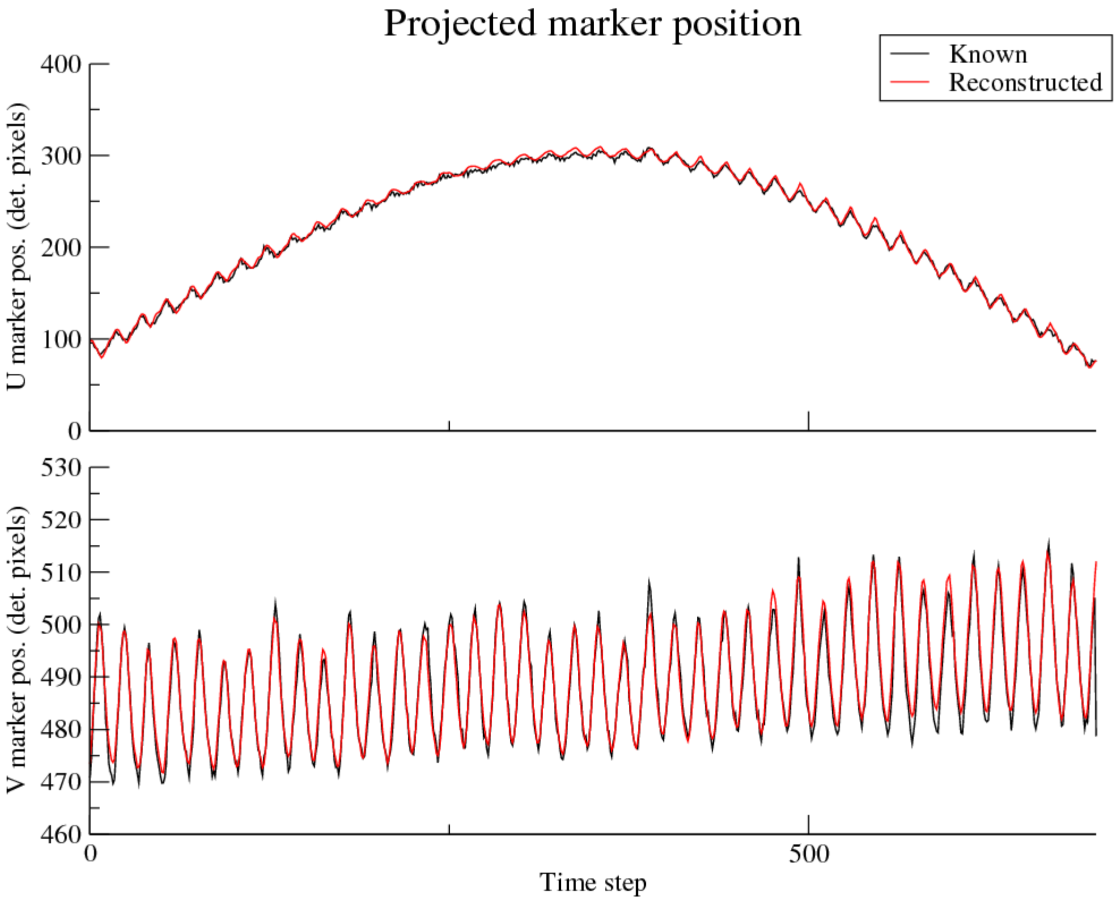


Figure 41: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 60 hybrid bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 5 modes for the motion model, and masked 4DCBCT projections as the fitting target (Test 18).

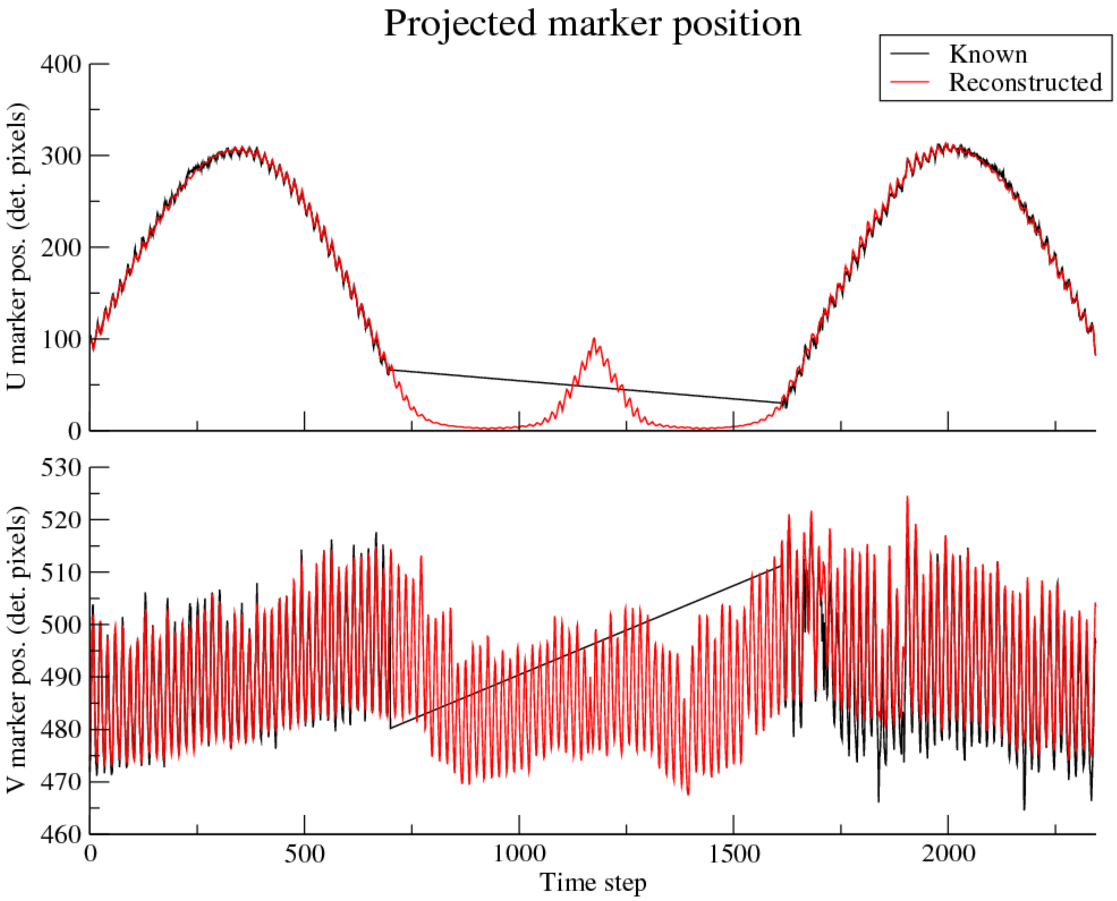


Figure 42: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the first 700 projections in the raw 4DCBCT data-set, and the fitting target projections were also sampled from the first 700 projections in the raw 4DCBCT data-set (Test 22).

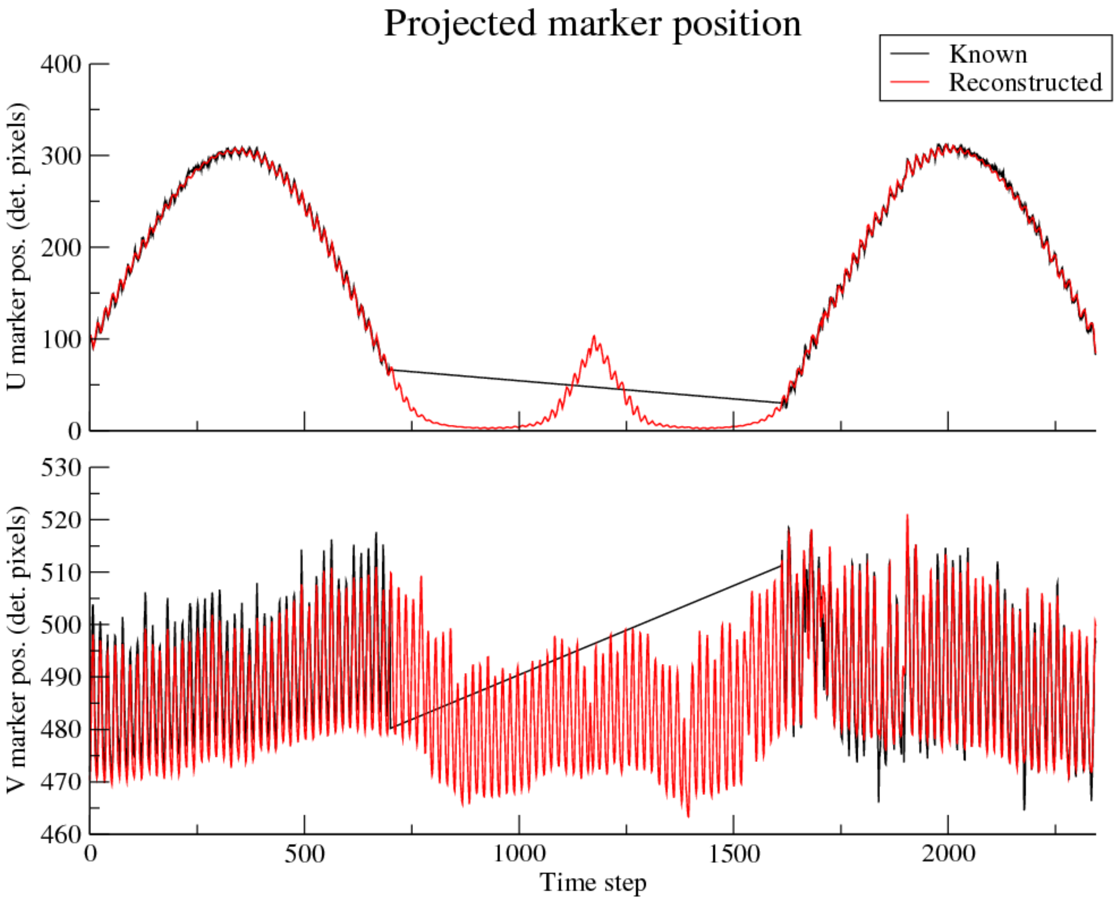


Figure 43: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the first 700 projections in the raw 4DCBCT data-set, and the fitting target projections were sampled from the last 700 projections in the raw 4DCBCT data-set (Test 23).

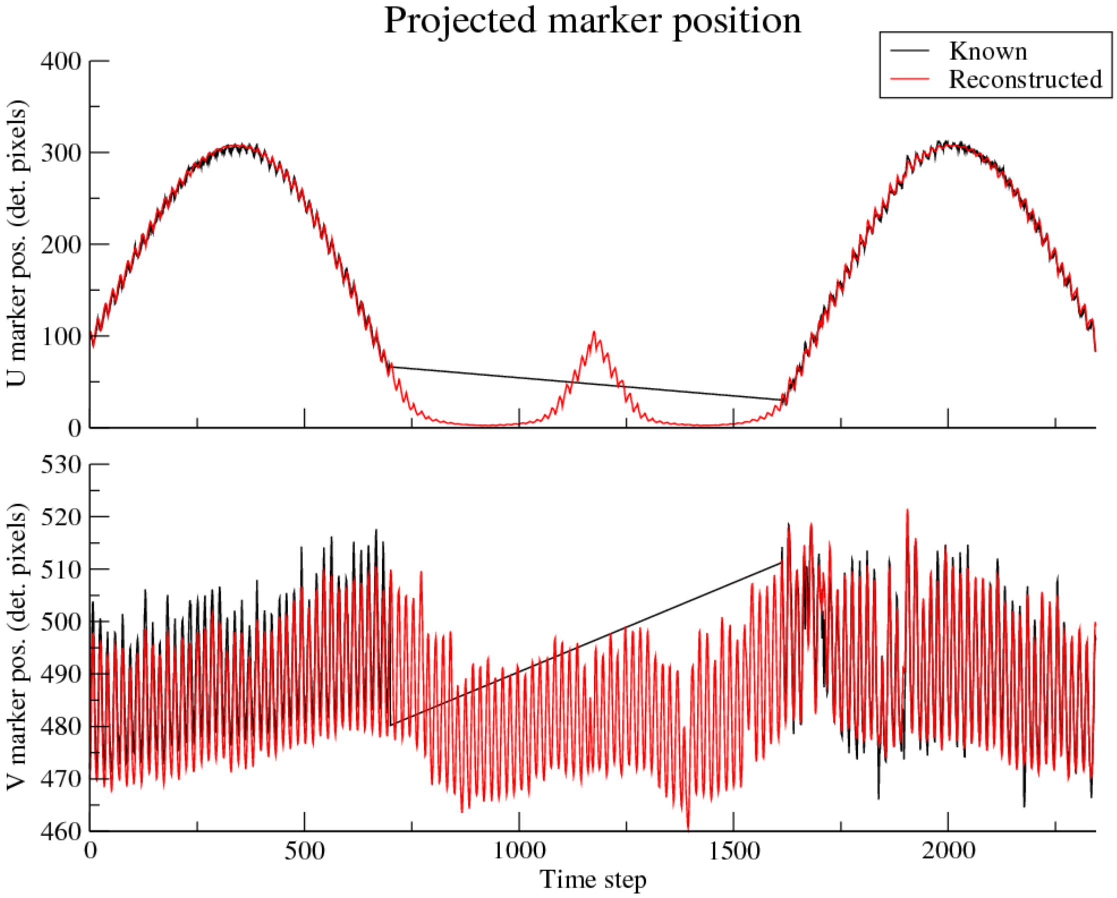


Figure 44: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the last 700 projections in the raw 4DCBCT data-set, and the fitting target projections were also sampled from the last 700 projections in the raw 4DCBCT data-set (Test 24).

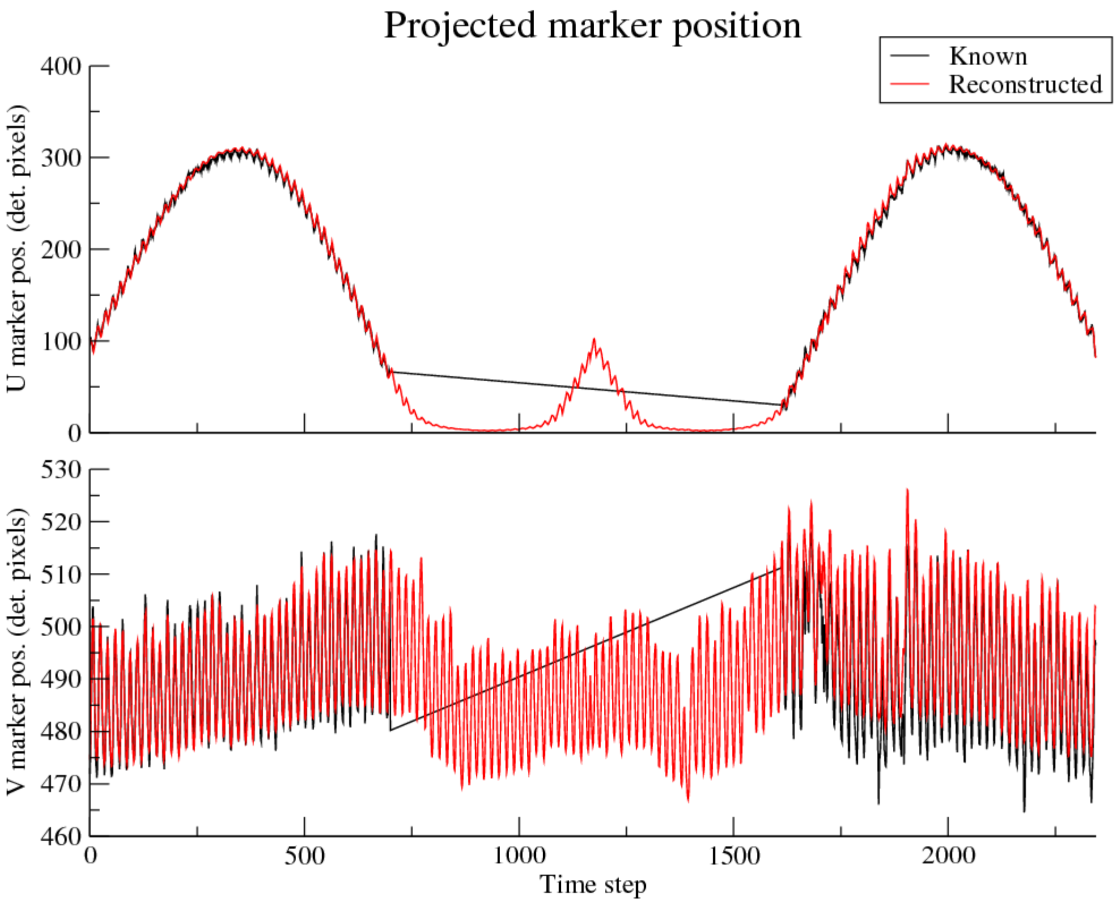


Figure 45: Comparison of 2D marker centroid locations in detector reference frame for known CBCT projections and DRRs cast through optimized FIPM-PCA 4DCBCT. FIPM-PCA 4DCBCTs were reconstructed using FIPM-bspline registration with 10 amplitude bins to generate the training set, parameterized functions of the breathing trace coupled to eigenvectors with 3 modes for the motion model, and masked 4DCBCT projections as the fitting target. Training data was limited to the last 700 projections in the raw 4DCBCT data-set, and the fitting target projections were sampled from the first 700 projections in the raw 4DCBCT data-set (Test 25).

4.7.2 Marker data tables

Table 5 describes the 3D marker trajectory results for all tests listed in Table 4. For each test it first lists a qualitative description of the plot of the reconstructed and gold standard data. It then lists the mean, standard deviation, and root mean square (RMS) of the disagreement between the two trajectories for the X, Y, and Z components of motion. The mean displays differences between the two data-sets due to systematic offsets, while the standard deviation captures differences in phase, amplitude, and waveform. The RMS serves as a combination of the two. For tests 19-25 three values are listed in each slot to provide temporally segmented results. The first value describes the mean, standard deviation, or RMS difference calculated over the first 700 data-points, the second value describes the same quantities calculated over only the final 700 data-points, and the third value calculates those quantities over all data-points (excluding times when the marker was out of frame).

Table 5: List of qualitative and quantitative results for real patient data test scenarios comparing reconstructed and gold standard 3D marker trajectories.

Test number	Qualitative marker trajectory observations	Trajectory difference mean (mm)	Trajectory difference standard deviation (mm)	RMS difference (mm)
Test 1	• Reconstructed trajectory tracks well with gold standard for all components, best for Z, worst for X.	• X: 5.94e-07 • Y: 2.37e-06 • Z: -6.78e-07	• X: 2.26e-01 • Y: 3.86e-01 • Z: 7.12e-01	• X: 0.226128 • Y: 0.385359 • Z: 0.711841
Test 2	• Reconstructed trajectory tracks well with gold standard for all components, best for Z, worst for X.	• X: -9.81e-08 • Y: 1.37e-06 • Z: -5.15e-07	• X: 2.27e-01 • Y: 6.66e-01 • Z: 4.07e-01	• X: 0.226721 • Y: 0.665752 • Z: 0.406924
Test 3	• Reconstructed trajectory tracks well with gold standard for all components, best for Z, worst for X.	• X: 1.25e-07 • Y: 6.87e-07 • Z: -2.34e-07	• X: 2.26e-01 • Y: 7.12e-01 • Z: 3.86e-01	• X: 0.225908 • Y: 0.711841 • Z: 0.385359
Test 4	• Reconstructed trajectory tracks well with gold standard for all components, best for Z, worst for X.	• X: -4.46e-06 • Y: 5.46e-07 • Z: -5.98e-06	• X: 2.48e-01 • Y: 6.62e-01 • Z: 4.43e-01	• X: 0.248054 • Y: 0.661302 • Z: 0.442314
Test 5	• Reconstructed trajectory tracks well with gold	• X: 1.50e-06 • Y: -1.07e-07	• X: 2.59e-01 • Y: 6.99e-01	• X: 0.258994 • Y: 0.698203

	standard for all components, best for Z, worst for X.	• Z: -7.04e-07	• Z: 4.17e-01	• Z: 0.417151
Test 6	• Reconstructed trajectory doesn't track quite as well with gold standard for all components, best for Z, worst for Y.	• X: 5.94e-06 • Y: -2.39e-07 • Z: -3.25e-06	• X: 2.27e-01 • Y: 4.07e-01 • Z: 6.66e-01	• X: 0.226544 • Y: 0.406912 • Z: 0.665781
Test 7	• Reconstructed trajectory tracks well with gold standard for Y and Z components. X component has significant waveforms differences.	• X: -3.58e-01 • Y: 4.97e-02 • Z:-1.47e-03	• X: 4.01e-01 • Y: 7.64e-01 • Z: 4.83e-01	• X: 0.537318 • Y: 0.764875 • Z:0.482190
Test 8	• Clear differences in all 3 directions, overshoot in X direction, undershoot in Y and Z directions.	• X: -4.07e-01 • Y: -3.64e-02 • Z: 1.34e-01	• X: 7.03e-01 • Y: 2.06e+00 • Z: 1.29e+00	• X: 0.811502 • Y: 2.063412 • Z: 1.298394
Test 9	• X is pretty good, maybe slight offset miss, Y looks decent but has small but definite offset miss, Z is good except for the usual offset miss.	• X: 1.30e-01 • Y: -1.07e+00 • Z: -1.14e+00	• X: 2.40e-01 • Y: 7.52e-01 • Z: 4.46e-01	• X: 0.272941 • Y: 1.308146 • Z: 1.221108
Test 10	• Total offset miss in the X direction, amplitudes are far off in Y and Z directions.	• X: -3.02e+00 • Y: -7.17e-01 • Z: -3.34e-01	• X: 5.00e-01 • Y: 1.44e+00 • Z: 2.25e+00	• X: 3.063418 • Y: 1.609857 • Z: 2.271602
Test 11	• Fairly large amplitude miss in X, smaller amplitude misses in Y and Z.	• X: -6.56e-01 • Y: -2.28e-01 • Z: -1.07e+00	• X: 6.21e-01 • Y: 1.13e+00 • Z: 8.33e-01	• X: 0.782153 • Y: 0.877719 • Z: 1.149748
Test 12	• Clear offset miss in X, Y looks good except for waveform differences, Z looks good except for usual offset.	• X: -8.64e-01 • Y: -5.15e-01 • Z: -1.02e+00	• X: 3.65e-01 • Y: 8.60e-01 • Z: 3.84e-01	• X: 0.938167 • Y: 1.001395 • Z: 1.090238
Test 13	• Amplitude and offset miss in X direction, waveform incorrect in some spots in Y direction, usual small offset in Z direction.	• X: -2.84e-01 • Y: -6.48e-01 • Z: -1.11e+00	• X: 3.00e-01 • Y: 7.66e-01 • Z: 3.87e-01	• X: 0.412556 • Y: 1.003235 • Z: 1.173118
Test 14	• Amplitude somewhat	• X: -2.74e-02	• X: 4.00e-01	• X: 0.400234

	wrong in X direction, waveform incorrect in some spots in Y direction, usual small offset in Z direction.	• Y: -4.13e-01 • Z: -1.13e+00	• Y: 9.74e-01 • Z: 4.04e-01	• Y: 1.057032 • Z: 1.198411
Test 15	• Amplitude and offset misses in X, amplitude miss in Y, amplitude and usual offset miss in Z.	• X: -4.60e-01 • Y: -4.90e-03 • Z: -8.69e-01	• X: 2.90e-01 • Y: 1.94e+00 • Z: 1.06e+00	• X: 0.543522 • Y: 1.937586 • Z: 1.367888
Test 16	• Offset miss in X, Y is pretty good beyond waveform differences, Z good besides usual offset.	• X: -4.99e-01 • Y: -6.60e-01 • Z: -1.08e+00	• X: 2.71e-01 • Y: 8.31e-01 • Z: 3.92e-01	• X: 0.567820 • Y: 1.060678 • Z: 1.149069
Test 17	• Mostly waveform miss in X and Y, good in Z except for usual offset.	• X: -1.91e-01 • Y: -8.91e-01 • Z: -1.11e+00	• X: 4.00e-01 • Y: 7.07e-01 • Z: 3.98e-01	• X: 0.443175 • Y: 1.137718 • Z: 1.182145
Test 18	• Amplitude and minor offset error in X, waveform and minor amplitude error in Y, usual offset error in Z.	• X: 3.50e-01 • Y: -8.42e-01 • Z: -1.23e+00	• X: 5.77e-01 • Y: 8.31e-01 • Z: 4.10e-01	• X: 0.674610 • Y: 1.182225 • Z: 1.295403
Test 19	• Offset miss in both sections in X, pretty good in first section with waveform differences more prominent for second section in Y, usual offset miss in Z but bigger for second section than first.	• X: (-8.67e-01, -7.37e-01, -8.02e-01) • Y: (-5.19e-01, -1.70e-01, -3.45e-01) • Z: (1.00e+00, -2.08e+00, -1.54e+00)	• X: (3.68e-01, 4.43e-01, 4.12e-01) • Y: (8.64e-01, 1.17e+00, 1.04e+00) • Z: (3.99e-01, 7.09e-01, 7.87e-01)	• X: (0.941319, 0.859742, 0.901482) • Y: (1.007100, 1.182243, 1.098110) • Z: (1.080324, 2.197397, 1.731063)
Test 20	• Offset undershoot in both sections for X, offset undershoot in first section with good fit in second section for Y, no offset in first section with typical offset in second section for Z.	• X: (-8.23e-01, -7.11e-01, -7.67e-01) • Y: (-1.10e+00 -5.74e-01, -8.36e-01) • Z: (1.78e-01, -8.64e-01, -3.43e-01)	• X: (3.04e-01, 3.76e-01, 3.46e-01) • Y: (7.14e-01, 9.69e-01, 8.90e-01) • Z: (4.37e-01, 7.28e-01, 7.95e-01)	• X: (0.876673, 0.804042, 0.841167) • Y: (1.309937, 1.125397, 1.221220) • Z: (0.471192, 1.129727, 0.865328)
Test 21	• Similar offset miss in both sections for X, small offset miss in first section and	• X: (-8.27e-01, -7.79e-01, -8.03e-01)	• X: (2.54e-01, 2.88e-01, 2.73e-01)	• X: (0.864791, 0.830461, 0.847812)

	pretty good in second section for Y, small offset miss in first section with larger offset miss in second section for Z.	• Y: (-1.01e+00, -4.65e-01, -7.37e-01) • Z: (-4.84e-01, -1.44e+00, -9.64e-01):	• Y: (7.14e-01, 9.84e-01, 9.01e-01) • Z: (3.99e-01, 7.09e-01, 7.49e-01)	• Y: (1.235932, 1.087428, 1.164101) • Z: (0.627017, 1.608610, 1.220508)
Test 22	• Offset miss in both sections but worse for first section in X, good fit for both in Y, offset miss in both sections but larger for second section for Z.	• X: (-2.75e-01, -2.91e-01, -2.83e-01) • Y: (-8.36e-01, -2.62e-01, -5.49e-01) • Z: (-9.78e-01, -1.94e+00, -1.46e+00)	• X: (2.44e-01, 3.13e-01, 2.81e-01) • Y: (7.00e-01, 9.79e-01, 8.98e-01) • Z: (4.09e-01, 7.23e-01, 7.60e-01)	• X: (0.367101, 0.427534, 0.398445) • Y: (1.089817, 1.012982, 1.052128) • Z: (1.060040, 2.072902, 1.645971)
Test 23	• Offset miss in both section but worse for second section in X, offset miss in first section and pretty good in second section for Y, no offset miss in first section with typical offset miss in second section for Z.	• X: (-4.06e-01, -4.36e-01, -4.21e-01) • Y: (1.53e+00, -9.15e-01, -1.22e+00) • Z: (6.15e-02, -8.76e-01, -4.07e-01)	• X: (2.55e-01, 3.34e-01, 2.97e-01) • Y: (7.17e-01, 9.93e-01, 9.19e-01) • Z: (4.32e-01, 7.41e-01, 7.66e-01)	• X: (0.479227, 0.548988, 0.515266) • Y: (1.689711, 1.349992, 1.529429) • Z: (0.435730, 1.147194, 0.867510)
Test 24	• No offset miss but bad phase miss in both sections for X, amplitude and small offset miss in both sections for Y, no offset miss in first section with typical offset miss in second section for Z.	• X: (2.32e-01, -6.22e-03, 1.13e-01) • Y: (-1.59e+00, -1.20e+00, -1.40e+00) • Z: (1.85e-01, -9.012e-01, -3.58e-01)	• X: (5.81e-01, 6.60e-01, 6.33e-01) • Y: (7.76e-01, 1.08e+00, 9.60e-01) • Z: (4.50e-01, 7.35e-01, 8.17e-01)	• X: (0.625314, 0.659519, 0.642632) • Y: (1.772427, 1.615938, 1.696041) • Z: (0.486621, 1.162829, 0.891126)
Test 25	• Offset miss in both sections for X, good fit for both sections with worse waveform difference for second section in Y, small offset for first section with bigger offset for second section in Z.	• X: (4.37e-01, 4.70e-01, 4.53e-01) • Y: (-7.75e-01, -5.26e-01, -6.51e-01) • Z: (-1.11e+00, -	• X: (3.43e-01, 4.08e-01, 3.77e-01) • Y: (7.58e-01, 1.08e+00, 9.42e-01) • Z: (4.20e-01, 7.28e-01,	• X: (0.555544, 0.622024, 0.589699) • Y: (1.084385, 1.202497, 1.144924) • Z: (1.184700, 2.288478,

		2.17e+00, -1.64e+00)	7.97e-01)	1.821817)

Table 6 describes the 2D projected marker position results for tests 9-25 listed in Table 4. For each entry it also gives a qualitative description of the corresponding plot followed by the mean, standard deviation, and RMS of the error of the reconstructed data with reference to the ground truth data. As before three values were supplied for each metric for tests 19-25.

Table 6: List of qualitative and quantitative results for comparison of marker projection centroids between ground truth 4DCBCT data and model generated DRRs.

Test number	Qualitative marker projection centroid observations	Marker projection centroid error mean (detector pixels)	Marker projection centroid error standard deviation (detector pixels)	Marker projection centroid error RMS (detector pixels)
Test 9	• Good fit with largest disagreement in the center for U, small offset in V.	• U: 4.10 • V: 1.79	• U: 4.46 • V: 1.91	• U: 6.05 • V: 2.62
Test 10	• Large phase and offset miss in U, large amplitude and offset miss in V.	• U: 8.86 • V: -17.22	• U: 30.25 • V: 8.81	• U: 31.50 • V: 19.34
Test 11	• Excellent agreement for both U and V.	• U: -0.25 • V: 0.07	• U: 3.13 • V: 1.75	• U: 3.14 • V: 1.75
Test 12	• Excellent agreement for both U and V.	• U: -0.83 • V: -0.12	• U: 3.06 • V: 1.64	• U: 3.17 • V: 1.65
Test 13	• Excellent agreement for both U and V.	• U: 0.12 • V: 0.23	• U: 2.44 • V: 1.65	• U: 2.44 • V: 1.67

Test 14	• Excellent agreement for both U and V.	• U: 0.88 • V: 0.31	• U: 3.00 • V: 1.72	• U: 3.12 • V: 1.75
Test 15	• Large amplitude misses in both U and V.	• U: 0.46 • V: -0.72	• U: 6.43 • V: 4.17	• U: 6.44 • V: 4.23
Test 16	• Excellent agreement for both U and V.	• U: -0.16 • V: 0.13	• U: 2.55 • V: 1.66	• U: 2.56 • V: 1.67
Test 17	• Excellent agreement for both U and V.	• U: 0.04 • V: 0.26	• U: 2.57 • V: 1.63	• U: 2.57 • V: 1.65
Test 18	• Good agreement in U, slight amplitude miss in V.	• U: 1.61 • V: 0.73	• U: 2.81 • V: 1.76	• U: 3.24 • V: 1.90
Test 19	• Excellent agreement for both sections in U, excellent agreement in first section with good agreement with small offset miss in second section for V.	• U: (-0.83, -0.36, -0.59) • V: (-0.12, 4.22, 2.10)	• U: (3.06, 4.13, 3.66) • V: (1.64, 3.05, 3.28)	• U: (3.17, 4.14, 3.70) • V: (1.65, 5.20, 3.90)
Test 20	• Excellent agreement for both sections in U, small offset miss in first section with good agreement in second section for V.	• U: (-1.03, -0.60, -0.81) • V: (-4.66, -0.48, -2.52)	• U: (2.54, 3.26, 2.94) • V: (1.73, 3.02, 3.24)	• U: (2.74, 3.31, 3.04) • V: (4.97, 3.05, 4.10)
Test 21	• Excellent agreement for both sections in U, small offset miss in different directions for both sections in V.	• U: (-0.96, -0.60, -0.77) • V: (-2.10, 1.76, -0.12)	• U: (2.49, 3.41, 3.00) • V: (1.61, 3.00, 3.09)	• U: (2.66, 3.46, 3.09) • V: (2.64, 3.47, 3.10)
Test 22	• Excellent agreement	• U: (0.19, 0.69, 0.45)	• U: (2.48, 3.82, 3.24)	• U: (2.49, 3.88, 3.27)

	for both sections in U, good agreement in first section with small offset miss in second section for V.	• V: (-0.14, 3.58, 1.76)	• V: (1.81, 3.19, 3.20)	• V: (1.81, 4.79, 3.65)
Test 23	• Excellent agreement for both sections in U, small offset miss in first section with good agreement in second section for V.	• U: (-0.44, -0.038, -0.23) • V: (-4.13, -0.55, -2.30)	• U: (2.70, 3.31, 3.03) • V: (1.87, 3.20, 3.19)	• U: (2.74, 3.31, 3.04) • V: (4.54, 3.25, 3.93)
Test 24	• Excellent agreement for both sections in U, small offset miss in first section with good agreement in second section for V.	• U: (0.72, 0.47, 0.59) • V: (-4.61, -0.43, -2.47)	• U: (3.55, 3.70, 3.63) • V: (1.93, 3.16, 3.36)	• U: (3.62, 3.73, 3.67) • V: (5.00, 3.19, 4.17)
Test 25	• Excellent agreement for both sections in U, good agreement in first section with small offset miss in second section for V.	• U: (1.44, 1.69, 1.57) • V: (0.37, 4.49, 2.48)	• U: (2.79, 3.63, 3.25) • V: (1.88, 3.21, 3.35)	• U: (3.14, 4.00, 3.61) • V: (1.92, 5.52, 4.17)

4.8 Discussion

Before beginning our analysis of the patient data reconstruction results it is important to note that our 3D validation data, the independently reconstructed 3D marker trajectory, is a gold standard rather than a ground truth. It itself is an estimate derived from the same projection data that feeds our algorithm. Therefore, we refer to discrepancies between those two data-sets as differences, rather than errors. It is impossible to know which data is truly correct, however we assert that agreement between the two sets is a valuable indicator that both methods are performing well. A further complication is that in order to compare the two sets of data their imaging reference frames must be precisely aligned. We believe that misalignment of this type is the cause of the persistent offset in the Z direction observed in tests 9-25. Indeed, in our 2D validation data comparing projected marker positions we see no evidence of such a shift in the V direction, which is essentially the same as the Z direction. Thus we believe that this is an artificial difference due to misalignment of the two geometries rather than a true discrepancy between the 3D validation and reconstructed data.

In addition we note that all tests are carried out on the same patient data-set. Differences in results are due solely to the different scenarios under which reconstructions were performed.

4.8.1 General observations

Based on the 3D marker trajectory difference data and plots, we can divide our tests into two rough groups: those fit to the gold standard data, and those fit to the masked 4DCBCT projections.

The first group encompasses test runs 1-8 and is characterized almost uniformly by tiny systematic offsets (difference mean), as well as submillimeter discrepancies due to phase/amplitude/waveform disagreement (difference standard deviation). The results plotted in Figure 23 confirm this story. Qualitatively we observe an excellent agreement for the X component of motion, with observable discrepancy only in the center where the gold standard trajectory becomes noisy. The Y component also displays close agreement, with small differences in amplitude apparent for some breathing cycles. The Z component shows the closest match of the three, with the two trends nearly indistinguishable. This tells us that our formulation of the motion model is sufficient to accurately represent the gold standard data. In doing so it helps lay the groundwork for ascertaining the effects of fitting to actual projection data by effectively ruling out one mode of failure. The lone outlier in this group is test 8 (results plotted in Figure 25), which uses eigenvectors trained on manually tracked marker locations within the backprojection reconstructed 4DCBCT, and fits to 2D geometric projections of the 3D gold standard marker trajectory. This test suffers from visible amplitude and phase discrepancies in the X and Y directions and a visible amplitude discrepancy in the Z direction. Its closest point of comparison, test 7 (plotted in Figure 24), which uses a training set based on manual tracking of the marker in the FIPM-bspline reconstructed 4DCBCT, shows some waveform disagreement in the X direction but is of high quality in all other aspects. It seems that a combination of degraded eigenvectors resulting from this training set generation method and the more complicated projection fitting target caused the drop in performance, as runs 4 and 6, where only one of these factors was present, both had results comparable to the other tests in the first group.

The second group covers test runs 9-18, and here the reconstructed 3D marker trajectories differ more widely from each other as well as from the gold standard trajectory. Discounting tests 10 and 15 for the moment as outliers, we see difference means ranging from -0.0274 to -0.864, -0.228 to -1.07, and -1.02 to -1.23 mm for the X, Y, and Z marker trajectory spatial components respectively. Difference standard deviations range from 0.24 to 0.621, 0.707 to 1.94, and 0.384 to 0.833 mm for the X, Y, and Z marker trajectory spatial components respectively. However, the qualitative assessments of the marker plots tell the most interesting story. Looking at Figure 27, Figure 28, Figure 29, and Figure 31 it becomes apparent that the X direction is the most volatile, with noticeable amplitude and offset differences present across all tests. The quantitative results for the X component are comparable or even favorable to the other two components, but the much smaller magnitude of motion in the X direction masks the ambiguity with which the algorithm is able to track the marker along this axis. Overall, the Y component appears to suffer mostly from mild differences in marker trajectory amplitude. The Z component plots display the systematic offset difference that is consistently observed across all test runs 9-18. Disregarding this difference as an artificial misalignment between the reconstruction and validation geometries the marker trajectory reconstruction along this axis appears nearly as good as that observed for tests 1-8. Intuitively these observations make sense; the poorest tracking occurs in the X direction, in which the motion is in the axial plane and has the smallest magnitude. Movement along an axis within the axial plane is resolved in the projection geometry with variable accuracy depending on the gantry angle, with the least sensitivity for motion normal to the detector plane and greatest sensitivity to motion parallel to the detector plane. The most accurate tracking occurs in the Z direction, in which the magnitude of motion is the greatest and is in a direction resolved equally well from all gantry angles.

The data for the 2D marker projection positions is somewhat more uniform, both qualitatively and quantitatively. Once again discounting tests 10 and 15 we see that Figure 37, Figure 38, Figure 39, and Figure 41 all appear similar for both the U and V directions and from Table 6 it is apparent that errors in U are generally bracketed between 2.5 to 3.5 detector pixels while errors in V tend to fall between 1.5 to 2.5 detector pixels. This homogeneity coupled with the variance observed in the 3D data implies that multiple 3D motion trajectories are capable of producing similar projection results. Since our optimization algorithm is driven by results in the projection domain it may indicate degeneracy within our solution search space.

Tests 10 and 15, which were deemed outlier tests, also present interesting cases. Test 10 is a testament to the importance of an effective binning strategy. The proprietary phase based binning method implemented automatically by the RPM system is a notoriously poor solution, and as such the results of the FIPM-bspline registrations to projections binned by this method are expected to be poor. As seen in Figure 21 there is no observable correlation between the average principal coefficient and breathing traces when using this method, producing random eigenvectors and a nonsensical initial parameter guess. In contrast, the average principal coefficient and breathing traces in Figure 22 show clear waveform similarities for the first two, dominant modes. The end results, displayed in Figure 26 and Figure 36, are huge misses in each spatial component. Test 10 display the consequences when the projection binning portion of the overall algorithm is not well executed. Test 15 is treated as an outlier for a different reason, namely because it includes only one mode in the motion model; it is discussed further in Section 4.8.5.

4.8.2 Comparison to literature

Within the literature there are a handful of investigators who have proposed methods to track either full moving anatomy or a moving lung tumor in a non-time averaged manner within the context of an image registration or reconstruction algorithm. Most of these algorithms are in the proof of concept stage and none compare reconstructed 3D motion results for patient data to a gold standard. Regardless, their experiences are still useful in comparison to our own.

One such method uses a motion model capable of rigid transformations to simultaneously reconstruct a reference image and its time dependent deformation. The temporal dependence of the motion model is based on a linear combination of scaled versions of motion surrogate amplitude and velocity traces. The intent of the algorithm is to reconstruct implanted marker or tumor motion. In a simulation study they documented an average marker tracking error of 0.08 mm and an average tumor tracking error of about 1.0 mm.[73]

Another method optimizes the principal coefficients of a PCA based deformation model to warp a reference image such that it matches individual x-ray projections. Though they reconstruct the full deforming anatomy with this method their principal goal is to track a tumor in 3D. In experiments using a numerical phantom they report a 3D tracking error of about 0.8 mm. They report similar error for a physical motion phantom moving in a single direction. They also reconstruct 3D tumor motion for patient data, however for validation they only assess tumor motion in the 2D detector reference frame. For this test they find an average error of about 1-2 mm along each axis of the detector plane when scaled back to the average 3D tumor location.[28]

A final method iteratively registers a prior image to a projection image acquired at a later time using a rigid deformation model. This algorithm is described as a 2D/3D registration approach and its intent is to track tumor motion in real time. In a physical motion phantom study with displacement in a single direction the authors reported a RMS tracking error of 2.6 mm. In a patient data study they did not provide numerical results but noted a visual correlation between the reconstructed 3D tumor path and the 1D motion of the patient diagram.[74]

Although somewhat sparse, these examples display roughly the same trends that we observe, namely successful motion tracking with errors on the order of a millimeter.

4.8.3 Binning strategy

For our investigation of the most effective binning strategy we tested three different algorithms based on breathing trace phase, amplitude, and a hybrid of phase and amplitude. As discussed at the end of Section 4.8.1 the phase based method (in the particular implementation that we considered) was known to be a poor solution and was never a true candidate. Regarding the comparative advantage of the amplitude vs. hybrid method, the results appear fairly inconclusive. Both deliver reconstructions that are quantitatively and qualitatively similar. The hybrid algorithm would appear to have several advantages over the simpler amplitude based method; it considers multiple criteria relevant to successful 4DCBCT registration/reconstruction when choosing bin boundaries, bins are calculated independently so that bin quality is not related to the number of bins generated, and it allows outlier projections to be rejected while including particularly useful projections in multiple bins. It also has some disadvantages in that its results are less predictable and it requires user supplied weighting factors that must be determined empirically.

Regardless, the distinctions appear moot in this case as there is no clear differentiation between results obtained with one method or the other. It is possible that the rigid, fairly regular, motion exhibited by the marker is simply not complicated enough to challenge either algorithm. This should be tested further in data-sets containing true deformable motion within a complete patient anatomy.

Of additional interest are the eigenvalue spectra associated with the eigenvectors generated from each method. The eigenvalue spectra tabulate the training set variance captured by each eigenvector; in other words they quantify the relative importance of each mode. A flat eigenvalue spectrum indicates feature vectors composed entirely of random noise, while noise free feature vectors that track motion coherent across all spatial components will result in a single non-zero eigenvalue. Feature vectors whose components contain relative phase shifts but follow an identical periodic waveform (disregarding amplitude and baseline offset differences) will be fully described by two modes, while for pseudo-periodic waveforms with multiple phase shifts 3 or more modes will be required. The eigenvalue spectra in Figure 6 from the simulated training data described in Chapter 2 display these trends. In our patient data testing feature vectors come from actual deformable registrations, and are therefore expected to contain a mixture of true motion signal and registration noise. Thus we expect the eigenvalue spectra generated by these tests to exhibit a few meaningful modes that fall off sharply in importance and are appended with long tails of non-zero modes dominated by noise. With this knowledge in hand we can use these spectra as rough diagnostic tools to assess the coherence of patient motion and the noisiness of deformable registrations. Figure 46 displays the eigenvalue spectra for feature vector training sets generated using the FIPM-bspline registration algorithm with different projection binning strategies. For reference it also includes a “clean”, or noise free, training set obtained by upscaling a group of displacement vectors sampled directly from the gold standard marker trajectory. All methods show the sharp falloff and long tail that we expect, with the shape of the curves being fairly smooth and similar for the amplitude and hybrid methods, while somewhat different and more jagged for the phase based method. Though the irregularity of the phase based spectrum may be an indicator of its poor quality, the effect does not appear to be drastic enough to be conclusive.

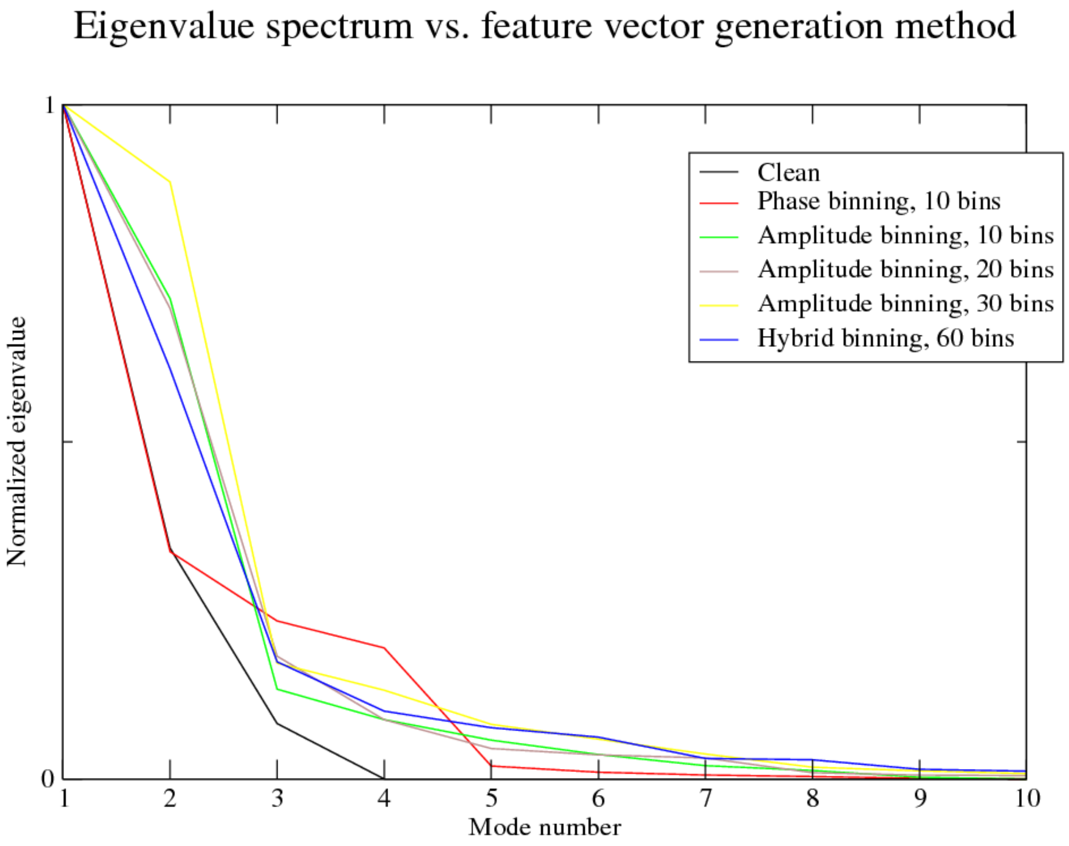


Figure 46: Eigenvalue spectra of feature vectors generated from FIPM-bspline registration of projections binned using different bin numbers and binning strategies.

4.8.4 Number of bins

The conclusion here is the same as that for overall binning strategy. Tests 11, 12, and 13 using 10, 20, and 30 amplitude bins respectively show similar results. In addition, the eigenvalue spectra in Figure 46 look more or less the same for different amplitude bin numbers. Once again, the simplicity of the marker motion may obscure effects that will be relevant for reconstruction of more complicated data. Based on previous work we expect 10 bins to act as a bare minimum, with an optimal number somewhere in the 20-30 range.[34]

4.8.5 Number of modes

Our investigation of the number of modes required for a complete motion model began with one mode and built up to the first five principal modes. The goal of this inquiry was to find the minimum number of modes required for an accurate reconstruction. In general, including a large number of modes will increase the dimensionality of the search space without introducing meaningful new motion data. Test 15, the run using only one mode, displayed visibly degraded marker trajectory results, as evidenced by Figure 30 and Figure 40. This was expected as it is mathematically impossible for one mode to fully capture hysteresis motion in which the spatial motion components are out of phase by anything other than 180 degrees. In an interesting result test 16 incorporating only 2 modes performed approximately evenly with tests 11-13, 17, and 18 which used 3, 4, and 5 modes respectively. This was approximately in line with our expectations based on the literature as well as our own past experiences.[24, 27, 28, 32] However, the simplicity of the motion contained in our segmented data-sets remains a factor to keep in mind when extrapolating these results to more general data.

4.8.6 Motion model temporal stability

Our present formulation of the PCA motion model assumes that the relative phase and amplitude relationships between the DVF motion vector components (which are represented by the eigenvectors) are constant for the duration of the reconstructed time series. It also assumes a fixed relationship in time between a tracked reference feature (the RPM breathing trace) and the moving anatomy to be reconstructed. We investigated the timeframe in which this assumption held in our chosen data-set by splitting it into two smaller sets and generating optimized motion models based solely on either set. We also examined hybrid scenarios in which training data was selected from one set and the optimization fitting target data was selected from the other set. Though for these tests we used limited data to train and optimize our motion model, we reconstructed the marker trajectory for all time points. Observing the reconstruction results for data not included in the motion model generation process gave us a picture of the degree to which the model was still valid for later (or earlier) time periods.

From tests 22 and 24 we can see the effect of reconstructing data from one time group using a motion model trained and optimized on data from a different time group. The general trend is most readily observable in the 2D marker projection data for tests 22 and 24 tabulated in Table 6 and plotted in Figure 42 and Figure 44. In test 22 the motion model was trained and optimized on the first group of time points, while for test 24 it was trained and optimized on the second set of time points. For the first group of time points going from test 22 to test 24 increased the projected marker centroid RMS error from 2.49 to 3.62 and 1.81 to 5.00 detector pixel units in the U and V directions respectively. Likewise, for the second group of time points going from test 24 to test 22 increased the projected marker centroid RMSE from 3.73 to 3.88 and 3.19 to 4.79 detector pixel units in the U and V directions respectively. Thus we see that for this metric, error increased for data reconstructed in a different time period than the data used to construct the motion model. When the two data-sets are one and the same, error is reduced. Even so, looking at Figure 42 and Figure 44 we still see a close correspondence between reconstructed and known marker positions for data not included in the training and optimization of the motion model. Indeed, using the approximate conversion factor of 0.259 mm/det. pixel described in section 4.7 the largest observed RMS error of 5.00 detector pixels translates to an error of only 1.295 mm at isocenter. The picture is murkier when viewed through the lens of the 3D marker trajectory data tabulated in Table 5 and plotted in Figure 32 and Figure 34. For this evaluation method the data in both groups exhibit closer agreement in test 22 than in test 24. It may be that noise and alignment errors in the gold standard data are large enough to wash out the small effect documented in the projected marker centroid results.

From tests 22-25 we also observe that fitting target data is the dominant factor influencing temporal stability of the motion model. This trend is on display in the 2D projected marker positions displayed in Figure 42, Figure 43, Figure 44, and Figure 45. For Figure 42 the training data was taken from the first group of time points, as was the fitting data. The plots, particularly in the V direction, show a superior agreement between the first group of time points as compared to the second. For Figure 43 the training data was taken from the first set of time points as well, but the target data was selected from the second group of time points. Compared to the plots in Figure 42 the agreement for the second section of time points is visibly improved while that for the first group is visibly degraded. Comparing Figure 44 and Figure 45, which are both trained on data from the second time group but fit to projections sampled from different groups, we observe the same effect. The best agreement occurs for the time points from which the fitting data was sampled regardless of the origin of the training data.

4.9 Conclusions

The goal of this chapter was to test our algorithm on real patient data in a variety of different scenarios. We simplified our data-sets to show only an implanted radio-opaque fiducial marker and compared the marker trajectory reconstructed by our FIPM-PCA 4DCBCT algorithm to a gold standard marker trajectory generated by an independent method. We found that our reconstructed data agreed with the validation data to within about a millimeter for all components of motion. We also investigated specific parameters of our algorithm, such as overall projection binning strategy and number of projection bins. These tests were largely inconclusive, with most options performing at a high level. In addition we confirmed two or three modes to be the optimal number for inclusion in the motion model for the tests considered. Finally, we studied the temporal stability of our motion model by using it to predict data on which it was not trained or optimized. In general we found that results were degraded in this scenario, however the effect was not large. These tests were carried out on only a single patient data-set. In particular our conclusions regarding the amplitude and hybrid binning algorithms and the test for the optimal number of modes should be revisited as more data-sets are examined. In the future, multiple patient data-set exhibiting different motion patterns should be considered.

5 Final thoughts

5.1 Summary of work

The purpose of this project was to build and validate a novel 4DCBCT reconstruction algorithm for the purpose of making lung cancer radiotherapy safer and more effective. In Chapter 2 we began by formulating our algorithm mathematically and implementing it in code. Achieving a stable, well organized, code base was one of the chief accomplishments in this development phase. We also devised a scheme for producing numerical 4DCBCT data-sets and tested our algorithm using this data. From these tests we learned several important lessons. For one, we saw only a modest increase in error as we increased the complexity of motion within our simulated data-sets. We also found that optimization of our motion model was frequently a tricky process, with gradient based optimizers proving unreliable. Based on this fact the downhill simplex algorithm became our method of choice. Finally, we compared results using ideal training set generation methods to results using realistic methods based on backprojection reconstructed 4DCBCT and FIPM-bspline 2D/3D registration. The clear take away from this test was the superiority of the FIPM-bspline method. This conclusion allowed us to confidently establish this method as an important piece of our algorithm moving forward.

In Chapter 3 we equipped our algorithm to perform reconstructions for real world data. Specifically, we developed an improved DRR algorithm meant to produce accurate emulations of actual CBCT projections. Having a DRR model that faithfully reproduces real projections is critical to our overall strategy of matching DRRs and projections to optimize our motion model, especially given our preference for a SSD based merit function. Our strategy was to combine an empirical characterization of our CBCT system with knowledge of the various physical effects important to image formation. In Chapter 3 we began by laying a mathematical foundation for our DRR algorithm based on the primary and secondary physical processes that influence projection image formation. We then described our methods for accounting for three secondary effects: scatter, beam hardening, and veiling glare. Next we introduced a procedure for obtaining a CT number to LAC conversion function based on measurements taken with the FBCT and CBCT systems. In addition, we discussed and demonstrated the effects of scatter, beam hardening, and veiling glare on these measurements. After this we tied all the pieces together by listing a step by step recipe for computing a DRR using our method. Finally we tested our algorithm by comparing DRRs to projections for both phantom and patient data. In these tests we compared both the complete version of our algorithm and a simplified version that ignored secondary physics effects to ground truth projection images. The results showed a close correspondence to the actual data for both methods, with a small but meaningful edge going to the full DRR algorithm. A final achievement from this stage was our successful transition to GPU based calculation, which improved the speed of our DRR algorithm by over 30x.

In Chapter 4 we took advantage of patient data-sets containing previously tracked implanted fiducial markers to test our algorithm on real data and compare the results to a ground truth. In addition, we investigated best practices and implemented various improvements to our algorithm with regards to both the motion model and optimization scheme. To do so we established a matrix of tests designed tease out the relationships between various algorithm parameters and our reconstruction results. One area of investigation was the method used to bin projections for performing 2D/3D registrations. The binning procedure is important because it can have an effect on the quality of results generated by the FIPM-bspline algorithm, which will in turn impact the quality of the motion model eigenvectors, and ultimately the success or failure of our 4DCBCT reconstructions. We investigated both breathing amplitude and hybrid amplitude/phase based approaches, and also investigated the effect of specifying different bin numbers for the amplitude method. Results for this test were somewhat inconclusive, with all methods performing at relatively the same level. We also tried including anywhere from one to five modes in our motion model. As expected the results indicated that one mode was insufficient and no benefit was gained beyond three modes. In addition, we implemented improvements to our optimization scheme, foremost of which was an automatic method for generating a good initial parameters guess. Finally we tested the temporal stability of our motion model and found only small degradation with separations of approximately three and a half minutes. Overall results were very promising with close agreement between reconstructed and gold standard marker trajectories resulting in largely submillimeter differences for individual components of motion.

5.2 Future directions

Immediately moving forward our algorithm should be tested on more patient data-sets containing implanted markers for further validation of the tests described in Chapter 4. Following this step, testing of the algorithm on full patient data-sets should be implemented. The chief obstacle to this phase will undoubtedly be designing a method to verify the results at many anatomic locations rather than a small set of fiducial or landmark positions. This is a difficult problem in its own right, and is typically resolved in the literature only through the use of numerical phantoms and simulated data-sets. One potentially valuable study might be to acquire a 4D series of high speed magnetic resonance (MR) images to form a ground truth picture of moving patient anatomy and then to create a simulated raw 4DCBCT data-set by casting projections through the MR images. The simulated raw 4DCBCT data-set could be reconstructed by our method and compared to the ground truth time series.

Further on, the speed of the algorithm must certainly be improved for clinical adoption to become a possibility. The most demanding application, onboard image reconstruction, likely requires that reconstructions be performed in a matter of minutes, while even the simplified tests documented in Chapter 4 required hours for algorithm convergence. Improvements in code efficiency and implementation on more powerful hardware will likely go a long way towards resolving this issue.

Once the algorithm has been suitably polished and verified it should be studied within the greater context of 4D radiotherapy delivery. Its potential for use in 4D pre-treatment planning, post-radiotherapy dose verification, and possibly even real-time radiotherapy adaptation and target tracking will need to be investigated alongside novel algorithms in those areas. Our FIPM-PCA 4DCBCT algorithm essentially provides a rich data platform upon which a host of new applications can be supported. Without proper integration into the overall radiotherapy process its full potential cannot be realized.

6 References

1. P. Keall, "4-Dimensional Computed Tomography Imaging and Treatment Planning," Seminars in Radiation Oncology **14**, 81-90 (2004).

2. A. Harsolia, G. D. Hugo, L. L. Kestin, I. S. Grills and D. Yan, "Dosimetric Advantages of Four-Dimensional Adaptive Image-Guided Radiotherapy for Lung Tumors Using Online Cone-Beam Computed Tomography," International Journal of Radiation Oncology, Biology, and Physics **70**, 582-589 (2008).

3. S. Dieterich, K. Cleary, W. D'Souza, M. Murphy, K. H. Wong and P. Keall, "Locating and targeting moving tumors with radiation beams," Med. Phys. **35**, 5684-5694 (2008).

4. T. Yamamoto, U. Langner, B. W. L. Jr., J. Shen and P. J. Keall, "Retrospective Analysis of Artifacts in Four-Dimensional CT Images of 50 Abdominal and Thoracic Radiotherapy Patients," International Journal of Radiation Oncology*Biology*Physics (2008).

5. L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T.-F. Li, G.-Y. Kim, G. Luxton and A. Koong, "Overview of image-guided radiation therapy," Medical Dosimetry **31**, 91-112 (2006).

6. S. S. Vedam, P. J. Keall, V. R. Kini, H. Mostafavi, H. P. Shukla and R. Mohan, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal," Physics in Medicine and Biology **48**, 45-62 (2003).

7. K. Taguchi, "Temporal resolution and evaluation of candidate algorithms for four-dimensional CT," Med. Phys. **30**, 640-650 (2003).

8. M. J. Fitzpatrick, G. Starkschall, J. A. Antolak, J. Fu, H. Shukla, P. J. Keall, P. Klahr and R. Mohan, "Displacement-based binning of time-dependent computed tomography image data sets," Medical Physics **33**, 235-246 (2006).

9. A. F. Abdelnour and et al., "Phase and amplitude binning for 4D-CT imaging," Phys. Med. Biol. **52**, 3515 (2007).

10. W. Lu, P. J. Parikh, J. P. Hubenschmidt, J. D. Bradley and D. A. Low, "A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT " Med. Phys. **33**, 2964 (2006).

11. U. W. Langner and P. J. Keall, "Prospective displacement and velocity-based cine 4D CT," Med. Phys. **35**, 4501 (2008).

12. J. J. Sonke, L. Zijp, P. Remeijer and M. v. Herk, "Respiratory correlated cone-beam CT," Med. Phys. **32**, 1176-1186 (2005).

13. T. Li, L. Xing, P. Munro, C. McGuinness, M. Chao, Y. Yang, B. Loo and A. Koong, "Four-dimensional cone-beam computed tomography using an on-board imager," Med. Phys. **33**, 3825-3833 (2006).

14. L. Dietrich and et al., "Linac-integrated cone beam CT: first experimental results " Phys. Med. Biol. **51**, 2939-2952 (2006).

15. T. Li, A. Koong and L. Xing, "Enhanced 4D cone-beam CT with inter-phase motion model " Medical Physics **34**, 3688-3695 (2007).

16. Z. Qi and G.-H. Chen, "Extraction of tumor motion trajectories using PICCS-4DCBCT: A validation study," Med. Phys. **38**, 5530 (2011).

17. S. Rit, J. W. H. Wolthaus, M. v. Herk and J.-J. Sonke, "On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion," Med. Phys. **36**, 2283 (2009).

18. S. Leng, J. Tang, J. Zambelli, B. Nett, R. Tolakanahalli and G. H. Chen, "High temporal resolution and streak-free four-dimensional cone-beam computed tomography," Physics in Medicine and Biology **53**, 5653-5673 (2008).

19. F. Bergner, T. Berkus, M. Oelhafen, P. Kunz, T. Pa, R. Grimmer, L. Ritschl and M. Kachelriess, "An investigation of 4D cone-beam CT algorithms for slowly rotating scanners," Med. Phys. **37**, 5044-5053 (2010).

20. E. Schreibmann, G. T. Y. Chen and L. Xing, "Image interpolation in 4D CT using a BSpline deformable registration model," International Journal of Radiation Oncology, Biology, Physics **64**, 1537-1550 (2006).

21. H. Gao, R. Li, Y. Lin and L. Xing, "4D cone beam CT via spatiotemporal tensor framelet," Med Phys **39**, 6943-6946 (2012).

22. J. R. McClelland, S. Hughes, M. Modat, A. Qureshi, S. Ahmad, D. B. Landau, S. Ourselin and D. J. Hawkes, "Inter-fraction variations in respiratory motion models," Phys. Med. Biol. **56**, 251-272 (2011).

23. D. Yang, W. Lu, D. A. Low, J. O. Deasy, A. J. Hope and I. E. Naqa, "4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling," Med. Phys. **35**, 4577-4590 (2008).

24. Q. Zhang, A. Pevsner, A. Hertanto and Y.-C. Hu, "A patient-specific respiratory model of anatomical motion for radiation treatment planning," Med. Phys. **34**, 4772-4781 (2007).

25. R. Zeng, J. A. Fessler and J. M. Balter, "Estimating 3-D Respiratory Motion From Orbiting Views by Tomographic Image Registration," IEEE Transactions on Medical Imaging **26**, 153-163 (2007).

26. A. Docef and M. J. Murphy, "Reconstruction of 4D deformed CT for moving anatomy," International Journal of CARS **3**, 591–598 (2008).

27. R. Li, X. Jia, J. H. Lewis, X. Gu, M. Folkerts, C. Men and S. B. Jiang, "Real-time volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy," Med. Phys. **37**, 2822-2826 (2010).

28. R. Li, J. H. Lewis, X. Jia, X. Gu, M. Folkerts, C. Men, W. Y. Song and S. B. Jiang, "3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy," Med. Phys. **38**, 2783-2794 (2011).

29. D. A. Low, P. J. Parikh, W. Lu, J. F. Dempsey, S. H. Wahab, J. P. Hubenschmidt, M. M. Nystrom, M. Handoko and J. D. Bradley, "Novel breathing motion model for radiotherapy," International Journal of Radiation Oncology, Biology, Physics **63**, 921-929 (2005).

30. R. Li, J. H. Lewis, X. Jia, T. Zhao, W. Liu, S. Wuenschel, J. Lamb, D. Yang, D. A. Low and S. B. Jiang, "On a PCA-based lung motion model," Physics in Medicine and Biology **56**, 6009-6030 (2011).

31. R. S. Brock, A. Docef and M. J. Murphy, "Reconstruction of a cone-beam CT image via forward iterative projection matching," Med. Phys. **37**, 6212-6220 (2010).

32. C. Vaman, D. Staub, J. Williamson and M. Murphy, "A method to map errors in the deformable registration of 4DCT images," Med. Phys. **37**, 5765-5776 (2010).

33. M. Galassi, J. Theiler, J. Davies, B. Gough, R. Priedhorsky, G. Jungman, M. Booth, J. O. Tähtinen, T. Walter, F. Rossi, C. Perassi, S. Jaroszewicz, N. Darnis, J. H. Stover, I. Alxneit, T. Keskitalo, L. Johnson, P. Alken, R. Ulerich, P. Holoborodko and P. Gonnet, "GNU Scientific Library Reference Manual," *Vol. 2011*.

34. C. Vaman, D. Staub, J. Williamson and M. J. Murphy, "A method to map errors in the deformable registration of 4DCT images," Medical Physics **37**, 5765-5776 (2010).

35. A. Docef, M. J. Murphy, P. Keall, J. Siebers and J. Williamson, "Forward CT reconstruction from limited projection data," in *Proceedings of the 19th Conference on Computer-Assisted Radiology and Surgery,* (Elsevier, Berlin, Germany, 2005), pp. 104 – 108.

36. A. Docef and M. J. Murphy, "Reconstruction of a deformed 4D CT for moving anatomy," Int. J. CARS **3**, 591 – 598 (2008).

37. D. Staub, A. Docef, R. S. Brock, C. Vaman and M. J. Murphy, "4D Cone-beam CT reconstruction using a motion model based on principal component analysis," Medical Physics **38**, 6697-6709 (2011).

38. R. S. Brock, A. Docef and M. J. Murphy, "Reconstruction of a cone-beam CT image via forward iterative projection matching," Med. Phys. **37**, 6212-6220 (2010).

39. A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm," Ultrason. Imaging **6**, 81-94 (1984).

40. 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. Med. Imaging **17**, 586-595 (1998).

41. D. Skerl, D. Tomazevic, B. Likar and F. Pernus, "Evaluation of similarity measures for reconstruction-based registration in image-guided radiotherapy and surgery," Int. J. Radiat. Oncol. Biol. Phys. **65**, 943-953 (2006).

42. D. Skerl, B. Likar, J. M. Fitzpatrick and F. Pernus, "Comparative evaluation of similarity measures for the rigid registration of multi-modal head images," Phys. Med. Biol. **52**, 5587-5601 (2007).

43. C. W. Helstrom, *Statistical theory of signal detection*, 2d ed. (Pergamon Press, Oxford, New York, 1968).

44. J. Wu and M. J. Murphy, "Assessing the intrinsic accuracy of 3D/3D rigid image registration results for patient setup in the absence of a ground truth," Med. Phys. **37**, 2501 – 2508 (2010).

45. J. P. W. Pluim, J. B. A. Maintz and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," IEEE Trans. Med. Imaging **22**, 986-1004 (2003).

46. R. L. Siddon, "Fast Calculation of the Exact Radiological Path for a 3-Dimensional CT Array," Med. Phys. **12**, 252-255 (1985).

47. D. B. Russakoff, T. Rohlfing, K. Mori, D. Rueckert, A. Ho, J. R. Adler and C. R. Maurer, "Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration," IEEE Trans. Med. Imaging **24**, 1441-1454 (2005).

48. C. S. Moore, G. P. Liney, A. W. Beavis and J. R. Saunderson, "A method to produce and validate a digitally reconstructed radiograph-based computer simulation for optimisation of chest radiographs acquired with a computed radiography imaging system," Brit. J. Radiol. **84**, 890-902 (2011).

49. W. Birkfellner, R. Seemann, M. Figl, J. Hummel, C. Ede, P. Homolka, X. H. Yang, P. Niederer and H. Bergmann, "Wobbled splatting - a fast perspective volume rendering method for simulation of x-ray images from CT," Phys. Med. Biol. **50**, N73-N84 (2005).

50. P. Lacroute and M. Levoy, "Fast volume rendering using a shear-warp factorization of the viewing transformation," in *Proceedings of the 21st annual conference on Computer graphics and interactive techniques,* (ACM, 1994), pp. 451-458

51. T. Totsuka and M. Levoy, "Frequency domain volume rendering," in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques,* (ACM, Anaheim, CA, 1993), pp. 271-278.

52. X. L. Li, J. Yang and Y. M. Zhu, "Digitally reconstructed radiograph generation by an adaptive Monte Carlo method," Phys. Med. Biol. **51**, 2745-2752 (2006).

53. J. Spoerk, H. Bergmann, F. Wanschitz, S. Dong and W. Birkfellner, "Fast DRR splat rendering using common consumer graphics hardware," Med. Phys. **34**, 4302-4308 (2007).

54. S. Mori, M. Kobayashi, M. Kumagai and S. Minohara, "Development of a GPU-based multithreaded software application to calculate digitally reconstructed radiographs for radiotherapy," Radiol. Phys. Technol. **2**, 40-45 (2009).

55. C. L. Yang, B. D. Yang and J. L. Wang, "Fast and High Precision Digitally Reconstructed Radiograph Generation on Graphic Processing Unit," Biomed. Eng.: App. Bas. C. **24**, 9-15 (2012).

56. H. Hong, K. Kim and S. Park, "Fast 2D-3D point-based registration using GPU-based preprocessing for image-guided surgery," Lect. Notes Comput. Sc. **4225**, 218-226 (2006).

57. G. J. Tornai, G. Cserey and I. Pappas, "Fast DRR generation for 2D to 3D registration on GPUs," Med. Phys. **39**, 4795-4799 (2012).

58. J. T. Bushberg, J. A. Seibert, E. M. L. Jr. and J. M. Boone, *The essential physics of medical imaging*, 2nd ed. (Lippincott Williams & Wilkins, Philadelphia, 2002).

59. E. P. Ruhrnschopf and K. Klingenbeck, "A general framework and review of scatter correction methods in x-ray cone-beam computerized tomography. Part 1: Scatter compensation approaches," Med. Phys. **38**, 4296-4311 (2011).

60. D. Lazos, D. Pokhrel, Z. Su, J. Lu and J. F. Williamson, "Experimental validation of a Monte Carlo-based kV x-ray projection model for the Varian linac-mounted cone-beam CT imaging system," in *Medical Imaging 2008: Physics of Medical Imaging. Proceedings of SPIE*, *Vol. 6913,* edited by J. Hsieh and E. Samei (2008), pp. 69133Q.

61. O. Nalcioglu and R. Y. Lou, "Post-reconstruction Method for Beam Hardening in Computerized Tomography," Phys. Med. Biol. **24**, 330-340 (1979).

62. P. M. Joseph and R. D. Spital, "A method for correcting bone induced artifacts in computed tomography scanners," J. Comput. Assist. Tomogr. **2**, 100-108 (1978).

63. C. H. Yan, R. T. Whalen, G. S. Beaupre, S. Y. Yen and S. Napel, "Reconstruction algorithm for polychromatic CT imaging: Application to beam hardening correction," IEEE Trans. Med. Imaging **19**, 1-11 (2000).

64. A.-K. Carton, R. Acciavatti, J. Kuo and A. D. A. Maidment, "The effect of scatter and glare on image quality in contrast-enhanced breast imaging using an a-Si/CsI(TI) full-field flat panel detector," Med. Phys. **36**, 920-928 (2009).

65. R. B. Benitez, R. Ning and D. Conover, "Cone beam CT image quality measurements: PSF de-convolution," in *Medical Imaging 2007: Physics of Medical Imaging*, *Vol. 6510,* (San Diego, CA, 2007), pp. U1925-U1936.

66. G. Poludniowski, P. M. Evans, A. Kavanagh and S. Webb, "Removal and effects of scatter-glare in cone-beam CT with an amorphous-silicon flat-panel detector," Phys. Med. Biol. **56**, 1837-1851 (2011).

67. D. Lazos and J. F. Williamson, "Monte Carlo evaluation of scatter mitigation strategies in cone-beam CT," Med. Phys. **37**, 5456-5470 (2010).

68. N. Milickovic, D. Baltast, S. Giannouli, M. Lahanas and N. Zamboglou, "CT imaging based digitally reconstructed radiographs and their application in brachytherapy," Phys. Med. Biol. **45**, 2787-2800 (2000).

69. W. Cai and G. Sakas, "DRR volume rendering using splatting in shear-warp context," in *Nuclear Science Symposium Conference Record, 2000 IEEE*, *Vol. 3,* (2000), pp. 19/12-19/17.

70. L. B. Lucy, "An iterative technique for the rectification of observed distributions," Astron. J. **79**, 745–754 (1974).

71. P. R. Poulsen, B. Cho and P. J. Keall, "A Method to Estimate Mean Position, Motion Magnitude, Motion Correlation, and Trajectory of a Tumor from Cone-Beam Ct Projections for Image-Guided Radiotherapy," Int J Radiat Oncol **72**, 1587-1596 (2008).

72. P. R. Poulsen, W. Fledelius, P. J. Keall, E. Weiss, J. Lu, E. Brackbill and G. D. Hugo, "A method for robust segmentation of arbitrarily shaped radiopaque structures in cone-beam CT projections," Medical Physics **38**, 2151-2156 (2011).

73. J. Martin, J. McClelland, C. Thomas, K. Wildermuth, D. Landau, S. Ourselin and D. Hawkes, presented at the Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012 IEEE Workshop on, 2012 (unpublished).

74. C. Gendrin, H. Furtado, C. Weber, C. Bloch, M. Figl, S. A. Pawiro, H. Bergmann, M. Stock, G. Fichtinger, D. Georg and W. Birkfellner, "Monitoring tumor motion by real time 2D/3D registration during radiotherapy," Radiother Oncol **102**, 274-280 (2012).

Vita

David Staub was born in Philadelphia, PA on March 20, 1986. He attended Duke University from 2004-2008, receiving a B.S. in physics. He then attended Virginia Commonwealth University from 2008-2013 where he received his Ph.D. in medical physics.