MOTION ESTIMATION IN COMPUTED TOMOGRAPHY

by

Caleb R. Rottman

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Bioengineering

The University of Utah

December 2016

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Caleb R. Rottman**

has been approved by the following supervisory committee members:

Sarang Joshi	, Chair	**August 11, 2016** Date Approved
Ross T. Whitaker	, Member	**August 11, 2016** Date Approved
Arvidas B. Cheryauka	, Member	**August 11, 2016** Date Approved
Amit Sawant	, Member	**August 11, 2016** Date Approved
Robert S. MacLeod	, Member	**August 11, 2016** Date Approved

and by **Patrick A. Tresco** , Chair/Dean of

the Department/College/School of **Bioengineering**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

X-ray computed tomography (CT) is a widely popular medical imaging technique that allows for viewing of in vivo anatomy and physiology. In order to produce high-quality images and provide reliable treatment, CT imaging requires the precise knowledge of the motion of the imaging system as well as the motion of the patient undergoing imaging.

Cone-beam CT is a specific CT modality in which the imaging system has a point X-ray source that emits diverging X-ray beams toward a flat detector. A mobile fluoroscopic C-arm has these properties, but since mobile C-arms are designed to be relatively inexpensive and mobile, the geometric parameters that describe the motion of these systems are imprecise. Applying standard CT reconstruction methods with data from these devices yields poor image quality, so three-dimensional (3D) reconstruction on mobile C-arms is rare. Patient motion in CT imaging is also clinically relevant. In thoracic CT, knowledge of the patient's breathing motion is used to inform dose calculations and treatment decisions in radiation therapy of lung cancer.

In this dissertation, I provide methods for estimating motion in both these cases. I introduce a CT reconstruction framework for mobile C-arms that is robust to imprecise geometries. In the proposed method, I jointly estimate the 3D image of attenuation coefficients and the geometry of the C-arm system. For thoracic CT imaging, I introduce a method for estimating breathing motion that models both the physics of CT acquisition and the physiology of respiration. This method models CT images as densities to account for the changing image intensity that corresponds to changing lung density. This method also incorporates a priori knowledge of tissue compressibility. In particular, it allows volume change inside the lungs while restricting volume change in the soft tissue and bones, yielding a highly accurate estimate of respiratory motion while remaining computationally efficient.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$\mathbb{R}$ ..... real numbers

I, I_i ..... image in $\mathbb{R}^3$

Ω ..... image domain, subset of $\mathbb{R}^3$

$\mathcal{L}^2(\mathbb{R}^3)$ ..... Hilbert space of square integrable functions on $\mathbb{R}^3$

$\|\cdot\|_{\mathcal{L}^2(\mathbb{R}^3)}$ ..... $\mathcal{L}^2$ norm

$\langle\cdot,\cdot\rangle_{\mathcal{L}^2(\mathbb{R}^3)}$ ..... $\mathcal{L}^2$ inner product

$P\{\cdot;\cdot\}$ ..... projection operator

$P^\dagger\{\cdot;\cdot\}$ ..... backprojection operator

$\dagger$ ..... adjoint operator

Δ, ∇, $\nabla\cdot$ ..... Laplacian, gradient, divergence

dx ..... volume element of $\mathbb{R}^3$

φ ..... diffeomorphism

$\mathrm{Diff}(\Omega)$ ..... diffeomorphism group on Ω

$\mathrm{Dens}(\Omega)$ ..... space of densities on Ω

$D\varphi$ ..... 3×3 Jacobian matrix of φ

$|D\varphi|$ ..... Jacobian determinant of φ

μ ..... linear attenuation coefficient

LIST OF ABBREVIATIONS

CT ..... computed tomography
LAC ..... linear attenuation coefficient
FDK ..... Feldkamp, Davis, and Kress reconstruction
EM ..... expectation maximization
OSEM ..... ordered subset expectation maximization
MLE ..... maximum likelihood estimate
MAP ..... maximum a posteriori
NCC ..... normalized cross-correlation
TV ..... total variation
ART ..... algebraic reconstruction technique
SART ..... simultaneous algebraic reconstruction technique
SIRT ..... simultaneous iterative reconstruction technique
LDDMM ..... large deformation diffeomorphic metric mapping
ANTs ..... Advanced Normalization Tools
GPU ..... graphics processing unit
RCCT ..... respiratory correlated computed tomography
GTV ..... gross tumor volume
PTV ..... planning target volume

CHAPTER 1

INTRODUCTION

Advancements in medical imaging over the past several decades have provided clinicians with noninvasive methods to view in vivo anatomy, make diagnoses, and inform treatment strategies. In traditional X-ray radiography, images are created as X-rays partially penetrate anatomy before being recorded on a detector. Image contrast at the detector is created as different anatomies attenuate X-rays at different rates: bones highly attenuate the X-ray beam, soft tissue moderately attenuates the beam, and air has negligible attenuation. This localized property of X-ray attenuation is known as the *linear attenuation coefficient*.

A more recent extension of X-ray radiography is X-ray computed tomography (CT). In X-ray CT, a scanner acquires multiple X-ray projections from different orientations surrounding the patient. These images are then combined to mathematically solve for the 3D volume of linear attenuation coefficients. This resulting 3D image of patient anatomy yields a much more natural and informative way of viewing internal anatomy than traditional two-dimensional (2D) X-ray radiographs, where different anatomical structures are in superposition. The process of solving for this underlying 3D volume given multiple X-ray projections is called CT reconstruction.

One of the greatest challenges in X-ray CT is dealing with uncertain motion. Uncertain motion in the acquisition system leads to miscalibration artifacts [1, 2, 3], diminishing the quality and diagnostic value of the reconstructed images. Uncertain motion after reconstruction (for example, between consecutive phases of the breathing cycle in thoracic CT) can lead to errors in radiation therapy of lung cancer, where a moving tumor is targeted with ionizing radiation. In this dissertation, I develop methods for estimating the uncertain motion in each of these cases, thereby enhancing the utility of X-ray CT.

1.1 X-ray Computed Tomography

The basic problem statement of X-ray computed tomography is as follows: given projection data (X-ray images with corresponding pose information), estimate the 3D volume of X-ray linear attenuation coefficients. CT reconstruction is the process of solving for this 3D volume.

As an X-ray photon passes through a material, it may interact with individual atoms via two different mechanisms: the photoelectric effect and Compton scattering. In the photoelectric effect, an X-ray photon is absorbed by an atom, causing the ejection of a photoelectron. In Compton scattering, a photon deflects at an angle after interacting with a valence electron. Both these effects contribute to the attenuation of an X-ray beam. The total attenuation of a narrow X-ray beam through a material is theoretically described by the Beer-Lambert law. If an X-ray beam with intensity i_0 is sent through a material, the attenuated X-ray intensity at the detector i_d is

$$i_d = i_0 \exp\left(- \int_l \mu(x') dx' \right), \tag{1.1}$$

where $\mu(x)$ is the linear attenuation coefficient at a location along the path, and this line integral is from the X-ray source to the detector element. Each of these line integrals in space is described by the geometric parameters (position and orientation of imaging components) of the imaging system at a point in time. The "forward problem" is easily modeled by this equation: given an image of attenuation coefficients, solve for the projection data. The "inverse problem" associated with CT imaging is more challenging: given projection data, solve for the image of attenuation coefficients.

Current CT reconstruction methods make several assumptions in their formulations. The first assumption is that the geometric information accompanying each X-ray projection is sufficiently accurate. The second assumption is that the 3D volume is sufficiently static throughout the scan.

The first assumption is reasonable for large, dedicated CT scanners. However, this assumption does not hold for mobile C-arms due to their variable and uncalibrated geometry. The second assumption is reasonable for imaging of the head and extremities as they can be forced to remain immobile, but in thoracic imaging, this assumption is inaccurate due to patient breathing. This dissertation develops methods for CT imaging that address

these two assumptions by accurately estimating the true motion of the CT scanner and the patient.

1.2 Mobile C-arm Reconstruction

A mobile fluoroscopic C-arm is a widely used imaging device that creates 2D X-ray images in real-time. Fluoroscopic C-arms are used in many procedures, including orthopedic applications (guiding of screws and plates, arthrography), cardiovascular applications (cardiac catheterization, stent placement, placement of intravenous catheters, angiography, bolus chasing), and gastroenterology procedures (barium X-rays, intravenous pyelograms). Their current widespread adoption is due to their good 2D image quality, relatively low cost, and mobility.

A fluoroscopic C-arm has an X-ray point source at one end of the "C" and a flat detector at the other end (see **Figure 1.1**). This C-arm gantry moves to acquire X-ray images from various positions and orientations. Although a mobile C-arm is designed primarily to be a 2D imaging device, it has the same basic geometry as a fixed-room cone-beam CT scanner. Because of these similarities, implementing 3D reconstruction appears feasible. The impact of implementing 3D reconstruction on a mobile C-arm is twofold: it improves the utility of mobile C-arms and decreases the cost of 3D imaging. Currently, pre- and postoperative

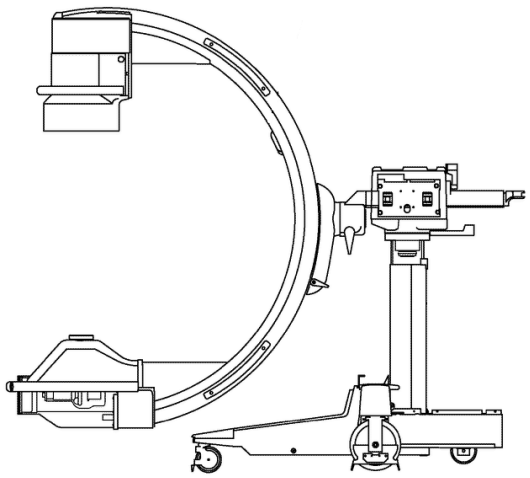


Figure 1.1. GE Mobile C-arm with the X-ray source (bottom of "C") and detector (top of "C").

CT scans are often done on a dedicated CT scanner. Implementing 3D imaging on a mobile C-arm will allow the 3D imaging and surgery to occur in the same location, thereby saving time. Also, since the price of a mobile C-arm is significantly less than a dedicated CT scanner, implementing 3D imaging on a mobile C-arm would decrease the cost of 3D imaging and thereby increase the access to 3D imaging, particularly in developing countries.

Since mobile C-arms are inherently designed for 2D imaging, they have several characteristics that make implementing 3D imaging especially challenging. Most mobile C-arms have limited angular range, nonisocentric orbits, and weak or no motorization. The largest barrier to 3D imaging in mobile C-arms is their variable and uncalibrated geometry. Dedicated CT scanners have precise internal motors and rigid components, which make the given geometric parameters at each projection quite accurate. Even with such rigidity, these scanners must be properly calibrated, as the accuracy of the given parameters degrades over time. These machines undergo routine maintenance to recalibrate the geometry, typically requiring specially designed calibration phantoms [4, 5, 6, 7].

Unlike fixed-room CT scanners, the geometry associated with mobile C-arms changes substantially from scan to scan, and only a small portion of this motion is repeatable. Furthermore, some mobile C-arms have no motorization, which requires the clinician to manually perform an orbital scan. This motion is clearly not repeatable from scan to scan, but even machines with small internal motors do not produce repeatable results. This uncertainty associated with the geometric parameters is significant, so a naive reconstruction using the given parameters produces a very poor reconstruction image.

CT reconstruction currently exists on many fixed-room C-arms (including the Siemens Artis Zeego and the GE Innova CT), but exists on only a select few mobile C-arms: the Siemens Siremobile Iso-C3D, the Philips BV Pulsera, and the Ziehm Vision RFD 3D. All these systems were built with 3D imaging as a primary feature: they all have isocentric orbits and significantly more physical stability than the majority of mobile C-arms. A reconstruction method that is robust to geometric uncertainties would have widespread adoption in mobile C-arms, as it would limit the need to expensively improve the hardware of the C-arm and could even be used to retrofit existing mobile C-arms. Additionally, this method could be used in fixed-room cone-beam scanners to eliminate their need for

routine calibration.

In this dissertation, I propose a method for CT reconstruction that is robust to geometric uncertainty. This method does not require the extensive hardware modifications that have been implemented in current mobile C-arms designed for 3D imaging. The proposed method can be implemented efficiently using graphics processing units (GPUs); therefore, the only increase to the cost is in computation. Furthermore, this technology can be used somewhere in between these two extremes: it can be used to improve upon imprecise sensor data combined with moderate hardware improvements to create an improved image.

1.3 Thoracic Imaging

According to the Centers for Disease Control and Prevention, lung cancer is the third most common form of cancer, yet is by far the leading cause of cancer death, accounting for approximately 27 percent of all cancer deaths in the United States [8]. Thoracic CT imaging is a valuable noninvasive technique for the diagnosis and treatment planning of lung cancer. Accurately assessing the local and global lung biomechanics from the image data is very important, and the registration of lung images at different phases of the breathing cycle plays a vital role in accurately describing respiratory motion.

The accurate estimation of respiratory motion has widespread applications, in particular, radiation therapy of lung cancer. In this application, the estimation of organ deformations during treatment impacts dose calculation and treatment decisions [9, 10, 11, 12]. The current state-of-the-art radiation treatment planning involves the acquisition of a series of respiratory correlated CT (RCCT) images to build four-dimensional (4D) (three spatial and one temporal) treatment planning datasets. Having imprecise knowledge of the tumor location leads to underradiation of the tumor while depositing damaging radiation in the surrounding healthy tissue.

The estimation of breathing motion is a challenging problem as the deformations are highly nonlinear. Inflexible motion models, such as rigid and affine transformations, fail to accurately represent complex biological deformations. Overly flexible models, on the other hand, overfit noisy images with nonsmooth and physiologically implausible deformations. The current state-of-the-art practice is to use the large deformation diffeomorphic metric mapping (LDDMM) framework, which uses a viscous fluid-flow model of these deforma-

tions [13, 14, 15, 16]. However, a simple fluid-flow model of lung motion is inadequate for the simple reason that the human body does not deform as a homogeneous material. During respiration, the lungs expand and contract as gas exchanges with the environment, while the rest of the body undergoes an incompressible deformation. The tissue surrounding the lungs consists of bone, fat, and soft tissue, all of which are incompressible tissues. Therefore, we expect the deformations in lung tissue to be highly compressible while the rest of the body has incompressible deformations. These physiological attributes are not modeled in LDDMM.

LDDMM also does not model the X-ray physics in CT imaging. In standard LDDMM, the action of a deformation on an image is the L^2 action, that is, it acts via function composition. In CT imaging, the image intensities correspond to the linear attenuation coefficient, or the ability of a material to absorb and scatter X-rays. A material's linear attenuation coefficient is directly related to its mass density. Therefore, instead of viewing this estimation problem as an image matching problem, it is best viewed as a density matching problem. These two improvements to deformable registration are significant contributions to motion estimation in thoracic CT imaging.

1.4 Main Contributions

The main contributions of this dissertation for estimating motion in CT imaging are organized into the following three chapters.

1.4.1 Chapter 2

A method for 3D cone-beam CT reconstruction for mobile C-arms is introduced. Current methods for cone-beam CT reconstructions assume a stable and well-known geometry of the imaging system. The hardware of mobile C-arms is considerably less geometrically robust than fixed-room C-arms. Therefore, the introduced method includes the estimation of the motion of the C-arm as well as the estimation of the underlying 3D image. This joint estimation of the image and the geometry is done in an alternating scheme, where the image is estimated in the expectation maximization (EM) framework and the geometry is estimated using a gradient-based energy minimization scheme to improve the correspondence between the acquisition data and estimated 3D image. This work is an expansion of

the work done in Rottman et al. [17] and Rottman et al. [18].

1.4.2 Chapter 3

This chapter provides analysis on the mass-preserving nature of thoracic CT images. Although image intensities in CT should be theoretically conserved, I show that in real thoracic CT datasets, the assumption of conservation of mass is not valid. However, I provide a method to transform these CT images into objects that are indeed mass preserving, and show the approach is applicable across multiple patient datasets acquired on the same CT scanner. This work is an expansion of the work done in Rottman et al. [19].

1.4.3 Chapter 4

A new method for thoracic CT image registration is introduced. The proposed diffeomorphic model acts on densities, thereby ensuring that the deformations preserve mass. These deformations are regularized by a locally-varying penalty on volume change. In incompressible areas of the body (including surrounding soft tissue and bones), the volume change is restricted. In much more compressible areas of the image (such as the lungs), the method allows considerable compression and expansion. The resulting algorithm is computationally efficient and accurately models respiratory motion. This work is an expansion of the work done in Rottman et al. [19] and Rottman et al. [20].

Finally, in Chapter 5, I discuss these contributions as well as the possibility for future work in these areas.

1.5 References

[1] N. T. Vo, M. Drakopoulos, R. C. Atwood, and C. Reinhard, "Reliable method for calculating the center of rotation in parallel-beam tomography," *Optics Express*, vol. 22, no. 16, pp. 19078–19086, 2014.

[2] C. Bodensteiner, C. Darolti, H. Schumacher, L. Matthäus, and A. Schweikard, "Motion and positional error correction for cone beam 3D-reconstruction with mobile C-arms," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, vol. 10, no. Pt 1, pp. 177–185, 2007.

[3] V. Patel, R. N. Chityala, K. R. Hoffmann, C. N. Ionita, D. R. Bednarek, and S. Rudin, "Self-calibration of a cone-beam micro-CT system," *Medical Physics*, vol. 36, no. 1, pp. 48–58, 2009.

[4] F. Noo, R. Clackdoyle, C. Mennessier, T. A. White, and T. J. Roney, "Analytic method

based on identification of ellipse parameters for scanner calibration in cone-beam tomography," *Physics in Medicine and Biology*, vol. 45, no. 11, pp. 3489–3508, 2000.

[5] F. Zhang, J. Du, H. Jiang, L. Li, M. Guan, and B. Yan, "Iterative geometric calibration in circular cone-beam computed tomography," *Optik – International Journal for Light and Electron Optics*, vol. 125, no. 11, pp. 2509–2514, 2014.

[6] D. Panetta, N. Belcari, A. Del Guerra, and S. Moehrs, "An optimization-based method for geometrical calibration in cone-beam CT without dedicated phantoms," *Physics in Medicine and Biology*, vol. 53, no. 14, pp. 3841–3861, 2008.

[7] N. K. Strobel, B. Heigl, T. M. Brunner, O. Schütz, M. M. Mitschke, K. Wiesent, and T. Mertelmeier, "Improving 3D image quality of X-ray C-arm imaging systems by using properly designed pose determination systems for calibrating the projection geometry," in *Proc. SPIE* (M. J. Yaffe and L. E. Antonuk, eds.), vol. 5030, pp. 943–954, 2003.

[8] CDC, "Basic information about lung cancer." `http://www.cdc.gov/cancer/lung/basic_info/index.htm`. Accessed: 2016-03-06.

[9] P. J. Keall, S. Joshi, S. S. Vedam, J. V. Siebers, V. R. Kini, and R. Mohan, "Four-dimensional radiotherapy planning for DMLC-based respiratory motion tracking," *Medical Physics*, vol. 32, no. 4, pp. 942–951, 2005.

[10] S. E. Geneser, J. D. Hinkle, R. M. Kirby, B. Wang, B. Salter, and S. Joshi, "Quantifying variability in radiation dose due to respiratory-induced tumor motion," *Medical Image Analysis*, vol. 15, no. 4, pp. 640–649, 2011.

[11] A. Sawant, P. Keall, K. B. Pauly, M. Alley, S. Vasanawala, B. W. Loo, J. Hinkle, and S. Joshi, "Investigating the feasibility of rapid MRI for image-guided motion management in lung cancer radiotherapy," *BioMed Research International*, vol. 2014, 2014.

[12] Y. Suh, W. Murray, and P. J. Keall, "IMRT treatment planning on 4D geometries for the era of dynamic MLC tracking," *Technology in Cancer Research and Treatment*, vol. 13, no. 6, pp. 505–515, 2014.

[13] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinetics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[14] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *International Journal of Computer Vision*, vol. 61, no. 2, pp. 139–157, 2005.

[15] S. Joshi, P. Lorenzen, G. Gerig, and E. Bullitt, "Structural and radiometric asymmetry in brain images," *Medical Image Analysis*, vol. 7, no. 2, pp. 155–170, 2003.

[16] M. I. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *International Journal of Computer Vision*, vol. 41, no. 1-2, pp. 61–84, 2001.

[17] C. Rottman, J. Hinkle, A. Cheryauka, R. Whitaker, and S. Joshi, "Joint cone-beam reconstruction and geometry estimation for mobile C-arms," in *Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 697–700, 2015.

[18] C. Rottman, L. McBride, A. Cheryauka, R. Whitaker, and S. Joshi, "Mobile C-arm 3D reconstruction in the presence of uncertain geometry," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part II* (N. Navab, J. Hornegger, M. W. Wells, and F. A. Frangi, eds.), pp. 692–699, Springer International Publishing, 2015.

[19] C. Rottman, M. Bauer, K. Modin, and S. Joshi, "Weighted diffeomorphic density matching with applications to thoracic image registration," in *5th MICCAI Workshop on Mathematical Foundations of Computational Anatomy (MFCA 2015)*, pp. 1–12, 2015.

[20] C. Rottman, B. Larson, P. Sabouri, A. Sawant, and S. Joshi, "Diffeomorphic density registration in thoracic computed tomography," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings* (S. Ourselin, A. Ercil, M. W. Wells, M. R. Sabuncu, L. Jowkowicz, and G. Unal, eds.), Springer International Publishing, 2016.

CHAPTER 2

JOINT CONE-BEAM RECONSTRUCTION AND GEOMETRY ESTIMATION FOR MOBILE C-ARMS

This chapter deals with the problem of X-ray CT reconstruction in the presence of uncertain system motion. In this case, the object being imaged is assumed to be stationary and the system motion is estimated during reconstruction. Although it is developed and tested with mobile C-arms, it is applicable to any cone-beam system that has uncertain motion.

2.1 Introduction

Mobile fluoroscopic C-arms are popular imaging systems that produce 2D X-ray images in real-time. However, 3D cone-beam tomographic reconstruction is mostly limited to larger, more expensive systems, such as fixed-room C-arms, isocentric C-arms, and O-arms. Since most mobile C-arms are inherently designed 2D imaging, the geometry of the acquisition scan is irregular and uncertain, so traditional cone-beam reconstruction techniques such as Feldkamp, Davis, and Kress (FDK) [1] do not work. In addition to geometric irregularities, mobile C-arms exhibit other non-ideal characteristics for 3D reconstruction, including nonisocentricity and limited angle coverage. Large isocentric C-arms and fixed-room C-arms have been developed that have much more consistent geometry, but smaller mobile C-arms traditionally have no need for such precision and are often moved by hand or by an imprecise internal motor. An acquisition sweep puts strain on the C-arm gantry, further decreasing the certainty of the geometry of an acquisition path. Image reconstruction with incorrect parameters yields poor image quality with miscalibration artifacts, so 3D computed tomography (CT) on mobile C-arms is currently limited.

Methods for geometric calibration of cone-beam CT systems can be divided into two

categories: *offline calibration* and *autocalibration*. Offline calibration methods correct for inconsistencies in the system geometry before the scanner is used in a clinical application. These methods are often necessary for fixed-room C-arms. Even though these systems have much more geometric certainty, these scanners typically need to undergo periodic calibration due to changes in geometry over the life of the scanner. This offline calibration is usually done using specifically designed calibration phantoms [2, 3, 4, 5]. These methods use precisely manufactured phantoms with high-attenuation markers and then analytically fit ellipses to the projection data. In these methods, the ground truth 3D image (calibration phantom) is well known, and the scan geometry is estimated using this prior knowledge. These methods are not viable for mobile C-arms, as much of the scanner geometry in mobile C-arms is not consistent between consecutive scans.

Autocalibration methods, on the other hand, estimate the true geometry on-the-fly using image-based updates. Some proposed methods use only the sinogram data of circular scans for autocalibration to estimate a subset of the geometric parameters [6, 7]. Kyriakou et al. use only the reconstructed image and seek to minimize its entropy with respect to the geometric parameters [8].

Finally, some methods use both the reconstructed image and the projection data and perform a 2D-3D registration between them. The proposed method fits into this category. Muders et al. use FDK reconstruction for a helical cone-beam system where they optimize over a 3D translation and 1D rotation by sampling along each of the four dimensions independently [9]. Bodensteiner et al. interleave ART image updates and parameter updates using a Hooke-Jeaves search to maximize mutual information. Similarly, Wein et al. use SIRT image updates with an amoeba direct search method to minimize the L^1 error [10]. In contrast, the proposed method implements the gradient with respect to all the geometric parameters, and then uses an expectation maximization (EM) image reconstruction framework that allows for a joint estimation of the geometry and the 3D image. For the geometry update, I implement the gradient of the normalized cross-correlation between the current estimate of the 3D image and the 2D projection data, which allows for any number of gradient-based optimization schemes.

In this chapter, I propose a method for 3D reconstruction in the presence of uncertain geometry. Instead of trying to get more precise geometry by expensively improving the

hardware, I simultaneously estimate the geometry and the 3D image. I have implemented the proposed method on the GPU. The results show that the geometric parameters are accurately estimated, and the resulting reconstruction images are free of miscalibration artifacts. The proposed method also does not require an isocentric trajectory, which does not exist in many mobile C-arms.

This method can be used to retrofit existing mobile C-arms with 3D imaging. In this case, the only added cost is in computation. Furthermore, the proposed method can be used with existing sensor data or with an a priori estimate of the acquisition scan, thereby improving the given geometry estimates.

2.2 Cone-Beam Reconstruction

2.2.1 X-ray Attenuation

The total attenuation of a narrow X-ray beam through a material is described by the Beer-Lambert law. Given an X-ray beam with intensity i_0 passing through a material, the attenuated X-ray intensity i_d recorded at a detector element is

$$i_d = i_0 \exp\left(-\int_l \mu(x')dx'\right), \tag{2.1}$$

where $\mu(x)$ is the linear attenuation coefficient at a location on the path, and this line integral is from the X-ray source to the detector element. The main challenge associated with CT reconstruction is solving for $\mu(x)$ given many detector recordings. By taking the natural log of both sides, we get an expression for the line integral:

$$\ln(i_0) - \ln(i_d) = \int_l \mu(x')dx'. \tag{2.2}$$

With this log transform, the X-ray system can be written as a linear system. Essential to describing this system is the knowledge of the precise locations of each of these line integrals. These locations at a point in time are described by the imaging system's geometric parameters. In order to define these geometric parameters in cone-beam systems, I will first define the applicable coordinate systems and the rigid transformations among them.

2.2.2 Cone-Beam Geometry

A point in space is described by its *world coordinates*, $\boldsymbol{x} = (x, y, z) \in \mathbb{R}^3$. This coordinate system is fixed and the reconstructed 3D image $I(\boldsymbol{x})$ of linear attenuation coefficients is defined in world coordinates.

The imaging system has a point X-ray source and a 2D X-ray detector that is assumed to be flat. A point in space is also described by its *camera coordinates*, $\boldsymbol{x}' = (x', y', z') \in \mathbb{R}^3$. The X-ray source is located at point $\boldsymbol{x}' = (0, 0, 0)$ in this coordinate system. The distance from the source to the detector is $l \in \mathbb{R}^+$, called the *source-to-image-detector distance* or *SID*. The unique point on the detector plane that is closest to the source is called the *piercing point*. The positive z' axis passes through the piercing point and is therefore orthogonal to the 2D detector plane. In camera coordinates, the piercing point is located at $(x', y', z') = (0, 0, l)$.

A point on the detector plane is described by its *projection coordinates*, $\boldsymbol{u} = (u, v) \in \mathbb{R}^2$. The piercing point is denoted as the point (u_0, v_0). The u and v axes have the same directions and scales as the x' and y' axes, respectively. Therefore, the point (u, v) in projection coordinates is the point $(u - u_0, v - v_0, l)$ in camera coordinates. The piercing point and SID are called the *intrinsic parameters*, which define the relationship between projection coordinates and camera coordinates. These parameters describe the relative alignment of the X-ray source and detector. Any nonrigid deformation of the C-arm gantry is reflected by a change in intrinsic parameters.

The transformation between camera coordinates and world coordinates is a rigid transformation in $\mathbb{R}^3$. I represent this transformation by a rotation matrix $R \in \mathrm{O}(3)$ combined with a translation $\boldsymbol{T} \in \mathbb{R}^3$, where $\mathrm{O}(3)$ is the orthogonal group of 3D rotation matrices. For any two matrices $R_0, R_1 \in \mathrm{O}(3)$, $|R_i| \in \{1, -1\}$, $R_i^{-1} = R_i^{\mathrm{T}}$, and $R_0 R_1 \in \mathrm{O}(3)$, where the group action is matrix multiplication. I define this rigid transformation between a point $\boldsymbol{p}'$ in camera coordinates and the same point $\boldsymbol{p}$ in world coordinates as

$$\boldsymbol{p} = R(\boldsymbol{p}' + \boldsymbol{T}). \tag{2.3}$$

This rotation can either be orientation preserving ($|R| = 1$) or orientation reversing ($|R| = -1$). This choice is based only on convention and will be consistent throughout the scan. The 3D rotation and 3D translation together are called the *extrinsic parameters*. The 3D image $I(\boldsymbol{x})$ defined in world coordinates can be equivalently described by camera coordinates:

$$I(\boldsymbol{x}) = I(R(\boldsymbol{x}' + \boldsymbol{T})) = I'(\boldsymbol{x}'). \tag{2.4}$$

The extrinsic and intrinsic parameters uniquely describe the system geometry at a point in time and are together represented as θ, consisting of nine scalar parameters (three

for translation, three for rotation, two for piercing point, and one for SID). These three coordinate systems are illustrated in Figure 2.1.

2.2.3 Projection and Backprojection Operators

The 2D projection image $P\{I;\theta\}$ evaluated at a point (u,v) is defined as the line integral of linear attenuation coefficients of $I(\boldsymbol{x})$ from the X-ray source to the point $\boldsymbol{u} = (u,v)$ on the detector. In camera coordinates, this is the line integral from $(0,0,0)$ to $(u-u_0, v-v_0, l)$. This line integral path is parameterized by $s \in [0,1]$ and written in camera coordinates as

$$\boldsymbol{p}'(s) = (s(u-u_0), s(v-v_0), sl)\,. \tag{2.5}$$

The length of this integral path is

$$\gamma(u,v;u_0,v_0,l) := \sqrt{(u-u_0)^2 + (v-v_0)^2 + l^2}\,. \tag{2.6}$$

Given the intrinsic and extrinsic parameters, the projection operator $P : L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^2)$ is

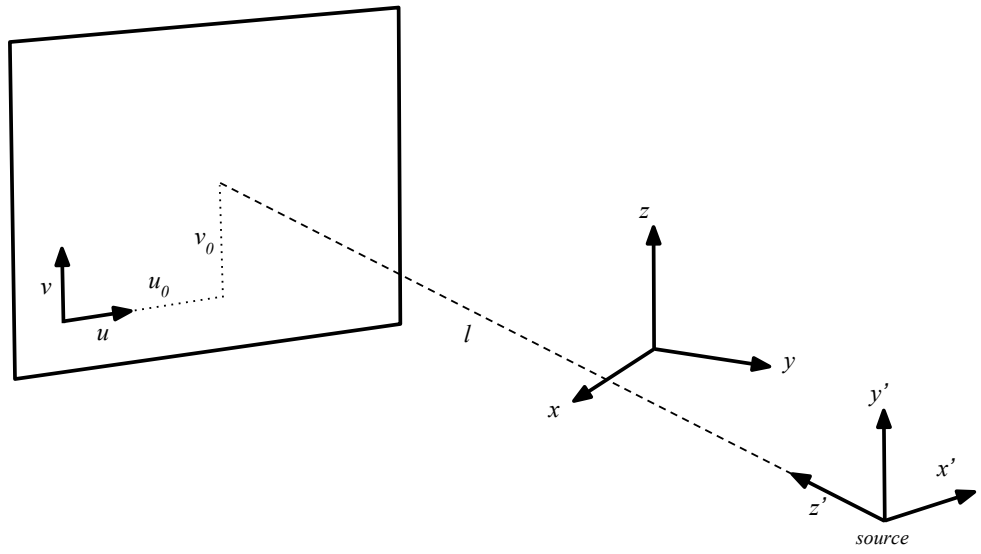


Figure 2.1. Cone-beam coordinate systems: world coordinates (x,y,z), camera coordinates (x',y',z'), and projection coordinates (u,v). In this setup, the rotation between world and camera coordinates is orientation reversing (by convention). The piercing point is the point (u_0,v_0) in the detector plane and the source-to-image-detector distance (SID) is l.

$$P\{I;\theta\}(\boldsymbol{u}) = \gamma \int_0^1 I'(\boldsymbol{p}'(s))\, ds \tag{2.7}$$

$$= \gamma \int_0^1 I(R(\boldsymbol{p}'(s) + \boldsymbol{T}))\, ds. \tag{2.8}$$

Since the projection operator is linear with respect to I, it has an adjoint $P^\dagger : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^3)$. This adjoint is called the backprojection operator and operates on a 2D projection image. Given any image $I(\boldsymbol{x}) \in L^2(\mathbb{R}^3)$ and any image on the detector $f(\boldsymbol{u}) \in L^2(\mathbb{R}^2)$, the adjoint relationship is

$$\Big\langle P\{I;\theta\}(\boldsymbol{u}),\, f(\boldsymbol{u}) \Big\rangle_{L^2(\mathbb{R}^2)} = \Big\langle I(\boldsymbol{x}),\, P^\dagger\{f;\theta\}(\boldsymbol{x}) \Big\rangle_{L^2(\mathbb{R}^3)}. \tag{2.9}$$

This backprojection operator (here defined in camera coordinates) is

$$P^\dagger\{f;\theta\}(\boldsymbol{x}') = \left(\frac{l}{z'}\right)^2 \frac{\|\boldsymbol{x}'\|}{z'} f\Big(\frac{x'l}{z'} + u_0, \frac{y'l}{z'} + v_0\Big). \tag{2.10}$$

The backprojection operator evaluated at a point $\boldsymbol{x}'$ is the value of f at the point $\boldsymbol{u}$, where $\boldsymbol{u}$ is the point on the detector that is on the same line connecting the source and $\boldsymbol{x}'$. This value is then weighted by the inverse of the squared distance from the source to $\boldsymbol{x}'$ to account for the diverging line integrals.

2.2.4 Iterative Cone-Beam Reconstruction

Although analytical methods such as filtered backprojection (i.e., Feldkamp et al. [1]) are widely popular for cone-beam reconstruction, algebraic and statistical methods have shown greatly improved reconstruction images that allow for nonstandard geometries and noisy data [11]. Iterative reconstruction methods can essentially be divided into two categories: algebraic reconstruction and statistical reconstruction. Algebraic methods seek to solve a linear system describing the imaging system. In contrast, statistical methods model noise characteristics and allow for prior information to be incorporated in the estimation. This section reviews some of the most popular methods for iterative CT reconstruction and discusses their properties.

2.2.4.1 Algebraic Reconstruction

In the previous section, I defined the projection and backprojection operators for the cone-beam system analytically. More generally, any X-ray system (including cone-beam, fan-beam, or parallel beam) can be described by a series of line integrals through a volume,

and these line integrals are linear with respect to the 3D volume (see Equation 2.2). Viewing this problem from a discrete standpoint, the entire dataset can be written as a linear system of equations. The unknown 3D image (with N unknown voxel values) is written as y_k, and the data (with M total detector pixel recordings) are written as b_i. Therefore, y_k is the vectorized discrete version of $I(\mathbf{x})$, and b_i is the vectorized discrete version of the full detector data, $\{f_j^*\}$. The forward projection model is written as the $N \times M$ matrix A_{ik}, called the system matrix, so the linear system is modeled as

$$\sum_k A_{ik} y_k = b_i \,, \tag{2.11}$$

where an element of A_{ik} describes the contribution of the kth voxel to the ith line integral (defined by the interpolation scheme used for the numerical integration). A single line integral passes through only a select number of voxels, so A_{ik} is a sparse matrix, and each element is nonnegative. Note that this matrix describes all the line integrals for all the projection images simultaneously. This matrix is very large, and thus a direct solution to Equation 2.11 is infeasible. Therefore, algebraic reconstruction schemes attempt to solve this linear system iteratively, and typically do so using a Landweber scheme, which is a special case of gradient descent.

The very first method for CT reconstruction, introduced by Gordon et al. in 1970, is called the algebraic reconstruction technique (ART) [12], which is equivalent to the Kaczmarz method [13]. This iterative method was introduced even before filtered backprojection, which was introduced by Shepp and Logan in 1974 [14]. The ART image update is

$$y_k^{n+1} = y_k^n + \sum_i A_{ik}^T \left(\frac{b_i - \sum_{k'} A_{ik'} y_{k'}^n}{\sum_{k'} A_{ik'}^2} \right). \tag{2.12}$$

Here, the transpose of the system matrix is the total backprojection operator. The image is updated one pixel measurement (one row of A_{ik}) at a time; that is, after M sequential updates, ART has finished a single iteration through all the projection data. Because of the sequential nature of this algorithm, it is not well suited for parallelization on GPUs.

An improvement to ART is the simultaneous iterative reconstruction technique (SIRT) [15, 16]. In SIRT, the 3D image is updated only after going through the entire dataset. This image update is

$$y_k^{n+1} = y_k^n + \frac{\omega}{A_{+k}} \sum_i A_{ik}^T \left(\frac{b_i - \sum_{k'} A_{ik'} y_{k'}^n}{A_{i+}} \right) \tag{2.13}$$

where

$$A_{i+} = \sum_{k}^{N} A_{ik} \tag{2.14}$$

and

$$A_{+k} = \sum_{i}^{M} A_{ik} \tag{2.15}$$

are the column sums and row sums, and ω is the relaxation term, typically set to one. Each sum over k is the full projection operator and each sum over i is the full backprojection operator. A_{+k} can be thought of as a full projection applied to a 3D image of ones, and A_{i+} can be thought of as a full backprojection applied to 2D images of ones. Assuming that every ray passes through the volume and every volume voxel is intersected with at least one ray, we have $A_{+k} > 0$ for all k and $A_{i+} > 0$ for all j. The SIRT image update can then be written in matrix notation. Defining R as the $N \times N$ diagonal matrix with values of $R_{ii} = 1/A_{i+}$ and C as the $M \times M$ diagonal matrix with $C_{kk} = 1/A_{+k}$, Equation 2.13 is then equivalently written as

$$y^{n+1} = y^n + \omega C A^T R(b - Ay^n), \tag{2.16}$$

where C and R are both positive definite diagonal matrices. This algorithm attempts to minimize the weighted least squares problem

$$L(y) = \|b - Ay\|_R^2 = (b - Ay)^{\mathrm{T}} R(b - Ay). \tag{2.17}$$

The gradient of L with respect to the image y is

$$\nabla_y L(y) = \nabla_y \left((b - Ay)^{\mathrm{T}} R(b - Ay) \right) \tag{2.18}$$

$$= -2A^T R(b - Ay). \tag{2.19}$$

This last step is just an application of the chain rule: by defining $c = b - Ay$, we have $\nabla_y c = -A^T$ and $\nabla_c c^T R c = 2Rc$, so the total gradient is $\nabla_y L = \nabla_y c \nabla_c L$. Equation 2.16 is then rewritten as

$$y^{n+1} = y^n - \frac{\omega}{2} C \nabla_y L(y). \tag{2.20}$$

This update is a form a gradient descent where the gradient is scaled by $\frac{\omega}{2}C$. It can easily be seen that L is convex. By weighting the gradient by this positive definite matrix, this update has been shown to monotonically converge to the weighted least squares solution

given that $\omega \leq 1$ (see [17, 18] for details). This weighted least squares solution is the maximum likelihood estimate (MLE) for a Gaussian noise model.

Despite these appealing theoretical convergence properties, SIRT has a relatively slow convergence since it updates only once after going through all the projection data. To speed up convergence, the simultaneous algebraic reconstruction technique (SART) was introduced [19]. SART has the same update equation as SIRT, but the image is updated after using a single projection. SART is parallelizable over a single projection image and converges very quickly in low noise data, but it is not guaranteed to converge to the MLE, which is especially problematic with inconsistent data (including high noise data).

There have been attempts to design an algorithm that incorporates both the fast "convergence" of SART and the stability of SIRT. One attempt is to use SART updates, but strongly underrelax (i.e., use a much lower value of ω [19]). Another such method is to update the image over disjoint ordered subsets of the data, which is called OS-SIRT [20] or OS-SART [21] (the terms are used interchangibly), where OS stands for ordered subsets. The image updates of SIRT, OS-SIRT, and SART can all be written generally as

$$I(\mathbf{x}) \mapsto I(\mathbf{x}) + \frac{\omega}{\sum_{j \in S} P_j^\dagger\{\mathbf{1}(\mathbf{u})\}} \sum_{j \in S} P_j^\dagger \left\{ \frac{f_j^*(\mathbf{u}) - P_j\{I(\mathbf{x})\}}{P_j\{\mathbf{1}(\mathbf{x})\}} \right\}. \tag{2.21}$$

Here, S is the current subset, $\mathbf{1}(\mathbf{u})$ is a 2D projection image of all ones, and $\mathbf{1}(\mathbf{x})$ is a 3D volume of all ones. In the case of SIRT, there is only one subset consisting of all projections. For SART, each subset contains a single projection image. For OS-SIRT, it is somewhere in between. Note that in none of the ordered subset methods (except SIRT) is the algorithm guaranteed to converge. However, OS-SIRT has been shown to exhibit large speedups in practice due to the relative consistency in the projection data. The choice in the subset size is a trade-off between convergence and speed, and more consistent datasets allow for a faster reconstruction with smaller subsets.

2.2.4.2 Statistical Reconstruction

Statistical reconstruction allows for a more complex modeling of the projection system, including a more accurate noise model as well as a prior belief of the reconstructed image. One of the most well-known statistical reconstruction methods is expectation maximization (EM) reconstruction.

The standard EM image update, originally introduced for emission tomography by Shepp and Vardi in 1982 [22], is a multiplicative update:

$$I(\boldsymbol{x}) \mapsto \frac{I(\boldsymbol{x})}{\sum_j P_j^\dagger\{\mathbf{1}(\mathbf{u})\}(\boldsymbol{x})} \sum_j P_j^\dagger \left\{ \frac{f_j^*}{P_j\{I\}} \right\}(\boldsymbol{x}) \,, \tag{2.22}$$

where each image update is known as a Richardson-Lucy iteration [23]. These updates converge to the maximum likelihood estimate (MLE) of the complete data with a Poisson model at the detector elements [24]. The log likelihood with the Poisson model is

$$\ell(y_k|b_i) = \sum_i \left(b_i ln\Big(\sum_k A_{ik} \Big) - ln(b_i!) - \sum_k A_{ik} y_k \right). \tag{2.23}$$

The EM algorithm is a general framework for maximizing the posterior or likelihood when some data are "missing". These "missing" data are some unknown data that would greatly simplify the estimation. In our case, the observed data b_i are the weighted sum of voxels along the line from the source to the detector. This sum of voxels inherent in the integration makes the estimation difficult, so we write the missing data as z_{ik}: the contribution of the kth voxel to the line integral measured at the ith detector element. Then, the observed data are written as

$$b_i = \sum_k z_{ik}. \tag{2.24}$$

Given the missing data, the maximum likelihood of y_k is just the weighted average of the missing data for each recording:

$$y_k = \frac{\sum_i z_{ik}}{\sum_i A_{ik}}. \tag{2.25}$$

This is the M-step of the EM algorithm. The E-step is also straightforward. Here, the conditional expectation of z_{ik} given the observed data b_i is

$$z_{ik} = \frac{b_i A_{ik} y_k}{\sum_{k'} A_{ik'}}. \tag{2.26}$$

Combining the E-step and the M-step yields

$$y_k = \frac{\sum_i \frac{b_i A_{ik} y_k}{\sum_{k'} A_{ik'}}}{\sum_i A_{ik}} = \frac{y_k}{\sum_i A_{ik}} \sum_i \frac{b_i A_{ik}}{\sum_{k'} A_{ik'}} \,, \tag{2.27}$$

which is just Equation 2.22 written in its discrete form.

In transmission tomography, the detector data are essentially Poisson distributed [25]. Like the algebraic case, the detector data are log-transformed to get the line integrals of the

attenuation (see Equation 2.2). This log-transformed data are neither Poisson nor Gaussian distributed. However, EM reconstruction can be applied to solve any linear inverse problem where the objection function is nonnegative and equivalently minimizes Csiszar's I-divergence [26, 27]. EM models for transmission have been introduced that maximize the true log-likelihood of the transmission data given the Poisson model [28, 29], but they converge very slowly [30] and are rarely used in practice. In these implementations, the projection data do not form an exponential family, so the M-step cannot be completed in closed form.

Unfortunately, the ML-EM algorithm (like other MLE algorithms) tends to amplify noise by overfitting to the data [23, 31, 32]. To regularize the estimation, I use a Bayesian reconstruction framework where I place a total variation (TV) prior on the 3D image:

$$U(I) = \int_{\Omega} |\nabla I(\boldsymbol{x})| \, d\boldsymbol{x} \,, \tag{2.28}$$

and solve for the *maximum a posteriori* (MAP) estimate, which maximizes the log posterior

$$ln \;\; p(I|f^*) \propto \ell(I|f^*) + \lambda U(I) \,, \tag{2.29}$$

where λ is a scalar parameter that controls the strength of the TV penalty. Adding a prior is very natural in the EM framework by using a "one-step-late" approximation proposed by Green by adding the derivative of the TV penalty of the current estimate in the denominator [33].

For the full EM implementation, $I(\boldsymbol{x})$ is updated after summing over all the data. For a faster convergence, the image is updated after summing over a single ordered (striated) subset S, before iterating through all the subsets of the data [34] in the same way as the algebraic case. The resulting algorithm is called ordered subset expectation maximization (OSEM) reconstruction.

The OSEM image update with the TV prior is then

$$I(\boldsymbol{x}) \mapsto \frac{I(\boldsymbol{x})}{\sum_{j \in S} P_j^{\dagger}\{\mathbf{1}(\mathbf{u})\}(\boldsymbol{x}) + \lambda \frac{\partial U(I)}{\partial I(\boldsymbol{x})}} \sum_{j \in S} P_j^{\dagger} \left\{ \frac{f_j^*}{P_j\{I\}} \right\}. \tag{2.30}$$

Here $\frac{\partial U(I)}{\partial I(\boldsymbol{x})}$ is the derivative of the TV penalty with respect to the image. I implement the TV derivative in the same manner as [26], using the following stencil for $\frac{\partial U(I)}{\partial I(\boldsymbol{x})}$:

$$\frac{\partial U(I)}{\partial I(\boldsymbol{x})} = \frac{\delta_{x^-} I(\boldsymbol{x})}{u(x-1,\cdot,\cdot)} + \frac{\delta_{y^-} I(\boldsymbol{x})}{u(\cdot,y-1,\cdot)} + \frac{\delta_{z^-} I(\boldsymbol{x})}{u(\cdot,\cdot,z-1)} - \frac{\delta_{x^-} I(\boldsymbol{x}) + \delta_{y^-} I(\boldsymbol{x}) + \delta_{z^-} I(\boldsymbol{x})}{u(\boldsymbol{x})}, \tag{2.31}$$

where

$$u(\boldsymbol{x}) = \sqrt{\delta_{x^+} I(\boldsymbol{x})^2 + \delta_{y^+} I(\boldsymbol{x})^2 + \delta_{z^+} I(\boldsymbol{x})^2 + \epsilon^2}, \tag{2.32}$$

where δ_{x^+} and δ_{x^-} are the partial derivatives solved by using forward and backward differences, respectively, and ϵ is a very small number to ensure divisibility.

2.2.4.3 Analysis of Reconstruction Techniques

To compare these reconstruction techniques, I implemented SART (both nonrelaxed and relaxed versions), OS-SIRT, SIRT, OSEM, and TV-OSEM. I used the public domain CT dataset from the University of North Carolina (`http://graphics.stanford.edu/data/voldata/`) to create a projection dataset simulated with Poisson noise. The ground truth dataset as well as an example projection can be seen in Figure 2.2. This dataset was simulated with 250 equally-spaced projections over a 223 degree scan. I used 640^2 projection images and reconstructed a 256^3 volume.

Since the intent of this section is to compare these reconstruction methods, the "static" version of each reconstruction method is used; for each reconstruction method, the ground truth geometric parameters were used, so there is no uncertain geometry in these experiments.

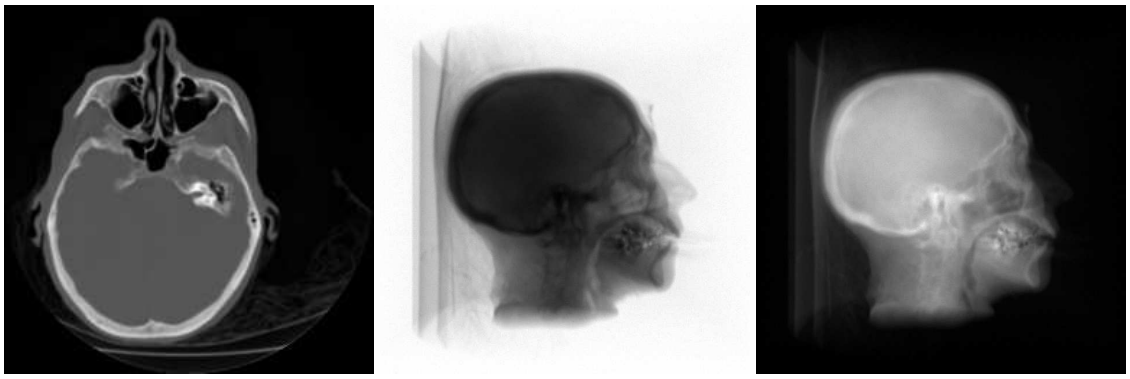


Figure 2.2. Setup for the 3D reconstruction algorithm analysis. On the left is the ground truth digital phantom. In the center is a projection simulated with Poisson noise. On the right is the center projection image that has been log-transformed.

The reconstructed images for each of the methods after 50 iterations are shown in Figure 2.3. The L^2 error between the reconstructed images and the ground truth image at each iteration is plotted in Figure 2.4. Note that none of these methods are necessarily expected to converge to the ground truth image. Rather, these methods seek to converge to the MLE of the reconstruction given the data (or the MAP estimate in the TV-OSEM case). However, the ground truth error gives insight into the accuracy of each reconstruction method and how each method handles noisy data.

As previously mentioned, SIRT reconstruction converges very slowly. Since it is not yet converged after 50 iterations, the SIRT reconstruction image is still overly smooth. The SART reconstruction image is very noisy, and it did not converge to the maximum likelihood estimate (as global convergence is not guaranteed in SART, especially for inconsistent systems). The two techniques that attempt to have better stability than SIRT, yet still seek

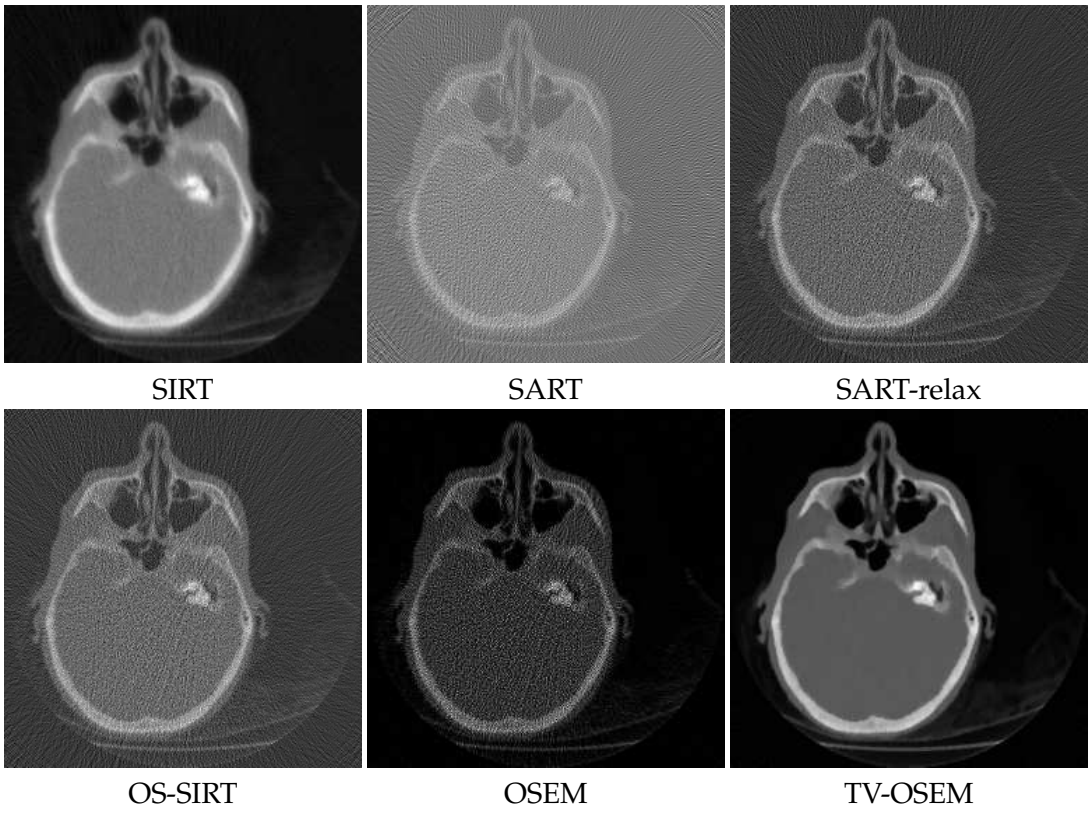


Figure 2.3. Algebraic and statistical reconstruction results for the six tested reconstruction methods. Each method was run for 50 iterations. Each image is displayed using the entire intensity range of the central axial slice. SART, SART-relax, and OS-SIRT all have negative values in the reconstructed image.

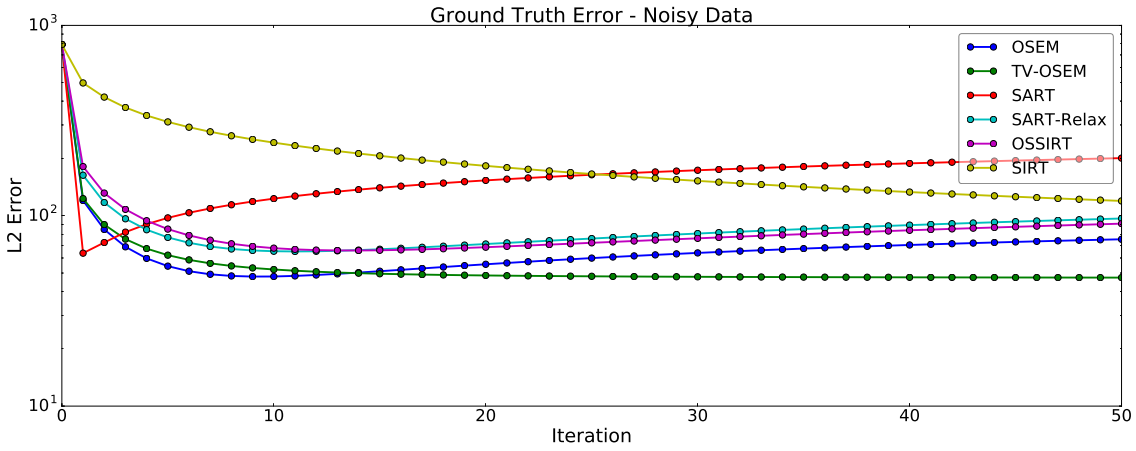


Figure 2.4. Convergence of the ground truth error for each of the reconstructed images, log scale. TV-OSEM converges quickly and is closest to the ground truth solution. All MLE methods converge to very noisy images, and the algebraic methods (which have no positivity constraint) have significant noise in regions of the image close to zero.

to converge to the maximum likelihood estimate, are SART with relaxation ($\omega = 0.2$), and OS-SIRT (using 20 ordered subsets). In all the ordered subset cases, bit-reversal ordering was used, which has been shown to improve convergence [35, 36]. Both relaxed SART and OS-SIRT converged to the maximum likelihood estimate and did so much faster than SIRT. In a separate experiment, SIRT was run for 2000 iterations. The resulting reconstruction looks identical to the reconstructions from OS-SIRT and relaxed SART, as all three converged to the maximum likelihood estimate. Therefore, the noise properties of relaxed SART and OS-SIRT reconstructions are due to the fact that the maximum likelihood estimate itself is very noisy.

The statistical methods (OSEM and TV-OSEM) were also implemented. OSEM converges quickly but is still quite noisy inside the skull, whereas of all the methods, TV-OSEM is by far the closest to the ground truth skull and looks most like the ground truth image.

All the maximum likelihood reconstruction methods produce noisy reconstructions by overfitting to the data. However, one significant difference between the algebraic techniques and OSEM is that OSEM naturally constrains the resulting reconstruction image to be nonnegative, as all its updates are multiplicative. The algebraic techniques have significant noise in areas of the image close to zero due to negative values in the reconstruction.

The most computationally complex parts of these reconstruction techniques are the

projection and backprojection operations, where the projection is the more complex of the two. Each of these methods has one backprojection operation and one projection operation per projection image. The projection of ones and the backprojection of ones is a fast operation, since no interpolation is required. Therefore, each of these methods has approximately the same computational complexity.

Overall, the maximum likelihood methods produce very noisy reconstruction results. The Bayesian reconstruction method (TV-OSEM) by far produces the best results: fast convergence, a reconstruction with minimal noise, and a reconstruction that is closest to the ground truth volume. For the rest of this chapter, I will use TV-OSEM for the image estimation.

2.2.5 Geometry Estimation

For each projection, I seek to minimize the mismatch between the projection data (f_j^*) and the associated projection operator applied to the current estimate of the reconstructed volume ($P_j\{I;\theta\}$). Non-ideal characteristics in real cone-beam data, such as X-ray scatter and limited field of view, create non-uniform intensity changes that do not become realized in the reconstructed image. Therefore, a local normalized cross-correlation (NCC) for the geometry energy functional was used. Using NCC has the effect of matching structures in the data to structures in the reconstructed image as opposed to strictly looking at intensity differences.

The normalized cross-correlation between two images $f(\boldsymbol{u})$ and $g(\boldsymbol{u})$ is defined as the cross-correlation between the mean-removed versions of f and g divided by their standard deviations:

$$NCC(f,g) = \frac{\int (f(\boldsymbol{u}) - \mu_f)(g(\boldsymbol{u}) - \mu_g)}{\sqrt{\int (f(\boldsymbol{u}) - \mu_f)^2}\sqrt{\int (g(\boldsymbol{u}) - \mu_g)^2}}. \tag{2.33}$$

To account for nonglobal intensity variations, I calculate these integrals over local regions of locally mean-removed images. I define the locally mean-removed images $\bar{f}(\boldsymbol{u}) = f(\boldsymbol{u}) - k * f(\boldsymbol{u})$, $\bar{g}(\boldsymbol{u}) = g(\boldsymbol{u}) - k * g(\boldsymbol{u})$, where k is a local normalized kernel (in my implementation, I use a Gaussian kernel). The local integration can be written as a convolution, so the local NCC at a single point $\boldsymbol{u}$ is then

$$NCC(f, g, \boldsymbol{u}) = \frac{k * (\bar{f}\bar{g})}{\sqrt{(k * \bar{f}^2)(k * \bar{g}^2)}}. \tag{2.34}$$

The energy functional at a single projection is

$$E_j(\theta) = \int_{\Omega_d} NCC(P\{I;\theta\}, f_j^*, \boldsymbol{u})\, d\boldsymbol{u}, \tag{2.35}$$

where Ω_d is the 2D region defined by the image detector bounds. The first variation of this energy functional with respect to the extrinsic and intrinsic parameters is

$$\delta E_j(\theta) = \int_{\Omega_d} \nabla_\theta NCC(P\{I;\theta\}, f_j^*, \boldsymbol{u})\, d\boldsymbol{u}, \tag{2.36}$$

where

$$\nabla_\theta NCC\left(P\{I;\theta\}, f_j^*, \boldsymbol{u}\right) = \frac{\left(k * \bar{P}^2\right)\left(k * \left(\bar{f}_j^* \nabla_\theta \bar{P}\right)\right) - \left(k * \bar{P}\bar{f}_j^*\right)\left(k * \bar{P}\nabla_\theta \bar{P}\right)}{\left(k * \bar{P}^2\right)^{3/2}\sqrt{k * \bar{f}_j^{*2}}} \tag{2.37}$$

is just an application of the quotient rule, and

$$\nabla_\theta \bar{P} = \nabla_\theta P\{I;\theta\} - k * \nabla_\theta P\{I;\theta\}. \tag{2.38}$$

I now analytically solve for $\nabla_\theta P\{I;\theta\}(\boldsymbol{u})$, the variation of the projection operator with respect to the extrinsic and intrinsic parameters. In the rest of this section, this variation is derived by solving for the variation of the translation, rotation, and intrinsic parameters separately.

2.2.5.1 Extrinsic Parameter Derivatives

I first solve for the derivative with respect to the translation parameter using the chain rule:

$$\frac{\partial}{\partial \boldsymbol{T}} P\{I\}(\boldsymbol{u}) = \gamma \int_0^1 \frac{\partial}{\partial \boldsymbol{T}} I(R(\boldsymbol{p}'(s) + \boldsymbol{T}))\, ds \tag{2.39}$$

$$= \gamma\, R^{\mathrm{T}} \int_0^1 (\nabla I)(R(\boldsymbol{p}'(s) + \boldsymbol{T}))\, ds. \tag{2.40}$$

Once again, $\gamma = \gamma(u, v; u_0, v_0 l)$ is the length factor defined in Equation 2.6.

The derivative with respect to R is slightly more complicated because I must constrain R to $\mathrm{O}(3)$. The update of R is written as an orientation preserving rotation applied to the left of R. These orientation preserving 3D rotation matrices form the special orthogonal group, $\mathrm{SO}(3)$. The tangent space of $\mathrm{SO}(3)$ at identity is $\mathfrak{so}(3)$, the space of skew-symmetric

matrices. The exponential map $\exp : \mathfrak{so}(3) \to \mathrm{SO}(3)$ is the matrix exponential, where the matrix exponential of any skew-symmetric matrix is guaranteed to be a rotation matrix. For any square matrix W, $W - W^{\mathrm{T}}$ is a skew-symmetric matrix. The rotation update is then written as:

$$R \mapsto \exp(W - W^{\mathrm{T}})R. \tag{2.41}$$

We now have an unconstrained optimization over W while assuring that R remains in $\mathrm{O}(3)$. I then solve for the Fréchet variation of the projection operator:

$$\begin{aligned}\left\langle \frac{\partial}{\partial R} P\{I\}(\boldsymbol{u}), W \right\rangle_{Frob} = \\ \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma \int_0^1 I\left(\exp\big(\epsilon(W - W^{\mathrm{T}})\big) R\big(\boldsymbol{p}'(s) + \boldsymbol{T}\big)\right) ds.\end{aligned} \tag{2.42}$$

Using the chain rule, we have

$$\frac{d}{d\epsilon}\exp(\epsilon(W - W^{\mathrm{T}})) = (W - W^{\mathrm{T}})\exp(\epsilon(W - W^{\mathrm{T}})), \tag{2.43}$$

and then

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \exp(\epsilon(W - W^{\mathrm{T}})) = (W - W^{\mathrm{T}}). \tag{2.44}$$

Using this with the chain rule on Equation 2.42 and evaluating the derivative at $\epsilon = 0$, we have

$$\left\langle \frac{\partial}{\partial R} P\{I\}(\boldsymbol{u}), W \right\rangle = \gamma \int_0^1 (\nabla I)(R(\boldsymbol{p}'(s) + \boldsymbol{T}))^{\mathrm{T}}(W - W^{\mathrm{T}})R(\boldsymbol{p}'(s) + \boldsymbol{T})\, ds. \tag{2.45}$$

For ease of notation, I introduce two vectors $\boldsymbol{w} := (\nabla I)(R(\boldsymbol{p}'(s) + \boldsymbol{T}))$ and $\boldsymbol{v} := R(\boldsymbol{p}'(s) + \boldsymbol{T})$, so the previous equation is written as

$$\left\langle \frac{\partial}{\partial R} P\{I\}(\boldsymbol{u}), W \right\rangle = \gamma \int_0^1 \boldsymbol{w}^{\mathrm{T}}(W - W^{\mathrm{T}})\boldsymbol{v}\, ds. \tag{2.46}$$

I then use the following identity of the Frobenius inner product:

$$\boldsymbol{w}^{\mathrm{T}} A \boldsymbol{v} = \langle \boldsymbol{w}\boldsymbol{v}^{\mathrm{T}}, A \rangle_{Frob}. \tag{2.47}$$

Equation 2.46 becomes

$$\left\langle \frac{\partial}{\partial R} P\{I\}(\boldsymbol{u}), W \right\rangle = \gamma \int_0^1 \langle \boldsymbol{w}\boldsymbol{v}^{\mathrm{T}}, W - W^{\mathrm{T}} \rangle\, ds \tag{2.48}$$

$$= \gamma \int_0^1 \langle \boldsymbol{w}\boldsymbol{v}^{\mathrm{T}} - \boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}, W \rangle\, ds. \tag{2.49}$$

The last step is just an expansion of the inner product ($\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}} = (\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}})^{\mathrm{T}}$). Here $\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}} - \boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$ is a skew-symmetric matrix, so the integrand is transformed via the star map

$$\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}} - \boldsymbol{v}\boldsymbol{w}^{\mathrm{T}} = *(\boldsymbol{w} \times \boldsymbol{v}), \tag{2.50}$$

where the star is the standard mapping between $\mathfrak{so}(3)$ and $\mathbb{R}^3$:

$$\begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \overset{*}{\rightleftarrows} \begin{pmatrix} a \\ b \\ c \end{pmatrix}. \tag{2.51}$$

Then, I write the derivative as a vector in $\mathbb{R}^3$:

$$\frac{\partial}{\partial R} P\{I\}(\boldsymbol{u}) = \gamma \int_0^1 \boldsymbol{w} \times \boldsymbol{v}\, ds \tag{2.52}$$

$$= \gamma \int_0^1 (\nabla I)(R(\boldsymbol{p}'(s) + \boldsymbol{T})) \times R(\boldsymbol{p}'(s) + \boldsymbol{T})\, ds. \tag{2.53}$$

Given an update vector $\boldsymbol{b} \in \mathbb{R}^3$, the update of R by $\boldsymbol{b}$ is

$$R \mapsto \exp(*\boldsymbol{b})R. \tag{2.54}$$

I apply the rotation update by calculating the exponential map in closed form using Rodrigues' rotation formula. Setting $\alpha = \|\boldsymbol{b}\|$ and $\hat{\boldsymbol{b}} = \boldsymbol{b}/\|\boldsymbol{b}\|$, I have

$$R \mapsto \left(\mathrm{I} + \sin(\alpha) * \hat{\boldsymbol{b}} + (1 - \cos(\alpha))(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}^{\mathrm{T}} - \mathrm{I})\right)R. \tag{2.55}$$

2.2.5.2 Intrinsic Parameter Derivatives

I now find the derivatives with respect to the piercing point and SID. For simplicity, the intrinsic parameters are combined into a single vector $\boldsymbol{\tau} = (u_0, v_0, -l)$. In this notation, the integral path is

$$\boldsymbol{p}'(s) = s(u - u_0, v - v_0, l) = s((u, v, 0) - \boldsymbol{\tau}) \tag{2.56}$$

and its derivative is

$$\frac{\partial}{\partial \boldsymbol{\tau}} \boldsymbol{p}'(s) = -s. \tag{2.57}$$

The length factor γ can be rewritten as

$$\gamma(u, v; \boldsymbol{\tau}) = \|(u, v, 0) - \boldsymbol{\tau}\| \tag{2.58}$$

and its derivative is

$$\frac{\partial}{\partial \boldsymbol{\tau}} \gamma(u, v; \boldsymbol{\tau}) = \frac{-((u, v, 0) - \boldsymbol{\tau})}{\|(u, v, 0) - \boldsymbol{\tau}\|}. \tag{2.59}$$

Since γ and the line integral are both functions of the intrinsic parameters, I solve for the derivative of $P\{I;\theta\}(\boldsymbol{u})$ using the product rule:

$$\frac{\partial}{\partial \boldsymbol{\tau}} \gamma \int_0^1 I(R(\boldsymbol{p}'(s) + \boldsymbol{T}))\, ds = \frac{-((u,v,0) - \boldsymbol{\tau})}{\|(u,v,0) - \boldsymbol{\tau}\|^2} P\{I\}(u,v) - \gamma(u,v;\boldsymbol{\tau}) R^{\mathrm{T}} \int_0^1 s(\nabla I)(R(\boldsymbol{p}'(s) + \boldsymbol{T}))\, ds. \quad (2.60)$$

The evaluation of the gradients with respect to all geometric parameters requires integrating the following:

$$I(R(\boldsymbol{p}'(s) + T)) \quad (2.61)$$

$$(\nabla I)(R(\boldsymbol{p}'(s) + T)) \quad (2.62)$$

$$(\nabla I)\big((R(\boldsymbol{p}'(s) + T)) \times (R(\boldsymbol{p}'(s) + T))\big) \quad (2.63)$$

$$s(\nabla I)(R(\boldsymbol{p}'(s) + T)). \quad (2.64)$$

The variations solved for in Equations 2.40, 2.53, and 2.60 are then individually plugged into Equation 2.38. The local normalized cross-correlation gradient is then integrated over $(u,v) \in \Omega_d$ (the bounds of the detector) as shown in Equation 2.36, yielding a vector in $\mathbb{R}^9$.

2.2.6 Implementation

I now describe the final algorithm for joint image and geometry estimation, as shown in Algorithm 1. For each subset of the projection data, I apply the multiplicative OSEM update as in Equation 2.30. Then, after all subsets apply their updates, I estimate all the geometric parameters using a gradient ascent with a line search using Brent's method. I repeat these updates until both the reconstructed image and geometric parameters converge.

Algorithm 1 Joint Reconstruction and Geometry Estimation

```
Divide data into k ordered subsets:
    {(0, k, 2k, ...), (1, k + 1, 2k + 1, ...)...}
I(x) ← 1
for i = 0...numIters do
    for S in subsets do
        Perform OSEM update for the current subset
    end for
    for projection j in all projections do
        Calculate parameter derivatives for f_j^*
        Update geometric parameters of f_j^*
    end for
end for
```

For the projection operator (Equation 2.8), I implement a single line integral for each detector element by ray-marching through $I(R(\boldsymbol{p}'(s) + \boldsymbol{T}))$ and evaluating the line integral at the points where the path intersects the faces of the image voxel grid (and only the intersections that are most perpendicular to the path). By doing this, I only need to perform bilinear interpolations of the 3D image. Otherwise, I would be interpolating at the voxel grid interior, requiring a more costly trilinear interpolation at each point on the path.

For the backprojection operator (Equation 2.10), I implement a voxel-based backprojection, where I perform a bilinear interpolation on the projection grid. This is not the numerical adjoint of the projection operator, but is considerably faster than a projection-splatting backprojection. Using mismatched projection and backprojection operators is supported in the literature [37] and has even been shown to improve convergence [38].

Evaluating all the integrals for the parameter gradients (Equations 2.62-2.64) requires integrating through ∇I. This line integral is implemented using a fixed step size ray-marching, and $(\nabla I)(R(\boldsymbol{p}'(s) + \boldsymbol{T}))$ is evaluated at each location using central differences while simultaneously evaluating $I(R(\boldsymbol{p}'(s) + \boldsymbol{T}))$. Computing the gradient on-the-fly has the advantage of having a much smaller memory footprint than precomputing ∇I, which is especially advantageous for GPU implementation where memory is limited.

Since iterative reconstruction algorithms often have limited field of view edge artifacts in the reconstructed image, I only evaluate the gradient integrals on the interior of the reconstructed image, which I define as locations in the 3D image where 90% of the projections pass through.

I implement the algorithm in a multiscale scheme, where I downsample the projection data and reconstructed volume by factors of 4, 2, and 1, keeping the estimate of the geometric parameters after each scale.

2.3 Results

I have joint reconstruction and geometry estimation results for four datasets using three different mobile C-arm systems. For the first dataset, I used orientation data from electromagnetic sensors attached to a C-arm and artificially created a projection dataset using these parameters as ground truth. With these ground truth parameters and a ground truth 3D image, I can compare the reconstruction results quantitatively. For the three remaining datasets, I used real projection data from mobile C-arms absent any ground truth 3D image. I compare results with a static TV-OSEM reconstruction and my joint reconstruction method, and visually compare the final reconstructions.

2.3.1 Ground Truth Geometric Parameters

The first dataset used tracking data during an orbital scan using a GE-OEC 9800 mobile C-arm with an image intensifier. A pair of electromagnetic position and orientation sensors were placed on the C-arm detector and the operating table. During a C-arm sweep, the orientation and displacement between the moving C-arm detector sensor and the fixed table sensor were measured. The location of the source and the intrinsic parameters for each projection were derived using a calibration target attached to the detector using precision high-attenuation markers [39]. This scan consisted of 144 acquisitions spaced 1 degree per frame, yielding a limited-angle scan.

Since there is no ground truth volume for this dataset, I used the previously mentioned digital skull phantom and simulated projections. The projection images were created using the given parameter values and were simulated with Poisson noise before being log-transformed.

I reconstructed this dataset in three scenarios: given the ground truth parameters, given nominal parameters with no geometry estimation, and given the nominal parameters while jointly estimating the geometry. The nominal trajectory was made by creating a circular trajectory approximating the true trajectory and using equally-spaced gantry

rotations. The reconstruction images of these experiments are shown in **Figure 2.5**. The reconstructed image without geometry estimation (given a nominal trajectory) is unrecognizable, whereas the reconstructed image with geometry estimation has image quality comparable to the scan given true parameters. The L^2 error between the reconstruction and the original ground truth image is 10.5 for the ground truth parameters, 233.2 for the nominal parameters, and 13.0 for the joint reconstruction given the nominal parameters.

I also evaluated the accuracy of the estimated parameters. **Figure 2.6** shows the ground truth, nominal, and estimated acquisition paths. In **Figure 2.7**, I quantified the accuracy of the estimated extrinsic parameters in contrast to the nominal parameters. Although I am optimizing over two separate objective functionals (log posterior for the image update, and NCC for geometry update), they both converge after 50 iterations (shown in **Figure 2.8**). The change in the reconstructed image is imperceptible after about 25 iterations, so in practice the algorithm is stopped then.

As a preprocessing step for the joint reconstruction, I performed an initial estimate of the translation in the x' and y' directions by performing a normalized cross-correlation 2D registration of consecutive projection images. Since the X-rays are divergent, the resulting translation is multiplied by z'/l for the initial estimates of $T_{x'}$ and $T_{y'}$. This preprocessing step leads to a more stable and faster convergence, since this 2D registration is considerably faster than integrating through ∇I.

In all three scenarios, I reconstructed a 256^3 volume using 640^2 projection images. The static OSEM reconstructions (without parameter estimation) took approximately two minutes to complete, whereas the reconstruction with pose estimation took approximately four minutes on a single Titan Z GPU.

2.3.2 Real C-arm Data

2.3.2.1 Physical Phantoms

The second and third datasets were acquired using a GE-OEC 6800 mobile C-arm experimentally retrofitted with improved acquisition and control components. The original low-watt monoblock was replaced with a pulse-capable high-power source block, and the image intensifier was replaced with a flat-panel detector. The acquisition scan of this retrofitted system is nearly isocentric. Although it has improved tube and detector

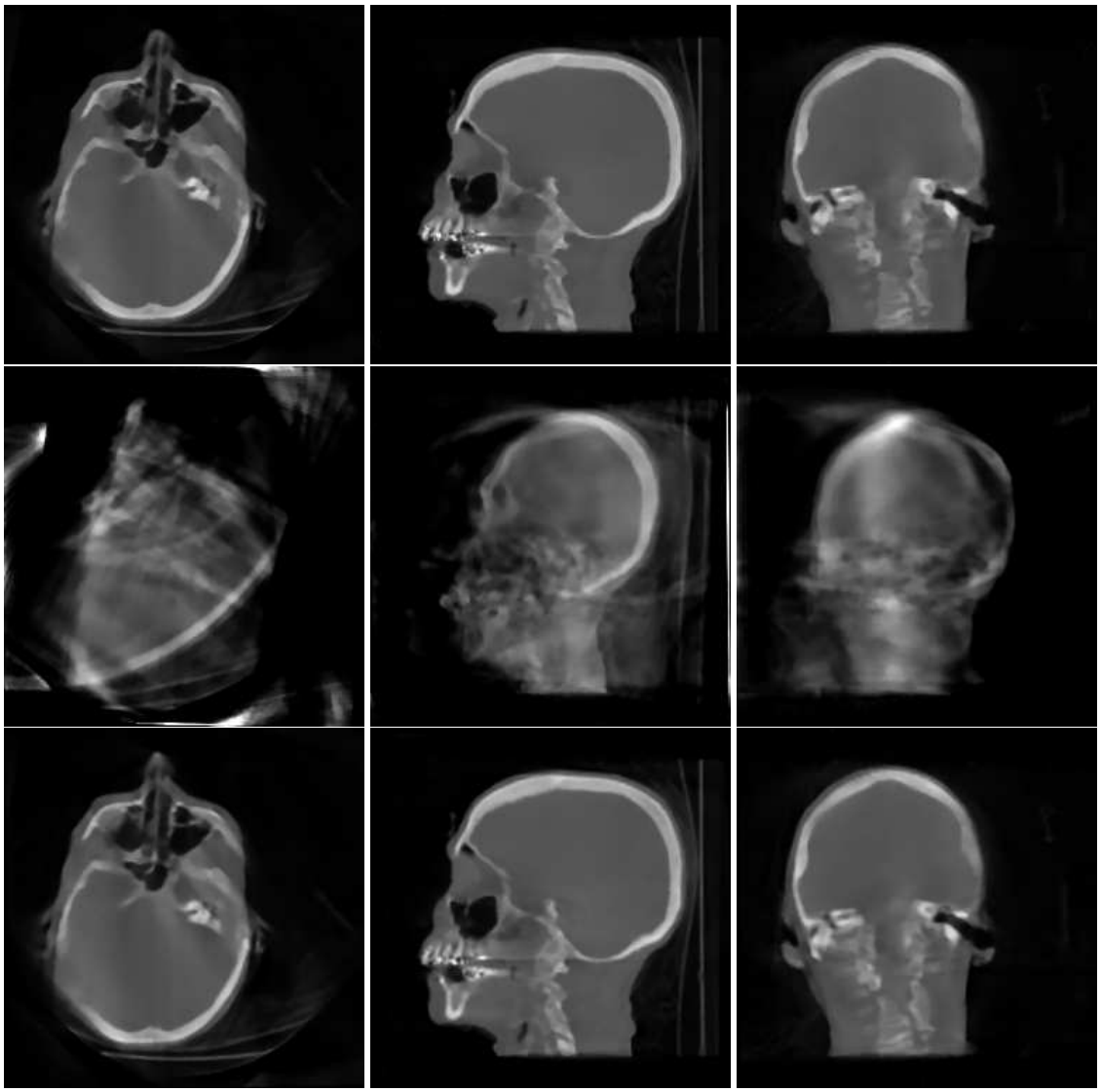


Figure 2.5. 3D reconstruction using digital skull phantom and ground truth geometric parameters (axial, sagittal, and coronal slices). Top row: static reconstruction with ground truth parameters. Middle row: static reconstruction with nominal parameters. Bottom row: joint image and geometry estimation given nominal parameters. Both the top and bottom results show limited angle artifacts, but the image quality of the ground truth parameter dataset and the geometry estimation dataset are comparable.

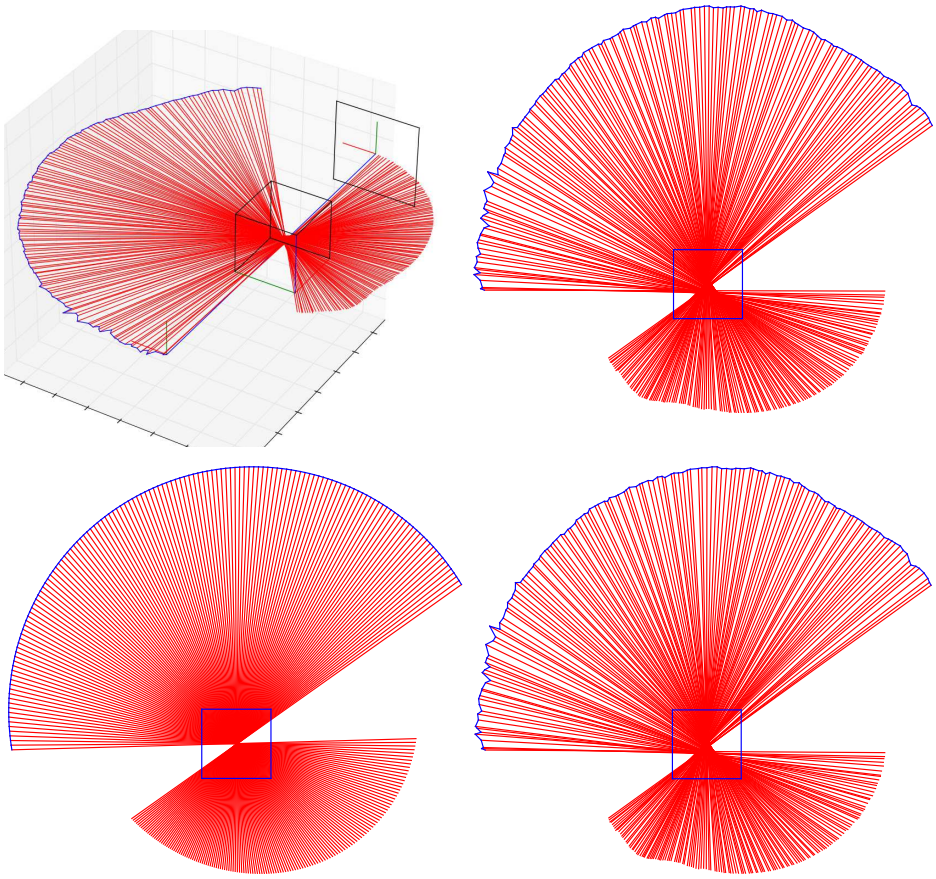


Figure 2.6. Cone-beam acquisition paths. The X-ray source of the C-arm follows the path along the blue curve, each red line shows the path from the X-ray source to the center of the detector, and the box represents the bounds of the reconstruction region. Upper left: 3D view of the ground truth acquisition path including the reconstruction region and the detector. Upper right: 2D plane of the ground truth acquisition path. Lower left: 2D plane of the nominal (given) path. Lower right: 2D plane of the estimated path (given the nominal path).

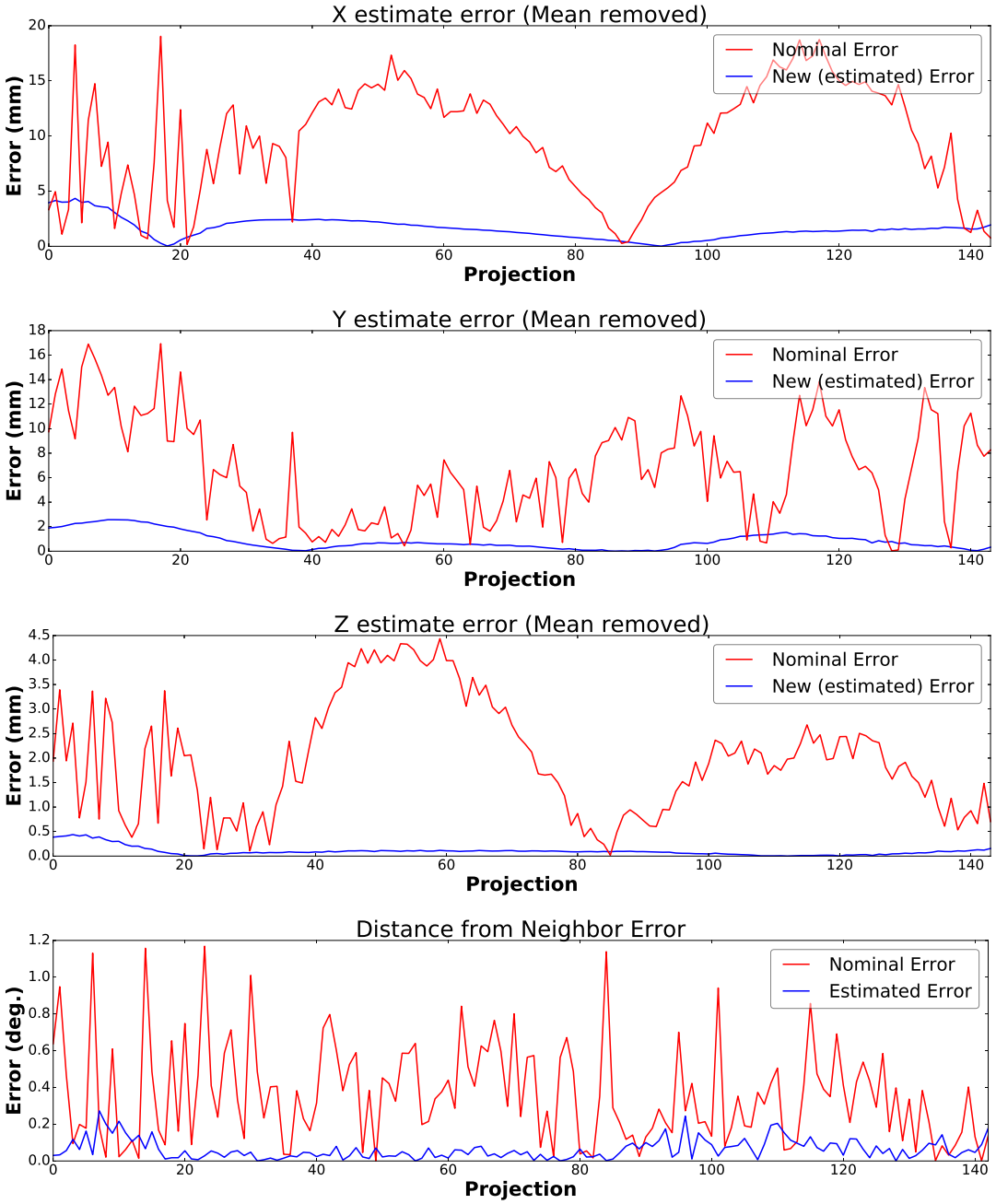


Figure 2.7. Nominal trajectory error and estimated geometry error for all 144 projections. From top to bottom, I show the x, y, and z world coordinate translation error (in mm) and the rotation error between neighboring rotation matrices (in degrees). The proposed method accurately estimates the true parameters. For the rotation estimation, I plot the error between the ground truth rotation step (rotation difference between neighboring acquisitions) and the estimated and nominal rotation steps. Note that if we have a global rotation or a global translation represented in all the parameters, we would reconstruct an identical image that is a rotated version or translated version of the true image. Therefore, I remove the global translation and rotation when evaluating these results.

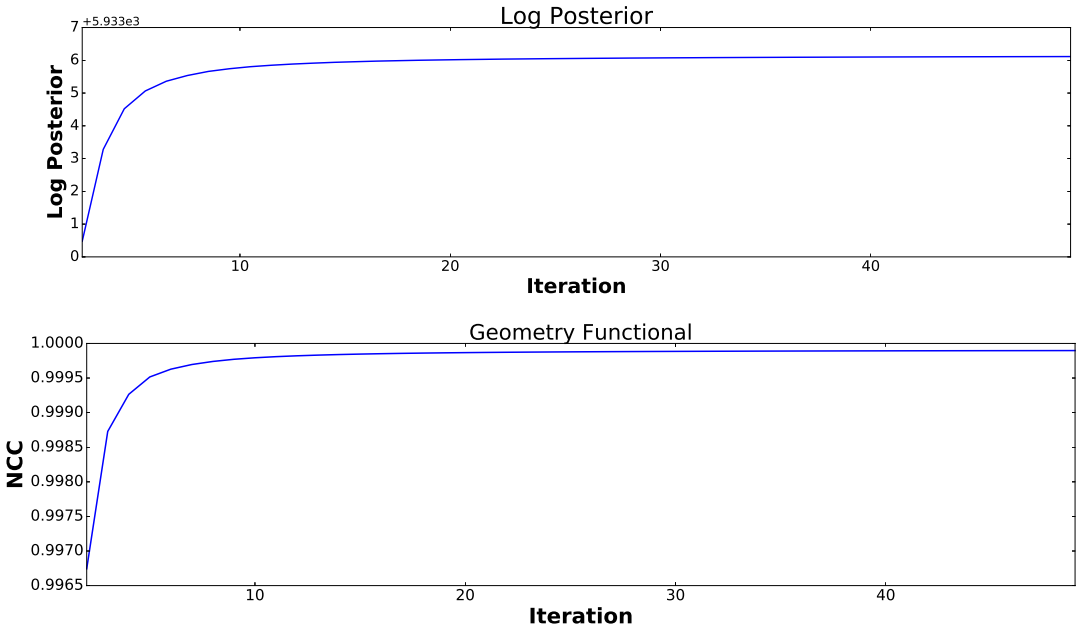


Figure 2.8. Joint reconstruction convergence of the two objective functionals. The top plot is the convergence of the log posterior (for the image updates). The bottom plot is the convergence of the average NCC (for the geometry updates). Both functionals have converged after 50 iterations.

components, the geometric parameters are still considered approximate. Even so, the nominal scan is much closer to the true scan than the other scenarios in this chapter. Because of this, the nominal reconstruction has only moderate miscalibration artifacts.

Orbital scans of a knee and a skull phantom were acquired. Both physical phantoms consist of cadaver bones encased in plastic. There are no given geometric parameters for this dataset, so I created a nominal projection scan using the nominal SID, assuming an isocentric scan, estimating the angular range visually, and using equally-spaced rotations.

I compared results from a static OSEM reconstruction to the joint reconstruction and geometry estimation method. These reconstruction results are in **Figure 2.9**. The proposed method eliminates miscalibration ghosting artifacts that exist in the static reconstruction.

The knee dataset consists of 274 756^2 projection images with an approximately 274 degree scan. The skull phantom dataset consists of 342 600^2 projection images with an approximately 196 degree scan. I performed the final reconstruction on a 512^3 grid. Both of the static reconstructions took approximately 10 minutes, and the reconstructions with geometry estimation took approximately 16 minutes.

Figure 2.9. Knee (left two columns) and skull (right two columns) physical phantoms reconstruction results. The top row consists of OSEM reconstruction without any geometry estimation. Both reconstructions exhibit ghosting artifacts where a structure appears twice in the reconstruction due to miscalibration. In particular, double edges exist on the front of the femur and back of the tibia in the knee phantom and around the skull boundary and the vertebrae on the skull phantom. The bottom row shows the OSEM reconstruction with geometry estimation. In these reconstructions, the ghosting artifacts are eliminated.

2.3.2.2 Cadaver Dataset

The final dataset was acquired from a developmental full-size mobile C-arm with a flat panel detector. An approximately 190 degree scan was acquired around a cadaver hand. Similar to the first dataset, an electromagnetic sensor was placed on the detector, giving an approximation of the geometric parameters.

I performed three reconstructions: static OSEM using the parameter values from the sensors, static OSEM given nominal parameters, and the proposed joint reconstruction method given the nominal parameters. The nominal parameters were created in the same way as the first dataset: a circular trajectory with evenly spaced rotations was used and the initial estimates of $T_{x'}$ and $T_{y'}$ were pre-estimated. Once again, the proposed method produced reconstructed images that are of similar quality to reconstruction given the geometric parameters from the sensors. These results can be seen in **Figure 2.10**.

This acquisition scan consisted of 179 1536^2 projection images spaced at one degree. The static reconstruction took approximately 15 minutes and the reconstruction with geometry estimation took 24 minutes.

2.4 Discussion

The joint image and geometry estimation problem is certainly not a convex problem. As such, it is always a concern for the estimation to converge in a local minimum. This issue is mitigated by estimating the geometric parameters at lower resolutions in the multiscale scheme, which helps the method find global mismatch first and then converge to smaller detailed mismatch. Furthermore, the most computationally expensive part of the algorithm is computing the line integrals of the gradient image and performing the 2D convolution. I typically then perform a static reconstruction at the final high-resolution scale and perform geometry estimation at the lower-resolution scales. For these two reasons, the multiscale formulation is essential to both the convergence and speed of the proposed method. However, global convergence is not guaranteed, and a somewhat reasonable initial estimate of the geometric parameters is required.

Although the geometry of a cone-beam system at a point in time is uniquely described by nine parameters, small variations in intrinsic parameters can be approximated by a change in the extrinsic parameters. For example, a small positive change in the piercing

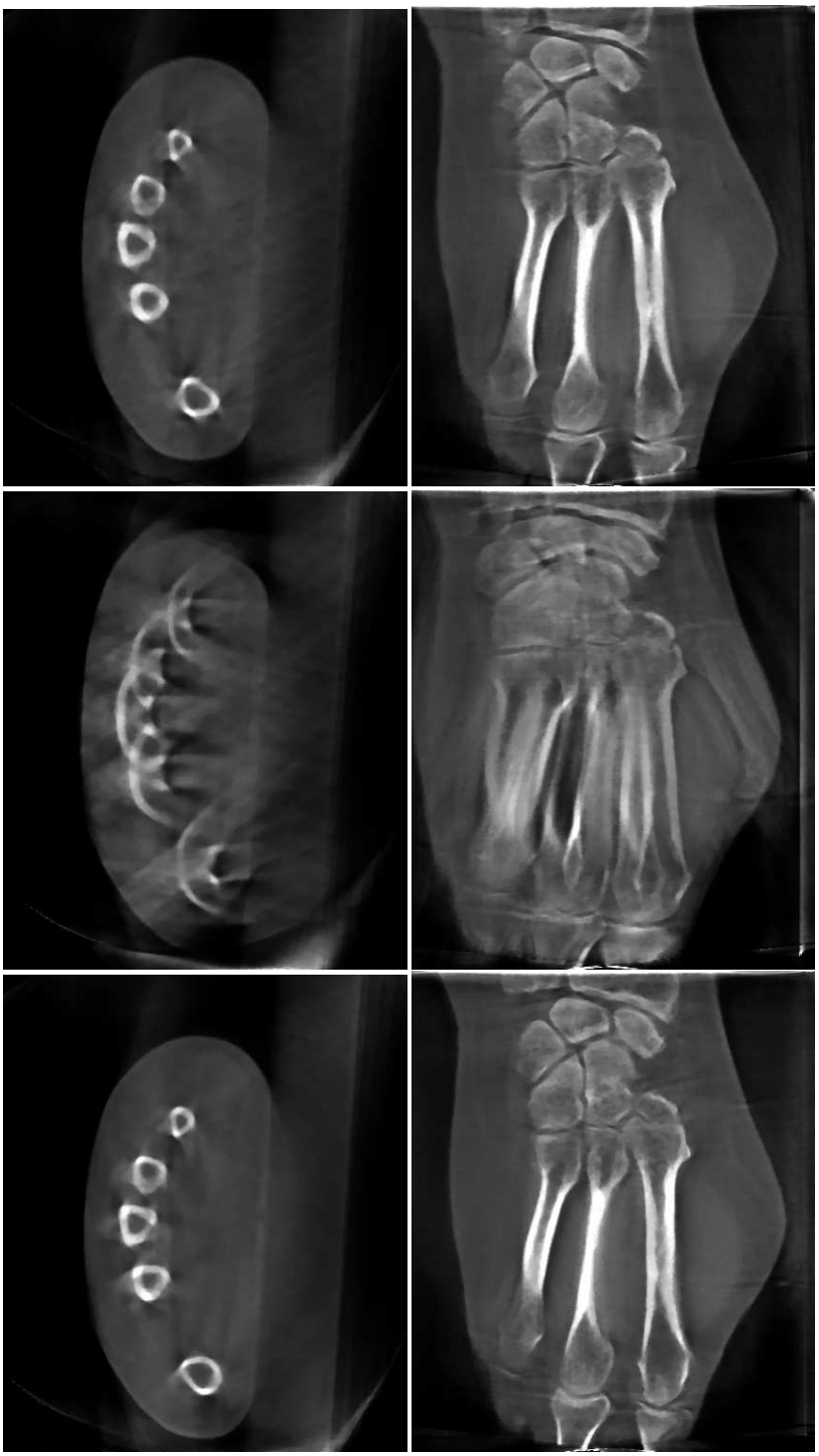


Figure 2.10. Hand cadaver reconstruction results. Top row: OSEM reconstruction with electromagnetic sensor data without geometry estimation. Middle row: OSEM reconstruction with nominal trajectory without geometry estimation. Bottom row: joint image and geometry estimation given the nominal trajectory.

point u_0 can be approximated by a small clockwise rotation around the y' axis and a small translation in the $-x'$ direction (as they are defined in **Figure 2.1**). Similarly, a small increase in the SID can be approximated by a small translation in the $+z'$ direction. Since these are nearly ambiguous, I optimize only rotation and translation. This has the benefit of being more computationally efficient. In my experiments, I found no perceptible drop in image quality compared to estimating all nine parameters. Additionally, a strong prior could be placed on the intrinsic parameters.

2.5 Conclusion

In this chapter, I have shown that I can accurately reconstruct a 3D volume using data from a mobile C-arm. My proposed method of jointly estimating the geometry and the image produces much improved results over the reconstruction using nominal parameters, making 3D reconstruction on a C-arm viable. I have implemented this method on the GPU. Since the geometry estimation was run at the lowest resolution, there is only a 50 percent increase in computation time compared to static OSEM reconstruction.

2.6 Acknowledgments

I would like to thank GE Healthcare for making this research possible. I would also like to thank Matt Simmons, Lance McBride, and Ganeshkumar M. R. who built the C-arm test system, and Arvidas Cheryauka for acquiring the data.

2.7 References

[1] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *Journal of the Optical Society of America A*, vol. 1, no. 6, pp. 612–619, 1984.

[2] F. Noo, R. Clackdoyle, C. Mennessier, T. A. White, and T. J. Roney, "Analytic method based on identification of ellipse parameters for scanner calibration in cone-beam tomography," *Physics in Medicine and Biology*, vol. 45, no. 11, pp. 3489–3508, 2000.

[3] F. Zhang, J. Du, H. Jiang, L. Li, M. Guan, and B. Yan, "Iterative geometric calibration in circular cone-beam computed tomography," *Optik – International Journal for Light and Electron Optics*, vol. 125, no. 11, pp. 2509–2514, 2014.

[4] N. K. Strobel, B. Heigl, T. M. Brunner, O. Schütz, M. M. Mitschke, K. Wiesent, and T. Mertelmeier, "Improving 3D image quality of X-ray C-arm imaging systems by using properly designed pose determination systems for calibrating the projection geometry," in *Proc. SPIE* (M. J. Yaffe and L. E. Antonuk, eds.), vol. 5030, pp. 943–954, 2003.

[5] Y. Cho, D. J. Moseley, J. H. Siewerdsen, and D. A. Jaffray, "Accurate technique for complete geometric calibration of cone-beam computed tomography systems," *Medical Physics*, vol. 32, no. 4, pp. 968–983, 2005.

[6] V. Patel, R. N. Chityala, K. R. Hoffmann, C. N. Ionita, D. R. Bednarek, and S. Rudin, "Self-calibration of a cone-beam micro-CT system," *Medical Physics*, vol. 36, no. 1, pp. 48–58, 2009.

[7] D. Panetta, N. Belcari, A. Del Guerra, and S. Moehrs, "An optimization-based method for geometrical calibration in cone-beam CT without dedicated phantoms," *Physics in Medicine and Biology*, vol. 53, no. 14, pp. 3841–3861, 2008.

[8] Y. Kyriakou, R. M. Lapp, L. Hillebrand, D. Ertel, and W. A. Kalender, "Image-based online correction of misalignment artifacts in cone-beam CT," in *Proc. SPIE* (E. Samei and J. Hsieh, eds.), vol. 7258, pp. 72581V–72581V–10, 2009.

[9] J. Muders and J. Hesser, "Stable and robust geometric self-calibration for cone-beam CT using mutual information," *IEEE Transactions on Nuclear Science*, vol. 61, no. 1, pp. 202–217, 2014.

[10] W. Wein, A. Ladikos, and A. Baumgartner, "Self-calibration of geometric and radiometric parameters for cone-beam computed tomography," in *11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 324–327, 2011.

[11] B. De Man and J. A. Fessler, "Statistical iterative reconstruction for X-ray computed tomography," in *Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning, and Inverse Problems* (Y. Censor, Jiang Ming, and G. Wang, eds.), pp. 1–28, Medical Physics Publishing, 2009.

[12] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *Journal of Theoretical Biology*, vol. 29, no. 3, pp. 471–481, 1970.

[13] Y. Censor, P. P. B. Eggermont, and D. Gordon, "Strong underrelaxation in Kaczmarz's method for inconsistent systems," *Numerische Mathematik*, vol. 41, no. 1, pp. 83–92, 1983.

[14] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, p. 23, 1974.

[15] J. Gregor and T. Benson, "Computational analysis and improvement of SIRT," *IEEE Transactions on Medical Imaging*, vol. 27, no. 7, pp. 918–924, 2008.

[16] A. C. Kak and M. Slaney, "Algebraic reconstruction algorithms," in *Principles of Computerized Tomographic Imaging*, ch. 7, pp. 275–296, Society for Industrial and Applied Mathematics, 2001.

[17] M. Jiang and G. Wang, "Convergence of the simultaneous algebraic reconstruction technique (SART)," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 957–961, 2003.

[18] A. van der Sluis and H. A. van der Vorst, "SIRT and CG-type methods for the iterative solution of sparse linear least-squares problems," *Linear Algebra and Its Applications*, vol. 130, pp. 257–303, 1990.

[19] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm," *Ultrasonic Imaging*, vol. 6, no. 1, pp. 81–94, 1984.

[20] W. Xu and K. Mueller, "A performance-driven study of regularization methods for GPU-accelerated iterative CT," in *Workshop on High Performance Image Reconstruction*, vol. 2, 2009.

[21] G. Wang and M. Jiang, "Ordered-subset simultaneous algebraic reconstruction techniques (OS-SART)," *Journal of X-ray Science and Technology*, vol. 12, pp. 169–177, 2004.

[22] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 113–122, 1982.

[23] R. L. White, "Image restoration using the damped Richardson-Lucy method," *The Restoration of HST Images and Spectra II*, pp. 104–110, 1994.

[24] S. Joshi and M. I. Miller, "Maximum a posteriori estimation with Good's roughness for three-dimensional optical-sectioning microscopy," *Journal of the Optical Society of America A*, vol. 10, no. 5, pp. 1078–1085, 1993.

[25] J. Nuyts, B. De Man, J. A. Fessler, W. Zbijewski, and F. J. Beekman, "Modelling the physics in iterative reconstruction for transmission computed tomography," *Physics in Medicine and Biology*, vol. 58, no. 12, pp. R63–R96, 2013.

[26] V. Y. Panin, G. L. Zeng, and G. T. Gullberg, "Total variation regulated EM algorithm," *IEEE Transactions on Nuclear Science*, vol. 46, no. 6, pp. 2202–2210, 1999.

[27] J. A. O'Sullivan and J. Benac, "Alternating minimization algorithms for transmission tomography," *IEEE Transactions on Medical Imaging*, vol. 26, no. 3, pp. 283–297, 2007.

[28] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *Journal of Computer Assisted Tomography*, vol. 8, no. 2, pp. 306–316, 1984.

[29] S. H. Manglos, G. M. Gagne, A. Krol, F. D. Thomas, and R. Narayanaswamy, "Transmission maximum-likelihood reconstruction with ordered subsets for cone beam CT," *Physics in Medicine and Biology*, vol. 40, no. 7, pp. 1225–1241, 1995.

[30] J. M. Ollinger, "Maximum-likelihood reconstruction of transmission images in emission computed tomography via the EM algorithm," *IEEE Transactions on Medical Imaging*, vol. 13, no. 1, pp. 89–101, 1994.

[31] S. Ahn and J. A. Fessler, "Globally convergent ordered subsets algorithms: Application to tomography," in *2001 IEEE Nuclear Science Symposium Conference Record*, vol. 2, pp. 1064–1068, 2001.

[32] H. Erdogan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Physics in Medicine and Biology*, vol. 44, no. 11, pp. 1064–1068, 1999.

[33] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Transactions on Medical Imaging*, vol. 9, no. 1, pp. 84–93, 1990.

[34] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[35] G. T. Herman and L. B. Meyer, "Algebraic reconstruction techniques can be made computationally efficient," *IEEE Transactions on Medical Imaging*, vol. 12, no. 3, pp. 600–609, 1993.

[36] D. Kim, S. Ramani, and J. A. Fessler, "Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 34, no. 1, pp. 167–178, 2015.

[37] G. L. Zeng and G. T. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," *IEEE Transactions on Medical Imaging*, vol. 19, no. 5, pp. 548–555, 2000.

[38] S. J. Glick and E. J. Soares, "Noise characteristics of MLEM SPECT reconstruction with a mis-matched projector-backprojector pair," *IEEE Transactions on Nuclear Science*, vol. 45, no. 4, pp. 2183–2188, 1998.

[39] A. Cheryauka, S. Breham, and W. Christensen, "Sequential intrinsic and extrinsic geometry calibration in fluoro CT imaging with a mobile C-arm," in *Proc. SPIE* (K. R. Cleary and R. L. Galloway, Jr., eds.), vol. 6141, pp. 61412H–61412H–8, 2006.

CHAPTER 3

CT IMAGES AS DENSITIES

In this chapter, I explore the relationship between CT images and densities. This connection is essential for my work in the following chapter where I introduce a density registration method to apply to thoracic CT images. This chapter will cover three aspects of this relationship. First, I will discuss why one would theoretically expect an image of linear attenuation coefficients to act like a density. Second, I will show that in practice, a CT image does not act exactly as a density. Finally, I will introduce a method for transforming CT images to act as densities, allowing them to be used in the density matching framework in the following chapter.

3.1 Mass Density and Linear Attenuation Coefficients

3.1.1 Intensity Changes in Thoracic CT

The motivation for this work is to be able to accurately account for intensity changes that occur in thoracic CT images during respiration. These intensity changes can be seen in a single subject CT dataset at different phase points of the breathing cycle. During inhalation, lung volume increases and the CT intensities inside the lung decrease. During exhalation, lung volume decreases and lung CT intensities increase. This effect can be seen in **Figure 3.1**. This relationship is also made clear by comparing the histograms of image intensities in the peak inhale and peak exhale images, seen in **Figure 3.2**. In both figures, I used a rat thoracic CT dataset composed of eleven phase points of an inhale-exhale breathing cycle. These data were acquired by placing a rat on a ventilator and acquiring data gated to the ventilator phase [1, 2].

3.1.2 Densities and Conservation of Mass

I will first introduce the mathematical definition of a density and then describe how it relates to material density and CT images. Mathematically, a 3D density $I(x)\,dx$ is a

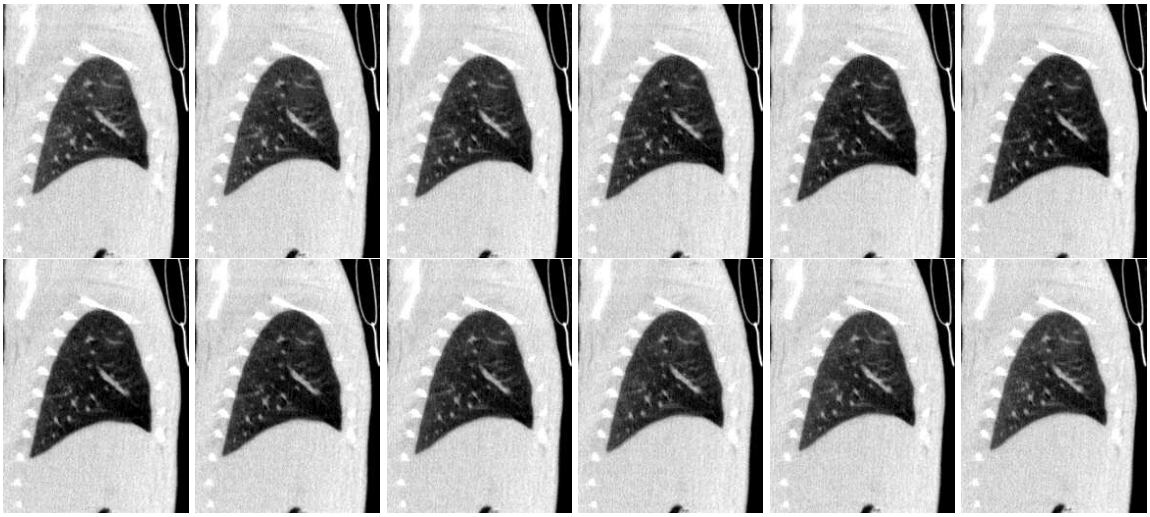


Figure 3.1. Rat breathing cycle. Top row: from left to right, inhale phase of the breathing cycle. Bottom row: from left to right, exhale phase of the breathing cycle. During inhalation, the lung volume increases and the CT intensities decrease. During exhalation, the lung volume decreases and the CT intensities increase.

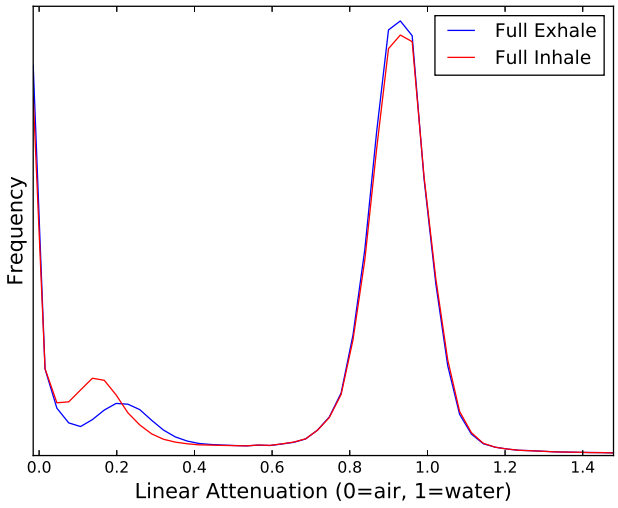


Figure 3.2. Voxel intensity histograms in rat thoracic CT. Here I show the CT histograms for the peak inhale and the peak exhale for the rat dataset. Each histogram has three peaks: the peak at 0.0 represents surrounding air, the middle peak represents lung tissue, and the peak at 1.0 represents soft tissue. For the lung tissue, the full inhale has higher volume but a lower image intensity than the full exhale, therefore showing some conservation of mass. For the soft tissue outside the lungs, the average intensity does not change because it is incompressible. The slight drop in frequency of the full inhale is due to soft tissue leaving the reconstruction boundary.

volume form on a domain $\Omega \subseteq \mathbb{R}^3$, where $I(x)$ is a nonnegative function on Ω and $dx = dx^1 \wedge dx^2 \wedge dx^3$ is the standard volume element on $\mathbb{R}^3$. The key distinction between a density and a function is how they deform. Given a deformation φ, a *function* $g(x)$ deforms via function composition (also called the L^2 image action):

$$\varphi_* g(x) = g \circ \varphi^{-1}(x). \tag{3.1}$$

In contrast, a *density* deforms via the density action:

$$\varphi_*(I(x)\,dx) = I \circ \varphi^{-1}(x)|D\varphi^{-1}(x)|\,dx, \tag{3.2}$$

where $|D\varphi^{-1}|$ is the Jacobian determinant of the diffeomorphism. The Jacobian determinant describes the volume change of φ. If φ is volume preserving, the Jacobian determinant is unitary. If φ increases (decreases) volume locally, then its value is greater than (less than) one. Therefore, the density action increases or decreases the intensities due to volume change.

A unique property of a density is that the total mass is conserved under the action of a deformation, where the total mass is defined as the integral of the density over Ω:

$$\int_\Omega \varphi_* I(x)\,dx = \int_\Omega I(x)\,dx. \tag{3.3}$$

This equality holds by performing a change of variables and is proven in the following chapter.

A probability density function (PDF) is a special case of this mathematical definition of a density. A PDF is nonnegative function of a continuous random variable, and the integral over the entire domain is always equal to one. In this dissertation, I consider only densities that exist on $\Omega \subseteq \mathbb{R}^3$, though the properties extend to other dimensions.

Perhaps the most common example of a density is a physical mass density $\rho(x)$, with units $\mathrm{g/cm^3}$. Physical mass density integrated over a closed system becomes the total physical mass (units g). The narrow beam X-ray linear attenuation coefficient (LAC) for a material is dependent on this physical mass density of the material:

$$\mu(x) = (\mu/\rho)_m \rho(x). \tag{3.4}$$

Here $(\mu/\rho)_m$ is a material specific property called the mass attenuation coefficient (units $\mathrm{cm^2/g}$). The linear attenuation coefficient describes the X-ray's intensity decrease per unit

distance as an X-ray beam passes through a region (units cm^{-1}). The X-ray mass attenuation coefficient is dependent on the X-ray energy. These mass attenuation coefficient values for specific materials at various energy levels have been experimentally tabulated by the National Institute of Standards and Technology (NIST) [3]. Like physical mass density, I define the LAC mass as the integral of the linear attenuation coefficient over a domain. As with physical mass densities, I am integrating over a 3D domain, so LAC mass has units cm^2. A summary of these different mass and density examples is found in **Table 3.1**.

Given a mixture of materials, the total linear attenuation coefficient is a combination of each material's mass attenuation coefficient weighted by the material's relative mass density:

$$\mu(x) = \sum_i (\mu/\rho)_i \rho_i(x). \tag{3.5}$$

Since these combine linearly, it can be shown that if we have conservation of physical mass in a closed system, we will also have conservation of LAC mass:

$$\begin{aligned}
\int_\Omega \varphi_* \mu(x)\, dx &= \int_\Omega \varphi_* \sum_i (\mu/\rho)_i \rho_i(x)\, dx \\
&= \sum_i (\mu/\rho)_i \int_\Omega \varphi_* \rho_i(x)\, dx \\
&= \sum_i (\mu/\rho)_i \int_\Omega \rho_i(x)\, dx \\
&= \int_\Omega \mu(x)\, dx.
\end{aligned} \tag{3.6}$$

Therefore, in a closed system, we have conservation of LAC mass, even with a mixture of materials.

Table 3.1. Density examples

	density	volume	mass
PDF	$f(x)$ (unitless)	$\int_A dx$ (unknown)	$\int_A f(x)dx = 1$ (unitless)
mass density	$\rho(x)$ (g/cm^3)	$\int_\Omega dx$ (cm^3)	$\int_\Omega \rho(x)dx$ (g)
LAC	$\mu(x)$ (cm^{-1})	$\int_\Omega dx$ (cm^3)	$\int_\Omega \mu(x)dx$ (cm^2)

3.2 Mass Density and Effective Linear Attenuation Coefficients

In the previous section, I showed that we expect conservation of LAC mass. However, this LAC is defined as the monoenergetic narrow beam linear attenuation coefficient. Instead, modern CT scanners use wide beams that yield secondary photon effects at the detector (mostly due to Compton scattering). CT image intensities reflect the *effective* linear attenuation coefficients as opposed to the true narrow beam linear attenuation coefficient. In this section, I will show that in modern CT systems, we do not have conservation of effective linear attenuation coefficient mass, which I will call CT mass.

In this section, I specifically look at the conservation of mass properties inside the lungs for two reasons. First, the field of view of the CT scanner is not a closed system, as portions of the abdomen leave and enter the field of view. The lungs are not technically a closed physical system as air enters and leaves the lungs, and oxygen and carbon dioxide is exchanged in the blood. However, the linear attenuation coefficient of air is essentially zero, and I assume the change in blood volume is negligible. Therefore, changes due to respiration should not change the total LAC mass inside the lungs. The second reason is that the majority of volume change occurs inside the lungs. Much of the anatomy surrounding the lungs is fluid-filled, and is therefore essentially incompressible. The CT intensity of this surrounding tissue remains constant throughout the breathing cycle, as can be seen in **Figure 3.2**.

First, I show experimental results which indicate that CT mass in the lungs is not conserved. In the rat dataset, I plot the average density, lung volume, and lung CT mass. This can be seen in **Figure 3.3**. Here, CT mass is clearly not conserved, as it is dependent on the average density and volume.

To see the relationship between effective LAC and true narrow beam LAC, a Monte Carlo simulation was run using the X-ray spectrum and geometry from a Philips CT scanner at various densities of water [4], since lung tissue is very similar to a mixture between water and air. In these results, the nonlinear relationship between effective LAC and narrow beam LAC is confirmed (see **Figure 3.4**).

Given conservation of mass within a single subject in a closed system, we expect an inverse relationship between average density in a region Ω and volume of that region:

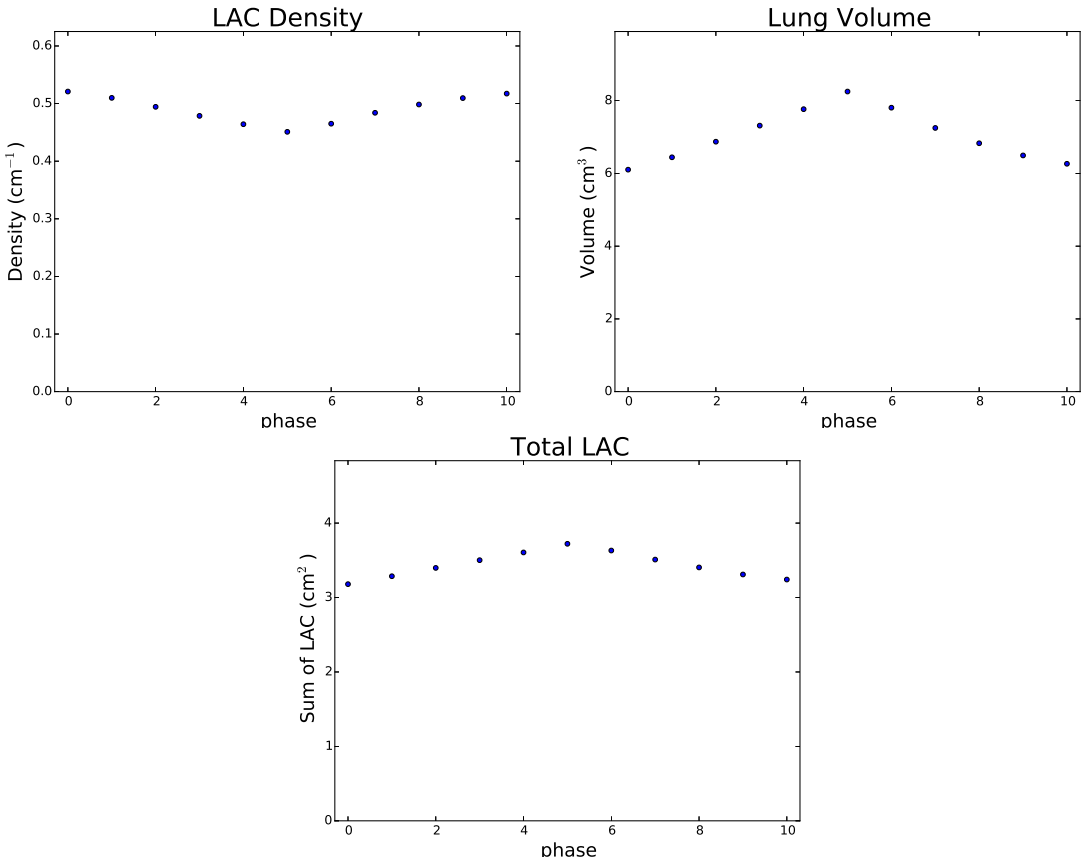


Figure 3.3. Density, volume, and CT mass at each phase of the rat breathing cycle. Using the given value for density, volume, and CT mass, it can be seen that CT mass is dependent on the average density and volume. Therefore, the nominal CT mass is not conserved.

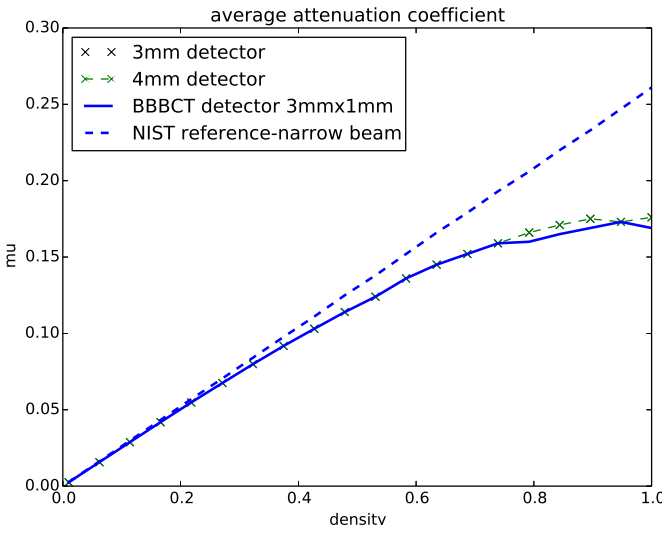


Figure 3.4. Effective LAC from Monte Carlo simulation (solid line) and NIST reference narrow beam LAC (dashed line). The true relationship between effective LAC and narrow beam LAC is nonlinear.

$$D_t = \frac{M}{V_t} \tag{3.7}$$

$$= \frac{M}{\int_{\Omega_t} 1 dx}. \tag{3.8}$$

Here Ω_t is the domain of the closed system (that moves over time) and t is a phase point of the breathing cycle. This relationship between density and volume becomes linear in log space with a slope of -1:

$$\ln(D_t) = \ln(M) - \ln(V_t). \tag{3.9}$$

These experimental results confirm the Monte Carlo simulation in that lungs imaged under CT do not follow this inverse relationship. Rather, the slope found in these datasets in log space is consistently greater (less negative) than -1 (see **Figure 3.5**). Because of this, I seek a nonlinear intensity transformation for CT images such that CT mass in the lungs is conserved.

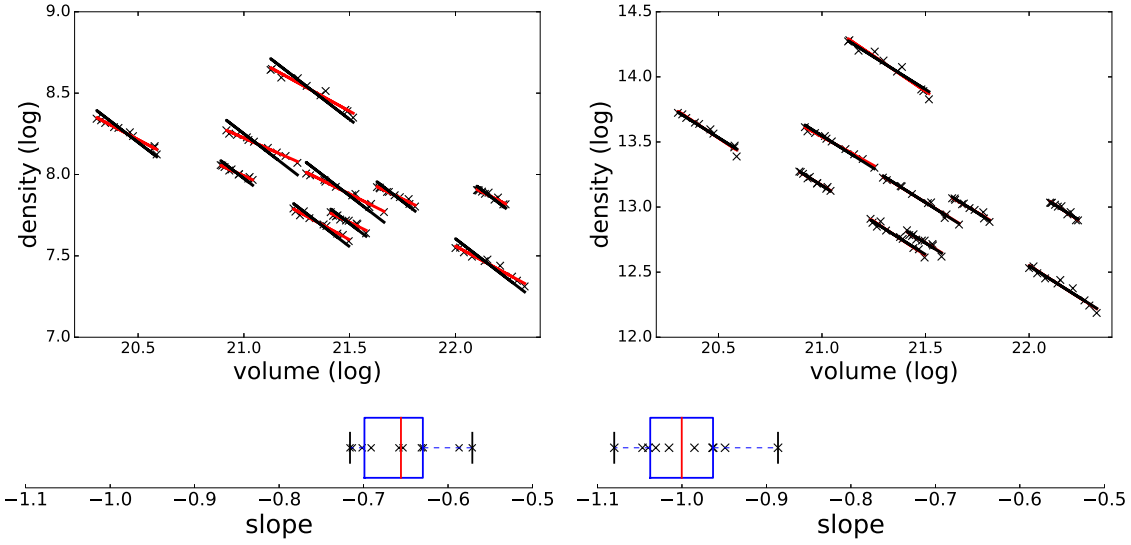


Figure 3.5. Density and volume log-log plots. Upper left: log-log plots without applying the power correction for all 10 DIR subjects. The best fit line to each dataset is in red and the mass-preserving line (slope = -1) is in black. Upper right: log-log plots after applying the power correction $I(x)^\alpha$ to the CT images. In this plot, the best fit line closely matches the mass-preserving line. Bottom row: corresponding box plots of the slopes found in the regression.

3.3 Conservation of Mass in CT Images

In this chapter, I model this intensity transformation as a power function, i.e., $I(x) \mapsto I(x)^{\alpha}$, and solve for the α that yields the best mass preservation. A power model was chosen because it is a single parameter monotonic function that preserves the density of air at zero. Furthermore, using a power function makes the analysis invariant to image scaling. That is, if $I(x)^{\alpha}$ exhibits conservation of mass, so does $(cI(x))^{\alpha}$, for $c \in \mathbb{R}^{+}$.

I evaluated this power model using images from the Deformable Image Registration (DIR) Laboratory dataset (`http://www.dir-lab.com/`) [5], which consists of 10 subjects with 10 phase images each. I also evaluated this model on a set of 18 subjects with 10 phase points each procured at University of Texas Southwestern Medical Center (UTSW). The lungs for each patient at each timepoint were segmented with active contours using ITK-SNAP [6] (`http://www.itksnap.org/`) combined with an intensity-based segmentation to remove high-density regions in the lungs and around the lung border due to imperfect initial segmentations.

For each subject, I performed a linear regression of the measured LAC density and calculated volume in log space. Let $\boldsymbol{d}(\alpha) = \log\left(\int_{\Omega_t} I_t(x)^{\alpha} dx / \int_{\Omega_t} 1 dx\right)$ (the log density) and $\boldsymbol{v} = \log(\int_{\Omega_t} 1 dx)$ (the log volume), where again t is a breathing cycle timepoint. The linear regression then models the relationship in log space as $\boldsymbol{d}(\alpha) \approx a\boldsymbol{v} + b$. Let $a_j(\alpha)$ be the slope solved for in this linear regression for the jth subject. To find the optimal α for the entire dataset, I solve

$$\alpha = \arg\min_{\alpha'} \sum_{j} (a_j(\alpha') + 1)^2, \tag{3.10}$$

which finds the value of α that gives an average slope closest to -1. I solved for α using a brute force search. Applying this power function to the CT data allows use of the density matching algorithm described in the next chapter.

3.4 Results

For the DIR dataset, I used the proposed method to solve for the exponent that yields conservation of mass. For this dataset, $\alpha = 1.64$ minimizes Equation 3.10. Without using the power fit, the average slope of the log-density log-volume plot was -0.66 ($\sigma = 0.048$, 95% confidence interval $[-0.69, -0.62]$). After applying the power function to the CT

intensities, the average slope is -1.0 ($\sigma = 0.054$). The log-log plots of all 10 patients in the DIR dataset as well as box plots of the slope are shown in **Figure 3.5**.

For the UTSW dataset, $\alpha = 1.90$ minimized Equation 3.10. Without using the power fit, the average slope of the log-log plot was -0.59 ($\sigma = 0.11$, 95% confidence interval $[-0.64, -0.53]$).

The difference between the intensity change of the given lung data and true mass preserving deformations was found to be statistically significant ($p \ll 0.001$ for the DIR dataset and $p \ll 0.001$ for the UTSW dataset). However, the differences between the mass-preserving nature of these two datasets were not statistically significant ($p = 0.11$). For these three statistical tests, I corrected for multiple comparisons using the Holm-Bonferroni method.

3.5 Discussion

In this chapter, I have shown that although the narrow beam linear attenuation acts as a density, the effective linear attenuation coefficient found in CT does not act as a density. However, applying a power function transforms the CT dataset into a set of images that exhibits conservation of mass. This simple nonlinear approximation yields excellent results even when using the same power function for multiple subjects in a single dataset.

I suspect that the biggest cause of the nonlinearity between true linear attenuation and effective attenuation is the presence of X-ray scatter, which creates secondary photons at the detector. The amount of X-ray scatter depends on the scanner geometry and the energy spectrum. For this reason, I do not necessarily expect that the same α parameter of the power functions to accurately correct for this across different CT scanners. However, I found that the same α parameter accurately corrects the nonlinearity across multiple subjects on the same scanner.

3.6 References

[1] R. E. Jacob and W. J. Lamm, "Stable small animal ventilation for dynamic lung imaging to support computational fluid dynamics models," *PLoS ONE*, vol. 6, no. 11, pp. 1–6, 2011.

[2] R. E. Jacob, J. P. Carson, M. Thomas, and D. R. Einstein, "Dynamic multiscale boundary

conditions for 4D CT of healthy and emphysematous rats," *PLoS ONE*, vol. 8, no. 6, pp. 1–9, 2013.

[3] NIST, "X-ray mass attenuation coefficients - water, liquid." `http://physics.nist.gov/PhysRefData/XrayMassCoef/ComTab/water.html`. Accessed: 2016-03-06.

[4] J. M. Boone and J. A. Seibert, "An accurate method for computer-generating tungsten anode X-ray spectra from 30 to 140 kv," *Medical Physics*, vol. 24, no. 11, pp. 1661–1670, 1997.

[5] R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A. K. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Physics in Medicine and Biology*, vol. 54, no. 7, pp. 1849–1870, 2009.

[6] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

CHAPTER 4

WEIGHTED DIFFEOMORPHIC DENSITY MATCHING WITH APPLICATIONS TO THORACIC IMAGE REGISTRATION

In Chapter 2, I explored the problem of uncertain motion in the imaging system while assuming the object being imaged is stationary. In this chapter, I explore the opposite scenario, where the motion of the imaging system is well known and the patient being imaged is undergoing normal physiological motion.

In this chapter, I consider the problem of tracking organs undergoing deformations as a result of breathing in the thorax and imaged via computed tomography (CT). This problem has wide-scale medical applications, in particular, radiation therapy of the lung. In this application, accurate estimation of organ deformations during treatment impacts dose calculation and treatment decisions [1, 2, 3, 4]. The current state-of-the-art radiation treatment planning involves the acquisition of a series of respiratory correlated CT (RCCT) images to build 4D (three spatial and one temporal) treatment planning datasets. Fundamental to the processing and clinical use of these 4D datasets is the accurate estimation of registration maps that characterize the motion of organs at risk as well as the target tumor volumes.

According to the Centers for Disease Control and Prevention, lung cancer is the third most common form of cancer, yet is by far the leading cause of cancer death, accounting for approximately 27 percent of all cancer related deaths in the United States [5]. Thoracic CT imaging is a valuable noninvasive technique for diagnosis and treatment planning of lung cancer. In this chapter, I introduce a novel method for estimating these registration maps. The proposed method uses the intimate link between densities and diffeomorphisms to create an efficient density matching algorithm. The proposed method includes a natural weighting penalty on volume change that models the varying compressibility properties

of different tissues. It restricts compression and expansion in incompressible areas of the body (such as bones and soft tissue surrounding the lungs) while allowing the lungs to expand and contract during respiration.

4.1 Introduction

Currently, the application of diffeomorphisms in medical image registration is mostly limited to the L^2 image action of the diffeomorphism group, which is not a mass-preserving transformation. Furthermore, the diffeomorphisms estimated from typical image registrations algorithms (such as LDDMM [6] or ANTs [7]) do not accurately model the varying compressibility of different tissue types. In thoracic datasets, the lungs are highly compressible. Conversely, the bronchial tubes and the tissue surrounding the lungs are nearly incompressible. During inhalation, as air enters, the lung volume increases and the lung density decreases, whereas during exhalation, the lung volume decreases and the lung density increases. But in both inhale and exhale, the lung mass is conserved.

In 2010, the EMPIRE10 [8] challenge compared the accuracy of many registration algorithms applied to intrapatient thoracic CT images. The winner of the competition, Advanced Normalization Tools (ANTs) [9, 10], used an LDDMM method with a normalized cross-correlation metric. The second place winner was NiftyReg [11, 12], a registration package using a cubic B-splines deformation model with a least trimmed square metric. Both registration methods are open-source tools and are considered to be state-of-the-art for registration accuracy. B-spline deformations are guaranteed to be smooth, but they do not guarantee diffeomorphic properties (such as invertibility). However, neither method models conservation of mass or spatially varying tissue compressibility.

More recently, some mass-preserving registration methods [13, 14] have been introduced. Both these methods use cubic B-splines, which do not have the flexibility of standard diffeomorphic models. These methods also do not model spatially varying tissue compressibility.

I present an image registration technique that incorporates conservation of mass and models organ compressibility. Instead of the L^2 image action of diffeomorphisms, I use the physically appropriate density action. I also regularize the diffeomorphism by using a space-varying penalty that allows for high compressibility of the lung tissue and at the

same time enforces incompressibility of high-density structures such as bone and soft tissue. The algorithm is based on the intimate link between the Riemannian geometry of the space of diffeomorphisms and the space of densities [15, 16, 17]. The resulting algorithm also has the added advantage that it is computationally efficient: it is orders of magnitude faster than existing diffeomorphic image registration algorithms.

4.1.1 Diffeomorphic Image Registration

The fundamental problem of image registration is as follows: given two images I_0 and I_1, find a realistic change of coordinates that brings these two images into correspondence. Attempting to solve this problem is often done by defining an energy functional that is the combination of two separate penalty functionals. The functional that quantifies the accuracy of the deformation (the degree of correspondence) is called the *data match* term $\bar{d}$. The functional that describes the prior probability of the deformation is called the *regularization* term d. The goal is to then solve for the ideal transformation that minimizes the total energy functional

$$E(T) = \sigma d^2(T, \mathrm{id}) + \bar{d}^2(T_* I_0, I_1)\,, \tag{4.1}$$

where T is some transformation acting on I_0 and σ is a scalar weighting that balances the relative strength of the two penalties. Essential to solving this problem is finding a suitable penalty that describes the data match, finding a suitable penalty for the regularization (along with the suitable class of transformations), and finding an accurate and efficient model to solve the minimization problem.

Recently, a great deal of interest has been focused on a special class of transformations called *diffeomorphisms*. A diffeomorphism φ is a bijective (one-to-one and onto) mapping from $\Omega \subseteq \mathbb{R}^d$ to $\Omega \subseteq \mathbb{R}^d$, where both φ and φ^{-1} are differentiable. In this chapter, I assume $d = 3$. The composition of any two diffeomorphisms is a diffeomorphism, and the identity transformation id is also a diffeomorphism. Therefore, the set of all diffeomorphisms forms a Lie group denoted $\mathrm{Diff}(\Omega)$.

A path in the group of diffeomorphisms is parameterized by $t \in [0, 1]$ and is written as $\varphi_t(x) = \varphi(x, t)$. This time-dependent diffeomorphism is defined by a time-dependent vector field $v_t(x) = v(x, t)$ subject to the O.D.E.

$$\frac{d}{dt}\varphi_t(x) = v_t \circ \varphi_t(x), \tag{4.2}$$

with suitable boundary conditions for $\varphi_0(x)$ and $\varphi_1(x)$. While the group of diffeomorphisms is not a vector space, its tangent space is. The Lie algebra $T_\varphi \operatorname{Diff}(\Omega)$ is the space of smooth vector fields. Since the time-varying velocity field must be smooth, a natural inner product space is the H^1 Hilbert space, where the V inner product is defined by

$$\langle v(x), u(x) \rangle_V = \langle Lv(x), u(x) \rangle_{L^2}. \tag{4.3}$$

Here L is some differential operator. In this chapter, I define the H^1 metric using the Hodge Laplacian on vector fields. Due to its connections to information geometry, I also refer to this metric as the *information metric*:

$$G^I(u, v) = \int_\Omega \langle -\Delta u, v \rangle \, dx. \tag{4.4}$$

Given the information metric, the distance metric in the diffeomorphism group is defined by the length of the minimum path $\varphi_t \in \operatorname{Diff}(\Omega)$:

$$d(\psi, \phi)^2 = \inf \left\{ \int_0^1 G^I(v_t, v_t) \, dt \quad \text{s.t.} \quad \varphi_0 = \psi, \varphi_1 = \phi \right\}. \tag{4.5}$$

This distance metric gives diffeomorphisms several ideal properties. First, this distance metric is right invariant: $d(\psi, \phi) = d(\psi \circ \varphi, \phi \circ \varphi)$ for all $\varphi \in \operatorname{Diff}(\Omega)$. From this, we can see that the lengths of a forward and inverse deformation are equal: $d(\mathrm{id}, \varphi) = d(\varphi^{-1}, \mathrm{id})$. Diffeomorphisms in image registration have been used extensively. For more details, see [18, 6, 19, 20]. The most common implementation of diffeomorphic image registration is LDDMM. In order to calculate the metric of the diffeomorphism, Equation 4.5 is discretized in time, so the computational complexity is dependent on the number of time points used.

4.2 Mathematical Formulation

Mathematically, the problem is to find a diffeomorphic transformation between two densities on a subset $\Omega \subseteq \mathbb{R}^3$. With a "density," I mean a volume form on Ω, i.e., an element of the form $\mu = I \, dx$ where $dx = dx^1 \wedge dx^2 \wedge dx^3$ is the standard volume element on $\mathbb{R}^3$ and $I = I(x)$ is a nonnegative function on Ω. The space of all densities on Ω is denoted $\operatorname{Dens}(\Omega)$. Of course, one might identify $I \, dx$ with its function I, and thereby think of $\operatorname{Dens}(\Omega)$ as the set of nonnegative functions on Ω. However, the invariance properties

and geometry of the problem are remarkably more transparent when viewing $\text{Dens}(\Omega)$ as a space of volume forms.

The group of diffeomorphisms $\text{Diff}(\Omega)$ acts from the right on $\text{Dens}(\Omega)$ by *pullback:* the right action of $\varphi \in \text{Diff}(\Omega)$ on $I\,dx \in \text{Dens}(\Omega)$ is given by

$$(\varphi, I\,dx) \mapsto \varphi^*(I\,dx) = (|D\varphi|\, I \circ \varphi)\; dx, \tag{4.6}$$

where $|D\varphi|$ denotes the Jacobian determinant of φ. The corresponding left action is given by *pushforward:*

$$(\varphi, I\,dx) \mapsto \varphi_*(I\,dx) = (\varphi^{-1})^*(I\,dx) = \left(|D\varphi^{-1}|\, I \circ \varphi^{-1}\right)\; dx. \tag{4.7}$$

The key characteristic of the density action is that it is mass preserving (the total integral remains constant). The proof of this is straightforward, although I will first prove an identity that will be used several times in this chapter.

Lemma 1. *Given the relationship $\varphi(y) = x$, we have $|D\varphi^{-1}| = \frac{1}{|D\varphi|}$.*

Proof. This can be derived using the chain rule:

$$D_y\varphi^{-1}(\varphi(y)) = D_y y = e \tag{4.8}$$

$$D_x\varphi^{-1}(x) \cdot D_y\varphi(y) = e \tag{4.9}$$

$$|D_x\varphi^{-1}(x)||D_y\varphi(y)| = |e| \tag{4.10}$$

$$|D_x\varphi^{-1}(x)| = \frac{1}{|D_y\varphi(y)|}. \tag{4.11}$$

□

Now I will prove that the density action is mass preserving.

Lemma 2. *For any $\varphi : \Omega \to \Omega \in \text{Diff}(\Omega)$ and any $I(x)\,dx \in \text{Dens}(\Omega)$, we have*

$$\int_\Omega \varphi_* I\,dx = \int_\Omega I\,dx\,. \tag{4.12}$$

Proof. I start with the definition of total mass after applying the density action

$$\int_\Omega \varphi_* I\,dx = \int_\Omega I(\varphi^{-1}(x))|D\varphi^{-1}(x)|\,dx. \tag{4.13}$$

Then, by performing the change of variables

$$x \mapsto \varphi(y), \quad dx \mapsto |D\varphi(y)|\,dy, \quad |D\varphi^{-1}(x)| \mapsto \frac{1}{|D\varphi(y)|}, \tag{4.14}$$

we have

$$\int_\Omega \varphi_* I \, dx = \int_\Omega I(\varphi^{-1}(\varphi(y))) \frac{1}{|D\varphi(y)|} |D\varphi(y)| \, dy = \int_\Omega I(y) \, dy. \tag{4.15}$$

□

The Riemannian geometry of the group of diffeomorphisms endowed with a suitable Sobolev H^1 metric is intimately linked to the Riemannian geometry of the space densities with the Fisher–Rao metric. This has been developed and extensively studied in [15, 16, 17]. The basic observation is that there are Sobolev H^1 metrics on the space of diffeomorphisms that descend to the Fisher–Rao metric on the space of densities.

The tangent space of $\mathrm{Dens}(\Omega)$ consists of the space of n-forms. Given a pair of tangent vectors $\alpha, \beta \in T_\mu \mathrm{Dens}(\Omega)$, the Fisher–Rao metric at the tangent space of μ is given by

$$\bar{G}^F_\mu(\alpha, \beta) = \int_\Omega \frac{\alpha}{\mu} \frac{\beta}{\mu} d\mu(x). \tag{4.16}$$

Now, the information metric defined is defined in Equation 4.4 with the Hodge Laplacian on vector fields which is defined by

$$\Delta u = \nabla \mathrm{div}(u) - \nabla \times \nabla u \, . \tag{4.17}$$

By the Helmholtz-Hodge decomposition, this vector field u can be split into a sum of a divergence free vector field and a curl free vector field. In this case, the divergence free vector field represents the vertical component of $T\,\mathrm{Diff}(\Omega)$ and the curl free vector field represents the horizontal component. Now, the infinitesimal action of a vector field u on a density μ is the Lie derivative

$$\mathcal{L}_u \mu = \mathrm{div}_\mu(u)\mu \, . \tag{4.18}$$

Given two vector fields $u, v \in T\,\mathrm{Diff}(\Omega)$, I define $\alpha = \mathcal{L}_u \mu, \beta = \mathcal{L}_v \mu$ where $\alpha, \beta \in T_\mu \mathrm{Dens}(\Omega)$. Then, the information metric on $T\,\mathrm{Diff}\,\Omega$ descends to the Fisher-Rao metric on $T\,\mathrm{Dens}\,\Omega$:

$$\begin{aligned} G^I(u, v) &= \int_\Omega \langle -\nabla \mathrm{div}(u), v \rangle d\mu(x) \\ &= \int_\Omega \langle \mathrm{div}(u), \mathrm{div}(u) \rangle d\mu(x) \\ &= \int_\Omega \langle \mathrm{div}(u), \mathrm{div}(u) \rangle d\mu(x) \\ &= \int_\Omega \frac{\alpha}{\mu} \frac{\beta}{\mu} d\mu(x) \\ &= \bar{G}^F(\alpha, \beta). \end{aligned} \tag{4.19}$$

Because of this, horizontal geodesics on $\mathrm{Diff}(\Omega)$ descend to Fisher-Rao geodesics on $\mathrm{Dens}(\Omega)$, and there is a unique horizontal lift of curves in $\mathrm{Dens}(\Omega)$ to $\mathrm{Diff}(\Omega)$.

It can be shown that $\mathrm{Dens}(\Omega)$ with the Fisher–Rao metric is isometric to the infinite-dimensional sphere $S^\infty(\Omega)$ [17]. Therefore, the geodesics and distance function associated with the Fisher–Rao metric are explicit. The distance associated with the Fisher–Rao metric is traditionally defined between *probability densities* (densities of total mass 1) and is given by

$$d_F(\mu_0, \mu_1) = \sqrt{\mathrm{vol}(\Omega)} \arccos\left(\frac{1}{\mathrm{vol}(\Omega)} \int_\Omega \sqrt{\frac{\mu_0}{dx} \frac{\mu_1}{dx}}\, dx \right) , \tag{4.20}$$

where μ_0 and μ_1 are probability densities. It naturally extends to the space of all densities and the case when $\mathrm{vol}(\Omega) = \infty$, for which it is given by

$$d_F^2(I_0\, dx, I_1\, dx) = \int_\Omega (\sqrt{I_0} - \sqrt{I_1})^2\, dx\,. \tag{4.21}$$

Notice that $d_F^2(\cdot, \cdot)$ in this case is the *Hellinger distance*. For details, see [17].

The Fisher–Rao metric is the unique Riemannian metric on the space of probability densities that is invariant under the action of the diffeomorphism group [21, 22]. This invariance property extends to the induced distance function, so

$$d_F^2(I_0\, dx, I_1\, dx) = d_F^2(\varphi_*(I_0\, dx), \varphi_*(I_1\, dx)) \qquad \forall \varphi \in \mathrm{Diff}(\Omega)\ . \tag{4.22}$$

This can be verified by applying the same change of variables used to prove Lemma 2 on Equation 4.21.

Motivated by the aforementioned properties, I develop a weighted diffeomorphic matching algorithm for matching two density images. The algorithm is based on the Sobolev H^1 gradient flow on the space of diffeomorphisms that minimizes the energy functional

$$E(\varphi) = d_F^2(\varphi_*(f\, dx), (f \circ \varphi^{-1})\, dx) + d_F^2(\varphi_*(I_0\, dx), I_1\, dx)). \tag{4.23}$$

This energy functional is only a slight modification of the energy functional studied in [17]. Indeed, if f in the above equation is a constant $\sigma > 0$, then Equation 4.23 reduces to the energy functional of Bauer, Joshi, and Modin [17, § 5.1]. Moreover, the geometry described in [17, §5.3] is valid also for the functional in Equation 4.23, and, consequently, the algorithm developed in [17, §5.2] can be used also for minimizing Equation 4.23. There the authors view the energy functional as a constrained minimization problem on the

product space $\mathrm{Dens}(\Omega) \times \mathrm{Dens}(\Omega)$ equipped with the product distance (cf. **Figure 4.1** and [17, §5] for details on the resulting geometric picture). Related work on diffeomorphic density matching using the Fisher–Rao metric can be found in [23, 24].

Using the invariance property of the Fisher–Rao metric and assuming infinite volume, the main optimization problem associated with the energy functional in Equation 4.23 is the following:

Given densities $I_0\,dx$, $I_1\,dx$, and $f\,dx$, find $\varphi \in \mathrm{Diff}(\Omega)$ minimizing

$$E(\varphi) = \underbrace{\int_\Omega (\sqrt{|D\varphi^{-1}|} - 1)^2 \, f \circ \varphi^{-1}\, dx}_{E_1(\varphi)} + \underbrace{\int_\Omega \left(\sqrt{|D\varphi^{-1}| I_0 \circ \varphi^{-1}} - \sqrt{I_1}\right)^2 dx}_{E_2(\varphi)} . \quad (4.24)$$

The invariance of the Fisher–Rao distance can be seen with the same change of variables as before: $x \mapsto \varphi(y)$, $dx \mapsto |D\varphi|\,dy$, and $|D\varphi^{-1}| \mapsto \frac{1}{|D\varphi|}$. Then, Equation 4.24 becomes

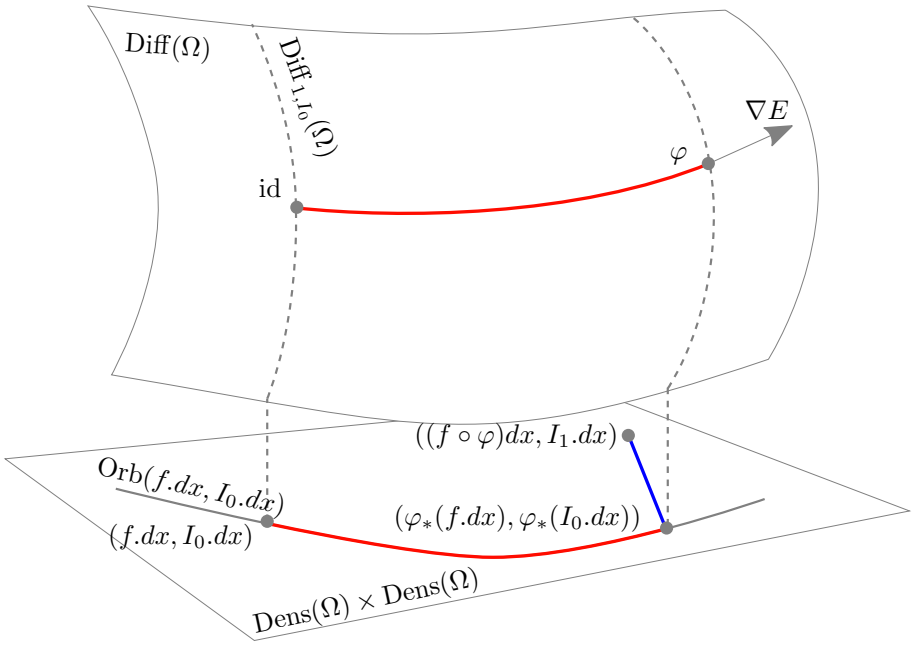


Figure 4.1. Illustration of the geometry associated with the density matching problem. The gradient flow on $\mathrm{Diff}(\Omega)$ descends to a gradient flow on the orbit $\mathrm{Orb}(f\,dx, I_0\,dx)$. While constrained to $\mathrm{Orb}(f\,dx, I_0\,dx) \subset \mathrm{Dens}(\Omega) \times \mathrm{Dens}(\Omega)$, this flow strives to minimize the product Fisher–Rao distance to $((f \circ \varphi)\,dx, I_1\,dx)$.

$$E(\varphi) = \int_\Omega (1 - \sqrt{|D\varphi|})^2 f \, dy + \int_\Omega \left(\sqrt{I_0} - \sqrt{|D\varphi| I_1 \circ \varphi} \right)^2 dy \, . \tag{4.25}$$

To better understand the energy functional $E(\varphi)$, I consider the two terms separately. The first term $E_1(\varphi)$ is a *regularity measure* for the transformation. It penalizes the deviation of the diffeomorphism φ from being volume preserving. The density $f\,dx$ acts as a weighting on the domain Ω. That is, change of volume (compression and expansion of the transformation φ) is penalized more in regions of Ω where f is large. The second term $E_2(\varphi)$ penalizes *dissimilarity* between $I_0\,dx$ and $\varphi^*(I_1\,dx)$. It is the Fisher–Rao distance between the initial density $I_0\,dx$ and the transformed target density $\varphi^*(I_1\,dx)$. Because of the invariance (Equation 4.22) of the Fisher–Rao metric, this is the same as the Fisher–Rao distance between $I_1\,dx$ and $\varphi_*(I_0\,dx)$.

Solutions to the problem in Equation 4.24 are *not* unique. To see this, let $\mathrm{Diff}_I(\Omega)$ denote the space of all diffeomorphisms preserving the volume form $I\,dx$:

$$\mathrm{Diff}_I(\Omega) = \{\varphi \in \mathrm{Diff}(\Omega) \mid \; |D\varphi| \, (I \circ \varphi) = I\}. \tag{4.26}$$

If φ is a minimizer of $E(\cdot)$, then $\psi \circ \varphi$ for any

$$\psi \in \mathrm{Diff}_{1,I_0}(\Omega) \coloneqq \mathrm{Diff}_1(\Omega) \cap \mathrm{Diff}_{I_0}(\Omega) \tag{4.27}$$

is also a minimizer. Notice that this space is not trivial. For example, any diffeomorphism generated by a *Nambu–Poisson vector field* (see [25]), with I_0 as one of its Hamiltonians, will belong to it. A strategy to handle the degeneracy was developed in [17, §5]. The fact that the metric is descending with respect to the H^1 metric on $\mathrm{Diff}(\Omega)$ can be used to ensure that the gradient flow is *infinitesimally optimal*, i.e., always orthogonal to the null-space. I employ the same strategy in this chapter. The corresponding geometric picture can be seen in **Figure 4.1**.

4.3 Gradient Flow Algorithm Development

I now derive in detail the algorithm used to optimize the functional defined in Equation 4.25. Let $\nabla^{G^I} E$ denote the gradient with respect to the information metric defined in Equation 4.4. My approach to minimize the functional of Equation 4.25 is to use a simple Euler integration of the discretization of the gradient flow:

$$\dot{\varphi} = -\nabla^{G^I} E(\varphi). \tag{4.28}$$

The resulting final algorithm (Algorithm 2) is order of magnitudes faster than LDDMM, since there is no requirement to time integrate the geodesic equations, as is necessary in LDDMM [26].

In the following theorem, I calculate the gradient of the energy functional:

Theorem 1. *The G^I–gradient of the matching functional in Equation 4.25 is given by*

$$\nabla^{G^I} E = -\Delta^{-1}\Big(-\nabla\big(f\circ\varphi^{-1}(1-\sqrt{|D\varphi^{-1}|})\big) - \sqrt{|D\varphi^{-1}|\, I_0\circ\varphi^{-1}}\nabla\big(\sqrt{I_1}\big) + \nabla\big(\sqrt{|D\varphi^{-1}|\, I_0\circ\varphi^{-1}}\big)\sqrt{I_1}\Big)\,. \tag{4.29}$$

Remark 1. *Notice that in the formula for $\nabla^{G^I} E$, I never need to compute φ, so in practice, I only compute φ^{-1}. I update this directly via $\varphi^{-1}(x) \mapsto \varphi^{-1}(x + \epsilon\nabla^{G^I} E)$ for some step size ϵ.*

Proof. I first calculate the variation of the energy functional. Therefore let φ_s be a family of diffeomorphisms parameterized by the real variable s, such that

$$\varphi_0 = \varphi \quad \text{and} \quad \frac{d}{ds}\Big|_{s=0}\varphi_s = v\circ\varphi. \tag{4.30}$$

I use the following identity, as derived in [27]:

$$\frac{d}{ds}\Big|_{s=0}\sqrt{|D\varphi_s|} = \frac{1}{2}\sqrt{|D\varphi|}\mathrm{div}(v)\circ\varphi. \tag{4.31}$$

The variation of the first term of the energy functional is

$$\frac{d}{ds}\Big|_{s=0} E_1(\varphi) = \int_\Omega f(y)(\sqrt{|D\varphi(y)|}-1)\sqrt{|D\varphi(y)|}\mathrm{div}(v)\circ\varphi(y)\,dy\,. \tag{4.32}$$

I perform a change of variable $y \mapsto \varphi^{-1}(x)$, $dy \mapsto |D\varphi^{-1}(x)|\,dx$, again using the fact that $|D\varphi(y)| = \frac{1}{|D\varphi^{-1}(x)|}$:

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=0} E_1(\varphi) &= \int_\Omega f\circ\varphi^{-1}(x)(1-\sqrt{|D\varphi^{-1}(x)|})\mathrm{div}(v(x))\,dx && (4.33)\\
&= \Big\langle f\circ\varphi^{-1}(1-\sqrt{|D\varphi^{-1}|}), \mathrm{div}(v)\Big\rangle_{L^2(\mathbb{R}^3)} && (4.34)\\
&= -\Big\langle \nabla\Big(f\circ\varphi^{-1}(1-\sqrt{|D\varphi^{-1}|})\Big), v\Big\rangle_{L^2(\mathbb{R}^3)}\,, && (4.35)
\end{aligned}$$

using the fact that the adjoint of the divergence is the negative gradient. For the second term of the energy functional, I expand the square

$$E_2(\varphi) = \int_\Omega I_0(y) - 2\sqrt{I_0(y) I_1\circ\varphi(y)|D\varphi(y)|} + I_1\circ\varphi(y)|D\varphi(y)|\,dy\,. \tag{4.36}$$

Now $\int_\Omega I_1 \circ \varphi(y)|D\varphi(y)|\,dy$ is constant (conservation of mass), so I only need to minimize over the middle term. The derivative is then

$$\frac{d}{ds}\Big|_{s=0} E_2(\varphi) = -\int_\Omega 2\sqrt{I_0(y)}(\nabla\sqrt{I_1}^T v)\circ\varphi(y)\sqrt{|D\varphi(y)|} - \sqrt{I_0(y)I_1\circ\varphi(y)|D\varphi(y)|}\mathrm{div}(v)\circ\varphi(y)\,dy. \tag{4.37}$$

I do the same change of variables as before:

$$\frac{d}{ds}\Big|_{s=0} E_2(\varphi) = -\int_\Omega \sqrt{I_0\circ\varphi^{-1}(x)}\frac{|D\varphi^{-1}(x)|}{\sqrt{|D\varphi^{-1}(x)|}}\left(2\nabla\sqrt{I_1(x)}^T v(x) + \sqrt{I_1(x)}\mathrm{div}(v)(x)\right)dx \tag{4.38}$$

$$= -\left\langle 2\sqrt{|D\varphi^{-1}|\,I_0\circ\varphi^{-1}}\nabla\sqrt{I_1}, v\right\rangle_{L^2(\mathbb{R}^3)} - \left\langle\sqrt{|D\varphi^{-1}|\,I_0\circ\varphi^{-1}I_1}, \mathrm{div}(v)\right\rangle_{L^2(\mathbb{R}^3)} \tag{4.39}$$

$$= \left\langle -\sqrt{|D\varphi^{-1}|\,I_0\circ\varphi^{-1}}\nabla\sqrt{I_1}, v\right\rangle_{L^2(\mathbb{R}^3)} + \left\langle\nabla\left(\sqrt{|D\varphi^{-1}|\,I_0\circ\varphi^{-1}}\right)\sqrt{I_1}, v\right\rangle_{L^2(\mathbb{R}^3)}. \tag{4.40}$$

From the above equations, I conclude that

$$-\Delta(\nabla^{G^I}E) = -\nabla\left(f\circ\varphi^{-1}(1-\sqrt{|D\varphi^{-1}|})\right) - \sqrt{|D\varphi^{-1}|\,I_0\circ\varphi^{-1}}\nabla\sqrt{I_1} + \nabla\left(\sqrt{|D\varphi^{-1}|\,I_0\circ\varphi^{-1}}\right)\sqrt{I_1}. \tag{4.41}$$

Since I am taking the Sobolev gradient of E, I apply the inverse Laplacian to the right-hand side of Equation 4.41 to solve for $\nabla^{G^I}E$. □

Algorithm 2 Weighted Diffeomorphic Density Matching

Choose $\epsilon > 0$
Set $\varphi^{-1} = \mathrm{id}$
Set $|D\varphi^{-1}| = 1$
for $iter = 0$...numIters **do**
  Compute $\varphi_* I_0 = I_0\circ\varphi^{-1}|D\varphi^{-1}|$
  Compute $u = -\nabla\big(f\circ\varphi^{-1}(1-\sqrt{|D\varphi^{-1}|})\big) - \sqrt{\varphi_* I_0}\nabla\sqrt{I_1} + \nabla(\sqrt{\varphi_* I_0})\sqrt{I_1}$
  Compute $v = -\Delta^{-1}(u)$
  Update $\varphi^{-1} \mapsto \varphi^{-1}(y + \epsilon v)$
  Update $|D\varphi^{-1}| \mapsto |D\varphi^{-1}|\circ\varphi^{-1}\mathrm{e}^{-\epsilon\mathrm{div}(v)}$
end for

Remark 2. *Algorithm 2 constructs the mapping φ^{-1} by numerically integrating the vector field v. Thus, for small enough ϵ, the computed transformation φ^{-1} is a diffeomorphism (as is also the case in LDDMM).*

4.4 Results

4.4.1 Evaluation on RCCT Dataset

I applied the proposed method to phase-binned human thoracic CT datasets acquired at the University of Texas Southwestern Medical Center. This dataset consists of three patients who are candidates for radiation therapy of a tumor in the lungs. Each patient dataset consists of a full breathing cycle with images at 10 phase points of the respiratory cycle. Each image has corresponding segmentations of certain regions of interest that were manually segmented by a clinical specialist. These segmentations include the gross tumor volume (GTV), the planning target volume (PTV), and a segmentation of the lungs. The GTV describes the primary tumor whereas the PTV consists of the primary tumor plus a small expansion (from 1.0 to 1.5 cm surrounding the tumor). These segmentations for a single phase of one of the patients are shown in Figure 4.2.

I used the proposed method to register each phase image to the full-exhale image (phase 5), where each image was first transformed with the power function introduced in the previous chapter. To show the properties of the resulting deformations, I show results from a registration of the full-exhale phase to the full-inhale phase of the first patient.

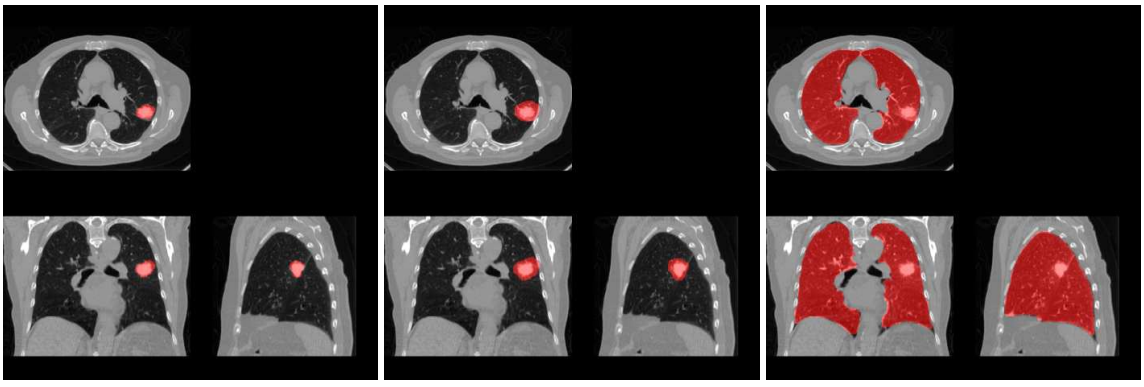
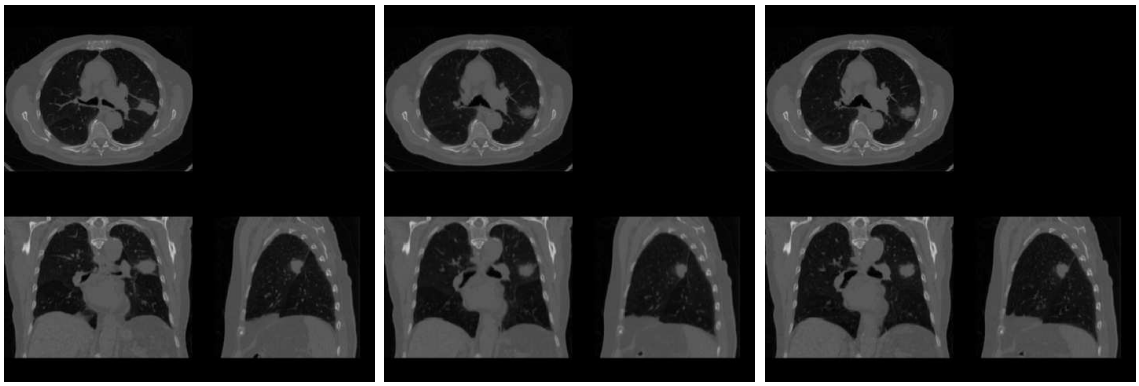


Figure 4.2. From left to right, the GTV, PTV, and lung segmentations are the red regions. The GTV segmentation contains the tumor, the PTV segmentation contains the tumor plus a surrounding region of interest, and the lung segmentation contains both lungs.

Shown in **Figure 4.3** are the full exhale, the full exhale deformed via the density action to the full inhale image, and the corresponding target image at full inhale. The data match term is also shown in **Figure 4.4**, where I show the change in the Fisher–Rao metric from the unregistered images to the density-registered images. The majority of the intensity mismatch in the unregistered image is at the diaphragm and the tumor.

For the compressibility penalty weighting f, I used a soft thresholding of the intensity values of the initial image using the logistic function. While any smooth positive function can be used, a soft thresholding of CT intensities provides a fairly accurate estimate of compressibility. High-intensity regions of the CT image (corresponding to bone and soft tissue) are given a high penalty ($f(x) = 10\sigma$), and low-intensity regions of the CT image (corresponding to air and lungs) are given a low penalty ($f(x) = 0.1\sigma$) (see **Figure 4.5**).

The Jacobian determinants of these methods plus the Jacobian determinants found in ANTs and NiftyReg are in **Figure 4.6**. The proposed method constrains the contraction and expansion to inside the lung and outside the body. In this figure, I also show the results of using the density action with a constant penalty function ($f(x) = \sigma$). The contraction and expansion of the Jacobian determinant from ANTS and NiftyReg is quite severe and does not represent normal physiological motion.



I_{in} $\varphi_*(I_{ex}\,dx)$ I_{ex}

Figure 4.3. Density action results. This figure shows the lung image at the full exhale, the full exhale deformed via the density action, and the corresponding image at full inhale.

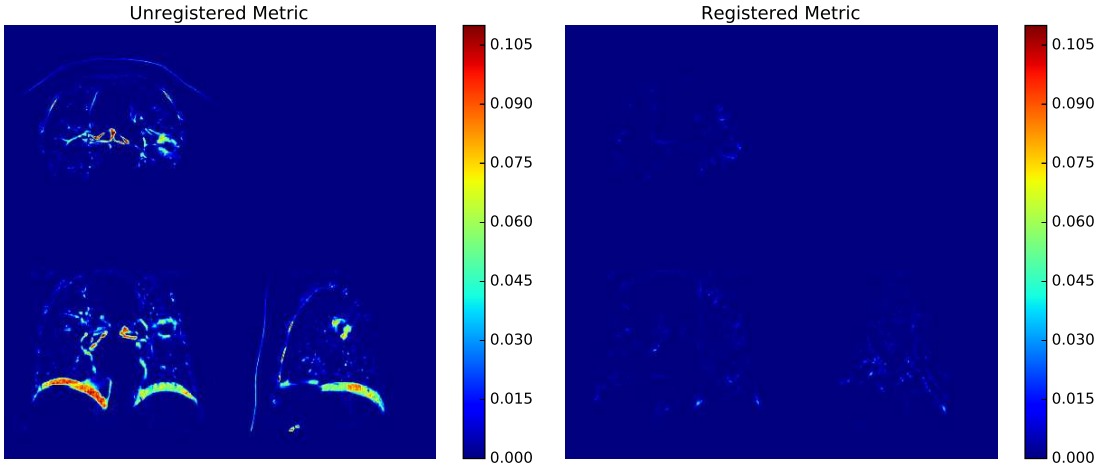


Figure 4.4. Fisher–Rao metric between the unregistered densities (left) and the Fisher–Rao metric between the deformed and target densities (right), using the proposed method.

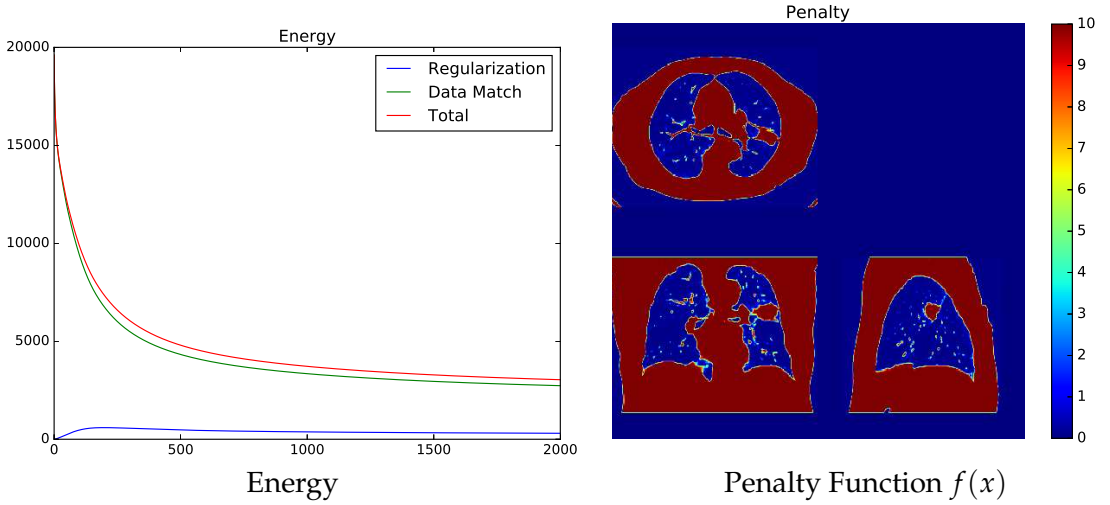


Figure 4.5. Energy plot and the weighting density. The energy is monotonically decreasing, and the penalty image places a high penalty on volume change outside the lungs.

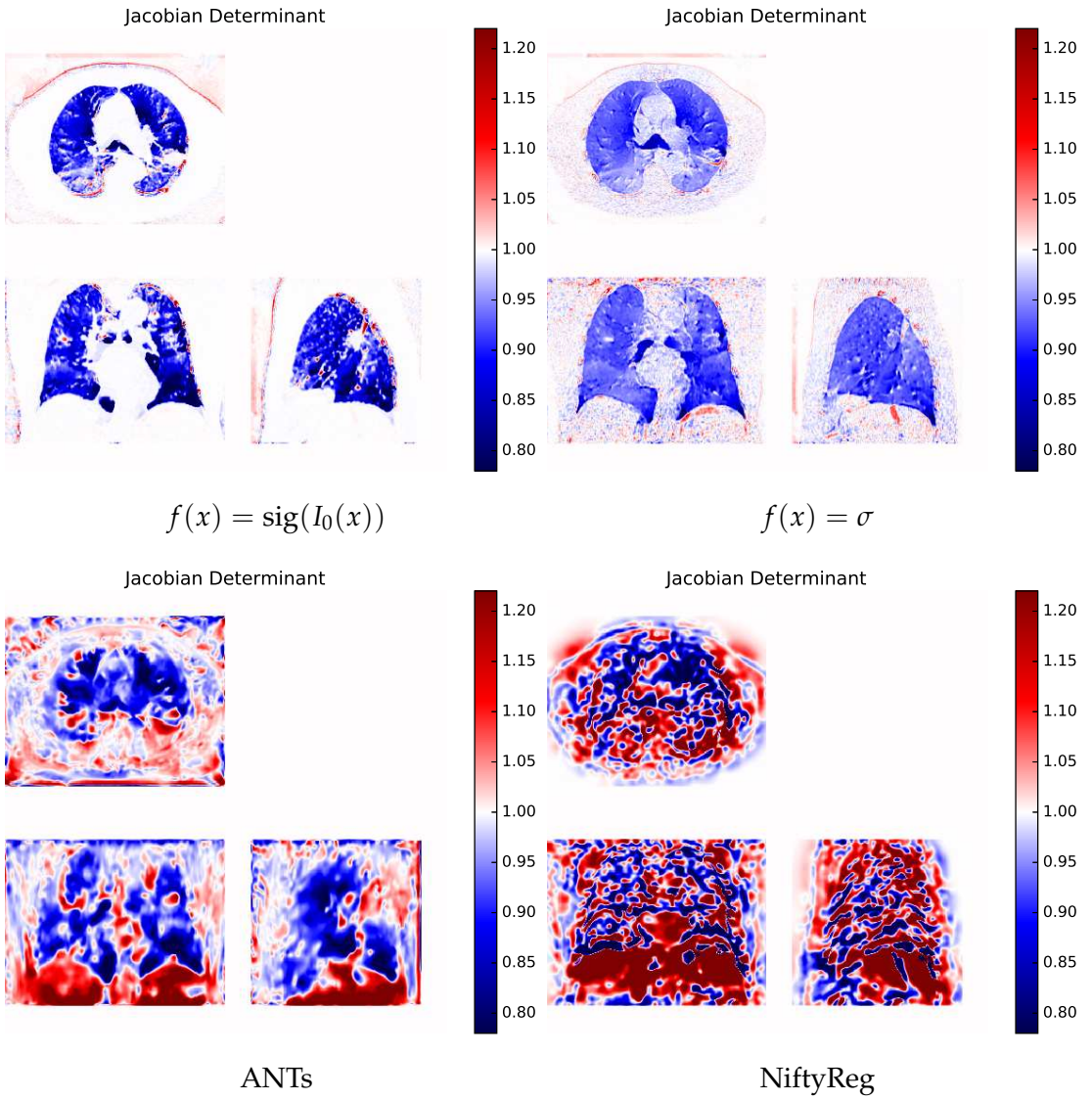


Figure 4.6. Jacobian determinants: In the upper left is the Jacobian determinant of the transformation estimated by the proposed method. Notice that the volume change is confined to inside the lungs and outside the body. In the upper right, I use the density action, but without a locally-varying penalty (i.e., $f(x) = \sigma$). Without the local-varying penalty, there is contraction and expansion outside of the lungs to match noise. On the bottom row are the Jacobian determinants of ANTs (left) and NiftyReg (right). In ANTs and NiftyReg, the contraction and expansion outside of the lungs are quite severe, and the mixture of contraction and expansion inside the lungs does not model known physiological motion.

4.4.2 Registration Accuracy

In order to evaluate the accuracy of the weighted diffeomorphic density matching algorithm, I compared the registration accuracy between the proposed method and two current state-of-the-art registration methods: ANTs and NiftyReg. To quantify the accuracy, I used the previously mentioned manually segmented GTV, PTV, and lungs, and compared the postregistration Dice coefficients of all the methods. All three methods showed a significantly higher average Dice coefficient compared to the unregistered images. Overall, the weighted density matching method had the highest overall accuracy. I looked at each of the segmentations separately and performed statistical analysis on the improvement. The three datasets consisted of 27 segmentations (3 of the 30 were missing), so there was a total of 24 registrations performed (as each image was registered to the full exhale image). The average Dice coefficients for each region can be seen in **Figure 4.7** and in **Table 4.1**. Overall, the proposed method is significantly better than the other two methods.

For the statistical analysis, I performed nine paired t-tests comparing my method to ANTs, NiftyReg, and unregistered for each of the three regions of interest. I corrected for multiple comparisons using the Holm-Bonferroni method.

To better understand the magnitude of the differences between registration methods,

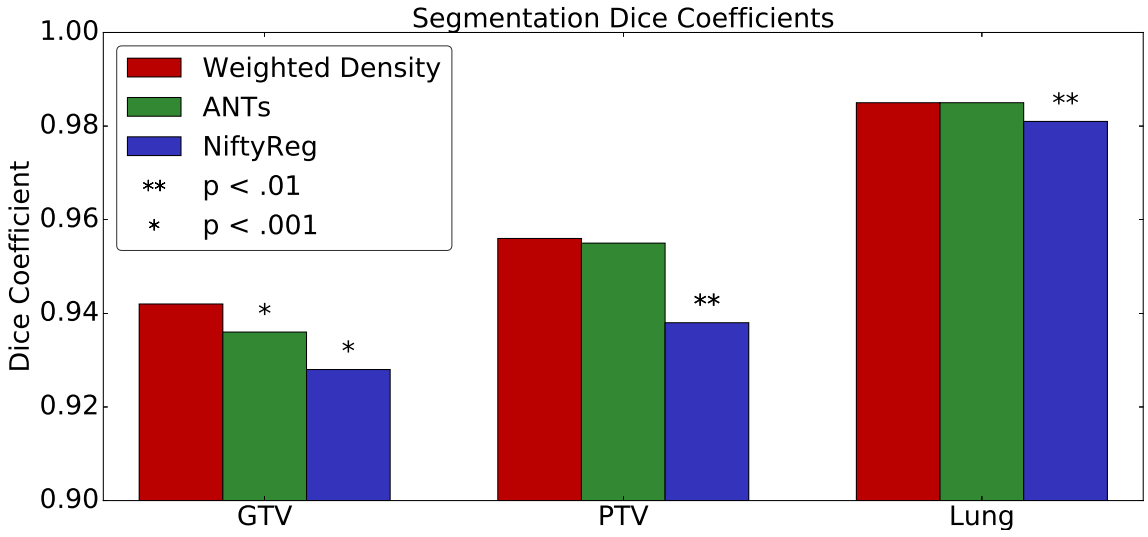


Figure 4.7. Dice coefficients after registration comparing three methods: The proposed weighted density matching, ANTs, and NiftyReg. Overall, the proposed method is better than the current state-of-the-art methods for deformable thoracic registration.

Table 4.1. Dice coefficient averages for all three datasets (24 independent registrations). The significance refers to paired t-tests between the proposed density registration and each individual method ($*$ is $p < 0.01$, $**$ is $p < 0.001$).

	Density	ANTs	NiftyReg	Unregistered
GTV	0.942	0.936*	0.928*	0.763**
PTV	0.956	0.955	0.938**	0.853**
Lung	0.985	0.985	0.981**	0.957**

I also computed the estimated displacement error of the tumor. For this, I computed the centers of mass of the deformed tumors and compared them to the reference center of mass at full exhale for each of the three patients. The mean displacement error was 4.77 mm ($\sigma = 3.36$ mm) for the unregistered tumors, 0.678 mm ($\sigma = 0.372$ mm) using ANTs, 0.698 mm ($\sigma = 0.475$ mm) using NiftyReg, and 0.452 mm ($\sigma = 0.239$ mm) using the proposed weighted density registration. Once again, the proposed method is more accurate than the proposed methods.

Both ANTs and NiftyReg packages are implemented on the CPU and have significant preprocessing steps. I implemented the proposed algorithm on a single Titan-Z GPU (using the PyCA software package [28] `http://bitbucket.org/scicompanat/pyca`). For a more relevant performance comparison to my method, I also implemented an intensity-based LDDMM method. The proposed algorithm is substantially faster than LDDMM-like methods. The proposed algorithm runs at 100 iterations per minute whereas LDDMM runs at 12 iterations per minute. I used 10 time steps to integrate the geodesic equations associated with the LDDMM formulation. Since I am not required to integrate the geodesic equations in the proposed algorithm, I have a nearly 10x speed-up compared to LDDMM.

4.5 Discussion

In this chapter, I introduced a computationally efficient method for estimating registration maps between thoracic CT images. The proposed solution accurately incorporates the fundamental property of mass conservation and the spatially varying compressibility of thoracic anatomy. CT mass was conserved by viewing the images as densities and applying the density action of a diffeomorphism instead of the typical L^2 action. The volume

change in incompressible organs was limited by placing a spacially-varying penalty on the Jacobian determinant of the diffeomorphism. Although any nonnegative function $f(x)$ can be used, I simply used a soft-thresholding function on the initial image. This choice is based on the assumption that low CT image values (such as the lungs and air) exhibit a large amount of volume change whereas high image values (such as other soft tissue and bone) are nearly incompressible.

The resulting method is more accurate than current state-of-the-art registration methods and is considerably more computationally efficient.

4.6 Acknowledgments

I would like to thank Pouya Sabouri and Amit Sawant for providing the patient data and segmentations.

4.7 References

[1] A. Sawant, P. Keall, K. B. Pauly, M. Alley, S. Vasanawala, B. W. Loo, J. Hinkle, and S. Joshi, "Investigating the feasibility of rapid MRI for image-guided motion management in lung cancer radiotherapy," *BioMed Research International*, vol. 2014, 2014.

[2] P. J. Keall, S. Joshi, S. S. Vedam, J. V. Siebers, V. R. Kini, and R. Mohan, "Four-dimensional radiotherapy planning for DMLC-based respiratory motion tracking," *Medical Physics*, vol. 32, no. 4, pp. 942–951, 2005.

[3] Y. Suh, W. Murray, and P. J. Keall, "IMRT treatment planning on 4D geometries for the era of dynamic MLC tracking," *Technology in Cancer Research and Treatment*, vol. 13, no. 6, pp. 505–515, 2014.

[4] S. E. Geneser, J. D. Hinkle, R. M. Kirby, B. Wang, B. Salter, and S. Joshi, "Quantifying variability in radiation dose due to respiratory-induced tumor motion," *Medical Image Analysis*, vol. 15, no. 4, pp. 640–649, 2011.

[5] CDC, "Basic information about lung cancer." `http://www.cdc.gov/cancer/lung/basic_info/index.htm`. Accessed: 2016-03-06.

[6] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *International Journal of Computer Vision*, vol. 61, no. 2, pp. 139–157, 2005.

[7] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ANTs similarity metric performance in brain image registration," *NeuroImage*, vol. 54, no. 3, pp. 2033–2044, 2011.

[8] K. Murphy, B. Van Ginneken, J. M. Reinhardt, S. Kabus, K. Ding, X. Deng, K. Cao, K. Du, G. E. Christensen, V. Garcia, T. Vercauteren, N. Ayache, O. Commowick,

G. Malandain, B. Glocker, N. Paragios, N. Navab, V. Gorbunova, J. Sporring, M. De Bruijne, X. Han, M. P. Heinrich, J. a. Schnabel, M. Jenkinson, C. Lorenz, M. Modat, J. R. McClelland, S. Ourselin, S. E. a. Muenzing, M. a. Viergever, D. De Nigris, D. L. Collins, T. Arbel, M. Peroni, R. Li, G. C. Sharp, A. Schmidt-Richberg, J. Ehrhardt, R. Werner, D. Smeets, D. Loeckx, G. Song, N. Tustison, B. Avants, J. C. Gee, M. Staring, S. Klein, B. C. Stoel, M. Urschler, M. Werlberger, J. Vandemeulebroucke, S. Rit, D. Sarrut, and J. P. W. Pluim, "Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge," *IEEE Transactions on Medical Imaging*, vol. 30, no. 11, pp. 1901–1920, 2011.

[9] G. Song and N. Tustison, "Lung CT image registration using diffeomorphic transformation models," in *Medical Image Analysis for the Clinic – A Grand Challenge: Workshop Proceedings from the 13th International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 23–32, 2010.

[10] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Medical Image Analysis*, vol. 12, no. 1, pp. 26–41, 2008.

[11] M. Modat, J. McClelland, and S. Ourselin, "Lung registration using the NiftyReg package," *Medical Image Analysis for the Clinic: A Grand Challenge EMPIRE 10*, pp. 33–42, 2010.

[12] M. Modat, G. Ridgway, Z. Taylor, M. Lehmann, J. Barnes, D. Hawkes, N. Fox, and S. Ourselin, "Fast free-form deformation using graphics processing units.," *Computer Methods and Programs in Biomedicine*, vol. 98, no. 3, pp. 278–284, 2010.

[13] Y. Yin, E. A. Hoffman, and C.-L. Lin, "Mass preserving nonrigid registration of CT lung images using cubic B-spline," *Medical Physics*, vol. 36, no. 9, pp. 4213–4222, 2009.

[14] V. Gorbunova, J. Sporring, P. Lo, M. Loeve, H. A. Tiddens, M. Nielsen, A. Dirksen, and M. de Bruijne, "Mass preserving image registration for lung CT," *Medical Image Analysis*, vol. 16, no. 4, pp. 786–795, 2012.

[15] B. Khesin, J. Lenells, G. Misiołek, and S. C. Preston, "Geometry of diffeomorphism groups, complete integrability and geometric statistics," *Geometry of Functional Analysis*, vol. 23, no. 1, pp. 334–366, 2013.

[16] K. Modin, "Generalized Hunter-Saxton equations, optimal information transport, and factorization of diffeomorphisms," *The Journal of Geometric Analysis*, vol. 25, no. 2, pp. 1306–1334, 2015.

[17] M. Bauer, S. Joshi, and K. Modin, "Diffeomorphic density matching by optimal information transport," *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 1718–1751, 2015.

[18] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinetics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[19] S. Joshi, P. Lorenzen, G. Gerig, and E. Bullitt, "Structural and radiometric asymmetry in brain images," *Medical Image Analysis*, vol. 7, no. 2, pp. 155–170, 2003.

[20] M. I. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *International Journal of Computer Vision*, vol. 41, no. 1-2, pp. 61–84, 2001.

[21] M. Bauer, M. Bruveris, and P. W. Michor, "Uniqueness of the Fisher–Rao metric on the space of smooth densities," *Bulletin of the London Mathematical Society*, vol. 48, no. 3, pp. 499–506, 2016.

[22] N. Ay, J. Jost, H. V. Le, and L. Schwachhöfer, "Information geometry and sufficient statistics," *Probability Theory and Related Fields*, vol. 162, no. 1, pp. 327–364, 2015.

[23] D. H. Seo, J. Ho, J. H. Traverse, J. Forder, and B. Vemuri, "Computing diffeomorphic paths with application to cardiac motion analysis," in *4th MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, pp. 83–94, 2013.

[24] D. Seo, J. Ho, and B. C. Vemuri, "Computing diffeomorphic paths for large motion interpolation," in *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1227–1232, June 2013.

[25] N. Nakanishi, "A survey of Nambu-Poisson geometry," *Lobachevskii Journal of Mathematics*, vol. 4, pp. 5–11 (electronic), 1999.

[26] L. Younes, F. Arrate, and M. I. Miller, "Evolution equations in computational anatomy," *NeuroImage*, vol. 45, no. 1, Supplement 1, pp. S40–S50, 2009.

[27] J. Hinkle and S. Joshi, "IDiff : Irrotational diffeomorphisms for computational anatomy," in *Information Processing in Medical Imaging*, pp. 754–765, Springer, 2013.

[28] J. S. Preston, J. Hinkle, N. Singh, C. Rottman, and S. Joshi, "PyCA - Python for computational anatomy." `https://bitbucket.org/scicompanat/pyca`, 2012-2016.

CHAPTER 5

DISCUSSION

In this chapter, I discuss the contributions of this dissertation and the potential for future work.

5.1 Summary of Contributions

Below, I will state each claim made in the introduction with a summary of how each claim was addressed.

5.1.1 Chapter 2

A method for 3D cone-beam CT reconstruction for mobile C-arms is introduced. Current methods for cone-beam CT reconstructions assume a stable and well-known geometry of the imaging system. The hardware of mobile C-arms is considerably less geometrically robust than fixed-room C-arms. Therefore, the introduced method includes the estimation of the motion of the C-arm as well as the estimation of the underlying 3D image. This joint estimation of the image and the geometry is done in an alternating scheme, where the image is estimated in the expectation maximization (EM) framework and the geometry is estimated using a gradient-based energy minimization scheme to improve the correspondence between the acquisition data and estimated 3D image.

For CT reconstruction, I showed that the ordered subset expectation maximization (OSEM) algorithm with a TV prior provides excellent image quality in addition to fast convergence. I derived and implemented the gradients with respect to the extrinsic and intrinsic parameters. I then used these to maximize the normalized cross-correlation (NCC) between the projection images and the projection operator applied to the current estimate of the volume. I showed that the resulting estimate of the geometry is very close to the true ground truth geometry. I also evaluated this method on four C-arm datasets. In each of these cases, the resulting reconstruction image is much improved and devoid of miscalibration artifacts. I implemented the proposed method on the GPU using a multiscale

approach, yielding a computationally efficient reconstruction algorithm that is robust to uncertain geometry.

5.1.2 Chapter 3

This chapter provides analysis on the mass-preserving nature of thoracic CT images. Although image intensities in CT should be theoretically conserved, I show that in real thoracic CT datasets, the assumption of conservation of mass is not valid. However, I provide a method to transform these CT images into objects that are indeed mass preserving, and show the approach is applicable across multiple patient datasets acquired on the same CT scanner.

I proved that in a closed system, the narrow-beam LAC mass is conserved. This underlying property causes the intensity changes that occur in the lungs due to respiration. However, I showed that the total intensity in CT images is *not* conserved, as CT data consist of images of *effective* linear attenuation coefficients, not the true narrow-beam linear attenuation coefficients. These raw CT images are not suitable for use in a density-matching algorithm. Therefore, I introduced a method to convert these raw CT images into images that are indeed mass preserving. By applying a power function to the CT images, they are converted into images that are mass preserving. Using two large sets of RCCT patient datasets, I solved for the optimal power function parameter and showed that applying this transformation yields consistent results across a single CT scanner.

5.1.3 Chapter 4

A new method for thoracic CT image registration is introduced. The proposed diffeomorphic model acts on densities, thereby ensuring that the deformations preserve mass. These deformations are regularized by a locally varying penalty on volume change. In incompressible areas of the body (including surrounding soft tissue and bones), the volume change is restricted. In much more compressible areas of the image (such as the lungs), the method allows considerable compression and expansion. The resulting algorithm is computationally efficient and accurately models respiratory motion.

The proposed method uses the recently discovered intimate link between diffeomorphisms and densities [1, 2, 3] that allows solving for the distance metric on diffeomorphisms in closed form and allows for a natural weighting on local tissue compressibility. I implemented this method and compared it to current state-of-the-art lung registration

techniques. The resulting deformations show much more physiologically realistic motion than previous state-of-the-art methods. I also quantified the accuracy of the deformations by comparing the accuracy for regions of interest in radiation oncology. The proposed method is significantly more accurate than the reference methods and is orders of magnitude faster.

5.2 Future Work

5.2.1 4DCT Reconstruction Using Weighted Density Registration

The focus of Chapter 4 is the registration of separate RCCT images. Recently, methods have been introduced for 4D reconstruction [4, 5] of respiratory CT datasets. These methods jointly estimate respiratory motion and the reconstructed image to help reduce artifacts common in phase-binned images. These methods already directly incorporate a motion model into the reconstruction framework, so adding the weighted diffeomorphic density model would be straightforward. One benefit to 4D reconstruction is that it no longer splits data into 10 distinct reconstruction images. Rather, 4DCT uses the full data to create a continuous motion model of the respiratory cycle that is indexed by the breathing amplitude or phase. The benefits of this addition include a reduction in binning artifacts and an increase in the signal-to-noise ratio.

One further improvement to 4DCT would be to model cycle-to-cycle phase variations. In this application, the 4D reconstruction would be parameterized by two surrogate signals: one that describes the location in the breathing phase and one that describes the magnitude of the current phase. This would further increase the accuracy of tumor tracking and improve treatment in radiation oncology.

5.2.2 Joint Reconstruction, Geometry Estimation, and Respiratory Estimation

In future work, both methods of motion estimation provided in this dissertation can be combined to yield a method for joint reconstruction, geometry estimation, and estimation of respiration for cone-beam thoracic CT. In this case, a C-arm would perform an orbital scan of the thorax. The reconstruction would then be done incorporating the estimation of the geometric parameters of the C-arm while also incorporating the proposed diffeomorphic motion model during the 4D reconstruction.

5.2.3 Quantitative Radiation Dose Calculation

CT imaging has traditionally had the primary purpose of being a visual tool, with which trained clinicians make diagnoses and treatment plans. In radiation therapy, it is necessary to quantitatively know the attenuation coefficients in the patient, as this information is needed to calculate the dose absorbed by the tumor and surrounding organs. The purpose of Chapter 3 was to transform an image of effective attenuation coefficients into an image that acts as a density, i.e., has conservation of mass. Further investigative work is needed to transform an image of effective attenuation coefficients into an image of true narrow-beam linear attenuation coefficients, which necessarily has conservation of mass. This improvement would allow more accurate dose calculations in radiation therapy of the lungs and therefore provide more reliable treatment.

5.3 References

[1] B. Khesin, J. Lenells, G. Misiołek, and S. C. Preston, "Geometry of diffeomorphism groups, complete integrability and geometric statistics," *Geometry of Functional Analysis*, vol. 23, no. 1, pp. 334–366, 2013.

[2] K. Modin, "Generalized Hunter-Saxton equations, optimal information transport, and factorization of diffeomorphisms," *The Journal of Geometric Analysis*, vol. 25, no. 2, pp. 1306–1334, 2015.

[3] M. Bauer, S. Joshi, and K. Modin, "Diffeomorphic density matching by optimal information transport," *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 1718–1751, 2015.

[4] J. Hinkle, M. Szegedi, B. Wang, B. Salter, and S. Joshi, "4D CT image reconstruction with diffeomorphic motion model," *Medical Image Analysis*, vol. 16, no. 6, pp. 1307–1316, 2012.

[5] J. Wang and X. Gu, "Simultaneous motion estimation and image reconstruction (SMEIR) for 4D cone-beam CT," *Medical Physics*, vol. 40, no. 10, p. 101912, 2013.