

HIGH-RESOLUTION QUANTITATIVE CONE-BEAM COMPUTED TOMOGRAPHY:
SYSTEMS, MODELING, AND ANALYSIS
FOR IMPROVED MUSCULOSKELETAL IMAGING

by

Qian Cao

A dissertation submitted to Johns Hopkins University
in conformity with the requirements for the degree of Doctor of Philosophy
Baltimore, Maryland
November 2020

© 2020 Qian Cao
All rights reserved

Abstract

This dissertation applies accurate models of imaging physics, new high-resolution imaging hardware, and novel image analysis techniques to benefit quantitative applications of x-ray CT in *in vivo* assessment of bone health. Imaging physics can be used to account for nonidealities in image formation and to optimize imaging systems; improved spatial resolution enables characterization of fine anatomical structures; novel image analysis techniques provide more consistent quantitative biomarkers. We pursue three Aims: 1. Characterization of macroscopic joint space morphology, 2. Estimation of bone mineral density (BMD), and 3. Visualization of bone microstructure. This work contributes to the development of extremity cone-beam CT (CBCT), a compact system for musculoskeletal (MSK) imaging.

Joint space morphology is characterized by a model which draws an analogy between the bones of a joint and the plates of a capacitor. Virtual electric field lines connecting the two surfaces of the joint are computed as a surrogate measure of joint space width, creating a rich, non-degenerate, adaptive map of the joint space. We showed that by using such maps, a classifier can outperform radiologist measurements at identifying osteoarthritic patients in a set of CBCT scans.

Quantitative BMD accuracy is achieved by combining a polyenergetic model-based iterative reconstruction (MBIR) method with fast Monte Carlo (MC) scatter estimation. On a benchtop system emulating extremity CBCT, we validated BMD accuracy and

reproducibility via a series of phantom studies involving inserts of known mineral concentrations and a cadaver specimen.

High-resolution imaging is achieved using a complementary metal-oxide semiconductor (CMOS)-based x-ray detector featuring small pixel size and low readout noise. A cascaded systems model was used to performed task-based optimization to determine optimal detector scintillator thickness in nominal extremity CBCT imaging conditions. We validated the performance of a prototype scanner incorporating our optimization result. Strong correlation was found between bone microstructure metrics obtained from the prototype scanner and μ CT gold standard for trabecular bone samples from a cadaver ulna.

Additionally, we devised a multiresolution reconstruction scheme allowing fast MBIR to be applied to large, high-resolution projection data. To model the full scanned volume in the reconstruction forward model, regions outside a finely sampled region-of-interest (ROI) are downsampled, reducing runtime and cutting memory requirements while maintaining image quality in the ROI.

Dissertation Committee

Wojciech Zbijewski, Ph.D. (advisor)
*Assistant Professor, Department of Biomedical Engineering
Johns Hopkins University*

Jeffrey H. Siewersen, Ph.D. (co-advisor)
*Professor, Department of Biomedical Engineering
Johns Hopkins University*

J. Webster Stayman, Ph.D.
*Associate Professor, Department of Biomedical Engineering
Johns Hopkins University*

Jerry L. Prince
*Professor, Department of Electrical and Computer Engineering
Johns Hopkins University*

Acknowledgements

I am deeply indebted to many years of phenomenal mentorship received at Hopkins. It is especially awe-inspiring to experience the inflationary expansion of the I-STAR universe and witness the ever-growing constellation of research projects. The faculty of I-STAR and AIAI are all brilliant engineer-scientists as well as educators. It is nothing short of a privilege to work in such a collaborative and stimulating environment. First, I would like to thank my advisor, Wojciech Zbijewski for leading me to many scenic views on this journey, and for being patient and nurturing along the way. I would like to thank my co-advisor, Jeffrey Siewersen, for setting high standards in research and communication, and for his highly contagious enthusiasm on all things medical imaging. I would also like to thank Webster Stayman for many enlightening discussions on modeling, image reconstruction and other interesting ideas. I would like to thank Jerry Prince for serving on my thesis committee.

Additionally, my projects were made possible by many clinical collaborators of the lab: Drs. Shadpour Demehri, Greg Osgood, and Gaurav Thawait. Prototyping of the extremity CBCT is made possible by a team of industry partners at Carestream, including Drs. John Yorkston, Xiaohui Wang, Weidong Huang, Bill Snyder, and many others. The work was also supported by grants NIH 1R01-EB-018896 and NIH 1R21-AR-062293. I would like to thank the Howard Hughes Medical Institute (HHMI) for supporting me through 3 years of my PhD.

I would also like to thank my friends and colleagues. Many have helped me deepen my understanding of the field: Jennifer Xu and Grace Gang, for a wealth of insights into cascaded systems analysis; Ali Uneri, for his speedy GPU projectors and tips on coding in CUDA; Alex Sisniega, for his Monte Carlo scatter simulation and motion correction libraries; Hao Dang and Steven Tilley II, who I have often turned to for questions about reconstruction; Sureerat (Ja) Reaugamornrat, who have helped me with image processing and registration. I need to thank: Niral Sheth, for staying late with me in the lab; Michael Brehler and Sarah Capostagno, for taking me on their morning runs; Tharindu De Silva, for the best Sri Lankan cuisine. I have also benefitted greatly from the company of Pengwei Wu, Wenying Wang, Esme Zhang, Runze Han, Shalini Subramanian, Matt Tivnan, Stephen Liu, Michael Ketcha, as well as visiting scholars Matthew Jacobson, Hao Zhang, Joseph Gorres, Amir Pourmorteza, Ashwin Mathews, Adam Wang, and Shiyu Xu.

Dedication:

For my parents.

Table of Contents

Abstract.....	ii
Dissertation Committee	iv
Acknowledgements.....	v
List of Tables	xiii
List of Figures.....	xiv
List of Abbreviations	xix
Chapter I Cone-Beam Computed Tomography: Systems and Models for Quantitative MSK Imaging.....	1
I.A Quantitative Biomarkers for MSK Imaging	1
I.B Dedicated CBCT Systems for MSK Imaging.....	2
I.B.1 Clinical Realizations	2
I.B.2 X-ray Production and Interaction.....	4
I.C Detector Characterization and Modeling	8
I.C.1 Indirect X-ray Detectors	8
I.C.2 Detector Characterization and Task-Based Assessment of Performance	11
I.C.2.1 Statistical Description of X-ray Detectors	11
I.C.2.2 Task-based Figures of Merit	15
I.C.2.3 Cascaded Systems Modeling	16
I.D Tomographic Reconstruction.....	20
I.D.1 Tomographic Reconstruction as an Inverse Problem	20
I.D.2 Analytical Reconstruction	21

I.D.3 Model-Based Iterative Reconstruction	23
I.D.3.1 Data Fidelity	23
I.D.3.2 Regularization.....	25
I.D.4 Application of MBIR in Quantitative Imaging.....	26
I.E Quantitative Image Analysis	28
I.F Thesis Outline.....	33
 Chapter II An Electrostatics-Inspired Model for Quantification and Analysis of Joint Macrostructure	35
II.A Introduction	35
II.B Methods	37
II.B.1 The Electrostatic Model for Joint Space Analysis.....	37
II.B.2 Application to Knee Joint Morphology	41
II.B.3 Characterization of Knee Osteoarthritis	43
II.C Results	46
II.C.1 Algorithm Performance	46
II.C.2 Joint Space Maps	47
II.C.3 Quantitative Analysis	48
II.D Discussion and Conclusion	52
 Chapter III Polyenergetic MBIR for Improved Quantification of Bone Composition	55
III.A Introduction	55
III.B. Methods	58
III.B.1 Experimental Setup	58
III.B.2 Polyenergetic PL	59
III.B.2.1 Object Model.....	59

III.B.2.2 Model Calibration	62
III.B.2.3 Reconstruction Workflow	64
III.B.3 Reproducibility Study	67
III.B.3.1 Water-Ca Phantom	67
III.B.3.2 Cadaver Specimen.....	68
III.C Results	69
III.C.1 Water-Ca Phantom	70
III.C.2 Cadaver Specimen.....	72
III.D Discussion	74
III.E Conclusion.....	78
Chapter IV Improving System Spatial Resolution for Quantification of Trabecular Microstructure.....	79
IV.A Introduction.....	79
IV.B Methods.....	82
IV.B.1 Task-Based Evaluation of CMOS Detectors	82
IV.B.1.1 Cascaded Systems Model.....	84
IV.B.1.2 System Gain	87
IV.B.1.3 System MTF, Detector Blur, and Focal Spot Blur.....	89
IV.B.1.4 Noise Power Spectrum in Projection Images.....	91
IV.B.1.5 Imaging Task.....	93
IV.B.1.6 Simulation Studies	94
IV.B.2 Benchtop Experimental Setup.....	95
IV.B.2.1 Measurement of Detector MTF, NPS, and Scan Dose ...	96
IV.B.2.2 Cadaver Study	98

IV.B.3 Bone Microstructure Imaging on a Prototype CMOS-based CBCT	99
.....
IV.B.3.1 A Prototype CMOS-Based Extremity CBCT Scanner	99
IV.B.3.2 Performance Evaluation	101
IV.C Results	103
IV.C.1 Detectability Index for Extremity CBCT	103
IV.C.2 Benchtop Experimental Studies	112
IV.C.2.1 Measured MTF and DQE	112
IV.C.2.2 Cadaver Imaging	115
IV.C.3 Bone Microstructure Imaging on a Prototype	118
IV.D Discussion and Conclusion	121
Chapter V A Multiresolution Model for High-Resolution MBIR in CBCT	129
V.A Introduction	129
V.B Methods	132
V.B.1 Multiresolution PWLS	132
V.B.2 Experimental Setup for Simulation Studies	138
V.B.3 Benchtop Study of Anthropomorphic Knee Phantom	142
V.B.4 Analysis of Runtime and Memory Footprint	143
V.C Results	145
V.C.1 Digital Phantom Study	145
V.C.2 Effect of the Location of the Fine Grid ROI	150
V.C.3 Reconstruction of an Anthropomorphic Knee Phantom	151
V.C.4 Computational Cost of Multiresolution PWLS	152
V.D Discussion and Conclusions	153

Chapter VI. Future Work	159
Bibliography	163

List of Tables

Table 1.1 Summary of key interactions relevant to diagnostic x-ray imaging. KE-Kinetic energy free-moving electron. BE-Binding energy of orbital electron. E-energy of incident photon. EM radiation-Electromagnetic radiation.	5
Table 1.2 Noise transfer properties for common stages in cascaded systems analysis....	16
Table 1.3 Summary of a 7-stage cascaded systems model for flat-panel x-ray detectors.	18
Table 2.1 Comparison of diagnostic accuracy (AUC) in discriminating normal and OA subjects. Radiologist reading corresponds to classification based on their manual measurement of the minimum distance in the medial compartment. The second columns shows the mean AUC in the region (± 5 principal components) about the optimal number of principal components for the three JSM methods. The p-value is that resulting from an unpaired heteroscedastic t-test comparing a given JSM method to the Radiologist Reading. For example, the improved mean AUC measured for the electrostatic JSM method was statistically significant ($p = 0.008$), but that for the CP-JSM method was not ($p = 0.071$).....	49
Table 3.1 Summary of water phantom configurations.....	68
Table 3.2 Summary of cadaver phantom configurations.....	69
Table 4.1A Glossary of terms and symbols in the cascaded systems model. Model constants include fundamental physical quantities, geometry and typical operating parameters of the extremity CBCT system, and detector parameters that are independent of CsI:Tl thickness.	85
Table 4.1B CMOS detector performance was analyzed as a function of quantities denoted as key variables. The derived quantities are functions of the model constants and key variables; their nominal values are given at two detector thicknesses corresponding to the CMOS sensors used in experimental studies, assuming all other parameters are at their nominal value.	86
Table 5.1 Glossary of terms and pseudocode of SPS optimization method for multiresolution PWLS. Ordered subsets are omitted for simplification.....	138

List of Figures

Figure 1.1 Commercial realizations of dedicated extremities x-ray CT systems [10], [17]–[20].	3
Figure 1.2 (A) X-ray tube output spectra in extremity CBCT (90 kVp with intrinsic filtration. [24] (B)-(D) Mass attenuation coefficient of CsI:Tl (detector scintillator), soft tissue and bone, along with contributions from each interaction type [25].	7
Figure 1.3 Schematic of an indirect x-ray detector featuring CsI:Tl scintillator and FOP [28].	10
Figure 1.4 Example of joint space width measurement in weight-bearing radiograph and extremity CBCT. Two possible joint space width estimates are shown in each image indicating the need for a more consistent measurement methodology.	30
Figure 1.5 Examples of BMD measurements of lumbar spine with (A) DXA [70] and (B) qCT [71].	31
Figure 1.6 Examples of trabecular metric maps derived from a μ CT image ROI of an ulna biopsy sample. (A) The reconstructed image, (B) segmentation, (C) trabecular spacing map and (D) trabecular thickness map.	32
Figure 2.1 Conventional methods of joint space analysis: (A) Distance along a longitudinal axis. (B) Distance to closest point on articulating surface.	36
Figure 2.2 Electrostatic joint space model. (a) Shows setup of the boundary value problem as well as the resulting field lines. (b) Lengths of field lines mapped onto the tibial surface.....	39
Figure 2.3 Flowchart illustration of the process for JSM calculation from CBCT scans.	42
Figure 2.4 Convergence of the JSM calculation as a function of iterations. The box-and-whisker plots show the median, quartiles, and range in RMSE for all JSMs computed over all 39 cases in the clinical study.....	47
Figure 2.5 Joint space maps computed using the electrostatic model for 39 subjects in the clinical study, grouped as normal and OA. The JSM is displayed as a heat map projected onto surface of the tibial plateau.....	48
Figure 2.6 Distribution of JSW as measured by the electrostatic JSM method, with box-and-whisker plots showing the median, first and third quartiles, and range in JSM for each subject. Also shown (asterisk) is the minimum distance in the medial compartment as measured by 3 MSK radiologists.....	49

Figure 2.7 Principal component weight for the first 24 components analyzed using the electrostatic JSM method.....	51
Figure 2.8 AUC in discriminating normal and OA subjects based on PCA of the JSM, analyzed as a function of the number of principal components. Curves are a simple polynomial fit to guide the eye. The Radiologist reading performance is marked by the mean and range (in gray).	51
Figure 2.9 Sensitivity and specificity curves comparing classification accuracy for the mean radiologist reading of the medial compartment and PCA of the electrostatic JSM, CP-JSM, and LA-JSM using the peak number of principal components from Figure 2.7. For ease of visualization, the points were fitted with ROC curves based on the assumption of binormal distribution [87].	52
Figure 3.1 Top-down view of a benchtop CBCT system emulating extremity CBCT. A water phantom (W5) is shown on the motion stage. The magenta box highlights the collection of Ca inserts used in this reproducibility study.....	59
Figure 3.2 Object model for the water- CaCO_3 binary mixture. Data points represent apparent CaCO_3 fraction for the phantom inserts and common tissue types in extremity CBCT.....	62
Figure 3.3 (A) MC scatter profile and (B) SPR for the same projection of the calibration phantom shown in (C-E). Bottom row, polyPL reconstructions with (C) no scatter correction and (D) no glare correction. (E) full correction.....	67
Figure 3.4 (A) Example reconstructions with FDK initialization (with scaled BMD values) and PolyPL. Yellow squares in W1 FDK image (top left) shows approximate locations of ROIs sampled for analysis. (B) Distribution of percent BMD error with apparent values in ROIs of the 200 mg/cm^3 insert across all configurations.	70
Figure 3.5 Distribution of BMD values of Ca insert ROIs (1620 total ROIs) vs apparent BMD, mean is marked by circles, error bars indicate 2x standard deviation from the mean. Dashed lines indicate apparent densities expected from spectral calibration and represents perfect accuracy.....	71
Figure 3.6 Sample of 2 configurations in the cadaveric knee phantom. C1 contains additional attenuators while C5 does not. The two BMD distributions are visibly similar.	73
Figure 3.7 Distribution of voxel-wise standard deviation across 5 cadaver phantoms in a tibial ROI, with all phantoms were registered to the ROI.	74
Figure 4.1 (A) Quantum detection efficiency, escape efficiency (left vertical axis, black lines), and scintillator gain (right vertical axis, gray line) computed as a function of scintillator thickness. (B) Zero-frequency P_K and integral of detector NNPS over the Nyquist frequency range (gray line) as a function of CsI:Tl thickness. (C) Examples of task functions (Eq. 4.14, assuming $C=1$) for three feature sizes: 0.05 mm (solid black	

line), 0.1 mm (solid dark gray line), and 0.2 mm (solid light gray line). The tasks emphasize distinct frequency bands depending on the underlying feature size.	89
Figure 4.2 CBCT test-bench used in experimental studies. The bench was configured with two CMOS sensors, one with CsI:Tl thickness of 0.4 mm (C400) and one with thickness of 0.7 mm (C700).....	96
Figure 4.3 (A) CMOS-based extremity CBCT system along with details of the CMOS CBCT gantry. (B) Key system and scan parameters for CMOS CBCT and a-Si:H FPD CBCT.	100
Figure 4.4 Task-based evaluation of CMOS detector performance in extremity CBCT. (A) Relative detectability for a range of feature sizes (vertical axis) as a function of pixel size. Scintillator thickness is assumed constant and equal to 0.7 mm. Detectability is normalized to the maximum value for each <i>aobj</i> . Dashed lines indicate maximum d' for each feature size. (B) Ratio of d'^2 achieved with the same scintillator as (A), but at increased electronic noise consistent with an a-Si:H FPD, to maximum d'^2 attained for each <i>aobj</i> by the low-noise CMOS detector of (A) . (C) Relative detectability of the CMOS detector as a function of scintillator thickness and imaging task, normalized by maximum detectability achieved for each <i>aobj</i> across the range of t_{CsI} . Pixel size is 0.099 mm. (D) Joint optimization of pixel size and scintillator thickness for a "trabecular" imaging task with feature size of 0.06 mm. The graph shows detectability of a CMOS detector (normalized by the maximum).....	107
Figure 4.5. (A) Ratio of detectability achieved with the C400 detector to that of C700 as a function of imaging task and imaging dose for the nominal extremity CBCT geometry. Top horizontal axis represents detector entrance dose (after attenuation by a simulated knee), and the bottom horizontal axis gives the corresponding measured central CTDI dose in a CBCT scan. (B) Detectability for the trabecular imaging task (<i>aobj</i> =0.06 mm) as a function of focal spot size and pixel size, normalized by the detectability achieved for each <i>aobj</i> using a 0.5 mm focal spot. C700 detector and nominal extremity CBCT geometry are assumed. (C) Pixel size and scintillator thickness yielding optimal detectability for the trabecular imaging task (<i>aobj</i> =0.06 mm) as a function of magnification and focal spot size. (D) Relative detectability values (normalized to detectability at the nominal CBCT geometry and <i>aspot</i> = 0.3 mm) at the optimal detector configurations found in (C).....	109
Figure 4.6 (A) Experimental measurements of detector MTF for CMOS sensors with 0.7 mm CsI:Tl (open circles) and 0.4 mm CsI:Tl (closed circles). Lines represent MTFs computed using the cascaded systems model. (B) Contact images of the GAMMEX 91437 resolution gauge obtained with 0.7 mm CsI:Tl (left) and 0.4 mm CsI:Tl (right). 113	
Figure 4.7. Measured (points) and simulated (lines) DQE for a range of dose levels for the CMOS detector with (A) 0.7 mm thick scintillator and (B) 0.4 mm scintillator.	115
Figure 4.8. (A) Magnified views of two trabecular regions in the subchondral bone of a cadaver knee imaged using CMOS detectors with different pixel sizes and scintillator thicknesses. High-resolution bone reconstruction was used. (A, top) Reconstructions of	

2x2 binned C700 projections, mimicking the pixel size of current a-Si:H FPDs. (A, middle) Reconstructions of C700 projections in 1x1 binning, showing the benefits of reduced pixel size provided by CMOS. (A, bottom) Images acquired with C400 in 1x1 binning, illustrating the visualization of trabecular detail using a thin scintillator. (B) A complete axial slice of C400 reconstruction obtained using high-resolution protocol (C) A C400 reconstruction obtained using a soft-tissue protocol with 4x4 pixel binning.....	117
Figure 4.9 (A) Detector MTF measurements. (B) Line profile through FBP reconstructions of a tungsten wire.	118
Figure 4.10 Reconstructions of a cadaveric ankle on a-Si:H FPD and CMOS CBCT systems in (A) high-resolution protocol and (B) high-contrast protocol.....	119
Figure 4.11 Reconstructions of the ulna (1st column) and sample trabecular ROIs with overlaid segmentations.....	120
Figure 4.12 (A) Distribution of Dice coefficient between the CBCTs and μ CT ROI segmentations and (B)(C)(D) correlation of trabecular metrics obtained with CBCT and μ CT. Dashed opaque lines denote confidence interval corresponding to $\alpha=0.05$. Grey solid line has unit slope and zero intercept (equivalence line representing absolute agreement).....	121
Figure 5.1 Schematic of multiresolution forward projection with detector binning. The estimated bone region parameterized using fine grid voxels (μF) is marked with a black dashed line. This region is projected onto native detector pixels (marked with lN) using projection operator AFN . The line integrals captured by the native projection pixels also include contributions from a subset of coarse voxels (μC), denoted by projection operator ACN . Line integrals for binned projection pixels (marked as lB) traverse only through coarse grid voxels, computed using projector operator ACB	135
Figure 5.2 Schematic of multiresolution regularization scheme illustrating how the regularization over boundaries between the fine and coarse voxel grids is performed using interpolation operators.....	137
Figure 5.3 Schematic of digital bone phantom with the boundary of a central fine voxel grid region marked with a thick dashed line. The ROIs used for measurement of artifacts in the fine grid region (μart) and noise ($\mu noise$) in the coarse grid region are indicated with boxes.	140
Figure 5.4 Convergence profile measured in reconstructions of the digital phantom. The estimate at the current iteration is compared with a converged image at 200 iterations. $\beta F = 103.5$; $\beta C = 104.5$. The RMSD curves for the fine grid region (dashed line) overlap for all values of $\eta \mu$	146
Figure 5.5 Multiresolution PWLS reconstructions for different downsampling factors and coarse grid regularization values. A central region of the digital phantom is shown for PWLS without detector binning (A) and with detector binning (B). The last column in each subfigure is a difference image between the case with maximal downsampling	

($\eta\mu = 10$) and the case of no downsampling ($\eta\mu = 1$). Arrows indicate artifacts due to strong coarse grid downsampling and regularization.148

Figure 5.6 (A) Magnitude of the artifact in fine grid ROI (RMSE of μ_{art} in Fig. 5.3, given by Eq. 5.11) as a function of noise in the coarse grid region. The noise metric quantifies the cumulative effect of regularization and voxel downsampling in μ_C . (B) The artifact ROI (μ_{art}) in the fine grid region for a range of downsampling factors (rows) and normalized coarse grid regularization strengths β (columns).150

Figure 5.7 Multiresolution PWLS with fine grid ROI conforming to bone boundaries (marked with dashed line in the leftmost mage) for three values of the normalized penalty strength β . The downsampling factor $\eta\mu$ is set to 10. The right-most subfigure shows a difference image between the downsampled case ($\beta = 106.5$) and the reference reconstruction with no downsampling. Projection binning is included.151

Figure 5.8 Multiresolution PWLS reconstruction of an anthropomorphic knee phantom acquired on a CBCT benchtop. Fine and coarse grid regions are delineated in A. B-D shows details of the high-resolution trabecular ROI in the sagittal, axial, and coronal planes. Boundaries of the ROI are marked with a dashed line.152

Figure 5.9 (A) Measured reconstruction speedup of multiresolution PWLS as a function of the size of fine grid FOV (for total volume size of 120x120x120 mm³). (B) Total memory consumption of downsampled cases compared to fine grid cases.153

List of Abbreviations

APS	active pixel sensor
a-Si:H	hydrogenated amorphous silicon
AUC	area-under-curve (of ROC)
BMD	bone mineral density
BvTv	bone volume fraction
CBCT	cone-beam computed tomography
CCD	charge-coupled device
CMOS	complementary metal-oxide semiconductor
CNR	contrast-to-noise ratio
CP	closest-point (JSW measurement method)
CPD	coherent-point drift (registration algorithm)
CSA	cascaded systems analysis
cSi	crystalline silicon
CsI:Tl	thallium-doped cesium iodide
CTDI	CT dose index (phantom or dose measurement)
CV	coefficient of variation
DQE	detective quantum efficiency
ESF	edge spread function
FBP	filtered backprojection
FDK	Feldkamp-Davis-Kress (reconstruction method)
FOP	fiber optic plate

FOV	field-of-view
FPD	flat-panel detector
FTCS	forward-time central-space (numerical derivative)
GPU	graphics processing unit
HR-pQCT	high-resolution peripheral quantitative computed tomography
HU	Hounsfield unit
HVL	half-value layer
JS	Joseph-Spital (beam-hardening correction method)
JSM	joint space map
JSW	joint space width
LA	longitudinal axis (JSW measurement method)
LSF	line spread function
MBIR	model-based iterative reconstruction
MC	Monte Carlo (scatter simulation)
MDCT	multidetector computed tomography
MSK	musculoskeletal
MTF	modulation transfer function
NNPS	normalized noise power spectrum
NPS	noise power spectrum
NPW	non-prewhitening
OA	osteoarthritis
OP	osteoporosis
OS	ordered subsets (optimization technique)

PCA	principal component analysis
PL	penalized likelihood (reconstruction method)
polyPL	polyenergetic penalized likelihood (reconstruction method)
PPS	passive pixel sensor
PSF	point spread function
PW	pre-whitening
PWLS	penalized weighted least-squares (reconstruction method)
QDE	quantum detection efficiency
RBF	radial basis function (kernel)
RMSE	root-mean-square error
RMSD	root-mean-square difference
ROC	receiver operating characteristic
ROI	region-of-interest
SAD	source-axis distance
SDD	source-detector distance
SNR	signal-to-noise ratio
SPR	scatter-to-primary ratio
SPS	separable parabolic surrogates (optimization algorithm)
SVM	support vector machine
TbSp	mean trabecular spacing
TbTh	mean trabecular thickness
TFT	thin film transistor

Chapter I

Cone-Beam Computed Tomography: Systems and Models for Quantitative MSK Imaging

The goal of this thesis is to develop new systems, models, and analysis methods that advance the spatial resolution and quantitative imaging capabilities of cone-beam computed tomography (CBCT). The work focuses on applications to imaging of the musculoskeletal (MSK) extremities to yield rich volumetric characterization of bone morphology, accurate determination of bone composition and enable high-resolution visualization of bone microstructure. In this chapter, we review the current state-of-the-art in dedicated CBCT systems for MSK imaging, the physics underlying the modality, prior work on modeling the physics, and algorithms for tomographic reconstruction.

I.A Quantitative Biomarkers for MSK Imaging

Development of quantitative MSK imaging is impactful for diagnosis and monitoring of many bone infirmities and abnormalities. It is estimated that 52.5 million U.S. adults aged ≥ 18 y suffer from some form of arthritis or related rheumatic condition [1]. Such prevalence places extensive demand on tools for management and monitoring of the disease. In recent years, novel therapeutic agents have emerged to treat arthritis and alter the progression of disease in the affected joints. The promise of such disease-modifying osteoarthritic drugs (DMOADs) [2] and antirheumatic drugs (DMARDs) [3] motivates the development of new methodologies to characterize subtle changes in biomarkers of joints and bones and provide reliable assessment of therapy response.

Prevalence statistics for osteoporosis (OP) is equally compelling. Recent estimates place the number of patients with OP and low bone mass at 53.6 million in 2010 [4], accounting for 54% of older adults in the US. The increased risk of fragility fractures in this population poses a significant healthcare burden.

On a macroscopic scale, quantitative imaging entails the characterization of bone and joint morphology. One example is weight-bearing joint space narrowing in osteoarthritis (OA), which correlates with cartilage degeneration and is used for OA diagnosis and staging [5]. Measurements of bone composition, specifically bone mineral density (BMD), are key for diagnosis of osteoporosis [6], assessment of fracture risk and monitoring of fracture healing [7]. On a microscopic scale, evidence suggests that changes in trabecular microstructure in OA precedes cartilage loss and the onset of symptoms [8], [9], providing a promising window for early diagnosis.

I.B Dedicated CBCT Systems for MSK Imaging

Advancements in spatial resolution and quantitative imaging capability introduced in this work are based on optimization of new detector technologies and modeling of imaging physics. These developments are made in the context of extremity CBCT [10], [11]. In this section, we introduce several clinical realizations of the modality and discuss the x-ray production and interaction physics underlying their operation.

I.B.1 Clinical Realizations

Dedicated x-ray CT systems have recently been developed to provide convenient logistics for streamlined orthopedic examinations. Figure 1.1A-C shows several CBCT systems for extremities imaging. A key feature for these systems is the ability to provide

3D evaluation of weight-bearing joints. Conventionally, such evaluation was only possible with 2D radiography [12]–[14]. Applications of weight-bearing imaging include assessment of OA, patellar malformation [15] and flatfoot deformity [16]. In this work, “extremity CBCT” refers to the Carestream OnSight 3D system shown in Fig 1.1C, unless otherwise specifically indicated. The scan dose for the extremity CBCT is 5-15 mGy, the field-of-view (FOV) is 20x20x20 cm³, and typical scan time is ~30 s [10]. Spatial resolution permits visualization of ~0.25-0.35 mm high-contrast features (frequency at which the modulation transfer function (MTF) declines to 10%, f₁₀, is approximately 1.5-1.8 mm⁻¹). A compact gantry with flexible positioning capability permits both weight-bearing as well as non-weight-bearing imaging. The extremity CBCT scanner operates in a circular orbit, with a source-detector-distance (SDD) of ~56 cm, rigidly rotating around the patient extremity in a 210° arc, capturing 420-600 projections.

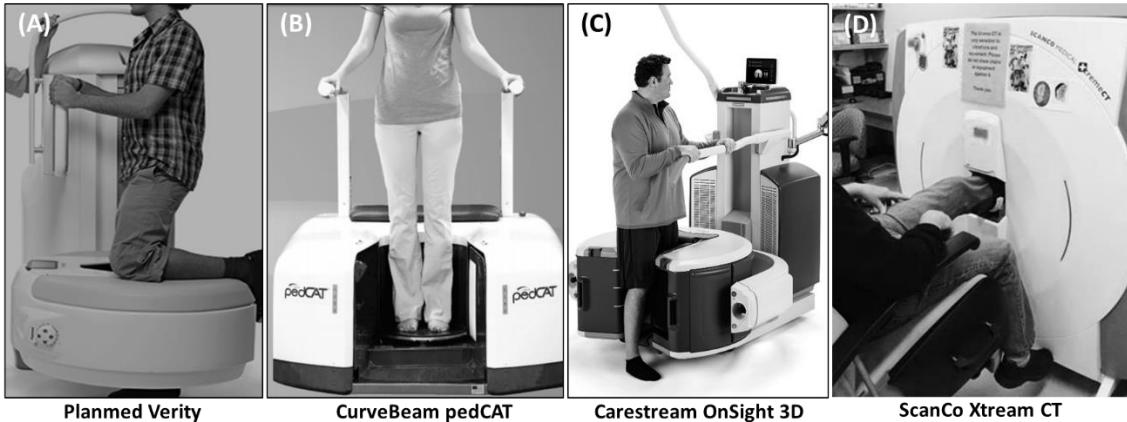


Figure 1.1 Commercial realizations of dedicated extremities x-ray CT systems [10], [17]–[20].

The SCANCO XtreamCT (Fig 1.1D) is representative of high-resolution peripheral quantitative CT (HR-pQCT) systems, which are not considered CBCT systems but are well-represented in MSK research. These systems emphasize high-resolution acquisition

(~95 um) for bone microstructure analysis but are limited in FOV and scan time. Additionally, these systems do not offer weight-bearing configurations.

I.B.2 X-ray Production and Interaction

Table 1.1 summarizes particle and photon interactions [21] relevant to diagnostic x-ray imaging (20-150 kV). These interactions are the basis of x-ray production, interaction with biological tissue, and capture by the detector. In each interaction, radiation (either a free moving electron or an x-ray photon) interacts with a target (orbital electrons or nucleus), producing freed electrons or emitting x-rays.

X-rays are produced in x-ray tubes by electrons from the cathode that are accelerated in vacuum by an electric field and interact with atomic nuclei of the target (anode). In a *bremsstrahlung* interaction, the cathode electrons are decelerated by the electric field of anode nuclei, emitting energy in the form of x-ray photons. The resulting spectrum (roughly representing the number of emitted x-rays per energy) linearly decreases with energy, with the maximum equal to the accelerating potential. Some low energy photons are also absorbed in the tungsten target and in other components in the x-ray tube (such as Cu or Al filters) after the interaction, resulting in a unimodal signature bremsstrahlung spectra (Fig. 1.2A) [22]. The x-ray spectrum also includes distinct peaks due to direct interactions between the electrons impinging the target and orbital electrons of the target nuclei. The x-ray beam is attenuated when propagated through biological tissue via photoelectric absorption, Compton scatter and Rayleigh scatter, and finally absorbed in the detector scintillator (CsI:Tl, see next section) predominately through the photoelectric and Compton effects.

Name	Radiation	Target	Outcome
Bremsstrahlung	e^-	Nucleus	Ionizing EM radiation (x-ray)
Rayleigh (Coherent) Scatter	Low E	e^- cloud	Scattered photon $E_{SC}=E$
Compton (Incoherent) Scatter	$E \gg BE$	Valence e^-	Scattered photon and ejected e^- where $E = E_{SC} + KE_{e^-}$ (assume BE small)
Photoelectric Absorption	$E \geq BE$	Orbital e^-	Photoelectron with $E_{PE}=E-BE$ and characteristic x-ray or Auger e^-

Table 1.1 Summary of key interactions relevant to diagnostic x-ray imaging. KE-Kinetic energy free-moving electron. BE-Binding energy of orbital electron. E-energy of incident photon. EM radiation-Electromagnetic radiation.

X-ray attenuation through an object of uniform material and density is described by

Beer's Law:

$$I(E) = I_0(E) \exp\left(-l\rho\left(\frac{\mu}{\rho}\right)_{tot}(E)\right) \quad (1.1)$$

where E denotes x-ray energy, $I_0(E)$ and $I(E)$ are the input and output x-ray spectra, l is the path integral of the x-ray beam through the object [mm], ρ is the density of the object [g/mm^3], and $\left(\frac{\mu}{\rho}\right)_{tot}(E)$ is the total mass attenuation of the material [mm^2/g]. The transmitted beam I is also known as the primary beam. The total mass attenuation of a material is the sum of contributions of each interaction type:

$$\left(\frac{\mu}{\rho}\right)_{tot}(E) = \left(\frac{\mu}{\rho}\right)_{PE}(E) + \left(\frac{\mu}{\rho}\right)_{Compton}(E) + \left(\frac{\mu}{\rho}\right)_{Rayleigh}(E) \quad (1.2)$$

Mass attenuation spectra for soft tissue and bone as well as a typical scintillator used in x-ray detectors (x-ray-to-light converter) are plotted in Fig. 1.2 [23]. For biological

tissue, attenuation is dominated by the photoelectric effect at low energies and the Compton effect at high energies. For CsI:Tl scintillators, since photoelectric absorption accounts for almost all of mass attenuation spectra in the diagnostic energy range (with photoelectric mass attenuation >2 orders of magnitude higher than that of Compton or Rayleigh). Given the polyenergetic attenuation described in Eq. 1.1 and decreasing mass attenuation spectra for biological tissue shown in Fig. 1.2CD, the mean energy of the beam increases as it propagates through tissue, an effect termed *beam hardening*. Thus, the same object may attenuate the x-ray beam differently depending on its position and surrounding attenuators.

Knowledge of the output spectra of the x-ray source is important for modeling of detector performance as well as for polyenergetic reconstruction algorithms. Throughout this work, half-value layer (HVL) measurements of Al and Cu were used to infer the x-ray source spectra. As described in IV.B.1, a solid-state exposure diode (Radcal Corp, Monrovia, CA, USA) was placed at the detector. The exposure was measured with successive layers of Al or Cu placed at the collimator of the x-ray source. The measurements were normalized by the exposure in air (without Al and Cu) to a 0-1 range. The thickness of Al or Cu which reduces the diode reading by half (HVL_{Al} or HVL_{Cu}) was found by interpolating the curve of normalized exposure with respect to mm Al or mm Cu and finding the intersection with 0.5. For modeling and reconstruction, simulated TASMICs spectra were generated via *Spektr* [24] and filtered with additional Al, Cu and W (representing inherent filtration) according to Beer's law until the HVLs computed from the simulated spectra matched that of the physical measurement. In III.B.2.2, the raw detector signal and estimated detector absorption spectra were used instead of dose diode measurements.

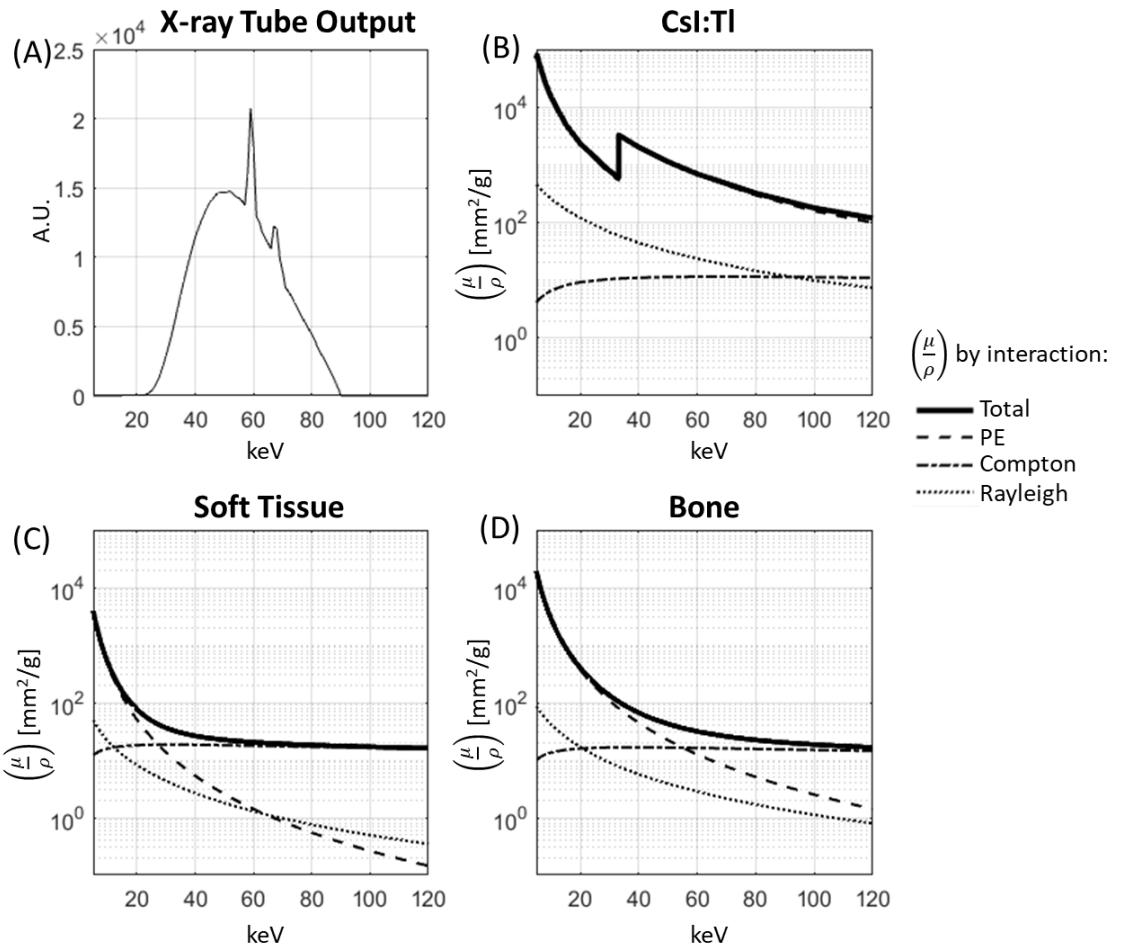


Figure 1.2 (A) X-ray tube output spectra in extremity CBCT (90 kVp with intrinsic filtration. [24] (B)-(D) Mass attenuation coefficient of CsI:Tl (detector scintillator), soft tissue and bone, along with contributions from each interaction type [25].

In extremity CBCT, the distance between the patient and the detector (air gap) is small. Thus, a significant portion of scattered photons are also captured in addition to the primary beam (see Fig. 3.3A). The signal from scattered photons, if uncorrected, leads to underestimated attenuation in the final reconstructed image. For a large object with high primary attenuation, the scatter-to-primary ratio (SPR) can be >1 . Detailed discussion of our correction approach can be found in III.B.2.3.

With respect to the process of x-ray detection using scintillator detectors, we note that the mass attenuation spectra for CsI:Tl exhibits several peaks in the diagnostic x-ray energy range. This arises from increases in photoelectric attenuation when x-ray energy surpasses the binding energy of an orbital shell (K, L, M, ...), allowing the x-rays to interact with electrons occupying the higher binding-energy, lower-level shell. Characteristic x-rays can be produced through this interaction. The K-shell energy of Cs is 35.98 keV. The characteristic x-ray produced from K-shell interaction is called *K-fluorescence*. The peak in mass attenuation corresponding to this energy is known as the *K-edge* of Cs. K-edges enhance the detection efficiency of the detector, but the emitted K-fluorescence photons can disperse laterally within the scintillator and be absorbed remotely, contributing to additional blur.

We do not explicitly consider the effects of electron transport (excitation and ionization) downstream of photon interactions, though this has been studied by Hajdok *et al.* [26], [27] in high-energy applications.

I.C Detector Characterization and Modeling

A major part of this thesis (Chapter IV) concerns the modeling and optimization of new detector technology to improve spatial resolution for bone microstructure characterization. Here we review principles underlying our detector modeling.

I.C.1 Indirect X-ray Detectors

X-ray detectors can be considered direct or indirect. In the first approach, the detector material converts absorbed x-rays directly to electron-hole pairs that are the source of the electric signal of the sensor. In indirect detection, x-rays are first absorbed in a

scintillator and converted to visible light photons (also known as secondary quanta), which are then captured by a photodiode. Most modern CBCT systems utilize a variety of indirect digital x-ray flat-panel detectors (FPDs). These sensors offer fine pixel size (150 μm and less), enabling high-resolution clinical and preclinical imaging at a level of detail not accessible to conventional multidetector CT (MDCT). Shown in Fig. 1.3, in this class of detectors, x-rays are first absorbed by a high-density scintillator material, converted to visible light which is then absorbed by a photodiode substrate. In contemporary detectors, the photodiode is typically hydrogenated amorphous silicon (a-Si:H), charge-coupled devices (CCDs) or more recently, complementary metal-oxide-semiconductors (CMOS). The scintillator usually consists of Thallium-doped Cesium Iodide (CsI:Tl) or Gadolinium oxysulfide ($\text{Gd}_2\text{O}_2\text{S:Tb}$), when paired with Si-based photodiodes. Systems shown in Fig 1.1A-C are all based on amorphous silicon (a-Si:H) digital flat-panel detectors (a-Si:H FPD), while Fig 1.1D is based on CCD technology.

The Carestream OnSight 3D system (Fig 1.1C) uses a Varex PaxScan2530 FPD (Varex, Salt Lake City UT) with 0.139 mm pixel pitch and 0.7 mm CsI:Tl scintillator thickness. Visible light photons generated in the scintillator can diffuse laterally before reaching the photodiode, blurring the output image. One benefit of the columnar needle structure of CsI:Tl (Fig. 1.3) is that the diffusion is reduced due to internal reflection within the needles. An in-depth discussion of this process can be found in Chapter IV.

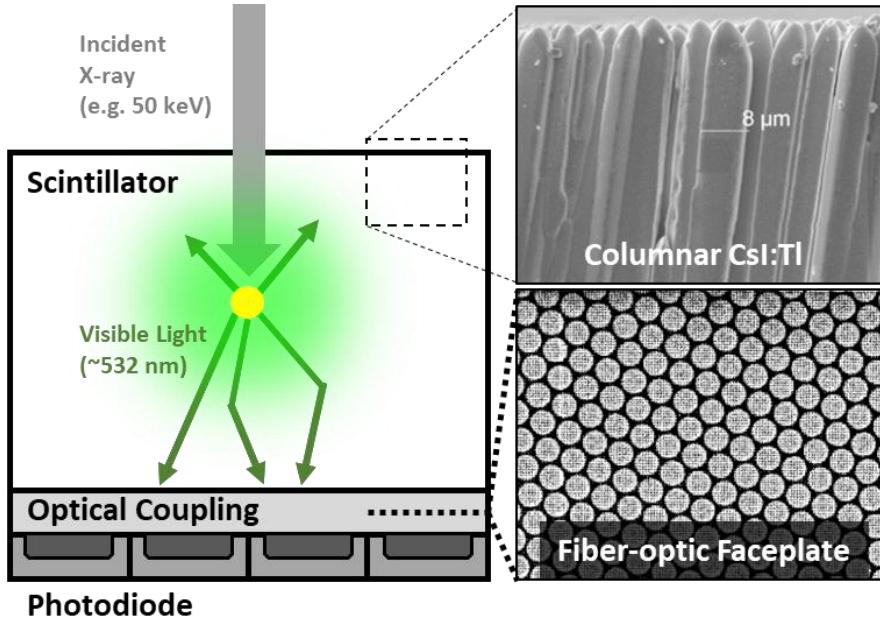


Figure 1.3 Schematic of an indirect x-ray detector featuring CsI:Tl scintillator and FOP [28].

In recent years, CMOS detectors have become an attractive option for use in indirect x-ray detectors. In contrast to conventional a-Si:H FPDs, CMOS sensors are manufactured on wafers of crystalline Si (cSi), which offer higher mobility for charge carriers [29]. Electron mobility is around $1400 \text{ cm}^2/\text{Vs}$ for cSi but $<1 \text{ cm}^2/\text{Vs}$ for a-Si:H. The rapid mobility translates to substantially faster (10x) detector readout speeds for CMOS compared to a-Si:H-based sensors. The higher mobility of cSi also leads to smaller transistor sizes and smaller pixels ($\sim 3 \mu\text{m}$), whereas a-Si:H FPD uses thin film transistors (TFT) limited to $50 \mu\text{m}$. CMOS detectors also exhibit lower noise than conventional a-Si:H FPDs. During detector readout in a-Si:H FPD, charge is transferred from the pixel to the dataline and converted to voltage via charge amplifiers at the end of the dataline. During this transfer, pixel noise is amplified by the ratio of dataline capacitance and pixel capacitance. This design is referred to as a passive pixel sensor (PPS), resulting in noise of 1000-2000 e⁻. CMOS detectors are typically active pixel sensors (APS), with charge

amplifiers integrated into each pixel. As a result, voltage is read out on the dataline instead of charge, mitigating the noise amplification problem in conventional a-Si:H PPS circuits.

In some modern CMOS FPDs, the circuitry on the cSi substrate is protected with a fiber optic plate (FOP). The FOP is doped with lead and intercepts x-ray photons not absorbed in the scintillator. In many cases, the columnar CsI:Tl is directly deposited onto the FOP, which is then coupled to the photodiode via an optical adhesive. The FOP is typically absent in conventional a-Si:H FPD due to the superior radiation hardness of amorphous silicon.

I.C.2 Detector Characterization and Task-Based Assessment of Performance

This section discusses how x-ray detectors can be modeled statistically and evaluated objectively, using conventional metrics of noise and resolution as well as with task-based metrics such as detectability.

I.C.2.1 Statistical Description of X-ray Detectors

Image acquisition by an x-ray detector can be described as [30], [31]:

$$g = \mathcal{H}_D f(x, y) + n \quad (1.3)$$

where random variable $g = [g_1 \dots g_M]^T$ is the vectorized form for a discrete output image with M pixels, x and y are spatial coordinates, $f(x, y)$ is a continuous function representing the 2D projection of the object attenuation as the input signal, and $n = [n_1 \dots n_M]^T$ is the vectorized form for pixel noise. The imaging operator for the detector \mathcal{H}_D defines contribution of input signal at point (x, y) to each output pixel in g . Note that the additive form for noise n does not limit our analysis to additive noise, since noise is the deviation of an instance of g from its expectation \bar{g} :

$$n = \bar{g} - g \quad (1.4)$$

Additionally, we also note that randomness in the image g results not only from measurement noise n but also variability of the imaged object f . However, in this work we consider only cases in which f is deterministic. This case is referred to as signal/background known-exactly (SKE/BKE).

If \mathcal{H}_D is *linear and shift-invariant*, the pixel output can be represented by system gain G and point spread function (*PSF*):

$$g_m = G \iint PSF(x, y)f(x + x_m, y + y_m)dx dy + n_m \quad (1.5)$$

where (x_m, y_m) is the location of output pixel m . Additionally, we let $\iint PSF(x, y)dx dy = 1$. Noise n is considered *stationary* if its mean and autocovariance are constant throughout the image. Note that the noise term – in particular, noise correlation – is typically also affected by PSF acting on the noise distribution of the input signal f .

The linear shift-invariance assumption greatly simplifies analysis of x-ray detectors by allowing the use of Fourier domain metrics. Specifically, blur and noise can be described by the modulation transfer function (MTF) and noise power spectrum (NPS), respectively. Most detectors are not linear or shift-invariant. For example, nonlinearity may occur at signal levels approaching the noise floor or saturation. On some systems, the projection of the focal spot of the x-ray source may vary significantly across the detector. Despite this, systems can often be approximated as linear and shift invariant in nominal operating conditions at the center of FOV.

The MTF is a frequency-domain decomposition of signal transfer characteristics and an indicator of system spatial resolution. The MTF can be derived from the PSF via:

$$MTF(u, v) = |\mathcal{F}\{PSF(x, y)\}| \quad (1.6)$$

where u and v are spatial frequencies. The Fourier transform is denoted by $\mathcal{F}\{\cdot\}$. However, this prescriptive definition is difficult to determine from measurement, since $PSF(x, y)$ is the system response to a delta function. In practice, a 1D MTF can be measured from the edge spread function (ESF) or line spread function (LSF). The LSF is the Radon transform of the PSF. The ESF is the integral of the LSF. Denoting polar coordinates in spatial and frequency domains as $(x, y) = (r \cos \theta, r \sin \theta)$ and $(u, v) = (\rho \cos \theta, \rho \sin \theta)$, the LSF and ESF at an angle θ are given by:

$$LSF_\theta(r) = \iint PSF(x, y) \delta(x \cos \theta + y \sin \theta - r) dx dy \quad (1.7.1)$$

$$LSF_\theta(r) = \int_{-\infty}^r LSF_\theta(r') dr' \quad (1.7.2)$$

where the $\delta(\cdot)$ is the Dirac delta function and $\delta(x \cos \theta + y \sin \theta - r)$ represents a line impulse. The 1D MTF corresponding to angle θ can be computed as:

$$MTF_{1D}(\rho) = \left| \frac{\mathcal{F}\{LSF_\theta(r)\}}{\int LSF_\theta(r') dr'} \right| \quad (1.8)$$

The denominator ensures $MTF(0) = 1$. The ESF is typically imaged from a sharp edge, such as a tungsten plate. When the edge is placed on the detector surface, the detector MTF is obtained; when placed elsewhere, the MTF contains magnification effects that may include the detector MTF and the x-ray focal spot MTF, the product of which is the system MTF for a particular system geometry. A detailed description of the edge method for detector MTF measurements can be found in Samei et al [32].

The NPS is a frequency-domain decomposition of variance in an image containing only stochastic fluctuations, such as image fluctuations owing to the Poisson-distributed

variation in the number of photons incident on a given pixel (also known as quantum noise). It can be defined as the Fourier transform of the autocovariance and applies to stationary noise. Practically, when considering N realizations of zero mean, noise only image $n(x, y)$, the NPS is computed as:

$$NPS(u, v) = \frac{1}{NA} \sum_N |\mathcal{F}\{n(x, y)\}|^2 \quad (1.9)$$

where A denotes the area in which the NPS is computed. The units of a 2D NPS are [signal]²[length]². For 3D NPS, A would take on units of volume. A property of the NPS is that the frequency-domain integration results in the variance of the noise-only image:

$$\iint NPS(u, v) dudv = \sigma^2 \quad (1.10)$$

In practice, the 2D NPS is computed from ROIs cropped from homogeneous areas of the projection image. In some cases, the normalized noise power spectrum (NNPS) is used:

$$NNPS(u, v) = \frac{NPS(u, v)}{(q_0 G)^2} \quad (1.11)$$

where q_0 is the input signal at the detector, and G is the system gain. Together, the denominator denotes the large area signal.

Resolution and noise characteristics described by MTF and NPS can be combined into the detective quantum efficiency (DQE):

$$DQE = \frac{(SNR)_{out}^2}{(SNR)_{in}^2} = \frac{\frac{q_0^2 G^2 MTF^2}{NPS}}{q_0} = \frac{MTF^2}{q_0 NNPS} \quad (1.12)$$

where SNR_{in} and SNR_{out} are signal-to-noise ratios of the input and outputs. This equation assumes that incident photons are Poisson-distributed ($SNR_{in} = \sqrt{q_0}$). Note that DQE is dimensionless and bounded between 0 and 1

I.C.2.2 Task-based Figures of Merit

Objective assessment of detector performance requires definition of imaging tasks and observers who perform these tasks, in addition to statistical descriptions of systems and images. While DQE represents frequency-dependent modulation in SNR, it does not describe performance relative to a specific imaging task. A *task* is the intended use of an image. In this work, we consider tasks as binary classification problems, in which an observer needs to decide between two hypotheses f_0 and f_1 based on image g . In frequency-domain, the task can be represented as a task-function:

$$W_{task} = |\mathcal{F}\{f_1 - f_0\}| \quad (1.13)$$

The two hypotheses may represent, for example, tumor-absent vs tumor-present. In this work, the two hypotheses take the form of small-feature vs large-feature. A detailed discussion of the task functions used in this work is described in section IV.B.1.

Two simple models of observer performance include the prewhitening (PW) and the nonprewhitening (NPW) model. Though these models do not rely on the assumption of linearity and shift-invariance, they have convenient representations using Fourier domain metrics:

$$d'_{PW}^2 = \iint \frac{MTF^2 W_{task}^2}{NNPS} dudv \quad (1.14)$$

$$d'_{NPW}^2 = \frac{(\iint MTF^2 W_{task}^2 dudv)^2}{\iint MTF^2 W_{task}^2 NNPS dudv} \quad (1.15)$$

Derivations for these indices can be found in ICRU report 54 [33] and the work of Wagner [34]. In the prewhitening model, division by NNPS decorrelates the noise. The two models are equivalent when the image contains only white noise. Other observers build upon these models to incorporate anatomical noise and models of human vision [35], [36].

I.C.2.3 Cascaded Systems Modeling

Signal and noise transfer through an imaging system can be described by cascaded systems analysis (CSA) [37]. In CSA, the image formation process is abstracted into various stages [38]. Each stage has a unique resolution and noise transfer characteristic. These stages are combined sequentially or in parallel to obtain aggregate gain and MTF for the imaging system as well as the output NPS.

The basic stages are summarized in Table 1.2. For clarity, the MTF for each stage is denoted T_i . The input and output image NPS for each stage are denoted S_i and S_o , respectively. Gain associated with each stage is denoted g . Mean and variance of gain are denoted \bar{g} and σ_g^2 , respectively. Mean input fluence is \bar{q}_i .

Stage Type	Noise Transfer
Gain	$S_o = \bar{g}S_i + \sigma_g^2\bar{q}_i$
Deterministic Spread	$S_o = S_i T_i^2$
Stochastic Spread	$S_o = (S_i - \bar{q}_i)T_i^2 + \bar{q}_i$
Sampling	$S_o = S_i * III_i$

Table 1.2 Noise transfer properties for common stages in cascaded systems analysis.

Similar to spread stages, gain stages can be either stochastic or deterministic depending on σ_g . A deterministic gain stage would require $\sigma_g^2 = 0$; a Poisson-distributed

gain stage would have $\sigma_g^2 = \bar{g}$; a gain stage corresponding to binomial selection would possess $\sigma_g^2 = \bar{g}(1 - \bar{g})$.

The frequency comb function III referenced in the sampling stage has the form:

$$III(u, v) = \sum_{i,j=-\infty}^{\infty} \delta(u - \frac{i}{a_x}, v - \frac{j}{a_y}) \quad (1.16)$$

where δ is the Dirac delta function, a_x and a_y are sampling intervals in spatial domain. From linear systems theory, MTF for sequential stages are multiplied, while MTF for stages connected in parallel are linearly combined, weighted by signal of the respective branches.

A CSA model of a-Si:H FPD is described by Siewerdsen *et al* [39]. This is a 7-stage model incorporating physical parameters of the detector, yielding detector gain and MTF as well as output NPS given Poisson-distributed x-ray photons as input. The 7 stages are connected sequentially. This model is summarized in Table 1.3:

Stage	Stage Type	Description
Poisson Input q_0	Poisson-Distributed X-rays at Detector	Poisson-distributed photons from x-ray tube attenuated via binomial selection in the patient
Quantum Detection Efficiency (QDE) g_1	Binomial Selection	Fraction of x-ray photons absorbed in scintillator
Quantum Gain g_2	Stochastic Gain	Optical photons generated per x-ray photon interacting
Scintillator Blur T_3	Stochastic Spread	Optical photon spread in scintillator
Coupling Efficiency g_4	Deterministic Gain	Fraction of optical photons lost from scintillator to photodiode
Pixel Aperture T_5	Deterministic Spread	Integration of optical photons on pixel photodiode
Pixel Sampling III_6	Sampling	Sampling - results in NPS aliasing
Readout Noise S_{add}	Additive Noise	Noise from readout circuitry

Table 1.3 Summary of a 7-stage cascaded systems model for flat-panel x-ray detectors.

Stages $g_1 - T_3$ describe the scintillator while stages $T_5 - S_{add}$ characterize the photodiode, with g_4 at the interface between the two components. The readout noise S_{add} also includes transistor and readout electronics (including amplifier and ADC) in addition to the photodiode. We assume the input to the detector is Poisson-distributed x-ray photons with mean fluence \bar{q}_0 [x-rays/mm²] as well as variance \bar{q}_0 [(x-rays/mm²)²] (note that units of variance is mean signal squared). Carrying out the algebra of CSA according the transfer characteristics of each stage, we arrive at the system gain G , MTF, output image NPS and DQE:

$$G = a_{pd}^2 \bar{g}_1 \bar{g}_2 \bar{g}_4 \quad (1.17.1)$$

$$MTF = T_3 T_5 \quad (1.17.2)$$

$$NPS = a_{pd}^4 \bar{q}_0 \bar{g}_1 \bar{g}_2 \bar{g}_4 (1 + \bar{g}_4 (\bar{g}_2 + \epsilon_{g_2}) T_3^2) T_5^2 * III + S_{add} \quad (1.17.3)$$

$$DQE = \frac{a_{pd}^4 \bar{q}_0 \bar{g}_1 \bar{g}_2^2 \bar{g}_4^2 T_3^2 T_5^2}{a_{pd}^4 \bar{q}_0 \bar{g}_1 \bar{g}_2 \bar{g}_4 (1 + \bar{g}_4 (\bar{g}_2 + \epsilon_{g_2}) T_3^2) T_5^2 * III + S_{add}} \quad (1.17.4)$$

where a_{pd} is the dimension of active photodiode within a pixel. The quantum gain stage g_2 is defined by a Poisson excess ϵ_{g_2} parameter: $\sigma_{g_2}^2 = \bar{g}_2(\epsilon_{g_2} - 1)$, which can be interpreted as deviation from Poisson-distribution.

An equivalent representation for Poisson excess is the Swank factor [40], conventionally denoted by I :

$$I = \frac{M_1}{M_0 M_2} = \frac{1}{1 + \frac{\epsilon_{g_2} + 1}{\bar{g}_2}} \quad (1.18)$$

where $M_0 = 1$, $M_1 = \bar{g}_2$ and $M_2 = \bar{g}_2^2 + \sigma_{g_2}^2$ are the zeroth, first, and second moments of g_2 . Detailed discussion of Swank factors for CsI:Tl scintillators can be found in the works of Zhao *et al.* [41] and Lubinski *et al.* [42]. In the absence of additive noise (S_{add}) and blurring ($T_3, T_5 = 1$) also assuming perfect x-ray absorption ($g_1 = 1$) and pixel sampling (no aliasing), the analytical form of DQE shown in Eq. 1.17.4 reduces to the Swank factor, corresponding to Swank's original interpretation.

The cascaded systems model was extended by Yao and Cunningham [43] to describe K-fluorescence in the scintillator via parallel cascades. Vedantham *et al.* [44] later reformulated the first two stages g_1 and g_2 to consider depth-dependent optical photon loss within the scintillator. We base our analysis of CMOS detectors on this updated model, described in detail in Chapter IV. Tward *et al.* [45] further extended the cascaded systems

model to 3D image reconstruction by filtered backprojection, but this is not considered in this work.

I.D Tomographic Reconstruction

The previous section described how projection data are acquired. In this section, we discuss algorithms for image reconstruction.

I.D.1 Tomographic Reconstruction as an Inverse Problem

Tomographic reconstruction methods are generally classified as analytical or iterative, beginning with filtered backprojection (FBP) [46] and algebraic reconstruction technique (ART) [47], respectively. Both were developed in parallel starting from the 1970s. Due to speed considerations, FBP is still a workhorse in commercial CT systems. However, whereas analytical reconstruction typically relies on idealized mathematical models for image formation, models of imaging physics are more easily incorporated into iterative methods. Physical models may include nonidealities such as system blur, spectral effects, non-standard geometries, etc. Iterative methods incorporating these models are often termed statistical iterative reconstruction (SIR) or MBIR. This section focuses on MBIR algorithms.

To describe the reconstruction problem, we extend the statistical description of x-ray detectors (Eq 1.3) to also consider the physics associated with x-ray projections of a 3D object:

$$g = \mathcal{H}f + n \quad (1.20)$$

As before, $g = [g_1 \dots g_M]^T$ is a vectorized representation of projection data (e.g. pixel values on x-ray detector); $f = [f_1 \dots f_N]^T$ is a vectorized discrete representation of the

imaged object (3D distribution of x-ray attenuation or material density); $n = [n_1 \dots n_M]^T$ is the noise vector, and \mathcal{H} is the imaging operator. In this work, f is always assumed to be parametrized by voxels, though many other basis representations are proposed, such as wavelets and spherical elements [48]. The *noise model* is the distribution of noise vector n , and as mentioned in I.C.2, is not limited to additive noise. The *forward model* describes how object f produces measurement g , this includes a model of the imaging operator \mathcal{H} and the noise model. Tomographic reconstruction is the *inverse problem* of estimating the object f from measurement g . All reconstruction algorithms, analytical or iterative, assume some form of forward model and noise model, either implicitly or explicitly.

I.D.2 Analytical Reconstruction

We give a brief description of the idealized model used for analytical reconstruction in 2D. Consider a continuous 2D axial attenuation image $\mu(x, y)$ with $(x, y) \in \mathbb{R}^2$ and a set of line integral measurements $l(s, \theta)$ related by:

$$l(s, \theta) = \iint_{-\infty}^{+\infty} \mu(x, y) \delta(x \cos \theta + y \sin \theta - s) dx dy \quad (1.23)$$

where s and θ denote position on the detector and projection angle, respectively. The Dirac Delta function $\delta(\cdot)$ is applied to integrate attenuation values in $\mu(x, y)$ along the line $s = x \cos \theta + y \sin \theta$. According to the Fourier slice theorem [49], the Fourier transform of the line integral measurement at θ equals the 1D slice of the Fourier transform of the attenuation image:

$$\mathcal{F}_s\{l(s, \theta)\} = \mathcal{F}_{xy}\{\mu\}(r \cos \theta, r \sin \theta) \quad (1.24)$$

where \mathcal{F}_s denotes the 1D Fourier transform along s , $\mathcal{F}_{xy}\{\mu\}$ denotes the 2D Fourier transform of attenuation image $\mu(x, y)$, r denotes the radial coordinates in spatial domain along angle θ . Thus, if a sufficient number and angular distribution of line integral measurements are known, then according to (1.24), the image μ can be obtained by reconstructing $\mathcal{F}_{xy}\{\mu\}$ from a series of slices and applying the 2D inverse Fourier transform.

This is the basis of the filtered backprojection algorithm, where the line integral measurements are filtered by a ramp $|r|$:

$$\mu(x, y) = \int_0^{2\pi} \int_0^\infty \mathcal{F}_{xy}\{\mu\}(r \cos \theta, r \sin \theta) e^{2\pi j r(x \cos \theta + y \sin \theta)} r dr d\theta \quad (1.25a)$$

$$= \int_0^\pi \int_{-\infty}^{+\infty} \mathcal{F}_s\{l(s, \theta)\} |r| e^{2\pi j r(x \cos \theta + y \sin \theta)} dr d\theta \quad (1.25b)$$

In practice, the ramp filter is multiplied by a low-pass apodization filter to reduce high-frequency noise. An analogue for volumetric reconstruction from 2D projections is the Feldkamp, Davis and Kress (FDK) algorithm [50], which contains additional weights to account for fan and cone angles. FDK provides fast baseline reconstructions with well-defined noise and resolution properties and is used in Chapter V. The algorithm implements row-wise ramp filtering, with added Hann apodization to control noise-resolution trade-off. Additionally, Parker weighting [51] was introduced to account for short scans (210° arc) generated by extremity CBCT prototypes.

I.D.3 Model-Based Iterative Reconstruction

Generally, MBIR is formulated as a maximum a posteriori estimation problem [52], [53]:

$$\hat{f} = \arg \max_f P(f|g) \quad (1.26.1)$$

$$= \arg \max_f P(g|f)P(f) \quad (1.26.2)$$

with f and g defined in Eq. 1.20. Bayes' Theorem is used to derive Eq. 1.26.2 from Eq. 1.26.1, discarding the marginal likelihood (independent of f). Distribution $P(f|g)$ is the *posterior*. Distribution $P(g|f)$ is the *likelihood function* (denoted by $L(y|\mu)$ in the examples below) and $P(f)$ is the image *prior*. The forward model and noise model are described by the likelihood function, thus the expression associated with the likelihood is also called the *data fidelity* term. In this work, the prior always takes the form of *regularization* (denoted by $R(\mu)$). However, one could incorporate other information about patient anatomy such as a prior image [54]. For simplicity, the examples throughout this section assume a monoenergetic case, in which the object f is a distribution of linear attenuation coefficients $\mu = [\mu_1 \dots \mu_M]^T$. Also, by convention, projection measurements are denoted as y instead of g .

I.D.3.1 Data Fidelity

In the forward model, the geometry of x-ray projection is described by the system matrix $A [M \times N]$, also referred to as the forward projection operator or projector. The elements of A , a_{ij} , describe the contribution of attenuation voxel μ_j to its projected line integral l_i :

$$l = A\mu \quad (1.21)$$

According to Beer's law, the mean measurement on the detector is given by:

$$\bar{y}(\mu) = I \exp(-A\mu) \quad (1.22)$$

where I is the intensity of ray without attenuation by the object. The transpose of the system matrix, A^T , is called the backprojection operator or backprojector. In practice, the system matrix is often too large to store in memory and is computed on the fly. Many methods for computation of A and A^T were proposed [55]–[58], with various tradeoffs of speed vs accuracy. Throughout this work, the separable footprints (SF) forward projector and the corresponding backprojector are used [58] for MBIR. In SF, projection of a voxel on the detector (“voxel footprint”) is approximated as a separable 2D trapezoid function. This approximation is relatively accurate compared to the previously cited alternatives, while separability allows integration of the footprint on the detector to be carried out efficiently. The projectors were implemented on graphical processing units (GPUs) using CUDA.

A commonly used noise model is the Poisson distribution with independent measurements:

$$y \sim \text{Poisson}\{\bar{y}\} \quad (1.27)$$

The likelihood term L based on the Poisson noise model is:

$$L(y|\mu) = \prod_{i=1 \dots M} \frac{\exp(\bar{y}_i) \bar{y}_i^{y_i}}{y_i!} \quad (1.28)$$

Since L is difficult to compute, instead of maximizing $L(y|\mu)$, the Poisson likelihood objective seeks to minimize $-\log L(y|\mu)$:

$$\hat{\mu} = \arg \min_{\mu} -\log L(y|\mu) \quad (1.29.1)$$

$$= \arg \min_{\mu} \sum_{i=1 \dots M} -y_i \log \bar{y}_i(\mu) + \bar{y}_i(\mu) \quad (1.29.2)$$

Here the term $y_i!$ is omitted since it is independent of μ .

Another useful noise model [59] assumes line integrals (note that this is different from the Poisson model above, which concerned detector intensity measurements) are Gaussian-distributed with diagonal covariance matrix of inverse projection measurements:

$$l_i \sim \mathcal{N}(\bar{l}_i, \frac{1}{y_i}) \quad (1.30)$$

Following the same derivation as above, the objective is:

$$\hat{\mu} = \arg \min_{\mu} -\log L(l|\mu) \quad (1.31.1)$$

$$= \arg \min_{\mu} \|l - A\mu\|_y^2 \quad (1.31.2)$$

Note that least squares are weighted by the projection data y , thus the objective is also called weighted-least-squares (WLS).

If the least squares are not weighted by projection data y , the solution is equivalent to the analytical Feldkamp-David-Kress (FDK) method [50] described above.

I.D.3.2 Regularization

Minimization of these objectives mentioned above will typically yield noisy images. This is because tomographic reconstruction is highly ill-conditioned. A standard technique is to apply regularization (an image prior in the formulation of Eq. 1.26) to enforce a desired texture in the final reconstruction. In this work, we only consider smoothness penalty terms in the form of:

$$R(x) = \frac{1}{2} \beta \sum_{j \in 1 \dots N} \sum_{k \in \mathcal{N}_j} \omega_{jk} \psi(x_j - x_k) \quad (1.32)$$

where β controls regularization strength, k is the index of neighboring voxels of x_j , ω_{jk} weights the contribution of the k th neighbor to j , and ψ is a potential function. In this work, \mathcal{N}_j always refers to the 6-connected neighborhood in 3D, with $\omega_{jk} = 1$. Two potential functions are used in this work. The first is a quadratic penalty (used in Chapter V):

$$\psi(x) = \frac{1}{2} x^2 \quad (1.33)$$

The second is the Huber penalty [60] (used in Chapter III)

$$\psi(x) = \begin{cases} x^2/2\delta, & |x| \leq \delta \\ |x| - \delta/2, & |x| > \delta \end{cases} \quad (1.34)$$

where the penalty is quadratic within $[-\delta, \delta]$, and linear elsewhere. Since a quadratic function penalizes large differences asymptotically more than a linear function, the Huber penalty enhances edge preservation for $|x| > \delta$.

I.D.4 Application of MBIR in Quantitative Imaging

Despite different image parametrizations, forward models, noise models and penalties used throughout this work, MBIR objectives in this dissertation have the form of:

$$\hat{f} = \arg \min_f -\log(P(g|f)P(f)) \quad (1.35.1)$$

$$= \arg \min_\mu -\log L(y|\mu) + R(\mu) \quad (1.35.2)$$

In this subsection, we review several applications of MBIR to quantitative imaging.

As described in Section I.B.2, CBCT images can be marred by beam hardening and scatter artifacts. In Chapter III, we apply the Poisson noise model (Eq. 1.27) with a forward

model incorporating basis material densities (which we call polyPL) [61], [62] for estimation of BMD. The forward model for polyPL is:

$$\bar{y}(\rho) = \int I_0(E) \exp\left(-\sum_{k=1...K} A\left(\frac{\mu}{\rho}\right)_k f_k(\rho)\rho\right) dE + r \quad (1.36)$$

where the unknown vector ρ describes the total density for each voxel, vector f_k describes the material fraction of the k th basis material for each voxel and $\left(\frac{\mu}{\rho}\right)_k(E)$ the mass attenuation of the k th basis material, r is x-ray scatter. Note that this forward model is parametrized with respect to density ρ instead of linear attenuation coefficient μ . Beam hardening artifacts are effectively mitigated if a suitable set of basis materials and material fraction functions f_k are selected. In Chapter III, we show that this forward model results in accurate and reproducible BMD in phantoms and cadaver scans.

It is worth noting that neither noise models mentioned in I.D.3.1 are physically accurate, despite their success in many reconstruction applications. As evident in our detector modeling work (Section I.C.2), the actual noise properties, just from the noise transfer characteristics of the detector alone, can be more complex than described above. Moreover, noise is also dependent on the object attenuating the beam. More sophisticated models, such as ones with correlated noise [63], [64], have been shown to further improve image quality at the cost of added computational complexity.

In Chapter V, we address issues in applying MBIR to high-resolution imaging systems such as CBCT incorporating CMOS detectors. In these cases, MBIR may not be feasible owing to slow projection/backprojection operators involving large system matrices. In addition, reconstruction may be hampered by limited memory. This is because MBIR requires the full support of the object to be modeled, not only the region-of-interest.

This is resolved by using a multiresolution scheme for the forward model and volume parametrization. High-resolution features such as trabecular bone are parametrized with fine voxels while the rest of the image is downsampled from native resolution to a level suitable for fast MBIR. Though PWLS is used in the chapter, multiresolution reconstruction is not limited to any specific forward model, noise model or penalty.

The multiresolution scheme can be particularly useful for high-fidelity forward models with added computational complexity. One example is a model incorporating shift-variant focal spot by Tilley *et al.* [64] mentioned previously. This is implemented with specialized projectors (e.g. “sourcelets”) which are more compute-intensive than standard SF projectors. For high-resolution data, multiresolution scheme could be applied such that only a selected ROI is modeled with these specialized projectors, reducing reconstruction time [65].

I.E Quantitative Image Analysis

The previous sections introduced imaging physics, detector modeling and reconstruction algorithms. To conclude this chapter, we introduce three examples of how high-resolution, weight-bearing CBCT and model-based reconstruction algorithms can enhance image analysis and improve estimation of biomarkers. These examples are developed and used in Chapters II, III and IV.

The first is characterization of weight-bearing joint space morphology (Fig. 1.4), a key biomarker for diagnosis and staging in OA indicative of articular cartilage loss. The measurement concerns the distance between the medial femoral condyle and medial tibial plateaus. A comparison of such a measurement in weight-bearing radiograph and extremity CBCT is shown below. A volumetric analysis provided by extremity CBCT should be able

to better delineate cortical bone surfaces by removing superpositioned anatomy. In addition, the reconstructed image in CBCT is not affected by perspective distortion, possibly making the measurement more consistent between operators. However, measurement of joint space in 3D is still a challenging problem. Due to the complex morphology of the tibial and femoral plateaus, degenerate points and asymmetric measurements are created when the closest-distance metric is applied from one surface to the other. When considering a projected distance along an axis, the selection of a consistent axis is difficult.

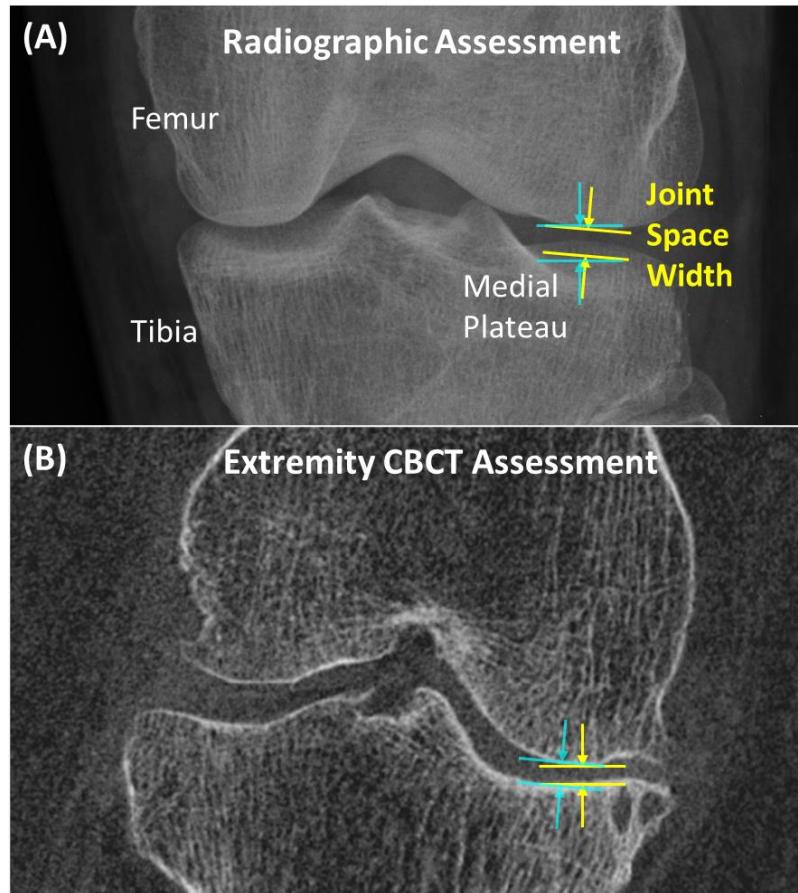


Figure 1.4 Example of joint space width measurement in weight-bearing radiograph and extremity CBCT. Two possible joint space width estimates are shown in each image indicating the need for a more consistent measurement methodology.

A second example is the estimation of BMD, pertinent to diagnosis and monitoring of osteoarthritis (OA) [66], osteoporosis (OP) [67] and bone healing. Conventionally, BMD is assessed via dual energy x-ray absorptiometry (DXA) or quantitative CT (qCT) [68], [69]. DXA is a 2D modality that measures areal BMD from low dose x-ray projections at two energy levels. Because of tissue superposition in radiographic views, areal BMD is dependent on size of the bone and on the composition of the surrounding tissues, and thus exhibits poor consistency. In contrast, qCT provides a 3D measurement derived from MDCT. This yields volumetric BMD, which are based on Hounsfield (HU) units calibrated with hydroxyapatite reference phantoms placed close to the patient. These calibrations

are patient and position-dependent due to physical artifacts caused by beam-hardening and scatter. In CBCT, HU values from such calibrations are less reproducible than in qCT primarily due to enhanced scatter.

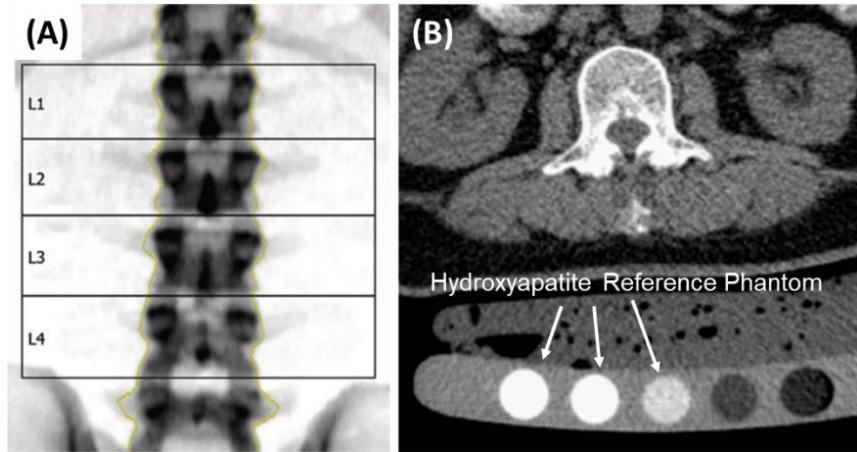


Figure 1.5 Examples of BMD measurements of lumbar spine with (A) DXA [70] and (B) qCT [71].

A third example is characterization of trabecular microstructure, which requires the capability to visualize features $<100 \mu\text{m}$. This resolution is beyond most conventional MDCT systems. The gold standard for imaging of bone microstructure is microCT (μCT), which is limited to *ex vivo* applications. The figure below shows derivation of trabecular spacing and trabecular thickness maps, which are two examples of microstructure metrics. An ROI is first cropped from the reconstructed image and the bone is segmented. For trabecular spacing, maximal spheres are fit between trabecular bone, each voxel is assigned the diameter of the maximal fitting sphere that includes it. Mean trabecular spacing (TbSp) is the average trabecular spacing over all non-trabecular bone voxels [72]. Mean trabecular thickness (TbTh) derived in the same way on the complementary volume. A simplified 2D case is illustrated in Fig. 1.6B, in which both the yellow and blue circles are the maximum fitting circles in the marrow and in the trabecular bone with radius Tb_s and Tb_t ,

respectively. In this 2D case, all pixels inside the yellow circle are assigned trabecular spacing of Tb_s and the pixels inside the blue circle are assigned trabecular thickness of Tb_t . Maps shown in Fig. 1.6CD are computed volumetrically, with fitting of maximal spheres instead of circles. Currently, the ability to perform absolute measurements of microstructure parameters *in vivo* is hindered by noise and resolution limitations of CT systems. However, it is possible to achieve adequate correlation of CBCT-derived trabecular metrics with μ CT gold standard [73], [74]. Further improvement to the microstructure imaging capability of extremity CBCT will offer unique insights into bone remodeling and is key to understanding OA and OP.

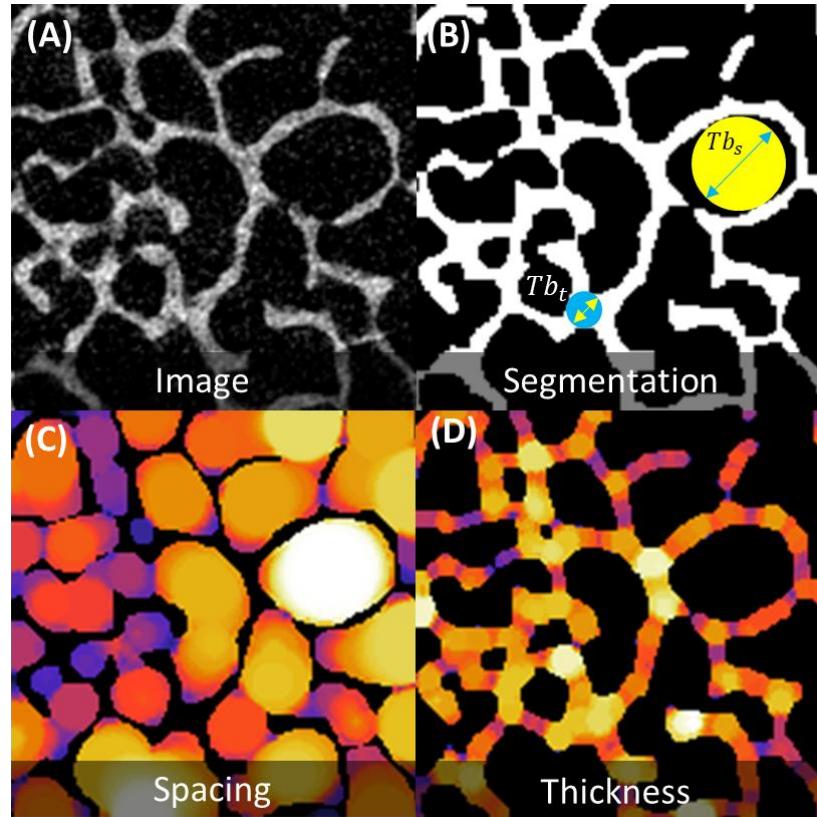


Figure 1.6 Examples of trabecular metric maps derived from a μ CT image ROI of an ulna biopsy sample. (A) The reconstructed image, (B) segmentation, (C) trabecular spacing map and (D) trabecular thickness map.

I.F Thesis Outline

Chapter II describes a novel physics-inspired method for characterizing joint space morphology with 3D image data obtained from extremities CBCT. A version of this chapter is published in:

- Cao, Q., Thawait, G., Gang, G., Zbijewski, W., Riegel, T., Brown, T., Corner, B., Shadpour, D., Siewerdsen, J.H. (2015). Characterization of 3D joint space morphology using an electrostatic model (with application to osteoarthritis). *Physics in Medicine and Biology*, 60, 947–960.

Chapter III demonstrates how BMD can be accurately and reproducibly estimated in extremity CBCT by integrating MBIR with additional artifact correction to account for imaging physics and system imperfections involved in image acquisition. The contents of this chapter can be found in:

- Cao, Q., Stayman, J. W., Siewerdsen, J. H., Zbijewski, W. (2020). Reproducible Bone Mineral Density Estimation in Extremity Cone-beam CT using Polyenergetic Model-based Reconstruction. (in preparation).
- Cao, Q., Sisniega, A., Stayman, J. W., Yorkston, J., Siewerdsen, J. H., Zbijewski, W. (2019). Quantitative cone-beam CT of bone mineral density using model-based reconstruction. *Proc. SPIE*, 10948.

Chapter IV shows how high-resolution imaging of bone microarchitecture can be achieved with the adoption of a custom CMOS detector resulting from task-based optimization incorporating a thin scintillator. This chapter assembles:

- Cao, Q., Sisniega, A., Brehler, M., Stayman, J. W., Yorkston, J., Siewerdsen, J. H., Zbijewski, W. (2018). Modeling and evaluation of a high-resolution CMOS detector for cone-beam CT of the extremities. *Medical Physics*, 45(1), 114–130.
- Cao, Q., Brehler, M., Sisniega, A., Tilley, S., Shiraz Bhurwani, M. M., Stayman, J. W. Zbijewski, W. (2018). High-resolution extremity cone-beam CT with a CMOS detector: Evaluation of a clinical prototype in quantitative assessment of bone microarchitecture. *Proc. SPIE*, 10573. Houston.

Chapter V highlights challenges associated with reconstructing high-resolution data and proposes a multiresolution scheme for model-based iterative reconstruction. This chapter is based on the publication:

- Cao, Q., Zbijewski, W., Sisniega, A., Yorkston, J., Siewersen, J. H., Stayman, J. W. (2016). Multiresolution iterative reconstruction in high-resolution extremity cone-beam CT. *Physics in Medicine and Biology*, 61, 7263–7281.

Chapter II

An Electrostatics-Inspired Model

for Quantification and Analysis of Joint Macrostructure

This chapter describes a novel method for producing joint space maps using volumetric imaging data. The method models the bones of a joint as plates of a capacitor, using the length of simulated electrostatic field lines between the articular bone surfaces as surrogate metric of joint space width. The result is a consistent, parameter-free method that fully samples both articular bone surfaces symmetrically. The method is shown to provide a potentially valuable quantitative predictor of OA.

II.A Introduction

Joint morphology is an important aspect of image-based diagnosis and management of osteoarthritis. Traditional semi-quantitative image-based measures include the Sharp-Larsen score [75], which grades the integrity of the joint space and presence or absence of erosions evident in a plain radiograph. In assessment of knee osteoarthritis, the joint space width (JSW) (i.e., distance between the femur and tibia) is a quantitative measure used by clinicians in diagnosis of degenerative disease. Accordingly, narrowing of JSW demonstrated in weight-bearing radiography can be used to assess the severity and progression of degenerative diseases of the knee. However, such measures carry a basic limitation associated with radiographic imaging (i.e., reducing 3D information to a plane) and are subject to variations in both patient setup and reader repeatability [76].

Modern 3D imaging techniques (including ultrasound, MR, CT, and tomosynthesis) enable more refined tomographic measures of joint space narrowing, such as a joint space map (JSM) [77], which characterizes the JSW across the articular surface. The high isotropic spatial resolution of extremities CBCT presents an opportunity for quantitative characterization of 3D joint space morphology beyond that of conventional measures and potentially more sensitive to subtle disease or treatment response. Possible methods to compute JSMs from such images include projection along a longitudinal axis (LA) of the joint (Fig. 2.1A) or the distance between closest points (CP) on articular surfaces (Fig. 2.1B). However, each of these methods entails a potential limitation: in the former, the definition of a LA is somewhat arbitrary (dependent on patient orientation) and is difficult to extend to multi-axis joints; in the latter, the measurement undersamples the surfaces at protrusions and concavities and is asymmetric from one articular surface to the other and vice versa. Unless otherwise stated, in the following sections, JSM refers to the electrostatic JSM resulting from the electrostatic model of section II.B.1. The maps resulting from the alternative LA and CP methods are denoted LA-JSM and CP-JSM, respectively.

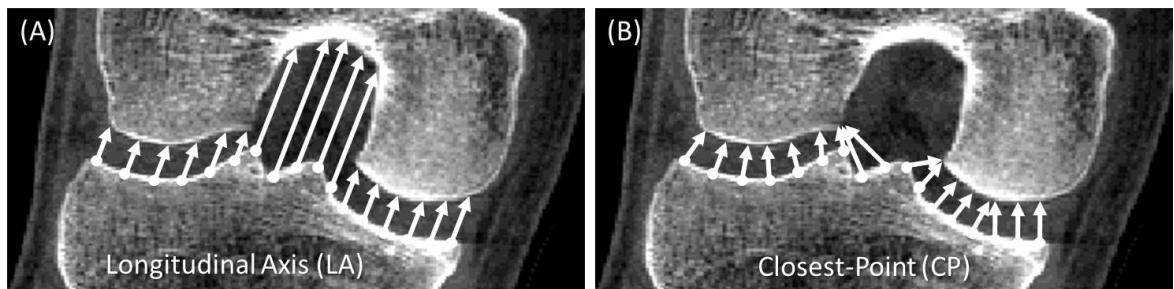


Figure 2.1 Conventional methods of joint space analysis: (A) Distance along a longitudinal axis. (B) Distance to closest point on articulating surface.

In this work, we describe a method for characterization of 3D joint space morphology that overcomes many of the basic limitations of the measures noted above.

The method is similar to that previously proposed to measure the thickness of tissues in the brain [78]. As detailed below, the method envisions the joint as an electrostatic capacitor bounded by equipotential surfaces (two or more bones). Elementary electrostatics yields a unique solution to the electric field lines within the capacitor (the intra-articular space) and a robust measure of joint space morphology. As a test case, we apply the method to the task of discriminating normal and OA subjects based on CBCT images of their weight-bearing knee.

II.B Methods

II.B.1 The Electrostatic Model for Joint Space Analysis

In the electrostatic model, articular surfaces of a synovial joint are conceptualized as conductors with different uniform charge densities, resulting in an electric potential at each boundary, analogous to the plates of a capacitor. Rather than using the Euclidean distance defined along an axis [LA in Fig. 2.1(A)] or closest point pairs [CP in Fig. 2.1(B)], the lengths of electric field lines between the two surfaces are used as a surrogate JSW within an unambiguous, non-degenerate curvilinear space defined by the joint space morphology. This method has a number of potential advantages: (i) it does not depend on an arbitrary axis as a reference direction; (ii) it is symmetric with respect to each of the articulating surfaces; and (iii) it densely samples each point on all surfaces.

Computationally, JSMs can be obtained by first finding the electric potential distribution within a ROI containing the joint space, taking its gradient to find the electric field, and computing the line integral along the field lines. We denote the volumetric ROI about the joint space as $\Omega \subset \mathbb{R}^2, \mathbb{R}^3$ and assume n articulating surfaces

denoted $\partial_1\Omega, \dots, \partial_n\Omega$. The electric field and electric potential in Ω are \vec{E} and Φ , with properties defined by elementary electrostatics:

$$\vec{E} = -\nabla\Phi \quad (2.1)$$

and Gauss' law:

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (2.2)$$

Conventionally, ϵ_0 denotes the permittivity of free space and ρ the charge density. Each surface $\partial_0\Omega, \partial_1\Omega, \dots, \partial_n\Omega$ is assigned a different charge density $\rho_0, \rho_1, \dots, \rho_n$, respectively. Alternatively, a particular surface of interest can be assigned charge density ρ_1 and all others ρ_0 . Substitution of (2.1) into (2.2) gives the Laplace equation

$$\nabla^2\Phi = 0 \quad (2.3)$$

Equation (2.3) can be solved subject to boundary conditions $\Phi(\partial_0\Omega) = V_0, \dots, \Phi(\partial_n\Omega) = V_n$, which are the electric potentials arising from the assigned charges. Finding Φ is a Dirichlet problem that has numerous well-documented solutions [79]. Once Φ is determined on Ω , \vec{E} can be found from (2.1), and the field line lengths are analyzed by line integrals between points on the boundaries $\partial_0\Omega \cup \dots \cup \partial_n\Omega$.

As shown in Fig. 2.2(A), we assigned $\Phi(\partial_1\Omega) = V_1$ and $\Phi(\partial_2\Omega) = V_2$ as the potential on two articulating surfaces of the knee (femur and tibia, respectively) and introduced $\Phi(\partial_0\Omega) = V_0$ on the boundary of the ROI so that Ω has compact support. The model is capable of handling multi-axis joints with any number of articulating surfaces, such as the wrist or ankle, by assigning a unique charge density to each surface and treating as a multi-conductor system, or by assigning one bone of interest (e.g., the scaphoid) a different charge from all other bones of the joint.

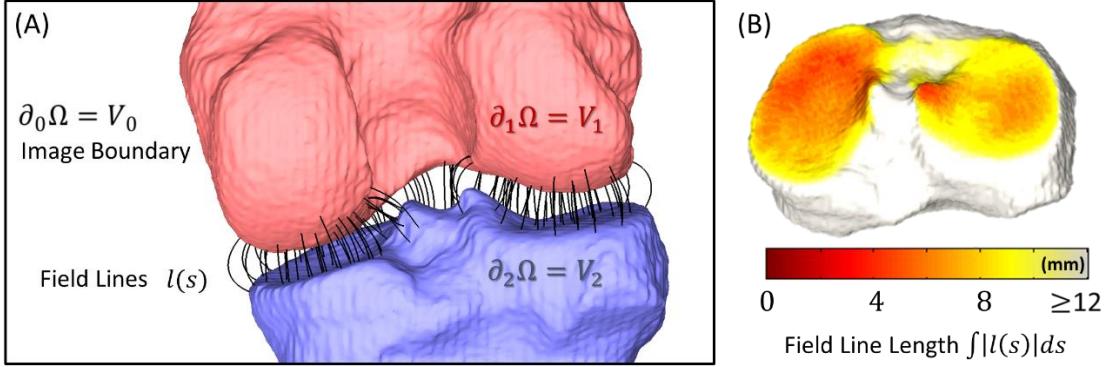


Figure 2.2 Electrostatic joint space model. (a) Shows setup of the boundary value problem as well as the resulting field lines. (b) Lengths of field lines mapped onto the tibial surface.

The electric potential \$\Phi\$ in the joint space \$\Omega\$ can be approximated using the Jacobi method over the Cartesian grid with nodes corresponding to the center of voxels. The basic formulation of the Jacobi method involves representing the partial differential equation (PDE) as a diffusion problem. In our case, the Laplace equation (2.3) can be solved for \$\Phi\$ by writing:

$$\frac{\partial \Phi}{\partial t} = \nabla^2 \Phi \quad (2.4)$$

where time, \$t\$, is analogous to the number of iterations. The boundary conditions, here also used as the initial condition, are set to the respective charges on \$\partial_1\Omega\$ and \$\partial_2\Omega\$. As \$t \rightarrow \infty\$, the initial conditions relax to the equilibrium solution satisfying \$\frac{\partial \Phi}{\partial t} = 0\$, yielding the solution \$\Phi\$. According to the forward-time central-space (FTCS) scheme, equation (4) can be discretized as:

$$\Phi_{(x,y,z)}^{n+1} - \Phi_{(x,y,z)}^n = v(\Delta\Phi_{(x,y,z)}^n) \quad (2.5)$$

where \$n\$ is the number of iterations and \$\Delta\Phi_{(x,y,z)}\$ is the discrete Laplacian:

$$\begin{aligned} \Delta\Phi_{(x,y,z)} &= \Phi_{(x+1,y,z)} + \Phi_{(x-1,y,z)} + \Phi_{(x,y+1,z)} + \Phi_{(x,y-1,z)} + \Phi_{(x,y,z+1)} \\ &\quad + \Phi_{(x,y,z-1)} - 6\Phi_{(x,y,z)} \end{aligned} \quad (2.6)$$

The parameter v governs the rate of relaxation, in this work set to $\frac{1}{6}$, which is theoretically the maximum step size that yields a stable solution [80]. The resulting discrete form for (2.5) is the iterative relation:

$$\begin{aligned} & \Phi_{(x,y,z)}^{n+1} \\ = & \frac{\Phi_{(x+1,y,z)}^n + \Phi_{(x-1,y,z)}^n + \Phi_{(x,y+1,z)}^n + \Phi_{(x,y-1,z)}^n + \Phi_{(x,y,z+1)}^n + \Phi_{(x,y,z-1)}^n}{6} \end{aligned} \quad (2.7)$$

According to equation (2.7), each voxel in the next iteration is the mean of its neighboring voxels in the current iteration. This form can be conveniently implemented as a convolution of a mean filter with the entire segmented image and resetting boundary conditions after each iteration. For the knee model in Fig. 2.2, voxels on $\partial_1\Omega$ and $\partial_2\Omega$ were assigned to +1 V and -1 V, respectively. Since approximating (nearly) infinite boundary conditions carries a large computational load (with little effect within the ROI of the joint space), we assume that the ROI of the joint space is far from the image border ($\partial_0\Omega$), so the potential field can be approximated accurately by zero padding.

From the resulting Φ , the electric field, \vec{E} , is solved via equation (2.1). The origin of each field line is arbitrary and nominally taken to be the voxels on surfaces $\partial_1\Omega$ and $\partial_2\Omega$. The length of each field line is then computed using Euler's method with linear interpolation between voxels from $\partial_1\Omega$ to $\partial_2\Omega$ at fixed step size d :

$$\vec{l}(s_{i+1}) = \vec{l}(s_i) + d \cdot \frac{\vec{E}(\vec{l}(s_i))}{|\vec{E}(\vec{l}(s_i))|} \quad (2.8)$$

for $i = 0, 1, \dots, N$, $\vec{l}(s_i)$ is a vector within the space Ω along the electric field at point s_i , and s_0 and s_N are points on the articulating surfaces. The JSM is then represented as a matrix with each element corresponding to the field line length.

The convergence of the solution was analyzed in terms of the RMSE of the map at the k^{th} iteration (JSM_k) from the converged JSM (JSM_c), where JSM_c was defined as the map computed from a very large number of iterations (25,000) in which field line lengths were changing negligibly (much less than the voxel size) with further iterations. The error at the k^{th} iteration was computed as:

$$\text{Error } (k) = \sqrt{\frac{\sum_{x \in JSM_k \cap JSM_c} (JSM_k(x) - JSM_c(x))^2}{\dim(JSM_k \cap JSM_c)}} \quad (2.9)$$

where $x \in JSM_k \cap JSM_c$ denotes voxels common to both the k^{th} and c^{th} iteration, and \dim is the number of voxels in that intersecting set.

II.B.2 Application to Knee Joint Morphology

The electrostatic model was employed to study knee joint morphology in images acquired using the extremity CBCT scanner as in Fig. 2.1. As detailed below and illustrated in Fig. 2.3, the process included steps associated with cropping to a ROI about the joint space, semiautomatic bone segmentation, calculation of the electric field using the Jacobi method and Gauss' Law, and calculation of field line lengths via Euler's method to yield a dense 3D point cloud labeling the JSW at each point on each surface.

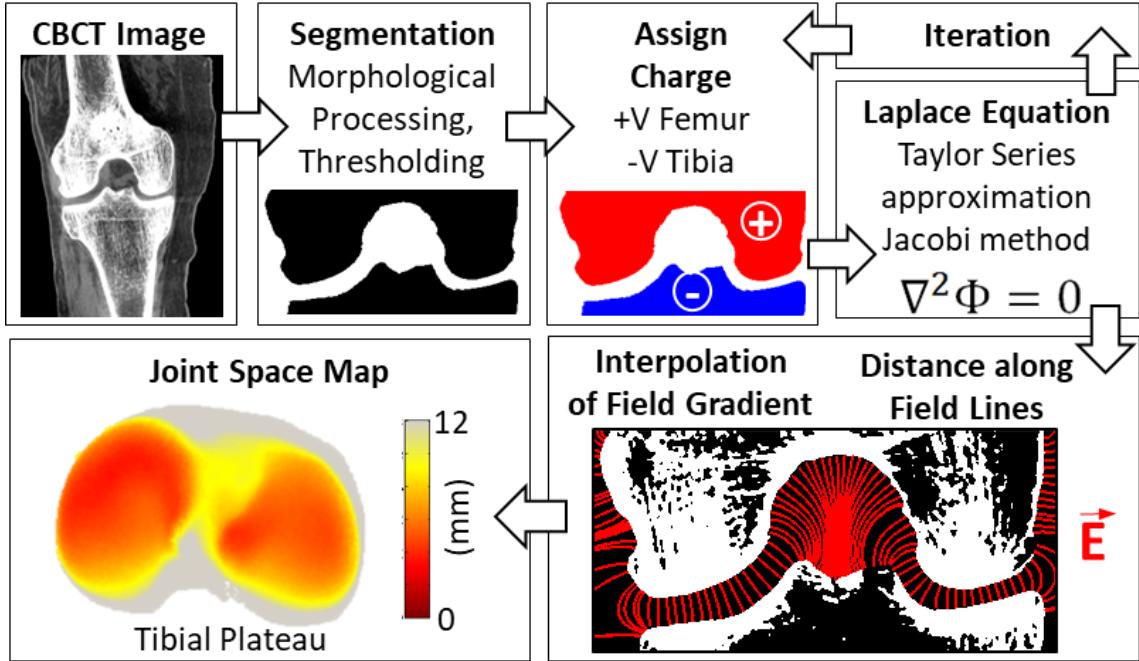


Figure 2.3 Flowchart illustration of the process for JSM calculation from CBCT scans.

The CBCT knee images were segmented in a semi-automatic pipeline to define the cortical bone surfaces. Each image was cropped to 256x256x128 voxels roughly centered about the intercondylar notch. Soft tissue was then subtracted from the image slice-by-slice by grayscale morphological opening and closing [81]. Bones were then segmented by local thresholding [82] based on the intensity range in a 3×3 neighborhood, and holes in the segmentation were filled by morphological closing. In instances of physical impingement of the femur and tibia (i.e., a “short circuit,” evident in 4/39 cases described below), the femur and tibia components were manually separated by one voxel. The femur and tibia were then identified as the two largest components in a connected component analysis.

Calculation of the electric potential and field lines was performed using the method in section II.B.1 carried out on a GPU (NVIDIA GTX470) using the *gpuArray* object in MATLAB (The Mathworks, Natick MA) for improved parallelization. The electric field at the $\partial_0\Omega$ boundary of the image was calculated with forward and backward differences, and

the field at interior voxels was computed using the central difference. Field line lengths were computed using MATLAB's *stream3c* function modified to Euler's method as described above with a step size of $d = 0.1$ voxels. Field lines terminating at the image border or with length greater than 3.6 cm were discarded, since the trajectories of these field lines were well outside the region of interest and no longer relevant to the JSW. Convergence of the solution was assessed in terms of equation (2.9). The JSM was represented as a colormap superimposed on the tibial plateau, which presented a simpler (flatter) surface than the femur and fairly intuitive visualization of the result.

II.B.3 Characterization of Knee Osteoarthritis

A study was conducted under an approved IRB protocol in which a total of N=39 (24 male, 15 female; 29-78 yo, mean 47 yo) subjects were recruited under informed consent, including 21 subjects presenting with OA and 18 subjects with no known symptoms or history of OA. All images were acquired using a scan protocol of 80 kV and 108 mA (~10 mGy dose to water at the center of a 14 cm water cylinderas described in [10]). A scan of the dominant knee was acquired for each subject in the standing position (Fig. 2.1) with their weight evenly distributed between both legs. Images were reconstructed by 3D filtered backprojection with a Hann kernel and a volume encompassing a $20 \times 20 \times 20 \text{ cm}^3$ FOV about the knee joint at isotropic voxel size of 0.52 mm. The OA population consisted of subjects diagnosed according to the current standard of clinical care based on symptoms, physical examination by an orthopedic specialist, and assessment by an MSK radiologist (independent of the JSM analysis herein) of joint space narrowing, osteophytes, effusion, and other established radiographic biomarkers of OA.

JSM features were analyzed to investigate whether the method could provide quantitative characterization of joint space morphology and possible discrimination of normal and OA knees. First, each point cloud was rigidly registered, and the coordinates of corresponding points identified using the coherent-point drift (CPD) algorithm [83] were averaged to yield a mean shape [84]. The registration yielded an $n \times p$ ($n = 39, p = 15934$) matrix $M = (\overrightarrow{m_1} \dots \overrightarrow{m_n})^T$ where p -dimensional column vectors $\overrightarrow{m_1}, \dots, \overrightarrow{m_n}$ denote the JSW at p corresponding points in a collection of n registered JSMs.

The centered data was $M_C = (\overrightarrow{m_1} - \overrightarrow{m_{avg}} \dots \overrightarrow{m_n} - \overrightarrow{m_{avg}})$, where $\overrightarrow{m_{avg}} = \frac{1}{n}(\overrightarrow{m_1} + \dots + \overrightarrow{m_n})$. Principal component analysis (PCA) was performed on the centered data to more fully analyze the spatial information conveyed by the JSM. The maps were decomposed into components characterizing the key features of variance between the registered maps using the *pca* function in MATLAB, returning a matrix $U_{p \times (n-1)}$ with columns corresponding to eigenvectors and a matrix $S_{n \times (n-1)}$ with rows corresponding to the weights of the principal components in descending order of variance, where $M_C = SU^T$.

To assess how well these principal component weights could be used to correctly classify normal and OA knees, truncated vectors $\overrightarrow{s'_1} \dots \overrightarrow{s'_n}$ from matrix $S = (\overrightarrow{s_1} \dots \overrightarrow{s_n})^T$ were chosen as input to a support vector machine (SVM) [85] following a leave-one-out scheme — treating each vector as testing data while using the others for training. A varying number of principal components [1-20, shown in Fig. 2.8, corresponding to different degrees of truncation] are considered. The SVM is a binary classifier that finds the hyperplane $\vec{w}^T \vec{z} + \vec{b} = 0$ that maximizes the margin $\frac{1}{\|\vec{w}\|}$ separating

the training data (\vec{z}_k, y_k) , where $y_k \in \{+1, -1\}$ is a binary label, and \vec{z}_k is the input vector under transformation of the radial basis function (RBF) kernel $K(\vec{s}'_k, \vec{s}'_l) = \vec{z}_k^T \vec{z}_l = \exp(-\frac{1}{2\sigma^2} \|\vec{s}'_k - \vec{s}'_l\|)$. The MATLAB functions *svmtrain* and *svmclassify* were used to train input vectors and classify test data, respectively. For this study, the L1-soft margin classifier was used:

$$\operatorname{argmin}_{w,b,\xi} (\vec{w}^T \vec{w} + C \sum_{k=1}^N \xi_k)$$

subject to the constraints $y_k(\vec{w}^T \vec{z}_k + b) > 1 - \xi_k$ and $\xi_k > 0$. The soft margin parameter, C , reflects the degree of error allowed in the training sample and reduces overfitting, empirically set to a value of $C = 100$. For each leave-one-out experiment conducted using a different number of input principal components, the RBF kernel σ was perturbed from 0 to 30, yielding sensitivity and specificity analogous to a receiver operating characteristic (ROC) curve. The JSMs from the electrostatic model were also compared with those derived from the LA and CP methods and analyzed with the same PCA, SVM, and leave-one-out training paradigm. For the LA method, the longitudinal axis was defined as the z-axis of the reconstruction, coinciding approximately with the long axis of the leg. For the CP method, closest points between the tibial and femoral surfaces were computed (down-sampled by a factor of 2 due to the large number of nearest-neighbor searches involved).

The classification accuracy resulting from these JSM features was compared to that associated with the JSW measured by three expert MSK radiologists using methods common in current diagnostic practice. Specifically, the closest distance between the femur and tibia in the medial compartment was assessed visually in triplanar views and measured with a digital ruler on a VuePACS 3D workstation (Carestream Health, Rochester NY). To

test the statistical significance in the difference between area-under-curve (AUC) values measured using JSMS and the radiologist JSW measurements, an unpaired heteroscedastic t-test was performed. The AUC for each radiologist was considered separately. For the various JSM methods, an ensemble of AUC measurements was formed from the 11 measurements about the peak performing number of principal components (shown in Fig. 2.8), better allowing for noise in the JSM classification results and providing a conservative estimate with respect to the optimal number of principal components. For the electrostatic JSM method, the optimal number of principal components was 10, giving an ensemble of AUC measurements resulting from 5-15 components; similarly for the LA-JSM and CP-JSM methods, the ensemble about the optimum included AUC measurements resulting from the first 10-20 components and 5-15 components, respectively.

II.C Results

II.C.1 Algorithm Performance

The JSM calculation was fully converged for all cases by 25,000 iterations, with no change in JSM values for further iterations. Calculation of the electric potential was the most computationally intensive aspect of the method (approximately 20 iterations per second in MATLAB code on the mid-range GPU used in this work, requiring ~20 minutes per case); however, a reasonably stable solution could be obtained with fewer iterations, as shown in Fig. 2.4. For example: RMSE less than the voxel size (0.52 mm) could be obtained in the first ~1000 iterations (1 min) and RMSE less than the step size ($d = 0.052$ mm) at 10,000 iterations (~8 min). Accelerating the computation time was not a focus of the current work, and there are numerous hardware (faster GPU) and software (compiled

C++ / CUDA code) solutions to be investigated in future work. The time required for field line length analysis was small by comparison (~45 s).

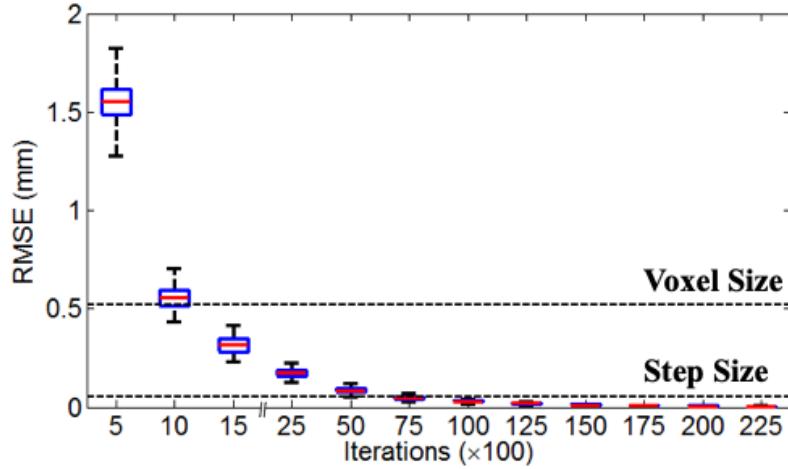


Figure 2.4 Convergence of the JSM calculation as a function of iterations. The box-and-whisker plots show the median, quartiles, and range in RMSE for all JSMs computed over all 39 cases in the clinical study.

II.C.2 Joint Space Maps

Osteoarthritis of the knee tends to present with narrowing of the joint space, with varying degrees in the medial and lateral compartment [86]—a pattern consistent overall with the JSM results presented in Fig. 2.5 for the 39 subjects in the clinical study. Some notable exceptions are apparent, such as subjects P23 and P38, which qualitatively demonstrate JSMs similar to the normal group; however, close inspection (and feature analysis below) reveals characteristics of the joint space morphology that are distinct. For example, P23 exhibits narrowing in the anterior medial aspect, and P38 demonstrates reduction of the intercondylar notch due to osteophyte growth. Such qualitative differences are investigated quantitatively below.

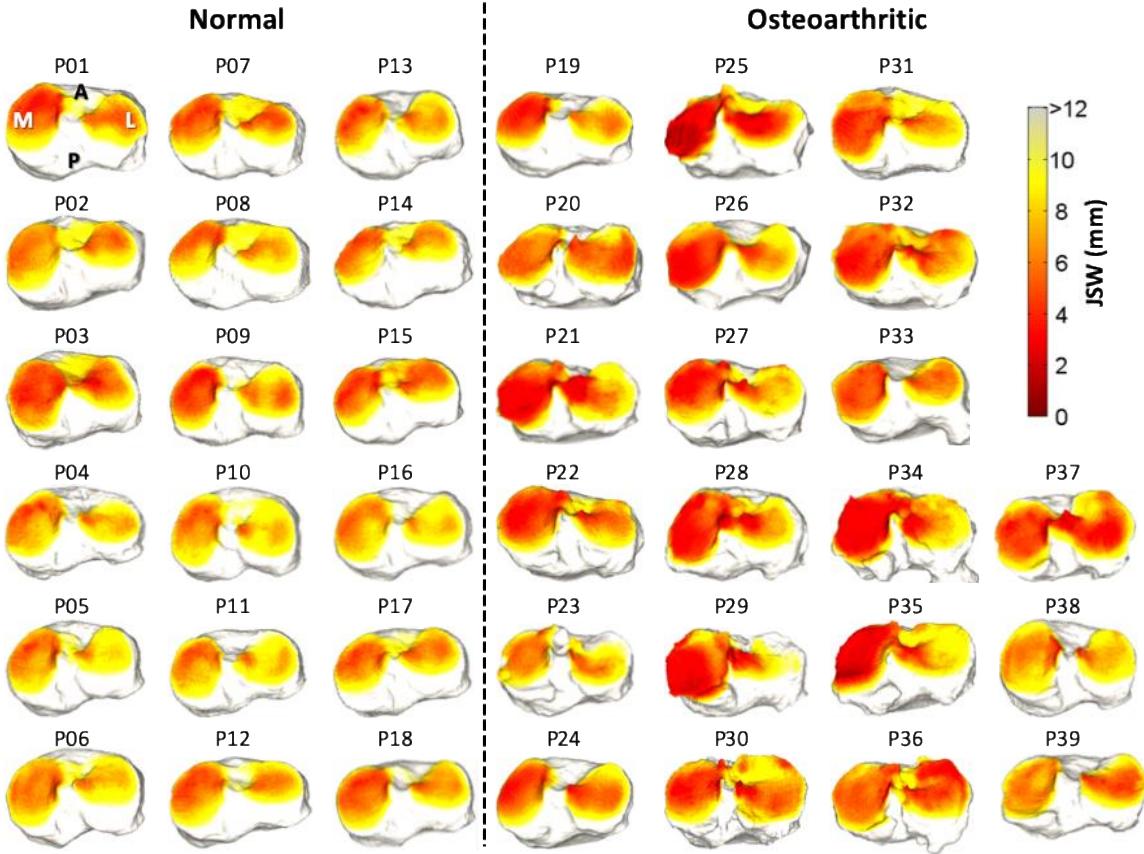


Figure 2.5 Joint space maps computed using the electrostatic model for 39 subjects in the clinical study, grouped as normal and OA. The JSM is displayed as a heat map projected onto surface of the tibial plateau.

II.C.3 Quantitative Analysis

Fig. 2.6 shows a simple analysis of the JSM in terms of the histogram of JSW values in each of the 39 subjects. Also shown are the radiologists' measurement of minimum distance in the medial compartment as in clinical diagnosis. Overall, the radiologist readings are similar to the minimum distance in the electrostatic JSM [denoted $\min(JSM)$], but there are notable exceptions in which $\min(JSM)$ is smaller than the radiologist reading (attributable to a minimum distance detected elsewhere than the medial compartment – e.g., in the intercondylar notch) and several in which $\min(JSM)$ is greater than the radiologist reading (attributable either to reader error or segmentation error). Of course,

equivalence between the two measures is not expected, since the field lines curve according to the electric potential, typically resulting in a longer distance within the curved space of the electrostatic model. Overall, the radiologist reading in the medial compartment was a good discriminator of normal and OA knees, yielding an overall area under the ROC curve (AUC) of 0.87. The $\min(JSM)$ value exhibited slightly better performance (AUC = 0.92), but the difference was not statistically significant ($p = 0.649$ via paired t-test on measurement data).

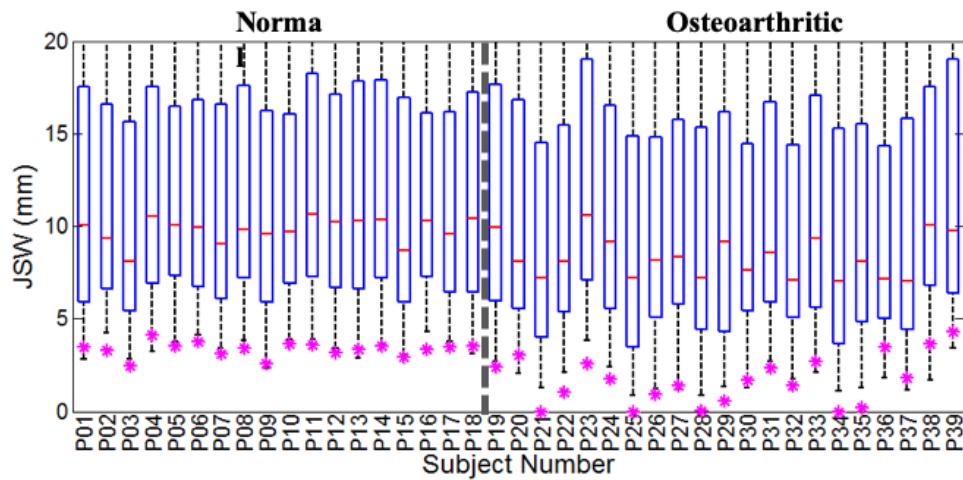


Figure 2.6 Distribution of JSW as measured by the electrostatic JSM method, with box-and-whisker plots showing the median, first and third quartiles, and range in JSM for each subject. Also shown (asterisk) is the minimum distance in the medial compartment as measured by 3 MSK radiologists.

Input for Classification	Mean AUC	p-value
Electrostatic JSM	0.98	0.008
CP-JSM	0.90	0.071
LA-JSM	0.76	0.007
Radiologist Reading	0.87	~

Table 2.1 Comparison of diagnostic accuracy (AUC) in discriminating normal and OA subjects. Radiologist reading corresponds to classification based on their manual measurement of the minimum distance in the medial compartment. The second column shows the mean AUC in the region (± 5 principal components) about the optimal number of principal components for the three JSM methods. The p-value is that resulting from an unpaired heteroscedastic t-test comparing a given JSM method to the Radiologist Reading. For example, the improved mean AUC measured for the electrostatic JSM method was statistically significant ($p = 0.008$), but that for the CP-JSM method was not ($p = 0.071$).

Analysis of the principal components of the spatial distribution evident in the electrostatic JSMs [illustrated in Fig. 2.7] yielded improved classification performance. Overall, the electrostatic JSMs for the OA group exhibited stronger variability in principal component weights than the normal group. Fig. 2.8 plots the AUC resulting from PCA of the JSMs as a function of the number of principal components included, yielding an AUC > 0.98 over a range of the first 10 to 15 principal components. The number of support vectors was in the range 10-15, which was below 50% of the sample size and is consistent with a modest fitting (i.e., not overfitting) in the SVM. Increasing the number of principal components reduced the classification accuracy in a manner attributed to noise in the JSM. Fig. 2.9 shows the sensitivity and specificity associated with the four tests considered: Radiologist Reading of the medial compartment; electrostatic JSM (using the first 10 principal components); CP-JSM (using the first 10 principal components); and LA-JSM (using the first 15 principal components). The LA-JSM method performed least favorably, consistent with the hypothesis described above regarding the arbitrariness in definition of the longitudinal axis. The CP-JSM method performed comparably to the Radiologist Reading method, which is sensible in that each method involves assessment of the closest distance between surfaces. The electrostatic JSM method performed best (mean AUC = 0.98) and the improvement with respect to the Radiologist Reading was statistically significant ($p = 0.008$).

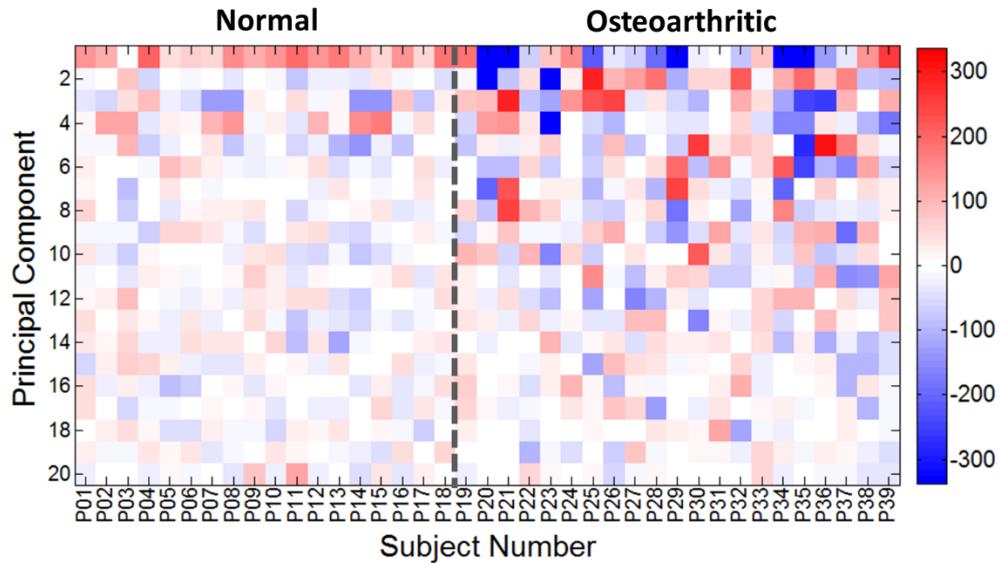


Figure 2.7 Principal component weight for the first 24 components analyzed using the electrostatic JSM method.

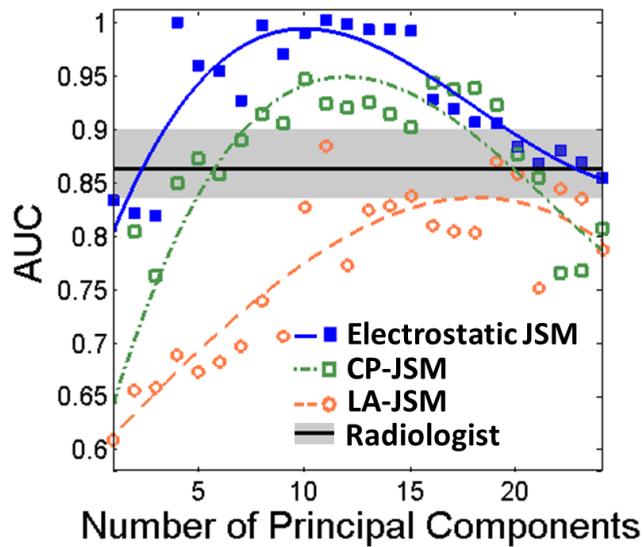


Figure 2.8 AUC in discriminating normal and OA subjects based on PCA of the JSM, analyzed as a function of the number of principal components. Curves are a simple polynomial fit to guide the eye. The Radiologist reading performance is marked by the mean and range (in gray).

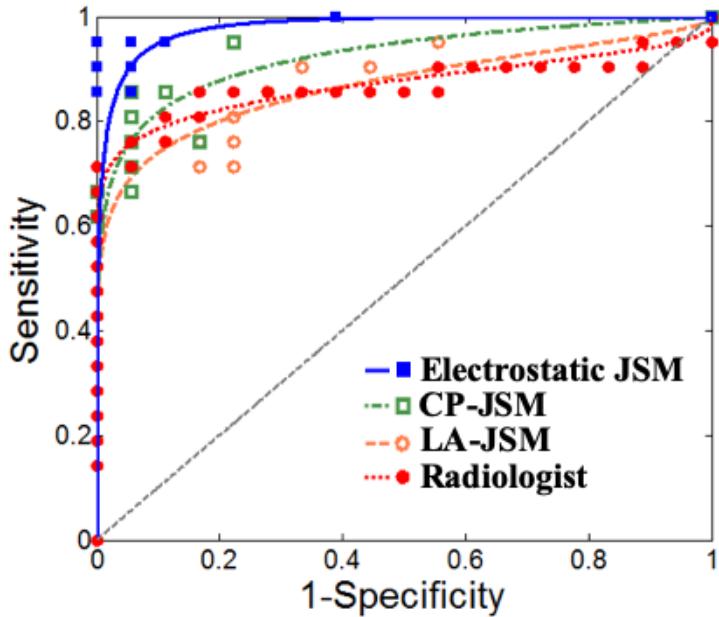


Figure 2.9 Sensitivity and specificity curves comparing classification accuracy for the mean radiologist reading of the medial compartment and PCA of the electrostatic JSM, CP-JSM, and LA-JSM using the peak number of principal components from Figure 2.7. For ease of visualization, the points were fitted with ROC curves based on the assumption of binormal distribution [87].

II.D Discussion and Conclusion

In summary, we have developed a method for morphological analysis of joints based on an abstraction of elementary electrostatics, envisioning the articular surfaces as conductors in a capacitor and characterizing the intra-articular space according to the distance along electric field lines. This method takes advantage of the isotropic 3D spatial resolution of CBCT extremity imaging and yields a unique, non-degenerate, symmetric, and dense characterization of the joint space in a manner that overcomes some of the limitations of other simple methods (e.g., distance along a longitudinal axis or distance between closest points). The method is also applicable to complex, multi-component joints and yields a rich “map” of the joint space. Simple analysis of the JSM (e.g., $\text{Min}(JSM)$) was comparable in knee OA classification performance to conventional clinical reading of

the medial compartment, and more sophisticated PCA of the map spatial distribution provided improved performance (nearly perfect in the limited sample of 39 subjects).

The current bottlenecks in applying such analysis of the joint space are the segmentation process and the iterative algorithm for solving the electrical potential. The first relies on a semiautomatic process for delineating the bones and required manual intervention to ensure smooth, well connected segmentations and to resolve impingements. We did not investigate more sophisticated segmentation methods such as active contour [88] or atlas-based [89] approaches that could alleviate the bottleneck and could perform well in segmenting high-contrast bone. The latter bottleneck is a product of the numerical method employed in the current work (i.e., the Jacobi approximation to the electric potential) and unoptimized implementation in MATLAB. Methods with faster convergence properties are the subject of future work as is implementation of the algorithm in C++ / CUDA better optimized for calculation on GPU.

The current study involves a variety of limitations. First, with respect to the task of discriminating normal and OA subjects, the method examined only the spatial relationship between the femur and the tibia and did not include other possible morphological characteristics (e.g., association with the patella or soft tissues, such as cartilage and meniscus) or material characteristics (e.g., bone mineral density). It may therefore be beneficial to complement the JSM with other measures that can also be derived from the CBCT image (e.g., soft tissue visibility [10] and bone mineral density [90]). Secondly, it is difficult to determine whether the SVM was overfitted for the current dataset and requires cross-validation in an expanded study beyond the fairly limited sample of 39 subjects before generalizing the results presented here to a larger population. Finally, to

compare the performance of CP-JSM and LA-JSM, although more sophisticated variations of these methods could be envisioned (e.g., for LA, computing the best fit to the tibial and/or femoral axis as the longitudinal axis), such is beyond the scope of this paper, which focuses on the concept of electrostatics as a model for morphological characterization. Moreover, the LA and CP methods exhibit some basic limitations even in more sophisticated forms: the former is limited to simple two-component joints (e.g., the knee), and the latter is subject to undersampling and yields a different answer depending on which surface is selected as the reference. The electrostatic method is free of these limitations and could be applied to complex joints (e.g., the wrist or ankle) and yields a unique solution.

Analysis of joint space morphology in terms of PCA of the JSMs computed from the electrostatic model appeared to yield a high level of classification accuracy in normal and OA knees. More sophisticated decomposition and classification methods can be envisioned, which might further improve the accuracy and sensitivity of morphological analysis. Future work also includes extension of the model to more complex, multi-component joints and application to other pathologies that could benefit from improved quantitative image analysis of joint morphology associated with disease progression or therapy response, including OA of the hand and wrist, rheumatoid arthritis, and the post-traumatic load-bearing ankle.

Chapter III

Polyenergetic MBIR for Improved Quantification of Bone Composition

This chapter describes the integration of polyenergetic MBIR, MC scatter simulation and glare correction for estimation of BMD in CBCT. The polyenergetic MBIR algorithm is parametrized with a mixture model for bone and water, achieving accurate and reproducible estimates of BMD.

III.A Introduction

Reconstructed attenuation values (Hounsfield Units, HU) suffer a variety of sources of inaccuracy in CBCT, primarily due to high SPR. Additionally, system blur resulting from finite focal spot, off-focal radiation [91], detector blur and veiling glare [92] may also affect HU accuracy. Though some of these factors have been addressed in the recent literature [64], [93], these corrections need to be combined with an efficient beam hardening artifact reduction scheme to achieve quantitative BMD characterization from polyenergetic data acquired with a single x-ray tube potential. (This work does not consider dual-energy imaging, which requires either specialized hardware - e.g. fast kV switching [94], photon counting detectors [95], dual layer detectors [96], multi-source system configurations [97] - or double scanning [98].)

Conventionally, beam hardening is addressed through linearization of the measured line integrals, followed by bone correction, typically involving an approach similar to that proposed by Joseph and Spital [99]. In Joseph-Spital (JS) correction, an initial

reconstruction is segmented into water and bone voxels. A reprojection is then performed to obtain the total amount of bone and water along each ray. A precomputed table is then used to correct the projection data given the amount of bone and water, replacing contribution of bone attenuation with equivalent attenuation by water. The corrected projections are backprojected to obtain the final reconstruction. However, for extremities imaging, the initial segmentation is often difficult in areas of trabecular bone, where voxels typically contain a mixture of marrow and bone. Secondly, material substitution (of bone by water, for example) does not fully model polyenergetic attenuation for bone. Thus, even with JS correction, interpolated BMD values obtained from an in-scan calibration phantom may still be biased by residual beam hardening artifacts.

Several MBIR algorithms have been developed to address beam hardening. These algorithms are typically formulated as optimization of the PL objective with respect to the image [100]. The primary challenge in adapting the PL framework to polyenergetic object models is that the number of unknowns (voxels times number of basis materials) exceeds the measured projection data. A common approach to address this difficulty is to parametrize each voxel with a single unknown and pose a physical model relating material composition of a voxel to its total attenuation. A notable example is IMPACT [101], in which the likelihood objective is expressed as a function of monochromatic linear attenuation for each voxel. In the forward model, a table look-up is performed to convert monochromatic attenuation into photoelectric and Compton attenuation components, which together constitute total attenuation along a line integral. Another example is polyquant [102], [103], which takes a similar approach to IMPACT but crafts the likelihood as function of electron density instead of monochromatic attenuation and incorporates a

kernel-based Compton scatter estimate. A third example is a polyenergetic penalized likelihood (polyPL) reconstruction framework introduced by Elbakri *et al* [61], [62], in which the likelihood is a function of total density. A fundamental assumption in this method is that material composition in biological tissue correlates with total density. A high-density voxel may consist of bone and water while a low-density voxel may be considered water-like. This is described by an object model, a function of material fractions of the basis materials with respect to total density.

We investigate whether a polyenergetic MBIR algorithm combined with an advanced, CBCT-specific artifact correction pipeline can enable accurate and reproducible measurements of BMD on an extremity CBCT system. Reproducibility is an essential aspect of any quantitative imaging biomarker [104]. Similar to prior studies of other densitometric biomarkers [105], [106], we have therefore evaluated the variability of BMD measurements of standardized test inserts obtained from CBCT images acquired under a variety of conditions, such as perturbations in object position, orientation, size and composition. To the best of our knowledge, this work is the first to provide such analysis for CBCT implementing polyenergetic MBIR and advanced artifact corrections.

We used the polyPL algorithm as the basis for the development of the BMD measurement capability. Compared to other similar iterative polyenergetic algorithms mentioned above, the object model in polyPL is formulated in terms of base materials instead of interaction cross-sections or electron densities as in other polyenergetic MBIR methods, and therefore can be conveniently adapted to yield direct estimation of bone mineral concentration. To this end, we propose to augment polyPL with a new ideal mixture object model. This modified polyPL is deployed in concert with accelerated MC-

based scatter correction [93] and deconvolution-based glare correction [92]. We developed a procedure to calibrate the key parameters of the resulting quantitative imaging pipeline to a specific CBCT system. Accuracy and reproducibility are validated through experimental studies of quantitative phantoms imaged in multiple configurations on a benchtop system emulating extremity CBCT.

We note that the goal of this work is not to compare the quality of reconstruction across different algorithms but rather on reproducibility of BMD values across the all phantoms imaged in all configurations with polyPL and the added correction pipeline.

III.B. Methods

III.B.1 Experimental Setup

All projection data were acquired on a benchtop system consisting an FPD (4030CB, Varex, USA) and a rotating anode x-ray source (DU694, Philips, Netherlands), shown in Fig. 3.1. The test bench was set up to reflect geometry of extremity CBCT with source-axis distance (SAD) of 43 cm and source-detector distance (SDD) of 56 cm). Scans of 420 frames were acquired over 210° at 90 kV and 0.2 mAs/frame, resulting in total scan dose comparable to a typical extremity CBCT scan at 15 mGy. Image acquisition was performed in 2x2 detector binning mode, with effective pixel size of 0.388 mm. Fig. 3.1 also shows an example phantom used in the validation studies. It consists of Ca inserts in a water cylinder. The inserts range from 50-500 mg/cm³ Ca (Gammex 472 DE TMM, Sun Nuclear, Melbourne, FL) and are designed to mimic a mixture of CaCO₃ in solid water. Further detailed description of the phantoms and validation experiments is given in section II.C.

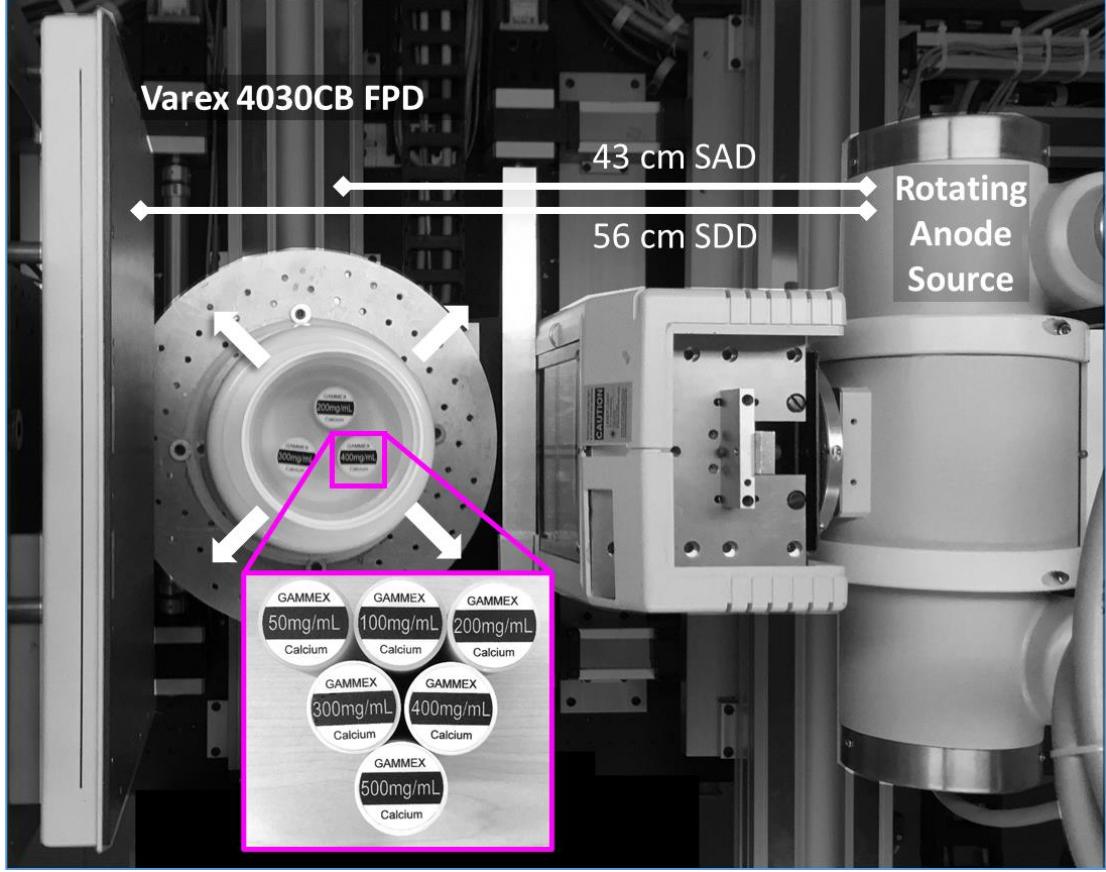


Figure 3.1 Top-down view of a benchtop CBCT system emulating extremity CBCT. A water phantom (W5) is shown on the motion stage. The magenta box highlights the collection of Ca inserts used in this reproducibility study.

III.B.2 Polyenergetic PL

III.B.2.1 Object Model

In polyPL [61], [62], the objective function with respect to total density $\rho = [\rho_1 \dots \rho_P]^T$ is defined by the likelihood term $L(\rho)$ and smoothness penalty $R(\rho)$:

$$\phi(\rho) = -L(\rho) + R(\rho) \quad (3.1)$$

The polyenergetic likelihood objective $L(\rho)$ is given by:

$$-L(\rho) = \sum_{i=1}^N h_i (\bar{Y}_i(\rho)) + r(\rho) \quad (3.2.1)$$

$$\bar{Y}_i = \int I(E) \exp\left(\sum_k^K - \left(\frac{\mu}{\rho}\right)_{k,E} s_{ik}\right) dE \quad (3.2.2)$$

$$s_{ik} = \sum_j^P a_{ij} \rho_j f_k(\rho_j) \quad (3.2.3)$$

where $h_i(t) = -Y_i \log(t) + t$ is the Poisson negative-log-likelihood function, $I(E)$ the source spectrum weighted by detector response, $\left(\frac{\mu}{\rho}\right)_{k,E}$ is the energy-dependent mass attenuation of the k th base material, and s_{ik} is the density integral for material k .

In order to compute s_{ik} we use an object model f_k that gives the fractional contribution of the k th base material to the attenuation of the j th voxel as a function of the total density of the voxel ρ_j . We assume $\sum_{k=1}^K f_k(\rho) = 1$; a_{ij} is the system matrix term corresponding to j th voxel and i th measurement. In this objective, a given total density value uniquely identifies a mixture of basis materials to compute the polyenergetic forward projection.

We adopt the ideal mixture as the model for f_k :

$$\frac{1}{\rho} = \sum_K \frac{m_k}{m} \cdot \frac{V_k}{m_k} = \sum_K \frac{f_k}{\rho_k} \quad (3.3)$$

where ρ_k denotes the pure densities of each base material, m the total mass, and m_k , V_k are the partial mass and volume of the k th base material, respectively. Specifically, we assume a binary mixture of water (denoted with subscript $k=1$) and Calcium Carbonate (CaCO_3 , $k=2$). In this formulation, BMD is computed from the CaCO_3 fraction as $\rho f_2(\rho) \frac{w_{Ca}}{w_{CaCO_3}}$, where w_{Ca} and w_{CaCO_3} denote molecular masses. As will be shown below, the pure densities of the base materials in Eq. 3.3 can be adjusted to match the spectral characteristics of the model to those of realistic human tissues, even though those tissues might contain small amounts of materials other than water and CaCO_3 .

Fig. 3.2 illustrates the proposed object model in application to the Ca inserts used in the experimental validation studies (Sec. III.B). The density and mass attenuation of each insert (denoted $\left(\frac{\mu}{\rho}\right)_{insert}$) were computed from its known elemental composition [107]. We then found the base material fractions f_1 and f_2 that yielded a binary water-CaCO₃ mixture with mass attenuation that best matched that of the inserts:

$$f_2 = \operatorname{argmin}_{\tilde{f}_2} \int \left((1 - \tilde{f}_2) \left(\frac{\mu}{\rho}\right)_1 + \tilde{f}_2 \left(\frac{\mu}{\rho}\right)_2 - \left(\frac{\mu}{\rho}\right)_k \right) I(E) dE \quad (3.4)$$

where the matching term was weighted by the spectral response of the system $I(E)$ (defined as the product of the x-ray spectrum and detector absorption efficiency).

We note that the BMD values estimated by this approach differ from the nominal Ca concentrations provided for the calibration inserts. We denote the BMDs obtained from the binary model as *apparent densities*. Apparent densities for the 50, 100, 200, 300, 400 and 500 mg/cm³ Ca inserts were 23, 74, 178, 280, 383 and 485 mg/cm³, respectively. Those values are consistent across energy levels (90-120 kV) in the $I(E)$ model.

The discrepancy between the apparent and nominal BMD values is likely due to a mismatch between the binary mixture of basis materials and the true composition of the inserts, which are formulated with solid water. Solid water contains significant amount of carbon (65-70%) and a small fraction of N, Mg and Ca. Despite this discrepancy, the binary mixture captures the attenuation properties of the inserts very well. At the effective energy of 52 keV, the mismatch between monochromatic mass attenuation of the insert $\left(\frac{\mu}{\rho}\right)_{k,52\text{ keV}}$ and that of their bases approximations $\left(f_1 \left(\frac{\mu}{\rho}\right)_{1,52\text{ keV}} + f_2 \left(\frac{\mu}{\rho}\right)_{2,52\text{ keV}}\right)$ was <0.41% with mean error of 0.18%. Furthermore, the CaCO₃ fractions obtained with the water-CaCO₃

model for the Ca inserts and for common tissue types encountered in orthopedic imaging [108] follow the $f_k(\rho)$ curve of an ideal mixture under reasonable assumptions for the pure base material densities ρ_1 and ρ_2 . This is demonstrated below.

Because polyPL only considers the apparent value in its forward model, evaluation of BMD accuracy in reconstructed images was based on apparent densities instead of nominal values (e.g. 50, 100, and 200 mg/cm³ Ca).

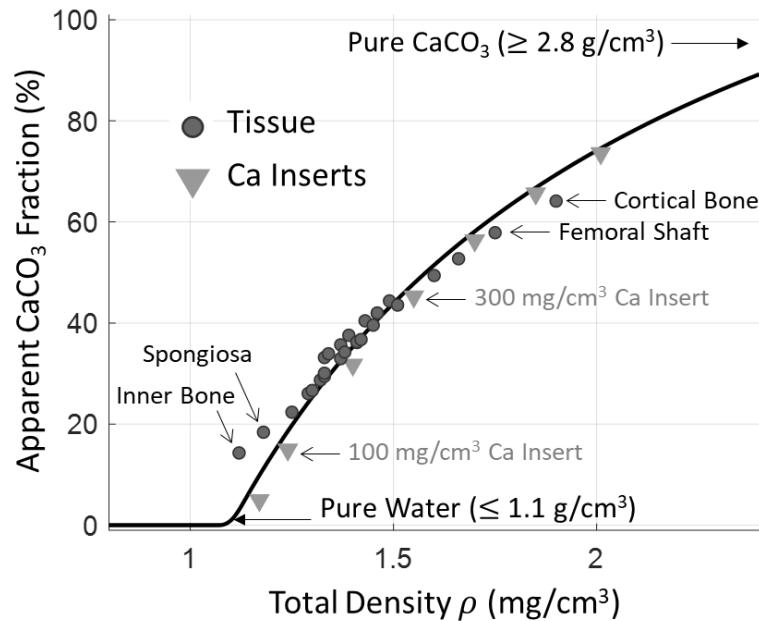


Figure 3.2 Object model for the water-CaCO₃ binary mixture. Data points represent apparent CaCO₃ fraction for the phantom inserts and common tissue types in extremity CBCT.

III.B.2.2 Model Calibration

Detector response, and attenuation spectra of basis materials were calculated from Spektr 3.0 [24]. The source spectrum was generated from Spektr and calibrated with HVL measurements at 90 kV by varying source filtration in the form of Cu, Al and W

thicknesses. Total filtration was estimated to be 0.25 mm Cu, 2 mm Al. The spectral response $I(E)$ was computed as:

$$\tilde{I}(E) = q(E) \left[1 - \exp \left(t_{CSI} \rho_{CSI} \left(\frac{\mu}{\rho} \right)_{CSI} \right) \right] \quad (3.5.1)$$

$$I(E) = \frac{\tilde{I}(E)}{\int \tilde{I}(E) dE} I_0 \quad (3.5.2)$$

where $q(E)$ is the calibrated source spectrum, ρ_{CSI} the density of CsI:Tl scintillator (4.51 g/cm³) and $\left(\frac{\mu}{\rho} \right)_{CSI}$ the mass attenuation spectra.

Prior to reconstruction, the projections were processed using a conventional deconvolution-based glare correction. We adopted an analytical form of the glare correction kernel [92], [93]:

$$PSF(r, a, b) = \frac{a}{2\pi b^2} \frac{1}{\left(1 + \frac{r^2}{b^2} \right)^{3/2}} \quad (3.6)$$

where r denotes the radius from the center. One possible way to derive this kernel is to apply inverse Fourier transform on a measured MTF. The MTF can be obtained from the LSF measured from an oversampled W edge positioned on the detector [109]. We opted to calibrate in the reconstruction domain since projection consistency is also impacted by long-range glare (off-focal radiation, etc) in addition to high frequency detector blur (>0.1 lp/mm).

We used one scan of an 11.7 cm water phantom with the 200, 300 and 400 mg/cm³ Ca inserts (phantom W5 in Sec. II.C) to calibrate the parameters of the ideal mixture model (the pure material densities ρ_1 and ρ_2), and of the glare response function used in the deconvolution-based correction.

The calibration of the ideal mixture model involved varying parameters ρ_1 from 0.95 to 1.2 g/cm³ and ρ_2 from 1.8 to 3.2 g/cm³ in polyPL reconstructions generated with scatter correction (Sec. III.B.2.3) using glare kernels from prior work performed on the same detector [93]. We then found the RMSE of the reconstructed insert densities compared to the apparent densities obtained from Eq. (3.4). The solid line in Fig. 3.2 represents the ideal mixture model $f_2(\rho)$ obtained from this calibration and used throughout this work. Note the smooth transition from water to the water-CaCO₃ mixture at ρ_1 (1.1 g/cm³). A polynomial was introduced in $f_2(\rho)$ over a small interval surrounding ρ_1 to ensure first-order differentiability of the objective function of Eq. 3.2.1–3.2.3, as required by the polyPL optimization algorithm (ordered-subsets separable parabolic surrogates (OS-SPS) algorithm with De Pierro’s additive convexity trick for Hessian separability [110], [111]).

A range of glare parameters (a, b) were used to perform polyPL. BMD (reconstructed vs. apparent density) was computed in ROIs at the centers of the 200, 300 and 400 mg/cm³ inserts. For each insert, the error was interpolated over the range of a and b using a cubic kernel. Parameters a and b were then selected to minimize the RMSE of all inserts yielding 0.084 and 3.105 for a and b , respectively.

The settings of the ideal mixture model and the glare model identified through this calibration were used in the processing of all other phantoms and configurations.

III.B.2.3 Reconstruction Workflow

Throughout the entire reconstruction pipeline, the voxel size was 0.4x0.4x0.4 mm³. Each dataset was first reconstructed with FDK using Hann filter with 0.8x Nyquist cutoff.

Prior to FDK, the projections were linearized assuming that all materials are water-like (“water correction”) [99].

The attenuation volumes obtained from FDK needed to be converted to density to perform MC scatter correction, and, subsequently, to be used as initialization for polyPL. To achieve this conversion, we assumed that the output of FDK was consistent with the effective attenuation defined by [112] and applied the ideal mixture object model of Eq. 3.3:

$$\mu(\rho) = \frac{\int \left(\left(\frac{\mu}{\rho} \right)_1 \rho f_1(\rho) + \left(\frac{\mu}{\rho} \right)_2 \rho f_2(\rho) \right) I(E) dE}{\int I(E) dE} \quad (3.7)$$

The ρ for each FDK voxel was found by interpolation in a table of μ values obtained by computing Eq. 3.9 over a range of f_1 and f_2 values.

Scatter distributions were computed with MC for every 10 projections in the acquisition (5° angular increment), with 10^8 photons/projection. For simplicity, the MC material models did not account for mixtures. Instead, voxels containing the water-bone mixture ($\rho_1 < \rho$) were simulated using CaCO_3 interaction cross-sections, but assuming a reduced density that was chosen to yield effective attenuation that matched the attenuation of the voxels’ mixture:

$$\rho_{2,MC} = \underset{\rho_{2,MC}}{\operatorname{argmin}} \int \left(\rho f_1 \left(\frac{\mu}{\rho} \right)_1 + \rho f_2 \left(\frac{\mu}{\rho} \right)_2 - \rho_{2,MC} \left(\frac{\mu}{\rho} \right)_2 \right) I(E) dE \quad (3.8)$$

A Gaussian kernel was used to smooth and interpolate the resulting scatter projections and to generate noiseless scatter projections for all 420 views ($r(\rho)$ in Eq. 3.2.1). A volume downsampling of 2 (with voxel size of $0.8 \times 0.8 \times 0.8 \text{ mm}^3$) was used to speed up simulations. Fig. 3.3(A)(B) shows an example of the smoothed scatter distribution

along with the corresponding SPR for the same projection. The scatter distribution was subtracted from the measured projections to obtain a new FDK reconstruction.

After 3 iterations of FDK-MC, the resulting density image was used to initialize polyPL. For polyPL, the noise-equivalent photon fluence I_0 was set to 10^5 photons/pixel, assumed to correspond to 80% detector saturation for the detector. We also implemented Nesterov updates to speed up convergence [113], [114]. Regularization strength β was $10^{5.5}$ and Huber δ was set to 10^{-4} g/mm³. The Huber penalty was used for $R(\rho)$ throughout this work. Optimization was terminated at 200 iterations with 6 subsets per iteration, when updates between iterations were verified to become negligible. MC scatter estimate was updated every 25 iterations.

Fig. 3.3C-E compares polyPL reconstructions of the calibration phantom for three scenarios: (C) without scatter correction but with glare correction, (D) with scatter correction but without glare correction, (E) with full suite of corrections. In C-D, polyPL compensates for projection inconsistencies but with limited success, resulting in residual cupping and streaking. This suggests that even though quantitative accuracy is most-impacted by scatter, the effect of glare correction is still noticeable, especially for parts of the image with sharp gradients (such as insert-water interfaces).

In terms of implementation, most operations were accelerated using precomputed tables and fast table-lookups (linear interpolation). This includes integration over energy bins (Eq. 3.2.2), which were computed as linear interpolation over a table of log attenuation values with respect to material density integrals. Additionally, the object model functions $f_1(\rho)$ and $f_2(\rho)$ were evaluated in parallel on each voxel using CUDA kernels. This cuts down on intermediate volumes for masking and minimizes host-GPU data transfer.

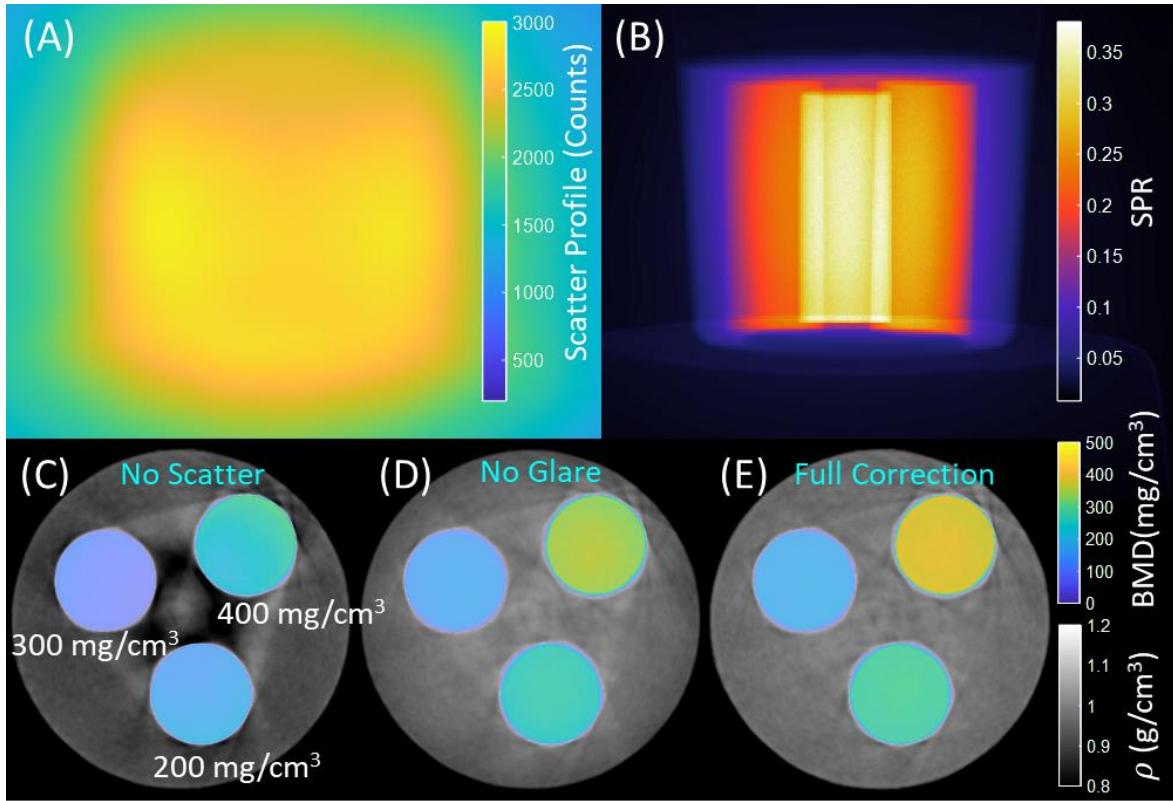


Figure 3.3 (A) MC scatter profile and (B) SPR for the same projection of the calibration phantom shown in (C-E). Bottom row, polyPL reconstructions with (C) no scatter correction and (D) no glare correction. (E) full correction.

III.B.3 Reproducibility Study

III.B.3.1 Water-Ca Phantom

To assess reproducibility and accuracy of polyPL at recovering BMD values, a family of phantoms consisting of Ca inserts in water cylinders were imaged on a test bench in extremity CBCT configuration (Fig. 3.2). Table 3.1 describes 5 water phantoms used in the study: Each phantom was scanned 4 times in different positions within the FOV, 2-4 cm apart, with rotation of 0-30 degrees. An array of 27 ROIs (3x3x3, each 6x6x6 mm³) were extracted from each insert image. This yields 20 scans, 60 insert realizations and 1620 ROIs. Mean BMD values of the ROIs were used for analysis. The ROIs were sorted by their nominal BMD values (50, 100, 200, etc mg/cm³ Ca). To assess BMD accuracy with

polyPL, mean BMD of each ROI is computed and compared with the apparent density of each insert. The coefficient of variation (CV) was computed as standard deviation normalized by mean.

Phantom	Inserts (mg/cm ³ Ca)	Water Diameter
W1	50, 100, 200, 300	15.6 cm
W2	200, 400, 500	-
W3	200, 500	11.7 cm
W4	50, 100, 200	-
W5	200, 300, 400 (calibration)	-

Table 3.1 Summary of water phantom configurations.

III.B.3.2 Cadaver Specimen

A fresh cadaver knee specimen (Maryland Anatomy Board) was scanned in a plastic container with added attenuators positioned near the knee joint, fixed in position with foam. The attenuators consisted of Gammex CaCO₃ inserts in the previous section as well as water inserts. After each scan, a new set of attenuators were put in. The position and flexion of the knee were also adjusted in between scans. A detailed listing of the configurations is shown in Table 3.2.

Phantom	Attenuators (mg/cm ³ Ca & water)
C1	100, 200, water
C2	50, 300, water
C3	200 x2, water
C4	100, 300
C5	No additional attenuators

Table 3.2 Summary of cadaver phantom configurations.

The cadaver datasets were also reconstructed using the pipeline of Sec. III.B. For analysis, we chose one of the datasets as the reference volume. The tibias of all other reconstructions were rigidly registered to the tibia of the reference volume[115]. A mask of the tibia was constructed using thresholding, morphological closing and hole-filling. An additional cortical ROI was derived via morphological erosion of the tibiae mask with radius 2.2 mm. The trabecular ROI was defined as the complement of the cortical ROI within the tibial mask. Within the trabecular bone ROI, some voxels contained negligible BMD, these voxels were removed from the analysis. To minimize the impact of reconstruction noise and possible registration errors on the analysis of reproducibility, we measured the average BMD values in 2.8 mm³ neighborhoods centered on each voxel of the trabecular and cortical ROI. For each voxel, the local reproducibility was then obtained as the standard deviation of the ROI-averaged BMD values, computed across all registered reconstructions.

III.C Results

III.C.1 Water-Ca Phantom

PolyPL reconstructions and FDK initialization for a configuration sampled from each water phantom are shown in Fig. 3.4A. The total density (in grayscale) is windowed to soft tissue ($0.8\text{-}1.2 \text{ g/cm}^3$) and overlaid with reconstructed BMD values from 0 to 500 mg/cm^3 . Visually, cupping effects from scatter and beam hardening typically seen in uncorrected CBCT images are absent. Additionally, the BMD insert profiles appear uniform with consistent values across all configurations. The mean BMD values for the 200 mg/cm^3 insert in all configurations are shown in Fig. 3.4B. In the boxplot, each bar represents a configuration (4 per phantom) and error bars denote distribution of mean BMD values of 27 ROIs extracted from the 200 mg/cm^3 insert. Though some phantom-dependent fluctuations in BMD values persist, the mean error throughout all ROIs was 4.4 mg/cm^3 relative to apparent BMD with standard deviation of all ROI means at 5.49 mg/cm^3 . The CV across the 200 mg/cm^3 inserts were 1.84%.

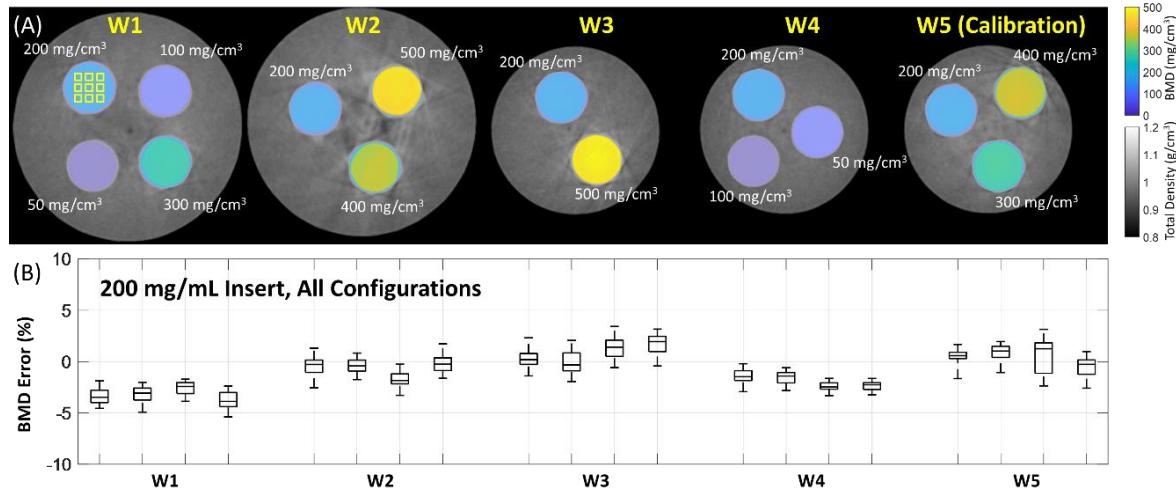


Figure 3.4 (A) Example reconstructions with FDK initialization (with scaled BMD values) and PolyPL. Yellow squares in W1 FDK image (top left) shows approximate locations of ROIs sampled for analysis. (B) Distribution of percent BMD error with apparent values in ROIs of the 200 mg/cm^3 insert across all configurations.

Figure 3.5 shows a similar analysis across all nominal insert densities. Each bar shows the mean BMD values from all configurations of the same insert. The reconstructed BMD is plotted against the apparent density, with dashed identity line indicating perfect reconstruction accuracy. Reconstruction error is highest for lower-density inserts, at ~6% for 50 mg/cm^3 , 4% for 100 mg/cm^3 inserts and <0.3% from 2 to 500 mg/cm^3 . In terms of reproducibility, from the 50 mg/cm^3 to the 500 mg/cm^3 insert, standard deviations of mean ROI values for each insert across all configurations increase from 1.66 mg/cm^3 to 11.19 mg/cm^3 , respectively. This results in CV from ~7.2% at 50 mg/cm^3 to ~2.3% at 500 mg/cm^3 .

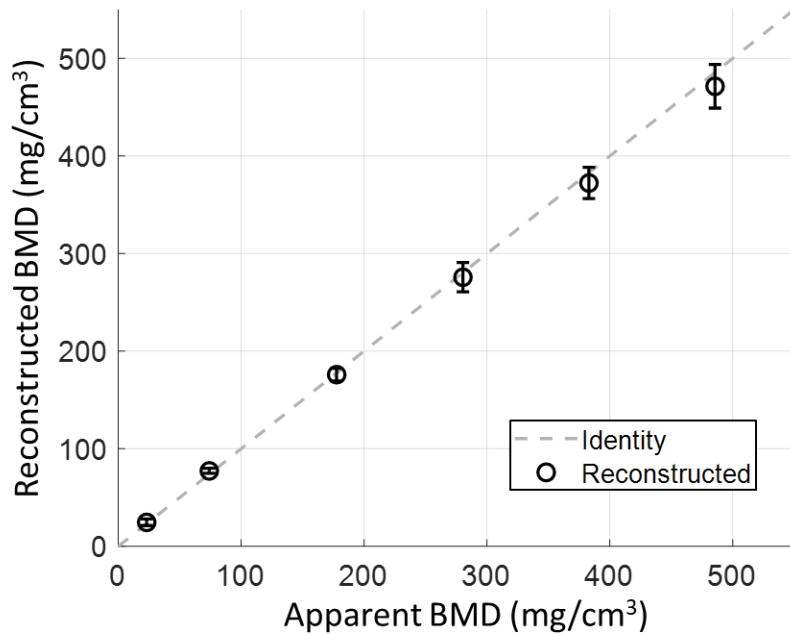


Figure 3.5 Distribution of BMD values of Ca insert ROIs (1620 total ROIs) vs apparent BMD, mean is marked by circles, error bars indicate 2x standard deviation from the mean. Dashed lines indicate apparent densities expected from spectral calibration and represents perfect accuracy.

III.C.2 Cadaver Specimen

Axial and coronal views of reconstructions of 2 cadaver phantom configurations overlaid with BMD are shown in Fig. 3.6. All reconstructions were registered to the tibia. Despite positioning of attenuators and deformations in surrounding soft tissue, reconstructed BMD appear similar across all imaging configurations. In the trabecular bone regions within both the tibia and femur, reconstructed BMD was $\sim 209 \text{ mg/cm}^3$, which is consistent with apparent densities derived from tissue composition in Fig. 3.2B, where inner bone and spongiosa amounted to 160 and 210 mg/cm^3 , respectively. None of the voxels in the reconstructed image approached the apparent density of cortical bone (1.9 g/cm^3). This is likely due to linear partial volume effects. Since this effect is modeled by the ideal mixture, no visible streaking is observed. On the contrary, the interface of Ca inserts and the water insert in C1 generated some streaking artifacts. This is expected since the interface contains a 3-material mixture of air, water and Ca, which is not modeled by the 2-material ideal mixture model. Additionally, the Huber penalty was not designed to handle the sharp gradients at the interface of the inserts in air. Moreover, streaking is localized to the water insert and the tibia ROI was unaffected.

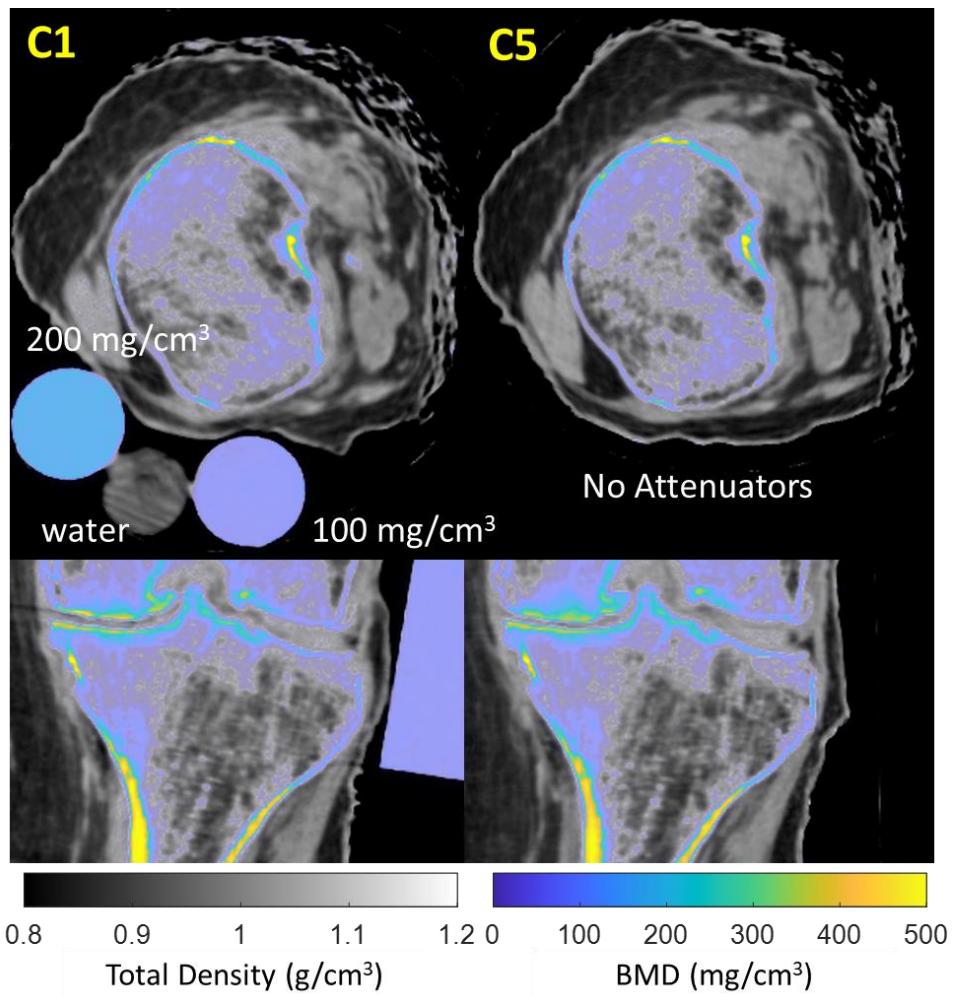


Figure 3.6 Sample of 2 configurations in the cadaveric knee phantom. C1 contains additional attenuators while C5 does not. The two BMD distributions are visibly similar.

Fig. 3.7 shows voxel-wise standard deviation between the configurations, along with a mask delineating tibial ROI (in dashed cyan). We restrict our analysis to the tibial ROI, which is partitioned into cortical and trabecular ROIs. Regions of high variation are visibly distributed around edges of cortical bone. This is consistent with the level of variability achieved in our water phantom study. Mean cortical bone BMD was 856 mg/cm^3 , with average standard deviation of 5.3 mg/cm^3 ; mean trabecular bone BMD was 209 mg/cm^3 , with average standard deviation of 2.5 mg/cm^3 . The mean standard deviation

for all tibial voxels were 4.0 mg/cc with mean BMD of 557 mg/cm³, suggesting high reproducibility of BMD.

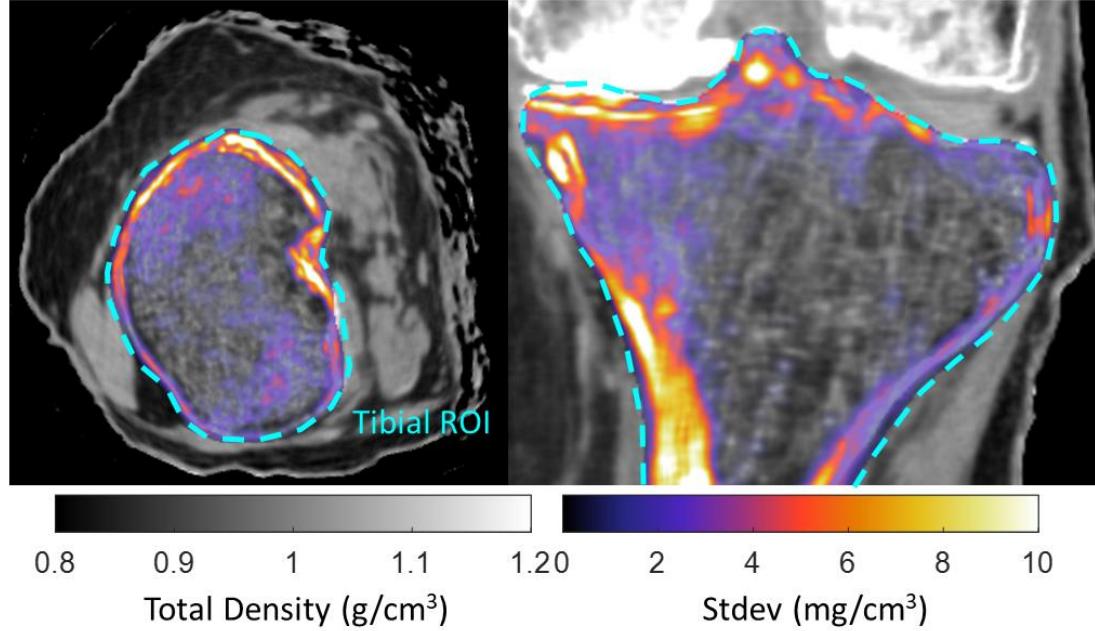


Figure 3.7 Distribution of voxel-wise standard deviation across 5 cadaver phantoms in a tibial ROI, with all phantoms were registered to the ROI.

III.D Discussion

We showed that apparent BMD can be reconstructed reproducibly with our polyPL-correction workflow across multiple families of phantoms. For all reconstructions, convergence was achieved at ~150 iterations, when image updates per iteration became trivially small and MC simulations returned stable, constant scatter profiles for every 25 iterations of polyPL.

Elbakri and Fessler proposed two object models with polyPL[62]. The first model is the displacement model, in which water/bone fractions for biological tissue were fit to their total densities with third-order polynomials. Since the model incorporates pure densities for bone, water and other biological tissues, the authors described it as suitable

for materials occupying distinct well-defined regions, such as cortical bone and soft tissue. Though a polynomial fit between the pure materials was used to provide convenient derivatives for optimization, the authors acknowledged that the model can be refined for quantitative accuracy. The second model the solution model, which describes solute in solvent (for example, KHPO₄ in water). The authors noted that this model is more useful when bone is not dense and exists in a water-bone mixture state. The interpretation of our ideal mixture model is similar to the displacement model, in which we assume anatomical structures are distinct and not physically mixed at a microscopic scale (e.g. lattice of trabecular bone in marrow). In this case, the volume of a voxel is the sum of the volumes of its components. By setting ρ_2 to infinity, our ideal mixture model also simplifies to Elbakri's solution model. This suggests that properties of the ideal mixture are applicable to both of Elbakri's mixture interpretations, given different parametrizations. In both our phantom and cadaver study, use of the solution model achieved artifact reduction and quantitative reproducibility.

In terms of performance, the bottleneck remains forward and backprojection operators. On an RTX 2080 (Nvidia, Santa Clara) graphics card, each iteration takes ~15s using our current optimizer and parameters. In practice this may be further optimized, either with more subsets or changing the weights in the Nesterov momentum term. One notable observation is that the two acceleration techniques are sometimes incompatible [114]. When the projection data is broken into 15 subsets or more, the momentum-driven image updates became unstable. The instability manifests in large updates to the density image, creating patches of hyperdense BMD as well as streaking surrounding these patches. Though some instable patches gradually converge with more iterations, others

persist even after reaching the high number of prescribed iterations. The specific number of ordered subsets to use as well as whether Nesterov acceleration is helpful for a given number of subsets warrants further investigation but is outside the scope of this proof-of-principle study. MC scatter simulation takes ~40s per volume.

We note several subtle aspects with our approach in its current form. First, our model allows for some ambiguity between dense soft tissue and sparse trabecular bone. For example, a voxel containing muscle may be as dense as a mixture of trabecular bone and marrow. This is difficult to distinguish without multi-energy scans or prior knowledge of gross anatomy. Fortunately, in practice, these tissues are spectrally similar (to water) and sufficiently low in density such that no visible streaks result from data consistency issues. However, this warrants careful interpretation of reconstructed density and making the distinction between apparent and absolute BMD. Another source of error results from the basis material approximation itself, described in section III.A. This is evident with our Ca inserts, where average error between actual mass attenuation and the best-fitting water- CaCO_3 approximation is 0.34% (Eq. 3.4). In biological tissue, the error may be higher, as composition is expected to be more varied, with additional elements such as P, N and Mg. This is also a contributor to the discrepancy between apparent and absolute values.

Secondly, we conclude that where glare [92] has not been accounted for, material density will be compromised. Not only does the glare correction impact sharpness of edge profiles in the final reconstruction, but also it affects reconstructed density in homogeneous areas via a low-frequency component, as demonstrated in Fig. 3.3(D)(E). This low frequency component is sometimes discarded in detector characterization literature by subtracting a linear fit to the tails of the LSF [109]. Nevertheless, we found it impactful for

quantitative reconstruction acting similar to that of scatter. The effects of glare are also evident in projection data containing high attenuating objects with sharp gradients, such as imaging in presence of metal implants, the inclusion of a well-tuned glare kernel may be very impactful. Unfortunately, this effect is dependent on collimation and the imaged object. However, we demonstrated that a reasonably simple calibration of glare could generate reproducible reconstructions for a wide range of phantoms without introducing significant streaking artifacts. The exact calibration techniques can be changed according to needs of the reader. The glare model we used is shift-invariant. However, we acknowledge that advanced blur models, such as shift-variant focal spot blur [64], may be important for high-resolution applications and specialized imaging geometries with high magnification.

Third, our MC scatter simulation does not consider material mixtures in its forward projection. Despite matching total effective attenuation to the object model used in polyPL, Compton scatter may be underestimated. This may account for residual streaking in phantom W2, where large water cylinder and high attenuation inserts resulted in high SPR of >1 . However, in realistic scans of the proximal extremities without implants, our simplified simulation should suffice.

In summary, artifact suppression and reproducibility rely on improving the consistency between reconstructed image and projection data. One way to facilitate this consistency is by augmenting the projection data to correct for physical imperfections in acquisition, such as in the case of glare and scatter. Another way is to improve the forward model to better account for how projections are generated by a given image, such as using a realistic object model for tissue composition in polyPL. We showed that by properly

accounting for the imaging physics, reproducibility can be achieved even without additional in-scan calibration phantoms, as done in conventional qCT. This should inform future work on quantitative reconstruction. For example, in the case of scans including orthopedic implants, the forward model could be revised such that voxels surrounding the implant could be parametrized with water-steel mixtures, while conventional water-bone mixture is used for surround tissue [116].

III.E Conclusion

We showed that extremity CBCT could be used for measuring BMD accurately and reproducibly. This was achieved using a polyenergetic forward model, an explicit object model for biological tissue composition, deconvolution-based glare correction, and MC scatter simulation. Quantitation was validated in the paradigm of a classic reproducibility study, with benchtop scans of phantoms containing Ca inserts of known concentration and cadaveric knee specimen in multiple configurations.

Chapter IV

Improving System Spatial Resolution for Quantification of Trabecular Microstructure

This chapter describes a task-based optimization of CMOS FPDs for visualization of bone microstructure. The optimization leverages noise-resolution tradeoff associated with scintillator thickness, given the small-pixel, low-noise characteristics of the CMOS sensor. The resulting detector configuration is incorporated into a prototype CMOS CBCT, which demonstrates improved correlation of bone microstructure metrics with that of gold standard (μ CT).

IV.A Introduction

Trabecular and cortical microarchitecture [90], [117]–[120] are biomarkers of bone health, with indices of bone microstructure found to improve prediction of fracture risk in OP [73], [74], [90], [121]–[127] compared to BMD. In OA, alterations in trabecular microarchitecture of subchondral bone often precede cartilage degeneration, [128]–[130] motivating investigation of structural metrics as an early biomarker of disease. However, trabecular features typically measure 0.05-0.2 mm, [131] so ultra-high-resolution pre-clinical micro-CT [132] remains the gold standard for bone morphometry. To enable *in vivo* measurements of microstructure in patients, a dedicated high-resolution peripheral quantitative CT (HR-pQCT) system has been developed [133], [134] for use in distal radius and ulna. Applications of HR-pQCT in more attenuating body sites, such as the knee,

require long scan times (~20 min) and have limited axial field-of-view (FOV) of ~6 cm [135]. Accurate characterization of bone morphometry via MDCT, CBCT, and/or MRI could greatly enhance the clinical utilization of such quantitative biomarkers.

Despite their limited spatial resolution, conventional MDCT and a-Si:H FPD CBCT systems have been shown to achieve statistically significant correlation with gold standard micro-CT for a variety of trabecular metrics.[73], [74], [126] In extremity CBCT, Pearson R-coefficient of 0.9 for Bone Volume (BV/TV), 0.66 for Trabecular Thickness (Tb.Th) and 0.68 for Trabecular Spacing (Tb.Sp) were found in comparison to micro-CT. [119] While promising, this result suggests that application of CBCT in quantitative bone imaging would benefit from further improvement in spatial resolution. Complementary metal-oxide semiconductor (CMOS) x-ray detectors offer a compelling alternative to a-Si:H FPDs for such high-resolution applications, owing to their higher readout speed, up to 10x lower electronic noise, and finer pixel pitch (~0.05-0.1 mm) [29], [136]–[138].

In the recent years, CMOS detectors emerged as an attractive option for indirect detection x-ray detectors. Early interest in medical applications of CMOS sensors has been primarily focused on breast imaging. In digital breast tomosynthesis (DBT), Choi *et al* [139] and Patel *et al.* [140] showed that the low electronic noise of CMOS detectors enables improved angular sampling by lowering the required dose per frame. In breast CBCT, Gazi *et al* [141] and Shen *et al* [138] demonstrated 125% increase in system limiting resolution and 45% enhancement in visibility of microcalcifications, respectively, compared to conventional a-Si:H FPDs. A model by Zhao *et al* [142] found that a commercial CMOS sensor operated in a low capacity, low noise mode can improve contrast-to-noise ratio (CNR) of small microcalcifications (<0.2 mm) at lower mean glandular dose levels

compared to an existing commercial DBT system. This analysis was extended to include models of DBT image reconstruction, image display and human observer performance. [143] Vedantham *et al* [144] used task-based modelling to investigate the tradeoffs between pixel size and scintillator thickness in breast CBCT and determined that a 0.525 mm thick CsI:Tl scintillator (somewhat thinner than the 0.6-0.7 mm CsI:Tl typically used in FPDs) is optimal for detecting 0.22 mm microcalcifications.

Results of modeling and optimization studies in breast imaging are not directly applicable to extremity CBCT due to differences in x-ray spectrum (~90 kV in extremity CBCT compared to ~30 kV in DBT), object composition, and system geometry. We use a cascaded systems model developed for CMOS-based extremity CBCT to evaluate performance in high-resolution applications using task-based detectability index framework [118], [145]. In particular, we investigate the potential benefits of reducing the thickness of the CsI:Tl scintillator below 0.6-0.7 mm as currently used in a-Si:H FPDs and CMOS sensors for CBCT. Thinner CsI:Tl reduces scintillator blur (potentially better matching the blur associated with the finer pixel aperture), but leads to increased noise due to diminished x-ray absorption, especially for the relatively high x-ray energies used in extremity imaging. For high-frequency tasks, however, the benefits of enhanced spatial resolution may outweigh the impact of elevated noise, resulting in improved detectability. Preliminary results [120] indicate that the visibility of trabeculae can indeed be improved using a 0.4 mm scintillator compared to the standard 0.7 mm thickness.

The work reported below extends the simulation and experimental studies reported in Ref. 15 to provide a more detailed analysis of CMOS detector performance across a broader scope of system parameters and imaging geometries using a new, continuously

tunable spectrum of imaging tasks representative of trabecular bone morphometry. The main contributions of this work include: (a) a model of a CMOS x-ray sensor that incorporates the effects of scintillator thickness; (b) a study of detectability in extremity CBCT as a function of feature size, pixel size, electronic noise, CsI:Tl thickness, focal spot size, dose, and system magnification; and (c) experimental validation in phantom and cadaver studies using two CMOS detectors, one with the current standard CsI:Tl thicknesses of 0.7 mm (denoted C700), and one (C400) custom-made with 0.4 mm thick CsI:Tl. This direct experimental comparison provides new insight into the tradeoffs between resolution, noise, and imaging task that govern the choice of scintillator thickness for high-resolution applications of CMOS sensors. The results motivated implementation of a CMOS detector on a prototype high-resolution extremity CBCT system.

IV.B Methods

IV.B.1 Task-Based Evaluation of CMOS Detectors

The performance of CMOS detectors in extremity CBCT imaging was investigated in terms of MTF, NPS, DQE, and detectability index (1.15):

$$d'^2 = \frac{(\int \int T_{system}^2 W_{task}^2 dudv)^2}{\int \int NNPS \cdot T_{system}^2 W_{task}^2 dudv} \quad (4.1)$$

The NPW observer model (Eq. 4.1) has been shown to agree reasonably well with human observer performance in CBCT for a wide range of simple imaging tasks [146]. The task function, W_{task} , is a frequency-domain specification of the feature(s) of interest in detection or discrimination. The $NNPS$ is the normalized NPS, and T_{system} is the MTF of

the system. The NPS and MTF were obtained from cascaded systems analysis of CMOS x-ray sensors, as described below. Since the focus of the current study is on the effects of detector design on imaging performance, the modeling was performed in the 2D projection domain and did not include signal and noise propagation in the 3D reconstruction cascade [147], [148]. The NPS in the denominator of Eq. (4.1) is normalized by the mean detector signal squared. As shown by Tward *et al* [147], the resulting NNPS is equivalent to the NPS of log-corrected projection data. The form in Eq. (4.1) therefore represents a detection task in log-corrected projection data. The detectability index was investigated as a function of scintillator thickness, pixel size, electronic noise, scan dose, focal spot size, and imaging task.

The description of theoretical methods is organized as follows: Sec. IV.B.1.1 introduces the basic cascaded systems model and defines fixed system parameters and key variables. Sec. IV.B.1.2 describes the modeling of scintillator thickness-dependent system gain, including quantum detection efficiency, generation of optical photons in the scintillator, depth-dependent scintillator escape fraction, and coupling efficiency. Models of thickness-dependent scintillator blur, focal spot blur, and system MTF are discussed in Sec. IV.B.1.3. Sec. IV.B.1.4 concerns the NPS and includes discussion of Swank factor and assumptions regarding Lubberts effect. Sec. IV.B.1.5 introduces task functions pertinent to high-resolution extremity imaging, and Sec. IV.B.1.6 outlines the simulation studies evaluating task-based detectability. The cascaded systems model is consistent with a significant body of previous work modeling a-Si:H FPDs [145], [149]–[152] and more recently CMOS systems [142], [153] with details below emphasizing aspects of the model related to scintillator thickness and pixel size.

IV.B.1.1 Cascaded Systems Model

The analytical detector model consisted of the following 7 stages: [145] (1) x-ray absorption in the scintillator (QDE), (2) generation of light quanta (quantum gain), (3) light spread in the scintillator (including the spread of K-fluorescence photons [154]), (4) coupling to the photodiode, (5) integration by pixel aperture, (6) sampling, and (7) readout with additive noise. Effects of scatter were not considered in the model. For input-quantum-limited conditions, scatter imparts a scaling of the NPS by a factor of (1+SPR) [155]. In the majority of studies performed here, various detector configurations are compared at the same system geometry and thus the same SPR. Consequently, it is assumed that the omission of scatter from the model has only minor impact on the metrics of relative detector performance investigated in this study. The parameters of the model are listed in Table 4.1A and 4.1B, with nominal values reflecting two Xineos-3030HR CMOS detectors (Teledyne DALSA) with CsI:Tl thicknesses of 0.4 mm and 0.7 mm used in experimental studies (Sec. IV.B.2).

Constants	Notation	Value	Ref.
Beam energy	E_{max}	90 kV	[151]
Fluence per exposure pre-object	\bar{q}_0/X	2.58×10^5 x-rays/mR/mm ²	[24]
Source-detector distance	SDD	560 mm	[152]
Source-axis distance	SAD	431 mm	
Total filtration (inh. + added)		3.4 mm Al+0.2 mm Cu	
Object	-	8 cm water + 7.3 cm spongiosa + 0.4 cm cortical bone	[25]
CsI:Tl density	ρ_{CsI}	4.51 g/cm ³	-
Packing fraction	f_{CsI}	0.7	-
K-edge energy	E_K	35 keV	
K-fluorescence probability	ξ	0.83	[156]
K-fluorescence yield	ω	0.87	
Work function	W	55.6 optical photons/keV	
Coupling efficiency (e ⁻ /photon)	\bar{g}_4	0.59	-
Fill factor	f_{pix}	0.85	
Electronics noise	σ_{add}	390 e ⁻	-
Conversion gain	k	139 e ⁻ /ADU	

Table 4.1A Glossary of terms and symbols in the cascaded systems model. Model constants include fundamental physical quantities, geometry and typical operating parameters of the extremity CBCT system, and detector parameters that are independent of CsI:Tl thickness.

Key Variables	Notation	Range (nominal value)	Ref.
Dose (mGy)	D	15-40 (15)	-
Focal spot size (mm)	a_{spot}	0.05-0.7 (0.5)	-
CsI:Tl thickness (mm)	t_{CsI}	0.35-0.75 (0.7)	-
Pixel size (mm)	a_{pix}	0.05-0.4 (0.099)	-
Magnification	M	1.1-2.1 (1.3)	-

Derived Quantities	Notation	Nominal Values (at $a_{pix}=0.099$ mm)	Ref.
		$@ t_{CsI} = 0.4 \text{ mm}$ $@ t_{CsI} = 0.7 \text{ mm}$	
Focal spot blur	T_{spot}	Eq. (4.10)	-
Quantum detection efficiency	\overline{g}_1	0.58	0.75
K-fluorescence reabsorption	f_K	0.66	0.77
Escape fraction	η_{esc}	0.40	0.37
Quantum gain (photons/X-ray)	\overline{g}_2	798	734
Photodiode aperture (mm)	a_{pd}	0.0913	-
Gain ($\text{mm}^3 \text{ e}^-/\text{photon}$)	G	2.28	2.71
Scintillator blur	T_3	Eq. (4.8)	-
Total K-fluor. blur	T_{Ktot}	Eq. (4.9)	-
Pixel aperture	T_5	-	-
Sampling function	III_6	-	-

Table 4.1B CMOS detector performance was analyzed as a function of quantities denoted as key variables. The derived quantities are functions of the model constants and key variables; their nominal values are given at two detector thicknesses corresponding to the CMOS sensors used in experimental studies, assuming all other parameters are at their nominal value.

IV.B.1.2 System Gain

Using the definitions in Table 4.1AB, the system gain G ($\text{mm}^3 \text{ e-}/\text{photon}$) is given by:

$$G(t_{CSI}) = a_{pd}^2 \bar{g}_1(t_{CSI}) \bar{g}_2(t_{CSI}) \bar{g}_4 \quad (4.2)$$

where t_{CSI} is scintillator thickness. The mean quantum detection efficiency, $\bar{g}_1(t_{CSI})$, is calculated by dividing the scintillator into a series of 200 slabs (thickness denoted Δz) and integrating over the slab depth (z) measured from the incident surface [150]:

$$g_1(E, z) = e^{-z \cdot \rho_{CSI} \cdot f_{CSI} \cdot \left(\frac{\mu}{\rho}\right)_{CSI}(E)} \left(1 - e^{-\Delta z \cdot \rho_{CSI} \cdot f_{CSI} \cdot \left(\frac{\mu}{\rho}\right)_{CSI}(E)}\right) \quad (4.3a)$$

$$g_1(E, t_{CSI}) = \int_0^{t_{CSI}} g_1(E, z) dz \quad (4.3b)$$

$$\bar{g}_1(t_{CSI}) = \frac{\int_0^{E_{max}} q_{det}(E) g_1(E, t_{CSI}) dE}{\int_0^{E_{max}} q_{det}(E) dE} \quad (4.3c)$$

where $\left(\frac{\mu}{\rho}\right)_{CSI}(E)$ is the mass attenuation coefficient of CsI:Tl and $q_{det}(E)$ is the spectrum at the detector (input spectrum $q_0(E)$ attenuated by the object). Within each slab, three parallel pathways for the quantum gain g_2 are considered: pathway A (gain denoted by \bar{g}_{2A}) describes local absorption of an X-ray photon without K-fluorescence; pathway B (\bar{g}_{2B}) accounts for locally absorbed energy resulting from K-fluorescent events; and pathway C (\bar{g}_{2C}) involves K-fluorescence x-ray photons produced locally but absorbed at a remote site, with the associated MTF denoted T_K [154]:

$$\bar{g}_{2A}(t_{CSI}) = \frac{\int_0^{E_{max}} q_{det}(E) \int_0^{t_{CSI}} g_1(E, z) \eta_{esc}(t_{CSI}, z) (1 - \xi\omega) E W dz dE}{\int_0^{E_{max}} q_{det}(E) g_1(E, t_{CSI}) (1 - \xi\omega) dE} \quad (4.4a)$$

$$\overline{g_{2B}}(t_{CSI}) = \frac{\int_0^{E_{max}} q_{det}(E) \int_0^{t_{CSI}} g_1(E, z) \eta_{esc}(t_{CSI}, z) \xi \omega (E - E_K) W dz dE}{\int_0^{E_{max}} q_{det}(E) g_1(E, t_{CSI}) \xi \omega dE} \quad (4.4b)$$

$$\overline{g_{2C}}(t_{CSI}) = \frac{\int_0^{E_{max}} q_{det}(E) \int_0^{t_{CSI}} g_1(E, z) \eta_{esc}(t_{CSI}, z) \xi \omega f_K(t_{CSI}) E_K W dz dE}{\int_0^{E_{max}} q_{det}(E) g_1(E, t_{CSI}) \xi \omega dE} \quad (4.4c)$$

The escape fraction $\eta_{esc}(t_{CSI}, z)$ refers to the fraction of optical photons that reach the scintillator exit surface and are subsequently coupled to the photodiode with efficiency $\overline{g_4}$. A linear fit to escape fraction estimated by Howansky *et al.* [157] for a scintillator with reflective backing was used to compute η_{esc} :

$$\eta_{esc}(t_{CSI}, z) = -0.185(t_{CSI} - z) + 0.312 \quad (4.5)$$

The quantum gain for each parallel pathway $\overline{g_{2A}}(t_{CSI})$, $\overline{g_{2B}}(t_{CSI})$ and $\overline{g_{2C}}(t_{CSI})$ are combined to form the total gain $\overline{g_2}(t_{CSI})$ [154]:

$$\overline{g_2}(t_{CSI}) = (1 - \xi \omega) \overline{g_{2A}}(t_{CSI}) + \xi \omega \overline{g_{2B}}(t_{CSI}) + \xi \omega f_K \overline{g_{2C}}(t_{CSI}) \quad (4.6)$$

The loss of photons in the FOP between the scintillator exit surface and the photodiode (e.g. due to FOP and optical glue) and the conversion efficiency from optical photons to electrons are described together by the optical coupling efficiency $\overline{g_4}$. The coupling efficiency was treated as a constant across all detector configurations. The value of $\overline{g_4}$ was determined empirically to yield a good match between the model and measurements of system gain and noise power spectra (Sec. IV.B.2). It was found that the estimates of the coupling efficiency were different for the two detectors: $\overline{g_4} = 0.7$ for C400 and $\overline{g_4} = 0.48$ for C700. This discrepancy is likely due to variations in CsI:Tl deposition techniques and in the quality of the interface between the CMOS and scintillator. Such variability represents differences in the manufacturing process rather than true dependence of $\overline{g_4}$ on

scintillator thickness. Therefore, the mean of the two empirical estimates ($\bar{g}_4 = 0.59$) was used in the detector model to approximate an average coupling efficiency of a CMOS detector of the type used in the experimental studies. The coupling efficiency values found for the two detectors were lower than those reported for a-Si:H FPDs ($\bar{g}_4 \approx 0.8$), likely due to light losses in the FOP, which is not used in a-Si:H FPDs.

Fig 4.1A illustrates the thickness-dependent QDE (black line), escape fraction $\bar{\eta}_{esc} = \int_0^{t_{CSL}} \eta_{esc}(t_{CSL}, z) dz$ (dashed black line), and quantum gain (gray line) as functions of CsI:Tl thickness. As anticipated, QDE increases with increasing thickness; however, quantum gain decreases for thicker CsI:Tl because of reduced escape efficiency. Overall, the $\bar{g}_1 \bar{g}_2$ product is $\sim 20\%$ higher for 0.7 mm CsI:Tl than 0.4 mm CsI:Tl.

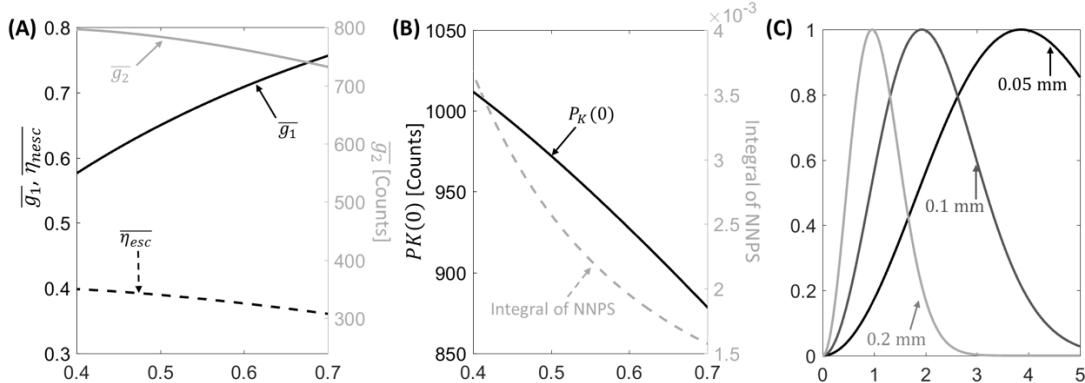


Figure 4.1 (A) Quantum detection efficiency, escape efficiency (left vertical axis, black lines), and scintillator gain (right vertical axis, gray line) computed as a function of scintillator thickness. (B) Zero-frequency P_K and integral of detector NNPS over the Nyquist frequency range (gray line) as a function of CsI:Tl thickness. (C) Examples of task functions (Eq. 4.14, assuming $C=1$) for three feature sizes: 0.05 mm (solid black line), 0.1 mm (solid dark gray line), and 0.2 mm (solid light gray line). The tasks emphasize distinct frequency bands depending on the underlying feature size.

IV.B.1.3 System MTF, Detector Blur, and Focal Spot Blur

System MTF is defined as the product of detector and focal spot MTFs:
 $T_{system}(u, v; t_{CSI}, a_{spot}) = T_{det}(u, v; t_{CSI}) \cdot T_{spot}(u, v; a_{spot})$. The detector MTF, T_{det} , is given by:

$$T_{det}(u, v; t_{CSI}) = T_3(u, v; t_{CSI})T_{Ktot}(u, v; t_{CSI})T_5(u, v) \quad (4.7)$$

where T_3 denotes scintillator blur, T_{Ktot} is the blur associated with the spread of K-fluorescence photons, and T_5 is the aperture function $T_5(u, v) = |sinc(a_{pd}u)sinc(a_{pd}v)|$. Both T_3 and T_{Ktot} are affected by scintillator thickness t_{CSI} . The scintillator blur is modeled using a two-component form that consists of an exponential term to account for a homogeneous, unstructured layer of CsI:Tl deposited on the surface of the scintillator, and a Lorentzian term describing the blur in the structured, columnar CsI:Tl layer [41]:

$$T_3(u, v; t_{CSI}) = Ae^{-|\sqrt{u^2+v^2}|/B} + \frac{1-A}{1+H(t_{CSI}) \cdot (u^2+v^2)} \quad (4.8)$$

where the terms A and B are independent of CsI:Tl thickness, and $H(t_{CSI})$ is a polynomial function of scintillator thickness. Each parameter was obtained from an empirical fit to measured presampling MTFs of CMOS detectors with 0.4 mm and 0.7 mm CsI:Tl (Sec. IV.B.1.3).

The total K-fluorescence blur, T_{Ktot} , weighs the K-fluorescence blur T_K by the gains of the individual pathways of g_2 [149]:

$$T_{Ktot}(u, v; t_{CSI}) = \frac{(1-\xi\omega)\overline{g_{2a}} + \xi\omega\overline{g_{2b}} + \xi\omega\overline{g_{2c}}f_K(t_{CSI})T_K(u, v; t_{CSI})}{\overline{g_2}} \quad (4.9)$$

where $f_K(t_{CSI})$ is the probability of remote absorption for an K-fluorescence photon, and $T_K(t_{CSI})$ is the spread associated with that absorption. Both $f_K(t_{CSI})$ and $T_K(u, v; t_{CSI})$ were computed from analytical models of Chan *et al.* and Que *et al.*[158], [159].

Similar to previous cascaded systems studies of extremity CBCT, a simplified Gaussian model of focal spot blur T_{spot} was adopted [151], [160]:

$$T_{spot}(u, v; a_{spot}) = \exp(-\pi(M - 1)^2 a_{spot}^2 (u^2 + v^2)) \quad (4.10)$$

where $M = SDD/SAD$ and the nominal value of parameter a_{spot} was set to 0.5 mm. Note that while the a_{spot} parameter in the simulations has a unit of mm, the x-ray source employed in the experiments is characterized using the focal spot index (IEC 336 standard, denoted FS), representing a range of focal spot size.

IV.B.1.4 Noise Power Spectrum in Projection Images

Using the gain and blur terms defined above, the NPS of a 2D projection image is:

$$S_{det} = \overline{q_0} a_{pd}^4 \overline{g_1} \overline{g_2} \overline{g_4} (1 + \overline{g_4} P_K T_3^2) T_5^2 * * III_6 + S_{add} \quad (4.11)$$

where the functional parameters were omitted for simplicity and $a_{pd} = \sqrt{f_{pix}} a_{pix}$ is the active pixel area. Convolution with the comb function $III_6 = \sum_N \delta(N/a_{pix})$ represents aliasing of quantum noise due to pixel sampling. S_{add} is the power spectrum of additive noise ($\sigma_{add}^2 a_{pix}^2$). P_K is a noise term associated with quantum gain as described by Richard *et al.* [161]. In the case of no K-fluorescence, $P_K = \overline{g_{2A}} + \varepsilon_{g_{2A}}$, where $\varepsilon_{g_{2A}}$ is the Poisson excess associated with local x-ray absorption. Considering K-fluorescence, P_K combines the gain and Poisson excess for the three parallel cascades of $\overline{g_2}$ (Eq. 4.4), as well as a cross term in noise between pathways B and C. The Poisson excess in each path is computed from the gain (Eq. 4.4) and the path-specific Swank factor. Each path-specific Swank factor

is computed as the product of a radiological Swank factor and an optical Swank factor. The CsI:Tl thickness-dependent radiological component is associated with polyenergetic absorption and was computed from the result of Zhao *et al.* [41] The optical component is due to light losses in the scintillator. For simplicity, it is assumed to be independent of CsI:Tl thickness and equal to 0.95, giving good agreement with the measured NPS on the C400 and C700 CMOS detectors. This value also agrees fairly well with an approximate estimate of the optical Swank factor I_{opt} given by the following equation from Lubinsky *et al* [42]:

$$I_{OPT}(t_{CsI}) = \frac{\langle \eta_{esc}(t_{CsI}, z) \rangle^2}{\langle \eta_{esc}(t_{CsI}, z)^2 \rangle} \quad (4.12)$$

This formula yields $I_{opt} = 0.98$ for 0.7 mm CsI:Tl and $I_{opt} = 0.99$ for 0.4 mm CsI:Tl. While the above equation is approximate and was derived under certain simplifying assumptions, it further justifies modeling I_{opt} as largely independent of t_{CsI} and supports the empirical value of 0.95 used in this study.

With the above definitions, the normalized NPS (NNPS) in Eq. 4.1 is given by:

$$NNPS = S_{det}/(a_{pd}^2 \bar{q}_0 \bar{g}_1 \bar{g}_2 \bar{g}_4)^2 \quad (4.13)$$

The noise model accounts for the effects of scintillator thickness on detector gain and *MTF*. However, the Lubberts effect [162], [163] (i.e., noise amplification due to the depth-dependent light spread from individual scintillation events) is assumed to be negligible. This assumption is supported by experimental results for columnar CsI:Tl scintillators [164], showing negligible contribution from the Lubberts effect (ratio of the square of detector *MTF* to *NNPS* very close to unity) over a broad range of scintillator thickness (~0.2 mm to ~0.6 mm).

The dashed gray line in Fig. 4.1B shows the integral of NNPS over the Nyquist frequency range as a function of scintillator thickness at an exposure of 0.126 mAs. Despite the increased escape fraction that partly compensates for the reduced QDE of thin scintillators (Fig. 4.1A) in mean signal, the projection noise substantially increases with decreasing CsI:Tl thickness. The plot of the zero-frequency P_K (black line) in Fig. 4.1B explains this behavior: the improved $\overline{g_2}$ of thin scintillators leads to increasing contribution of the associated conversion noise encapsulated by P_K .

IV.B.1.5 Imaging Task

We considered task functions representing discrimination of two Gaussian stimuli of different widths, defined in the detector plane as:

$$W_{task}(u, v) = CM^2 \left[e^{-2\pi^2(kMa_{obj})^2(u^2+v^2)} - e^{-2\pi^2(Ma_{obj})^2(u^2+v^2)} \right] \quad (4.14)$$

where C is a contrast term, M is the system magnification, a_{obj} is the feature size measured in the object plane, and k is a parameter that determines the relative width of the Gaussian stimuli. Here, k was set at 0.8, and a_{obj} was varied to simulate a range of anatomical feature sizes and associated spatial frequency bands (Fig. 4.1C). In the context of trabecular bone imaging, an increase in detectability of the task in Eq. (4.14) indicates improved ability to discriminate trabeculae (or trabecular cavities) that differ in size by 20% and could thus indicate improved assessment of change in trabecular thickness (or trabecular spacing).

Holding other parameters of Eq. (4.14) fixed, a change in feature size, a_{obj} , implies a change in the area under W_{task}^2 - i.e., signal power. To account for this scaling, the results below are reported in terms of relative d^2 , where the detectability for a given feature size is normalized to a reference value (e.g., the value at nominal system parameters) for the

same feature size. Detector performance for various tasks (feature sizes) is then compared in terms of this relative detectability, and the contrast term C cancels out for all cases (can be ignored) and was set to unity.

IV.B.1.6 Simulation Studies

Detectability for the task function of Eq. (4.14) for feature sizes ranging from 0.03 to 0.2 mm [131] was studied in relation to five system parameters (Table 4.1, Key Variables): pixel size (a_{pix}), scintillator thickness (t_{CSI}), focal spot blur (a_{spot}), magnification (M) and bare-beam x-ray exposure.

The x-ray spectrum ($q_0(E)$) was obtained using the Spektr 3.0 [24] implementation of TASMICS [22] for a beam energy of 90 kV. Tube inherent filtration was found experimentally by estimating Al thickness that achieved best match of detector entrance dose between Spektr simulations and measurements on the benchtop x-ray source (Sec. 2.2) performed with a Si diode (AccuDose, RadCal Corp., Monrovia, CA) for 70, 80, 90, and 100 kV beams. The added filtration (Table 4.1) was the same as the filters applied in the experimental setup. Additionally, differences in tube output (mGy/mAs) between the Spektr 3.0 model and the source used in the experiments were accounted for by fitting a constant scale factor.

The nominal SDD was set to 560 mm and nominal magnification was set to 1.3 to emulate the extremity CBCT scanner [151] and test bench setup. The x-ray spectrum at the detector, $q_{det}(E)$, was obtained from the input spectrum, $q_0(E)$, attenuated by material approximating a human knee: 8 cm water, 7.6 cm spongiosa and 0.4 cm cortical bone. To investigate the effects of system geometry in hypothetical future system designs, additional geometries with varying magnification were simulated by holding SAD fixed and varying

SDD. Tube output was kept the same in all configurations so that the patient dose was constant.

IV.B.2 Benchtop Experimental Setup

Experimental evaluation of CMOS-based extremity CBCT was performed on an imaging test bench (Fig. 4.2). Two CMOS detectors (Xineos-3030HR, Teledyne DALSA, Eindhoven, NL) were tested, one with CsI:Tl thickness of 0.4 mm (C400) and one with thickness of 0.7 mm (C700). The detectors have a 30 x 30 cm² FOV, pixel size of 0.099 x 0.099 mm², 14-bit digitization, and frame rate of up to 30 fps for full resolution readout. The scintillators are deposited on FOP made with leaded glass. The same set of experiments was performed with both detectors, keeping the geometric configuration and acquisition parameters fixed to provide head-to-head performance comparison for two CsI:Tl thicknesses. The detectors were operated in high-full-well sensitivity mode.

The test bench employed an IMD RTM 37 rotating anode x-ray source (IMD, Grassobbio, IT) with 3 kW power and nominal focal spot of 0.3 FS (IEC336). The beam energy in all experiments was 90 kV (0.2 mm Cu + 2 mm Al added filtration). The SDD was fixed at ~560 mm and the SAD was ~430 mm, consistent with the geometry of a clinical extremity CBCT developed at our institution.[165] The object to be imaged was rotated using a Velmex B4800 rotation stage (Velmex, Bloomfield, NY), and projections were obtained in a step-and-shoot mode of x-ray pulses synchronized with detector readout.

We observed evidence of possible direct x-ray interaction in the crystalline Si (due to the small fraction of x-rays that were not absorbed in the lead glass of the FOP) as bright, single-pixel outliers (“speckles”) apparent after gain and offset corrections. The speckles were isolated by applying a Laplacian filter to select pixels that deviated from the mean of

the filtered frame by more than 2 standard deviations. Such speckles identified by filtering (typically accounting for <2% of total pixels) were corrected by linear interpolation of nearest neighbors.

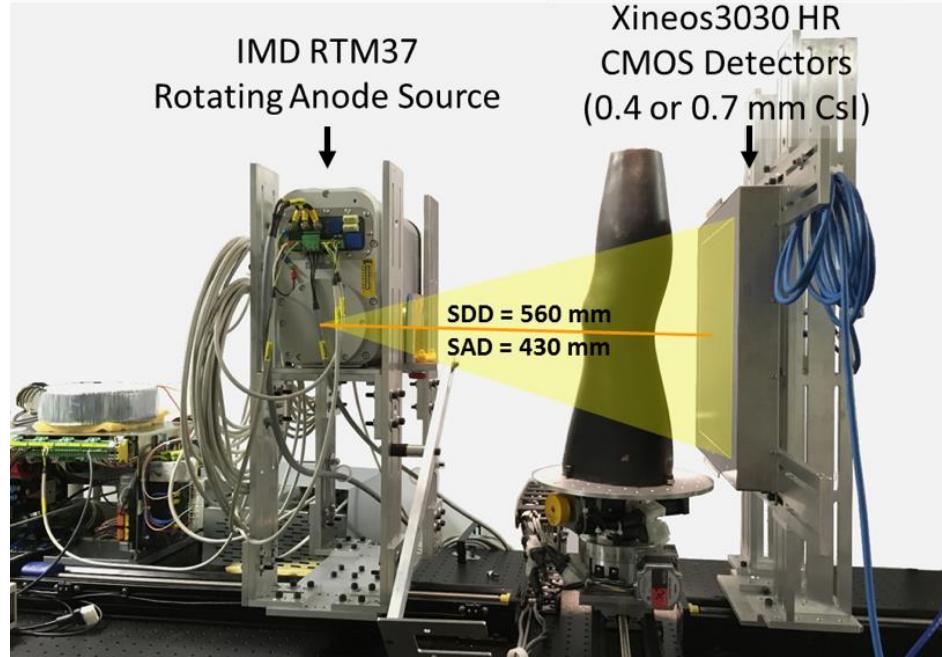


Figure 4.2 CBCT test-bench used in experimental studies. The bench was configured with two CMOS sensors, one with CsI:Tl thickness of 0.4 mm (C400) and one with thickness of 0.7 mm (C700).

IV.B.2.1 Measurement of Detector MTF, NPS, and Scan Dose

Measurements of detector MTF and NPS were performed with additional 2.5 mm Cu and 2 mm Al placed in the beam to simulate attenuation by 15 cm of water. Detector MTF was measured using a tungsten edge placed on the detector surface and imaged at $\sim 15^\circ$ angle to the pixel matrix. The exposure was set to deliver $\sim 90\%$ detector saturation and 100 frames were acquired, gain and offset corrected, and averaged. An oversampled ESF was obtained from the projection of the tilted tungsten edge following Samei *et al.* [109]. The ESF was oversampled at $20 \mu\text{m}$ intervals via bilinear interpolation, and the LSF was computed by numerical differentiation of the ESF. The presampling detector MTF was

obtained as the absolute value of a Fourier transform of the LSF, divided by a sinc function corresponding to the 20 μm binning.

The parameters of the scintillator blur model $T_3(u, v; t_{CSI})$ in Eq. (4.8) were estimated by a least-squares fit to the measured MTFs of the two detectors (denoted MTF_{C400} for the C400 sensor and MTF_{C700} for the C700 sensor):

$$(A^*, B^*, H_{400}^*, H_{700}^*) \\ = \arg \min_{A, B, H_1, H_2} \left[\left(T_3|_{A, B, H_{400}} - \frac{MTF_{C400}}{T_5 T_{Ktot400}} \right)^2 + \left(T_3|_{A, B, H_{700}} - \frac{MTF_{C700}}{T_5 T_{Ktot700}} \right)^2 \right] \quad (4.15)$$

where $T_{Ktot400}$ and $T_{Ktot700}$ are the K-fluorescence blur functions for the two CsI:Tl thicknesses computed from Eq. (4.9). The dependence of T_3 on CsI:Tl thickness is encapsulated in the function $H(t_{CSI})$ in Eq. (4.8), assumed to follow a polynomial of the form $H(t_{CSI}) = h_1 \cdot t_{CSI}^2 + h_2 \cdot t_{CSI}$. The values of the function $H(t_{CSI})$ at the CsI:Tl thicknesses of the two detectors, H_{400} and H_{700} , were obtained through the fit in Eq. (4.15). The parameters of the polynomial model h_1 and h_2 were estimated from H_{400} and H_{700} by an additional fitting step, resulting in the thickness-dependent $H(t_{CSI}) = 0.35t_{CSI}^2 + 0.18t_{CSI}$. The factors A^* and B^* were found to be 0.20 (unit-less) and 1.42 mm^{-1} , respectively.

NNPS was computed for detector exposure ranging from 0.019-0.15 mAs/frame (from 3 ms/frame to 24 ms/frame at 6.3 mA). One hundred projections were obtained at each exposure, offset-corrected using a mean of 50 dark images and normalized by the mean of all air projections. Ninety-nine difference images (denoted Δ) were then obtained

from pairs of consecutive frames to remove residual structure or low-frequency trends. Effects of detector lag were assumed negligible based on manufacturer specification of 1st frame lag of 0.1%. [166] In each difference image, 144 non-overlapping regions of interest (ROIs) of 81x81 pixels were drawn. The NNPS was given by:

$$NNPS(u, v) = \frac{a_{pix}^2}{2n^2N} \sum_N |FFT[\Delta(x, y)]|^2 \quad (4.16)$$

where n is the side-length of each ROI, N is the total number of ROIs across all difference images, and the factor of 2 accounts for the use of difference images. DQE was computed as:

$$DQE = \frac{MTF^2}{\overline{q}_{det} \cdot NNPS} \quad (4.17)$$

The x-ray fluence \overline{q}_{det} was obtained from Spektr simulation using the source model described in Sec. IV.B.1.

Dose measurements were performed using three 16 cm diameter CTDI phantoms stacked to cover the longitudinal FOV of the benchtop CBCT. A Farmer chamber (AccuDose, Radcal Corp., Monrovia CA) was placed at the center of the x-ray beam in the central hole of the CTDI phantom. The central CTDI dose per projection was obtained by averaging 110 exposures of the CTDI phantom for x-ray techniques ranging 0.032-0.16 mAs/exposure (from 5 ms/fame to 25 ms/frame at 6.3 mA).

IV.B.2.2 Cadaver Study

CBCT scans of a cadaver knee were acquired using 420 projections evenly distributed over 220° (equal to 180° + fan angle and thus providing complete sampling). The x-ray technique was varied from 0.038–0.15 mAs/frame (from 6 ms/fame to 24 ms/frame at 6.3 mA). Image reconstruction used the FDK algorithm with a Hann-apodized

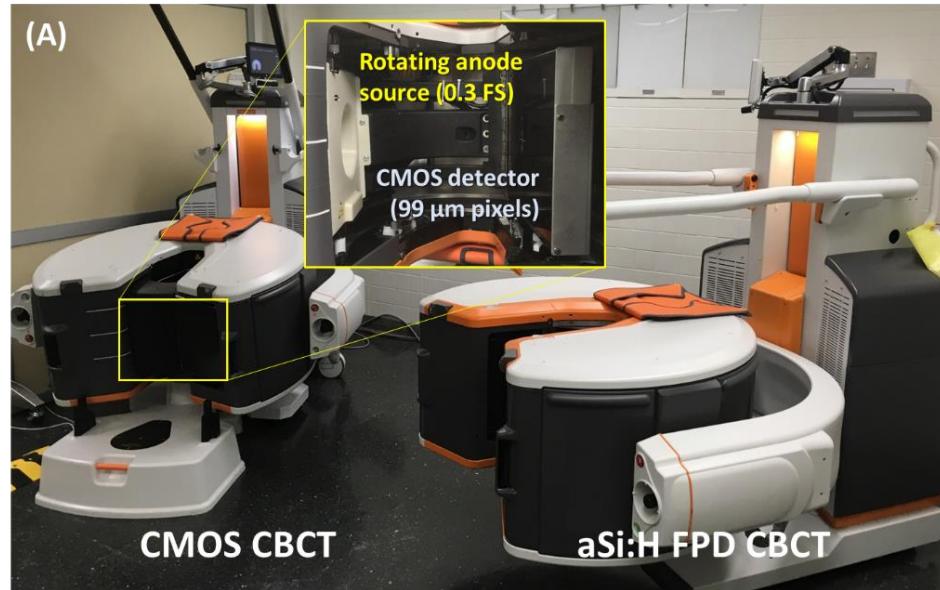
filter and Parker short scan weights [51], [167]. For high-resolution bone reconstruction, the filter cutoff was 0.9 of the Nyquist frequency (f_{Nyq}) and voxel size was 0.025 mm. Soft tissue reconstruction involved 4x4 software projection binning, filter cutoff at $0.8 \times f_{Nyq}$, and 0.3 mm voxel size. Scatter correction was performed assuming that the scatter in each projection view is uniform and equal to a fraction of the mean projection value (empirically set to 0.4) in a 5x5 pixel ROI at the center of the projection.

IV.B.3 Bone Microstructure Imaging on a Prototype CMOS-based CBCT

IV.B.3.1 A Prototype CMOS-Based Extremity CBCT Scanner

The CMOS-based extremity CBCT scanner was built on a gantry that is largely identical to a commercially available a-Si:H FPD extremity scanner (OnSight3D, Carestream Health, Rochester NY). Both are shown in Fig. 4.3(A). The gantry allows weight-bearing and non-weight-bearing imaging of the extremities and provides a compact geometry with source-detector distance of 560 mm and magnification of 1.3. The CMOS-based scanner utilizes a compact rotating anode source with 0.3 FS focal spot (IMD RTM 37) and a Dalsa Xineos3030 CMOS detector with 99 μm pixel size and a custom 400 μm thick CsI:Tl scintillator, selected in task-based optimization to maximize detectability in high-resolution imaging [7]. The imaging protocol involved 90 kV beam energy, 420 frames over a 210° short scan trajectory, and dose of ~15 mGy (central dose in a 16 cm CTDI phantom). The CMOS-based extremity CBCT was compared to an a-Si:H FPD CBCT (prototype OnSight3D system) using a Varian PaxScan2530 detector with pixel size of 0.137 mm and CsI:Tl scintillator thickness of 0.7 mm, and a stationary anode x-ray source with 0.5 FS focal spot. The imaging protocol also involved 420 frames and a 90 kV

beam. The dose measured from an ion chamber at the center of a CTDI phantom was 12 mGy.



(B)	CMOS CBCT	aSi:H CBCT
Detector	Dalsa Xineos3030 CMOS FPD	Varex PaxScan2530 aSi:H FPD
Pixel size	99 μm	139 μm
Scintillator thickness	\sim 0.4 mm	\sim 0.7 mm
e ⁻ noise	\sim 380 e ⁻	\sim 2000 e ⁻
Frame Rate	30 fps	7.5 fps
X-ray source	Single source, Rotating anode	Three source Stationary anode
Focal spot	0.3	0.5
Scan Dose	\sim 15 mGy	

Figure 4.3 (A) CMOS-based extremity CBCT system along with details of the CMOS CBCT gantry. (B) Key system and scan parameters for CMOS CBCT and a-Si:H FPD CBCT.

The CBCT scans for both systems were acquired at the native pixel size of their detectors. This acquisition mode is not commonly used in clinical imaging on the a-Si:H

FPD system, where 2x2 binning is typically employed to reduce scan time. As shown in Fig. 4.3B, the higher frame rate of the CMOS detector results in dramatic reduction in scan time, from ~60 sec for a-Si:H FPD CBCT to ~17 seconds for CMOS CBCT.

IV.B.3.2 Performance Evaluation

The spatial resolution of the prototype CMOS system was evaluated in comparison to the a-Si:H FPD CBCT in both projection and 3D reconstruction domains. Detector MTF was measured for both systems using a tungsten edge placed at the center of the detector surface at a tilt of ~2° and imaged for 100 frames with the gantry held stationary. An oversampled edge-spread profile was obtained to derive the line-spread functions and MTFs. The detector presampling MTF was then obtained by dividing out a sinc function corresponding to the bin size of the oversampled edge. For characterization of 3D resolution, a ~127 μm tungsten wire tensioned within a hollow plastic cylinder was scanned at the center of FOV of both systems and reconstructed with a ramp filter with Nyquist cutoff frequency at a fine voxel size of 25 μm.

Visual evaluation of image quality for the two systems was performed by an MSK radiologist in scans of a cadaveric ankle. The high-resolution reconstruction protocols used the Feldkamp algorithm and 25 μm voxel grid. The reconstruction filter was a Hann-apodized ramp. The cutoff frequency was the same for both systems and equal to ~3.3 mm⁻¹. The high-contrast reconstruction protocol involved binning of the projection data on each system to ~0.3 mm pixel size (2x2 binning on a-Si:H FPD CBCT, identical to the current clinical extremity CBCT protocol, and 3x3 binning on CMOS CBCT). The high-contrast reconstructions also utilized a Hann filter. The cutoffs were ~1.7 mm⁻¹ and the voxel size was 200 μm. The images from the two scanners were registered to facilitate visual

comparison. First, a rigid image-domain pre-registration was performed using initial reconstructions. Next, the transform found in the pre-registration was applied to image coordinate system during backprojection to match the position of the two reconstructed volumes with respect to anatomy. This approach avoids interpolation and blurring that would be present if the coordinate transform was applied directly to the reconstructed volume.

Finally, the performance of the two CBCT systems was compared in the following metrics of trabecular microarchitecture: BvTv, TbTh and TbSp. As mentioned in Chapter I, mean thickness is computed from assigning each voxel of the corresponding segmentation the diameter of the maximum fitting sphere (to either trabecular ridges or spacing) that includes it. An excised human ulna was scanned on a μ CT system (Skyscan1176, Bruker, BE) to obtain a gold standard characterization of trabecular detail at 28 μ m voxel size. The ulna was then imaged on the a-Si:H FPD extremity CBCT and on the CMOS-based extremity CBCT prototype. High-resolution CBCT reconstructions using 28 μ m voxels and Hann filters with cutoffs set at the Nyquist frequency of each system were obtained. The coordinate transform method described above was applied to obtain mutually registered CBCT and μ CT volumes without interpolation.

Trabecular analysis was performed in 16 corresponding ROIs (4x4x4 mm³) randomly selected in the CBCT and μ CT images. To segment bone voxels for the analysis, global thresholding (0.2 of max-min-normalized intensity for the entire image) followed by morphological closing (in a 5x5x5 neighborhood) was applied to μ CT ROIs. CBCT ROIs were segmented by sweeping the threshold level for each ROI to find a value that gave maximal Dice similarity index with the μ CT segmentation of the same ROI. This

approach uses a fairly simple segmentation algorithm to focus the comparison on the effects of baseline imaging performance of the two systems. Metric values were derived from the ROI segmentations and the correlation between CBCT and μ CT was assessed using Pearson's coefficient.

IV.C Results

IV.C.1 Detectability Index for Extremity CBCT

Fig. 4.4 presents the evaluation of task-based detectability with respect to detector pixel size, electronic noise, and scintillator thickness. The nominal extremity CBCT scanner geometry was assumed. Recognizing the multi-dimensional character of this evaluation, a set of two-dimensional optimizations holding other parameters fixed at nominal values was chosen to summarize key findings. The simulated bare-beam x-ray exposure was 0.126 mAs / projection. Based on Farmer chamber measurements (Sec. IV.B.2.1), the dose measured at the center of a CTDI phantom D for this exposure is estimated to be 15 mGy (assuming 420 projections/scan). This value is within the typical dose range for scan protocols in FPD-based extremity CBCT [10].

Fig. 4.4A shows relative detectability for a CMOS detector with $t_{CSI} = 0.7$ mm as a function of detector pixel size and task. For each task, detectability was normalized by the maximum detectability for that task across all values of a_{pix} . The model of focal spot blur was not included in this analysis to elucidate the impact of detector parameters. The optimal pixel size, indicated with a dashed line in Fig. 4.5A, varies from 0.07-0.11 mm

across the investigated range of a_{obj} (0.03-0.2 mm), with smaller feature size favoring smaller a_{pix} .

Fig. 4.4B shows the performance of a hypothetical a-Si:H FPD assumed identical to the CMOS detector except for 5x increased electronic noise. At each value of a_{obj} , the detectability values across all values of a_{pix} were normalized by the maximum detectability attained by the CMOS detector for the same task. The optimal pixel size for any feature size is larger for the a-Si:H FPD than for CMOS, compensating the increased contribution of electronic noise. For coarse features, this increased a_{pix} has only a minor effect on the maximum detectability, which is comparable to that achieved with the CMOS. However, for small feature sizes corresponding to imaging tasks in trabecular bone (0.05-0.1 mm), the maximum detectability of the a-Si:H FPD is only 20%-60% of the maximum d^2 of the CMOS. Due to its lower electronic noise, a CMOS detector is able to better exploit the increased Nyquist frequency and the improved aperture MTF (T_5) associated with reduced pixel size, resulting in better performance in high-resolution tasks.

Based on Fig. 4.4A, the pixel size of the CMOS sensor used in the experimental studies ($a_{pix} = 0.099$ mm) is seen to provide optimal or nearly optimal performance for a broad range of feature sizes. In Fig. 4.4C, the pixel size was thus fixed at 0.099 mm to investigate the effects of scintillator thickness. The detectability for each a_{obj} was normalized by their maximum value across the investigated range of t_{CSI} . For small feature size, the improvement in spatial resolution provided by a thin scintillator compensates for the increase in quantum noise resulting from reduced QDE (see Sec. IV.B.1.2 and Fig. 4.1), and optimal d^2 is achieved at CsI:Tl thickness of ~0.4 mm for a_{obj} of ~0.07 mm. As the feature size increases, there is less benefit to the improved MTF provided by thin t_{CSI} , and

the optimal d'^2 shifts toward thicker CsI:Tl to minimize quantum noise (increase QDE). Detectability at a_{obj} of ~ 0.13 mm is optimized for CsI:Tl thickness of ~ 0.7 mm, typical for current FPD and CMOS detectors for CBCT applications. This scintillator thickness delivers a balanced performance over a broad range of feature size, achieving d'^2 within $\sim 10\%$ of its maximum value (obtained at optimized t_{CsI}) for a_{obj} from $0.08 - 0.2$ mm. Using a thin scintillator benefits high-frequency tasks, but results in a relatively steep decline in detectability for coarse features ($\sim 20\%$ reduction compared to the value at optimized thickness). This effect, however, can be partly mitigated by additional projection binning and filtering to generate a lower resolution “soft-tissue” image separate from the full resolution “bone” reconstruction.

In Fig. 4.4D, the detectability for fine feature size ($a_{obj} = 0.06$ mm) consistent with visualization of trabecular bone is analyzed as a function of t_{CsI} and a_{pix} . Recognizing that different diagnostic tasks may be sensitive to features of different size, we chose 0.06 mm (as measured in the object domain in the nominal extremity CBCT geometry) as representative of the lower range of typical human trabeculae [168] (0.05 mm – 0.2 mm thick). The values of d'^2 are normalized to the maximum over the investigated range of t_{CsI} and a_{pix} , achieved at ~ 0.075 mm detector pixel size and ~ 0.35 mm CsI:Tl thickness. The dashed line marks the optimal scintillator thickness for each a_{pix} . Consistent with Fig. 4.4A, a pixel size of 0.05-0.1 mm is favored across the entire range of t_{CsI} , providing an optimal tradeoff between noise (increasing for smaller pixels) and aperture size and Nyquist frequency (improving with smaller pixels). There is a sharp drop in detectability at pixel size > 0.1 mm, approximately corresponding to the magnified feature size. The relative benefit of using small a_{pix} is diminished for thicker t_{CsI} , where scintillator blur

dominates the MTF. For the pixel size of current CMOS sensors (0.05-0.1 mm), the detectability for 0.06 mm feature size is maximized using CsI:Tl thicknesses of ~0.4 mm. The optimal thickness is relatively constant for pixels in the 0.05-0.1 mm range (despite the 2-fold change in a_{pix}), indicating diminishing benefits of improved scintillator sharpness compared to the increase in quantum noise (reduction in QDE). For the CMOS detector used in the experimental studies ($a_{pix} = 0.099$ mm), a ~20% improvement in d'^2 is anticipated at the optimal CsI:Tl thickness of ~0.35 mm compared to the commonly employed ~0.7 mm CsI:Tl. Beyond a_{pix} of ~0.1 mm, the optimal t_{CsI} changes almost linearly with pixel size, indicating that once the system resolution drops below this value, the optimization of d'^2 is driven primarily by reduction of projection quantum noise attained using thicker scintillators.

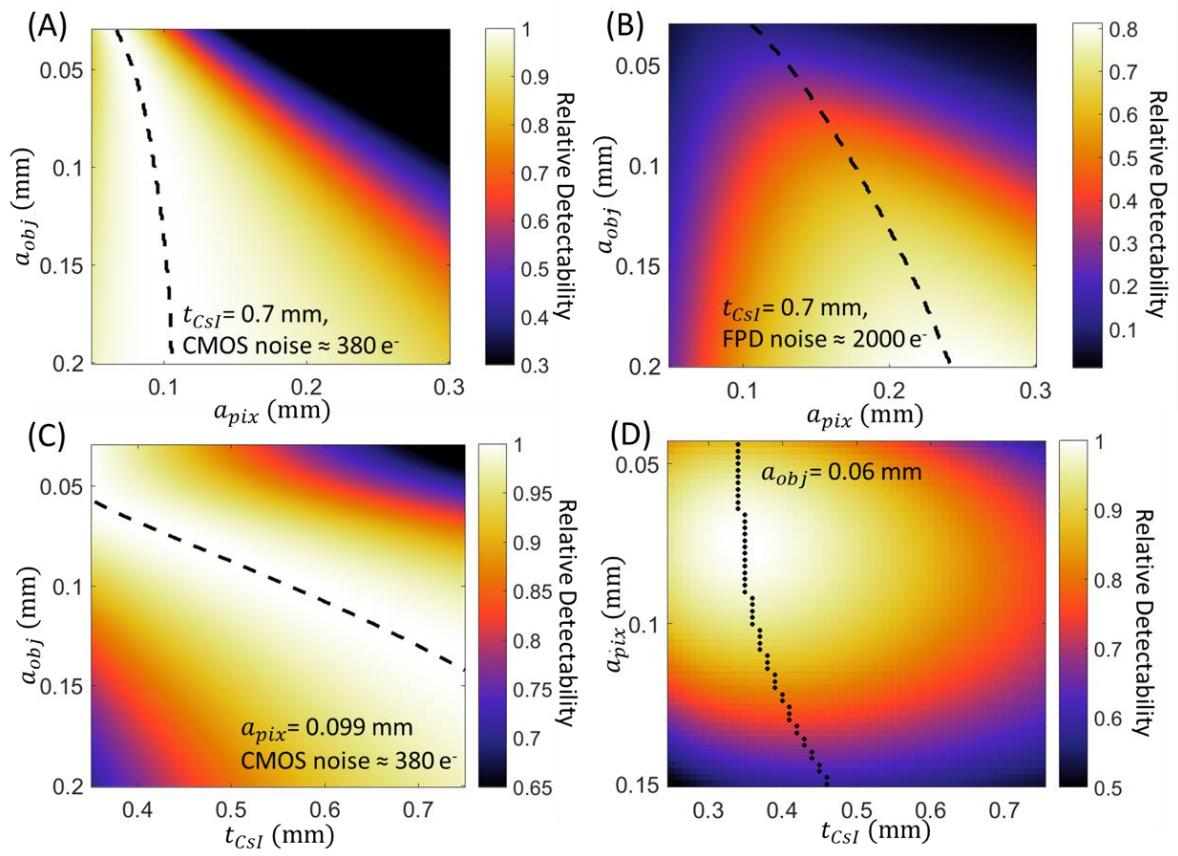


Figure 4.4 Task-based evaluation of CMOS detector performance in extremity CBCT. (A) Relative detectability for a range of feature sizes (vertical axis) as a function of pixel size. Scintillator thickness is assumed constant and equal to 0.7 mm. Detectability is normalized to the maximum value for each a_{obj} . Dashed lines indicate maximum d' for each feature size. (B) Ratio of d'^2 achieved with the same scintillator as (A), but at increased electronic noise consistent with an a-Si:H FPD, to maximum d'^2 attained for each a_{obj} by the low-noise CMOS detector of (A). (C) Relative detectability of the CMOS detector as a function of scintillator thickness and imaging task, normalized by maximum detectability achieved for each a_{obj} across the range of t_{CSI} . Pixel size is 0.099 mm. (D) Joint optimization of pixel size and scintillator thickness for a "trabecular" imaging task with feature size of 0.06 mm. The graph shows detectability of a CMOS detector (normalized by the maximum).

Fig. 4.5 uses the theoretical system model of Sec. IV.B.1.1 to investigate effects of imaging dose, focal spot size, and system geometry in CMOS-based extremity CBCT. The ratio of d'^2 of a CMOS detector with 0.099 mm pixels and $t_{CSI} = 0.4$ mm (C400) to d'^2 of the same detector with $t_{CSI} = 0.7$ mm (C700) is shown in Figure 4.5A for the nominal extremity CBCT system geometry. A dose range typical of extremity imaging is considered (5 - 30 mGy central CTDI scan dose). The d'^2 ratio is fairly independent of dose over a broad range. A thin scintillator is preferred for feature size < 0.1 mm, and a slight change in relative detectability of the two detectors occurs at very low dose (likely below practical imaging dose levels) and is due to the effects of electronic noise. At such low dose, the electronic noise terms favor the detector with an even thinner scintillator.

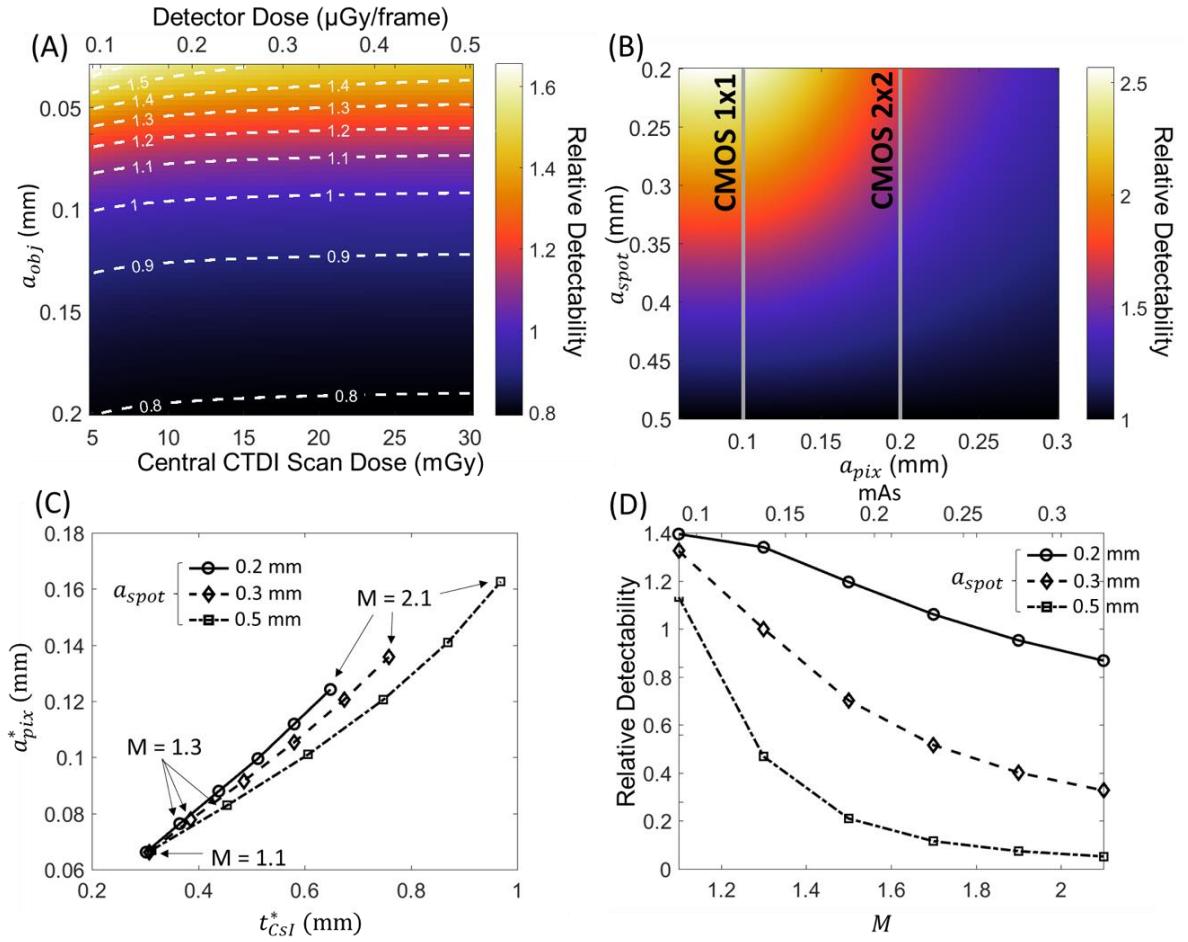


Figure 4.5. (A) Ratio of detectability achieved with the C400 detector to that of C700 as a function of imaging task and imaging dose for the nominal extremity CBCT geometry. Top horizontal axis represents detector entrance dose (after attenuation by a simulated knee), and the bottom horizontal axis gives the corresponding measured central CTDI dose in a CBCT scan. (B) Detectability for the trabecular imaging task ($a_{obj}=0.06 \text{ mm}$) as a function of focal spot size and pixel size, normalized by the detectability achieved for each a_{obj} using a 0.5 mm focal spot. C700 detector and nominal extremity CBCT geometry are assumed. (C) Pixel size and scintillator thickness yielding optimal detectability for the trabecular imaging task ($a_{obj}=0.06 \text{ mm}$) as a function of magnification and focal spot size. (D) Relative detectability values (normalized to detectability at the nominal CBCT geometry and $a_{spot} = 0.3 \text{ mm}$) at the optimal detector configurations found in (C).

The current FPD-based extremity CBCT scanner implements a stationary anode x-ray source with a 0.5 FS (IEC336). Clinical protocols use 2x2 detector binning with 0.388 mm pixels to maintain a ~30 sec scan time. In this configuration, there is little benefit from using a smaller focal spot, since detector blur associated with relatively large pixels then dominate the system MTF [151]. A CMOS detector provides finer pixels (with similar or faster scan time), motivating reconsideration of the optimal x-ray focal spot size. Fig. 4.5B shows the detectability for a 0.06 mm feature size as a function of pixel size and focal spot size. To enable comparison with the current system, CsI:Tl thickness of 0.7 mm was assumed (as in the a-Si:H FPD sensor) and d'^2 at each a_{pix} was normalized by the value at $a_{spot}=0.5$ mm. For $a_{pix} = 0.1$ mm, the system MTF is dominated by source blur at $a_{spot}=0.5$ mm, and a 2x – 2.5x improvement in d'^2 can be achieved by adopting an x-ray source with a focal spot of 0.2 – 0.3 mm. The benefits of smaller a_{spot} are less pronounced for pixel size >0.25 mm used in the current extremity CBCT (typically 0.139 mm pixel size).

The selection of the x-ray source for benchtop experimentation and implementation on the CMOS-based prototype balanced the need for small focal spot against design requirements such as x-ray power (≥ 1 kW) and the size of the x-ray unit able to fit inside the CBCT gantry. Based on these considerations, the compact (~40x20x30 cm³) rotating anode IMD RTM 37 tube with 0.3 FS (IEC336) was chosen after survey of commercially available systems.

Fig. 4.5C expands the investigation of CMOS detector configurations to hypothetical system geometries that differ from the current extremity CBCT system in terms of system magnification. Pixel size and scintillator thicknesses yielding optimal d'^2 for a 0.06 mm feature size are shown for a range of focal spot sizes (each a_{spot} is one line)

and magnifications (each M is a marked as a data point). Since the x-ray exposure was kept constant at 0.126 mAs/projection (as in Fig. 4.3), the fluence on the detector decreases with M . At $M=1.1$, the impact of source blur is minimized and the detector fluence is high, and thus the optimization is primarily driven by detector resolution. This results in the same optimal configuration with t_{CSI} of ~0.3 mm and a_{pix} of ~0.65 mm for all focal spot sizes. As the magnification increases, the detector input dose decreases, but the resolution requirements for the task diminish because the feature is magnified. The net result is that detector MTF becomes less of a factor at higher magnifications and the optimization shifts towards thicker CsI:Tl and larger pixel sizes to counteract the loss of input quanta. This effect is most noticeable for the focal spot of 0.5 mm (dash-dot line), where source blur dominates at high magnifications and d'^2 is primarily driven by reduction in quantum noise with thicker scintillators and larger pixels. For a_{spot} of 0.2 mm (solid line), on the other hand, system MTF is not as strongly affected by focal spot blur. The optimal detector configuration appears to be still partly driven by system resolution, favoring thinner CsI:Tl than for larger focal spots. Configurations with a_{spot} of 0.3 mm (dashed line consistent with the source used in experimental studies) represent a somewhat intermediate case. For magnifications of up to 1.3, the optimal a_{pix} and t_{CSI} are similar to those for focal spot of 0.2 mm. At those magnifications, the system MTF is dominated by detector blur for both x-ray focal spots. (Note that the optimal configurations at $M=1.3$ agree well with the optimum in Fig. 4.4D, where source blur was not included in the simulation). As M increases above 1.3, source blur becomes more prominent and the optimization switches toward using thicker scintillator and larger pixels to improve the noise component of d'^2 . In all cases, the diminishing need for high spatial resolution due to magnification of the

feature of interest allows for substantial increase in optimal a_{pix} and t_{CSI} at high values of M .

Figure 4.5D shows the detectability as a function of M and t_{CSI} for the optimal detector configurations identified in Fig. 4C. The detectabilities are normalized by the optimal value at the nominal CBCT configuration and a_{spot} of 0.3 mm. Consistent with Fig. 4.6B, d'^2 generally improves with decreasing focal spot size. However, the detectability at optimal detector configuration decreases with increasing M . Despite the diminishing resolution requirements at high magnifications, the increase in optimal a_{pix} and t_{CSI} required to compensate for the loss of input quanta associated with long SDDs leads to an unfavorable tradeoff between system MTF and NNPS. The resulting decrease in d'^2 is especially pronounced at a_{spot} of 0.5 mm, where system resolution is increasingly dominated by source blur at high values of M . For fixed source output and patient dose, configurations with smaller magnification and thus improved detector input fluence are preferred for the high-resolution tasks considered in this work.

IV.C.2 Benchtop Experimental Studies

IV.C.2.1 Measured MTF and DQE

Figure 4.6A shows MTF measurements for the C400 and C700 configurations. The frequency at 50% modulation (f_{50}) is $\sim 1.4x$ higher for the detector with thinner scintillator. Solid lines indicate empirical fits to detector MTF used in the cascaded systems model (Eqs. 4.7, 4.8 and 4.15). The parameterization of T_3 in Eqs. 4.8 and 4.15 appears to adequately capture the thickness-dependent component of scintillator blur, as indicated by good quality of the fit for both configurations. Superior spatial resolution of the C400 is

confirmed in Fig. 4.6B, which show images of a Gammex 91437 (Gammex, WI) radial resolution gauge placed on the surface of each detector (mean of 50 frames acquired at 6.3 mA and 20 ms). Compared to C700 (left panel), C400 (right panel) maintains modulation of the line pattern (i.e., is free from signal aliasing) up to higher spatial frequencies.

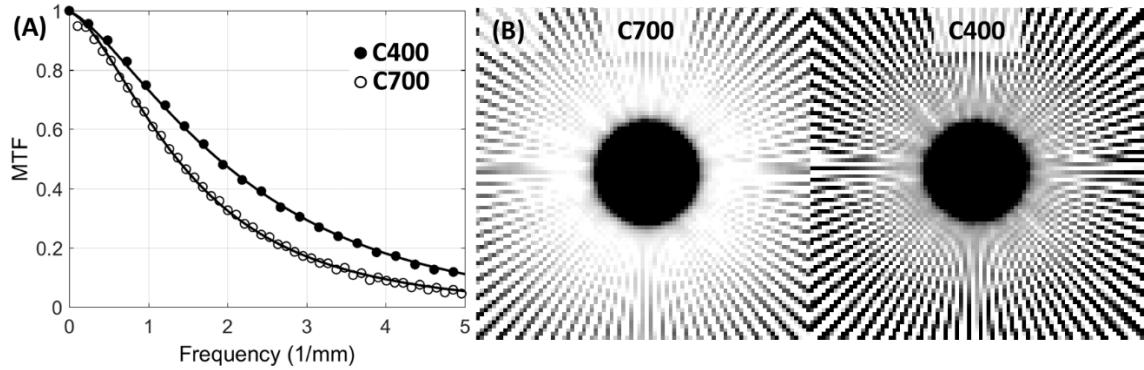


Figure 4.6 (A) Experimental measurements of detector MTF for CMOS sensors with 0.7 mm CsI:Tl (open circles) and 0.4 mm CsI:Tl (closed circles). Lines represent MTFs computed using the cascaded systems model. (B) Contact images of the GAMMEX 91437 resolution gauge obtained with 0.7 mm CsI:Tl (left) and 0.4 mm CsI:Tl (right).

Fig. 4.7 shows measured DQE for the two detectors (points) along with simulated DQE obtained from the cascaded systems model (lines). DQE is presented for two values of detector entrance dose in the quantum-limited range and at a low entrance dose of ~ 0.5 mGy/frame, where the impact of electronic noise becomes visible at high frequencies. (All doses were measured behind additional filtration emulating attenuation by 15 cm water; see Sec. IV.B.2.1). The C700 outperforms C400 for spatial frequencies up to ~ 3 mm $^{-1}$, where the DQE of both detectors begins to converge, indicating a regime where the improved MTF of the C400 detector overcomes the elevated noise (reduced QDE) associated with the thinner scintillator. Near the Nyquist frequency, C400 achieves better DQE than C700 for all considered dose levels (0.127 for C400 vs 0.069 for C700 at 0.470 μ Gy/frame; 0.117 vs 0.057 at 0.293 μ Gy/frame; 0.049 vs 0.017 at 0.045 μ Gy/frame). The

noise-equivalent dose, the dose at which quantum noise reaches the same magnitude as electronics noise [169], is $\sim 0.03 \mu\text{Gy}/\text{frame}$.

We observed fair agreement between simulated and measured DQE. Most significant discrepancies were observed at low frequencies for the C700 detector, likely due to inaccuracies in estimation of packing fraction f_{CSI} and K-fluorescence reabsorption blur T_K . Since direct measurements of those parameters for the two sensors were not available, their implementation in the model were based on approximations, namely: the same value of f_{CSI} was used for C400 and C700, chosen empirically to yield fair agreement with measured NNPS(0) for both detectors; however, a better overall fit in DQE could be potentially achieved by adjusting the packing fraction individually for each detector. With respect to T_K , the model uses an analytical formula adapted from the work of Que *et al.* [159]. The impact of T_K in the total K-fluorescence MTF T_{Ktot} (Eq. 4.9) becomes negligible at $\sim 3 \text{ mm}^{-1}$, reflecting the relatively long range of K-fluorescence photons. This implies that potential inaccuracies in the analytical estimate of T_K (originally derived for amorphous selenium) are most likely to affect low- to mid- frequency DQE. Further refinement of the estimates of f_{CSI} and T_K requires dedicated experimental techniques and more sophisticated scintillator models (e.g., MC simulations [170], [171]) that are beyond the scope of this study. In the present work, the impact of these approximations is limited because we are primarily concerned with tasks emphasizing mid- to high-frequencies ($> 2 \text{ mm}^{-1}$), where the simulation agrees well with measurements.

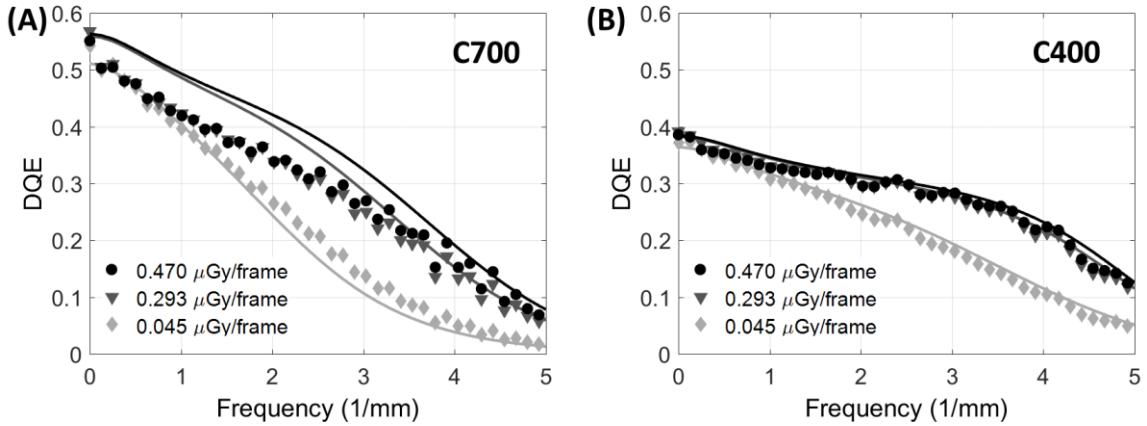


Figure 4.7. Measured (points) and simulated (lines) DQE for a range of dose levels for the CMOS detector with (A) 0.7 mm thick scintillator and (B) 0.4 mm scintillator.

IV.C.2.2 Cadaver Imaging

Reconstructions of a cadaveric knee obtained using the C400 and C700 configurations are shown in Fig. 4.8. The central CTDI dose was 15 mGy in both scans.

Fig. 4.8A shows details of two trabecular regions in high-resolution bone reconstructions of the knee. The images in the top row of were generated using the high-resolution reconstruction protocol, but after digital 2x2 pixel binning of C700 projections to mimic the ~0.2 mm pixels of a-Si:H FPDs operated at full resolution. Comparison with images obtained using C700 and the native pixel size of 0.1 mm in the center row reveals the improvement in visualization of trabecular detail enabled by the fine pixel size of CMOS detectors.

Bone protocol reconstruction of projections acquired using the custom CMOS with 0.4 mm CsI:Tl are presented in the bottom row of Figs. 4.8A (trabecular ROI) and 4.8B (complete axial slice). Adoption of a thin scintillator further enhances the visibility of the trabecular pattern compared to C700. As expected, based on the simulation studies, where the increase in d^2 with C400 was ~10-20%, the improvement in visualization of high

frequency detail is perceptible, but modest, and accompanied by slightly increased noise. This increased noise might affect visualization of low contrast soft-tissue structures. However, since very high spatial resolution is not essential in soft tissue evaluation, the impact of elevated noise can be mitigated at least in part by binning the projections, using larger voxels and adjusting the reconstruction filter. As shown in Fig. 4.8C, soft-tissue reconstructions obtained using C400 and the soft-tissue protocol with 4x4 digital binning (Sec. IV.B.2.2) achieve adequate visualization of soft-tissues.

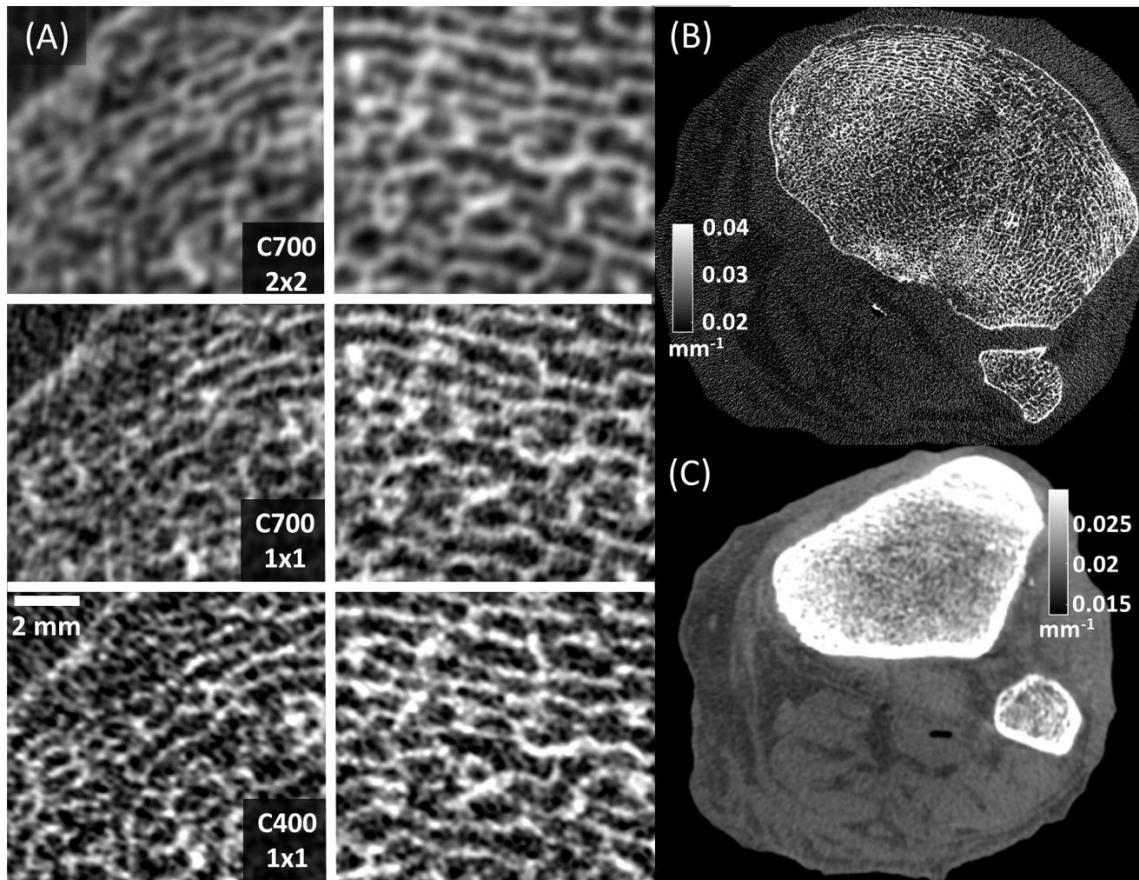


Figure 4.8. (A) Magnified views of two trabecular regions in the subchondral bone of a cadaver knee imaged using CMOS detectors with different pixel sizes and scintillator thicknesses. High-resolution bone reconstruction was used. (A, top) Reconstructions of 2x2 binned C700 projections, mimicking the pixel size of current a-Si:H FPDs. (A, middle) Reconstructions of C700 projections in 1x1 binning, showing the benefits of reduced pixel size provided by CMOS. (A, bottom) Images acquired with C400 in 1x1 binning, illustrating the visualization of trabecular detail using a thin scintillator. (B) A complete axial slice of C400 reconstruction obtained using high-resolution protocol (C) A C400 reconstruction obtained using a soft-tissue protocol with 4x4 pixel binning.

IV.C.3 Bone Microstructure Imaging on a Prototype

Fig. 4.9A compares the detector MTF for a-Si:H FPD and CMOS CBCT systems. In addition to extended Nyquist frequency, the frequency at 20% modulation for the CMOS detector is ~ 1 lp/mm larger than that of the a-Si:H FPD. Fig. 4.9B shows the highest possible resolution reconstruction (ramp filter, Nyquist cutoff) of the 127 μm tungsten wire, with the FWHM for the CMOS system ~ 0.1 mm smaller than that of the a-Si:H FPD system.

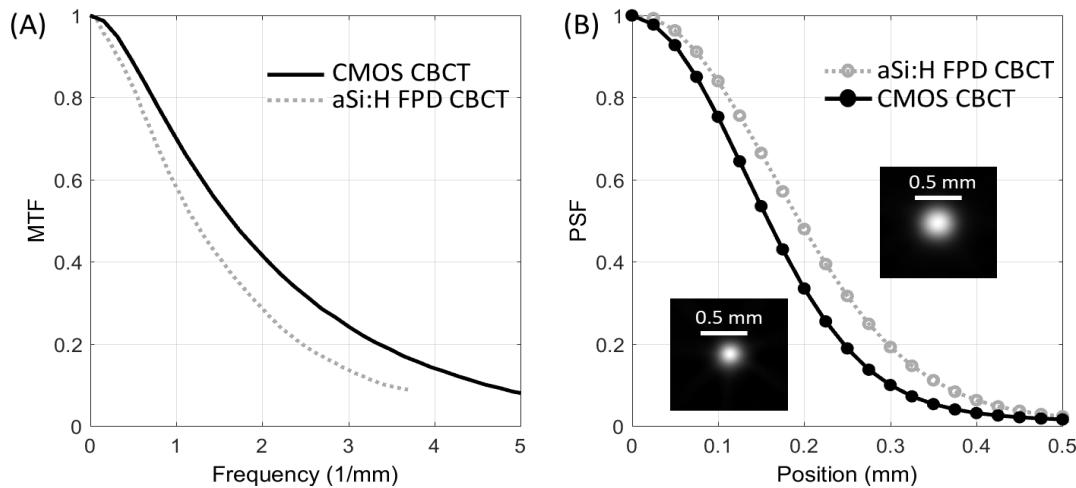


Figure 4.9 (A) Detector MTF measurements. (B) Line profile through FBP reconstructions of a tungsten wire.

CMOS and a-Si:H FPD systems are compared in a realistic clinical scenario in Fig. 4.10A using high-resolution reconstructions of the cadaveric ankle obtained at approx. equal dose. A modest but perceptible improvement in delineation of bony detail with a CMOS detector is apparent, especially in the improved modulation for fine trabecular features in the enlarged view of the calcaneus. Fig. 4.10B shows high-contrast reconstructions of the ankle. Despite the smaller pixel size and thinner scintillator, the CMOS image maintains sufficient contrast-to-noise performance to achieve comparable visualization of soft-tissue structures to the a-Si:H FPD system.

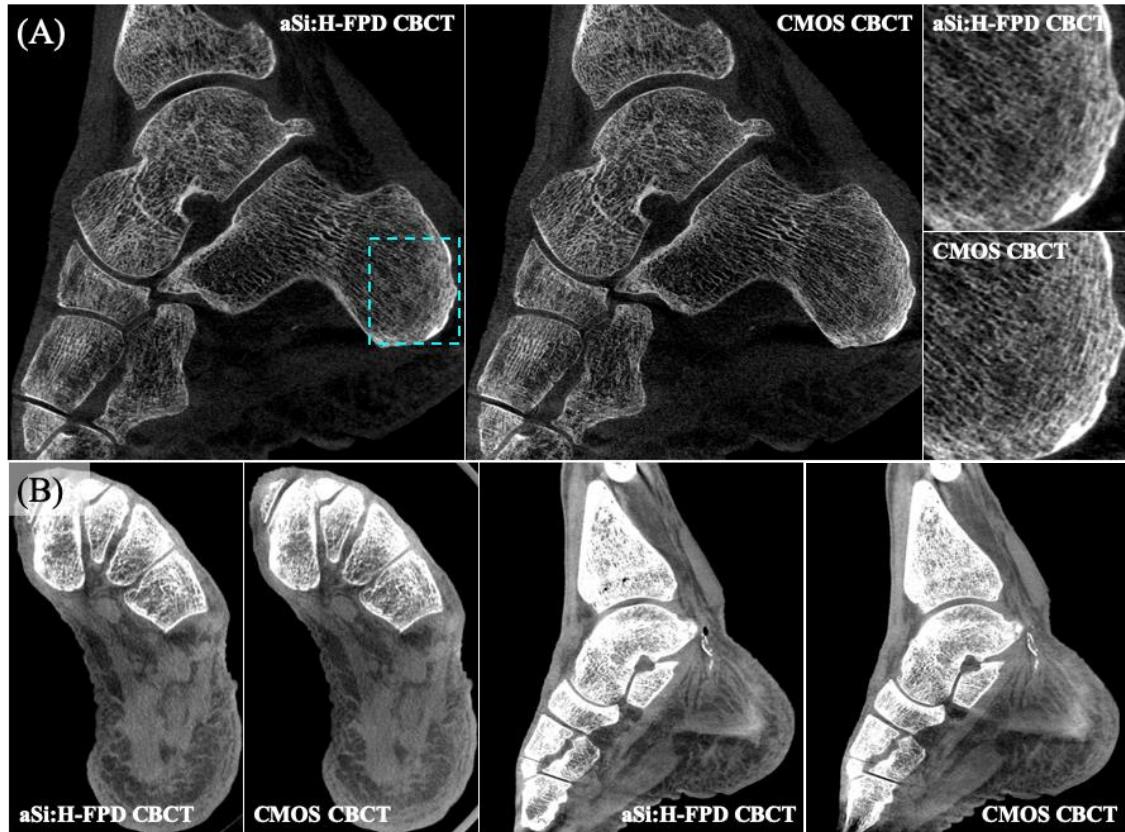


Figure 4.10 Reconstructions of a cadaveric ankle on a-Si:H FPD and CMOS CBCT systems in (A) high-resolution protocol and (B) high-contrast protocol.

As shown in Figs. 4.11 and 4.12, the intrinsic resolution advantage of CMOS CBCT leads to improved performance in quantitative evaluation of trabecular microstructure. Visual analysis of sample ROIs reveals that CMOS-CBCT yields a segmentation that better preserves the topology of trabecular microstructure compared to the a-Si:H FPD. This is supported by the 10% improvement in the mean Dice value of the ROIs in Fig. 4.11. This improved segmentation leads to better correlation of the trabecular metrics with gold standard μ CT (Fig. 4.12B-D). Both systems achieve >0.9 correlation coefficient for TbSp and BvTv, where the CMOS detector provides $\sim 10\%$ improvement over a-Si:H FPD. The benefit of CMOS is more pronounced for TbTh, where the correlation coefficient is

increased from 0.49 to 0.74. Trabecular “ridges” are typically thinner than the spaces that separate them and thus TbTh is most sensitive to improved spatial resolution.

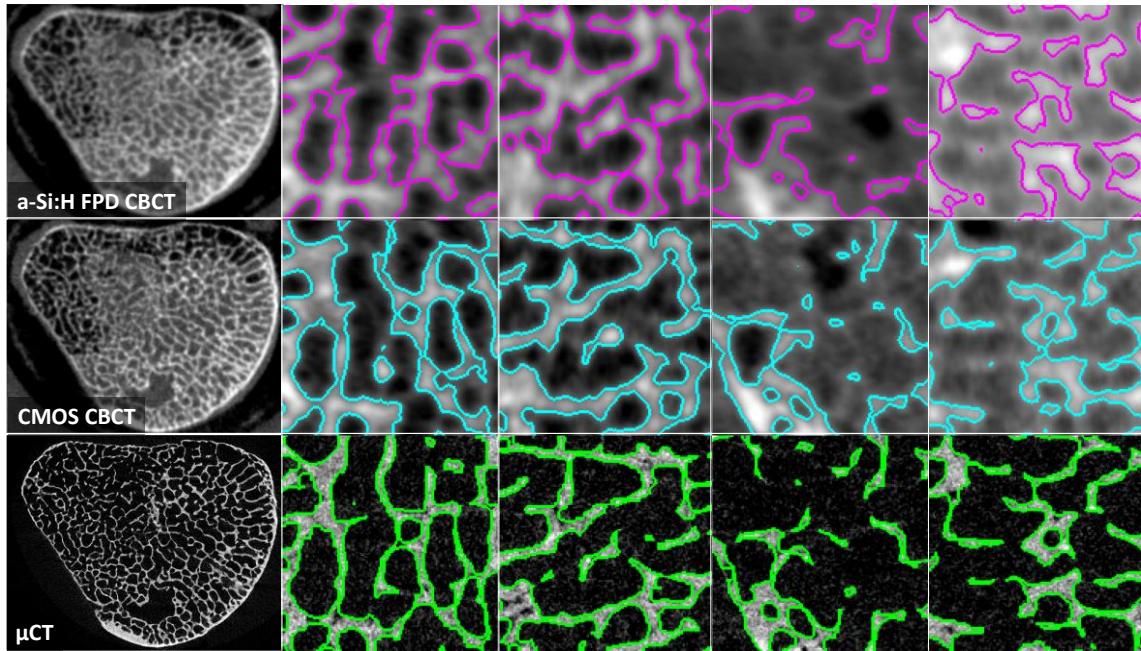


Figure 4.11 Reconstructions of the ulna (1st column) and sample trabecular ROIs with overlaid segmentations.

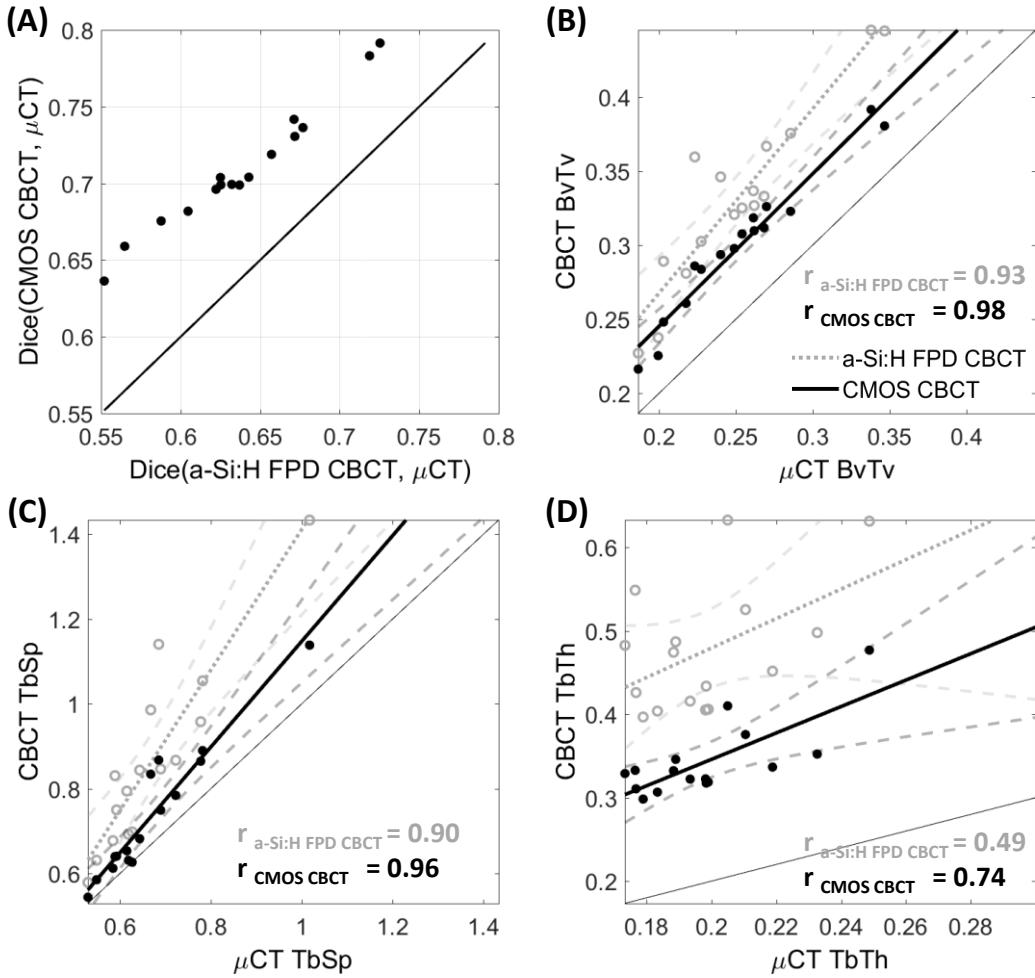


Figure 4.12 (A) Distribution of Dice coefficient between the CBCTs and μ CT ROI segmentations and (B)(C)(D) correlation of trabecular metrics obtained with CBCT and μ CT. Dashed opaque lines denote confidence interval corresponding to $\alpha=0.05$. Grey solid line has unit slope and zero intercept (equivalence line representing absolute agreement).

IV.D Discussion and Conclusion

Analysis of task-based detectability points to the benefits of CMOS detectors in high-resolution applications of extremity CBCT. A CMOS detector was found to achieve $\sim 4x$ better detectability (d^2) for 0.06 mm features compared to an otherwise identical detector ($a_{pix} = 0.1$ mm) that exhibited electronic noise comparable to an a-Si:H FPD. This is because the low electronic noise of CMOS detectors allows the system to take advantage of smaller apertures and improved sampling associated with fine detector pixels.

Pixel size <0.1 mm was preferred for all tasks considered ($a_{obj} = 0.03$ to 0.2 mm). In addition to higher spatial resolution, CMOS detectors offer 3-4x faster frame rate than a-Si:H FPDs. Scan times of 20-30 sec are anticipated for CMOS-based extremity CBCT operated at full detector resolution, reducing the risk of patient motion during acquisition.

Since the optimal pixel size for high-frequency tasks is smaller for CMOS detectors compared to a-Si:H FPDs, such applications also benefit from using a scintillator that is thinner than the 0.6-0.7 mm CsI:Tl that is commonly employed. In simulation studies, a ~0.4 mm scintillator was shown to provide 10-20% better detectability for high-frequency tasks compared to a 0.7 mm scintillator for conditions typical of extremity CBCT. DQE measurements of CMOS detectors with 0.4 mm CsI:Tl (C400) and 0.7 mm CsI:Tl (C700) showed that while C700 showed better DQE at low frequencies, C400 provided improved DQE at frequencies >3 lp/mm. CBCT imaging of a cadaveric knee confirmed a modest but perceptible improvement in delineation of trabecular detail with the custom CMOS detector with 0.4 mm CsI:Tl.

The decrease in DQE and detectability for low-frequency tasks may hamper soft-tissue imaging in a system based on a detector with thin scintillator. However, software binning of projection data obtained with C400 at imaging dose of 15 mGy yielded adequate delineation of soft tissue (muscles, tendons, and fat). Overall, the prototype CMOS detector with 0.4 mm scintillator provided improved performance in imaging of trabecular bone compared to a conventional screen without major detriment to soft-tissue visualization.

The results discussed above were obtained assuming the geometry of current generation extremity CBCT. This compact configuration was developed specifically to enable weight bearing imaging of a single extremity. The primary constraint resulting from

this consideration is that the detector needs to rotate between the legs of the subject, limiting system magnification. It is interesting to consider potential benefits of other system configurations, free of this restriction. To this end, additional study was performed to analyze the impact of altering the magnification by increasing the SDD and keeping the SAD constant. The current value of SAD (~400 mm) cannot be significantly shortened in a realistic system configuration considering spatial constraints (size of the extremity and the collimator box) and the available x-ray source cone angles. When x-ray source output and patient dose are kept constant and SDD (and thus magnification) is increased, the detectability is generally maximized at thicker CsI:Tl and larger pixel sizes. This is partly because larger magnifications exhibit somewhat reduced demands on detector resolution, allowing the optimization to be driven by noise reduction. However, the detectability at optimal detector configurations decreases with M , indicating that the overall tradeoff between system MTF and NPS still favors configurations with high enough input fluence to allow using thin scintillators and small pixels to minimize detector blur.

The tradeoffs associated with system magnification would be different in a study where the source output was adjusted to yield constant detector input fluence across all detector configurations. Under such conditions, systems with higher magnification would likely be favored since the optimization would not contend with the loss in input quanta. However, the constraint of fixed source output is a reasonable starting point with respect to fixing the dose to the patient. This is not only due to patient safety considerations, but also because of the limited power available in x-ray tubes with focal spot sizes small enough (≤ 0.2 mm) not to dominate system resolution at high magnifications.

There are recognized limitations of the CMOS detector model. A slight discrepancy between simulated and measured DQE was attributed primarily to inaccuracy in the model of K-fluorescence blur [159]. This discrepancy is unlikely to influence the general conclusions, in particular with respect to medium and high-frequency tasks where there was good agreement between simulation and measurement. Furthermore, the thickness-dependent model of scintillator MTF was parameterized based on measurements on C400 and C700. This parameterization may include effects unrelated to scintillator blur, such as differences in optical coupling between the two sensors. The thickness of the scintillators is also not exactly known and subject to manufacturing tolerances. Nevertheless, we believe that the model properly captures general trends associated with reduced scintillator thickness, as confirmed in experimental studies with the two detectors.

The task model in the current work represents discrimination of two stimuli by an NPW observer. It is recognized that this choice may not exactly represent the performance of a computer algorithm in quantitative analysis of bone quality. Rather, it is intended as a reasonable approximation of the fundamental capability of the system to resolve high-resolution details, under an assumption that any improvement in this capability will likely benefit the analysis of trabecular microstructure. Conventionally, such analysis involves thresholding and binarization to extract the bone voxels and measure structural indices [172], which may be more closely related to estimation tasks than detection / discrimination tasks [173]. It is anticipated that improved system performance in the discrimination task will translate also to improved performance in the threshold-based quantitative measurements. A potentially more challenging extension of this work might involve predicting the performance of a texture classification algorithm applied to trabecular

regions [174]–[177]. The investigation of textural biomarkers of bone quality is still in early stages and there is no consensus yet as to which of the textural features are best suited for diagnostic applications. For some features, such as those based on the grey-level co-occurrence matrix [178], improved conspicuity of fine trabecular detail may improve the sensitivity of classification. However, it is possible that other textural features may be less sensitive to this aspect of system design. Overall, new task functions and observer models specific to the features of interest will need to be developed to enable task-based analysis of textural biomarkers. Such development is beyond the scope of this work.

Finally, the optimization studies were performed in projection domain to focus the analysis on the fundamental aspects of imaging performance. Achieving best possible baseline imaging capability, as reflected in the 2D metrics evaluated in this work, is an essential first step in the development of the new capability for high-resolution analysis of bone microstructure. Numerous practical design considerations, more amendable to analysis using models of 3D imaging performance, were not investigated. Such considerations include x-ray scatter, shift-variant x-ray source blur, blur due to gantry motion in continuous pulsed acquisition and reconstruction algorithm.

The current generation extremity CBCT relies on algorithmic scatter correction methods [93], [179], [180] to achieve adequate soft tissue contrast despite the relatively high SPR of ~0.5-1 [180]. Previous work indicated a modest benefit in CNR from an anti-scatter grid in extremity CBCT geometry [180]. Ongoing work on a prototype CMOS-based system will determine whether the improvement in CNR due to a grid translates to improvement in quantitative metrics of bone microstructure and whether such benefit outweighs potential practical challenges in calibration and grid line artifact correction.

A fairly general, simplified model of x-ray focal spot was used in this work. This model was not intended to simulate any particular x-ray tube, but to provide an adjustable model to investigate “first-order” effects of source blur. This approach assumes that the blur is isotropic and shift-invariant. In practice, neither of those assumptions is perfectly satisfied. Pinhole measurements [181] of the focal spot on the RTM37 tube used in this work revealed a complex, non-isotropic shape with full width at half maximum (FWHM) of ~0.2 mm in the cathode-anode direction, close to the stated nominal value. In systems that are not dominated by focal spot blur, such as the proposed CMOS CBCT using ~0.1 mm detector pixels, the shape of the focal spot is unlikely to have substantial impact on system resolution. A potentially more significant effect is the variation in the apparent focal spot size seen in different areas of the detector. For the extremity CBCT geometry, the apparent focal spot width for a 0.3 mm x-ray source is ~0.16 mm at the edge of the 300 mm detector on the cathode side of the source, ~0.1 mm at the center and ~0.025 mm on the anode side. Inside a projection view of ~100 mm diameter object (e.g. human tibia), the apparent focal spot varies from ~0.12 mm on the cathode side to ~0.06 mm on the anode side. In 3D imaging, magnification changes throughout the FOV, resulting in even more complex combination of shift variant source and detector blurs, additionally affected by angular sampling and the imaging orbit (short or full scan). A study evaluating local 3D imaging performance in the presence of shift variant blurs was recently reported [181]. An extremities CBCT system with 0.3 mm focal spot size, 0.1 mm detector pixels, and short scan acquisition was simulated using a geometric forward projection model. The FDK reconstructions for the trabecular bone phantom was assessed throughout the FOV using RMSE. There was a ~7% change in RMSE between the best and worst values.

Interestingly, the short scan orbit provided a potential advantage over a full scan in that a region of the FOV was sampled primarily by the cathode side of the source, thereby improving local resolution. Overall, however, the fidelity of the reconstruction was only slightly affected by the shift variant source blur, likely because scintillator blur dominates in this system geometry.

Another practical consideration not investigated in this study is the effect of continuous gantry rotation. The Xineos3030 detector operates at 30 fps, resulting in total scan time ~17 seconds for a 210° trajectory. For 5 ms x-ray pulse length (attainable by the RTM37 tube used here), the distance travelled by an object at the radius of 50 mm from the axis of rotation is 0.055 mm/pulse. After considering system magnification, this distance is less than the pixel size. The resulting motion blur is thus minimal, but detailed investigation of this effect is left to future experimental studies.

Based on findings of the benchtop study, the custom CMOS detector with 0.4 mm scintillator and the compact rotating anode x-ray source with 0.3 mm focal spot have been translated to the gantry of an extremity CBCT system (OnSight 3D) to develop a prototype high-resolution scanner for clinical studies. Parallel work involves development of motion correction [182], [183] and advanced reconstruction algorithms with model-based deburring to further enhance system resolution discussed above [184]. Evaluation against gold-standard micro-CT in metrics of bone microstructure using bone core samples is ongoing. A clinical pilot study of test-retest reproducibility in bone morphometry in OA patients is in preparation. It is anticipated that the improved performance in high-frequency imaging tasks provided by the CMOS detector will enable robust measurements of bone

microarchitecture *in vivo*, benefiting image-based assessment of osteoporosis, osteoarthritis, and monitoring of fracture healing.

The prototype CMOS CBCT system achieved ~40% improved spatial resolution, better visualization of bony detail at equal scan dose, and improved correlation with μ CT in metrics of bone microarchitecture compared to the system with conventional a-Si:H FPD. Moreover, the scan time in high-resolution mode is ~3x shorter using CMOS than with the a-Si:H FPD operated in 1x1 binning, greatly mitigating the risk of patient motion. Overall, the CMOS-based scanner delivers promising performance in quantitative trabecular imaging.

Chapter V

A Multiresolution Model for High-Resolution MBIR in CBCT

This chapter describes how MBIR can be performed for high-resolution data, given the limitation that an object always needs to be reconstructed in its entirety for MBIR. The multiresolution solution presented here reparametrizes the image and projection data into coarse and fine ROIs, modeling the complete support of the reconstructed object, resulting in consistent forward projections while achieving native-spatial resolution in the fine ROI.

V.A Introduction

Images of bone microstructures in the previous chapter are reconstructed from analytical reconstruction. Application of MBIR to high-resolution CBCT imaging is challenged by the computational burden of iterative algorithms. Whereas analytical methods typically require only one backprojection operation, each iteration of MBIR generally requires at least one forward and one backprojection. Furthermore, since MBIR relies on a data fidelity term that matches image estimates to the measured data, a complete transaxial region of support for the image volume must be reconstructed to capture all contributions from the object for each ray path. Thus, ROI reconstructions that are often straightforward to implement using analytical methods cannot be realized in standard iterative algorithms. The requirement to reconstruct the entire FOV, typically as large as 20x20x20 cm³, using very fine voxels ($\leq 75 \mu\text{m}$), can make the application of MBIR to

high-resolution imaging prohibitively slow because the speed of the projection and backprojection operators is generally proportional to the number of voxels.

Projection operators can be accelerated by optimized implementation on parallel hardware using GPUs [185]–[188] or distributed computing [189]. However, for large-FOV high-resolution CBCT data, both approaches suffer from latency of data transfer – either between host computer and GPU or between computational nodes in a distributed network. Furthermore, while such methods improve the baseline speed of the projection operator, the performance gains remain dependent on the size of the FOV and resolution. The benefits of parallelization diminish when the number of required processes exceeds the number of available computational nodes due to very fine discretization of the FOV.

One approach to overcome the decreased performance of MBIR in high-resolution applications relies on the observation that while iterative reconstruction of an isolated high-resolution ROI is not possible, the model of the volume can be altered so that a fine parameterization is only used in the ROI. Some versions of such multiresolution schemes forego the traditional square voxel basis functions and employ a sparse representation of the volume using heterogeneous voxels, wavelets [190], [191] or meshes [192], [193]. Here we will focus on traditional voxel bases, where multiresolution reconstruction can be realized by using voxels of different size to parameterize different subvolumes (i.e. coarse grid regions and fine grid regions) in the image. These bases are straightforward and map well to physical memory in computing hardware.

Early examples of multiresolution methods include those that enable application of the iterative algorithm to a limited internal high-resolution ROI and use an analytical method for the surrounding volume. The full FOV is first reconstructed with an analytical

method. The ROI subvolume is set to zero and the surrounding volume is reprojected. The resulting sinogram is subsequently subtracted from the measured data to isolate the line integrals through the ROI. The line integrals of the ROI are then iteratively reconstructed [194], [195]. An additional step consisting of smoothing of the reprojected coarse grid line integrals was found to yield reduced noise in the ROI reconstruction [196]. However, artifacts attributable to edge degradation and beam-hardening artifacts in the initial FBP were found in the iteratively reconstructed ROI obtained using this general approach [197].

A multiresolution approach where subvolumes with different levels of discretization are reconstructed using an iterative algorithm was proposed in [198]. The fine and coarse grids are jointly reconstructed using a maximum-likelihood (ML) algorithm applied in an alternating scheme, updating one grid while holding the other fixed. In contrast to this alternating optimization, simultaneous optimization of both coarse and fine grids using general purpose optimization algorithms was developed in [199], [200]. Similar to our work, this approach relies on a natural factorization of the line integrals into the projections of the fine and coarse grids, yielding a joint objective function for the entire volume. This formulation was also applied to reduce metal artifacts by applying high fidelity forward models (finer parameterization of metallic edges and polychromatic beam modeling) only in ROIs containing metallic objects [201], [202].

This work investigates the application of multiresolution reconstruction in the context of extremity CBCT, in particular for visualization and quantitative assessment of bone microarchitecture. In this application, the fine trabecular detail ($\leq 75 \mu m$) needs to be reconstructed at very high-resolution only over a relatively small ROI, covering the subchondral bone in the knee joint. The surrounding anatomy (cartilage, ligaments, muscle,

fat, etc.) can be reconstructed at lower resolution over the larger FOV. We introduce a new multiresolution MBIR algorithm based on the PWLS framework [59] with a SPS optimizer [61], [203]. We propose a penalty function specifically designed to provide regularization across the boundaries between fine and coarse voxel grids. We extend the multiresolution framework to include the possibility of regional binning of the projection images, resulting in variable pixel size in addition to the variable voxel size. We focus our evaluation on the tradeoffs between artifacts in the fine grid ROI and voxel size and regularization applied in the surrounding coarse grid region.

V.B Methods

V.B.1 Multiresolution PWLS

The conventional forward model for a transmission tomography system, assuming a voxelized object representation, is given by:

$$\bar{y}_i = g_i \exp(-[\mathbf{A}\mu]_i) \quad (5.1)$$

where \bar{y}_i denotes the mean of the i^{th} measurement, g_i represents the system gain (including bare beam fluence and detector sensitivity for measurement i), μ is a vector of voxel attenuation values, and \mathbf{A} is the forward projection operator. To formulate an objective function for model-based reconstruction, a noise model needs to be chosen. For quantum-limited projection data with negligible contribution of electronic noise, Poisson noise is often assumed. For systems with non-negligible contributions of electronic noise (such as the FPDs used in extremity CBCT) and when the projection data is post-processed with artifact correction algorithms prior to reconstruction (e.g. beam hardening and scatter

correction), a more general Gaussian noise model is often more appropriate, leading to a weighted-least-squares objective. We adopt the Gaussian model and PWLS in this work as better suited for the statistics of FPD CBCT projection data. However, the multiresolution methodology presented here could be incorporated into any forward model and easily generalized to MBIR algorithms based on the Poisson noise model. PWLS reconstruction [59] employs the following objective function:

$$\begin{aligned}\hat{\mu}_{PWLS} &= \underset{\mu>0}{\operatorname{argmin}} \frac{1}{2} \|A\mu - l\|_W^2 + \beta R(\mu) \\ &= \underset{\mu>0}{\operatorname{argmin}} \frac{1}{2} \|A\mu - l\|_W^2 + \beta \sum_{k \in K} \psi([\mathbf{C}\mu]_k)\end{aligned}\quad (5.2)$$

where l denotes the vector of line integrals $l_i = -\log y_i / g_i$. The traditional choice for \mathbf{W} is a diagonal weighting by the inverse variance, approximated using the measurements: $w_i = y_i$. The second term (R) is a roughness penalty specified by the finite differencing operator \mathbf{C} [204] and the potential function ψ . Here, \mathbf{C} is a matrix and applies the 6 nearest-neighbor pairwise voxel difference (excluding diagonal neighbors). K is the total number of pairwise differences in the volume ($K \sim 6 \times$ number of voxels) and ψ is the quadratic penalty.

The above system model and reconstruction objective is general and permits many options for the parameterization of the image volume. We will use this framework and define forward projection and differencing operators for a multiresolution representation of the object. Specifically, we choose to represent the total volume μ as a union of a set of fine grid voxels μ_F and a set of coarse grid voxels μ_C :

$$\mu = \begin{bmatrix} \mu_C \\ \mu_F \end{bmatrix} \quad (5.3)$$

The ratio of the coarse grid voxel size to the fine grid voxel size defines a downsampling factor η_μ . The linearity of the forward projection operator permits definition of the following multiresolution projector \mathbf{A} , consisting of coarse and fine forward projectors \mathbf{A}_C and \mathbf{A}_F :

$$\mathbf{A}\boldsymbol{\mu} = [\mathbf{A}_C \quad \mathbf{A}_F] \begin{bmatrix} \boldsymbol{\mu}_C \\ \boldsymbol{\mu}_F \end{bmatrix} \quad (5.4)$$

Integrating (3) and (4) into (1), we may rewrite the forward model as:

$$\bar{\mathbf{y}} = g \circ \exp(-\mathbf{A}\boldsymbol{\mu}) = g \circ \exp(-\mathbf{A}_C\boldsymbol{\mu}_C) \circ \exp(-\mathbf{A}_F\boldsymbol{\mu}_F) \quad (5.5)$$

where operator \circ denotes the Hadamard (elementwise) matrix product. This forward model provides a convenient mathematical form for the development of iterative reconstruction algorithms that treat the reconstructed volume as a sum of non-overlapping regions, each with its own distinct forward model. In this work, we focus on forward models with different discretization of the image volume; however, this kind of decomposition has also been used in reconstruction methods that parameterize the object into known foreground components and an unknown background [205] for reconstruction in the presence of objects known to be in the FOV (eg, high-density surgical implants).

In addition to the large memory footprint of the reconstructed volume, projection datasets acquired with high-resolution FPD can also be prohibitively large (up to $\sim 3000 \times 3000$ pixels and ~ 1000 projections). Thus, we extend the multiresolution approach to include multiresolution binning of projection data. The binning ratio η_y is defined as the ratio of the coarse pixel size to fine pixel size. As illustrated in Fig. 5.1, we assume that the binning is applied (via appropriate pre-processing) in such a manner that the fine grid regions of the volume are projected entirely onto the fine, natively sampled (unbinned)

regions of the projection image. The forward projector now comprises \mathbf{A}_{FN} (fine grid voxels to native pixels), \mathbf{A}_{CN} (coarse grid voxels to native pixels) and \mathbf{A}_{CB} (coarse grid voxels to binned pixels):

$$\hat{l} = \begin{bmatrix} \hat{l}_B \\ \hat{l}_N \end{bmatrix} = \mathbf{A}\hat{\mu} = \begin{bmatrix} \mathbf{A}_{CB} & 0 \\ \mathbf{A}_{CN} & \mathbf{A}_{FN} \end{bmatrix} \begin{bmatrix} \hat{\mu}_C \\ \hat{\mu}_F \end{bmatrix} \quad (5.6)$$

where l_N and l_B denote the line integrals associated with native and binned pixels, respectively.

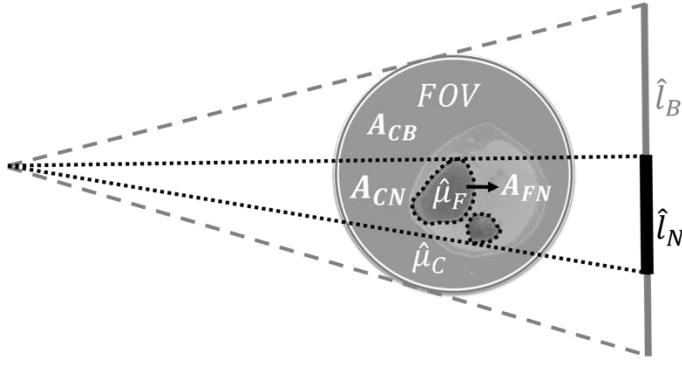


Figure 5.1 Schematic of multiresolution forward projection with detector binning. The estimated bone region parameterized using fine grid voxels ($\hat{\mu}_F$) is marked with a black dashed line. This region is projected onto native detector pixels (marked with \hat{l}_N) using projection operator \mathbf{A}_{FN} . The line integrals captured by the native projection pixels also include contributions from a subset of coarse voxels ($\hat{\mu}_C$), denoted by projection operator \mathbf{A}_{CN} . Line integrals for binned projection pixels (marked as \hat{l}_B) traverse only through coarse grid voxels, computed using projector operator \mathbf{A}_{CB} .

Using the above multiresolution definitions, the data fidelity term in Eq. 5.2 is nearly complete. It remains to define the weighting matrix:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_B & 0 \\ 0 & \mathbf{W}_N \end{bmatrix} \quad (5.7)$$

with the corresponding diagonal inverse variance weights, \mathbf{W}_B and \mathbf{W}_N , which represent the measurement weights of the binned and native pixels, respectively. Such weighting can accommodate, for example, reduced noise in the binned measurements.

Implementation of the regularization term in the multiresolution PWLS of Eq. 5.2 requires specification of how the differencing operator \mathbf{C} acts across the boundary between fine and coarse grid voxels. We propose the following general multiresolution penalty:

$$\tilde{\mu}_c = \begin{bmatrix} \mu_c \\ \mathbf{D}\mu_F \end{bmatrix} \quad \tilde{\mu}_F = \begin{bmatrix} \mathbf{U}\mu_c \\ \mu_F \end{bmatrix} \quad (5.8)$$

$$R(\mu) = \beta_C \sum_{k \in K_C} \psi([\mathbf{C}_C \tilde{\mu}_c]_k) + \beta_F \sum_{k \in K_F} \psi([\mathbf{C}_F \tilde{\mu}_F]_k) = \beta_C R_C + \beta_F R_F$$

where we have defined augmented coarse and fine image vectors, $\tilde{\mu}_c$ and $\tilde{\mu}_F$, respectively. These vectors include boundary voxel values from the adjacent ROI that have been resampled to the appropriate voxel grid. That is, we use \mathbf{D} , an interpolating operator that downsamples fine grid voxels neighboring the μ_F/μ_C boundary to coarse grid voxels, to allow application of a coarse pairwise voxel difference operator [204], \mathbf{C}_C , across the boundary of the coarse and fine ROIs (Fig. 5.2). Alternately, we use \mathbf{U} , an operator that upsamples the coarse grid voxels on the μ_F/μ_C interface, to apply a fine pairwise voxel difference operator, \mathbf{C}_F , across the ROI boundary in the opposite direction. In this work, we choose \mathbf{D} to be a binning operation that computes the mean over fine voxels, and \mathbf{U} to upsample from coarse to fine voxels using linear interpolation. We have two regularization parameters, β_F and β_C , that control the penalty strengths for the fine and coarse grid ROIs, respectively. The sets, K_C and K_F , are rows of $\tilde{\mathbf{C}}\hat{\mu}$ which enumerate all pairwise voxel differences contributing to the penalty terms for the coarse and fine ROIs, respectively.

Eq. 5.8 amounts to applying the conventional penalty to a region (fine or coarse grid) that was virtually expanded to include boundary voxels from the other region, which were appropriately resampled using interpolation. This procedure is general and can be

applied to arbitrary boundaries, except perhaps for highly degenerate cases where interpolation is not applicable (e.g. ROIs consisting of a single voxel).

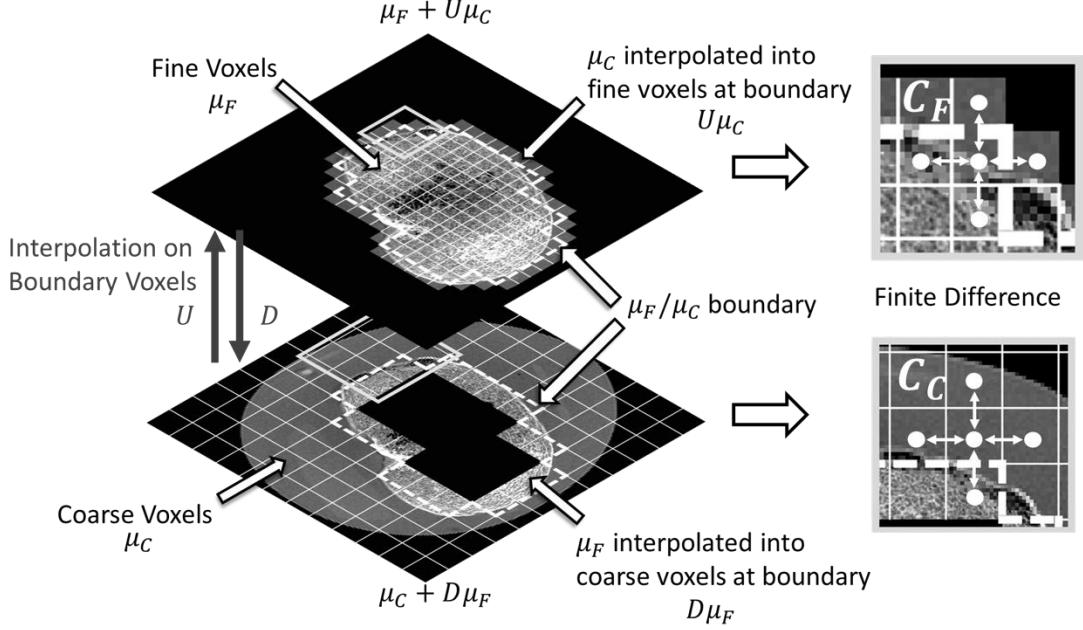


Figure 5.2 Schematic of multiresolution regularization scheme illustrating how the regularization over boundaries between the fine and coarse voxel grids is performed using interpolation operators.

The objective function shown in Eq. 5.2 can be optimized using the SPS approach [61], [203]. We note that the same data-fit surrogate as in Eq. 5.17 of Elbakri and Fessler (2002) can be used here, with the standard projector operator replaced by the multiresolution projector. Crucially, the surrogate in the SPS algorithm is separable across voxels, i.e. the update for each voxel depends on the surrounding volume only through the line integrals of the previous iterate of the reconstruction. This yields a multiresolution algorithm that allows independent updates of the fine- and coarse-grid ROIs, as detailed in Table 1. Similarly, the separable surrogate for the penalty of [206] can be applied directly to the multiresolution objective, as the construction of the surrogate is independent of

discretization. The algorithm presented in Table 1 can be easily generalized to an ordered subset form in a manner analogous to [203].

Definitions	
\hat{l}_B, \hat{l}_N	forward-projected line integrals onto binned and native detector pixels
l_B, l_N	binned and native preprocessed line integral data
w_B, w_N	diagonal entries in weighting matrices \mathbf{W}_B and \mathbf{W}_N .
c	precomputed PWLS curvature, $c = [\mathbf{W}]_+$
$\mathbf{1}$	volume covering the entire reconstruction FOV of all 1's.
\dot{L}_C, \dot{L}_F	derivatives of data-fit surrogates for voxels of $\hat{\mu}_C$ and $\hat{\mu}_F$.
d_C, d_F	curvatures of data-fit surrogates for voxels of $\hat{\mu}_C$ and $\hat{\mu}_F$.
$\dot{\psi}, \omega_{\psi_k}$	gradient and curvature of the penalty function ψ .
\dot{R}_C, \dot{R}_F	derivatives of penalty surrogates for voxels of $\hat{\mu}_C$ and $\hat{\mu}_F$.
r_C, r_F	curvatures of penalty surrogates for voxels of $\hat{\mu}_C$ and $\hat{\mu}_F$.
Algorithm	
Precompute $d_C = A_{CB}^T [\tilde{A} \mathbf{1} \circ c]_B + A_{CN}^T [\tilde{A} \mathbf{1} \circ c]_N \quad d_F = A_{FN}^T [\tilde{A} \mathbf{1} \circ c]_N$ for iteration $n=1, \dots, N$	
<i>Coarse Grid Update</i> $(j \in \text{coarse voxels})$ $\hat{l}_B = A_{CB} \hat{\mu}_C$ $\dot{h}_B = w_B \circ (\hat{l}_B - l_B)$ $\dot{L}_C = A_{CB}^T \dot{h}_B + A_{CN}^T \dot{h}_N$ $\dot{R}_{Cj} = \sum_{k \in K_C} c_{kj} \dot{\psi}([\tilde{C} \hat{\mu}]_k)$ $r_{Cj} = \sum_{k \in K_C} c_{kj}^2 \omega_{\psi_k} ([\tilde{C} \hat{\mu}]_k)$ $\hat{u}_C = \max (\hat{u}_C - \frac{\dot{L}_C + \beta_C \dot{R}_C}{d_C + \beta_C r_C}, 0)$	
<i>Fine Grid Update</i> $(i \in \text{fine voxels})$ $\hat{l}_N = A_{CN} \hat{\mu}_C + A_{FN} \hat{\mu}_F$ $\dot{h}_N = w_N \circ (\hat{l}_N - l_N)$ $\dot{L}_F = A_{FN}^T \dot{h}_N$ $\dot{R}_{Fi} = \sum_{k \in K_F} c_{ki} \dot{\psi}([\tilde{C} \hat{\mu}]_k)$ $r_{Fi} = \sum_{k \in K_F} c_{ki}^2 \omega_{\psi_k} ([\tilde{C} \hat{\mu}]_k)$ $\hat{u}_F = \max (\hat{u}_F - \frac{\dot{L}_F + \beta_F \dot{R}_F}{d_F + \beta_F r_F}, 0)$	
end	

Table 5.1 Glossary of terms and pseudocode of SPS optimization method for multiresolution PWLS. Ordered subsets are omitted for simplification.

V.B.2 Experimental Setup for Simulation Studies

The proposed multiresolution approach was evaluated in a simulation study. The primary goal of the study was to evaluate the impact of different coarse voxel

downsampling factors and the effect of regularization strength in the coarse and fine grid volumes on image artifacts in the fine grid region. Figure 5.3 shows the digital phantom used in simulations. The phantom included soft tissues (emulating fat and muscle) and two bone regions. The bones consisted of a cortical shell and fine trabecular features based on a segmented micro-CT image. The attenuation values of the simulated tissues corresponded to effective beam energy of 60 keV, which approximates that of the 90 kV (4 mm Al, 0.3 mm Cu added filtration) spectrum of a dedicated extremity CBCT at our institution [165], [207]. The digital phantom used 0.075 mm voxels and simulated projections were generated on a 0.194 mm pixel grid. A bare beam fluence of 10^5 photons per detector pixel was assumed and Poisson noise was added to the projections. Simulated detector pixels had 100% efficiency. System geometry emulated the extremities CBCT scanner: SAD of 436 mm and SDD of 560 mm.

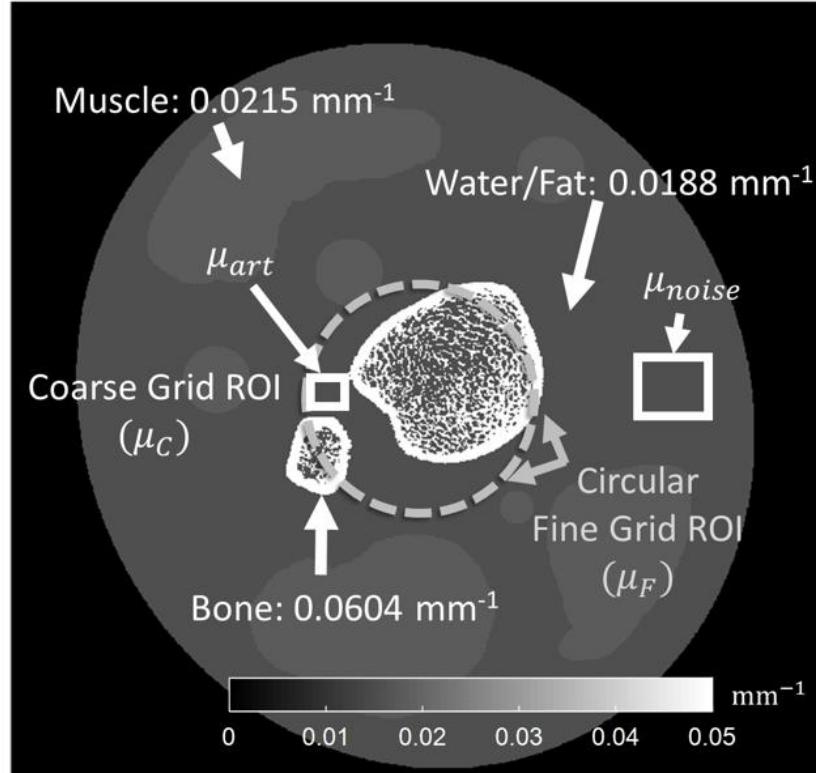


Figure 5.3 Schematic of digital bone phantom with the boundary of a central fine voxel grid region marked with a thick dashed line. The ROIs used for measurement of artifacts in the fine grid region (μ_{art}) and noise (μ_{noise}) in the coarse grid region are indicated with boxes.

In the multiresolution reconstruction, the regularization strength of the interior fine grid ROI β_F was fixed at 10^3 based on visual assessment of spatial resolution. 60 subsets were used. The reconstructed voxel size in the fine grid ROI was 0.15 mm, whereas the voxel size in the coarse grid region was varied by changing the downsampling factor η_μ from 1 (i.e., the same voxel size for μ_F and μ_C) to 10 (i.e., 10x larger voxels for μ_C). In cases where detector binning was considered in addition to a multiresolution volume representation, the binning factor (η_b) applied to the simulated projections was kept equal to the volume downsampling factor. The forward and back-projection operators in the simulation and reconstruction used an in-house GPU implementation of separable

footprints algorithm using trapezoidal functions [58]. The separable footprint algorithm is voxel-driven and thus can easily accommodate arbitrary ROI shapes. However, our current implementation of the projector relies on GPU texture memory and thus only handles rectangular volumes. Non-rectangular fine grid ROIs were simulated using a minimum bounding box and masking operations.

In addition to varying the coarse voxel size, the regularization strength β_C in the coarse region was also varied. Since image downsampling imparts a degree of smoothing, the effect of a specific β_C value will differ between reconstructions with different coarse voxel sizes. To facilitate comparison of the effects of regularization across a range of voxel sizes (downsampling factors), a normalization that accounts for the effects of sampling was applied to the quadratic penalty [208]. Following the derivation by Yu, for quadratic penalties:

$$\psi = \frac{t^2}{2} = \frac{(\mu_i - \mu_k)^2}{2} \approx \frac{[\nabla\mu \cdot \Delta_{ki}]^2}{2} = \Delta_{ki}^2 \frac{\nabla\mu^2}{2} \quad (5.9)$$

where Δ_{ki} is the distance between neighboring voxels k and i . Eq. (5.9) implies that within a neighborhood where the change in underlying attenuation volume is smooth, the action of the penalty (for a fixed penalty strength) scales approximately as square of the voxel size or, equivalently, as square of η_μ . Henceforth, reconstructions are compared across a range of η_μ in a manner that accounts for additional blur due to downsampling by using the following normalized penalty strength, denoted as β :

$$\beta = (1/\eta_\mu^2)\beta_C \quad (5.10)$$

In our studies we varied β from 0 (no regularization) to 10^8 for each value of η_μ .

To assess the artifacts in the fine grid region, a small (3 mm x 3.8 mm) ROI, μ_{art} , was selected in a uniform area of soft tissue adjacent to two bone edges (Fig. 5.3), where initial evaluation indicated pronounced streaking in reconstructions with high downsampling and regularization of the coarse grid region. To quantify artifacts, reconstruction of noiseless projection data were generated for each set of parameters of the multiresolution reconstruction. RMSE was calculated in the ROI between the noiseless multiresolution reconstruction and a truth image that was obtained by downsampling the digital phantom (0.075 mm voxels) to the voxel size of the fine grid region (0.15 mm):

$$RMSE = \sqrt{\frac{\sum_j (\hat{\mu}_{art_j} - \mu_{art_j})^2}{N}} \quad (5.11)$$

where $\hat{\mu}_{art_j}$ are voxels in the artifact ROI of the noiseless multiresolution reconstruction, μ_{art_j} are the corresponding voxels in the downsampled digital phantom and N is the total number of voxels in μ_{art} .

Increase in regularization strength and increase in voxel downsampling both impart resolution loss and noise reduction in the coarse grid region. The noise in μ_C (denoted as SD_c) was used as a metric quantifying the cumulative effect of coarse discretization and regularization. The noise was measured as RMSE between a flat region in the coarse grid subvolume of a noisy reconstruction, denoted as μ_{noise} (Fig. 5.3), and the same region in the truth image.

V.B.3 Benchtop Study of Anthropomorphic Knee Phantom

An anthropomorphic lower extremity phantom was scanned on a CBCT test bench [209] simulating a dedicated extremity CBCT (SAD=435.7 mm, SDD=559.2 mm). The

detector was a PaxScan4030CB (Varian, Palo Alto, CA) operated at 1.5 fps and 0.194 mm pixel size. The x-ray source was a rotating anode DU694 x-ray tube with 14° anode angle, enclosed in EA10 housing (Dunlee, Aurora, IL). The scan consisted of 360 projections acquired over 360° at 90 kV (+0.2 mm Cu, 2 mm Al), with 0.4 mAs per projection.

The total reconstructed volume was 18x18x18 cm³. The irregularly shaped fine grid ROI contained a region of subchondral bone identified using an initial Feldkamp-Davis-Kress (FDK) analytical reconstruction and consisted of 31,623,168 voxels of 0.15 mm size. There were 26,505,888 voxels of 0.6 mm size in the coarse grid region ($\eta_\mu = 4$). Multiresolution detector binning with $\eta_y = 4$ was applied, with the region of fine pixels in each view matching a forward projection of the subchondral bone mask. Regularization strengths were set to $\beta_F = 10^{-0.5}$ and $\beta_C = 10^4$.

V.B.4 Analysis of Runtime and Memory Footprint

To estimate the savings in computation time and memory consumption provided by the multiresolution method, benchmarking was performed for a 12x12x12 cm³ volume. This volume is smaller than the one used in the anthropomorphic phantom experiments of Sec. V.B.3. This is because a fine grid reconstruction of the entire grid was needed for benchmarking. The memory available on the GPU used in the experiments was not sufficient to fit a fine voxel grid larger than 12x12x12 cm³. Note however that this size of the volume is sufficient to cover a typical knee joint on the extremity CBCT system.

Fine grid ROIs with 0.15 mm voxels and varying volume (5x5x5 cm³ to 12x12x12 cm³) were considered. For each fine grid ROI, the remainder of the FOV was parameterized using a coarse grid over a range of upsampling factors ($\eta_\mu = 1, 2, 4$). Since computation

times for projection and backprojection vary and can be dependent on implementation, we considered the computational cost associated with projection-backprojection pairs. Average total projection-backprojection times were obtained separately for each fine grid volume (denoted as t_f) and for coarse grid volumes covering the entire FOV (t_c) at the various coarse grid voxel sizes. Furthermore, for each fine grid ROI, the projection-backprojection time for a volume equal in size to the fine grid ROI, but parameterized using coarse voxels, was measured (t_Δ). The coarse grid projection-backprojection times were computed with and without projection binning. The estimate of projection-backprojection time for a given multiresolution parameterization (defined by a combination of fine grid volume size and upsampling factor) is $t_f + t_c - t_\Delta$, assuming that projection and backprojection times are linear with the number of voxels. The validity of this assumption was confirmed for the range of volume discretizations considered here for both projection and backprojection operations. The reconstruction speedup was computed as the ratio of the average measured time for a projection-backprojection pair of our multiresolution implementation to the same average measured time using only the fine grid ($\eta_\mu = 1$).

Each PWLS iteration requires storing the following variables: the current image μ , derivatives of data fit and regularization surrogates, \dot{L} and \dot{R} , the curvatures of data fit and regularization surrogates, d_j and r , and the measured projections. (Here we ignored potential memory savings due to ordered subsets that may be offset by increased memory transfer times.) Memory footprint was thus calculated as the size of the projection dataset plus five times the size of the image volume, assuming all volumes are kept in memory. The memory footprint could be reduced by reusing the arrays associated with some of the variables, however at the expense of increased data transfer latency.

V.C Results

V.C.1 Digital Phantom Study

Fig. 5.4 illustrates the convergence properties of multiresolution PWLS reconstruction as a function of the downsampling factor. For each value of η_μ , we computed root-mean-square difference (RMSD) between the volume estimate at the n^{th} iteration and a nearly converged estimate (approximated by the solution after 200 iterations). RMSD computed within the fine grid ROI (dashed line) and within the coarse grid ROI (solid lines) are shown separately. 50 iterations of multiresolution PWLS were sufficient to achieve an RMSD of less than 10^{-4} mm^{-1} with this stable solution in both ROIs and across all values of η_μ . The RMSD behavior in μ_F was the same across all downsampling factors. Convergence was generally faster in μ_C than in μ_F , even at $\eta_\mu = 1$ (equal voxel size in both grids), likely because the phantom contains fewer high-frequency structures in the coarse grid ROI. The convergence rate in μ_C increases with higher downsampling. This behavior is intuitive since coarser grids (with fewer parameters to estimate) generally have better condition numbers, leading to faster convergence. Overall, the number of iterations needed to arrive at a stable solution in multiresolution PWLS is determined by the convergence rate in fine grid ROI and is thus independent of η_μ . The crucial advantage of using higher downsampling is in improved time per iteration. Based on this analysis, 50 iterations of multiresolution PWLS were used throughout this study.

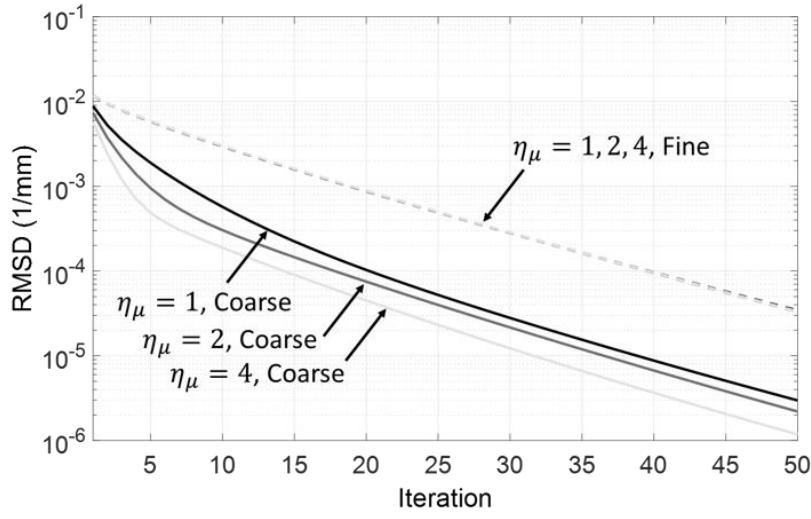


Figure 5.4 Convergence profile measured in reconstructions of the digital phantom. The estimate at the current iteration is compared with a converged image at 200 iterations. $\beta_F = 10^{3.5}$; $\beta_C = 10^{4.5}$. The RMSD curves for the fine grid region (dashed line) overlap for all values of η_μ .

Fig. 5.5 shows a comparison of multiresolution PWLS reconstructions (without projection binning in (A) and with projection binning in (B)) across a range of downsampling factors and normalized regularization strengths. Note that reconstructions with matched β but unmatched η_μ use different β_C , as given by Eq. 5.12. The case of $\eta_\mu = 1$ corresponds to the same voxel size in μ_F and μ_C , but different values of β_F and β_C . A zoomed ROI is shown for each reconstruction that covers the fine grid ROI. Note that an increase in coarse grid regularization increases streak artifacts in the fine grid region. The artifacts become pronounced at large values of the downsampling factor. Additionally, when detector pixels are binned (Fig. 5.5B), noticeable streak artifacts are apparent in the coarse grid near the ROI boundary in the case where $\eta_\mu = 10$. The right-most column in Fig. 5.5AB shows a difference image computed between the case of $\eta_\mu = 10$ and a reference image at $\eta_\mu = 1$ (for this computation, the coarse grid voxels for $\eta_\mu = 10$ were upsampled using nearest neighbor interpolation). Even for this high downsampling factor,

there is no distortion in the trabecular features contained in the fine grid ROI μ_F (note that the same fine grid regularization strength β_F is applied in the downsampled and reference reconstructions).

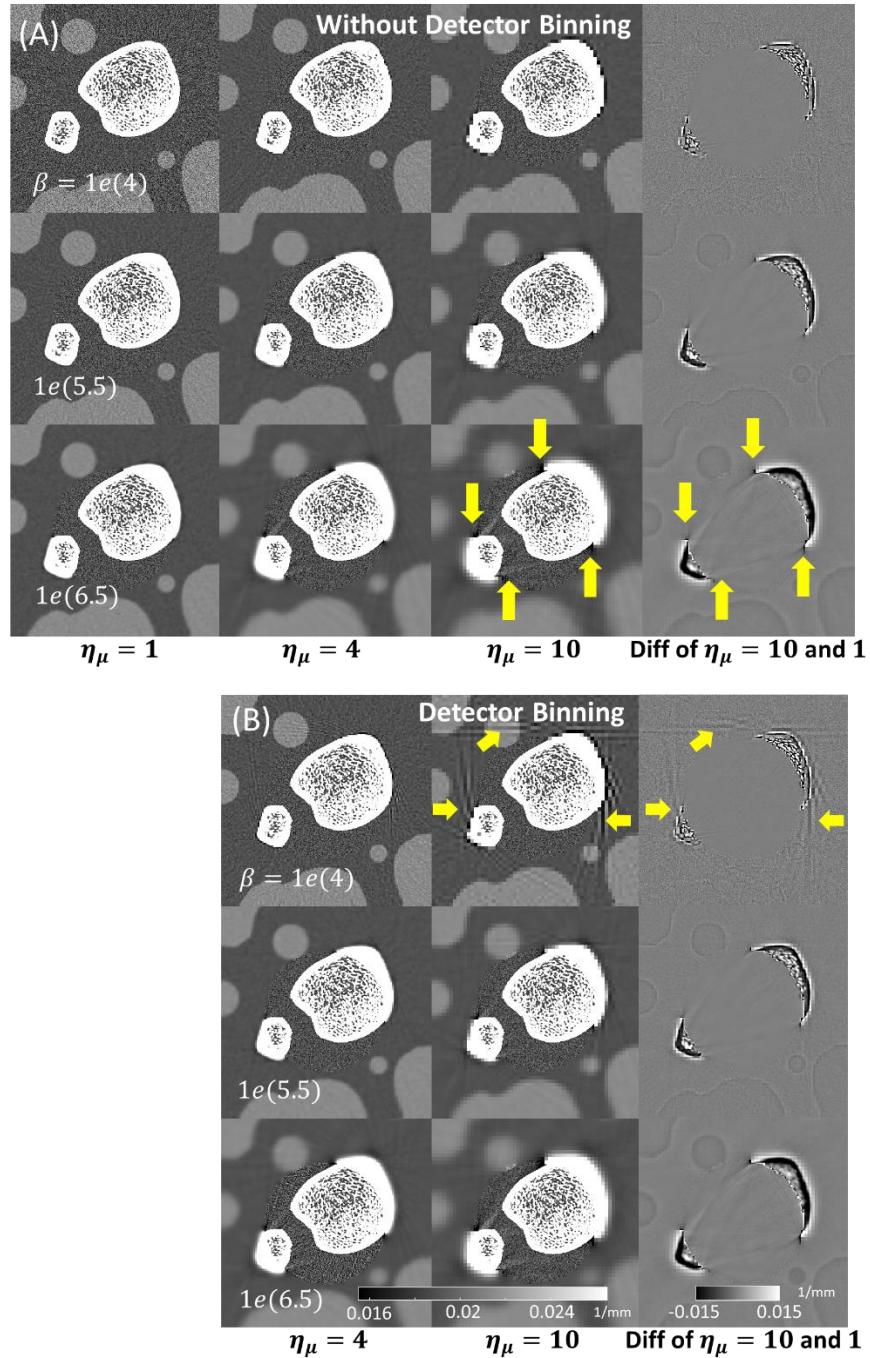


Figure 5.5 Multiresolution PWLS reconstructions for different downsampling factors and coarse grid regularization values. A central region of the digital phantom is shown for PWLS without detector binning (A) and with detector binning (B). The last column in each subfigure is a difference image between the case with maximal downsampling ($\eta_\mu = 10$) and the case of no downsampling ($\eta_\mu = 1$). Arrows indicate artifacts due to strong coarse grid downsampling and regularization.

Fig. 5.6A investigates the trade-off between the strength of artifacts in μ_F , measured using Eq. 5.11, and the cumulative effect of regularization and downsampling in μ_C , represented by the level of noise in the coarse grid region. The results are very similar for cases with and without detector binning; for clarity, results are shown only for reconstructions with detector binning.

For all downsampling levels, the artifact metric exhibits a plateau across a range of low-to-moderate noise levels (standard deviations of 10^{-4} to 10^{-3} mm^{-1} compared to muscle attenuation value of $\sim 2 \times 10^{-2} \text{ mm}^{-1}$). In this regime, regularization strength in μ_C can be adjusted without introducing significant artifacts in the fine grid ROI. The RMSE rapidly increases at lower noise levels, corresponding to increased smoothing in μ_C . The graphs in Figure 5.6A are largely overlapping, indicating that the magnitude of the artifacts is a function of the overall blur in the coarse grid ROI, regardless of whether the blur results from downsampling or regularization. Consequently, when a lower value of the downsampling factor is used, there is more room to adjust the regularization in the coarse grid ROI, i.e. β_C can be varied over a broader range of values without introducing artifacts in the fine grid region. If the coarse grid region is of clinical interest, this ability to tune the resolution-noise tradeoff in μ_C when the downsampling factor is relatively low may be relevant for optimizing the performance of the reconstruction.

Figure 5.6B illustrates the artifacts in the fine grid region using a zoom on the artifact ROI (μ_{art}). As the normalized penalty strength (β) and η_μ increase, streaking becomes apparent. The more downsampling, the lower the value of normalized β where the streaking becomes pronounced.

Further examination of Figure 5.6A reveals a slight increase in RMSE at the lowest regularizations (highest noise levels) in the presence of downsampling ($\eta_\mu \geq 4$). This is attributed to bias in the forward projection of μ_C introduced by sharp edges in the coarse discretization of fat-muscle interfaces of the phantom. This effect is diminished when sufficiently high levels of regularization are applied in the coarse grid region, blurring the discretization-induced sharp intensity transitions. The artifact due to sharp edges in weakly regularized μ_C is visible as a pronounced cross-hatch pattern at $\eta_\mu \geq 4$ and low β values in Figure 5.6B. Note that a similar, but much less conspicuous pattern is visible at $\eta_\mu = 1$, where it is likely to represent a combination of various discretization artifacts that are often present in noiseless reconstruction of digital phantoms with sharp edges.

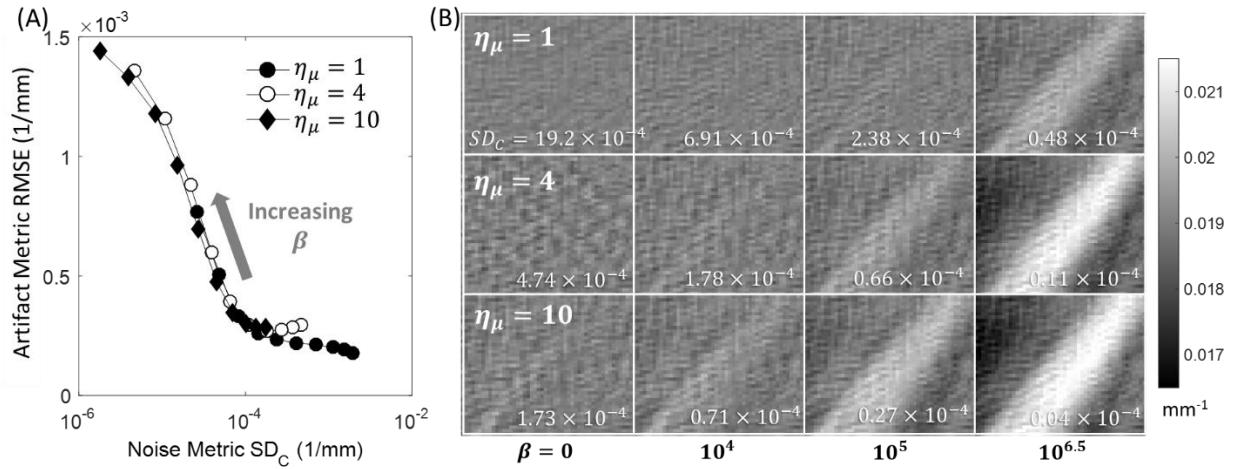


Figure 5.6 (A) Magnitude of the artifact in fine grid ROI (RMSE of μ_{art} in Fig. 5.3, given by Eq. 5.11) as a function of noise in the coarse grid region. The noise metric quantifies the cumulative effect of regularization and voxel downsampling in μ_C . (B) The artifact ROI (μ_{art}) in the fine grid region for a range of downsampling factors (rows) and normalized coarse grid regularization strengths β (columns).

V.C.2 Effect of the Location of the Fine Grid ROI

Examination of Fig. 5.5 suggests that the most pronounced artifacts in the fine grid ROI emerge from areas where the interface between μ_F and μ_C crosses regions of high

contrast and sharp intensity transitions. The appearance and magnitude of the artifacts will depend on the location of the fine and coarse grid ROIs, as illustrated in Fig. 5.7 for a case with both voxel downsampling and projection binning. Here, the boundaries of μ_F (marked with an orange dashed line) were designed to avoid cortical bone boundaries and other high contrast gradients. The downsampling factor was set to 10, which leads to severe artifacts in the reconstructions of Fig. 5.5 (5th column). With the fine grid ROI conforming to the bone boundaries, however, no streaking artifacts are visible in the space between the two bones. A strong mismatch in the magnitude of noise in μ_C and μ_F is apparent at higher values of normalized regularization strength. Similar to Fig. 5.5, the right-most column of Fig. 5.7 investigates a difference image between the downsampled case and the reference image at $\eta_\mu = 1$ (equal voxel size in the coarse and fine grids). No visible distortions in the trabecular structure were introduced using multiresolution reconstruction.

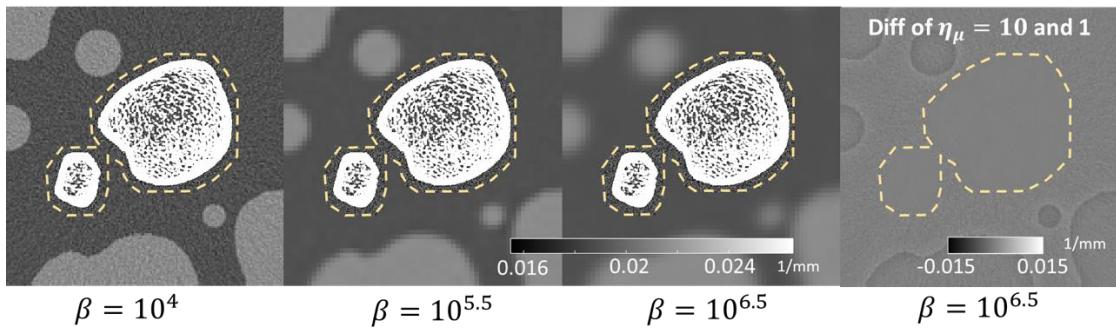


Figure 5.7 Multiresolution PWLS with fine grid ROI conforming to bone boundaries (marked with dashed line in the leftmost mage) for three values of the normalized penalty strength β . The downsampling factor η_μ is set to 10. The right-most subfigure shows a difference image between the downsampled case ($\beta = 10^{6.5}$) and the reference reconstruction with no downsampling. Projection binning is included.

V.C.3 Reconstruction of an Anthropomorphic Knee Phantom

Figure 5.8 shows a reconstruction of an anthropomorphic knee phantom acquired on the CBCT test bench. Fig. 5.8A illustrates the boundaries of the fine grid ROI. Four-

fold downsampling of the volume and projections was employed. The multiresolution PWLS used 3.36% of the number of voxels and 17.7% of the number of detector pixels compared to a full FOV fine grid reconstruction. No artifacts arise from downsampling, and high-resolution trabecular features are visible in the fine grid ROI [outlined with orange dashed line in Fig. 5.8B-D]. The resulting multiresolution image can be stored in less than 0.5 GB, whereas for the full-field fine grid image, approximately 7 GB would be required.

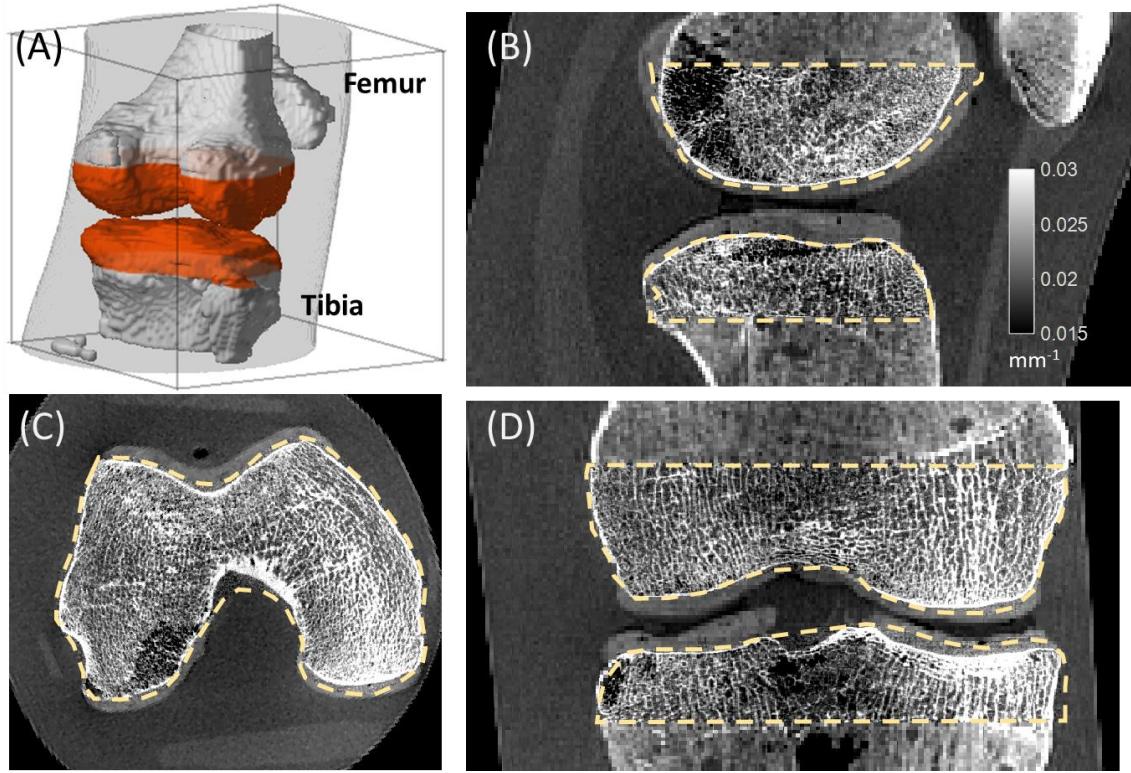


Figure 5.8 Multiresolution PWLS reconstruction of an anthropomorphic knee phantom acquired on a CBCT benchtop. Fine and coarse grid regions are delineated in A. B-D shows details of the high-resolution trabecular ROI in the sagittal, axial, and coronal planes. Boundaries of the ROI are marked with a dashed line.

V.C.4 Computational Cost of Multiresolution PWLS

Fig. 5.9 summarizes the investigation of computational savings provided by multiresolution PWLS. As expected, the speedup factor (compared to reconstruction using

fine grid voxels throughout the entire FOV) can be as high as a 5x - 10x when the fine grid ROI is a relatively small fraction of the total size of the FOV. Most of the speedup is from image downsampling, rather than detector binning, which is to be expected with the voxel-driven forward projection method used here.

Fig. 5.9B examines the reduction in memory footprint provided by multiresolution PWLS. As anticipated, the memory savings are enhanced when projection binning is used in conjunction with volume downsampling. For a fine grid ROI of a similar size as the one used in the anthropomorphic knee phantom of Sec. V.C.3, the memory footprint reduction is approximately 20% compared to reconstruction using only the fine grid voxels.

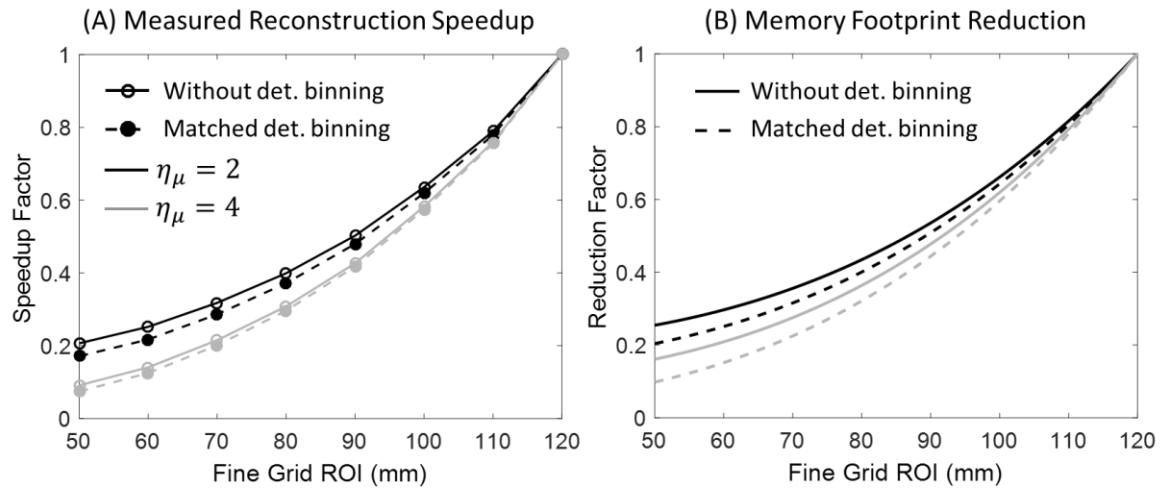


Figure 5.9 (A) Measured reconstruction speedup of multiresolution PWLS as a function of the size of fine grid FOV (for total volume size of 120x120x120 mm³). (B) Total memory consumption of downsampled cases compared to fine grid cases.

V.D Discussion and Conclusions

We developed a PWLS reconstruction algorithm implementing a multiresolution voxelized parameterization of the object. The algorithm was evaluated for application in accelerated iterative reconstruction of finely sampled ROIs (voxel sizes <100 μm) from high-resolution extremity CBCT projection data. It was assumed that the fine grid region

is of the primary clinical interest and thus the investigation was focused on mitigation of artifacts in the fine grid region and quantification of the reconstruction speedup.

Streaking artifacts found in the fine grid ROI are likely caused by inconsistencies between the simulated forward projections of the coarsely and finely sampled ROIs. Artifact magnitude is thus a function of the cumulative blur due to downsampling and regularization of the coarse grid region. For example, downsampling factors as high as 10x were used without introducing visible streaking in the fine grid region by applying a relatively weak regularization in the coarsely sampled sub-volume. However, the tradeoff was that the μ_C exhibited a “blocky” appearance, which could limit the diagnostic utility of the coarsely sampled region. This may however be acceptable if the clinical interest is primarily in the fine grid ROI and achieving maximal speedup is a priority (and the coarsely sampled region may be reconstructed separately by other means using parameters suitable to soft-tissue visualization).

Additional memory savings can be attained when the multiresolution representation is augmented with binning of projection data. This work evaluated a scenario where projection binning was applied to pixels receiving the line integrals that traversed only through μ_C . This approach results in similar performance with respect to artifacts in the fine grid region as the method using only volume downsampling. In the coarse grid region, however, streaks tangential to the ROI boundary were found when using high levels of projection binning ($\eta_y = 10$) in conjunction with volume downsampling. The artifacts are hypothesized to emerge from mismatches in simulated projections of some of the coarse grid voxels located close to the ROI boundary. Such voxels are projected both on the binned and native pixels, depending on whether a given line integral traverses through μ_F for each

projection angle. Since this inconsistency is not present for μ_F voxels, the fine grid ROI was unaffected.

The simulation study of Sec. V.C.1 involves an ROI boundary that crosses through high contrast bone regions, resulting in a challenging scenario for multiresolution PWLS. Sec. 3.2 illustrates that both types of streaking artifacts discussed above can be effectively mitigated using a fine grid ROI that conforms to high contrast edges.

Overall, the optimal value of the downsampling factor will depend on the location of the fine and coarse grid regions and on the clinical application (i.e. whether both ROIs are of clinical interest, or only the finely sampled sub-volume). The results indicate that downsampling factors of $\sim 4x$ are possible without perceptible artifacts in either of the ROIs. At this level of η_μ , the regularization strength in the coarsely sampled sub-volume can be adjusted over a relatively broad range without adversely affecting the fine grid region. For a typical knee volume, this downsampling corresponds to more than 5-fold acceleration of the iterative reconstruction. The current execution time for the multiresolution reconstruction of the knee phantom of Figure 8, obtained with $\eta_\mu=4$, is ~ 2 min per iteration (or ~ 100 min for the complete reconstruction of 50 iterations) using an un-optimized MATLAB-CUDA implementation. A naïve approach utilizing fine voxels throughout the whole volume would take ~ 10 hours for the reconstruction to complete and require $\sim 15x$ more memory, and thus the multiresolution algorithm provided a significant step towards achieving clinically practical runtimes for iterative reconstruction of the trabecular detail in human studies. We anticipate that an additional 5x acceleration over the execution time stated here will be possible with an optimized implementation using a compiled executable with improved memory management on a multi-GPU workstation.

For certain configurations of the coarse and fine grid regions, downsampling factors as high as 10x can be used. For the typical knee CBCT (e.g. the sample reconstruction in Fig. 5.8), the 10x downsampling should yield ~52x speedup in PWLS, assuming a projection time that scales linearly with number of voxels. The acceleration factor will increase if the high-resolution ROI is a smaller fraction of the total volume.

Similar to the selection of the downsampling factor, the location, shape and the procedure for delineating the fine grid ROI will depend on the clinical application. The knee phantom reconstruction in Figure 8 represents a scenario where the scan parameters were selected to yield projection data that supports high-resolution reconstruction of the entire knee joint. In this case, a high fidelity FDK reconstruction could be obtained using the same projection data (thus at no additional dose), enabling precise definition of the fine grid ROI. However, such close adherence to bony boundaries is not always necessary. As shown in Fig. 5.7, image artifacts due to downsampling will be minimized as long as the boundaries of the fine grid ROI do not cross high contrast edges. Detection of such high-contrast edges will typically be possible in an initial FDK reconstruction even if the projection data is noisy or sparsely sampled, as in low-dose imaging protocols intended for use with iterative reconstruction algorithms. In such cases, the initial analytical reconstruction may be non-diagnostic, but will likely be sufficient to delineate the fine grid ROI for the iterative algorithm. Furthermore, while the use of precisely defined fine grid ROIs (conforming with the object of interest) optimizes reconstruction time by minimizing the volume parameterized with fine voxels, the low dose applications are unlikely to yield data supporting reconstructions on very fine voxels. The use of sub-optimal fine grid ROIs

(e.g. slightly larger than the object of interest) for low dose data will thus not be as detrimental for reconstruction performance as in the case of high-resolution imaging.

Other factors to be considered in choosing the shape of the fine grid ROI involve computational burden and ease of implementation. In principle, the shape of the ROIs is only limited by the requirement that the fine grid voxel size must be an integer fraction of the coarse voxel size to ensure complete coverage of the volume. A fine grid ROI that is tightly matched to the shape of area of high-resolution reconstruction has the advantage of minimizing the memory usage. At the same time, the execution speed may be hampered by non-contiguous memory access patterns that are likely to emerge when the fine grid ROI consists of multiple disjoint patches. Rectangular ROIs are easier to map to contiguous memory blocks and may thus yield improved computation speed compared to more complex ROI shapes when executed using general purpose software. In practice, the use of regular, rectangular or cylindrical fine grid ROIs loosely following the boundaries of the objects of interest and avoiding sharp tissue boundaries, combined with downsampling factors of 4-5, is likely to yield robust performance in typical clinical applications of multiresolution PWLS.

In conclusion, the proposed multiresolution algorithm for PWLS was tested in application to extremity CBCT. Artifact-free reconstructions of finely sampled ROIs were achieved at computation times that are 5x - 10x shorter compared to a brute force solution that applies fine voxel parameterization to the entire volume. The multiresolution framework can be further expanded to support other applications where computationally expensive forward models (for example, models with modeling of detector blur [184]) are applied only to a sub-region of the field-of-view. The algorithm permits application of

different regularization strengths in the regions of coarse and fine grid, enabling reconstructions in which the sampling and regularization are individually optimized depending on the clinical task in each image region.

Chapter VI. Future Work

Chapters of this dissertation span the development of new systems (high-resolution CMOS detectors), models (imaging system models for advanced reconstruction methods and evaluation of image quality), and analysis techniques (quantitative evaluation of bone microstructure, density, and joint space morphology). Each of these developments helped advance CBCT spatial resolution and quantitative imaging capability, focusing specifically on application to MSK imaging.

One of the most difficult bottlenecks for implementing the electrostatic model described in Chapter II is the accurate segmentation of the femur and tibia. This process is especially difficult for patients with severe cartilage degradation or significant osteophyte growth, where parts of the femoral and tibial surfaces are <1 pixel apart and must be manually segmented. Since the publication of this work, deep-learning models have emerged that could provide accurate segmentations various anatomical structures [210], [211]. These models could potentially be applied for automatic segmentation in challenging cases.

The polyPL algorithm in Chapter III can be extended for quantitative imaging in the presence of Ti/Stainless Steel implants for fracture healing applications. This can be achieved by varying the basis material and material fraction functions of voxels in the vicinity of the implant. Preliminary results using Ti implants showed significant artifact reduction using polyPL [116]. One interesting observation is that artifact reduction in this case is strongly dependent on the glare correction kernel, especially terms associated with

high-frequency blur. This may suggest potential benefits from incorporating more sophisticated models of blur and glare into polyPL [64] for reconstruction of highly attenuating objects. These more sophisticated models could be implemented with multiresolution reconstruction described in Chapter V to maintain competitive reconstruction time. Additionally, there are artifacts not corrected in the current workflow, notably patient motion. More work is needed to integrate polyPL with existing motion-correction algorithms for extremity CBCT [183]. Finally, quantitative accuracy was calculated on an ROI on the order of few millimeters. It may be of interest to perform quantitative BMD measurements of bone microstructure on high-resolution images, such as ones obtainable on a CMOS detector (Chapter IV). In this case, quantitation may be impacted by noise, and regularization design will be important.

Chapter IV addressed the optimization and integration of CMOS technology in a dedicated CBCT extremity scanner to achieve high-resolution acquisition and enhanced visualization of bone microstructure. We constructed a cascaded systems model of the detector and performed task-based optimization of its scintillator thickness, leveraging the trade-off between resolution and detection efficiency. At the time of writing this dissertation, efforts have been made to laser-machine a finely-pixelated structure ($\sim 90 \mu\text{m}$) in CsI:Tl scintillators to further restrict lateral spread of optical photons [212]. Despite decreased packing fraction from pixilation, the net gain in spatial resolution may yield better detectability for trabecular bone than optimizing CsI:Tl thickness alone.

Additionally, the high-resolution capability offered by CMOS extremity CBCT may be beneficial for inferring biomechanical properties of bone. The concept of combining CT with finite element modeling (FEM) of bone is not new [213], [214].

However, due to limited spatial resolution of conventional MDCT, prior literature has mostly regarded trabecular bone as an isotropic material. Visualization of trabecular features may be helpful in performing more accurate anisotropic FEM.

Data-driven methods with models such as encoders-decoders [215] and generative adversarial networks (GAN) [216] incorporating convolutional layers [217], [218] have already been applied to image reconstruction problems. These models make use of large sets of training data by propagating training-input into the model and finding model parameters that minimize error with training-output (*training phase*). Once trained, projection measurements are propagated through the model to generate a reconstructed output (*inference phase*). In contrast to MBIR, sophisticated knowledge of patient anatomy and physics is encoded implicitly into the model through training. Due to extensive use of priors, acquisition could be very sparse (a few projections in the case of CBCT), reducing patient dose and mitigating motion artifacts. Given sufficient training data and model capacity (# of parameters), even a simple data-driven model (for example, perceptrons) can theoretically approximate any arbitrary function that describes the reconstruction problem [219]. In addition, data-driven models often result in a faster solution, due to efficient hardware-accelerated propagation of input through the model.

In practice, whether sufficient training data is attained to achieve a generalizable inverse transform is often questioned. Moreover, the abundance of parameters in the model obfuscates first-principle physics inherent in image formation, making such models difficult to interpret. Finally, the limitations of these method are not well-studied and there are concerns such models are not generalizable, and may generate additional artifacts pertinent to clinical diagnosis [220]. For simple quantitative measurements such as

volumetric joint space width described in Chapter II, sparse reconstruction leveraging data-driven models of bone morphology may suffice. However, sparse reconstruction for high-resolution characterization of bone microarchitecture is more challenging, given information-theoretic limits and patient variability.

Bibliography

- [1] Centers for Disease Control and Prevention, “Prevalence of doctor-diagnosed arthritis and arthritis-attributable activity limitation--United States, 2010-2012.,” *Morb. Mortal. Wkly. Rep.*, vol. 62, no. 44, pp. 869–873, Nov. 2013.
- [2] G. Verbruggen, “Chondroprotective drugs in degenerative joint diseases,” *Rheumatology*, vol. 45, no. 2, pp. 129–138, 2006, doi: 10.1093/rheumatology/kei171.
- [3] G. Li *et al.*, “Identical subchondral bone microarchitecture pattern with increased bone resorption in rheumatoid arthritis as compared to osteoarthritis,” vol. 22, 2014, doi: 10.1016/j.joca.2014.08.015.
- [4] N. C. Wright *et al.*, “The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine,” *J. Bone Miner. Res.*, vol. 29, no. 11, pp. 2520–2526, 2014, doi: 10.1002/jbmr.2269.
- [5] R. D. Altman and G. E. Gold, *Atlas of individual radiographic features in osteoarthritis, revised.*, vol. 15 Suppl A. 2007.
- [6] J. E. Adams, “Quantitative computed tomography,” *Eur. J. Radiol.*, vol. 71, no. 3, pp. 415–424, 2009, doi: 10.1016/j.ejrad.2009.04.074.
- [7] P. Geusens *et al.*, “High-resolution in vivo imaging of bone and joints: a window to microarchitecture.,” *Nat. Rev. Rheumatol.*, vol. 10, no. 5, pp. 304–13, May 2014, doi: 10.1038/nrrheum.2014.23.

- [8] H. Weinans, M. Siebelt, R. Agricola, S. M. Botter, T. M. Piscaer, and J. H. Waarsing, “Pathophysiology of peri-articular bone changes in osteoarthritis,” *Bone*, vol. 51, no. 2, pp. 190–196, 2012, doi: 10.1016/j.bone.2012.02.002.
- [9] G. Zhen *et al.*, “Inhibition of TGF- β signaling in subchondral bone mesenchymal stem cells attenuates osteoarthritis,” *Nat. Med.*, vol. 19, no. 6, pp. 704–712, 2013, doi: 10.1038/nm.3143.Inhibition.
- [10] J. A. Carrino *et al.*, “Dedicated cone-beam CT system for extremity imaging,” *Radiology*, vol. 270, no. 3, pp. 816–24, Mar. 2014, doi: 10.1148/radiol.13130225.
- [11] S. Demehri *et al.*, “Assessment of image quality in soft tissue and bone visualization tasks for a dedicated extremity cone-beam CT system,” *Eur. Radiol.*, vol. 25, no. 6, pp. 1742–1751, 2015, doi: 10.1007/s00330-014-3546-6.
- [12] M. Kothari *et al.*, “Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis,” *Eur. Radiol.*, vol. 14, no. 9, pp. 1568–1573, 2004, doi: 10.1007/s00330-004-2312-6.
- [13] Q. Cao *et al.*, “Characterization of 3D joint space morphology using an electrostatic model (with application to osteoarthritis),” *Phys. Med. Biol.*, vol. 60, pp. 947–960, 2015.
- [14] N. A. Segal, M. C. Nevitt, J. A. Lynch, J. Niu, J. C. Torner, and A. Guermazi, “Diagnostic performance of 3D standing CT imaging for detection of knee osteoarthritis features,” *Phys. Sportsmed.*, vol. 43, no. 3, pp. 213–220, 2015, doi: 10.1080/00913847.2015.1074854.
- [15] J. Marzo, M. Kluczynski, A. Notino, and L. Bisson, “Comparison of a novel weightbearing cone beam computed tomography scanner versus a conventional

- computed comography scanner for measuring patellar instability,” *Orthop. J. Sport. Med.*, vol. 4, no. 12, pp. 1–7, 2016, doi: 10.1177/2325967116673560.
- [16] C. C. Netto *et al.*, “Flexible adult acquired flatfoot deformity: comparison between weightbearing and nonweightbearing cone-beam CT examinations,” in *AOFAS Annual Meeting*, 2016.
- [17] A. Hirschmann, C. W. A. Pfirrmann, G. Klammer, N. Espinosa, and F. M. Buck, “Upright Cone CT of the hindfoot: Comparison of the non-weight-bearing with the upright weight-bearing position,” *Eur. Radiol.*, vol. 24, no. 3, pp. 553–558, 2014, doi: 10.1007/s00330-013-3028-2.
- [18] A. J. Huang, C. Y. Chang, B. J. Thomas, P. J. MacMahon, and W. E. Palmer, “Using cone-beam CT as a low-dose 3D imaging technique for the extremities: initial experience in 50 subjects,” *Skeletal Radiol.*, vol. 44, no. 6, pp. 797–809, 2015, doi: 10.1007/s00256-015-2105-9.
- [19] M. Richter, B. Seidl, S. Zech, and S. Hahn, “PedCAT for 3D-imaging in standing position allows for more accurate bone position (angle) measurement than radiographs or CT,” *Foot Ankle Surg.*, vol. 20, no. 3, pp. 201–207, 2014, doi: 10.1016/j.fas.2014.04.004.
- [20] A. M. Cheung *et al.*, “High-resolution peripheral quantitative computed tomography for the assessment of bone strength and structure: A review by the canadian bone strength working group,” *Curr. Osteoporos. Rep.*, vol. 11, no. 2, pp. 136–146, 2013, doi: 10.1007/s11914-013-0140-9.
- [21] J. Bushberg, J. A. Seibert, E. M. Leidholdt, and J. M. Boone, *The essential physics of medical imaging*, 2nd Editio. Philadelphia: Lippincott Williams & Wilkins,

2002.

- [22] A. M. Hernandez and J. M. Boone, “Tungsten anode spectral model using interpolating cubic splines: unfiltered x-ray spectra from 20 kV to 640 kV.,” *Med. Phys.*, vol. 41, no. 4, p. 042101, 2014, doi: 10.1118/1.4866216.
- [23] J. H. Hubbell and S. M. Seltzer, “Tables of x-ray mass attenuation coefficients and mass energy absorption coefficients,” Gaithersburg, MD, 1996.
- [24] J. Punnoose, J. Xu, A. Sisniega, W. Zbijewski, and J. H. Siewerdsen, “Technical Note: spektr 3.0—A computational tool for x-ray spectrum modeling and analysis,” *Med. Phys.*, vol. 43, no. 3057, pp. 4711–41906, 2016, doi: 10.1118/1.1758350.
- [25] M. Berger *et al.*, “XCOM: photon cross sections database,” *NIST Stand. regerence database*, vol. 8, no. 1, pp. 3587–3597, 1998.
- [26] G. Hajdok, J. J. Battista, and I. A. Cunningham, “Fundamental x-ray interaction limits in diagnostic imaging detectors : Spatial resolution,” *Med. Phys.*, vol. 35, no. 3180, 2008, doi: 10.1118/1.2924219.
- [27] G. Hajdok, J. J. Battista, and I. a Cunningham, “Fundamental x-ray interaction limits in diagnostic imaging detectors: frequency-dependent Swank noise.,” *Med. Phys.*, vol. 35, no. 7, pp. 3194–3204, 2008, doi: 10.1118/1.2936412.
- [28] L. M. Cook, “Fractal fiber optics,” *Appl. Opt.*, vol. 30, no. 36, p. 5220, 1991, doi: 10.1364/ao.30.005220.
- [29] G. Zentai and I. Fellow, “Comparison of CMOS and a-Si flat panel imagers for x-ray imaging,” in *IEEE International Conference on Imaging Systems and Techniques*, 2011, pp. 194–200.

- [30] H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, “Model observers for assessment of image quality.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 90, no. 21, pp. 9758–9765, 1993, doi: 10.1109/NSSMIC.2002.1239406.
- [31] K. J. Myers, “Ideal observer models of visual signal detection,” in *Handbook of Medical Imaging, Volume 1: Physics and Psychophysics*, SPIE, 2000, pp. 559–592.
- [32] E. Samei and M. J. Flynn, “An experimental comparison of detector performance for direct and indirect digital radiography systems,” *Med. Phys.*, vol. 30, no. 4, pp. 608–622, 2003, doi: 10.1118/1.1561285.
- [33] ICRU, “ICRU report 54,” 2015.
- [34] R. F. Wagner and D. G. Brown, “Unified SNR analysis of medical imaging systems,” *Phys. Med. Biol.*, vol. 30, no. 6, pp. 498–515, 1985, doi: 10.1088/0031-9155/30/6/001.
- [35] S. Richard and J. H. Siewerdsen, “Comparison of model and human observer performance for detection and discrimination tasks using dual-energy x-ray images.,” *Med. Phys.*, vol. 35, no. 11, pp. 5043–5053, 2008, doi: 10.1118/1.2988161.
- [36] G. J. Gang, D. J. Tward, J. Lee, and J. H. Siewerdsen, “Anatomical background and generalized detectability in tomosynthesis and cone-beam CT,” *Med. Phys.*, vol. 37, no. 5, pp. 1948–1965, 2010, doi: Doi 10.1118/1.3352586.
- [37] I. Cunningham, M. Westmore, and A. Fenster, “A spatial-frequency dependent quantum accounting diagram and detective quantum efficiency model of signal and noise propagation in cascaded imaaina systems,” *Med. Phys.*, vol. 21, no. 3,

- pp. 417–427, 1994, doi: 10.1118/1.597401.
- [38] M. Rabbani, R. Shaw, and R. Van Metter, “Detective quantum efficiency of imaging systems with amplifying and scattering mechanisms,” *J. Opt. Soc. Am. A*, vol. 4, no. 5, p. 895, May 1987, doi: 10.1364/JOSAA.4.000895.
- [39] J. H. Siewerdsen *et al.*, “Empirical and theoretical investigation of the noise performance of indirect detection, active matrix flat-panel imagers (AMFPIs) for diagnostic radiology,” *Med. Phys.*, vol. 24, no. 1, pp. 71–89, 1997, doi: 10.1118/1.597919.
- [40] R. K. Swank, “Absorption and noise in x-ray phosphors,” *J. Appl. Phys.*, vol. 44, no. 9, pp. 4199–4203, 1973, doi: 10.1063/1.1662918.
- [41] W. Zhao, G. Ristic, and J. Rowlands, “X-ray imaging performance of structured cesium iodide scintillators.,” *Med. Phys.*, vol. 31, pp. 2594–2605, 2004, doi: 10.1118/1.1782676.
- [42] A. R. Lubinsky, W. Zhao, G. Ristic, and J. A. Rowlands, “Screen optics effects on detective quantum efficiency in digital radiography: zero-frequency effects.,” *Med. Phys.*, vol. 33, no. 5, pp. 1499–1509, 2006, doi: 10.1118/1.2188082.
- [43] J. Yao and I. A. Cunningham, “Parallel cascades: New ways to describe noise transfer in medical imaging systems,” *Med. Phys.*, vol. 28, no. 10, pp. 2020–2038, 2001, doi: 10.1118/1.1405842.
- [44] S. Vedantham, A. Karellas, and S. Suryanarayanan, “Solid-state fluoroscopic imager for high-resolution angiography: Parallel-cascaded linear systems analysis,” *Med. Phys.*, vol. 31, no. 5, pp. 1258–1268, 2004, doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

- [45] D. J. Tward and J. H. Siewerdsen, “Cascaded systems analysis of the 3D noise transfer characteristics of flat-panel cone-beam CT,” *Med. Phys.*, vol. 35, no. 12, pp. 5510–5529, 2008, doi: 10.1118/1.3002414.
- [46] L. A. Shepp and B. F. Logan, “Fourier reconstruction of a head section,” *IEEE Trans. Nucl. Sci.*, vol. NS-21, no. 3, pp. 21–43, 1974, doi: 10.1109/tns.1974.6499235.
- [47] R. Gordon and G. T. Herman, “Reconstruction of Pictures from Their Projections,” *Commun. ACM*, vol. 14, no. 12, pp. 759–768, 1971, doi: 10.1145/362919.362925.
- [48] R. M. Lewitt, “Alternatives to voxels for image representation in iterative reconstruction algorithms.,” *Phys. Med. Biol.*, vol. 37, no. 3, pp. 705–16, 1992, doi: 10.1088/0031-9155/37/3/015.
- [49] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts and Recent Advances*, 2nd ed. SPIE, 2009.
- [50] L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” *J. Opt. Soc. Am.*, vol. 1, no. 6, pp. 612–619, 1984.
- [51] D. L. Parker, “Optimal short scan convolution reconstruction for fanbeam CT.,” *Med. Phys.*, vol. 9, no. 2, pp. 254–257, 1982, doi: 10.1118/1.595078.
- [52] B. De Man and J. Fessler, “Statistical Iterative Reconstruction for X-Ray Computed Tomography,” *Biomed. Math. Promis. Dir. Imaging, Ther. Planning, Inverse Probl.*, no. m, pp. 1–28, 2009.
- [53] J. A. Fessler, “Medical image reconstruction: a brief overview of past milestones and future directions,” pp. 1–5, 2017.
- [54] J. W. Stayman, H. Dang, Y. Ding, and J. H. Siewerdsen, “PIRPLE: a penalized-

- likelihood framework for incorporation of prior images in CT reconstruction.,” *Phys. Med. Biol.*, vol. 58, no. 21, pp. 7563–82, Nov. 2013, doi: 10.1088/0031-9155/58/21/7563.
- [55] T. M. Peters, “Algorithms for Fast Back- and Re-Projection in Computed Tomography,” *IEEE Trans. Nucl. Sci.*, vol. 28, no. 4, pp. 3641–3647, 1981.
- [56] R. Siddon, “Fast calculation of the exact radiological path for a three-dimensional CT array,” *Med. Phys.*, vol. 12, no. 2, pp. 252–255, 1985.
- [57] B. De Man and S. Basu, “Distance-driven projection and backprojection in three dimensions,” *Phys. Med. Biol.*, vol. 49, no. 11, pp. 2463–2475, 2004, doi: 10.1088/0031-9155/49/11/024.
- [58] Y. Long, J. a Fessler, and J. M. Balter, “3D forward and back-projection for X-ray CT using separable footprints.,” *IEEE Trans. Med. Imaging*, vol. 29, no. 11, pp. 1839–50, Nov. 2010, doi: 10.1109/TMI.2010.2050898.
- [59] K. Sauer and C. Bouman, “A local update strategy for iterative reconstruction from projections,” *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 534–548, 1993, doi: 10.1109/78.193196.
- [60] K. Lange and J. a Fessler, “Globally convergent algorithms for maximum a posteriori transmission tomography.,” *IEEE Trans. Image Process.*, vol. 4, no. 10, pp. 1430–8, Jan. 1995, doi: 10.1109/83.465107.
- [61] I. A. Elbakri and J. A. Fessler, “Statistical image reconstruction for polyenergetic X-ray computed tomography,” *IEEE Trans. Med. Imaging*, vol. 21, no. 2, pp. 828–831, 2002, doi: 10.1109/ISBI.2002.1029387.
- [62] I. A. Elbakri and J. A. Fessler, “Segmentation-free statistical image reconstruction

- for polyenergetic x-ray computed tomography with experimental validation,” *Phys. Med. Biol.*, vol. 48, no. 15, pp. 2453–2477, 2003, doi: 10.1088/0031-9155/48/15/314.
- [63] S. Tilley, J. H. Siewerdsen, and J. W. Stayman, “Model-based iterative reconstruction for flat-panel cone-beam CT with focal spot blur, detector blur, and correlated noise,” *Phys. Med. Biol.*, vol. 61, no. 1, pp. 296–319, 2015, doi: 10.1088/0031-9155/61/1/296.
- [64] S. Tilley *et al.*, “Penalized-likelihood reconstruction with high-fidelity measurement models for high-resolution cone-beam imaging,” *IEEE Trans. Med. Imaging*, vol. 37, no. 4, pp. 988–999, 2018, doi: 10.1109/TMI.2017.2779406.
- [65] S. Tilley II, “High-quality computed tomography using advanced model-based iterative reconstruction,” Johns Hopkins University, 2019.
- [66] S. A. Hardcastle, P. Dieppe, C. L. Gregson, G. Davey Smith, and J. H. Tobias, “Osteoarthritis and bone mineral density: are strong bones bad for joints?,” *Bonekey Rep.*, vol. 4, no. January, pp. 1–8, 2015, doi: 10.1038/bonekey.2014.119.
- [67] J. Richards *et al.*, “Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study,” *Lancet*, vol. 371, no. 9623, pp. 1505–1512, 2008, doi: 10.1016/S0140-6736(08)60599-1.
- [68] P. J. Pickhardt, G. Bodeen, A. Brett, J. K. Brown, and N. Binkley, “Comparison of femoral neck BMD evaluation obtained using lunar DXA and QCT with asynchronous calibration from CT colonography,” *J. Clin. Densitom.*, vol. 18, no. 1, pp. 5–12, 2015, doi: 10.1016/j.jocd.2014.03.002.
- [69] E. M. Lewiecki and N. E. Lane, “Common mistakes in the clinical use of bone

- mineral density testing,” *Nat. Clin. Pract. Rheumatol.*, vol. 4, no. 12, pp. 667–674, 2008, doi: 10.1038/ncprheum0928.
- [70] L. Oei, F. Koromani, F. Rivadeneira, M. C. Zillikens, and E. H. G. Oei, “Quantitative imaging methods in osteoporosis,” *Quant. Imaging Med. Surg.*, vol. 6, no. 6, pp. 680–698, 2016, doi: 10.21037/qims.2016.12.13.
- [71] J. S. Bauer and T. M. Link, “Advances in osteoporosis imaging,” *Eur. J. Radiol.*, vol. 71, no. 3, pp. 440–9, 2009.
- [72] T. Hildebrand and P. Ruegsegger, “A new method for the model-independent assessment of thickness in three-dimensional images,” *J. Microsc.*, vol. 185, no. 1, pp. 67–75, Jan. 1997, doi: 10.1046/j.1365-2818.1997.1340694.x.
- [73] E. Klintström, Ö. Smedby, R. Moreno, and T. B. Brismar, “Trabecular bone structure parameters from 3D image processing of clinical multi-slice and cone-beam computed tomography data,” *Skeletal Radiol.*, vol. 43, no. 2, pp. 197–204, 2014, doi: 10.1007/s00256-013-1766-5.
- [74] L. Mulder, B. Van Rietbergen, N. J. Noordhoek, and K. Ito, “Determination of vertebral and femoral trabecular morphology and stiffness using a flat-panel C-arm-based CT approach,” *Bone*, vol. 50, no. 1, pp. 200–208, 2012, doi: 10.1016/j.bone.2011.10.020.
- [75] T. Sokka, “Radiographic Scoring in Rheumatoid Arthritis,” *Bull. NYU Hosp. Jt. Dis.*, vol. 66, no. 2, pp. 166–168, 2008.
- [76] J. T. Sharp *et al.*, “Measurement of joint space width and erosion size.,” *J. Rheumatol.*, vol. 32, no. 12, pp. 2456–61, Dec. 2005.
- [77] B. Kalinosky, J. M. Sabol, K. Piacsek, B. Heckel, and T. Gilat Schmidt,

- “Quantifying the tibiofemoral joint space using x-ray tomosynthesis.,” *Med. Phys.*, vol. 38, no. 12, pp. 6672–82, Dec. 2011, doi: 10.1118/1.3662891.
- [78] A. J. Yezzi and J. L. Prince, “An Eulerian PDE approach for computing tissue thickness.,” *IEEE Trans. Med. Imaging*, vol. 22, no. 10, pp. 1332–9, Oct. 2003, doi: 10.1109/TMI.2003.817775.
 - [79] J. D. Jackson, “Classical Electrodynamics,” 3rd ed., Wiley, 1998, pp. 57–135.
 - [80] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, “Numerical Recipes in C,” 2nd ed., Cambridge University Press, 1992, pp. 836–871.
 - [81] S. R. Sternberg, “Grayscale Morphology,” *Comput. Vision, Graph. Image Process.*, vol. 35, no. 3, pp. 333–355, 1986, doi: DOI: 10.1016/0734-189X(86)90004-6.
 - [82] C. K. Chow and T. Kaneko, “Automatic Boundary from Detection of the Cineangiograms ” Left Ventricle,” *Comput. Biomed. Res.*, vol. 5, no. 4, pp. 388–410, 1972.
 - [83] A. Myronenko and X. Song, “Point set registration: coherent point drift.,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–75, Dec. 2010, doi: 10.1109/TPAMI.2010.46.
 - [84] G. Chintalapani, L. M. Ellingsen, O. Sadowsky, J. L. Prince, and R. H. Taylor, “Statistical atlases of bone anatomy: construction, iterative improvement and validation.,” *Med. Image Comput. Comput. Assist. Interv.*, vol. 10, no. Pt 1, pp. 499–506, Jan. 2007.
 - [85] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

- [86] M. van der Esch *et al.*, “Osteoarthritis of the knee: multicompartimental or compartmental disease?,” *Rheumatology (Oxford)*., vol. 53, no. 3, pp. 540–6, Mar. 2014, doi: 10.1093/rheumatology/ket393.
- [87] C. E. Metz, B. A. Herman, and J. Shen, “Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously distributed data,” *Stat. Med.*, vol. 17, no. 1998, pp. 1033–1053, 1998.
- [88] A. Yezzi, Z. Lilla, and T. Kapur, “A variational framework for joint segmentation and registration,” in *Mathematical Methods in Biomedical Image Analysis, 2001. MMBIA 2001. IEEE Workshop on*, 2001, pp. 44–51, doi: 10.1109/MMBIA.2001.991698.
- [89] J. Schmid, J. Kim, and N. Magnenat-Thalmann, “Robust statistical shape models for MRI bone segmentation in presence of small field of view.,” *Med. Image Anal.*, vol. 15, no. 1, pp. 155–68, Feb. 2011, doi: 10.1016/j.media.2010.09.001.
- [90] A. A. Muhit *et al.*, “Peripheral quantitative CT (pQCT) using a dedicated extremity cone-beam CT scanner,” in *Proc. SPIE*, 2013, vol. 8672, p. 867203, doi: 10.1117/12.2006939.
- [91] B. Van der Heyden *et al.*, “Modelling of the focal spot intensity distribution and the off-focal spot radiation in kilovoltage X-ray tubes for imaging,” *Phys. Med. Biol.*, vol. accepted, no. December 2016, 2019.
- [92] G. Poludniowski, P. M. Evans, A. Kavanagh, and S. Webb, “Removal and effects of scatter-glare in cone-beam CT with an amorphous-silicon flat-panel detector,” *Phys. Med. Biol.*, vol. 56, no. 6, pp. 1837–1851, 2011, doi: 10.1088/0031-9155/56/6/019.

- [93] A. Sisniega *et al.*, “High-fidelity artifact correction for cone-beam CT imaging of the brain,” *Phys. Med. Biol.*, vol. 60, no. 4, p. 1415, 2015, doi: 10.1088/0031-9155/60/4/1415.
- [94] J. M. S. Wait, D. Cody, A. K. Jones, J. Rong, V. Baladandayuthapani, and S. C. Kappadath, “Performance evaluation of material decomposition with rapid-kilovoltage-switching dual-energy CT and implications for assessing bone mineral density,” *Am. J. Roentgenol.*, vol. 204, no. 6, pp. 1234–1241, 2015, doi: 10.2214/AJR.14.13093.
- [95] K. Taguchi, M. Zhang, E. C. Frey, J. Xu, W. P. Segars, and B. M. W. Tsui, “Image-domain material decomposition using photon-counting CT,” *Proc. SPIE*, vol. 6510, no. March 2007, p. 651008, 2007, doi: 10.1117/12.713508.
- [96] M. Lu *et al.*, “Dual energy imaging with a dual-layer flat panel detector,” in *Proc. of SPIE*, 2019, vol. 1094815, no. March, p. 40, doi: 10.1117/12.2513499.
- [97] T. R. C. Johnson *et al.*, “Material differentiation by dual energy CT: Initial experience,” *Eur. Radiol.*, vol. 17, no. 6, pp. 1510–1517, 2007, doi: 10.1007/s00330-006-0517-6.
- [98] R. E. Alvarez and A. MacOvski, “Energy-selective reconstructions in X-ray computerised tomography,” *Phys. Med. Biol.*, vol. 21, no. 5, pp. 733–744, 1976, doi: 10.1088/0031-9155/21/5/002.
- [99] P. Joseph and R. Spital, “A method for correcting bone induced artifacts in computed tomography,” *J. Comput. Assist. Tomogr.*, vol. 2, no. 1, pp. 100–8, 1978.
- [100] J. Hsieh, B. Nett, Z. Yu, K. Sauer, J.-B. Thibault, and C. Bouman, “Recent

- advances in CT image reconstruction,” *Curr. Radiol. Rep.*, vol. 1, pp. 39–51, 2013, doi: 10.1007/s40134-012-0003-7.
- [101] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, “An iterative maximum-likelihood polychromatic algorithm for CT,” *IEEE Trans. Med. Imaging*, vol. 20, no. 10, pp. 999–1008, 2001, doi: 10.1109/42.959297.
- [102] J. H. Mason, A. Perelli, W. H. Nailon, and M. E. Davies, “Polyquant CT: Direct electron and mass density reconstruction from a single polyenergetic source,” *Phys. Med. Biol.*, vol. 62, no. 22, pp. 8739–8762, 2017, doi: 10.1088/1361-6560/aa9162.
- [103] J. H. Mason, A. Perelli, W. H. Nailon, and M. E. Davies, “Quantitative cone-beam CT reconstruction with polyenergetic scatter model fusion,” *Phys. Med. Biol.*, vol. 63, no. 22, 2018, doi: 10.1088/1361-6560/aae794.
- [104] D. C. Sullivan *et al.*, “Metrology standards for quantitative imaging biomarkers,” *Radiology*, vol. 277, no. 3, pp. 856–866, 2015, doi: 10.1021/ma301990u.
- [105] C. C. Glüer, G. Blake, Y. Lu, B. A. Blunt1, M. Jergas1, and H. K. Genant1, “Accurate assessment of precision errors: How to measure the reproducibility of bone densitometry techniques,” *Osteoporos. Int.*, vol. 5, no. 4, pp. 262–270, 1995, doi: 10.1007/BF01774016.
- [106] L. Wang *et al.*, “Validation of asynchronous quantitative bone densitometry of the spine: Accuracy, short-term reproducibility, and a comparison with conventional quantitative computed tomography,” *Sci. Rep.*, vol. 7, no. 1, pp. 1–7, 2017, doi: 10.1038/s41598-017-06608-y.
- [107] Gammex, “Tissue Mimicking Materials (TMM) Nominal Characteristics.” 2016.

- [108] H. Zhou, P. J. Keall, and E. E. Graves, “A bone composition model for Monte Carlo x-ray transport simulations,” *Med. Phys.*, vol. 36, no. 3, pp. 1008–1018, 2009, doi: 10.1118/1.3077129.
- [109] E. Samei, M. J. Flynn, and D. A. Reimann, “A method for measuring the presampled MTF of digital radiographic systems using an edge test device.,” *Med. Phys.*, vol. 25, no. 1, pp. 102–113, 1998, doi: 10.1118/1.598165.
- [110] A. R. De Pierro, “A Modified Expectation Maximization Algorithm for Penalized Likelihood Estimation in Emission Tomography,” *IEEE Trans. Med. Imaging*, vol. 14, no. 1, pp. 132–137, 1995.
- [111] H. Erdogan and J. A. Fessler, “Ordered subsets algorithms for transmission tomography,” *Phys. Med. Biol.*, vol. 2835, pp. 2835–2851, 1999.
- [112] P. S. Tofts, “Definitions of effective energy in computed tomography,” *Phys. Med. Biol.*, vol. 26, no. 2, pp. 313–317, 1981, doi: 10.1088/0031-9155/26/2/010.
- [113] Y. Nesterov, “Smooth minimization of non-smooth functions,” *Math. Program.*, vol. 103, no. 1, pp. 127–152, 2005, doi: 10.1007/s10107-004-0552-5.
- [114] A. S. Wang, J. W. Stayman, Y. Otake, S. Vogt, G. Kleinszig, and J. H. Siewerssen, “Accelerated statistical reconstruction for C-arm cone-beam CT using Nesterov’s method,” *Med. Phys.*, vol. 42, no. 5, pp. 2699–2708, 2015, doi: 10.1118/1.4914378.
- [115] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. Pluim, “elastix: A Toolbox for Intensity-Based Medical Image Registration,” *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 196–205, 2010, doi: 10.1109/TMI.2009.2035616.
- [116] Q. Cao *et al.*, “Cone-beam CT of load-bearing surgical hardware using a

- mechanical model of implant deformation,” in *AAPM Annual Meeting*, 2019.
- [117] E. Marinetto *et al.*, “Quantification of bone microarchitecture in ultra-high resolution extremities cone-beam CT with a CMOS detector and compensation of patient motion,” in *Computer Assisted Radiology and Surgery: 30th International Congress and Exhibition*, 2016.
- [118] Q. Cao *et al.*, “High-resolution cone-beam CT of the extremities and cancellous bone architecture with a CMOS detector,” in *AAPM Annual Meeting*, 2016, p. 3797.
- [119] M. Brehler *et al.*, “Quantitative assessment of trabecular bone microarchitecture using high-resolution extremities cone-beam CT,” in *RSNA Scientific Assembly and Annual Meeting*, 2016.
- [120] Q. Cao *et al.*, “High-resolution extremity cone-beam CT with a CMOS detector: Task-based optimization of scintillator thickness,” in *Proc. of SPIE*, 2017, vol. 10132, p. 1013210, doi: 10.1117/12.2255695.
- [121] S. Majumdar *et al.*, “Correlation of trabecular bone structure with age, bone mineral density, and osteoporotic status: in vivo studies in the distal radius using high resolution magnetic resonance imaging.,” *J. Bone Miner. Res.*, vol. 12, no. 1, pp. 111–118, 1997, doi: 10.1359/jbmr.1997.12.1.111.
- [122] M. A. Bredella *et al.*, “Distal radius in adolescent girls with anorexia nervosa: trabecular structure analysis with high-resolution flat-panel volume CT,” *Radiology*, vol. 249, no. 3, pp. 938–946, 2008, doi: 10.1148/radiol.2492080173.
- [123] A. S. Issever *et al.*, “Assessment of trabecular bone structure using MDCT: Comparison of 64- and 320-slice CT using HR-pQCT as the reference standard,”

Eur. Radiol., vol. 20, no. 2, pp. 458–468, 2010, doi: 10.1007/s00330-009-1571-7.

- [124] C. J. Walsh *et al.*, “Women with anorexia nervosa: finite element and trabecular structure analysis by using flat-panel volume CT,” *Radiology*, vol. 257, no. 1, pp. 167–174, 2010, doi: 10.1148/radiol.10100222.
- [125] A. J. Burghardt, T. M. Link, and S. Majumdar, “High-resolution computed tomography for clinical imaging of bone microarchitecture,” *Clin. Orthop. Relat. Res.*, vol. 469, no. 8, pp. 2179–2193, 2011, doi: 10.1007/s11999-010-1766-x.
- [126] J. Van Dessel, Y. Huang, M. Depypere, I. Rubira-Bullen, F. Maes, and R. Jacobs, “A comparative evaluation of cone beam CT and micro-CT on trabecular bone structures in the human mandible,” *Dentomaxillofacial Radiol.*, vol. 42, no. 8, 2013, doi: 10.1259/dmfr.20130145.
- [127] N. C. Harvey *et al.*, “Trabecular bone score (TBS) as a new complementary approach for osteoporosis evaluation in clinical practice,” *Bone*, vol. 78, pp. 216–224, 2015, doi: 10.1016/j.bone.2015.05.016.
- [128] C. Buckland-Wright, “Subchondral bone changes in hand and knee osteoarthritis detected by radiography,” *Osteoarthr. Cartil.*, vol. 12, no. SUPPL., pp. 10–19, 2004, doi: 10.1016/j.joca.2003.09.007.
- [129] H. Weinans, M. Siebelt, R. Agricola, S. M. Botter, T. M. Piscaer, and J. H. Waarsing, “Pathophysiology of peri-articular bone changes in osteoarthritis.,” *Bone*, vol. 51, no. 2, pp. 190–6, Aug. 2012, doi: 10.1016/j.bone.2012.02.002.
- [130] C. Chappard *et al.*, “Subchondral bone micro-architectural alterations in osteoarthritis: A synchrotron micro-computed tomography study,” *Osteoarthr. Cartil.*, vol. 14, pp. 215–223, 2006, doi: 10.1016/j.joca.2005.09.008.

- [131] J. E. Adams, “Advances in bone imaging for osteoporosis,” *Nat. Rev. Endocrinol.*, vol. 9, no. 1, pp. 28–42, 2012, doi: 10.1038/nrendo.2012.217.
- [132] Y. Jiang, J. Zhao, E. Y. Liao, R. C. Dai, X. P. Wu, and H. K. Genant, “Application of micro-ct assessment of 3-d bone microstructure in preclinical and clinical studies,” *J. Bone Miner. Metab.*, vol. 23, no. SUPPL. 1, pp. 122–131, 2005, doi: 10.1007/BF03026336.
- [133] S. Boutroy, M. L. Bouxsein, F. Munoz, and P. D. Delmas, “In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography,” *J. Clin. Endocrinol. Metab.*, vol. 90, no. 12, pp. 6508–6515, 2005, doi: 10.1210/jc.2005-1258.
- [134] J. A. Macneil and S. K. Boyd, “Accuracy of high-resolution peripheral quantitative computed tomography for measurement of bone quality,” vol. 29, pp. 1096–1105, 2007, doi: 10.1016/j.medengphy.2006.11.002.
- [135] A. Kroker, Y. Zhu, S. L. Manske, R. Barber, N. Mohtadi, and S. K. Boyd, “Quantitative in vivo assessment of bone microarchitecture in the human knee using HR-pQCT,” *Bone*, vol. 97, pp. 43–48, 2016, doi: 10.1016/j.bone.2016.12.015.
- [136] A. C. Konstantinidis, M. B. Szafraniec, R. D. Speller, and A. Olivo, “The Dexela 2923 CMOS X-ray detector: A flat panel detector based on CMOS active pixel sensors for medical imaging applications,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 689, pp. 12–21, Oct. 2012, doi: 10.1016/j.nima.2012.06.024.
- [137] M. Esposito *et al.*, “DynAMITe: a wafer scale sensor for biomedical applications,”

J. Instrum., vol. 6, no. 12, pp. C12064–C12064, 2011, doi: 10.1088/1748-0221/6/12/C12064.

- [138] Y. Shen, Y. Zhong, C.-J. Lai, T. Wang, and C. C. Shaw, “Cone beam breast CT with a high pitch (75 μm), thick (500 μm) scintillator CMOS flat panel detector: Visibility of simulated microcalcifications.,” *Med. Phys.*, vol. 40, no. 10, p. 101915, 2013, doi: 10.1118/1.4820440.
- [139] J.-G. Choi, H.-S. Park, Y. Kim, Y.-W. Choi, T.-H. Ham, and H.-J. Kim, “Characterization of prototype full-field breast tomosynthesis by using a CMOS array coupled with a columnar CsI(Tl) scintillator,” *J. Korean Phys. Soc.*, vol. 60, no. 3, pp. 521–526, 2012, doi: 10.3938/jkps.60.521.
- [140] T. Patel, K. Klanian, Z. Gong, and M. B. Williams, “Detective quantum efficiency of a CsI-CMOS x-ray detector for breast tomosynthesis operating in high dynamic range and high sensitivity modes,” in *International Workshop on Digital Mammography*, 2012, pp. 80–87.
- [141] P. M. Gazi, K. Yang, G. W. Burkett, S. Aminololama-Shakeri, J. Anthony Seibert, and J. M. Boone, “Evolution of spatial resolution in breast CT at UC Davis.,” *Med. Phys.*, vol. 42, no. 4, p. 1973, 2015, doi: 10.1118/1.4915079.
- [142] C. Zhao, J. Kanicki, A. C. Konstantinidis, and T. Patel, “Large area CMOS active pixel sensor x-ray imager for digital breast tomosynthesis: Analysis, modeling, and characterization,” *Med. Phys.*, vol. 42, no. 11, pp. 6294–6308, 2015, doi: 10.1118/1.4932368.
- [143] C. Zhao and J. Kanicki, “Task-based modeling of an 5k ultra-high-resolution medical imaging system for digital breast tomosynthesis,” *IEEE Trans. Med.*

- Imaging*, vol. PP, no. 99, pp. 1–1, 2017, doi: 10.1109/TMI.2017.2695982.
- [144] S. Vedantham, S. Shrestha, L. Shi, G. Vijayaraghavan, and A. Karella, “Task-specific optimization of scintillator thickness for CMOS-detector based cone-beam breast CT,” in *AAPM Annual Meeting*, 2016, vol. 43, no. 6, p. 3346, doi: 10.1118/1.4955660.
- [145] J. H. Siewerdsen *et al.*, “Empirical and theoretical investigation of the noise performance of indirect detection, active matrix flat-panel imagers (AMFPIs) for diagnostic radiology,” *Med. Phys.*, vol. 24, no. 1, p. 71, 1997, doi: 10.1118/1.597919.
- [146] G. J. Gang *et al.*, “Analysis of Fourier-domain task-based detectability index in tomosynthesis and cone-beam CT in relation to human observer performance.,” *Med. Phys.*, vol. 38, no. 4, pp. 1754–1768, 2011, doi: 10.1118/1.3560428.
- [147] D. J. Tward and J. H. Siewerdsen, “Cascaded systems analysis of the 3D noise transfer characteristics of flat-panel cone-beam CT,” *Med. Phys.*, vol. 35, no. 12, p. 5510, 2008, doi: 10.1118/1.3002414.
- [148] D. J. Tward and J. H. Siewerdsen, “Noise aliasing and the 3D NEQ of flat-panel cone-beam CT: effect of 2D/3D apertures and sampling.,” *Med. Phys.*, vol. 36, no. 8, pp. 3830–3843, 2009, doi: 10.1118/1.3166933.
- [149] I. A. Cunningham, J. Yao, and V. Subotic, “Cascaded models and the DQE of flat-panel imagers: noise aliasing, secondary quantum noise and reabsorption,” in *Proc. SPIE*, 2002, vol. 4682, pp. 61–72, doi: 10.1117/12.465610.
- [150] S. Vedantham, A. Karella, and S. Suryanarayanan, “Solid-state fluoroscopic imager for high-resolution angiography: Parallel-cascaded linear systems

- analysis,” *Med. Phys.*, vol. 31, no. 5, pp. 1258–1268, 2004, doi: 10.1016/j.biotechadv.2011.08.021.Secreted.
- [151] W. Zbijewski *et al.*, “A dedicated cone-beam CT system for musculoskeletal extremities imaging: design, optimization, and initial performance characterization,” *Med. Phys.*, vol. 38, no. 8, p. 4700, 2011, doi: 10.1118/1.3611039.
- [152] P. Prakash *et al.*, “Task-based modeling and optimization of a cone-beam CT scanner for musculoskeletal imaging.,” *Med. Phys.*, vol. 38, no. 10, pp. 5612–29, Oct. 2011, doi: 10.1118/1.3633937.
- [153] C. Zhao, A. C. Konstantinidis, Y. Zheng, and T. Anaxagoras, “50 μm pixel pitch wafer-scale CMOS active pixel sensor x-ray detector for digital breast tomosynthesis,” *Phys. Med. Biol.*, vol. 60, p. 8977, 2015, doi: 10.1088/0031-9155/60/23/8977.
- [154] J. Yao and I. A. Cunningham, “Parallel cascades: new ways to describe noise transfer in medical imaging systems.,” *Med. Phys.*, vol. 28, no. 10, pp. 2020–2038, 2001, doi: 10.1118/1.1405842.
- [155] J. H. Siewerdsen and D. A. Jaffray, “Optimization of x-ray imaging geometry (with specific application to flat-panel cone-beam computed tomography).,” *Med. Phys.*, vol. 27, no. 8, pp. 1903–1914, 2000, doi: 10.1118/1.1286590.
- [156] C. E. Metz and C. J. Vyborny, “Wiener spectral effects of spatial correlation between the sites of characteristic x-ray emission and reabsorption in radiographic screen-film systems.,” *Phys. Med. Biol.*, vol. 28, no. 5, pp. 547–564, 1983, doi: 10.1088/0031-9155/28/5/008.

- [157] A. Howansky, B. Peng, A. R. Lubinsky, and W. Zhao, “Deriving depth-dependent light escape efficiency and optical Swank factor from measured pulse height spectra of scintillators,” *Med. Phys.*, no. November, pp. 1–14, 2017, doi: 10.1002/mp.12083.
- [158] H. P. Chan and K. Doi, “Energy and angular dependence of x-ray absorption and its effect on radiographic response in screen--film systems.,” *Phys. Med. Biol.*, vol. 28, no. 5, pp. 565–79, 1983, doi: 10.1088/0031-9155/28/5/009.
- [159] W. Que and J. Rowlands, “X-ray imaging using amorphous selenium: Inherent spatial resolution,” *Med. Phys.*, vol. 22, no. 4, p. 365, 1995, doi: 10.1118/1.597471.
- [160] H. Barrett and W. Swindell, *Radiological Imaging: The Theory of Image Formation, Detection, and Processing*. New York: Academic, 1981.
- [161] S. Richard, “Optimization of imaging performance and conspicuity in dual-energy x-ray radiography,” University of Toronto, 2008.
- [162] G. Lubberts, “Random noise produced by x-Ray fluorescent screens,” *J. Opt. Soc. Am.*, vol. 58, no. 11, p. 1475, 1968, doi: 10.1364/JOSA.58.001475.
- [163] A. Badano, R. M. Gagne, B. D. Gallas, R. J. Jennings, J. S. Boswell, and K. J. Myers, “Lubberts effect in columnar phosphors.,” *Med. Phys.*, vol. 31, no. 11, pp. 3122–3131, 2004, doi: 10.1118/1.1796151.
- [164] A. Howansky, A. Lubinsky, S. Ghose, K. Suzuki, and W. Zhao, “Direct measurement of Lubberts effect in CsI : Tl scintillators using single x-ray photon imaging,” in *Proc. SPIE*, 2017, vol. 10132, p. 1013209, doi: 10.1117/12.2255561.
- [165] W. Zbijewski *et al.*, “Design and optimization of a dedicated cone-beam CT

- system for musculoskeletal extremities imaging,” *Med. Phys.*, vol. 38, no. 8, p. 4700, 2011, doi: 10.1118/1.3611039.
- [166] Teledyne DALSA, “Xineos-3030HR Specifications.” pp. 8–11, 2015.
- [167] M. D. Silver, “A method for including redundant data in computed tomography.,” *Med. Phys.*, vol. 27, no. 4, pp. 773–774, Apr. 2000, doi: 10.1118/1.598939.
- [168] J. E. Adams, “Advances in bone imaging for osteoporosis,” *Nat. Rev. Endocrinol.*, vol. 9, no. 1, pp. 28–42, 2013, doi: 10.1038/nrendo.2012.217.
- [169] W. H. Maes, I. M. Peters, C. Smit, Y. Kessener, and J. Bosiers, “Low-dose performance of wafer-scale CMOS-based X-ray detectors,” in *Proc. SPIE*, 2015, vol. 9412, p. 94120C, doi: 10.1117/12.2081996.
- [170] A. Badano and J. Sempau, “MANTIS: combined x-ray, electron and optical Monte Carlo simulations of indirect radiation imaging systems.,” *Phys. Med. Biol.*, vol. 51, pp. 1545–1561, 2006, doi: 10.1088/0031-9155/51/6/013.
- [171] M. Freed, S. Miller, K. Tang, and A. Badano, “Experimental validation of Monte Carlo (MANTIS) simulated x-ray response of columnar CsI scintillator screens,” *Med. Phys.*, vol. 36, no. 11, pp. 4944–4956, 2009, doi: 10.1118/1.3233683.
- [172] M. L. Bouxsein, S. K. Boyd, B. A. Christiansen, R. E. Guldberg, K. J. Jepsen, and R. Muller, “Guidelines for assessment of bone microstructure in rodents using micro-computed tomography.,” *J. Bone Miner. Res.*, vol. 25, no. 7, pp. 1468–1486, Jul. 2010, doi: 10.1002/jbmr.141.
- [173] M. A. Kupinski, E. Clarkson, K. Gross, and J. W. Hoppin, “Optimizing imaging hardware for estimation tasks,” in *Proc. SPIE*, 2003, vol. 5034, pp. 309–313, doi: 10.1117/12.480337.

- [174] A. Valentinitisch *et al.*, “Computational identification and quantification of trabecular microarchitecture classes by 3-D texture analysis-based clustering,” *Bone*, vol. 54, no. 1, pp. 133–140, 2013, doi: 10.1016/j.bone.2012.12.047.
- [175] A. Valentinitisch *et al.*, “Trabecular bone class mapping across resolutions: translating methods from HR-pQCT to clinical CT,” in *Proc. SPIE*, 2015, p. 94131D, doi: 10.1117/12.2081187.
- [176] B. Paniagua, A. C. Ruellas, E. Benavides, S. Marron, L. Wolford, and L. Cividanes, “Validation of CBCT for the computation of textural biomarkers,” in *Proc. SPIE*, 2015, vol. 9417, p. 94141B, doi: 10.1117/12.2081859.
- [177] J. W. MacKay *et al.*, “Quantitative analysis of tibial subchondral bone: Texture analysis outperforms conventional trabecular microarchitecture analysis,” *J. Magn. Reson. Imaging*, vol. 43, no. 5, pp. 1159–1170, 2016, doi: 10.1002/jmri.25088.
- [178] R. M. Haralick, “Statistical and structural approaches to texture,” *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, 1979, doi: 10.1109/PROC.1979.11328.
- [179] W. Zbijewski *et al.*, “Dose and scatter characteristics of a novel cone beam CT system for musculoskeletal extremities,” in *Proc. SPIE*, 2012, vol. 8313, p. 831318, doi: 10.1117/12.911843.
- [180] A. Sisniega *et al.*, “Monte Carlo study of the effects of system geometry and antiscatter grids on cone-beam CT scatter distributions,” *Med. Phys.*, vol. 40, no. 5, p. 051915, 2013, doi: 10.1118/1.4801895.
- [181] S. Tilley II, W. Zbijewski, and J. W. Stayman, “High-Fidelity Modeling of Shift-Variant Focal-Spot Blur for High-Resolution CT,” *Proc. Int. Meet. Fully Three-Dimensional Image Reconstr. Radiol. Nucl. Med.*, pp. 752–759, 2017.

- [182] A. Sisniega, J. W. Stayman, Q. Cao, J. Yorkston, J. H. Siewerdsen, and W. Zbijewski, “Image-based motion compensation for high-resolution extremities cone-beam CT,” in *Proc. SPIE*, 2016, vol. 9783, p. 97830K, doi: 10.1117/12.2217243.
- [183] A. Sisniega, J. Stayman, J. Yorkston, J. Siewerdsen, and W. Zbijewski, “Motion compensation in extremity cone-beam CT using a penalized image sharpness criterion,” *Phys. Med. Biol.*, vol. 62, pp. 3712–3734, 2017, doi: 10.1088/1361-6560/aa6869.
- [184] S. Tilley II, J. H. Siewerdsen, and J. W. Stayman, “Model-based iterative reconstruction for flat-panel cone-beam CT with focal spot blur, detector blur, and correlated noise,” *Phys. Med. Biol.*, vol. 61, pp. 296–319, 2016, doi: 10.1088/0031-9155/61/1/296.
- [185] B. Keck and H. Hofmann, “GPU-accelerated SART reconstruction using the CUDA programming environment,” in *Proc. SPIE Medical Imaging*, 2009, p. 72582B, doi: 10.1117/12.811559.
- [186] J. Brokish, P. Sack, and Y. Bresler, “Combined algorithmic and GPU acceleration for ultra-fast circular conebeam backprojection,” in *Proc. SPIE Medical Imaging* 2010, 2010, p. 7622:762256, doi: 10.1117/12.844028.
- [187] E. Papenhausen, Z. Zheng, and K. Mueller, “GPU-accelerated back-projection revisited : squeezing performance by careful tuning,” in *International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2011, pp. 19–22.
- [188] G. Pratx and L. Xing, “GPU computing in medical physics: a review.,” *Med.*

Phys., vol. 38, no. 5, pp. 2685–2697, 2011, doi: 10.1118/1.3578605.

- [189] X. Liu, S. Boons, and A. Sasov, “Strategies for GPU-based cone-beam CT reconstruction for very large data volumes,” in *International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 529–532.
- [190] T. Frese, C. A. Bouman, and K. Sauer, “Adaptive wavelet graph model for Bayesian tomographic reconstruction.,” *IEEE Trans. Image Process.*, vol. 11, no. 7, pp. 756–70, 2002, doi: 10.1109/TIP.2002.801586.
- [191] S. Degirmenci *et al.*, “Acceleration of iterative image reconstruction for X-ray imaging for security applications,” in *Proc. SPIE Computational Imaging XIII*, 2015, p. 94010C, doi: 10.1117/12.2082966.
- [192] J. G. Brankov, Y. Yang, and M. N. Wernick, “Tomographic Image Reconstruction Based on a Content-Adaptive Mesh Model,” *IEEE Trans. Med. Imaging*, vol. 23, no. 2, pp. 202–212, 2004, doi: 10.1109/TMI.2003.822822.
- [193] A. Sitek, R. H. Huesman, and G. T. Gullberg, “Tomographic reconstruction using an adaptive tetrahedral mesh defined by a point cloud,” *IEEE Trans. Med. Imaging*, vol. 25, no. 9, pp. 1172–1179, 2006, doi: 10.1109/TMI.2006.879319.
- [194] A. Ziegler, T. Nielsen, and M. Grass, “Iterative reconstruction of a region of interest for transmission tomography.,” *Med. Phys.*, vol. 35, no. 4, pp. 1317–1327, 2008, doi: 10.1118/1.2870219.
- [195] C. W. Stearns, R. M. Manjeshwar, and S. D. Wollenweber, “An efficient algorithm for targeted reconstruction of tomographic data,” *IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 5, pp. 2808–2811, 2007, doi: 10.1109/NSSMIC.2006.356462.

- [196] E. A. Rashed, H. Kudo, T. Zeniya, and H. Iida, “Practical statistical models for region-of-interest tomographic reconstruction and long object problem,” *IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 5, pp. 3505–3511, 2007, doi: 10.1109/NSSMIC.2007.4436885.
- [197] B. Hamelin, Y. Goussard, J.-P. Dussault, G. Cloutier, G. Beaudoin, and G. Soulez, “Design of iterative ROI transmission tomography reconstruction procedures and image quality analysis.,” *Med. Phys.*, vol. 37, no. 9, pp. 4577–4589, 2010, doi: 10.1118/1.3447722.
- [198] P. J. La Rivière, “Monotonic iterative reconstruction algorithms for targeted reconstruction in emission and transmission computed tomography,” *IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 5, no. 2, pp. 2924–2928, 2007, doi: 10.1109/NSSMIC.2006.356488.
- [199] B. Hamelin, Y. Goussard, and J. P. Dussault, “Penalized-likelihood region-of-interest CT reconstruction by local object supersampling,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2007, no. 4, pp. 739–742, doi: 10.1109/IEMBS.2007.4352396.
- [200] Z. Yu, J.-B. Thibault, C. Bouman, K. Sauer, and J. Hsieh, “Edge-localized iterative reconstruction for computed tomography,” *10th Intl. Mtg. fully 3D image Recon. Rad. Nucl. Med.*, no. 1, pp. 255–258, 2009.
- [201] K. Van Slambrouck and J. Nuyts, “A patchwork (back)projector to accelerate artifact reduction in CT reconstruction,” *IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, no. 1, pp. 2625–2629, Oct. 2010.
- [202] K. Van Slambrouck and J. Nuyts, “Metal artifact reduction in computed

- tomography using local models in an image block-iterative scheme.,” *Med. Phys.*, vol. 39, no. 11, pp. 7080–93, Nov. 2012, doi: 10.1118/1.4762567.
- [203] H. Erdoğan and J. A. Fessler, “Ordered subsets algorithms for transmission tomography.,” *Phys. Med. Biol.*, vol. 44, no. 11, pp. 2835–51, Nov. 1999.
- [204] J. A. Fessler, “Grouped Coordinate Descent Algorithms for Robust Edge-Preserving Image Restoration,” in *Proc. SPIE Image Reconstruction and Restoration*, 1997, vol. 3170, no. 3, pp. 184–194, doi: 10.1117/12.279713.
- [205] J. W. Stayman, Y. Otake, J. L. Prince, a J. Khanna, and J. H. Siewerdsen, “Model-based tomographic reconstruction of objects containing known components.,” *IEEE Trans. Med. Imaging*, vol. 31, no. 10, pp. 1837–48, Oct. 2012, doi: 10.1109/TMI.2012.2199763.
- [206] H. Erdoğan and J. a Fessler, “Monotonic algorithms for transmission tomography.,” *IEEE Trans. Med. Imaging*, vol. 18, no. 9, pp. 801–14, Sep. 1999, doi: 10.1109/42.802758.
- [207] J. H. Siewerdsen, A. M. Waese, D. J. Moseley, S. Richard, and D. a Jaffray, “Spektr: A computational tool for x-ray spectral analysis and imaging system optimization,” *Med. Phys.*, vol. 31, no. 11, pp. 3057–3067, 2004, doi: 10.1118/1.1758350.
- [208] Z. Yu, C. A. Bouman, J. Thibault, and K. D. Sauer, “Image Grid Invariant Regularization for Iterative Reconstruction,” in *International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 517–520.
- [209] Z. Zhao, G. J. Gang, and J. H. Siewerdsen, “Noise, sampling, and the number of

- projections in cone-beam CT with a flat-panel detector.,” *Med. Phys.*, vol. 41, no. 6, p. 061909, 2014, doi: 10.1118/1.4875688.
- [210] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch, and G. Langs, “Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem,” *Eur. Radiol. Exp.*, vol. 4, no. 1, 2020, doi: 10.1186/s41747-020-00173-2.
- [211] W. Zhu *et al.*, “AnatomyNet: Deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy,” *Med. Phys.*, vol. 46, no. 2, pp. 576–589, 2019, doi: 10.1002/mp.13300.
- [212] V. V Nagarkar *et al.*, “New scintillators for medical imaging,” in *APS March Meeting*, 2019, p. E34.002.
- [213] J. A. MacNeil and S. K. Boyd, “Bone strength at the distal radius can be estimated from high-resolution peripheral quantitative computed tomography and the finite element method,” *Bone*, vol. 42, no. 6, pp. 1203–1213, 2008, doi: 10.1016/j.bone.2008.01.017.
- [214] M. Viceconti, M. Qasim, P. Bhattacharya, and X. Li, “Are CT-Based Finite Element Model Predictions of Femoral Bone Strength Clinically Useful?,” *Curr. Osteoporos. Rep.*, p. 1, 2018, doi: 10.1007/s11914-018-0466-4.
- [215] L. Shen, W. Zhao, and L. Xing, “Harnessing the power of deep learning for volumetric CT imaging,” in *Proc. SPIE*, 2019, vol. 1094826, no. March, doi: 10.1117/12.2513032.
- [216] M. Mardani *et al.*, “Deep Generative Adversarial Networks for Compressed Sensing (GANCS) Automates MRI arXiv : 1706 . 00051v1 [cs . CV] 31 May

2017,” pp. 1–12.

- [217] K. Fukushima, “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biol. Cybern.*, vol. 36, no. 4, pp. 193–202, 1980, doi: 10.1007/BF00344251.
- [218] Y. LeCun *et al.*, “Backpropagation applied to digit recognition,” *Neural computation*, vol. 1, no. 4. pp. 541–551, 1989.
- [219] B. Hanin, “Universal function approximation by deep neural nets with bounded width and relu activations,” *Arxiv*, no. 1708.02691, pp. 1–9, 2017.
- [220] V. Antun, F. Renna, C. Poon, B. Adcock, and A. C. Hansen, “On instabilities of deep learning in image reconstruction - Does AI come at a cost ?,” *Arxiv*, no. 1902.05300, 2019.

