

LOW DOSE COMPUTED TOMOGRAPHY RECONSTRUCTION METHODS

by

SayedMasoud Hashemi Amroabadi

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of The Institute of Biomaterials and Biomedical
Engineering
University of Toronto

© Copyright 2015 by SayedMasoud Hashemi Amroabadi

Abstract

Low Dose Computed Tomography Reconstruction Methods

SayedMasoud Hashemi Amroabadi

Doctor of Philosophy

Graduate Department of The Institute of Biomaterials and Biomedical Engineering

University of Toronto

2015

This study examines potential methods of achieving a reduction in X-ray radiation dose of Computer Tomography (CT) without significantly degrading the quality of the CT images. Assuming that the effective X-ray dose in CT is proportional to the product of the exposure area, the number of projections and the dose for each projection, four approaches are proposed to lower the radiation dose. Each of the proposed approaches controls one of these factors. Increasing the slice thickness increases the number of detected X-ray photons, thereby reducing the image noise. However, it also reduces the detectability and measurement accuracy of the small structures. An optimum slice thickness/overlap for low dose thoracic-CT images is proposed to compromise between the detectability and characterization accuracy of small pulmonary nodules.

Compressed Sensing (CS) based CT reconstruction methods have shown significant potential for low dose CT imaging by reducing the number of projections. The connection between the statistical iterative CT reconstruction and compressed sensing is explored, based on maximum a posteriori formulation of CT reconstruction. It is shown that this formulation can be transformed into a weighted CS problem, the weights of which are direct consequences of the geometry and the CT statistics. A computationally efficient solution is proposed, enabling lower dose CT scans, while substantially reducing the computational load in comparison to conventional CS-based CT reconstruction methods.

An iterative denoising algorithm is proposed to improve the CNR of the low dose CT images. To achieve this goal, a noise confidence region evaluation algorithm is proposed to adaptively update the regularization parameters, by validating the statistical properties of the denoised

images at the end of each iteration. The algorithm stops when the statistical properties of the denoising residual resemble those of the considered noise.

Finally, an innovative combination of a newly developed accelerated wavelet inverse Radon transform and non-convex CS formulation is proposed to recover the wavelet subbands of CT images from a reduced number of locally scanned projections. The proposed method reduces the radiation dose by simultaneous reduction of the number of projections and the area of exposure.

Dedication

To my parents, Mohammad and Fatema.

Acknowledgements

I would like to thank all people who have helped and inspired me during my study. Foremost, I would like to express my sincere gratitude to my co-advisors Dr. Narinder Paul and Professor Richard S.C. Cobbold for the continuous support of my study and research. Without their guidance and assistance, this thesis would not have been possible. I have been very fortunate to have the advisers who gave me the freedom to explore my area on my own. I would like to thank Dr. Paul for his unwavering confidence in my abilities, and for his generous financial support. He always made time for me to talk about ideas, despite an overwhelming schedule. I would like to thank Prof. Cobbold for his valuable guidance to learn and enjoy the real essence of research.

I would like to thank Dr. Patrick R. Gill for his guidance, collaboration, and motivating feedbacks. I am extremely grateful to my former supervisor, Prof. Soosan Beheshti, for her insightful comments and constructive criticisms. I am grateful to Dr. Hatem Mehrez for being my mentors during my internship at Toshiba medical systems, for his feedbacks and help even after finishing the internship.

I thank Prof. M. Joy and Prof. R. Adve for serving as my committee members and for their valuable guidance and feedbacks. I thank Prof. G. Wright and Prof. I. Cunningham for serving as my examiners, for their comments and valuable time.

I am grateful to Amir Manbachi for the friendship and the enormous fun of working with him, which are just hard to express in few sentences. I would also like to thank many friends of mine, who have helped me and made my life more fun. I would like to thank Hisham Assi for the joyful and helpful discussions about different ideas.

Finally, my special thanks go to my family: my father, Mohammad, my mother, Fatema, my brothers, Mahdi and Meisam, my step-sister, Sheida and my beloved nephew, Ali. I thank them all for their patience and unreserved support, which are the cornerstones of my Ph.D. life.

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Thesis Objectives	3
1.3	Thesis Organization	5
2	Computed Tomography: Background Review	7
2.1	Generations of Computed Tomography	7
2.1.1	First Generation	8
2.1.2	Second Generation	8
2.1.3	Third Generation	9
2.1.4	Fourth Generation	10
2.1.5	Electron Beam CT (EBCT)	10
2.1.6	Modern CT scanners	11
2.2	Single Step CT imaging Reconstruction	12
2.2.1	Filtered Back Projection of Parallel and Fan beam CT Reconstruction .	13
2.2.2	Helical Cone Beam Reconstruction	15
2.3	Central Slice Theorem	17
2.3.1	Pseudo-Polar Fourier Transform	18
2.4	Rebinning Process	19
2.5	Iterative Reconstruction	20
2.5.1	Algebraic Reconstruction Technique (ART)	20
2.5.2	Statistical Iterative Reconstruction Methods	21
2.6	X-ray Dose measurement	22
2.6.1	Clinical Dose Reduction Strategies	24
3	Compressed Sensing in Computed Tomography	27
3.1	Historical Background	27
3.2	Basic Concepts	29
3.2.1	Optimization Methods	30

3.3	Signal Measurement and Recovery Conditions	31
3.4	Recovery Algorithms	33
3.4.1	ℓ_1 minimization algorithms	33
3.4.2	Greedy Algorithms	35
3.4.3	Reweighted ℓ_1 Minimization	36
3.5	Total Variation minimization	37
3.5.1	Isotropic TV minimization with Split Bregman Algorithm	37
3.6	Application of Compressed Sensing in Computed Tomography	38
4	Image Reconstruction of Small Pulmonary Nodules¹	42
4.1	Materials and methods	43
4.1.1	Image reconstruction parameters and nodule measurements	44
4.1.2	Statistical Analysis	47
4.2	Results	47
4.2.1	Computer simulated data: Virtual nodules	47
4.2.2	Phantom data: In vitro synthetic nodules	49
4.2.3	Patient data: In vivo nodules	49
4.3	Discussion and Conclusion	50
4.4	Summary	51
5	CT Reconstruction through MAP Based Compressed Sensing Modeling	59
5.1	Fan Beam and Helical Cone Beam Image Reconstruction	60
5.2	Maximum a Posteriori (MAP) Model of CT	62
5.2.1	MAP Model of CT imaging and Weighted-CS formulation	62
5.2.2	Sparsity and Piecewise Linearity of the CT Images in MAP Model	63
5.2.3	Proposed CS-based CT Formulation incorporating Measurement Noise and Interpolation Error	64
5.2.4	Calculation of ϵ	65
5.3	Expectation Maximization based Solution for the Proposed Weighted CS-based CT Reconstruction	66
5.3.1	Latent Image and Expectation Maximization Method for ℓ_1 and TV minimization	67
5.3.2	Fourier Based E-Step Formulation	68
5.3.3	FCSA-LEM Algorithm in a Nutshell	69
5.4	Simulation Methods and Results	69
5.4.1	Phantoms and CT measurements	69
5.4.2	ART-TV based method	71

5.4.3	Results and Discussion	72
6	An Adaptive Method for Updating Regularization Parameters	79
6.1	Importance of the Regularization Parameter Selection	80
6.2	NCRE Stopping Criterion and Regularization Parameter Updating Method	81
6.2.1	Noise Confidence Region Evaluation (NCRE)	81
6.2.2	Robust Convergence Step	83
6.3	NCRE in Denoising and Compressed Sensing	84
6.3.1	Soft Thresholding (NCRE-ST)	84
6.3.2	TV Denoising (NCRE-TV)	84
6.3.3	Iterative Soft Thresholding Algorithm (NCRE-ISTA)	84
6.3.4	CT Image Denoising (NCRE-BM3D)	85
6.4	Results	87
6.5	Summary	90
7	Compressed Sensing in Local CT Reconstruction	94
7.1	Wavelet Transform and CT image Reconstruction	95
7.1.1	Wavelet based Local CT Reconstruction	96
7.1.2	Compressed Sensing based CT Reconstruction	97
7.2	Proposed method: AWiR-SISTA	98
7.2.1	Accelerated Wavelet inverse Radon (AWiR) Transform	98
7.2.2	AWiR-SISTA steps	99
7.3	Results	101
7.4	Summary	101
8	Conclusions, Summary and Future Directions	106
8.1	Contributions	108
8.2	Suggestions for Future Work	109
APPENDICES		112
A	Noise in Computed Tomography Reconstructed Images	112
A.1	Noise in CT images	112
A.2	Notes on Noise Statistics in reconstructed CT images	115
A.3	Noise Considerations in our proposed Algorithm (NCRE-BM3D)	115
A.4	Summary	115
B	Measured CT number in Small Pulmonary Nodules	121

B.1	Volume of Spherical Segment	122
B.2	Volume of Spherical Cap	122
B.3	Importance of ROI in measurements	124

List of Tables

6.1 Comparison of BayesShrink and NCRE-ST with 1D test signals. SNR values are in dBs and are the average of 50 runs.	89
-------------------------------------------------------------------------------------------------------------------------------	----

List of Figures

2.1 Three X-ray source geometries: (A) Parallel Geometry, (B) Fan beam Geometry, and (C) Cone beam Geometry.	8
2.2 First Generation of CT scanners which completed the scan in two repeated translation and rotation steps. Average scan time was about 6 min for one 360° rotation.	9
2.3 Second Generation of CT scanners which completed the scan task in two simultaneously repeated translation and rotation steps.	10
2.4 (A) Third Generation of CT scanners which includes a rotation step only. (B) Ring Artifact caused by problematic detector elements or uncalibrated detectors.	11
2.5 Fourth Generation of CT scanners with a 360° detector ring.	12
2.6 Schematic of an Electron Beam CT (EBCT). Reproduced, with permission, from [39].	12
2.7 Helical (Spiral) CT trajectory.	13
2.8 (A) Shepp-Logan phantom, (B) reconstructed by Back Projection, and (C) with Filtered Back Projection.	14
2.9 (A) Parallel beam geometry and (B) 3 rd generation fan beam geometry with a flat detector (C) 3 rd generation fan beam geometry with a curved detector, (D) Rebinning illustration.	15
2.10 (A) Helical trajectory and (B) the fan beams in parallel z-slices.	16
2.11 Central Slice Theorem and polar to Cartesian interpolation.	18

2.12 (A) Pseudo-polar grids: red lines are Basically Horizontal (BH) and the black lines are Basically Vertical (BV). (B) Polar Grids (red dots) on the pseudo-polar grids (gray dots).	19
2.13 Projection matrix modeling in Algebraic Reconstruction Technique (ART).	21
2.14 (A) Angular modulation in which the x-ray tube current is lowered as the X-rays are aimed in the anteroposterior directions, and increased when the X-rays are aimed in the lateral-medial directions, when there will be more X-ray attenuation. (B) Z-axis modulation in which, for example, fewer X-rays are required in the abdominal region superior to the pelvic bones compared with the pelvic region. Reproduced, with permission, from [68].	25
2.15 Graph of tube current (mA) superimposed on a CT projection radiograph showing the variation in tube current as a function of time (and, hence, table position along the z-axis) at spiral CT in a 6-year-old child. An adult scanning protocol and an AEC system (CareDose 4D; Siemens Medical Solutions) were used with a reference effective tube current-time product of 165 mAs. The mean effective tube current-time product for actual scanning was 38 mAs (effective tube current-time product-tube current-time product/pitch). Reproduced, with permission, from [69].	26
2.16 The linear attenuation coefficients of iodine, cortical bone and water as a function of X-ray energy. The yellow arrows show the contrast between water and iodine in 80kV and 140kV, which are more distinct in 80kV. Reproduced, with permission, from [72].	26
3.1 Radiation dose reduction in CS-based CT reconstruction methods: (A) Regular CT system with many equiangular X-ray projections, and (B) Sparse view CT with few X-ray projections. Reproduced, with permission, from [75].	28
3.2 Sparse representation of a CT image with wavelet transform: (A) A chest CT image, (B) Subbands of the image with one level Haar wavelet, and (C) sorted absolute values of the wavelet coefficients of the image with 5 level Haar wavelet transform.	30
3.3 Comparison of the different ℓ_p norm minimization problems in sparse recovery problems. Reproduced, with permission, from [76].	31

3.4 Comparison of CT image reconstruction methods using a greatly reduced number of projections with that obtained using Compressed Sensing. (A) Simulated 128×128 Shepp-Logan phantom reconstructed with 36 projections with: (B) Filtered Back Projection, (C) Algebraic Reconstruction Technique, and (D) Compressed Sensing.	39
3.5 Angular independent coded aperture CT. The white regions indicate the sampled part of the data. Reproduced, with permission, from [74].	40
3.6 Different sampling strategies being considered, as demonstrated on a fan-beam geometry. The blue circles denote source locations (corresponding to view angles), and lines denote source-detector pairs (for clarity, only the lines for one source location are shown for each case). The solid and dashed lines correspond to measured and non-measured line integrals, respectively. The large circle represents the source trajectory, and the small circle represents the imaged domain. (A) Uniform view sub-sampling, (B) random view angle sub-sampling, (C) uniform detector sampling, and (D) random detector sub-sampling. Reproduced, by permission, from [114].	41
3.7 Schematic of a coded aperture CT scanner proposed in [117]. Black regions block the X-ray rays. Reproduced, with permission, from [117].	41
4.1 Illustrating the influence of effective slice-thickness on lung nodules: (A) Simulated nodules 5, 8, 10 mm diameter (left to right in each row) and 100 HU, -630 HU, -800 HU CT# (top to bottom row in each image) reconstructed with 0.5/0.5, 3.0/1.5, 4.0/2.0, 5.0/2.5 mm slice thickness/overlap (left to right). Qualitatively, there is a decrease in edge sharpness and image noise with larger slice thickness. (B) Solid \sim 5 mm nodules (top row) and \sim 5 mm ground glass nodules (bottom row) from in vivo (left column) and in vitro (right column) image data sets. Qualitatively the in vivo and in vitro datasets have comparable nodule appearances for low dose CT. Slice thickness/ Overlap = 2.0/2.0 mm and WW/WL = -550/1600 for both in-vitro and in-vivo nodules.	45
4.2 Variation in measured CT# and CNR with changes in the effective slice thickness (left: non-overlapped and Right: with 50% overlapped) and scan starting position for a 5 mm solid simulated nodule (Red bars show the STD of the changes due to the ten different simulations performed and scan starting position): (A) Measured mean CT attenuation number (CT#) and (B) Contrast to Noise Ratio (CNR).	48

4.3	Variation in CT# and CNR with change in effective slice thickness (left: non-overlapped and Right: with 50% overlapped) for simulated nodules: (A, B) Measured mean CT attenuation number (CT#) and (C, D) Contrast to Noise Ratio (CNR). (A, C) correspond to the simulated solid nodules and (B, D) correspond to the simulated ground glass nodules.	53
4.4	Variation in measured CT# and CNR with change in effective slice thickness (left: non-overlapped and Right: with 50% overlapped) of the in vitro Nodules: (A, B) Measured mean CT attenuation number (CT#) and (C, D) Contrast to Noise Ratio (CNR). (A, C) correspond to the synthetic solid nodules and (B, D) correspond to the synthetic ground glass nodules.	54
4.5	Variation in measured CT# and CNR with change in effective slice thickness (left: non-overlapped and Right: with 50% overlapped) for in vivo nodules: (A, B) Measured mean CT attenuation number (CT#) and (C, D) normalized Contrast to Noise Ratio (CNR). (A, C) correspond to two solid nodules and (B, D) correspond to two ground glass nodules.	55
4.6	Histogram of the measurement error for CT# of the in vivo nodules at three different slice-thicknesses /overlap. Measurements were made at 1.0/1.0, 2.0/2.0, and 5.0/2.5 mm, and the error is calculated based on the gold standard values of the CT# obtained from the 0.5/0.5 mm reconstruction.	56
4.7	A single trans-axial CT images reconstructed at the same table position through the maximum dimension of a sub 5 mm nodule using two different slice-thickness/overlap reconstructions; 5.0/2.5 mm (left column) and 2.0/2.0 mm (right column), with lung (top row) and mediastinal kernel (bottom row) reconstructions. For the larger slice-thickness reconstructions the nodule appears to be a GG but, from the images with a smaller effective slice thickness a solid nodule is clearly indicated.	57
4.8	Comparison of the changes in CNR with Iterative Reconstruction (AIDR3D) and FBP as a function of effective slice thickness (left: non-overlapped and Right: with 50% overlapped): Mean values of the normalized Contrast to Noise Ratio (CNR) of in vivo nodules as a function of reconstruction slice thickness/overlap for: (A) solid nodules and (B) ground glass nodules.	58
5.1	Flowchart of the proposed weighted CS-based CT reconstruction method.	61

5.2	Illustration of the Error Adaptation Weight (EAW) for a simple example. The equiangular lines are plotted with red dashed lines and smaller red circles show the measurements. The solid lines and the gray larger circles represent the pseudo-polar grids.	66
5.3	Reconstruction time comparison using a standard desktop computer, for: (1) fan beam filtered back projection (FBP) reconstruction, (2) the proposed method (FCSA-LEM), and (3) a fan beam ART-TV based method.	72
5.4	ART-TV Reconstruction results: (A) Shep-Logan phantom reconstructed from 200 projections, and (B) Reconstruction Error.	73
5.5	Simulation results for a Shepp-Logan phantom. The bottom row shows an expanded view of the region marked in panel (A). (A) Original phantom image. Reconstructions using 128 projections with (B) inverse pseudo-polar Fourier transform using the least squares method (normalized error ≈ 0.9), (C) an iterative soft thresholding based method (TwIST) (normalized error $\approx 10^{-1}$), and (D) the proposed FCSA-LEM method (normalized error $\approx 10^{-2}$). Rebinned parallel rays were used in all three methods to recover the image.	74
5.6	Normalized reconstruction error for the simulated Shepp-Logan phantom reconstructed with the inverse pseudo polar Fourier transform (PPFT) using least squares (LS) method, using an iterative soft thresholding based method (TwIST), and the proposed FCSA-LEM method. Rebinned parallel rays were used in all three methods to recover the 512×512 image.	74
5.7	Effect of Error Adaptation Weight (EAW) on the PSNR of the reconstructed image. Solid red line shows the results when EAW was included and dashed blue line shows the reconstructed image results when EAW is not included.	75
5.8	Comparison of FBP and the proposed method for a custom built cardiac plaque phantom. (A) Image reconstructed from 900 projections using FBP, (B) Image reconstructed from 200 equiangular projections with FBP, (C) Image reconstructed from 200 equiangular projections with the proposed method, FCSA-LEM, (D) scan protocol used in images (E&F), in which the projections are taken within the white areas, (E) FBP reconstructed image from the 450 projections gathered from the mask shown in image (D), and (F) FCSA-LEM reconstructed image from projections shown in image (D).	76

5.9	Comparison of FBP and the proposed method for a Lungman CT phantom. (A) Image reconstructed from 900 projections with FBP (noise standard deviation 50HU), (B) Image reconstructed from 200 projections with FBP, (C) Image reconstructed from 200 projections with the proposed method, FCSA-LEM (noise standard deviation 22HU).	77
5.10	Comparison of FBP and the proposed method for the chest CT scan data from a patient. (A) Image reconstructed from 900 projections with FBP, (B) Image reconstructed from 200 projections with FBP, (C) Image reconstructed from 200 projections with the proposed method, FCSA-LEM.	77
5.11	Helical scan tested on a simple simulated phantom. (A) The original phantom. (B) Image reconstructed with the proposed method. (C) Difference between the true image and the reconstructed image.	78
6.1	Three possible regions for the residual at the end of each iteration. If it lies in Region II (Noise Confidence Region) denoising is stopped.	83
6.2	Flowchart of BM3D denoising algorithm. The operations surrounded by dashed lines are repeated for each processed block, marked with R . Reproduced, with permission, from [187].	86
6.3	Grouping the blocks of noisy images corrupted by a zero mean white Gaussian noise with standard deviation 15. Reproduced, with permission, from [187].	86
6.4	Schematic of the proposed CT image denoising scheme (NCRE-BM3D).	87
6.5	Comparison of regular TV denoising and NCRE-TV: (A) Original Image, (B) Noisy image ($\text{SNR} = 19.94 \text{ dB}$), (C) Denoised with BayesShrink ($\text{SNR} = 21.93 \text{ dB}$), (D) Denoised with NCRE-ST ($\text{SNR} = 23.09 \text{ dB}$), (E) Denoised by regular TV denoising ($\text{SNR} = 22.49 \text{ dB}$), and (F) Denoised by NCRE-TV ($\text{SNR} = 25.56 \text{ dB}$).	88
6.6	(A) Original Signal, and (B) Recovery error of ISTA (Top) and NCRE-ISTA (Bottom).	90
6.7	Compares the number of measurements needed for exact recovery by ℓ_p and ℓ_1 minimizations, using NCRE-ISTA.	91
6.8	Low contrast slice of CATPHAN denoised by the proposed method and Adaptive Iterative Dose Reduction (AIDR3D). Top: scanned with 50mAs/120kVp and Bottom: scanned with 300mAs/120kVp. Left: image reconstructed with FC52, Middle: reconstructed with AIDR3D, and Right: reconstructed with FC52 and denoised by BM3D-NCRE. In all images Window-Width/Window-Level is: WW=100/WL=70.	91

6.9 Top: Line pair slice of the CATPHAN scanned with 50mAs and 120kVp (WW=400/WL=60). Bottom: Red rectangular ROI of the images, the red lines show the cut off line pair resolution (WW=400/WL=500). Left: Phantom reconstructed with FC52 (noise standard deviation=64HU), Middle: Phantom reconstructed with FC52+AIDR3D (noise standard deviation=31HU), and Right: Phantom reconstructed with FC52 and denoised with the proposed method (noise standard deviation=22HU).	92
6.10 (A) CT lung image, noise std in red rectangle is 60HU, (B) Denoised with BM3D, noise 38HU, and (C) Denoised with NCRE-BM3D, noise 15HU. WW/WL = 1400/-450HU, (D) The difference between the noisy and denoised image, and (E) Top: Original, Middle: BM3D, Bottom: NCRE-BM3D, WW/WL = 1500/-600HU.	93
7.1 Region of Interest scanning method.	95
7.2 Scaling and wavelet functions of Daubechies 4.	96
7.3 Wavelet inverse Radon transform methods. (A) Wavelet inverse Radon transform according to [31], and (B) Accelerated wavelet subbands recovery. The blocks marked by (1) and (2) in (A) are replaced by the corresponding ones in (B).	98
7.4 Proposed AWiR-SISTA method for local CT image reconstruction.	99
7.5 Comparison of full Shepp-Logan phantom images reconstructed from 200 projections by WiR and AWiR-SISTA: (A) Subbands recovered by AWiR, (B) image recovered by AWiR with a normalized error of 2×10^{-1} , (C) image subbands reconstructed with AWiR-SISTA, and (D) phantom reconstructed by AWiR-SISTA with a normalized error of 1.02×10^{-3}	102
7.6 An ROI with radius of 1/2 of the phantom: (A) Subbands recovered by WiR from full data, (B) ROI image recovered by WiR with a normalized error of 2×10^{-1} (C) Subbands recovered by AWiR-SISTA, and (D) ROI image recovered by AWiR-SISTA from 200 projections with a normalized error of 0.8×10^{-2}	103
7.7 Patient ROI with radius of 1/2 that of a lung CT image recovered by. (A) Original Image, (B) Difference between the ROI image reconstructed by WiR and the original image, (C) Difference between the ROI image reconstructed with the proposed method, (D) WiR with normalized error of 4×10^{-1} , (E) AWiR-SISTA with normalized error of 0.7×10^{-2} , and (F) Difference of the original ROI image and the ROI recovered by AWiR-SISTA.	104

7.8 A non-centered ROI in a Shepp-Logan phantom with a radius 1/2 phantom reconstructed from 200 projections by: (A) WiR with a normalized error of 2.5×10^{-1} , and (B) AWiR-SISTA with a normalized error of 1.1×10^{-2}	105
8.1 The projections angles that minimize the fan beam to pseudo-polar parallel beams interpolation error.	111
A.1 Streak artifact caused by the presence of high attenuating bone structures in the image. Left: Image taken from Toshiba Aquilion one CT scanner with corrections, and Right: image without the corrections.	113
A.2 Phantom scanned with eight different X-ray source currents: top left to right 5, 10, 25, 50 mAs and bottom left to right 100, 150, 200, 250 mAs.	114
A.3 Compares the noise distribution of the noise in the six regions of the phantom shown in figure A.2 with white Gaussian noise.	117
A.4 Shows the noise variance changes for different CT#’s in the six regions of the phantom.	118
A.5 Shows the noise variances of each region for different scan doses. Each connected line shows the result of a fixed dose (tube current). The figure at the bottom shows the 100, 150, 200, 250 mAs data which are compact in the upper figure. The region with CT# of 0 HU has smaller variances, since this is the CT# associated to water and the scanners are well calibrated for its variations. .	119
A.6 Soft tissue region around the lung, which is used to evaluate the noise characteristics in the NCRE algorithm. The statistical distribution of the noise in this region is shown, with blue solid line, and is compared with Gaussian distribution with same variance, red dashed line.	119
A.7 Depicts the noise statistics of two small regions used in BM3D denoising and compares their noise distribution with that of similar Gaussian noises.	120
B.1 (A) Spherical Segment, and (B) Spherical Cap.	121
B.2 Three cases with largest cross section of a nodule, when the slice thickness is less than the nodule radius.	123
B.3 CT# changes of a nodule with diameter of 5 mm and CT# of 100 HU, as a function of slice thickness.	124
B.4 A 66% ROI at the center of a nodule.	124

List of Abbreviations and Symbols

Symbols	Description
$f(x,y)$	Computed Tomography image in (x,y) coordinate system.
$\bar{f}(x,y)$	Noiseless Computed Tomography image in (x,y) coordinate system.
$\hat{f}(x,y)$	Noiseless Computed Tomography image in (x,y) coordinate system.
Δf	Denoising residual.
\mathbf{x}	Vectorized image.
\mathbf{y}	Measured X-ray projections.
$\bar{\mathbf{y}}$	Expected value (noiseless) of the X-ray projection measurements.
$\sigma_{\mathbf{y}_i}^2$	Variance of the projection data.
\mathbf{n}	Measurement noise.
σ_e^2	Variance of electrical noise in CT measurements.
σ_n^2	Variance of the CT Image noise.
D^{-1}	Covariance matrix of the measurement noise.
d_i 's	Diagonal elements of the noise covariance matrix D .
\mathbf{A}	Measurement matrix, X-ray projection matrix.
\mathbf{TV}	Total variation norm.
W^T, W	Wavelet transform matrix and its inverse.
$\lambda, \lambda_1, \lambda_2$	Regularization parameters in CS problems.
θ	2D-wavelet coefficients of image $f(x,y)$.
∇^h, ∇^v	First derivatives in the horizontal and vertical directions.
$\delta(x,y)$	Dirac delta function.
$\xi(l, \varphi)$	CT detector photon count at angle φ and distance l .
$\bar{\xi}(l, \varphi)$	Expected value of the photon count of the CT detector.
ξT	Number of radiated photons from the X-ray source.
s, q	GGD parameters.
$\ \cdot\ _p$	p-norm.
\cdot^T	Conjugate transpose.
$\rho, \Psi_j^n, v_{\rho,n}, a_n,$	Parameters used in ρ -variation distribution.
$\varphi_{BH}, \varphi_{BV}$	Projection angles on Basically Horizontal (BH) and Basically Vertical (BV) lines.
BV, BH	Basically Horizontal and Basically Vertical grids on pseudo-polar coordinate.
ω_x, ω_y	Spatial angular frequencies in the x and y directions.
$l, \gamma, \gamma_T, R, \varphi, \beta$	Fan and Parallel beam geometry parameters.
$g(l, \varphi)$	Parallel beam projection data.

$\mathcal{R}(\gamma, \beta)$	Fan beam projection data.
$(u, v), \phi_{ss}, \psi, P, P_d,$	
$w(\phi_{ss}, u), \Delta_z, \mathcal{D}$	Helical Cone Beam geometry parameters.
e_i	Interpolation error.
\mathbf{C}	Confidence matrix with elements of \mathbf{c}_i .
Δ	FCSA combination factor.
\mathbf{z}	Latent image used in EM.
$\mathcal{F}_{1D}, \mathcal{F}_{2D}$	1D and 2D Fourier transform matrix.
Φ	Sparsifying transform.
α	EM regularization parameter.
n_d	number of detectors.
$n_{k\varphi}$	number of projections.
k	number of non-zero elements in sparse signals.
m	number of measurements.
$\mu(a, b)$	Coherence between a and b .
Λ	set of regularization parameters.
$\mathcal{G}(z, \mathbf{n})$	Ordering function used in NCRE.
ζ	Upper and Lower boundary width parameter of NCRE.
ϕ	scaling function.
ψ	wavelet function.
Acronyms	Description
1D	One Dimensional.
2D	Two Dimensional.
3D	Three Dimensional.
ADMM	Alternating augmented Lagrangian methods.
AIDR3D	Adaptive Iterative Dose Reduction 3D.
ART	Algebraic Reconstruction Technique.
AWGN	Additive White Gaussian Noise.
AWiR	Accelerated Wavelet inverse Radon Transform.
BM3D	Block Matching and 3D filtering.
CB-SSRB	Cone Beam Single Slice ReBinning.
CoSaMP	Compressive Sampling Matching Pursuit.
CNR	Contrast to Noise Ratio.
CS	Compressed Sensing.
CST	The Central Slice Theorem.
CT	Computed Tomography.

DFR	Direct Fourier Reconstruction.
EAW	Error Adaptation Weight.
EBCT	Electron Beam CT.
EM	Expectation Maximization.
EST	Equally Sloped Tomography.
FBP	Filtered Back Projection.
FCSA	Fast Composite Splitting Algorithm.
FFT	Fast Fourier Transform.
GGD	Generalized Gaussian Distribution.
GO	Ground Glass Opacity.
IHT	Iterative Hard Thresholding.
ISTA	Iterative Soft Thresholding Algorithm.
kVp	peak voltage of X-ray source in kilo volt.
LASSO	Least Absolute Selection and Shrinkage Operator.
LDCT	Low-Dose CT.
LEM	Latent Expectation Maximization.
MAD	Median Absolute Deviation.
MAP	Maximum a Posterior.
mAs	product of rotation time (second) and X-ray source current (Milli-Amps)
MDCT	Multi-slice Detectors CT.
ML	Maximum Likelihood.
MP	Matching Pursuit.
MSE	Mean Squared Error.
NCPN	Non-Calcified Pulmonary Nodule.
NCRE	Noise Confidence Region Evaluation.
OMP	Orthogonal Matching Pursuit.
OSEM	Ordered Subset Expectation Maximization.
PICCS	Prior image constrained compressed sensing.
PPFT	Pseudo-polar Fourier transform.
PSNR	Peak Signal to Noise Ratio.
PPS	Plasma Panel Sensor.
PWLS	Penalized weighted least-squares.
ROI	Region of Interest.
SIR	Statistical Iterative Reconstruction.
SSIM	Structural Similarity Index.
SISTA	Subband Iterative Hard Thresholding.

StOMP	Stage-wise Orthogonal Matching Pursuit.
SWM	Switching Median filter algorithm.
TFT	Thin Film Transistor.
TwIST	Two Step Iterative Soft Thresholding.
WiR	Wavelet inverse Radon.
WW/WL	Window Width/Window Level.

Table 1

Chapter 1

Introduction

For more than a century conventional X-ray radiography has offered a valuable non-invasive means of diagnosis, but, it suffers from a number of limitations. For instance, imaging the brain with radiography typically yields insufficient details. This is due to the fact that each picture element displays the sum of all attenuation components along the ray lines. As a result, the contrast of the images are dominated by the structures with high attenuation, such as bone, and thereby, the low intensity objects, such as soft tissues, are completely hidden in most cases. This issue is solved through the development of cross-sectional imaging modalities, such as Computed Tomography (CT), in which the neighboring and superimposed structures have very little influence on the image.

Over the past two decades, there has been a major increase in the clinical use of CT primarily due to its unsurpassed speed and the fine detail that can be obtained in cross-sectional views of the internal organs. However, compared to conventional radiography, CT results in a relatively large radiation dose to patients. Studies over the past decade have shown that the higher radiation dose is of serious long-term concern for increasing the potential risk of developing cancer. Based on comparable risk estimates along with information on CT scan usage from 1991 through 1996, about 0.4% (or 4 in 1000) of all cancers in the United States may be related to radiation from CT scans [1, 2]. For older people, this issue is not very serious, since cancers that could result from CT scan-related radiation exposure would take years to develop. However, given the fact that there is increased use of CT scans in the pediatric population, this could be a problem since these children can be expected to be alive many years after a scan was performed. As a result, low dose CT imaging that maintains the resolution and achieves good contrast to noise ratio has been the goal of many CT developments over the past decade.

1.1 Motivation

The main focus of this thesis is to explore the primary mechanisms that can be used to reduce the radiation dose in CT examinations without degrading the diagnostic quality of the images. To achieve a better understanding of the possible X-ray dose reduction approaches, two concepts are briefly reviewed in this section: image quality and the effect of radiation dose reduction on the image quality.

CT image quality is usually described in the sense of contrast, spatial resolution, image noise, and artifacts. Different metrics have been used to measure the performance of the physical systems and the reconstruction algorithms. Signal to noise ratio (SNR) and contrast to noise ratio (CNR) are usually used as surrogate measures for CT image quality. Other metrics are also used to measure the effect of detectors, reconstruction and post-processing algorithms on the images, such as modulation transfer function, noise power spectrum, noise equivalent quanta, and detector quantum efficiency. More accurate metrics have been recently used to model the observer performance with a good compatibility with the human observer studies, such as model observer methods and detectability index [3].

One of the most important applications of CT is in detecting low contrast structures, a task that is limited primarily by image noise and therefore by the radiation dose: a higher radiation dose results in a lower noise, thereby improving the contrast detectability. In one of the simplest definitions, the CT image noise is the standard deviation of voxel values in a homogeneous (typically water) phantom, and this is influenced by a large number of parameters, including:

1. Electronic noise: caused by electronic circuits and the X-ray detectors.
2. Statistical noise: resulting from fluctuations in the the number of X-ray photons reaching the detectors and detected by them. This can be controlled by:
 - Radiation dose, which itself is a function of X-ray source peak voltage (kVp), the product of X-ray source current and exposure time (mAs), table speed in helical CT exams (pitch), X-ray collimation and exposure area.
 - Reconstruction slice thickness, changing the thickness changes the number of detected X-ray photons. For example, compared with a slice thickness of 5 mm, a thickness of 10 mm approximately doubles the number of X-ray photons entering each detector.
3. The choice of reconstruction and filtering algorithms.
4. Artifactual noise: the presence of artifacts might be viewed as a form of noise that interfere with the interpretation of the CT image, for instance beam hardening, streak artifacts

and photon starvation effects.

Spatial resolution can be improved and electronic noise can be reduced by using smaller and better X-ray detectors [4]. Commercially available Plasma Panel Sensors (PPS) used in CT scanners are typically based on amorphous silicon (a-Si) Thin Field Transistors (TFT) and photodiode array technology coupled with a cesium iodine (CsI) scintillator layer. The advantage of these sensors is the compact size, which enables high resolution imaging. However, their outputs are highly affected by thermal noise and the electrical noises caused by amplifier circuits, which limits their application to higher dose radiations. Alternatively, X-ray sensitive photodetectors materials, such as amorphous selenium (a-Se), can be used that directly convert X-rays into electronic charge. As a result, improving the electronics and the detector materials can greatly affect the quality of low radiation dose images. But, this is beyond the scope of this thesis.

1.2 Thesis Objectives

The focus of this thesis is on the influence of statistical noise, and the improvement of reconstruction algorithms in low-dose CT imaging. The primary objective of this thesis is to reduce the X-ray radiation dose in CT without a significant degradation in the clinical diagnostic quality of the CT images, with a specific focus on thoracic CT images. The effective X-ray dose in CT scanners is proportional to the product of the X-ray exposure at each projection view and the number of projection views (which is a function of gantry rotation time and depends on the ability and speed of CT scanners to control the X-ray source). Low dose CT images reconstructed with conventional Filtered Back Projection (FBP), which directly calculates the image in a single reconstruction step, suffer from low contrast to noise ratios. A reduced radiation dose decreases either the number of emitted photons (by controlling the X-ray source current) or their energy (by controlling the voltage of the X-ray source). This increases the amount of photon noise in CT images and degrades the image quality. Several methods have been proposed for lowering the relative amount of noise in a low dose CT scan. These methods can be grouped as follows: 1) improving scan protocols and the hardware of the scanner [5], 2) those that employ new reconstruction methods [6, 7, 8, 9], 3) adding denoising algorithms [10, 11], and 4) reducing the X-ray exposure area. The challenges and limitations of available low dose imaging techniques are briefly discussed below. These challenges are addressed by proposing four different approaches discussed in chapters 4 to 7 of this thesis.

New iterative reconstruction approaches have been proposed to improve the reconstruction quality and to decrease image artifacts. They can be grouped as: Algebraic Reconstruction

Techniques (ART) [6, 12, 13] and; Statistical Iterative Reconstruction (SIR) [14, 15]. By using a suitable imaging model that includes the statistical properties of the CT images, iterative reconstruction algorithms can increase the quality of low dose CT images. However, genuine iterative methods [16] are computationally intensive, which has hindered widespread clinical application. For example, VEO [16] is a model based iterative reconstruction algorithm introduced by GE health care but, by comparison to FBP based methods, it is relatively slow. Hybrid techniques have been developed to gain lower computational load with some benefits of true iterative methods (e.g. AIDR3D from Toshiba, and ASiR from GE). However, there is no evidence of improved lesion detectability in sub-mSv CT imaging using these hybrid algorithms.

Iterative reconstruction methods have progressed with the introduction of Compressed Sensing (CS). This is a relatively recent innovation in signal processing that allows recovery of images from fewer projections than that required by the Nyquist sampling theorem [17, 18]. While conventional iterative CT reconstruction methods focus on reducing the X-ray exposure in each projection, CS permits reconstructions from fewer projections. Such methods are capable of reconstructing high quality images from a substantially smaller number of views than those needed in FBP [19, 20], thereby permitting the use of a much lower dose scanning protocol than that needed in conventional reconstruction methods. However, current CS-based reconstruction algorithms suffer from two drawbacks: their high computational load makes them impracticable for routine clinical use [8, 9, 21, 22, 23] and, the statistics of CT measurements [24, 25, 26, 27] are not usually incorporated in the problem formulation. As a result, it seems unlikely that these methods could be used directly for clinical CT systems.

Utilization of noise removal techniques can improve the quality of the low dose CT images. Different approaches have been proposed to reduce the effect of noise, such as projection noise removal before image reconstruction [28], and statistical CT reconstruction [15], which suppress the noise during reconstruction by the use of more realistic assumptions compared to FBP. Another common area of research includes the development of algorithms for noise reduction in reconstructed CT images. These methods are required to reduce noise while preserving edges, as well as small structures, that might be important for diagnosis. Partial differential equation based methods [29, 7] have been used as the standard edge-preserving methods in image domain denoising.

The state of the art CT denoising methods can typically be derived by optimizing an objective function, which includes a data-fidelity and a weighted regularization term, solved iteratively. A general limitation of denoising methods is the loss of contrast within small structures. One straight forward approach to achieve the best combination of noise removal and feature preservation is to tune the regularizing parameters. Manually fixed parameters are usually used in

the denoising algorithms, which combined with a good stopping criterion offers a good trade off between noise removal performance and feature preservation [30]. However, the choice of suitable values of these regularization parameters is still a challenge. If these parameters are not appropriately updated, the optimization gets trapped in a local minimum. Therefore, aside from the numerical techniques needed for computing the solution, there are two important issues in solving an image denoising problem: 1) estimating and updating the regularization parameter, and 2) iteration stopping criterion.

Another potential method for reducing the dose is to reduce the exposure area, by limiting it to just those regions that need detailed clinical examination. However, the FBP reconstruction method, as widely used in clinical applications, is a non-local transform because of the use of Fourier transform as its theoretical foundation. This means that the recovery of any pixel in a CT image requires knowledge of all projections thereby exposing the patient to a relatively large amount of radiation even when only a small region needs to be examined.

Several methods for local CT reconstruction (also known as interior tomography) have been previously studied. These algorithms are usually based on the localized nature of wavelet transforms [31, 32, 33], iterative reconstruction methods [34], or generalized FBP and truncated Hilbert transform [35, 36]. Combinations of local tomography methods with CS leads to huge dose reductions. However, since classical CS-based CT reconstruction methods use the discrete Radon matrix, they are not suitable for local CT reconstruction.

1.3 Thesis Organization

This thesis explores and reports on different possible methods of reducing the effective X-ray dose without sacrificing the image quality. Four algorithms are proposed here, each of which answers some of the challenges mentioned above.

Chapter 2 discusses the computed tomography hardware, its physics and the classical reconstruction algorithms. Chapter 3 presents the idea of compressed sensing, its application in CT reconstruction, and the challenges of the conventional CS-based CT reconstruction algorithms. In chapter 4 we find an optimum slice thickness/overlap for reconstruction of low-dose thoracic-CT images and study the effect of iterative reconstruction on detectability and characterization of small non-calcified pulmonary nodules.

In chapter 5 we explore the connections between the statistical iterative CT reconstruction and compressed sensing based CT reconstruction algorithms. The first part of this chapter leads to our proposed weighted CS model. This new CS optimization model is solved by a fast and computationally efficient method, proposed in the second part of the chapter.

A noise confidence region evaluation (NCRE) algorithm is proposed in chapter 6. NCRE pro-

vides an stopping criterion and a parameter updating approach to improve the performance of the iterative CT image denoising algorithms.

Chapter 7 presents a new CS-based algorithm that reconstructs high quality images from few locally scanned X-ray projections.

Chapter 2

Computed Tomography: Background Review

X-ray Computed Tomography (CT) is a medical imaging technique that enables the reconstruction of cross-sections of an object, using a series of X-ray measurements taken from many angles around the object. The first CT scanner was introduced by Sir Godfrey Hounsfield (1972) at EMI (London, England), for which, Alan Cormack and Hounsfield were awarded the Nobel Prize at 1979 [37]. Since then, CT scanners have been improved both mechanically and electronically enabling high resolution CT images to be generated and displayed very quickly. The current quality and speed with which they are generated is in part due to improvements in efficient image reconstruction algorithms. Although the problem of tomographic image reconstruction was mathematically solved by Johann Radon in 1917 [38], the first scanner used a simple algebraic reconstruction algorithm, since the connection between the new modality and the Radon transform was unknown at that time. This field is still evolving and new algorithms are currently used that adapt to the new projection geometries and new mathematical theories. In addition, iterative reconstruction algorithms, which enable major radiation dose reductions, have been developed. A variety of CT projection geometries, shown in figure 2.1, and the classical single step and iterative CT reconstruction algorithms are briefly discussed in this chapter.

2.1 Generations of Computed Tomography

A CT scanner has an X-ray source and a detector array, which are located in a gantry that can rotate at high speed to gather projection data from a subject that lies on a table. The peak voltage of the X-ray source (kVp) controls the energy of the photons and the X-ray current

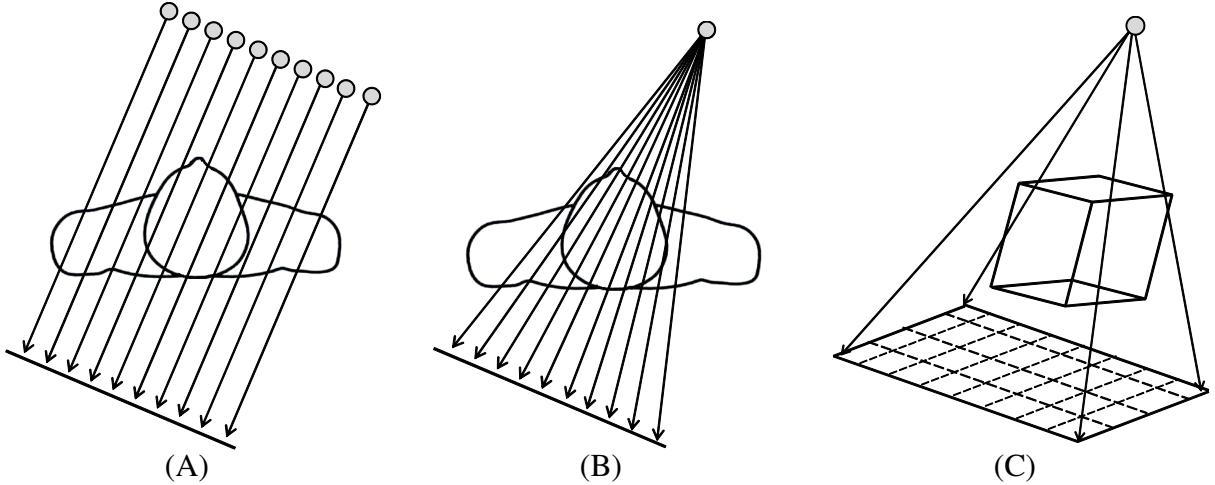


Figure 2.1: Three X-ray source geometries: (A) Parallel Geometry, (B) Fan beam Geometry, and (C) Cone beam Geometry.

(mA) controls the photon count. A nonlinear combination of these two parameters controls the X-ray dosage.

2.1.1 First Generation

The first generation CT scanners consisted of a dedicated head scanner with a water filled box. Pencil-width collimated beams of X-ray were used and these were detected on the opposite side using a single detector. The width of the X-ray beam was 3 mm within the plane of the image slice and 13 mm wide along the axis of the object (perpendicular to the slice). The X-ray source and the single element detector translate through the slice and rotate around the object with 1° rotation steps, as shown in figure 2.2. A simple implementation of Algebraic Reconstruction Technique (ART) was used to reconstruct the cross-sectional images in the first generation [39, 40].

2.1.2 Second Generation

The very long time to complete a scan was one of the most important concerns about the first generation CT systems. The scan time was reduced in the second generation with the use of multiple narrow beams and multiple detectors, in 1974. This scanner was the first waterless full-body CT scanner, that was introduced at Georgetown University [41]. This device introduced several innovations which are still used in all CT scanners: table movement through the gantry, gantry tilting, and a laser indicator to position slices.

A Fourier-based reconstruction algorithm was used to reconstruct the images in this scanner

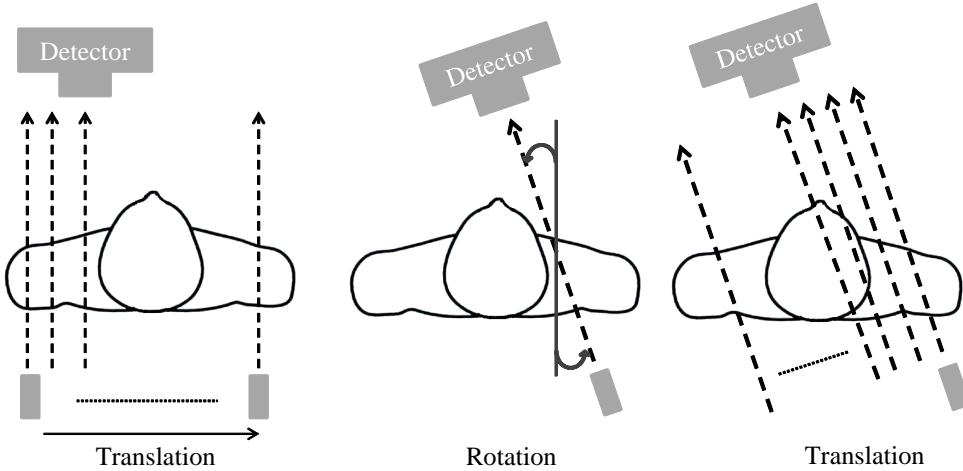


Figure 2.2: First Generation of CT scanners which completed the scan in two repeated translation and rotation steps. Average scan time was about 6 min for one 360° rotation.

and scan time was in the order of 5 to 6 minutes. As a result, patient motion was still a serious challenge in this generation. A second generation scanner with 3 narrow beams and 3 detectors is shown in Figure 2.3 (20 or more narrow beams and detectors were usually used in this generation).

Data is acquired via simultaneous measurements at each point along a translation. If there is a 1° angle between each of the 3 narrow beams, then one translation collects data for views at, for example, 30° , 31° , and 32° . As a result, only one third of the translations are required, and a 3-fold scan time reduction would be achieved. In fact, the second generation CT systems reduced the scan time by a factor of $\frac{1}{n_d}$, where n_d is the number of detectors. Further speed improvements were limited by the mechanical complexity of rotate and translate geometry. The mechanical tolerances and complexities involved indicated the need to eliminate translation motion.

2.1.3 Third Generation

Faster scans required the elimination of translation motion. This was achieved in third generation CT scanners that utilized a pure rotational motion by widening the X-ray beam into a fan beam to cover the entire patient width with an array of detectors, as shown in figure 2.4. To have a sufficient number of measurements a large number of detector elements were used. The early third generation CT scanners, installed in late 1975, could scan in less than 5 sec. This generation, which is still used in the modern scanners, decreased the scan time to about 0.3 sec: modern slip ring design enables data to be transferred while the gantry is rotating.

The most important problem in the third generation is called ring artifact. First and second

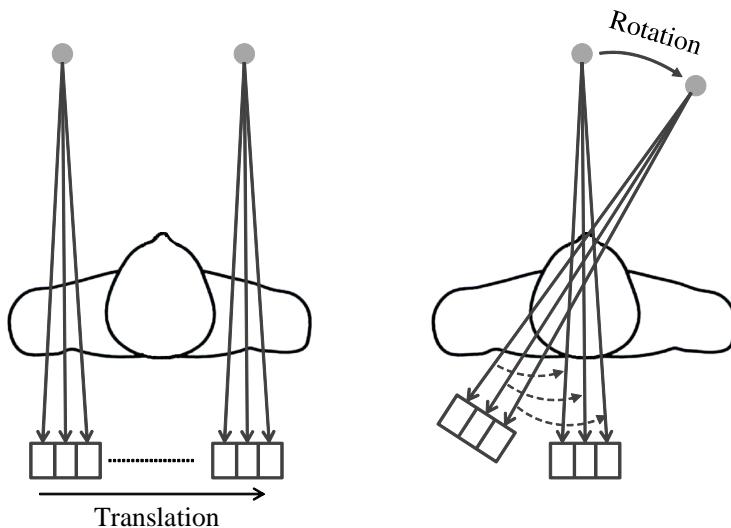


Figure 2.3: Second Generation of CT scanners which completed the scan task in two simultaneously repeated translation and rotation steps.

generation detectors were dynamically recalibrated at the beginning of each translation. In addition, each detector measures the rays passing through all voxels. Therefore, any detector error or drift was spread across the image evenly and was not visible. However, since in third generation CT X-ray source and detectors move together, each detector measures rays passing only at a specific distance from the center of rotation, depending on the location of the detector in the array. As a result, any error or drift in the calibration of a detector is back projected along a ring, creating a ring artifact, as illustrated in figure 2.4-B. This problem has never been solved, but has been corrected by image processing algorithms in modern scanners.

2.1.4 Fourth Generation

The fourth generation was designed to address the ring artifact in the third generation. It used a ring of stationary detector elements, as shown in figure 2.5. Unlike third generation detectors, each fourth generation detector can measure rays at any distance from the center of rotation and can be dynamically calibrated, so that ring artifact is not a problem anymore. However, many detectors were needed in this generation, which significantly increased the manufacturing cost.

2.1.5 Electron Beam CT (EBCT)

Cardiac CT requires ultrafast scans to minimize the cardiac motion. Electron Beam CT (EBCT) was designed in 1984 to address this problem. To accelerate the scanning process in EBCT, the mechanical motion is replaced by an electron beam which is swept along an arc of 360° so

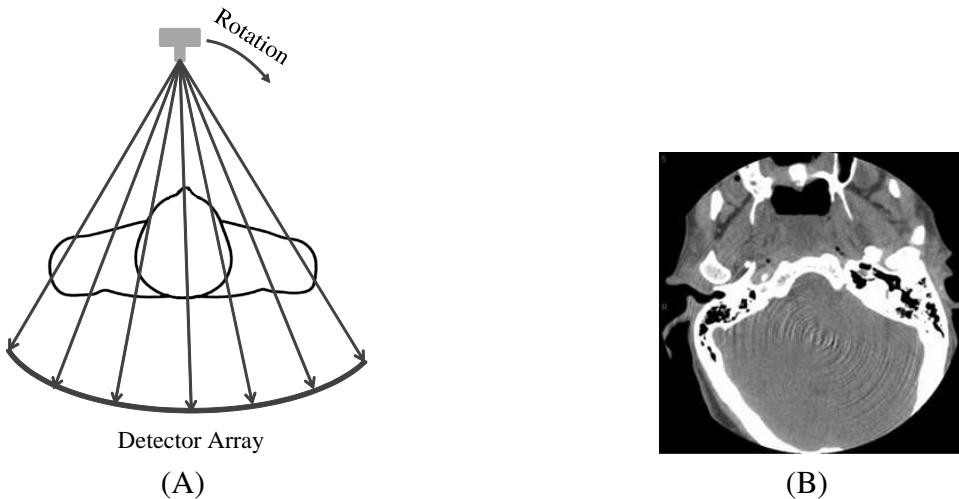


Figure 2.4: (A) Third Generation of CT scanners which includes a rotation step only. (B) Ring Artifact caused by problematic detector elements or uncalibrated detectors.

as to create an X-ray beam without mechanical movement, enabling a full set of projections to be achieved in around 10-20 ms. However, because of the low mAs values (number of emitted electrons), the quality of images in EBCT is usually less than the conventional CT scanners and has been replaced by multi-slice scanners.

2.1.6 Modern CT scanners

Most modern CT scanners are based on the 3rd generation scanners, but with multi-slice detectors (MDCT). Slip rings were introduced to increase the temporal resolution and to solve the problem with inter-scan delay caused by the cable connections after 360 degrees rotation. Slip ring passes the power to the X-ray source and detectors by brushes and the data is transmitted from detectors via wireless technologies allowing continuous rotation to occur with much reduced interference. Since slip ring technology allows the complete elimination of inter-scan delays, a new scan strategy was made possible, called Helical or Spiral CT scan to continuously acquire data as the table moves through the gantry, as shown in figure 2.7.

Helical scan CT systems are the standard body CT imaging since the mid-1990's. An important factor in helical CT is "Pitch Factor" which shows how fast the table slides through the gantry relative to the rotation time and slice thicknesses. Pitch factor is defined as the table movement per rotation divided by the slice thickness (e.g., if the slice thickness is 10 mm and the table moves 10 mm during one tube rotation, then the Pitch = $10/10=1.0$).

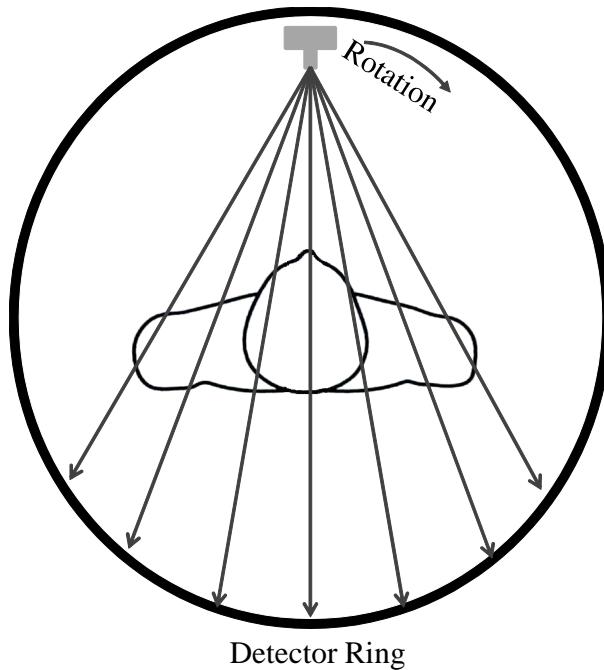


Figure 2.5: Fourth Generation of CT scanners with a 360° detector ring.

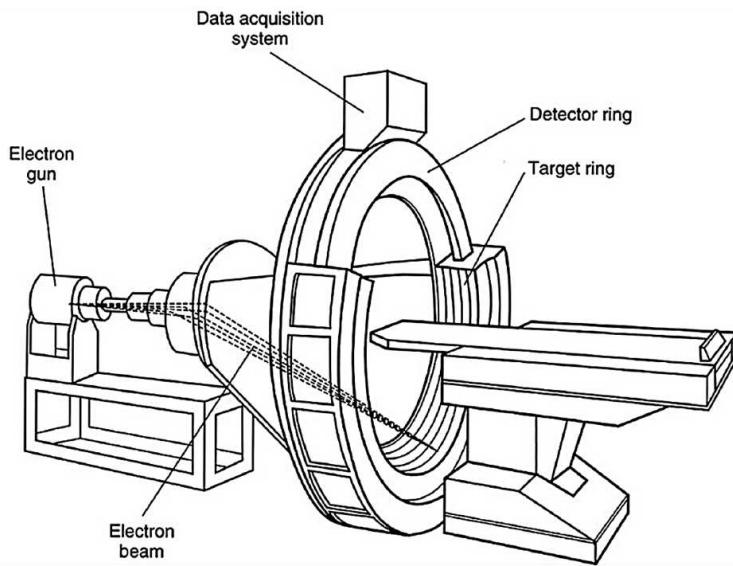


Figure 2.6: Schematic of an Electron Beam CT (EBCT). Reproduced, with permission, from [39].

2.2 Single Step CT imaging Reconstruction

Filtered Back Projection (FBP) is widely used in contemporary CT scanners, which is based on the Radon transform introduced in 1917 [42]. In this section the back projection based reconstruction algorithms for different geometries are presented beside some alternative single

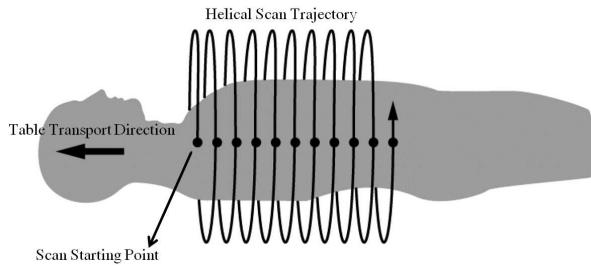


Figure 2.7: Helical (Spiral) CT trajectory.

step methods, which are closely related to back projection and Radon transform.

2.2.1 Filtered Back Projection of Parallel and Fan beam CT Reconstruction

X-ray projections of the parallel beam CT scanners, including first and second generations, can be expressed as the Radon transform of the object. The Radon transform is defined as [42]:

$$g(l, \varphi) = \mathcal{R}(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \varphi + y \sin \varphi - l) dx dy, \quad (2.1)$$

which is the integral along a ray at angle φ and at the distance l from the origin, $\delta(x, y)$ is Dirac delta function, and $f(x, y)$ is the image attenuation at (x, y) . However, this is not what the scanners directly measure. A wide variety of X-ray detectors are available, some counting single photons, some providing measurements of count rate or total flux, others measuring the energy, position, and/or incidence time of each X-ray photon. The statistical properties of the signal and noise depend on the type of detector. For instance, photon counting detectors measure the number of photons that hit the detector, $\xi(l, \varphi)$, which can be modeled by Poisson distribution with an expected value of $\bar{\xi}(l, \varphi)$ [43, 14]. The relation between $g(l, \varphi)$ and $\bar{\xi}(l, \varphi)$ is

$$g(l, \varphi) = -\log\left(\frac{\bar{\xi}(l, \varphi)}{\xi_T}\right), \quad (2.2)$$

where ξ_T is the number of radiated photons from the X-ray source. To recover the scanned object from the acquired X-ray projections the inverse problem of equation (2.1), known as inverse Radon transform or back projection, should be solved, i.e.,

$$f(x, y) = \int_0^\pi g(x \cos \varphi + y \sin \varphi, \varphi) d\varphi. \quad (2.3)$$

Back projection is equivalent to calculating the summation of all the rays that pass through a single pixel at different angles. While back projection is conceptually simple, it does not

correctly solve the inverse Radon transform problem. As shown in figure 2.8, a back projected image is very blurry. In more formal terms, the point spread function of back projection is circularly symmetric, and decreases as the reciprocal of its radius. To correct for this, the projections are first filtered in the Fourier domain by a ramp filter and then the filtered projections are back projected. This process can be expressed by:

$$\begin{aligned}
 f(x, y) &= \mathcal{F}_{2D}^{-1}(F(\omega_x, \omega_y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\omega_x, \omega_y) e^{+j2\pi(\omega_x \times x + \omega_y \times y)} dk_x dk_y \\
 f(x, y) &= \int_0^{2\pi} \int_0^{\infty} F(l \cos \varphi, l \sin \varphi) e^{+j2\pi l(x \cos \varphi + y \sin \varphi)} l dl d\varphi \\
 &= \int_0^{2\pi} \int_0^{\infty} G(l, \varphi) e^{+j2\pi l(x \cos \varphi + y \sin \varphi)} l dl d\varphi \\
 &= \int_0^{\pi} \left[\int_{-\infty}^{\infty} |l| G(l, \varphi) e^{+j2\pi l \rho} dl \right]_{\rho=x \cos \varphi + y \sin \varphi} d\varphi \\
 &= \int_0^{\pi} \mathcal{F}_{1D}^{-1}\{|l| G(l, \varphi)\}_{\rho=x \cos \varphi + y \sin \varphi} d\varphi,
 \end{aligned} \tag{2.4}$$

in which $|l|$ characterizes the ramp filter in the frequency domain, $G(l, \varphi)$ is the 1D Fourier coefficients of the projections $g(l, \varphi)$, $F(\omega_x, \omega_y)$ is the 2D Fourier transform of the image, \mathcal{F}_{2D}^{-1} is the inverse 2D Fourier transform, and \mathcal{F}_{1D}^{-1} is the inverse 1D Fourier transform.

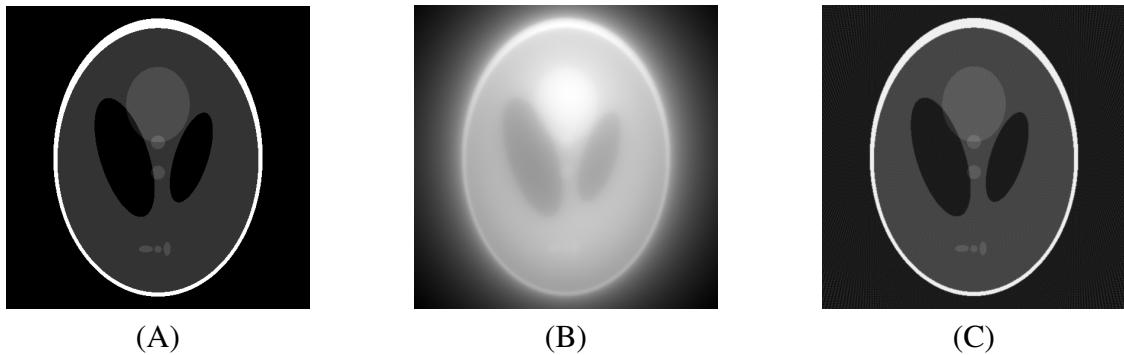


Figure 2.8: (A) Shepp-Logan phantom, (B) reconstructed by Back Projection, and (C) with Filtered Back Projection.

For the fan beam geometry shown in figure 2.9, reconstruction of the CT images from the projections $\mathcal{R}(\gamma, \beta)$ is very similar to the parallel FBP reconstruction given by equation (2.4). In this geometry each ray at a projection angle β has a different length, which is a function of the angle between the ray and the central beam. To correct this, each detector signal at a position γ must be scaled by $R \cos \gamma$, so that the back projection is expressed by

$$f(x, y) = \frac{1}{2} \int_{-\gamma}^{2\pi - \gamma} \int_{-\gamma_T}^{\gamma_T} \mathcal{R}(\gamma, \beta) c(r \cos(\beta + \gamma - \varphi) - R \sin(\gamma)) R \cos(\gamma) d\gamma d\beta, \tag{2.5}$$

in which $c(\cdot)$ is the ramp filter in spatial domain, and $2\gamma_T$ is the maximum fan angle. Other geometrical parameters are shown in figure 2.9. When the projection data is complete and noiseless, filtered back projection is exact. However, in real applications data is always corrupted with noise, in addition, because of the digital nature of the systems, the number of samples are limited.

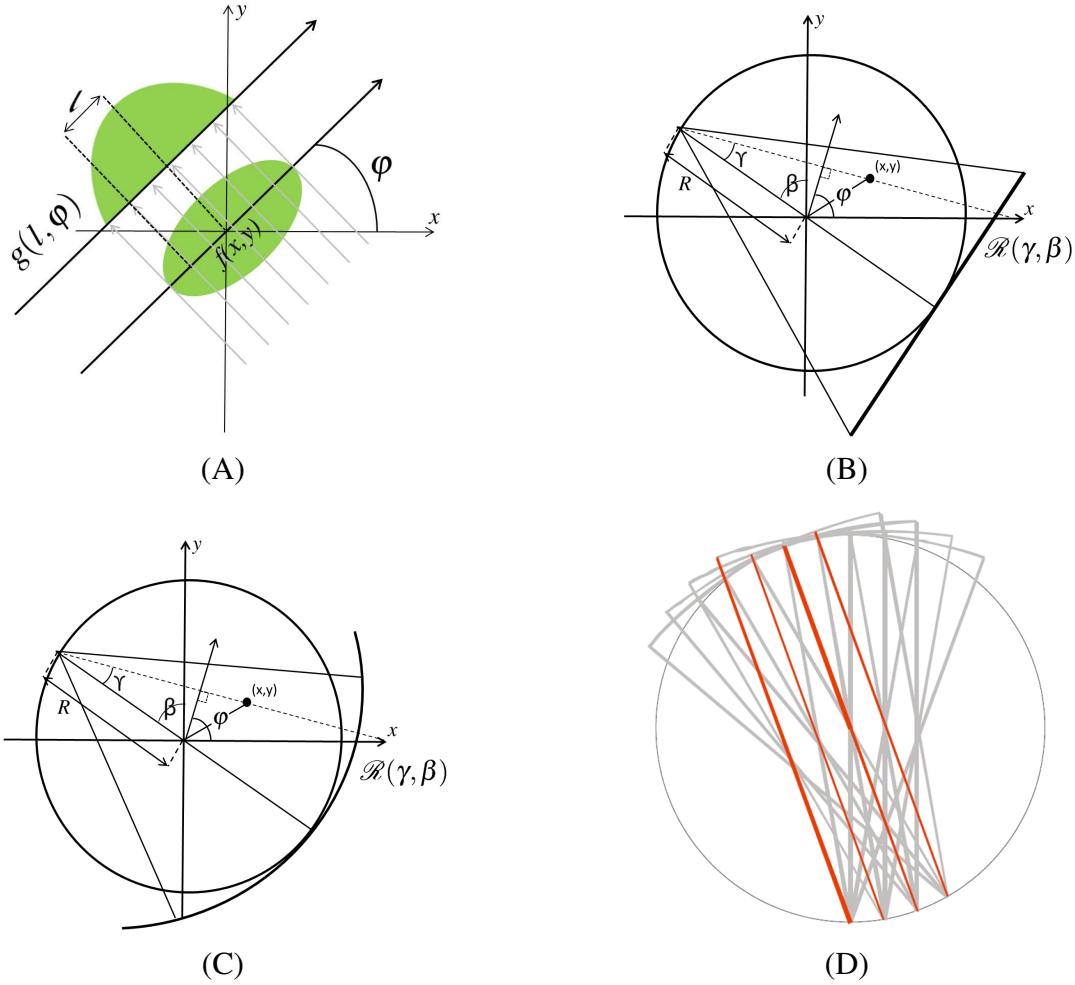


Figure 2.9: (A) Parallel beam geometry and (B) 3rd generation fan beam geometry with a flat detector (C) 3rd generation fan beam geometry with a curved detector, (D) Rebinning illustration.

2.2.2 Helical Cone Beam Reconstruction

Almost all the methods proposed for reconstruction of the helically scanned cone beam objects are based on converting the 3D data to 2D fan beam data and then reconstructing the images from these estimated fan beam projections [44, 45, 46, 47]. As an example of such methods,

Cone Beam Single Slice ReBinning (CB-SSRB) algorithms is briefly described here [46].

Sufficiency Condition for Cone beam Geometry

In 2D geometries, the sufficiency condition for an exact reconstruction is very simple: every projection line intersecting the object should contain at least one source point. However, for 3D geometries such as helical cone beam the sufficiency conditions are more complicated. Tuy [45] has shown that to have an exact reconstruction in cone beam and helical cone beam geometries every plane that intersects the object must contain at least one source point. These conditions may not be satisfied in real scanners, where the whole patient body is not covered with the X-ray beams. This problem is known as long object problem. Therefore, the helical cone beam reconstruction algorithms are usually inexact, but still accurate enough for obtaining high quality images.

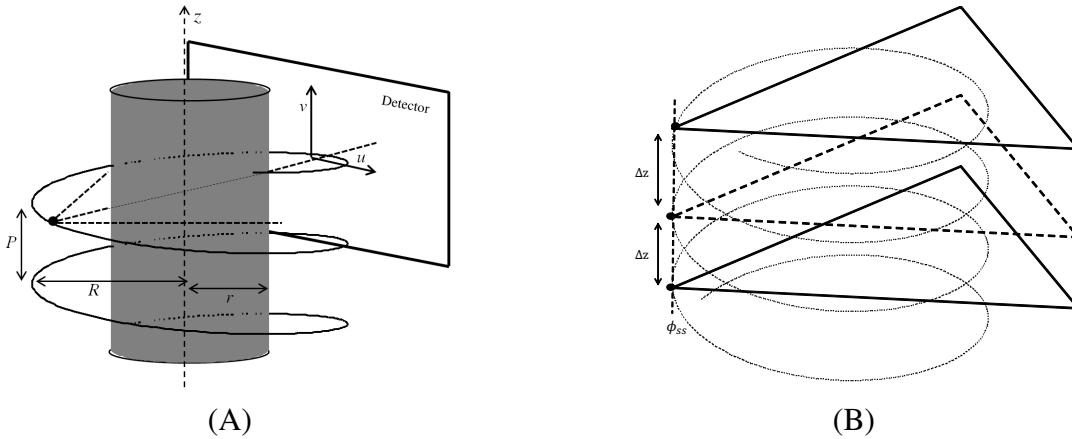


Figure 2.10: (A) Helical trajectory and (B) the fan beams in parallel z -slices.

Cone Beam Single Slice ReBinning(CB-SSRB)

Helical cone beam images are usually reconstructed by converting the helical cone beam projections to fan beam data estimated at different axial locations z . The CT slices are then reconstructed using fan beam reconstruction methods. Here, we use Cone Beam Single Slice ReBinning (CB-SSRB) proposed by Noo et al. [46] to estimate the fan beam projections. In this method, the fan beam data $p_z(\varphi, u)$ at an axial location z and at an azimuthal angle φ is estimated from

$$p_z(\varphi, u) \simeq \frac{\sqrt{u^2 + \mathcal{D}^2}}{\sqrt{u^2 + v^2 + \mathcal{D}^2}} g_\psi(u, v), \quad (2.6)$$

in which $v = (u^2 + \mathcal{D}^2) \Delta z / (R\mathcal{D})$. In these expressions $g_\psi(u, v)$ is the cone beam projections at helical angular position ψ , \mathcal{D} is the distance from the source to the detector, Δz is the signed distance between the cone and the desired fan beam sources, and u , v , and R are geometry parameters defined in figure 2.10. As also shown by Noo et al. each interpolated fan beam projection should be weighted by:

$$w(\phi_{ss}, u) = \begin{cases} \sin^2\left(\frac{\pi\phi_{ss}}{2(2\gamma_T+2\gamma)}\right) & \text{if } \phi_{ss} \in [0, 2\gamma_T + 2\gamma] \\ 1 & \text{if } \phi_{ss} \in [2\gamma_T + 2\gamma, \pi + 2\gamma] \\ \sin^2\left(\frac{\pi(\pi+2\gamma_T-\phi_{ss})}{2(2\gamma_T-2\gamma)}\right) & \text{if } \phi_{ss} \in [\pi + 2\gamma, \pi + 2\gamma_T] \end{cases}$$

where $\phi_{ss} = (\frac{\pi}{2} + \gamma_T)(1 - \frac{\Delta z}{P_d})$, $P_d = 0.5P(\pi/2 + \gamma_T)/(2\pi)$, P is the pitch of the helical trajectory, and $2\gamma_T$ is the maximum fan angle. The weighted fan beams $w(\phi_{ss}, u) \times p_z(\varphi, u)$ can then be used to reconstruct the image slices, by using the fan beam reconstruction methods.

2.3 Central Slice Theorem

To reduce the computational load of back projection (which is of $O(N^3)$ for an $N \times N$ image), Fourier based reconstruction methods have been proposed [48, 49, 50]. The Central Slice Theorem (CST) or Direct Fourier Reconstruction (DFR) relates the 1D Fourier transform of the projections to the 2D Fourier transform of the image, as shown in figure 2.11. In parallel beam geometry, the 1D Fourier transform of each projection (in the ideal case where the data is noiseless and continuous) is the same as the 2D Fourier coefficients of the object along a line that passes through the center of the frequency domain with the same angle as the corresponding projection, as can be seen in equation (2.4). Therefore, this can be used as an alternative to reconstruct the CT images by computing the 1D Fourier coefficients of the projection and putting them along the corresponding line in the 2D Fourier domain. This method requires the interpolation of polar data onto a Cartesian grid followed by an inverse 2D Fourier transform on the same grid to reconstruct the CT image.

If the projections were continuous with infinite number of measurements, DFR would be an exact reconstruction algorithm. However, in reality the scan continuity condition is not possible due to a limited number of projections. Thus, to reconstruct a perfect image, the missing coefficients must be interpolated. This is a source for additional artifacts, which needs an over sampled projection to decrease its effect. However, oversampled projections means that the patient is exposed to a higher radiation dose.

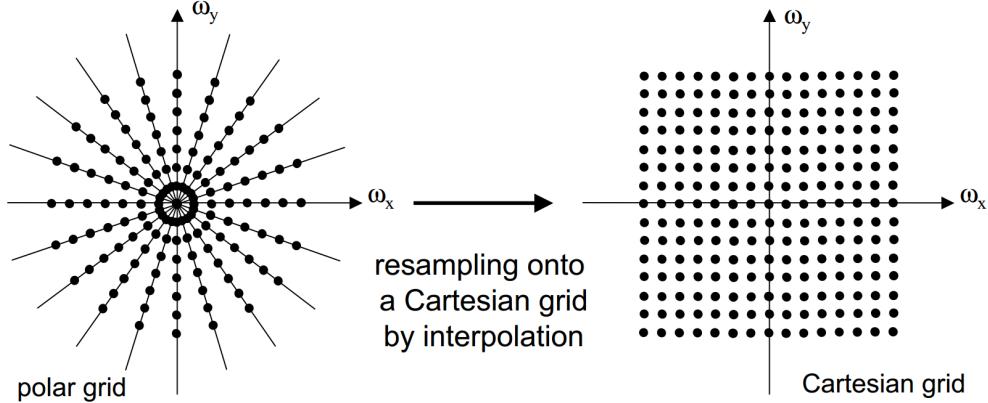


Figure 2.11: Central Slice Theorem and polar to Cartesian interpolation.

2.3.1 Pseudo-Polar Fourier Transform

To address the interpolation problem in DFR based methods, an Equally Sloped Tomography (EST) method as originally proposed for Electron Microscopy [51, 52, 53] can be used. This is an iterative method that makes use of the pseudo-polar Fourier transform (PPFT) [54]. The PPFT computes the Fourier coefficients directly on pseudo-polar grids, which can be described by

$$\begin{aligned} BV &\equiv \{\omega_y = \frac{\pi l}{N} \text{ for } -N \leq l < N, \omega_x = \omega_y \cdot \frac{2m}{N} \text{ for } -N/2 \leq m < N/2\} \\ BH &\equiv \{\omega_x = \frac{\pi l}{N} \text{ for } -N \leq l < N, \omega_y = \omega_x \cdot \frac{2m}{N} \text{ for } -N/2 < m \leq N/2\}. \end{aligned} \quad (2.7)$$

On the BV grids the Fourier transform is given by

$$F(\omega_x, \omega_y) = \sum_{i_1=0}^{N-1} \tilde{f}_1[i_1, l] \exp\left(-\frac{i2\pi i_1 m}{N} \cdot \frac{l}{N}\right), \quad (2.8)$$

where $\tilde{f}_1[i_1, l] = \sum_{i_2=0}^{2N-1} f_Z[i_1, i_2] (-1)^{i_2} \exp\left(-\frac{i2\pi i_2 l}{2N}\right)$ is the 1D Fourier transform of the zero-padded columns of the image (f_Z). In fact, the same equation can be written for BH by applying the same equation on rows rather than columns. Equation (2.8) can be interpreted as the fractional Fourier transform of the 1D Fourier transform of the zero-padded columns of the image weighted by $(-1)^{i_2}$. To reconstruct an $N \times N$ image from its PPFT coefficients, $4N^2$ samples are needed: $2N$ samples on $2N$ equally sloped radial lines that are at either basically horizontal projection angles of φ_{BH} or basically vertical projection angles of φ_{BV} , as illustrated

in figure 2.12. These are given by

$$\begin{aligned}\varphi_{BH} &= \tan^{-1}(2m/N), \quad -N/2 \leq m < N/2, \\ \varphi_{BV} &= \tan^{-1}(2m/N + \pi/2), \quad -N/2 \leq m < N/2.\end{aligned}\quad (2.9)$$

A fast algorithm has been proposed by Averbuch et al. [54] to calculate the PPFT and its

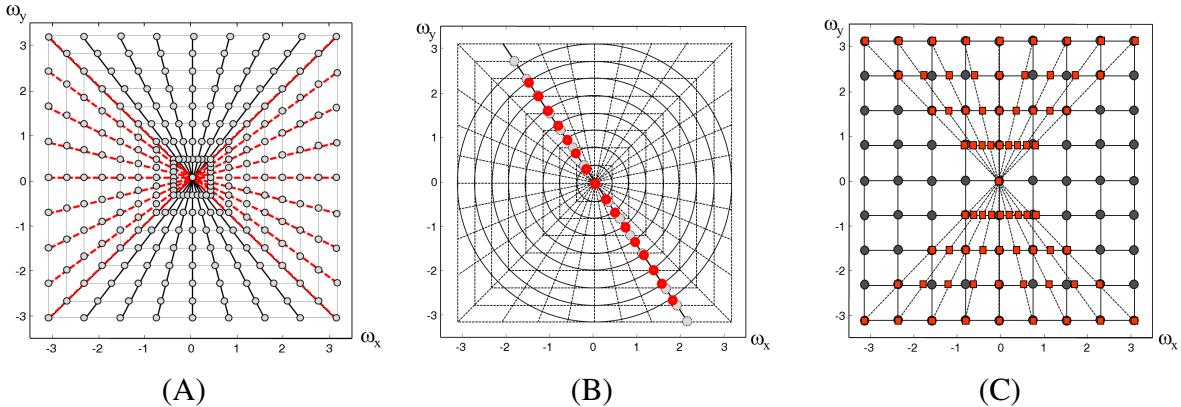


Figure 2.12: (A) Pseudo-polar grids: red lines are Basically Horizontal (BH) and the black lines are Basically Vertical (BV). (B) Polar Grids (red dots) on the pseudo-polar grids (gray dots).

adjoint. To make use of this for CT image reconstruction, the projection rays should be measured or interpolated on the equally slopped lines, as shown in figure 2.12-A, and the PPFT algorithm can be used to reconstruct the CT images. Three important properties of the PPFT make it a good alternative to the classic DFR methods: 1) it is closer to polar (equiangular line) grids compared to Cartesian grids, which decreases the interpolation error significantly, figure 2.12-B and 2.12-C, 2) it can be computed with a fast algorithm proposed in [54], and 3) unlike interpolating the polar data on a Cartesian grid in regular DFR methods, it has an analytical adjoint function, that makes it usable in iterative applications. However, it should be noted that DFR algorithms, including EST, are only valid for parallel projection beam geometries.

2.4 Rebinning Process

An alternative method to reconstruct the fan beam images is to rearrange the fan beam rays to parallel beams and then use equation (2.4) to reconstruct the images: a process called rebinning [55, 46]. In addition, rebinning enables the use of central slice theorem and direct Fourier reconstruction in fan beam geometry. This step makes use of a geometric relationship between fan and parallel beams that is illustrated in figure 2.9-D. Specifically, if $\mathcal{R}(\gamma, \beta)$ is the fan beam

projected data and $g(R \sin \gamma, \beta + \gamma)$ is the corresponding rebinned parallel ray, then

$$\mathcal{R}(\gamma, \beta) = g(R \sin \gamma, \beta + \gamma), l = R \sin \gamma, \varphi = \beta + \gamma, \quad (2.10)$$

where γ , R , φ and β are geometric parameters defined in this figure.

2.5 Iterative Reconstruction

The quality of the images reconstructed by standard FBP algorithms depend on several assumptions about scanner geometry and raw data, such as continuity of the projections and noiseless measurements, which are typically violated. The FBP method involves a compromise between high reconstruction speed and image quality. To achieve better image quality from the same raw data, more realistic assumptions about scanner geometry and noise statistics must be made. This is done in the more computationally complex iterative reconstruction methods. Such iterative reconstruction methods may result in longer reconstruction times but also in substantially less image noise from the same raw data through more complex modeling of detector response and the statistical behavior of the measurements. Simplifications of iterative methods are used by CT vendors (e.g. AIDR3D from Toshiba, and ASiR from GE) based on iterative application of FBP and smoothing functions that shorten the reconstruction time, while maintaining much lower image noise than single step FBP with no significant impact on spatial or contrast resolution [56].

In this section two classical iterative algorithms, Algebraic Reconstruction Techniques (ART) and Statistical Iterative Reconstructions (SIR), are briefly described.

2.5.1 Algebraic Reconstruction Technique (ART)

Algebraic reconstruction is based on the model shown in figure 2.13 [6, 12, 13], in which the problem is to estimate the attenuation value of the pixels f_j from the measured projections g_i :

$$g_i = \sum_{j=1}^N w_{i,j} f_j, \quad i = 1, \dots, M \quad (2.11)$$

in which M is the total number of rays (in all projections) and $w_{i,j}$ is the weighting factor that represents the contribution of the j^{th} pixel to the i^{th} ray integral. In other words $w_{i,j}$ is equal to the fractional area of the j^{th} image pixel covered by the i^{th} ray. When the image size is very large (e.g. 512×512), the computational task of finding the inverse matrix and calculating the the pixel values is daunting. As a result, iterative methods are used to estimate the pixel values

(2.11) and thereby estimate the CT image. The iterative procedure can proceed as follows: knowing the $w_{i,j}$'s, in the $l + 1^{th}$ iteration of ART, each pixel of the image is updated by:

$$\begin{aligned}\hat{f}_j^{l+1} &= \hat{f}_j^l + \left[\frac{g_i - \check{f}_i^l}{\sum_{k=1}^N w_{i,k}} \right] w_{i,j}, \quad j = 1, \dots, N^2, \\ \check{f}_i^l &= \sum_{k=1}^N \hat{f}_k^{l-1} w_{i,k} \quad \text{for } i = 1, \dots, M.\end{aligned}\quad (2.12)$$

Since the calculation of exact $w_{i,j}$ is not straight forward, in many ART implementations the $w_{i,j}$'s are simply replaced by 1 and 0, depending on whether the center of the k^{th} image cell is within the i^{th} ray or not [6]. It should be noted that ART methods are able to deal with the noisy and limited view image reconstruction, as the imaging geometry and prior information about the object can be incorporated into the reconstruction procedure.

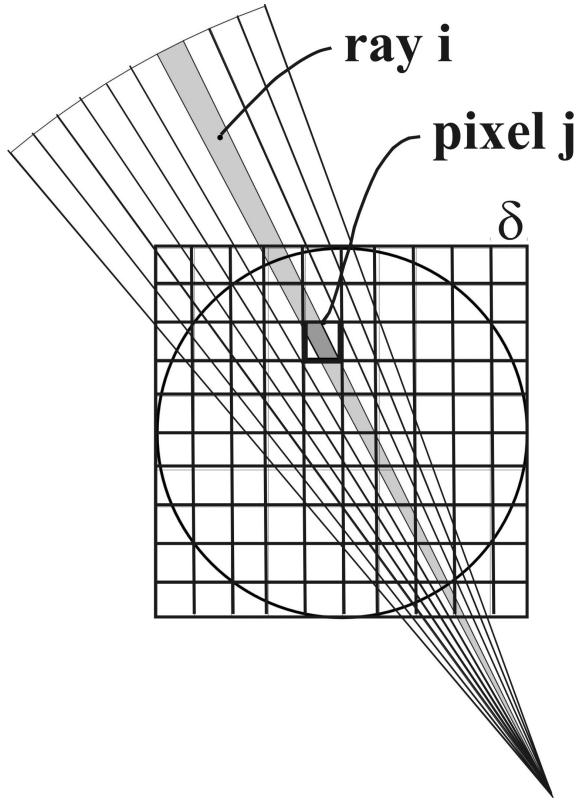


Figure 2.13: Projection matrix modeling in Algebraic Reconstruction Technique (ART).

2.5.2 Statistical Iterative Reconstruction Methods

Statistical methods assume a certain distribution of the measured photons and find an attenuation function that maintains the projection data as close as possible to the measured data

and at the same time follows the assumed prior knowledge (probabilistic distribution). The photon detection statistics are usually modeled with Poisson distribution. If the measured data is represented by $\xi = [\xi_1, \xi_2, \dots, \xi_M]$, in which ξ_i is a random variable characterizing the number of collected photons by i^{th} detector, based on Beer's law $E[\xi_i] = \xi_T e^{-\int_L f(x,y,z)dl}$ for $i = 1, 2, \dots, M$, where L defines the line along which the photons are attenuated with attenuation coefficients shown by f 's. To simplify the equations, the line integration are re-written as $\int_{L_i} f(x,y,z)dl = \sum_{j=1}^{N^2} a_{i,j} f_j = A \times f$ where $a_{i,j} = \int_{L_i} w(x,y,z)dl$, and w shows the fraction of each pixel (voxel) that affects the line integration. With above notations, the maximum likelihood (ML) estimate of the attenuations can be written as [15]:

$$\begin{aligned}\hat{f} &= \underset{f}{\operatorname{argmax}} \log P(E[\xi] = \xi; f) \\ P[E[\xi] = \xi; f] &= \prod_{i=1}^M P[E[\xi]_i = \xi_i; f] = \prod_{i=1}^M \frac{e^{-E[\xi]_i(f)} (E[\xi]_i(f))^{\xi_i}}{\xi_i!}\end{aligned}\quad (2.13)$$

in which $E[\xi]_i$ is the expected number of photons. Since we deal with large matrices, the computational load of image reconstruction with SIR algorithms is very high. In $l + 1^{th}$ iteration SIR estimates the attenuation of the j^{th} pixel by means of

$$\hat{f}_j^{l+1} = \frac{\hat{f}_j^l}{\sum_{j=1}^{N^2} a_{i,j}} \sum_{i=1}^M \frac{\xi_i a_{i,j}}{\sum_{j=1}^{N^2} \hat{f}_j^l a_{i,j}}. \quad (2.14)$$

Generally, there is no unique solution to the problem presented by equation (2.13), especially when the problem is ill-posted, i.e., not enough samples are available. Different methods have been proposed to increase the quality of reconstruction and to decrease the reconstruction time by incorporating additional penalties [57] or by simplifying the optimization problem. One method that may increase the reconstruction quality and the convergence speed is to use second order Taylor's expansion of equation (2.13) to make it a quadratic problem [58]. One of the best approaches proposed to decrease the computational complexity of SIR is through the Ordered Subset Expectation Maximization (OSEM) [59] method. OSEM divides the projections into different subsets: each subset can reconstruct the image with smaller number of projections and its likelihood can be maximized for each iteration independent from other subsets.

2.6 X-ray Dose measurement

In this section we briefly review different definitions, units, and measurement methods of X-ray dose in CT imaging. Dose is a general term used to quantify how much radiation exposure a

person (or other material) has received. Radiation dose quantities are described in three ways: absorbed, equivalent, and effective. The amount of energy deposited in a substance is called the absorbed dose. The absorbed dose is measured in a unit called the gray (Gy). A dose of one gray is equivalent to a unit of energy (Joule) deposited in a kilogram of a substance. Equal absorbed doses will not necessarily produce equal biological effects in a living matter. In fact, the effect depends on the type of the radiation (alpha, beta, gamma, etc.) and also the tissue that receives the radiation. In order to be able to compare the biological effect of different types of radiations, a weighting factor is used. The weighted absorbed quantity is then called the equivalent dose and is measured in a unit called Sievert (Sv). Hence, 1 mSv of any type of radiation will have the same biological effectiveness. Another quantity that describes the dose is the effective dose. Different tissues and organs in the body have different sensitivity to radiation. To achieve an indication of how exposure affects each tissue, the equivalent dose is multiplied by a conversion factor, which is related to the weight and sensitivity of a particular tissue or organ, giving the effective dose in units of Sv.

The entrance skin dose is usually measured using an ionization chamber. Such a chamber is the simplest of all gas-filled radiation detectors, and is widely used for the detection and measurement of ionizing radiation (X-ray, gamma, beta, etc.). It is a solid envelope surrounding a gas (usually air) filled cavity in which an electric field is established to collect the ions formed by the radiation. These devices usually measure air kerma. Air kerma is the energy released per unit mass of a small volume of air when an X-ray beam irradiates it and the resulting dose is normally expressed in mR. To compare the doses of X-ray CT examinations, the dose measure should be converted to mSv. The Roentgen (R, also Röntgen) is a legacy unit of measurement for the kerma of X-rays [43]. It is named after the German physicist Wilhelm Conrad Röntgen, the man who discovered the X-ray; an achievement that earned him the first Nobel Prize in Physics in 1901. One Röntgen is the amount of radiation that will produce, under normal conditions of pressure, temperature, and humidity, in 1 kg of air, an amount of positive or negative ionization charge equal to 2.58×10^{-4} coulomb.

One Röntgen of air kerma deposits 0.00877 Gy (0.877 rad) of absorbed dose in dry air. Hence, in order to convert the measured dose (mR) from the device to mGy, we should multiply it by 0.0087. After converting the dose to mGy, a conversion factor must be used to obtain the effective dose in mSv (e.g., the conversion coefficients published by the National Radiological Protection Board (NRPB-R262) [60]).

The computed tomography dose index (CTDI) is a commonly used radiation exposure index in CT. Several variations of the CTDI have been defined. For example, the CTDI₁₀₀ reflects the dose contribution from a 100-mm range centered on the index slice, measured by ionization chamber. The weighted CTDI (CTDI_w in mGy) is the average radiation dose in scan volume

measured in standard CT phantoms. It reflects the weighted sum of two-thirds peripheral dose and one-third central dose in a 100-mm range in a cylindrical acrylic phantom (either 11 cm in diameter for a head phantom or 21 cm in diameter for a body phantom). CTDIw is directly proportional to radiation dose in unit volume, influenced by pitch factor. To make the dose independent from pitch, CTDIw is defined as CTDIvol divided by pitch. Dose-length product or DLP (mGy.cm) is equal to the product of CTDIvol and the scan length. It is directly proportional to total scanned radiation dose.

2.6.1 Clinical Dose Reduction Strategies

Some clinically used strategies for dose optimization will be briefly discussed in this subsection. It is notable that for dose optimization, the exam must be performed using doses that are As Low As Reasonably Achievable (ALARA), consistent with the diagnostic task at hand [61]. A better understanding of possible strategies can be obtained by examining the following CT dose parameters:

1. Tube voltage (kV) is the potential difference between cathode and anode of the X-ray tube. The radiation dose is proportional to square of tube voltage.
2. Current-rotation time product (mAs) represents the photon fluence, which is the number of photons that enter an imaginary sphere, with a cross-sectional area of 1 m^2 , and is directly proportional to the radiation dose.
3. Pitch is the ratio of table feed per gantry rotation to nominal width of beam collimation and it is inversely proportional to radiation dose. It is notable that an increase in the pitch decreases the duration of radiation exposure.

A fundamental part of CT dose optimization is based on the use of body size-adapted CT protocols. Minimal required radiation dose would be variable for different patients depending on body size and habitus. Hence, the optimal tube voltage and current should be determined for the adapted radiation dose. There have been different studies on finding the body size-adapted CT protocols for children and adults [62, 63]. In these studies, body mass index or body weight has been applied to the adaptive protocols. However, researchers have found that cross-sectional dimensions can be better used to find the body size-adapted CT protocols [64, 65, 66, 67]. It is notable that parameters other than body weight or body mass index have not been commonly used for CT protocols. The reason is the difficulty in clinical implementation of them.

Another strategy for dose reduction is tube current modulation. The basic idea is to adapt the tube current to attenuation of the body region. This means to increase the current for a more

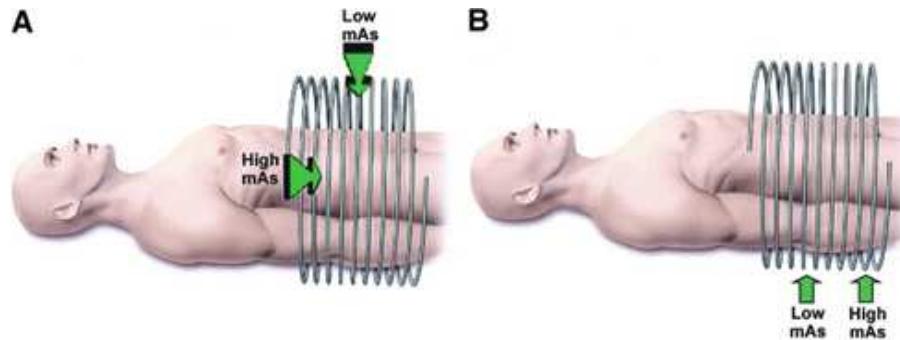


Figure 2.14: (A) Angular modulation in which the x-ray tube current is lowered as the X-rays are aimed in the anteroposterior directions, and increased when the X-rays are aimed in the lateral-medial directions, when there will be more X-ray attenuation. (B) Z-axis modulation in which, for example, fewer X-rays are required in the abdominal region superior to the pelvic bones compared with the pelvic region. Reproduced, with permission, from [68].

attenuating area and decrease it for a less attenuating area. The overall goal is to reduce dose while maintaining the image quality. The tube current may be adjusted in the x-y plane, the z-axis, or a combination of both. Figure 2.14 shows principles of tube current modulation [68]. In addition, a graphic illustration of tube current modulation is shown in figure 2.15 [69]. With other scanning parameters held constant, a reduction in tube current decreases radiation exposure but increases image noise or mottle, a principle determinant of image quality. The use of tube current modulation has helped in CT dose reduction of up to 15-50% in children and adults [70, 71].

In addition to the previous strategies, the importance of optimal tube voltage has been emphasized in recent studies [72, 73]. Figure 2.16 shows the linear attenuation coefficients of iodine, cortical bone and water as a function of X-ray energy [72]. The graph shows that the attenuation coefficient is different for different materials and depends on the X-ray energy. This affects image contrast, meaning that at lower tube voltages the contrast increases. However, lower tube voltages have some disadvantages including higher noise in the image. At higher tube voltages the photon penetration will be increased which results in having less severe artifacts from metallic objects or thick bones such as the skull base. Hence, there is a trade off between image noise and contrast enhancement. Therefore, the optimal tube voltage should be determined for patient size and each type of CT examination to achieve an optimal trade off between contrast and noise [72, 73].

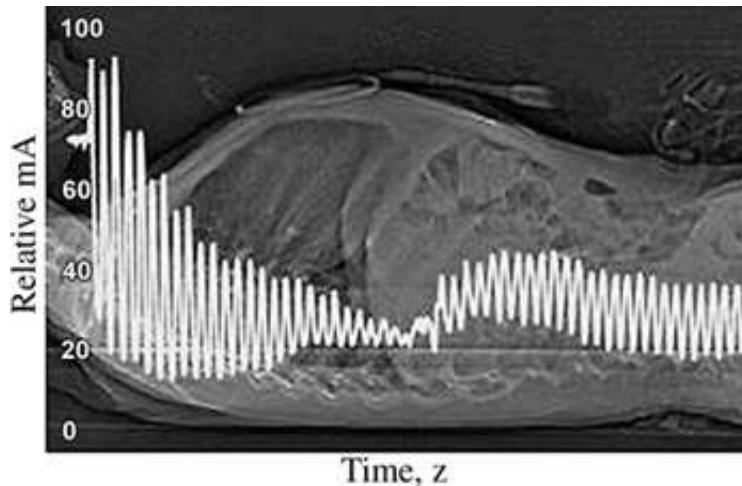


Figure 2.15: Graph of tube current (mA) superimposed on a CT projection radiograph showing the variation in tube current as a function of time (and, hence, table position along the z-axis) at spiral CT in a 6-year-old child. An adult scanning protocol and an AEC system (CareDose 4D; Siemens Medical Solutions) were used with a reference effective tube current-time product of 165 mAs. The mean effective tube current-time product for actual scanning was 38 mAs (effective tube current-time product-tube current-time product/pitch). Reproduced, with permission, from [69].

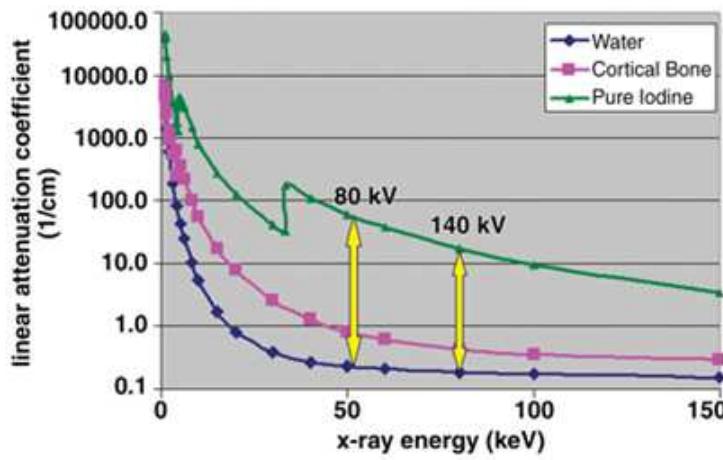


Figure 2.16: The linear attenuation coefficients of iodine, cortical bone and water as a function of X-ray energy. The yellow arrows show the contrast between water and iodine in 80kV and 140kV, which are more distinct in 80kV. Reproduced, with permission, from [72].

Chapter 3

Compressed Sensing in Computed Tomography

The introduction of Compressed Sensing (CS) in 2006 provided a means of avoiding the limitations imposed by the Nyquist sampling theorem. It has been widely used in medical image processing by making use of the fact that many signals can be accurately represented by only a few non-zero coefficients in a suitable basis or dictionary. Nonlinear optimization enables the recovery of such signals from few measurements. Based on this idea a new class of low dose CT imaging methods have been developed enabling high quality CT images to be reconstructed from a reduced number of projections, as illustrated in figure 3.1. Fan-beam CT scanners typically measure the X-ray projections from 800-2000 X-ray source positions that are distributed uniformly over the angular range of $180 < \varphi \leq 360$. The most recently proposed CS-based CT reconstruction techniques can reconstruct the CT images from about 1/10th of the number of projections needed by conventional methods for the same image quality [74]. Compressed sensing and its application in CT reconstruction are briefly described in this chapter.

3.1 Historical Background

The key idea of compressed sensing is to recover a sparse signal from few linear measurements by utilizing convex optimization algorithms. From a different viewpoint, it enables an exact recovery of high-dimensional sparse signals after a dimension reduction step. Sparsity of the signals has been used in compression and dimension reduction, before the introduction of CS, to address the computational challenges involved in dealing with high-dimensional data. The goal of the data compression algorithms is to find the most accurate representation of a signal within an acceptable level of distortion from a few samples. Transform coding is one

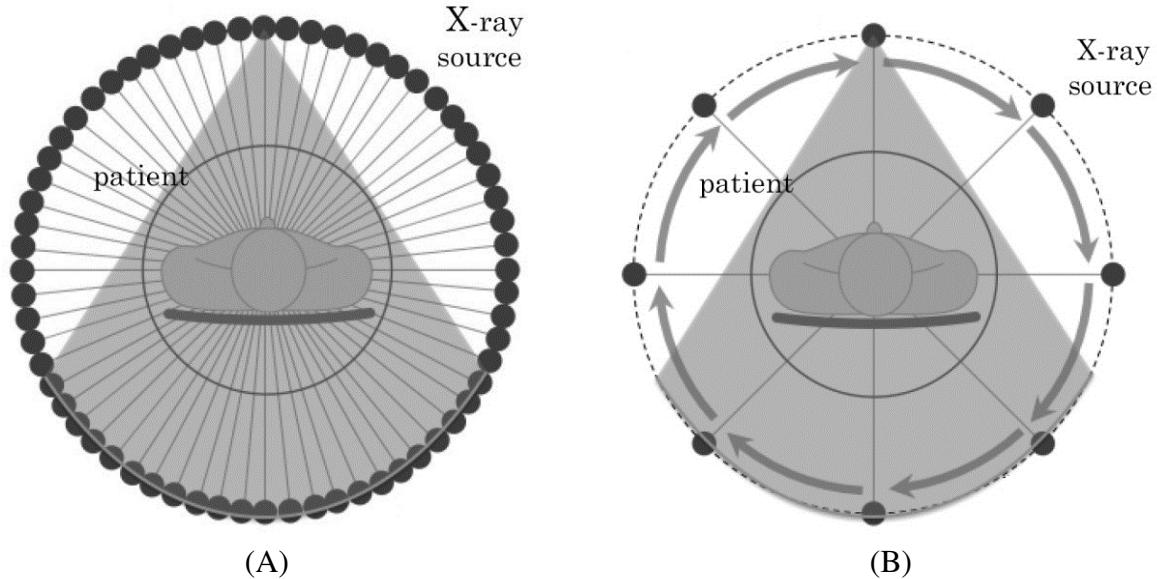


Figure 3.1: Radiation dose reduction in CS-based CT reconstruction methods: (A) Regular CT system with many equiangular X-ray projections, and (B) Sparse view CT with few X-ray projections. Reproduced, with permission, from [75].

of the most popular data compression techniques that relies on finding a basis that provides sparse or compressible representations for signals [76, 77]. A sparse signal of length N , can be represented exactly with k nonzero coefficients, where $k \ll N$. Compressible signals, on the other hand, are well-approximated by just the k largest coefficients, while the other coefficients can be non-zero but with very small values. Both sparse and compressible signals can be represented with a very small error by preserving only the values and locations of the k largest coefficients. This process, called the sparse approximation, forms the foundation of many compression schemes such as the JPEG, JPEG2000, MPEG, and MP3 standards.

While the transform coding techniques focus on the compression of the signal measured in many samples, i.e., with Nyquist sampling rate, and then removing the redundant unnecessary coefficients, compressed sensing introduces a new framework for signal acquisition. Rather than sampling at a high rate and then compressing the sampled data, CS directly senses the data in a compressed form, i.e., at a lower sampling rate. As a result, CS offers the potential for large reductions in the sampling cost of the signals that have sparse or compressible representation. While recovery of the sparse signals from few samples has recently attracted a considerable attention, particularly following the publications of Candès and Taro [17] and Donoho [18], it should be noted that this field existed before CS. For example ℓ_1 minimization was used in seismology by Claerbout and Muir [78] in 1973, as an alternative to ℓ_2 minimization as used in least squares error solutions. Also, application of ℓ_1 -norm to recover sparse spike trains was

proposed by Santosa and Symes [79] in 1986. Moreover, Rudin, Osher and Fatemi [80] used Total Variation (TV) minimization in image processing which is very close to the idea of ℓ_1 minimization.

3.2 Basic Concepts

We assume that $\mathbf{x} = \{\mathbf{x}_i, i = 1, \dots, N\} \in \mathbb{R}^N$ is a k -sparse vector, i.e., with only k non-zero elements. This means that its ℓ_0 -norm (which counts the number # of non-zero elements) is small and equal to k :

$$\|\mathbf{x}\|_0 = \sum_{i=1}^N |\mathbf{x}_i|^0 = \#\{i : \mathbf{x}_i \neq 0\}, \quad (3.1)$$

where

$$|\mathbf{x}_i|^0 = \begin{cases} 1, & \text{if } \mathbf{x}_i \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Using this notation, and considering the goal of CS to recover the signal \mathbf{x} from few samples \mathbf{y} , CS can be formulated as the Least Absolute Selection and Shrinkage Operator (LASSO formulation [81, 82]). Specifically, if \mathbf{x} is sparse,

$$\mathbf{y} = \mathbf{Ax}, \text{ such that } \|\mathbf{x}\|_0 = k. \quad (3.2)$$

If \mathbf{x} is not sparse, an orthonormal basis Φ may exist such that $\mathbf{x} = \Phi\theta$, with θ being sparse. In this case the LASSO formulation is:

$$\mathbf{y} = \mathbf{A}\Phi\theta, \text{ such that } \|\theta\|_0 = k, \quad (3.3)$$

in which, \mathbf{A} is an $m \times N$ sensing or measurement matrix (which may include the sparse representation dictionary), with $m \ll N$ and $m \approx k$, which makes this problem an under-determined linear system, whose sparsity provides the prior information needed to recover the vector. To recover the sparse signal from just a few samples, CS theory provides: signal and sparsity models, conditions for suitable sensing matrices, recovery algorithms, convergence conditions and provides information about the accuracy of the recovery and the number of required samples. It has been shown that [76, 83] the wavelet transform provides a good sparse or compressible representation of many types of images. Other sparsifying transforms have also been used, including: Curvelets [84], Shearlets [85], and noiselet [86]. As an example figure 3.2 shows a CT image and its wavelet subbands with many small and few large coefficients. In this section,

for simplicity and without loss of generality, it is assumed that \mathbf{x} is sparse, unless otherwise noted.

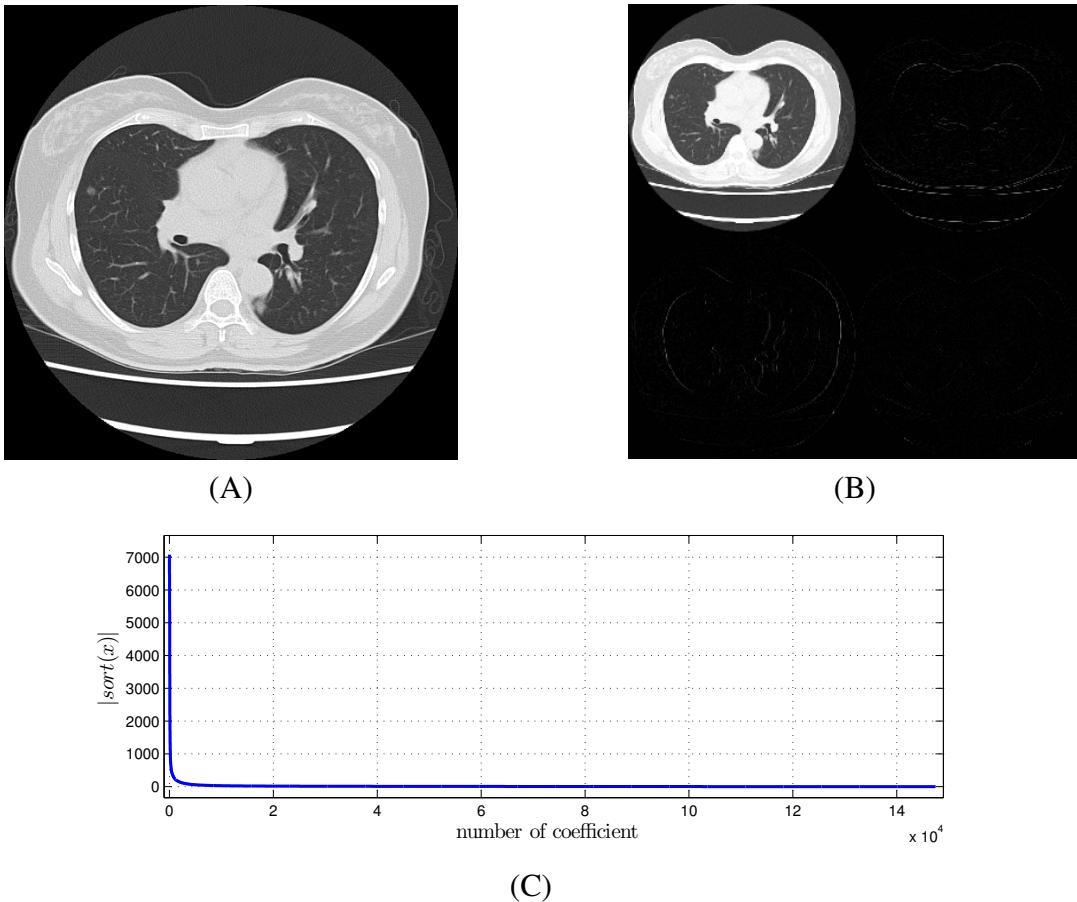


Figure 3.2: Sparse representation of a CT image with wavelet transform: (A) A chest CT image, (B) Subbands of the image with one level Haar wavelet, and (C) sorted absolute values of the wavelet coefficients of the image with 5 level Haar wavelet transform.

3.2.1 Optimization Methods

Compressed sensing estimates the sparsest signal that provides the best match between the measured data and the recovered signal. This sparsest signal can be found by solving

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_0, \text{ such that, } \mathbf{y} = \mathbf{Ax}. \quad (3.4)$$

which is closely related to LASSO formulation. However, this is a non-convex, NP-hard problem [76]. The key idea of compressed sensing is to substitute the ℓ_0 -norm with the closest convex ℓ_1 -norm. It can be shown that using ℓ_1 in the following formulation (called the Basis

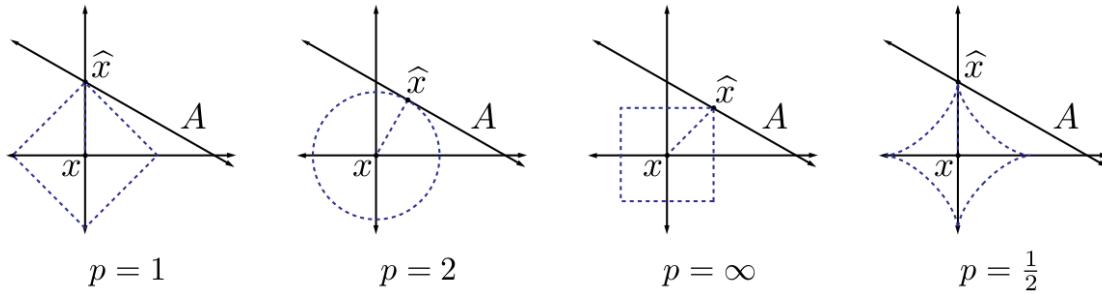


Figure 3.3: Comparison of the different ℓ_p norm minimization problems in sparse recovery problems. Reproduced, with permission, from [76].

Pursuit), the sparsest signal \mathbf{x} can be estimated from the noiseless measurements

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \text{ such that, } \mathbf{y} = \mathbf{Ax}. \quad (3.5)$$

In presence of the noise, i.e., $\mathbf{y} = \mathbf{Ax} + \mathbf{n}$, in which \mathbf{n} is a zero mean white Gaussian noise with variance of σ_n^2 , the basis pursuit problem can be expressed as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \text{ such that, } \|\mathbf{y} - \mathbf{Ax}\|_2^2 < \varepsilon. \quad (3.6)$$

CS theory provides the necessary and sufficient conditions on the measurement matrix \mathbf{A} and the number of needed samples to recover the signal exactly or with a limited error in the noisy or compressible cases. Figure 3.3 intuitively demonstrates the power of ℓ_1 and non-convex $\ell_p, 0 < p \leq 1$ norms, compared to ℓ_2 norm as used in conventional least squares error minimization algorithms, to pick the sparsest signals in \mathbb{R}^2 space (which is a point on one of the axes). In \mathbb{R}^2 , $\mathbf{y} = \mathbf{Ax}$ is a line and the solution of ℓ_p minimization is the point at which the ℓ_p ball intersects the line. As can be seen, because of the pointiness of the ℓ_p ball for $0 < p \leq 1$, the sparse signal is fully recovered, unlike ℓ_2 .

To solve the ℓ_1 minimization problem used in CS formulations (3.5) and (3.6), many different algorithms have been proposed. In section 3.4 some of these recovery methods are described.

3.3 Signal Measurement and Recovery Conditions

Although ℓ_1 minimization has been used in many other applications, such as deconvolution [87] and regression [88], in CS applications the system is typically under-determined, and therefore, tighter conditions are usually needed. One popular condition on the measurement matrix \mathbf{A} , which guarantees the sparse recovery of CS, is the restricted isometry property (RIP)

[89]. RIP of order k is satisfied if there exists a $\delta_k \in (0, 1)$, such that

$$(1 - \delta_k) \|\mathbf{x}_S\|_2^2 \leq \|\mathbf{A}_S \mathbf{x}_S\|_2^2 \leq (1 + \delta_k) \|\mathbf{x}_S\|_2^2, \quad (3.7)$$

in which S is a set of column indices corresponding to the k largest entries of \mathbf{x} , \mathbf{x}_S is a vector obtained by keeping only the elements corresponding to the k largest values of \mathbf{x} , and \mathbf{A}_S is a matrix with columns of \mathbf{A} chosen corresponding to S . In terms of eigenvalues of the matrix $\mathbf{A}^T \mathbf{A}$, with $.^T$ being the transpose of the matrix, RIP is equivalent to having all the eigenvalues in the range $[1 - \delta_k, 1 + \delta_k]$. In a non-mathematical language, RIP means that the measurement matrix \mathbf{A} preserves the norm of the vectors and the distances between the points in \mathbb{R}^N space. Therefore, if two points are far away in \mathbb{R}^N , their projections to \mathbb{R}^m by $\mathbf{A} \in \mathbb{R}^{N \times m}$ are not any closer.

Assuming $\hat{\mathbf{x}}$ to be the solution to (3.5) and $\delta_{2k} < \sqrt{2} - 1$, then the recovery error is [90]:

$$\begin{aligned} \|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 &\leq C_0 \|\mathbf{x} - \mathbf{x}_S\|_2^2 / \sqrt{k} \\ \|\hat{\mathbf{x}} - \mathbf{x}\|_1 &\leq C_0 \|\mathbf{x} - \mathbf{x}_S\|_1, \end{aligned} \quad (3.8)$$

for some constant C_0 . As a result, for a k -sparse signal ($\mathbf{x} = \mathbf{x}_S$) this recovery will be exact. On the other hand, if $\hat{\mathbf{x}}$ is the solution to (3.6) with $\delta_{2k} < \sqrt{2} - 1$, then the error is given by:

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 \leq C_0 \|\mathbf{x} - \mathbf{x}_S\|_2^2 / \sqrt{k} + C_1 \varepsilon \quad (3.9)$$

for some constants C_0 and C_1 . Although CS guarantees the exact recovery of the sparse signals that are measured by a matrix that satisfies the RIP condition, design of such measurement matrices is still an open problem.

Another important factor in CS is the correlation between the measurement and representation basis [91]. If the vector \mathbf{x} is not sparse, but is sparsely represented by $\mathbf{x} = \Phi \theta$, the basis pursuit is in the following form:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \|\theta\|_1, \text{ such that, } \mathbf{y} = \mathbf{A} \Phi \theta. \quad (3.10)$$

To be able to recover the sparse signal θ , the measurement matrix and the representation basis should be incoherent, with coherence defined as the maximum correlation between the elements of \mathbf{A} and Φ

$$\mu(\mathbf{A}, \Phi) = \sqrt{N} \max_{1 \leq i, j \leq N} \langle \Phi_i, \mathbf{A}_j \rangle. \quad (3.11)$$

For a k-Sparse signal, the number of measured samples m should satisfy

$$m \geq C\mu^2(\mathbf{A}, \Phi).k.\log N, \quad (3.12)$$

for a constant C . As can be seen, the higher the coherence between \mathbf{A} and Φ , the higher the number of required samples. It has been shown that a matrix with a i.i.d. Gaussian random entries is incoherent to any matrix and satisfies the RIP condition. Using random measurement matrix, for an exact recovery the number of required samples should satisfy:

$$m \geq C.k.\log(N/k). \quad (3.13)$$

However, random measurement is not possible in many applications, such as computed tomography in which the measurement method is governed by the geometry of the CT scanner.

3.4 Recovery Algorithms

In this section a number of numerical approaches to the problem of signal recovery in CS are discussed. The methods are divided into three groups: ℓ_1 minimization, Greedy algorithms, and iterative re-weighted ℓ_1 minimization. Moreover, Total Variation (TV) minimization is presented as an alternative method used widely in medical imaging applications.

3.4.1 ℓ_1 minimization algorithms

CS theory provides necessary and sufficient conditions to recover the sparse signals from few samples, using ℓ_1 minimization. Nonlinear optimization algorithms are used to recover the sparse signals from the measured samples. These algorithms are either based on solving equation (3.6) or equivalently, by solving:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \text{ such that, } \|\mathbf{A}^T(\mathbf{y} - \mathbf{Ax})\|_\infty < \lambda. \quad (3.14)$$

Both of these minimization problems are convex with conic constraint [92, 93].

However, most proposed algorithms use an unconstrained form of the problem using Lagrange multipliers, i.e.,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda\|\mathbf{x}\|_1, \quad (3.15)$$

in which λ is a positive constant. There exists some λ 's that make this problem equivalent to (3.6). Although (3.15) is not very sensitive to small changes of λ , finding appropriate λ values is very challenging and is still an open problem, which is usually done by trial and error. In chapter 6 we have proposed a method that tunes the regularization parameter adaptively to minimize the recovery error.

Many different algorithms have been proposed to solve the ℓ_1 minimization problem (3.15), including: Homotopy based methods [94], Iterative Soft Thresholding Algorithm (ISTA) [95, 96, 97], Alternating augmented Lagrangian methods (ADMM) [98, 24], and the Split Bregman recovery [99, 100]. In this section iterative soft thresholding and ADMM algorithms are described with more details.

At each iteration, ISTA consists of a gradient descent update step followed by a soft thresholding, defined by $S_\lambda(x) = \text{sign}(x) \max\{|x| - \lambda, 0\}$ [95]. A pseudo-code of the ISTA is shown in Algorithm 1. ADMM works very well in optimizing a summation of simple but sometimes

Algorithm 1 Pseudo-code for Iterative Soft Thresholding.

Inputs: measurement matrix (\mathbf{A}), measurement vector (\mathbf{y}), Sparsity level (k)

Initialize: $\hat{\mathbf{x}}_0 = 0$, $i = 0$

while Stopping Criterion is not met **do**

- 1 $\hat{\mathbf{x}}_i = S_\lambda(\hat{\mathbf{x}}_{i-1} + \mathbf{A}^T(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_{i-1}))$
- 2 $i \leftarrow i + 1$

endwhile

Outputs: Sparse representation $\hat{\mathbf{x}}$.

non-smooth convex functions [98]. Consider the following optimization problem

$$\underset{\mathbf{x}, \mathbf{z}}{\operatorname{argmin}} f(\mathbf{x}) + g(\mathbf{z}), \text{ such that, } M\mathbf{x} = \mathbf{z}, \quad (3.16)$$

in which f and g are convex functions. The augmented Lagrangian is defined as

$$L_\lambda(\mathbf{x}, \mathbf{z}, \mathbf{v}) = f(\mathbf{x}) + g(\mathbf{z}) + \mathbf{v}^T(M\mathbf{x} - \mathbf{z}) + \frac{c}{2} \|M\mathbf{x} - \mathbf{z}\|_2^2. \quad (3.17)$$

Using this definition, at the i^{th} iteration, ADMM applies the following three steps:

$$\begin{aligned} \mathbf{x}^{i+1} &= \underset{\mathbf{x}}{\operatorname{argmin}} L_\lambda(\mathbf{x}, \mathbf{z}^i, \mathbf{v}^i), \\ \mathbf{z}^{i+1} &= \underset{\mathbf{z}}{\operatorname{argmin}} L_\lambda(\mathbf{x}^i, \mathbf{z}, \mathbf{v}^i), \\ \mathbf{v}^{i+1} &= \mathbf{v}^i + c(M\mathbf{x}^{i+1} - \mathbf{z}^{i+1}). \end{aligned} \quad (3.18)$$

As can be seen, this method reduces the joint (x, z) minimization to two simple Lagrangian multiplier problems. In equation (3.15) assuming $f(x) = \frac{1}{2}\|\mathbf{Ax} - \mathbf{y}\|_2^2$, $g(z) = \lambda\|z\|_1$, and $M = I$, for solving ℓ_1 minimization in CS recovery, ADMM consists of:

$$\begin{aligned}\mathbf{x}^{i+1} &= (\mathbf{A}^T \mathbf{A} + cI)^{-1}(\mathbf{A}^T \mathbf{y} + cz^i - \mathbf{v}^i), \\ z^{i+1} &= \text{sign}(\mathbf{x}^{i+1} + 1/c\mathbf{v}^i) \max\left\{\left|\mathbf{x}^{i+1} + \frac{1}{c}\mathbf{v}^i\right| - \lambda/c, 0\right\}, \\ \mathbf{v}^{i+1} &= \mathbf{v}^i + c(\mathbf{x}^{i+1} - z^{i+1}).\end{aligned}\quad (3.19)$$

3.4.2 Greedy Algorithms

Greedy algorithms estimate the support and coefficients of the signal in an iterative approach. At each iteration the estimate of the signal is improved by updating its support. Two well known Greedy algorithms are Matching Pursuit (MP) [101] based methods and Iterative Hard Thresholding (IHT) [102, 103]. MP is based on updating the dictionary at each iteration by adding the vectors on which the residual has the largest projection. Then, the selected vectors of the dictionary are removed from the residual and this is repeated until the norm of the residual is smaller than a preset threshold. The pseudo-code of an Orthogonal Matching Pursuit (OMP) algorithm is shown in Algorithm 2 [104]. At each iteration it finds the most correlated column of the measurement matrix with the measurement residual and adds it to the support. In this algorithm \mathbf{A}^\dagger is the Moore-Penrose pseudo inverse of \mathbf{A} .

OMP has been improved and extended by Stage-wise Orthogonal Matching Pursuit (StOMP) [105], and Compressive Sampling Matching Pursuit (CoSaMP) [106], which allow multiple coefficients to be added to the support at each iteration with tighter bounds on convergence and performance. Iterative Hard thresholding is a simple and straightforward algorithm consisting

Algorithm 2 Pseudo-code of Orthogonal Matching Pursuit.

Inputs: measurement matrix (\mathbf{A}), measurement vector (\mathbf{y})

Initialize: $\hat{\mathbf{x}}_0 = 0$, $r_0 = \mathbf{y}$, $S = \emptyset$, $i = 0$

while Stopping Criterion is not met **do**

- 1 $g_i \leftarrow \mathbf{A}^T r_i$ {Form signal estimate from residual}
- 2 $\Lambda \leftarrow \Lambda \cup \underset{j=1, \dots, N}{\text{argmax}} g_j$ {Add largest residual entry to the support}
- 3 $\hat{\mathbf{x}}_{i|\Lambda_i} \leftarrow \mathbf{A}_{\Lambda_i}^\dagger \mathbf{y}$, $\hat{\mathbf{x}}_{i|\Lambda_i^c} \leftarrow 0$ {Update signal estimate}
- 4 $r_i \leftarrow \mathbf{y} - \mathbf{A} \hat{\mathbf{x}}_i$ {Update measurement residual}
- 5 $i \leftarrow i + 1$

endwhile

Outputs: Sparse representation $\hat{\mathbf{x}}$.

of a gradient descent update followed by hard thresholding applied iteratively until the stopping criterion is met. Pseudo-code of the IHT algorithm is given in Algorithm 3. In this pseudo-code $H_k(x)$ is a hard thresholding functional, which keeps the largest k entries and maps the smaller elements to zero.

Algorithm 3 Pseudo-code of Iterative Hard Thresholding.

Inputs: measurement matrix (\mathbf{A}), measurement vector (\mathbf{y}), Sparsity level (k)

Initialize: $\hat{\mathbf{x}}_0 = 0$, $i = 0$

while Stopping Criterion is not met **do**

- 1 $\hat{\mathbf{x}}_i = H_k(\hat{\mathbf{x}}_{i-1} + \mathbf{A}^T(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_{i-1}))$
- 2 $i \leftarrow i + 1$

endwhile

Outputs: Sparse representation $\hat{\mathbf{x}}$.

3.4.3 Reweighted ℓ_1 Minimization

Although CS theory has shown that ℓ_1 minimization is an appropriate alternative for ℓ_0 minimization, a big difference between ℓ_1 and ℓ_0 is in the dependence of ℓ_1 on the magnitude of the coefficients, such that the larger ones are penalized more heavily, while in ℓ_0 all the coefficients are penalized the same. In [107] a weighted ℓ_1 minimization is proposed to equalize the non-zero coefficients, i.e.,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \sum_{i=1}^N w_i |\mathbf{x}_i| \text{ such that, } \mathbf{y} = \mathbf{Ax}. \quad (3.20)$$

In an ideal scenario, such that the weights are inversely related to the magnitude of the coefficients, the weighted ℓ_1 minimization is exactly the same as ℓ_0 minimization. However, since the true value of the coefficients are unknown, the weights are updated at each iteration with the most updated estimate of the coefficients. For instance, at $l+1^{th}$ iteration with $\hat{\mathbf{x}}^l$ being the most updated coefficients estimate, the weights are defined as $w_i^{l+1} = \frac{1}{|\hat{\mathbf{x}}_i^l| + \varepsilon}$. It has been shown that weighted CS can recover signals from fewer number of samples than regular CS. In addition, weighted CS has been used to solve non-convex ℓ_p minimization [108] problems by utilizing new weights such that $w_i^{l+1} = \frac{1}{(|\hat{\mathbf{x}}_i^l| + \varepsilon)^{1-p}}$.

Algorithm 4 Pseudo-code for the split Bregman for TV recovery.

Inputs: measurement matrix (\mathbf{A}), measurement vector (\mathbf{y})

Initialize: $u^0 = f$, $d_h = d_v = b_h = b_v = 0$

while $\|u^i - u^{i-1}\| > tol$ **do**

$$\begin{aligned} 1 \quad u^{i+1} &= \operatorname{argmin}_{\mathbf{u}} \frac{\lambda}{2} \|\mathbf{Au} - \mathbf{y}\|_2^2 + \frac{\mu}{2} \|d_h^i - \nabla^h u - b_h^i\|_2^2 + \frac{\mu}{2} \|d_v^i - \nabla^v u - b_v^i\|_2^2 \\ 2 \quad d_h^{i+1} &= S_{\frac{1}{\mu}}(\nabla_h u^{i+1} + b_h^i) \\ 3 \quad d_v^{i+1} &= S_{\frac{1}{\mu}}(\nabla_v u^{i+1} + b_v^i) \\ 4 \quad b_h^{i+1} &= b_h^i + (\nabla_h u^{i+1} - d_h^{i+1}) \\ 5 \quad b_v^{i+1} &= b_v^i + (\nabla_v u^{i+1} - d_v^{i+1}) \\ 6 \quad i &\leftarrow i + 1 \end{aligned}$$

endwhile

Outputs: Sparse representation $\hat{\mathbf{x}}$.

3.5 Total Variation minimization

Total Variation image restoration was introduced by Rodin, Osher, and Fatemi [80], which estimates the best denoised image by minimizing the ROF functional, defined by

$$\hat{f} = \operatorname{argmin}_{\mathbf{u}} \mathbf{TV}(u) + \frac{\lambda}{2} \|u - f\|_2^2, \quad (3.21)$$

in which f is the image to be denoised and $\mathbf{TV}(x)$ is the TV-norm with two anisotropic ($|\nabla^h x| + |\nabla^v x|$) and isotropic ($\sqrt{|\nabla^h x|^2 + |\nabla^v x|^2}$) definitions. As can be seen, since TV norm can be described by ℓ_1 norm of the image gradient, it has a very close relation with CS and ℓ_1 minimization. In many applications [7, 9, 30, 109], TV minimization has been successfully used to recover the images from few samples. In the next subsection the split Bregman method is described for solving the TV image recovery, as given by

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{u}} \mathbf{TV}(u) + \frac{\lambda}{2} \|\mathbf{Au} - \mathbf{y}\|_2^2. \quad (3.22)$$

3.5.1 Isotropic TV minimization with Split Bregman Algorithm

The split Bregman algorithm is an special case of ADMM. For isotropic TV minimization, instead of solving $\operatorname{argmin}_{\mathbf{u}} |\nabla^h u| + |\nabla^v u| + \frac{\lambda}{2} \|\mathbf{Au} - \mathbf{y}\|_2^2$ the split Bregman solves the following problem:

$$\operatorname{argmin}_{\mathbf{u}} |d_h| + |d_v| + \frac{\lambda}{2} \|\mathbf{Au} - \mathbf{y}\|_2^2 \text{ such that, } d_h = |\nabla^h u|, d_v = |\nabla^v u|. \quad (3.23)$$

Equivalently by applying the Bregman iteration [99], this can be expressed as

$$\operatorname{argmin}_{u, d_v, d_h} |d_h| + |d_v| + \frac{\lambda}{2} \|\mathbf{A}u - \mathbf{y}\|_2^2 + \frac{\mu}{2} \|d_h - \nabla^h u - b_h^I\|_2^2 + \frac{\mu}{2} \|d_v - \nabla^v u - b_v^I\|_2^2. \quad (3.24)$$

To solve (3.24), Algorithm 4 can be used.

3.6 Application of Compressed Sensing in Computed Tomography

CT image recovery from incomplete projections is an ill-conditioned problem without a unique solution. Classically, to be able to recover images from incomplete data, different prior models of the images have been used in statistical iterative reconstruction algorithms [110, 15]. However, after the introduction of CS, the research focus changed. Currently, the focus is to recover high quality low dose images from fewer projection angles, as demonstrated in figure 3.4.

Two different approaches are currently used for employing CS in CT imaging, depending on the primary objective, and these are as follows:

1) *Radiation dose reduction*: to reduce the X-ray dose, CS-based CT reconstruction algorithms are used to reconstruct high quality CT images either from few high dose projections or from many low dose projections. However, to date, there is no clear evidence as to which approach is best.

2) *Motion artifact reduction*: CS improves the temporal resolution of the CT imaging by reducing the number of projections or by limiting the projection angle. This higher temporal resolution directly improves the performance of CT scanners for imaging moving objects, i.e., respiratory motions in chest CT scans and cardiac motions in cardiac-CT.

Current CS-based reconstruction algorithms suffer from two drawbacks: their high computational load makes them impracticable for routine clinical use [7, 8, 9, 111, 112] and the CT statistics and scanning geometries [24, 25, 26, 27] have not been incorporated in the problem formulation. As a result, there are considerable hurdles to be overcome for these methods to be used in clinical CT systems. The main challenge in solving CS optimization problem within a reasonable amount of time in CT imaging arises from the size of the measurement matrix \mathbf{A} . Currently, in most available CS-based reconstruction methods used for modern CT geometries the measurement matrix \mathbf{A} is a Radon sampling matrix, similar to ART, which models the rays going through the patient. To reconstruct a 512×512 pixel image from 900 sensors and 1200 projection angles, \mathbf{A} would be a 1080000×262144 matrix. As typical iterations each usually require two multiplications by \mathbf{A} and \mathbf{A}^T , it would take several hours of computation on a typ-

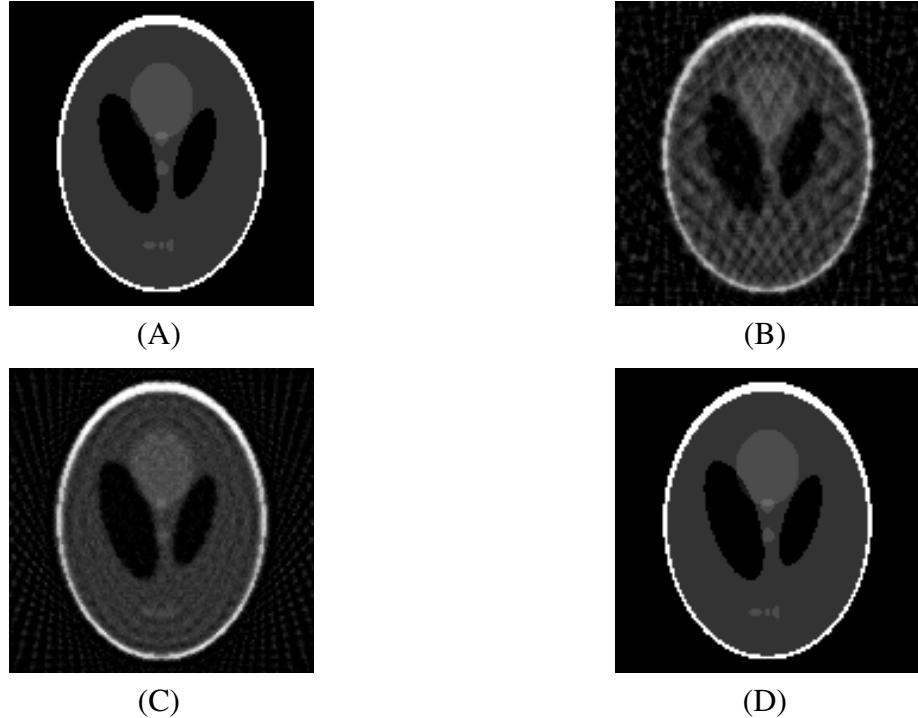


Figure 3.4: Comparison of CT image reconstruction methods using a greatly reduced number of projections with that obtained using Compressed Sensing. (A) Simulated 128×128 Shepp-Logan phantom reconstructed with 36 projections with: (B) Filtered Back Projection, (C) Algebraic Reconstruction Technique, and (D) Compressed Sensing.

ical desktop computer to reconstruct the image [8, 9, 113].

To use CS in CT imaging, the measurements need to be maximally sensitive to as many of the used dictionary basis as possible. Moreover, the images that differ in a small number of basis must be as distinguishable as possible in measurement space. Therefore, optimal measurements are as unstructured as possible in terms of their sensitivity to basis images, but they should also maximally separate sparse representations in measurement space. This property is quantitatively measured by incoherency of the measurement basis. However, the measurement techniques in CT imaging are not very flexible, due to the rotational geometry. These limitations increase the coherency between the measurements, which increases the number of samples needed by CS in CT reconstruction [114]. To reduce the number of needed samples two approaches have been used: 1) utilization of better sparsifying methods, such as group sparsity [115, 9], and 2) improving the measurement matrices to reduce the coherency of the measurements [114, 111, 116].

As an example of the first group, Prior Image Constrained Compressed Sensing (PICCS) [9, 8] uses a prior image \mathbf{x}_p in its formulation to represent the image with more sparsity. The prior image \mathbf{x}_p is a low quality image computed by smoothing the image reconstructed using FBP

from few projections. To reconstruct the CT images, PICCS uses the following optimization expression,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} [a\|\Phi_1(\mathbf{x} - \mathbf{x}_p)\|_1 + (1-a)\|\Phi_2\mathbf{x}\|_1], \text{ s.t., } \mathbf{y} = \mathbf{Ax}, \quad (3.25)$$

in which Φ_1 and Φ_2 are two sparsifying transforms, and a is a positive constant. In PICCS, total variation is used for both Φ_1 and Φ_2 and, in addition, $a = 0.91$. Equation (3.25) is solved based on an ART reconstruction followed by a TV denoising. In each iteration ART is used to recover the image from $\mathbf{y} = \mathbf{Ax}$, then a steepest descent algorithm is used to minimize the total variation of the weighted summation in equation (3.25).

To reduce the coherency of the measurements in [74] an improved detector arrangement, as shown in figure 3.5, was proposed, but not tested. In [114] different sampling strategies, as shown in figure 3.6, are investigated for tomography. It has been shown that random view angle sub-sampling and uniform view sub-sampling are not very different. However, random detector sub-sampling has improved the incoherency. Using this property, coded aperture X-ray imaging [117], as shown in figure 3.7, has been proposed that optimizes the CT hardware for CS application.

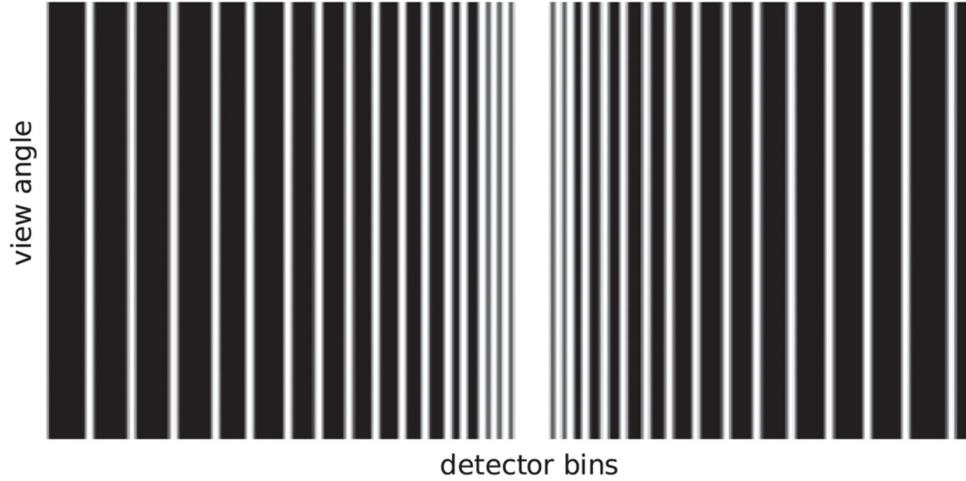


Figure 3.5: Angular independent coded aperture CT. The white regions indicate the sampled part of the data. Reproduced, with permission, from [74].

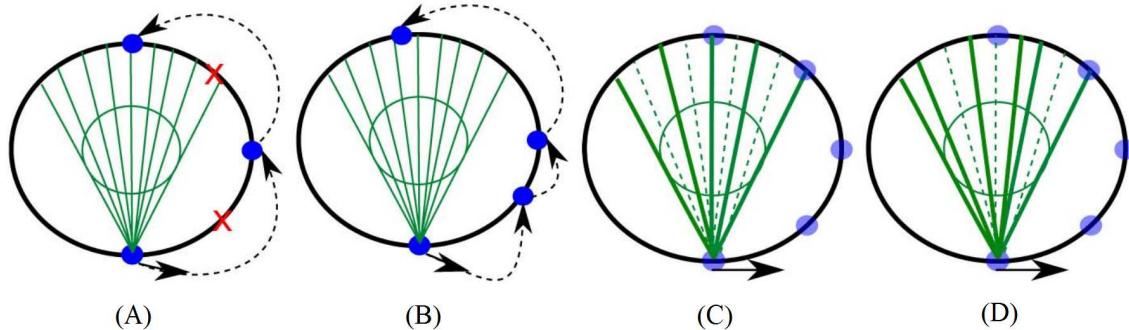


Figure 3.6: Different sampling strategies being considered, as demonstrated on a fan-beam geometry. The blue circles denote source locations (corresponding to view angles), and lines denote source-detector pairs (for clarity, only the lines for one source location are shown for each case). The solid and dashed lines correspond to measured and non-measured line integrals, respectively. The large circle represents the source trajectory, and the small circle represents the imaged domain. (A) Uniform view sub-sampling, (B) random view angle sub-sampling, (C) uniform detector sampling, and (D) random detector sub-sampling. Reproduced, by permission, from [114].

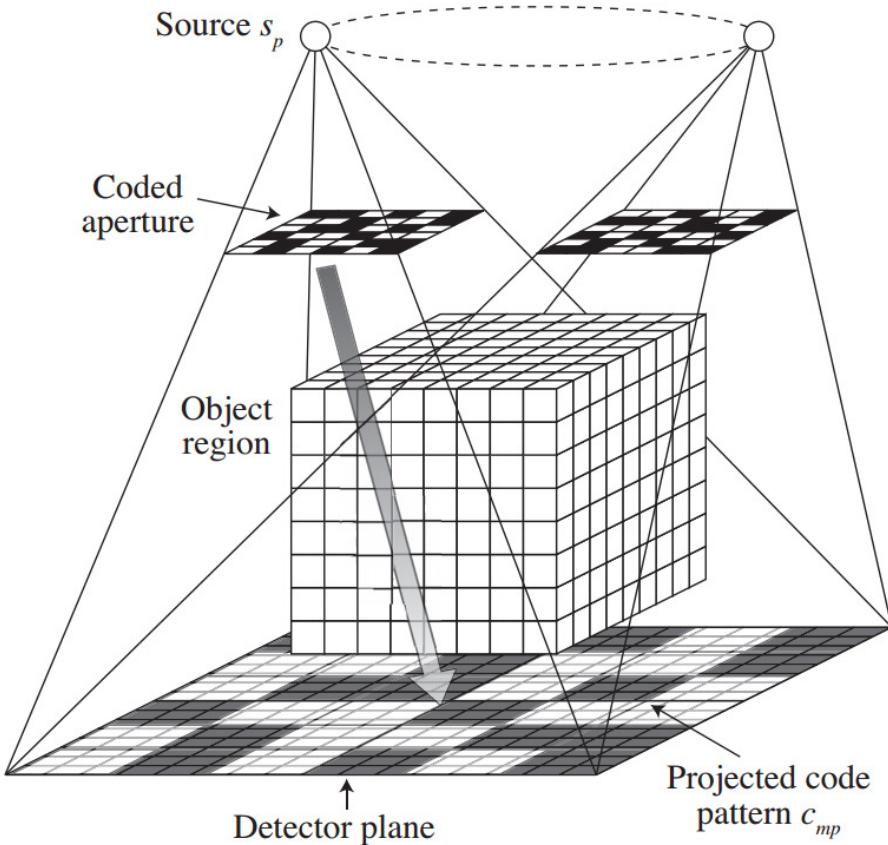


Figure 3.7: Schematic of a coded aperture CT scanner proposed in [117]. Black regions block the X-ray rays. Reproduced, with permission, from [117].

Chapter 4

Image Reconstruction of Small Pulmonary Nodules¹

Indeterminate non-calcified pulmonary nodules are generally regarded as potentially malignant lesions and require close monitoring with a series of follow-up CT examinations. Because such imaging, needs to be repeated from time to time, a low dose protocol should be used that does not sacrifice the image quality needed for nodule detection and characterization. The ratio of the image slice thickness to reconstruction interval (overlap) has an important effect on nodule detectability. For example, a small slice thickness can result in a poor Signal to Noise Ratio (SNR), while a larger thickness implying a higher degree of overlap can result in a significant loss in spatial resolution. This chapter describes the methods and results of an investigation as to the optimal protocol to be used, in terms of the slice thickness/overlap parameters and the potential advantages of using an iterative reconstruction algorithm. The two aims of this investigation are: 1) to determine the effect of slice thickness/overlap on nodule size and density with specific focus on nodule detectability and CT number accuracy for low dose CT scans and, 2) to optimize the reconstruction slice thickness/overlap of low dose CT for pulmonary nodule detection and characterization.

Lung cancer is the leading cause of global cancer related mortality [118]. At present, most patients have advanced disease which explains the low overall 5-year survival rate of 14% that has remained stable for the past two decades. Early stage lung cancer has 5-year survival rates of 60-70%. Therefore, early detection of lung cancer is a critical step in reducing mortality [119]. CT has high sensitivity for detection of pulmonary nodules; but as the majority of detected nodules are benign, CT remains of limited specificity [120]. Indeterminate pulmonary

¹This chapter is based on: S.M. Hashemi, H. Mehrez, R.S.C. Cobbold, and N.S. Paul, “Optimal image reconstruction for detection and characterization of small pulmonary nodules during low-dose CT,” European radiology, vol. 24, no. 6, pp. 1239–1250, 2014.

nodules often undergo serial surveillance CT scans to detect any suspicion changes such as increase in nodule size or in nodule density [121].

As serial CT scans can result in a relatively large radiation exposure to patients, there has been increased interest in utilizing Low-Dose CT (LDCT) scans for nodule detection. Screening trials for early detection of lung cancer [122, 123] using thoracic LDCT have demonstrated survival benefits when screening high risk populations and this has resulted in increased application of LDCT for nodule detection and characterization. Thin slice image reconstructions are important in reducing partial volume effects and thereby obtaining accurate nodule density and size measurements especially in smaller nodules [124, 125]. However, thin section LDCT results in significant increases in image noise [126] which in turn reduces the measured contrast to noise ratio (CNR) of lung nodules, and potentially leads to difficulties in accurate detection and characterization particularly of low contrast nodules that are purely or predominantly of Ground Glass Opacity (GGO).

Malignant change in ground glass nodules is reflected by an increase in measured nodule CT attenuation from GGO to soft tissue density [127, 128, 129]. Therefore, it is important that pulmonary nodule LDCT studies be performed using parameters that optimize the accuracy of measuring nodule attenuation or CT# and CNR. The slice thickness is an important parameter in determining nodule density and size accurately [124, 125, 130, 131], and the image reconstruction overlap can improve nodule detectability [132, 133, 134]. In addition, there is increasing interest in the use of Iterative Reconstruction (IR) in thoracic CT to facilitate improved image quality at reduced patient dose [135]. Therefore, it is important to evaluate the effect of clinically available iterative methods on detection and characterization of small pulmonary nodules.

4.1 Materials and methods

The aim of this study is to determine the optimal slice thickness/overlap of image reconstruction and to examine the effect of IR on the detectability and characterization of small non-calcified pulmonary nodules (NCPN) during LDCT. These goals were achieved by computer simulations², phantom studies, and patient scans.

²Although the computer simulations were performed numerically using MATLAB, it should be noted that an analytical approach can also be used, as detailed in Appendix B.

4.1.1 Image reconstruction parameters and nodule measurements

All images were reconstructed using the following combinations of reconstruction slice thickness/overlap: 0.5/0.5, 1/1, 2/1, 2/2, 3/1.5, 3/3, 4/2, 4/4, 5/2.5, and 5/5 mm. For each of these 10 combinations, we have also simulated the effect of table position (scan starting position) on the reconstructed image. Due to partial volume artifact, the scan starting position will impact the location of the nodule with respect to the selected slices for reconstruction, which in turn influences measured attenuation and as a result the nodule classification. To take this effect into account, the location of the first reconstructed slice in every configuration is shifted by $n \times 0.25$ mm, where n is an integer between 0 and $overlap/0.25$ and 0.25 mm corresponds to half of the thinnest possible slice thickness. For example, when considering 2.0/2.0 mm reconstructions we generate eight different series with the first slice located at 0.0, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, and 1.75 mm; respectively. Therefore, for each nodule 90 image series are reconstructed consisting of [2 of (0.5/0.5 mm), 4 of (1.0/1.0 mm), 4 of (2.0/1.0 mm), 8 of (2.0/2.0 mm), 6 of (3.0/1.5), 12 of (3.0/3.0 mm), 8 of (4.0/2.0 mm), 16 of (4.0/4.0 mm), 10 of (5.0/2.5 mm), and 20 of (5.0/5.0 mm)]. For every series, all nodule locations were identified, and the slice containing the largest nodule cross section was used to determine CT# and CNR values [136]. Measurement of the CT# was performed by prescribing a region of interest (ROI) in the center of the nodule to cover about 66% of the greatest nodule dimension³. The background image noise was measured by selecting two ROIs at two tissue locations on the same slice as the nodule with maximum visibility. The CNR was defined by the following relationship [137, 138]:

$$CNR = \frac{Nodule\ CT\ number - Background\ CT\ number}{STD\ of\ background\ noise}. \quad (4.1)$$

Analysis was carried out in 3 stages: computer simulation, phantom evaluation, and patient data. These three stages are illustrated in figure 4.1. Our reference (gold) standard CT# is known for the simulation and in vitro data; and the measurements obtained from a slice thickness/overlap of 0.5/0.5 mm are considered as the reference standard for the in vivo nodules. CT# measurement error is defined as the difference between the measured value and the reference standard CT#. For the purpose of our analysis, solid nodules were defined as near spherical opacities with a CT# of approximately 100 HU and GGO were defined as opacities with a CT# of less than -400 HU, that were imperceptible on Soft-Tissue window settings (WL = 40, WW = 400) [139].

Computer simulations: Virtual nodules

Our computer simulations used a simplified model of high and low contrast nodules (solid

³Appendix B also discusses the importance of the ROI selection.

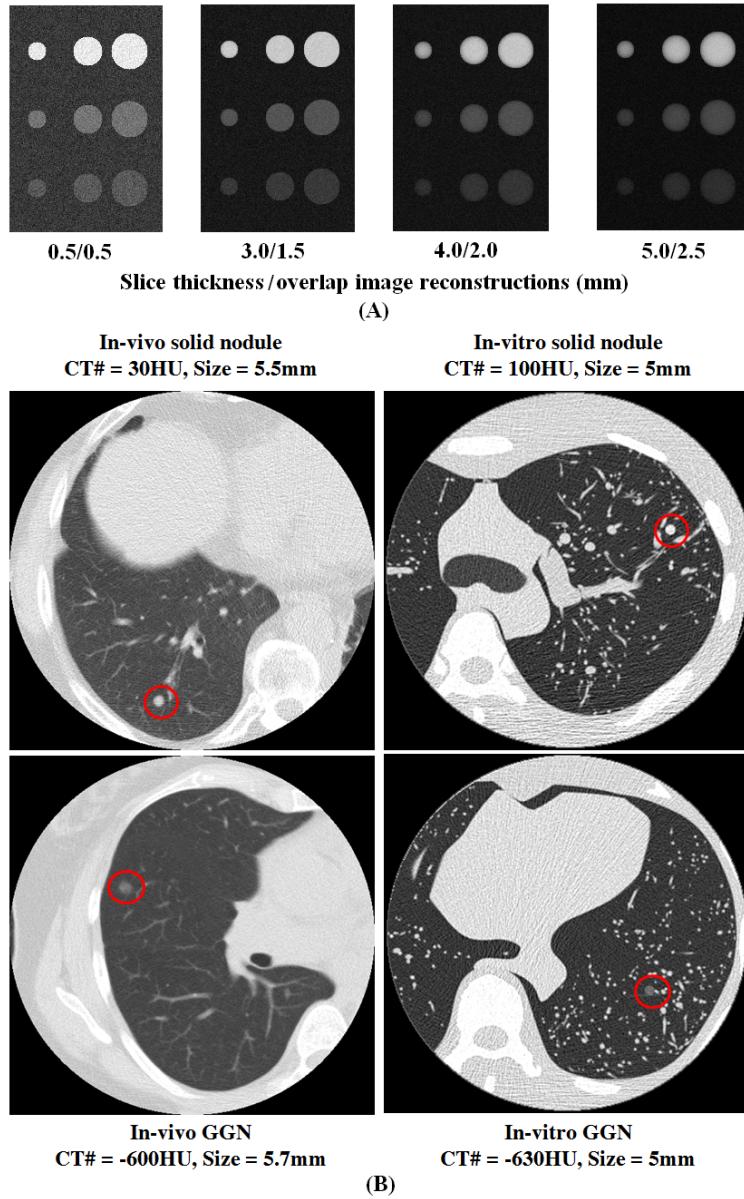


Figure 4.1: Illustrating the influence of effective slice-thickness on lung nodules: (A) Simulated nodules 5, 8, 10 mm diameter (left to right in each row) and 100 HU, -630 HU, -800 HU CT# (top to bottom row in each image) reconstructed with 0.5/0.5, 3.0/1.5, 4.0/2.0, 5.0/2.5 mm slice thickness/overlap (left to right). Qualitatively, there is a decrease in edge sharpness and image noise with larger slice thickness. (B) Solid ~5 mm nodules (top row) and ~5 mm ground glass nodules (bottom row) from in vivo (left column) and in vitro (right column) image data sets. Qualitatively the in vivo and in vitro datasets have comparable nodule appearances for low dose CT. Slice thickness/ Overlap = 2.0/2.0 mm and WW/WL = -550/1600 for both in-vitro and in-vivo nodules.

nodules and GGO) on a smooth background that provided a reference for the measured CT# and CNR changes of the phantom and patient scans. Nine virtual nodules corresponding to

3 different diameters (5, 8, and 10 mm) and 3 different CT# (100, -630, and -800 HU) were simulated using MATLAB (Mathworks, Natick, MA, USA) by placing every nodule in a three dimensional mesh. The mesh was represented by an array of $512 \times 512 \times 320$ corresponding to a volume of $400 \times 400 \times 80 \text{ mm}^3$, resulting in a voxel size of $0.78 \times 0.78 \times 0.25 \text{ mm}^3$. The mesh background CT# was set to -1000 HU corresponding to air. Gaussian noise [140] was added to each mesh voxel including points where the nodule was located. The mean noise value was set to 0 HU with a standard deviation (STD) of $50 \times \sqrt{\frac{5\text{mm}}{0.25\text{mm}}} \approx 223 \text{ HU}$, where 50 HU is within the range of image noise measured during clinical thoracic LDCT performed at the Toronto General Hospital using 5 mm image reconstruction. Because the image noise is proportional to $\sqrt{\frac{1}{SliceThickness}}$ [141, Proposition 2.11], the factor $\sqrt{\frac{5\text{mm}}{0.25\text{mm}}}$ was used to determine the noise corresponding to images whose slice thickness was 0.25 mm.

Phantom evaluation: In vitro synthetic nodules

Nine synthetic nodules corresponding to 3 diameters (5, 8, and 10mm) and 3 attenuations (100, 630, and 800HU) were placed in a commercially available adult thoracic anthropomorphic phantom, Lungman, Kyoto Kakagu, Japan [142]. The positions of the nodules were randomly selected but covered the entire lung region from the apex to the base. The phantom was scanned using the LDCT protocol, as described in the next section. A total of 90 image series were reconstructed, as described earlier.

Patient data: In vivo nodules

The clinical protocol for thoracic LDCT, as used at the Toronto General Hospital, consists of $64 \times 0.5\text{mm}$ detector row collimation, 120 kV, 50 mA, 0.5sec gantry rotation speed, and a helical pitch factor of 0.83 [143]. The scan FOV was adjusted to the patient size and the display FOV was adjusted to encompass the entire patient chest.

Patient data were reconstructed using Toshiba's lung kernel (FC55) for Filtered Back Projection (FBP) and mediastinal (FC04) kernels with Adaptive Iterative Dose Reduction 3D technique (AIDR3D, Toshiba Medical Systems [144]). Raw data for all patients undergoing thoracic LDCT scans over a six month period (October 2010 to April 2011) were stored, and the reconstructed axial images were retrospectively reviewed to identify 4-7 mm non-calcified pulmonary nodules. A total of 20 NCPN were identified in 15 patients. The raw data was reconstructed into 180 different image series using 90 different configurations of reconstruction slice-thickness/overlap and each series was reconstructed with FBP and AIDR3D. Each image series was anonymized and reviewed independently in random order by two readers who were asked to identify all nodules and measure the nodule CT# and CNR. The measured values were averaged to provide mean nodule attenuation. Examples of simulated, in vitro, and in vivo solid

and GGO nodules are shown in figure 4.1.

4.1.2 Statistical Analysis

Spearman's rank correlation coefficient was used to assess the relationship between the errors in nodule attenuation measurement and slice-thickness. The Student's T-test for paired samples was used to determine the significance of the difference between measured attenuation and actual values for synthetic and in vitro data. A p-value threshold of 0.05 was used as cut-off for statistical significance.

4.2 Results

Our results indicate that the variations in measured CT# and CNR of a nodule depend highly on slice thickness/overlap, nodule size, and scan starting position. Figure 4.2 shows the CT# and CNR values as a function of slice thickness/overlap for a 5 mm diameter virtual solid nodule. The effect of the scan starting position is displayed in this figure by demonstrating the minimum, maximum and the standard deviation for each CT# and CNR value. It is clear that the reconstructions with overlap are less affected by the starting position. In practice, because the starting point of each scan is changed randomly, this effect should be minimized by the use of overlapped reconstructions. As very similar results were obtained for other nodule sizes, in order to simplify the figures, only the mean values are shown for all subsequent figures.

4.2.1 Computer simulated data: Virtual nodules

The measured CT# and CNR of simulated solid (100 HU) and GGO (-630 HU) nodules of 5, 8 and 10 mm diameter at different reconstruction slice thickness/overlap are demonstrated in figure 4.3. The behavior of the nodules with CT# of -800 HU is exactly the same as the nodules with CT# of -630HU and is not shown in the figure. For solid nodules, there was an overall decrease in measured nodule CT# with increase in reconstruction slice thickness. This effect was more pronounced with 5 mm nodules which maintained $\pm 10\%$ accuracy of CT# up to a reconstruction slice thickness/overlap of 2.0/2.0 mm. However, there was a major reduction in measured CT# for the 5 mm solid nodule above a slice thickness/overlap of 3.0/1.5 mm, such that the measured CT# was negative (figure 4.3-A). This trend was less noticeable for GGO of -680 and -800 HU which demonstrated a more stable measured CT# across all of the tested image reconstruction combinations (figure 4.3-B). The Spearman's rank correlation coefficient between measured CT# and slice thickness was -1 ($p < 0.05$) demonstrating strong correlation

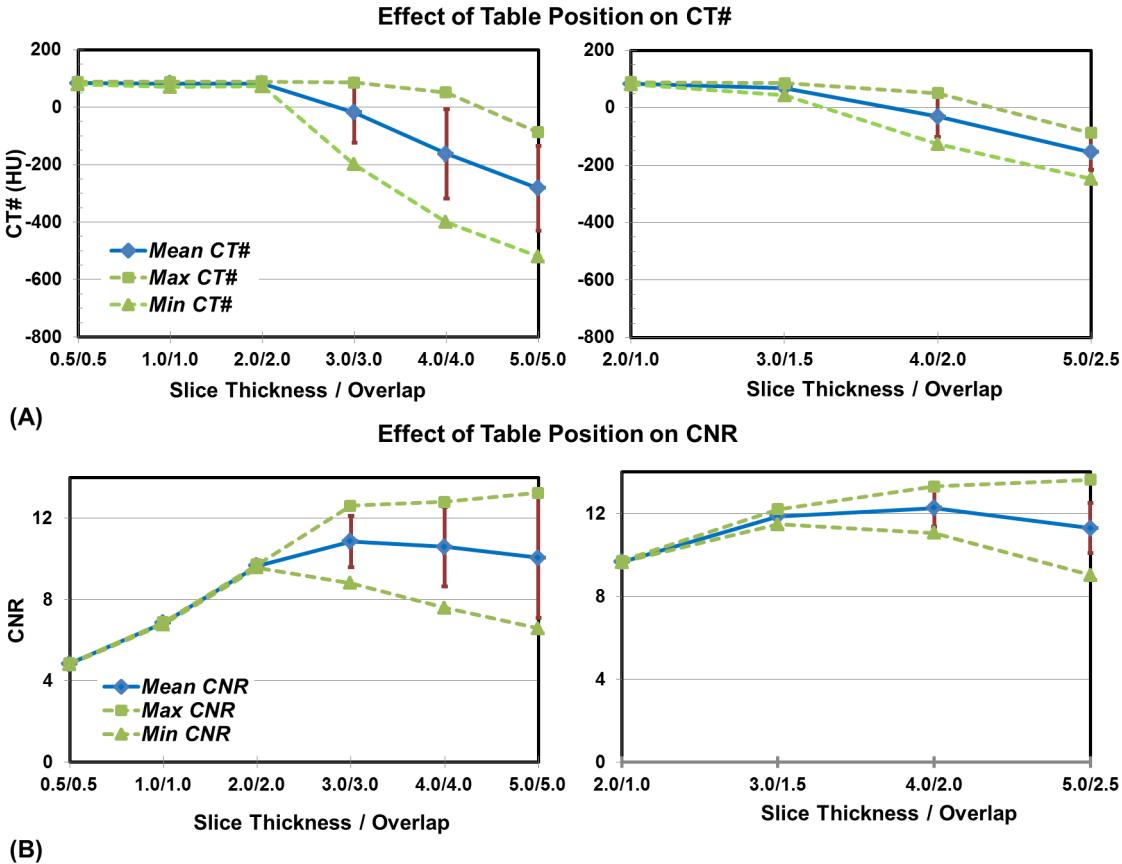


Figure 4.2: Variation in measured CT# and CNR with changes in the effective slice thickness (left: non-overlapped and Right: with 50% overlapped) and scan starting position for a 5 mm solid simulated nodule (Red bars show the STD of the changes due to the ten different simulations performed and scan starting position): (A) Measured mean CT attenuation number (CT#) and (B) Contrast to Noise Ratio (CNR).

between the error in CT# and slice thickness. The Paired Student's T-test demonstrated a significant difference between the measured CT# and the actual value only for the 5 mm nodules ($p < 0.05$).

The change in CNR was more consistent for solid and ground glass nodules. The measured CNR increases as a function of reconstruction slice-thickness. The measured CNR for 5 mm nodules maximized in value at a slice-thickness/overlap of 4.0/2.0 mm. Image reconstructions performed without overlapping slices had higher variations in the measured CT# and lower CNR values compared to images with equivalent slice thicknesses and 50% overlap (figure 4.3-C and 4.3-D).

4.2.2 Phantom data: In vitro synthetic nodules

The measured CT# and CNR of in vitro solid (100 HU) and GGO (-630 HU) with diameters of 5, 8, and 10 mm at different reconstruction slice thickness/overlap is demonstrated in figure 4.4. It can be seen that these in vitro results have a similar pattern to the simulated results of figure 4.3. Specifically, increasing the reconstruction slice thickness/overlap decreases the accuracy in measured CT# particularly for the 5 mm solid nodule which measures close to 0 HU at a reconstruction slice thickness/overlap of 3.0/1.5 mm. Also the CT# for the 5 mm nodule decreases by 400 HU at a reconstruction slice thickness/overlap of 5.0/5.0 mm (figure 4.4-A). There is little change in the measured CT# for the GGO (figure 4.4-B). Similarly, the measured nodule CNR tends to increase with higher reconstruction slice thickness/overlap and peaks at 4.0/2.0 mm reconstructions for 5 mm nodules (figure 4.4-C and 4.4-D). Statistical analysis on the synthetic in vitro nodules has shown identical results to the simulated nodules. The Spearman's rank correlation coefficient between the measured CT# and slice-thickness was -1 ($p < 0.05$). Paired Student's T-test demonstrated a significant difference between the measured CT# and the actual value only for the 5 mm nodules ($p < 0.05$).

4.2.3 Patient data: In vivo nodules

Analysis was performed using 20 small pulmonary nodules with a mean maximum diameter of 5.1 ± 1.2 mm (range 4.0-6.9 mm), median 4.5 mm, and mean CT# of -21 ± 296 HU (range 380 HU to -664 HU), median 17.5 HU. The initial study cohort was 50 consecutive patients with low dose thoracic CT performed for clinical reasons. The criterion for study inclusion was that each study should contain at least one non calcified lung nodule ranging in size from 4-7 mm. The nodules were identified and the CT# and CNR were measured by two readers. The correlation between the two readers was 0.98 ± 0.001 . As shown in figure 4.5, the reduction in measured nodule CT# and the general increase in nodule CNR with larger reconstruction slice thickness/overlap for solid and GGO was similar to the trend for simulated and in vitro nodules. Furthermore, figure 4.5 demonstrates that for a given slice thickness, CT# and CNR are higher when 50% reconstruction overlap is employed compared to reconstructions where there is no overlap. Finally, we note that the CNR increases monotonically from a slice thickness of 0.5 mm to 2.0 mm and reaches a maximum at 4.0/2.0 mm reconstruction slice thickness/overlap. Conversely, for 3.0/3.0, 4.0/4.0, and 5.0/5.0 mm reconstructions, there is a reduction in CNR of more than 20% from the maximal measured CNR values. The estimated accuracies, with which in vivo nodules can be classified, are summarized in the histogram of figure 4.6. These show that: with a slice thickness/overlap of 1.0/1.0 mm, the CT# measurement error is likely less than 50 HU and that for a slice thickness/overlap of 2.0/2.0 mm, should be less than 250

HU. In both of these cases the measurement error does not cause reclassification of the solid nodules. Much larger CT# measurement errors can occur with certain slice thickness/overlap values. For example, it was found that with 5.0/2.5 mm the CT# measurement error ranged up to 500 HU. Based on the measured CT# of 5.0/2.5 mm reconstructions, the effect of this measurement error is that 8 of the 20 solid in vivo nodules (40%) would have been classified as GGO. However, with a slice thickness/overlap of 2.0/2.0 mm they would have been properly classified as solid nodules. This is illustrated in a clinical example shown in figure 4.7.

Figure 4.8 shows the slice thickness/overlap CNR results obtained using FBP and AIDR-3D for 20 in vivo nodules. It can be seen that the overall increase in nodule CNR is maintained with increasing slice thickness with AIDR3D. It increases the CNR of solid nodules by 40% and of GGO by 60% (Paired Student's T-test, $p < 0.05$). The use of IR was found to have no significant influence on the estimation of the CT# values of the nodules (Paired Student's T-test, $p = 0.2$) and it did not change the position of the slice thickness/overlap CNR peaks for different nodules.

4.3 Discussion and Conclusion

Nodule detection and characterization based on clinical LDCT is usually performed in two stages. The first stage is focussed on nodule detection and requires a review of the trans-axial CT images that are reconstructed using a lung kernel algorithm and displayed on lung window settings (WL: 550, WW: 1600). The second stage requires characterization of detected nodules. This is performed preferentially using images reconstructed by a soft tissue kernel algorithm and viewed on a soft tissue window display (WL: 40, WW: 400); a region of interest may be prescribed in the nodule to quantify the CT#. Optimal slice thickness /overlap reconstruction is required for each of these steps to ensure accurate nodule detection and characterization. Our results from the simulated, in vitro, and in vivo nodules demonstrate that slice thickness/overlap reconstructions do not have a significant effect on the accuracy of CT# estimates of larger nodules, i.e. ≥ 8 mm in diameter. However, larger slice thicknesses result in a significant increase in CNR which is likely to improve the conspicuity of nodules that have a pure GGO composition.

The reconstructed slice thickness significantly influences the accuracy of measured CT# in nodules smaller than 5mm. Furthermore, solid nodules demonstrate a much larger variation in measured CT# with reconstructed slice thickness compared to GGO. For example, a 5 mm nodule that is measured 100 HU in a 0.5/0.5 mm image reconstruction could be measured 400 HU less than that obtained with a 5.0/5.0 mm reconstruction. The measurement error with reconstructions of 2.0/2.0 mm and below is less than 10% for simulated, in vitro and in vivo

nodules. Therefore, accurate measurement of nodule CT# can be determined by employing thin slices to overcome partial volume artefact. However, this approach has significant limitations for clinical LDCT as the resultant image noise adversely affects nodule detectability due to the reduction in nodule CNR. For example, for nodules smaller than 5 mm, we found that CNR reaches a maximum at slice thickness/overlap of 4.0/2.0 mm. Consequently, the optimal combination of slice reconstruction parameters for lung nodule assessment during LDCT should include a 4.0/2.0mm reconstruction of the whole chest for nodule detection and a 2.0/2.0 mm reconstruction for nodule characterization. As compared to the FBP algorithm, we have also found that the IR algorithm, AIDR3D, did not affect estimation of the CT#. Because the use of AIDR3D increased the nodule CNR, it is likely to improve nodule detectability and consequently, the use of AIDR3D is recommended during routine LDCT exams.

Various studies [124, 125, 131, 145] have demonstrated that decreasing slice thickness improves nodule characterization, especially for small nodules (<5mm diameter). While these results are in agreement with our findings, we have demonstrated that thin slice reconstructions (0.5mm) have significant challenges for use in routine clinical LDCT protocols. CT image noise, which is inversely proportional to the square root of slice thickness, decreases the detectability of the nodule for these images. Recent evaluation of thin slice image data [13], suggested that nodules smaller than 5 mm in size were significantly better depicted with 1.25 mm compared to 5 mm slice thicknesses. While this data appears to contradict our findings, the origin of the discrepancy is due to the use of a high dose protocol of 140 kV and 120 mAs [13], which facilitated accurate nodule detection even in images reconstructed with a 1.25 mm slice thickness. Finally, we note that to the best of our knowledge our investigations are the first to evaluate a LDCT protocol for simulated, in vitro, and in vivo nodules with both FBP and IR. Therefore, our conclusions and recommendations for translation to the clinical setting are also different from previous reports.

Our study is limited by the use of quantitative metrics to assess parameters that affect nodule detection and characterization using CT. These metrics are important when performing quantitative analysis of pulmonary nodules, for example in multi-spectral analysis and nodule perfusion studies using CT. In routine clinical practice, most pulmonary nodules are characterized qualitatively, but quantitative analysis is utilized in equivocal cases. Therefore, the findings from this research are relevant to the practicing radiologist.

4.4 Summary

In conclusion it can be stated that the best detection of small pulmonary nodules (with \sim 5 mm diameter) requires slice thickness/overlap of 4.0/2.0 mm, and that slice thickness/overlap of

2.0/2.0 mm is required for optimal nodule characterization. Moreover, the use of iterative reconstruction improves conspicuity of small ground glass nodules through a significant increase in nodule CNR.

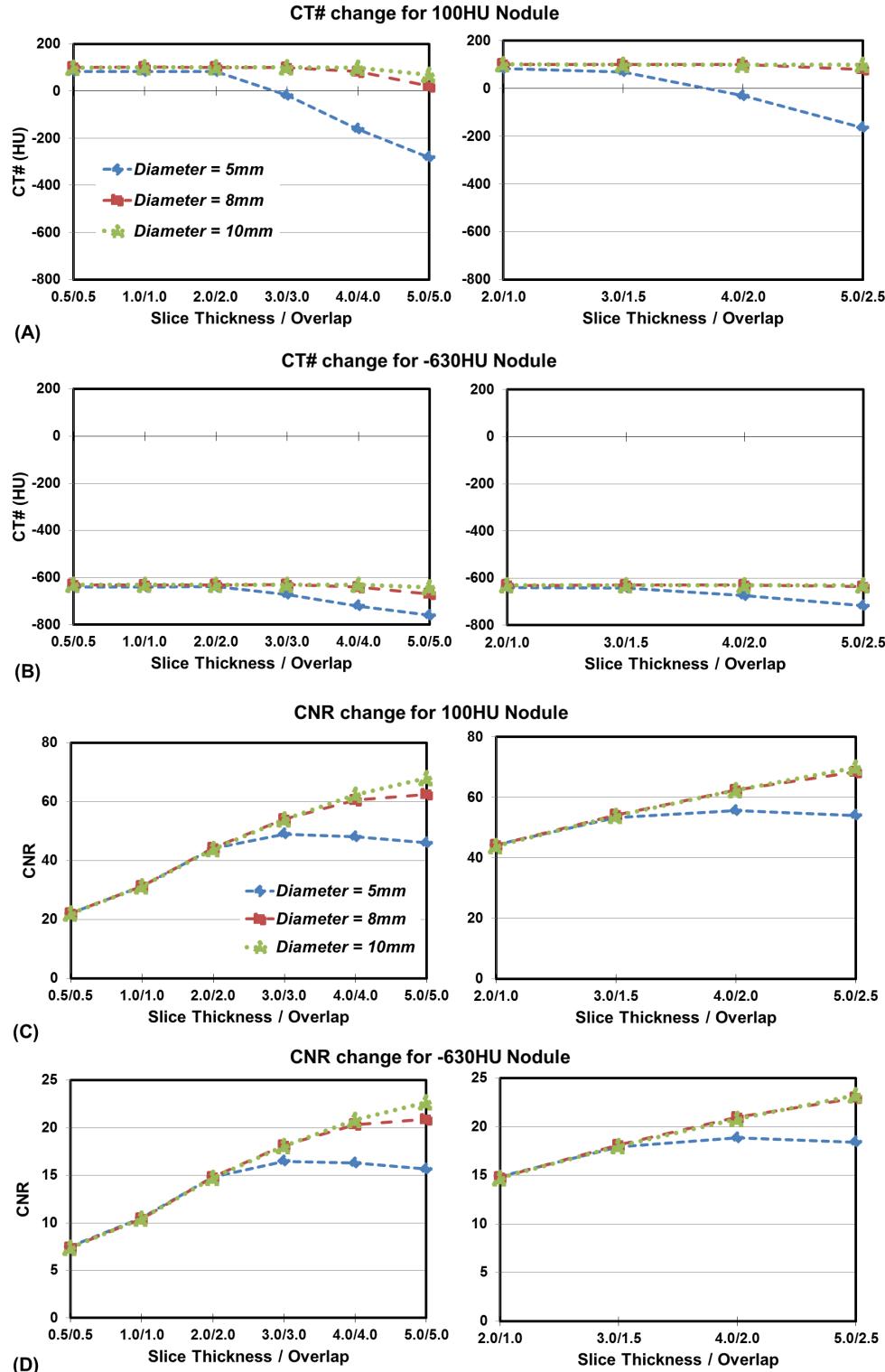


Figure 4.3: Variation in CT# and CNR with change in effective slice thickness (left: non-overlapped and Right: with 50% overlapped) for simulated nodules: (A, B) Measured mean CT attenuation number (CT#) and (C, D) Contrast to Noise Ratio (CNR). (A, C) correspond to the simulated solid nodules and (B, D) correspond to the simulated ground glass nodules.

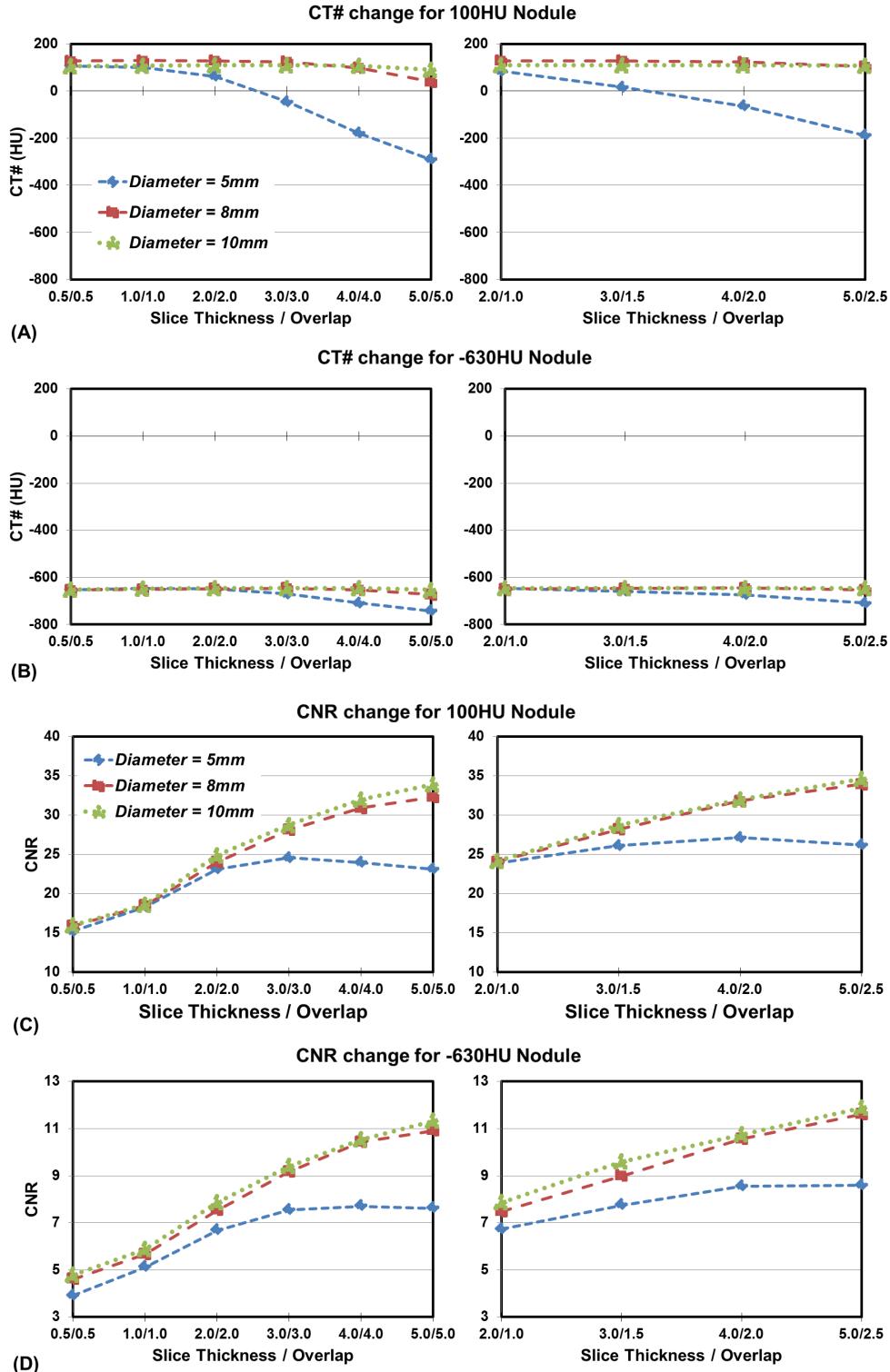


Figure 4.4: Variation in measured CT# and CNR with change in effective slice thickness (left: non-overlapped and Right: with 50% overlapped) of the in vitro Nodules: (A, B) Measured mean CT attenuation number (CT#) and (C, D) Contrast to Noise Ratio (CNR). (A, C) correspond to the synthetic solid nodules and (B, D) correspond to the synthetic ground glass nodules.

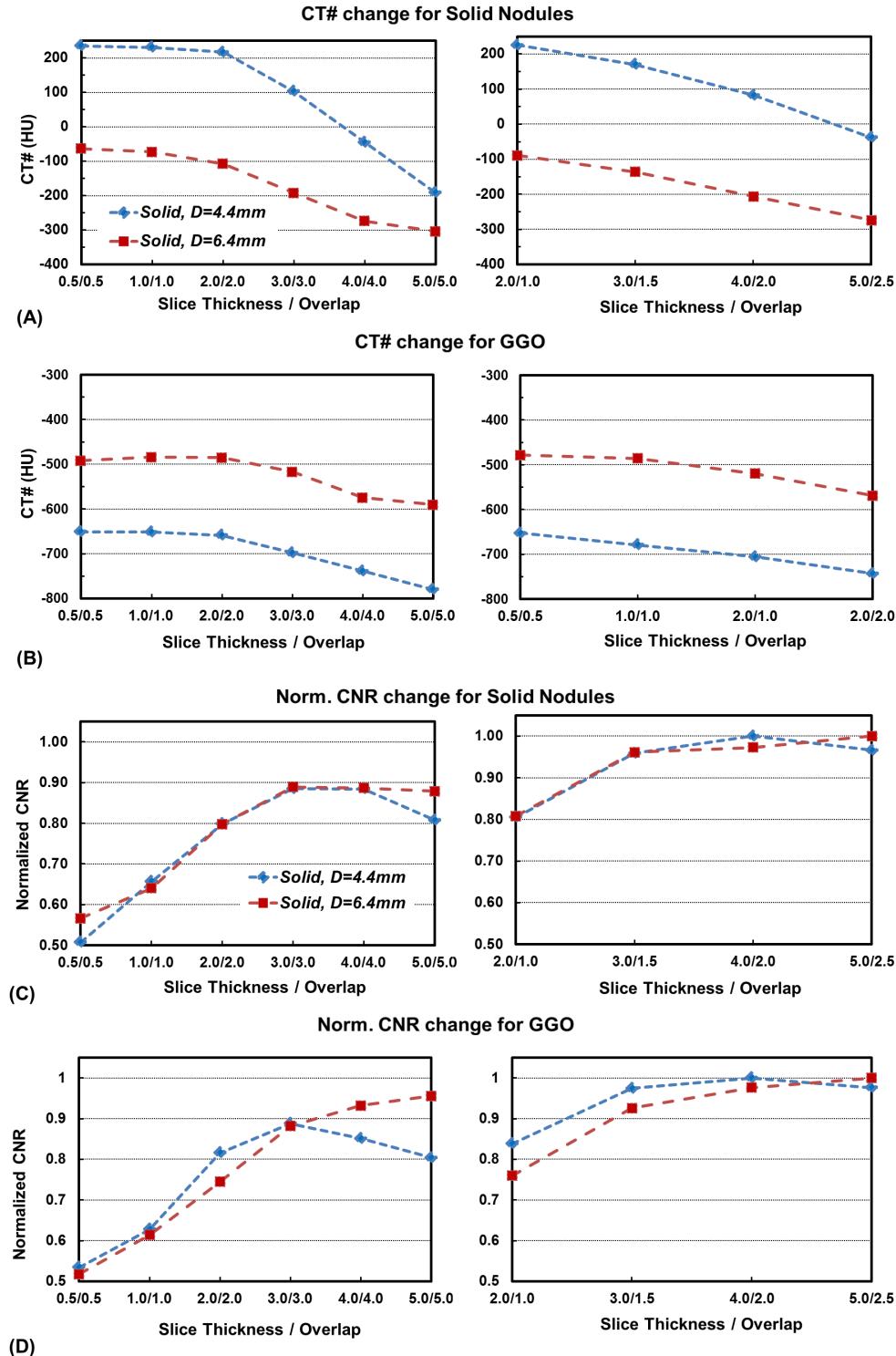


Figure 4.5: Variation in measured CT# and CNR with change in effective slice thickness (left: non-overlapped and Right: with 50% overlapped) for in vivo nodules: (A, B) Measured mean CT attenuation number (CT#) and (C, D) normalized Contrast to Noise Ratio (CNR). (A, C) correspond to two solid nodules and (B, D) correspond to two ground glass nodules.

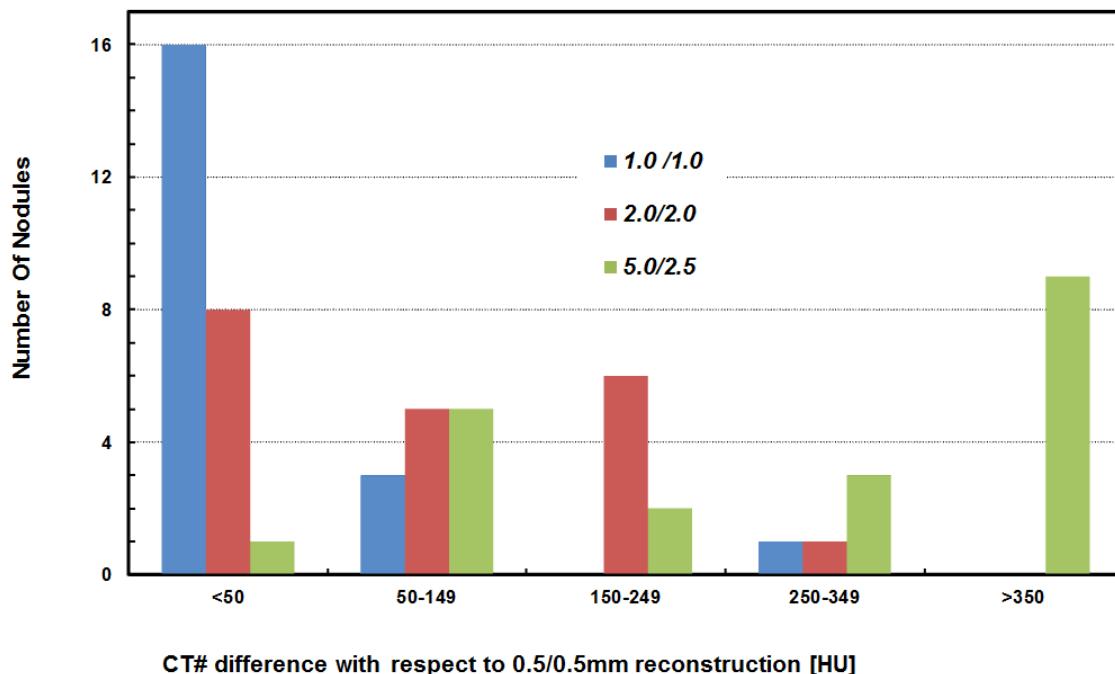


Figure 4.6: Histogram of the measurement error for CT# of the in vivo nodules at three different slice-thicknesses /overlap. Measurements were made at 1.0/1.0, 2.0/2.0, and 5.0/2.5 mm, and the error is calculated based on the gold standard values of the CT# obtained from the 0.5/0.5 mm reconstruction.

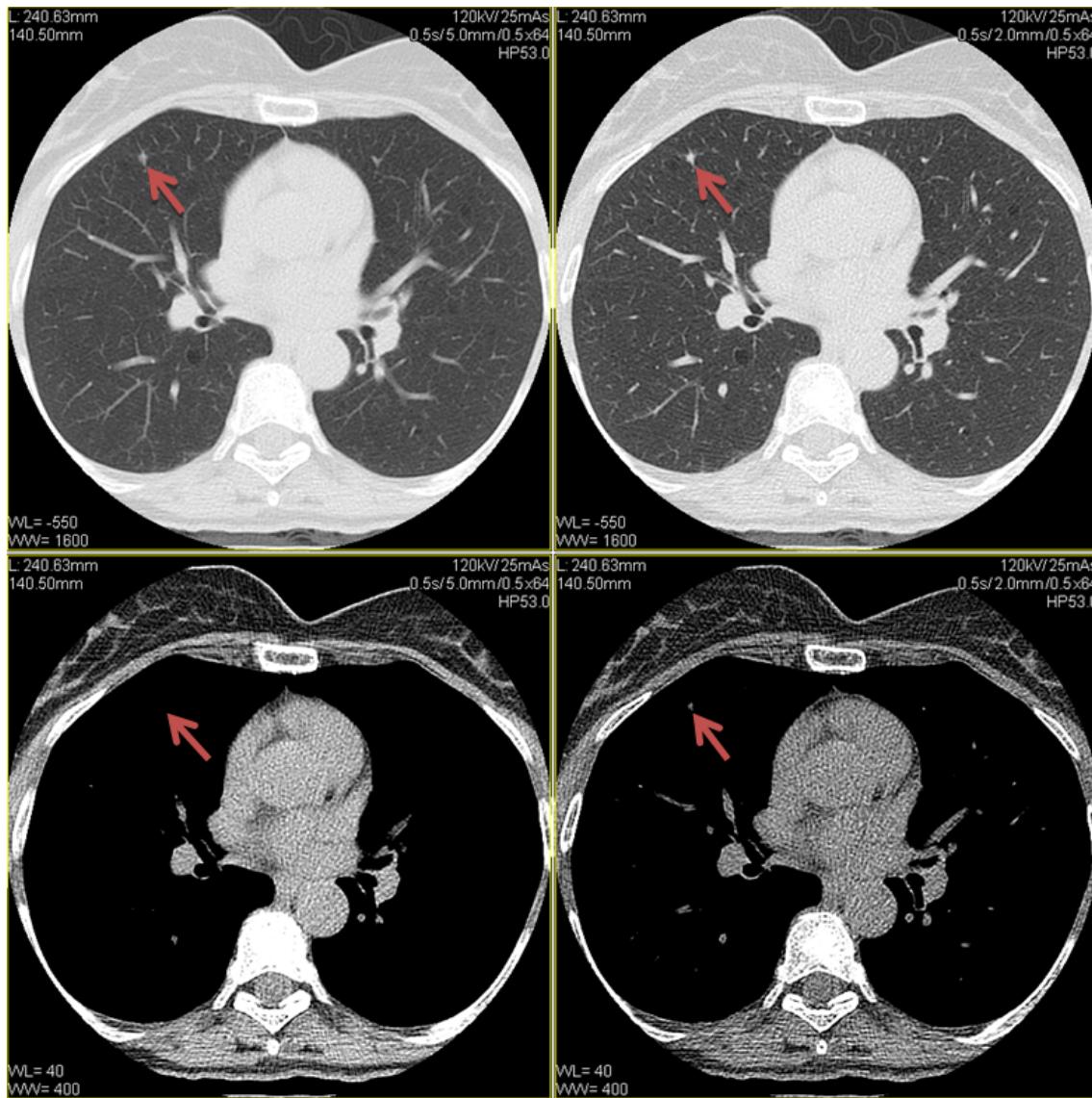


Figure 4.7: A single trans-axial CT images reconstructed at the same table position through the maximum dimension of a sub 5 mm nodule using two different slice-thickness/overlap reconstructions; 5.0/2.5 mm (left column) and 2.0/2.0 mm (right column), with lung (top row) and mediastinal kernel (bottom row) reconstructions. For the larger slice-thickness reconstructions the nodule appears to be a GG but, from the images with a smaller effective slice thickness a solid nodule is clearly indicated.

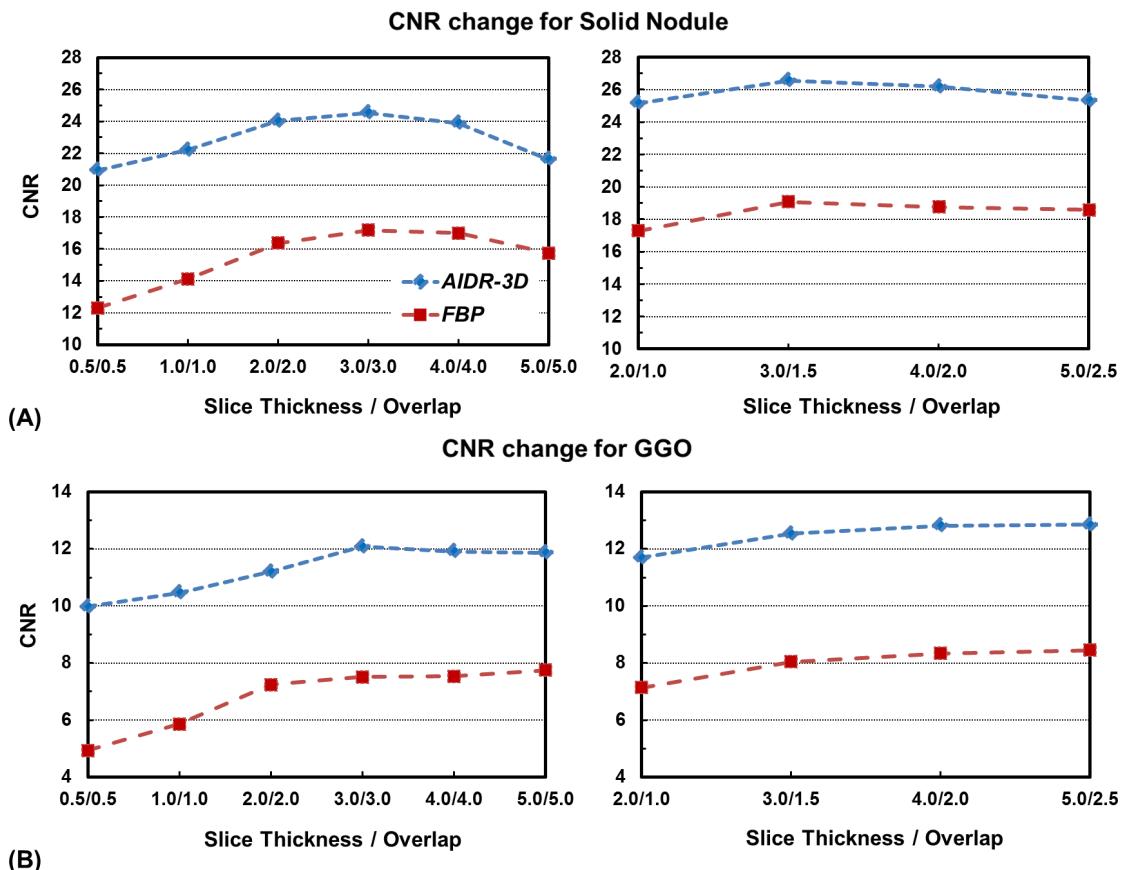


Figure 4.8: Comparison of the changes in CNR with Iterative Reconstruction (AIDR3D) and FBP as a function of effective slice thickness (left: non-overlapped and Right: with 50% overlapped): Mean values of the normalized Contrast to Noise Ratio (CNR) of in vivo nodules as a function of reconstruction slice thickness/overlap for: (A) solid nodules and (B) ground glass nodules.

Chapter 5

CT Reconstruction through MAP Based Compressed Sensing Modeling

This chapter analyzes CT reconstruction process by using a Maximum a Posteriori estimation approach transformed into a weighted-CS problem. The method makes use of the Pseudo-polar Fourier Transform (PPFT) (Section 2.3.1), enabling the Fourier coefficients to be computed directly on pseudo-polar grids, thereby decreasing the interpolation error and greatly decreasing the computation burden. Extension of this approach to fan and helical cone beam scans is made possible by rebinning the projections to parallel rays. The method described incorporates the effects of both the measurement noise and interpolation error.

Measurement noise in CT scanners can best be modeled by a Poisson distribution [43], while the noise considered in classical CS formulations, such as those given by equations (3.15) and (3.22), is white additive Gaussian noise [17, 18]. Consequently, to enable a more accurate low dose CS-based CT reconstruction, the classical CS formulations should be modified. Different approaches have been used in prior studies to address this problem. For instance, in [146] to consider the statistical properties of the low-dose measurements, an iterative SIR based technique followed by TV denoising was used. The penalized weighted least-squares (PWLS) formulation of statistical CT reconstruction was used in [147, 148] to improve the quality of reconstructed images. This PWLS formulation can be characterized as a weighted ℓ_1 minimization problem as proposed by Candes et al. [149], who showed that by using appropriate weights the quality of the recovered signal can be improved. In this modified CS formulation the weights could account for the statistical characteristics of the signals. For example, weighting methods have been proposed by Khajehnejad et al. [150] and by Vaswani and Wei [151] for recovery of the signals with a partial known support and with a priori information about the probability of each entry of the signal being non-zero.

We propose a weighted CS formulation in the framework of statistical CT image reconstruc-

tion algorithms. The weight, denoted by Error Adaptation Weight (EAW), is a function of the rebinning interpolation error [152] and the Poisson noise of the CT projections [148] calculated from a maximum a posteriori (MAP) model of CT reconstruction. Simulation results are shown for phantoms and a patient. For example, a 512×512 Shepp-Logan phantom when reconstructed from 128 rebinned projections using a conventional CS method had 10% error, whereas with the proposed method the reconstruction error was less than 1%. Computation times of less than 30 sec were obtained on a standard desktop computer, compared to couple of hours using the conventional CS-based CT reconstruction algorithms.

5.1 Fan Beam and Helical Cone Beam Image Reconstruction

A brief description of the proposed algorithm is provided in this section. More details are provided in succeeding sections. As mentioned in section 2.3, the computational complexity of CT reconstruction can be significantly reduced by using direct Fourier reconstruction [50]. This avoids the need for multiplication of large matrices but potentially incurs other errors. Since the regular 2D Fourier transform is on Cartesian grids, a polar to Cartesian interpolation is needed. To reduce the interpolation error, PPFT will be used in our proposed method. Therefore, the sampling matrix \mathbf{A} should measure the X-ray projections on the equally sloped radial lines given by equation (2.9).

Although CT scanners typically collect data along equally spaced angles, they have the flexibility to collect data along the angles of a pseudo-polar grid. Then, the equally distant measured data should be interpolated to the pseudo-polar grids, as shown in figure 2.12-B and equation (2.7). The resulting interpolation error can be limited by oversampling the Fourier data by zero-padding the projections on the equally sloped radial lines. In addition, fan beam and helical geometries should be rebinned to equally sloped parallel radial lines, which adds extra interpolation errors to the reconstruction problem, as described in section 2.4. At each interpolation step, the error is tracked to be included in the proposed compressed sensing model along with the measurement noise. The calculated weights and the prepared data are then fed into the proposed FCSA-LEM solver. Consequently, the proposed CT reconstruction method, as shown in figure 5.1, can be summarized by the following two major stages:

- Data preparation and rebinnning: fan or helical projections are mapped to parallel equally sloped radial lines. The output \mathbf{y} of this stage is the X-ray projection samples on pseudo-polar grids, or the Fourier transform of the calculated parallel rays $\mathcal{F}_{1D}\mathbf{y}$, which is the under-sampled pseudo-polar Fourier transform of the CT image. This step is marked by (1) in figure 5.1.

- Image Reconstruction: the CT image is reconstructed with the proposed FCSA-LEM method. The measurement matrix is the discrete Radon transform performed on pseudo-polar grids or its Fourier transform which is the fast pseudo-polar Fourier transform function \mathcal{F}_{1DA} ; and the input data is \mathbf{y} or $\mathcal{F}_{1D}\mathbf{y}$, correspondingly, from the first stage. This step is marked by (2) in figure 5.1 and is described in section 5.3.

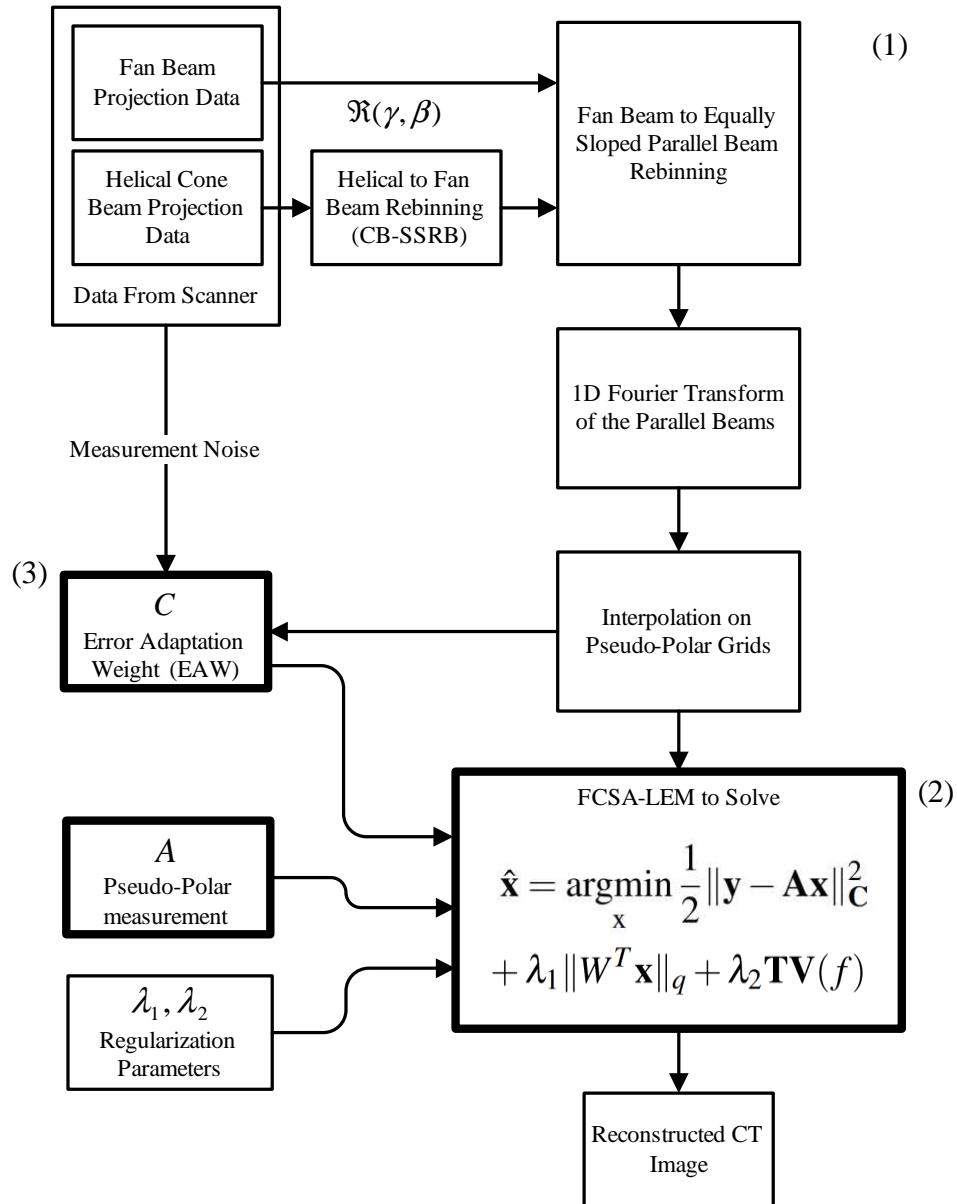


Figure 5.1: Flowchart of the proposed weighted CS-based CT reconstruction method.

5.2 Maximum a Posteriori (MAP) Model of CT

We propose a CT reconstruction method based on maximum a posteriori (MAP) optimization of the projection data with two classes of a priori models: sparsity of the wavelet coefficients and piecewise linearity of the images. We show that the MAP model can be transformed into a weighted CS model with rigorously calculated weights, which depend on the magnitude of the noise sources that arise from both measurement and interpolation.

5.2.1 MAP Model of CT imaging and Weighted-CS formulation

Measured X-ray projection photon numbers $\xi(l, \varphi)$ in equation (2.2) are usually corrupted by two kinds of noise: electrical noise in the detectors and associated circuits (with variance of σ_e^2) and the photon counting noise for which the observed counts are a Poisson distribution about the mean $\bar{\xi}$. If we consider the discrete formulation of equation (2.2), using the second order Taylor series expansion of the Poisson distribution, the log likelihood of the measurements is [153, 154]

$$\log p(\mathbf{y}|\mathbf{x}) \approx -\frac{1}{2}(\mathbf{y} - \mathbf{Ax})^T D(\mathbf{y} - \mathbf{Ax}) + \mathcal{O}(\mathbf{y}^3), \quad (5.1)$$

in which \mathbf{y} denotes the vectorized results of the Radon transform $g(l, \varphi)$, \mathbf{x} denotes the vectorized $f(x, y)$, \mathbf{A} is the projection matrix, and $\mathcal{O}(\mathbf{y}^3)$ is a function which depends upon measured data only. Assuming that the measurement at each detector is independent from other detectors, D would be a diagonal matrix, i.e., ignoring the scattered photons. For the purpose of the MAP estimation $O(\mathbf{y}^3)$ may be ignored since it does not depend on \mathbf{x} . Ignoring this term, (5.1) describes a simplified CT model that can be written as

$$\mathbf{y} = \mathbf{Ax} + \mathbf{n}, \quad (5.2)$$

in which \mathbf{n} is Gaussian distributed noise with a covariance matrix D^{-1} and d_i is the i^{th} diagonal element of D , which is proportional to the detector counts. The latter corresponds to the maximum likelihood of the inverse of the variance of the projection measurements, i.e., to $1/\sigma_{y_i}^2$. From equation (2.2) the i^{th} measured projection y_i is given by

$$y_i = \log \left(\frac{\xi_T}{\xi_i} \right) = \log \left(\frac{\xi_T}{\bar{\xi}_i} \right) + \log \left(\frac{\bar{\xi}_i}{\xi_i} \right) \approx \bar{y}_i + \left(1 - \frac{\xi_i}{\bar{\xi}_i} \right), \quad (5.3)$$

where $\bar{\mathbf{y}}_i$ is noiseless and ξ_i follows the Poisson distribution with $\sigma_\xi^2 = \bar{\xi}_i$. As a result, the variance of projection data can be estimated from

$$\sigma_{\mathbf{y}_i}^2 \approx (\sigma_\xi^2 + \sigma_e^2)(\bar{\xi}_i)^{-2}. \quad (5.4)$$

Using ξ_i as an unbiased estimate of $\bar{\xi}_i$ the diagonal elements of D can be expressed as

$$d_i = \frac{1}{\sigma_{\mathbf{y}_i}^2} = \frac{\xi_i^2}{\sigma_e^2 + \xi_i}. \quad (5.5)$$

To reconstruct the image from the projections, the MAP estimator can be used and this can be written as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} \log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x}). \quad (5.6)$$

Here $h(x) = \log p(\mathbf{x})$ acts as a penalty function, which is used to statistically model the wavelet coefficients distribution and the piecewise linearity of the CT images.

5.2.2 Sparsity and Piecewise Linearity of the CT Images in MAP Model

Many studies [155, 156] have shown that the wavelet transform of a variety of images, $\theta = W^T \mathbf{x}$, can be modeled by Generalized Gaussian Distribution (GGD), i.e., by

$$p(\theta_i) = K(s, q) \cdot \exp\left(-\left|\frac{\theta_i}{s}\right|^q\right), \quad (5.7)$$

where W^T is the wavelet transform, W is its inverse, s and q are the parameters of the GGD and $K(s, q)$ is the normalization parameter. It should be noted that when $q = 1$, the GGD is equivalent to Laplacian distribution and when $q = 2$ it describes a Gaussian distribution. By using equations (5.1), (5.6), and (5.7) the MAP model for CT images can be expressed as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_D^2 + \lambda \|W^T \mathbf{x}\|_q, \quad (5.8)$$

where $\|\mathbf{y} - \mathbf{Ax}\|_D^2 = (\mathbf{y} - \mathbf{Ax})^T D(\mathbf{y} - \mathbf{Ax})$, λ is a function of GGD parameters. Typically, q is chosen to be in the range $0 < q \leq 1$, and θ is a sparse representation of the image $\mathbf{x} = W\theta$, and $\|\mathbf{x}\|_q = \sum_i |x_i|^q$. Another prior on $p(x)$ is the piecewise linearity of the images. A ρ -variation distribution is proposed to describe piecewise constant functions [157]. If $x_n(t) = \sum_{j=1}^n x_j^n \Psi_j^n(t)$ is a piecewise function spanned by the roof-top basis $\Psi_j^n(t)$, the following class of probability

distribution can be used to describe it:

$$p(x_1^n, \dots, x_n^n) = v_{\rho,n} \exp \left(-\frac{a_n}{(n+1)^{1-\rho}} \sum_{j=1}^{n+1} |x_j^n - x_{j-1}^n|^\rho \right), \quad (5.9)$$

where $a_n > 0$, $x_0^n = x_{n+1}^n = 0$, $v_{\rho,n}$ is normalizing factor, and $[x_1^n, \dots, x_n^n]^T$ is a \mathbb{R}^n -valued random vector. When $\rho = 1$, this yields the total variation norm. By making use of this, equations (5.1) and (5.6) enable the following MAP model to be obtained for CT images:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_D^2 + \lambda \mathbf{TV}(f). \quad (5.10)$$

As can be seen, equations (5.10) and (5.8) are generalized forms of the CS models given by equations (3.22) and (3.15), respectively.

It has been previously shown that the quality of the reconstructed image can be improved by combining the sparsity and total variation penalty terms [24]. As a result, the CT images can be recovered from undersampled data by using:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_D^2 + \lambda_1 \|W^T \mathbf{x}\|_q + \lambda_2 \mathbf{TV}(f), \quad (5.11)$$

where λ_1 and λ_2 are the regularization parameters. *In summary, MAP estimation of CT images is equivalent to a weighted CS problem, as shown in equation (5.11), with weights calculated in equation (5.5).*

5.2.3 Proposed CS-based CT Formulation incorporating Measurement Noise and Interpolation Error

In section 5.2.1 it was shown that the MAP estimator of CT is a form of weighted CS problem given by (5.11), in which the weight is a function of noise variance and is denoted by D in equation (5.5). To avoid the computational burden associated with the huge projection matrix of CS-based CT, our proposed algorithm makes use of the fast PPFT-based Radon transform described in Section 2.4. This not only accelerates the computations by reducing the computational complexity, but also substantially reduces the gridding error and eliminates the regridding step. The latter is possible because of the fast inverse algorithm which eliminates the need to regrid the updated image to pseudo-polar grids at each iteration.

As described in Section 2.4 the proposed method is generalized for use with non-parallel geometries by rebinding the X-ray beams onto equally sloped radial lines as given by (2.9). The rebinding step causes interpolation error to the measured data, which propagates in each it-

eration of CS-based CT reconstruction. This problem has not received much attention in the literature. Fahimian et al. [111] proposed an EST method for reconstructing fan beam and helical cone beam images, in which they overcome the rebinning interpolation problem at each iteration by using a non-local total variation minimization smoothing step. In the method proposed by Hashemi et al. [152] an ℓ_2 -TV optimization scheme was used to reconstruct the CT images from fan beam projections. To compensate for the interpolation error, a confidence matrix was added to the CS scheme, enabling control of the error propagation in successive iterations.

To control the interpolation error we make use of a MAP model of the CT reconstruction process. Denoting the variance of interpolation error by e_i , the variance of the measurements is $\sigma_{y_i}^2 \approx (\sigma_\xi^2 + \sigma_e^2 + e_i)(\bar{\xi}_i)^{-2}$. Using (5.3), the effects of both the noise variance and interpolation error can be lumped together into the form of an Error Adaptation Weight (EAW), denoted by a diagonal matrix \mathbf{C} with diagonal elements $c_i = \frac{\xi_i^2}{\sigma_e^2 + \xi_i + e_i}$. Assuming a nearest neighbor interpolation, the interpolation error is linearly dependent on the distance between the desired and the measured grids, i.e., $e_i = \varepsilon_i \times \xi_i$ in which $\varepsilon_i \in [0, \infty)$ models the interpolation distance. Consequently, if the dose at each projection is high enough to ignore the electric noise σ_e , the EAW can be rewritten as

$$c_i = \frac{\xi_i^2}{\xi_i + \varepsilon_i \times \xi_i} = \xi_i \times \frac{1}{1 + \varepsilon_i}. \quad (5.12)$$

Using this definition, our proposed CS formulation to reconstruct the CT images can be expressed by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_{\mathbf{C}}^2 + \lambda_1 \|W^T \mathbf{x}\|_q + \lambda_2 \mathbf{TV}(f). \quad (5.13)$$

This implicitly models the effect of polar and non-parallel projections to pseudo-polar gridding, which in turn affects the noise in the data. This error is considered to be linearly dependent on the interpolation distance, while it is typically smaller using more accurate interpolation methods, i.e., in Kaiser-Bessel based interpolation [158]. Therefore, the proposed CS formulation can be thought of as a minimization of an upper bound of the pseudo-polar rebinning error. In addition, \mathbf{C} in this formulation acts similar to a Jacobi preconditioner and therefore could accelerate the convergence rate.

5.2.4 Calculation of ε

The value of ε_i represents the error of the interpolated samples. If the interpolated sample is close to the original measurements, the value of ε_i is small and the confidence about the interpolated value is high. If the angular distance of the measured data from the interpolated line is more than the angular difference of the equally sloped lines, the interpolation error is

considered to be high: this follows from the fact that the distance of points on the line from the true measured values are maximal and therefore the error is maximal. Using equation (5.12), this condition corresponds to $c_i \rightarrow 0$. The closer the equally sloped lines are to the rays on which the measurements are made, the smaller the interpolation error is; and ε_i s on that line get closer to zero. Finally, if the desired equally sloped rays are exactly on the polar lines, the interpolation error e_i is zero, which is equivalent to $\varepsilon_i = 0$. This process is illustrated in figure 5.2.

In practice, $\left(\frac{1}{1+\varepsilon_i}\right)$ can be estimated by rebinning an all-ones matrix with the same size as the measured data onto the equally sloped radial angles following by an interpolation on the pseudo-polar grid. Note that this has to be calculated only once before the reconstruction.

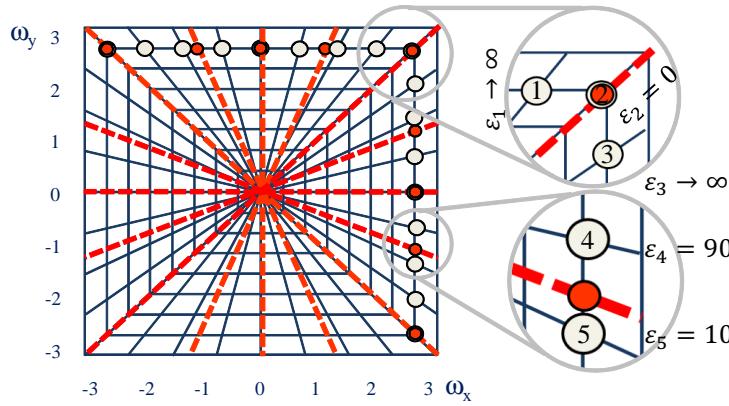


Figure 5.2: Illustration of the Error Adaptation Weight (EAW) for a simple example. The equiangular lines are plotted with red dashed lines and smaller red circles show the measurements. The solid lines and the gray larger circles represent the pseudo-polar grids.

5.3 Expectation Maximization based Solution for the Proposed Weighted CS-based CT Reconstruction

In solving the proposed optimization problem, Fast Composite Splitting Algorithm (FCSA) [159, 160, 26] was used to decompose equation (5.13) into two simpler sub-problems given by

$$\begin{aligned}\hat{\mathbf{x}}_1 &= \underset{\mathbf{x}}{\operatorname{argmin}} f_1(\mathbf{x}), \quad f_1(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_{\mathbf{C}}^2 + \lambda_1 \|\mathbf{W}^T \mathbf{x}\|_q \\ \hat{\mathbf{x}}_2 &= \underset{\mathbf{x}}{\operatorname{argmin}} f_2(\mathbf{x}), \quad f_2(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_{\mathbf{C}}^2 + \lambda_2 \mathbf{TV}(f),\end{aligned}\tag{5.14}$$

which are similar to equations (5.10) and (5.8). By calculating $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$, the FCSA method proposes that the solution to the problem can be obtained by a linear combination of the solu-

tions of the two sub-problems, *i.e.*,

$$\hat{\mathbf{x}} = \Delta \hat{\mathbf{x}}_1 + (1 - \Delta) \hat{\mathbf{x}}_2, \quad (5.15)$$

in which $\Delta = f_2/(f_1 + f_2)$ is a function of the value of the objective functions of the two subproblems. Thus, for recovering CT images from the X-ray projections we propose an FCSA based EM method, called the FCSA-LEM algorithm.

5.3.1 Latent Image and Expectation Maximization Method for ℓ_1 and TV minimization

Here, we propose an efficient method to solve the ℓ_1 and TV sub-problems given by equation (5.14). Specifically, an EM algorithm is used to solve these two optimizations problems [161, 162] for a CT modeled by equation (5.2). A latent variable \mathbf{z} is defined such that the model described by equation (5.2) can be written as¹:

$$\mathbf{y} = \mathbf{Az} + n_1, \quad (5.16)$$

in which \mathbf{z} is chosen to be:

$$\mathbf{z} = \mathbf{x} + n_2 = W\theta + \alpha n_2. \quad (5.17)$$

The noise is split into two parts: $\mathbf{n} = \alpha \mathbf{An}_2 + n_1$ with $p(n_1) = N(n_1; 0, D^{-1} - \alpha^2 \mathbf{AA}^T)$, $p(n_2) = N(n_2; 0, I)$. Using this notation, the EM algorithm consists of two steps:

E-step: Compute the conditional expectation of the log likelihood, given the observed data and the current estimate $\theta^{(t)} = W^T \mathbf{x}^{(t)}$ *i.e.*,

$$Q(\theta, \theta^{(t)}) = E[\log p(\mathbf{y}, \mathbf{z}|\theta)|\mathbf{y}, \theta^{(t)}]. \quad (5.18)$$

M-Step: Update the estimate using

$$Q^{(t+1)} = \underset{\theta}{\operatorname{argmax}} \{Q(\theta, \theta^{(t)}) - h(\theta)\}. \quad (5.19)$$

In the E-step, $\mathbf{z}^{(t)} = E[\mathbf{z}|\mathbf{y}, \theta^{(t)}]$ needs to be calculated and substituted into equation (5.18), in which we need to calculate the likelihood $p(\mathbf{y}, \mathbf{z}|\theta) = p(\mathbf{y}|\mathbf{z}, \theta)p(\mathbf{z}|\theta) = p(\mathbf{y}|\mathbf{z})p(\mathbf{z}|\theta)$. Since $p(\mathbf{z}|\mathbf{y}, \hat{\theta}^{(t)}) \propto p(\mathbf{y}|\mathbf{z})p(\mathbf{z}|\hat{\theta}^{(t)})$, where both $p(\mathbf{y}|\mathbf{z})$ and $p(\mathbf{z}|\hat{\theta}^{(t)})$ are Gaussian, therefore $p(\mathbf{z}|\mathbf{y}, \hat{\theta}^{(t)})$ is also Gaussian with mean value of $\mathbf{z}^{(t)} = E[p(\mathbf{z}|\mathbf{y}, \hat{\theta}^{(t)})]$ [163, 110]. Using the fol-

¹The definition and the effect of this latent variable in the final solution is similar to the variable splitting strategy used in alternating direction methods (ADM) [98, 24].

lowing equations $p(\mathbf{y}|\mathbf{z}) \sim N(\mathbf{y}; \mathbf{Az}, D^{-1} - \alpha^2 \mathbf{AA}^T)$, $p(\mathbf{z}|\theta) \sim N(\mathbf{z}; 0, \alpha^2 I)$, and the Maximum Penalized Likelihood Estimate, the E-step results in

$$\mathbf{z}^{(t)} = \mathbf{x}^{(t)} + C_2 \mathbf{A}^T (\mathbf{AC}_2 \mathbf{A}^T + C_1)^{-1} (\mathbf{y} - \mathbf{Ax}^{(t)}), \quad (5.20)$$

in which $C_2 = \alpha^2 I$ and $C_1 = D^{-1} - \alpha^2 \mathbf{AA}^T$. Therefore, the E-step can be summarized by the calculation of:

$$\mathbf{z}^{(t)} = \mathbf{x}^{(t)} + \alpha^2 \mathbf{A}^T D (\mathbf{y} - \mathbf{Ax}^{(t)}). \quad (5.21)$$

By including the rebinning error, *i.e.* using the EAW introduced in equation (5.13), this step consists of calculating

$$\begin{aligned} \mathbf{z}^{(t)} &= \mathbf{x}^{(t)} + \alpha^2 \mathbf{A}^T (D \text{diag}(1 + \varepsilon)^{-1}) (\mathbf{y} - \mathbf{Ax}^{(t)}) \\ &= \mathbf{x}^{(t)} + \alpha^2 \mathbf{A}^T \mathbf{C} (\mathbf{y} - \mathbf{Ax}^{(t)}), \end{aligned} \quad (5.22)$$

which is followed by an M-step. To find $Q(\theta, \theta^{(t)})$ in the M-step, using $\log p(\mathbf{y}, \mathbf{z}|\theta) = -1/(2\alpha^2)(\mathbf{x}-\mathbf{z})^T(\mathbf{x}-\mathbf{z}) + K'$, the following equation will be used: $Q(\theta, \theta^{(t)}) = -1/(2\alpha^2)\|\mathbf{x} - \mathbf{z}\|_2^2 + K'$, in which K' does not depend on \mathbf{x} or θ . Consequently, in the M-Step the following optimization problem must be solved:

$$\begin{aligned} \mathbf{x}^{(t+1)} &= \underset{\mathbf{x}}{\operatorname{argmax}} \left\{ \frac{1}{2\alpha^2} \|\mathbf{x} - \mathbf{z}\|_2^2 + h(\mathbf{x}) \right\} \\ \theta^{(t+1)} &= \underset{\theta}{\operatorname{argmax}} \left\{ \frac{1}{2\alpha^2} \|W\theta - \mathbf{z}^{(t)}\|_2^2 + h(\theta) \right\}. \end{aligned} \quad (5.23)$$

Equation (5.23) is a denoising problem, which can be solved for some special cases, e.g. soft thresholding if $h(\theta) = \lambda_1 \|\theta\|_1$ [164] and the TV denoising problem if $h(\theta) = \lambda_2 \mathbf{TV}(W\theta)$ [99].

5.3.2 Fourier Based E-Step Formulation

The computational complexity of CS recovery can be reduced from $O(N^3)$ to $O(N^2 \log N)$ by utilizing DFR and fast PPFT algorithm. Using the EM algorithm enables the inclusion of the pseudo-polar Fourier transform in the formulation together with the effects of the measurement noise and interpolation error. This is possible due to the orthogonality of Fourier transform. If

\mathcal{F}_{1D} denotes the 1D Fourier transformed matrix, we can re-write $\log p(\mathbf{y}|\mathbf{z})$ as

$$\begin{aligned}\log p(\mathbf{y}|\mathbf{z}) &= -\frac{1}{2\alpha^2}(\mathbf{y} - \mathbf{Az})^T(D^{-1} - \alpha^2\mathbf{AA}^T)^{-1}(\mathbf{y} - \mathbf{Az}) + K \\ &= -\frac{1}{2\alpha^2}(\mathbf{y} - \mathbf{Az})^T(\mathcal{F}_{1D}^T\mathcal{F}_{1D})(D^{-1} - \alpha^2\mathbf{AA}^T)^{-1}(\mathcal{F}_{1D}^T\mathcal{F}_{1D})(\mathbf{y} - \mathbf{Az}) + K \\ &= -\frac{1}{2\alpha^2}(\mathcal{F}_{1D}\mathbf{y} - \mathcal{F}_{1D}\mathbf{Az})^T\mathcal{F}_{1D}(D^{-1} - \alpha^2\mathbf{AA}^T)^{-1}\mathcal{F}_{1D}^T(\mathcal{F}_{1D}\mathbf{y} - \mathcal{F}_{1D}\mathbf{Az}) + K,\end{aligned}\tag{5.24}$$

where $\mathcal{F}_{1D}\mathbf{y}$ is the Fourier transform of the measurements and K does not depend on \mathbf{z} . Since the measurements \mathbf{y} are on the pseudo-polar grids, $\mathcal{F}_{1D}\mathbf{y}$ can be calculated by fast pseudo-polar Fourier algorithm. This means that the combination of $\mathcal{F}_{1D}\mathbf{A}$ is the forward pseudo-polar Fourier transform that is used in our algorithm and replaces \mathbf{A} in equation (5.22); and $(\mathcal{F}_{1D}\mathbf{A})^T$ is the conjugate transpose of pseudo-polar transform. Replacing \mathbf{A} with $\mathcal{F}_{1D}\mathbf{A}$, \mathbf{A}^T with $(\mathcal{F}_{1D}\mathbf{A})^T$, and C_1^{-1} with $\mathcal{F}_{1D}(D^{-1} - \alpha^2\mathbf{AA}^T)^{-1}\mathcal{F}_{1D}^T$, equation (5.22) will remain the same.

5.3.3 FCSA-LEM Algorithm in a Nutshell

The pseudo-code shown in Algorithm 5 outlines the FCSA-LEM method used to solve the generalized CS problem as derived in section 5.2.3 and expressed by equation (5.13). It uses the method introduced in section 5.3.1, which uses equations (5.22) and (5.23) to reconstruct the CT images. In this algorithm $\text{prox}\{g(x), z\} = \underset{x}{\operatorname{argmin}} g(x) + \frac{1}{2}\|x - z\|_2^2$. To find $\hat{\mathbf{x}}_1$, the optimization problem in step 2 of Algorithm 5 is solved using a wavelet soft thresholding algorithm proposed by Hashemi et al. [155]. Moreover, to calculate $\hat{\mathbf{x}}_2$ in step 3, split Bregman TV based denoising algorithm proposed in [99], discussed in section 3.5 and Algorithm 4, is used. Then, $\hat{\mathbf{x}}_k$ is estimated in k^{th} iteration by equation (5.15). In this algorithm, \mathbf{A} is a fast pseudo-polar Fourier based Radon transform, \mathbf{A}^T is its adjoint, and \mathbf{y} is the interpolated measurements on the pseudo-polar grids.

5.4 Simulation Methods and Results

5.4.1 Phantoms and CT measurements

Fan beam simulations were performed using a Shepp-Logan software phantom that is available in MATLAB, two types of experimental phantoms, and a clinical patient. The two phantoms were: a custom made phantom that mimics different cardiac plaques, and a Lungman chest

Algorithm 5 Pseudo-code of **FCSA-LEM** algorithm used to solve the optimization problem.

Initialize: $\alpha, \mu_1, \mu_2, \mathbf{c}, r_1 = 0, t_1 = 1, \text{maxiter, tol}$

while $\frac{\|\hat{x}_k - \hat{x}_{k-1}\|_2}{\|\hat{x}_k\|_2} > \text{tol}$ or $k < \text{maxiter}$ **do**

```

1       $\mathbf{z} = r_k + \alpha^2 \mathbf{A}^T \mathbf{C}(\mathbf{y} - \mathbf{A}r_k)$ 
2       $\hat{\mathbf{x}}_1 = W(\text{prox}\{\lambda_1 \|W^T \mathbf{x}\|_1, W^T \mathbf{z}\})$ 
3       $\hat{\mathbf{x}}_2 = \text{prox}\{\lambda_2 \mathbf{TV}(\mathbf{x}), \mathbf{z}\}$ 
4       $\hat{\mathbf{x}}_k = \Delta \hat{\mathbf{x}}_1 + (1 - \Delta) \hat{\mathbf{x}}_2$ 
5       $t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$ 
6       $r_{k+1} = \hat{\mathbf{x}}_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_{k-1})$ 
7       $k \leftarrow k + 1$ 

```

endwhile

phantom. X-ray projections of the phantoms and the patient were taken using a Toshiba Aquilion ONE[©] scanner (Toronto General Hospital, Canada). This cone-beam system has 320 rows of detectors along the z-axis, with each row consisting of 896 elements. The scanner gathers data from 900 projection angles in each 360° rotation, and this is achieved in 0.5 sec. Unless otherwise stated, when the images were reconstructed from fewer than the 900 projections the projection views were selected equiangularly. For all the scan protocols the X-ray tube current-exposure time product was 50 mAs and the peak voltage was 120 kV. This current/voltage is high enough to ignore the electric noise in the simulations. Data from the central row of a volumetric scan on one single rotation served as the fan beam data.

A 3D cylindrical software phantom is used to test the helical cone beam reconstruction. This consist of a 80 cm long cylinder, centered at $(x = 0, y = 0, z = 0)$, with a diameter of 30 cm, and having a CT# of 0.1². It contains three spheres whose diameters are 3.0 cm and one sphere with a diameter of 5.0 cm. The centers of the three smaller spheres and CT#'s are: $(x = -4, y = -14, z = -4)$, CT#=0.4; $(x = -4, y = 14, z = -4)$, CT#=0.2; and $(x = -2, y = 10, z = -4)$, CT#=0.3. It should be noted that all three have the same z-location, and two of them partially overlap. The largest sphere is at $(x = 16, y = 2, z = 2)$ and has CT#=0.2.

For all the simulations, Daubechies wavelets with four vanishing moments in 5 levels were used. Moreover, by manual tuning it was found that the best values for λ_1 and λ_2 in (5.14) are given by $\lambda_1 = 0.05 \|W^T(\mathbf{A}^T \mathbf{y})\|_\infty$ and $\lambda_2 = 1 \times 10^{-3} \mathbf{TV}(\mathbf{A}^T \mathbf{y})$. Clinically high quality thoracic CT images were achieved with these values. For chest scans it was found that the numerical factors in both could be varied over a wide range without affecting the final image quality. However, it should be noted that for other organs it seems that these parameters could be ad-

²Since this is a software phantom the CT numbers given are relative contrast numbers.

justed in a relatively straightforward manner to achieve sharper or smoother images, similar to the different CT reconstruction kernels in the clinical CT scanners which are optimized for different applications and organs.

5.4.2 ART-TV based method

The proposed method is compared with a conventional ART-TV based CT reconstruction algorithm. The convergence of this method is justified by Projection on Convex Set (POCS) algorithm. As an example in [165] an ART-TV based CT reconstruction method is proposed denoted by adaptive-steepest-decent POCS (ASD-POCS), which has been used by many other researchers [166, 167]. Here, we use a simple method in which the updates are calculated using an ART based method described below. This step is followed by a TV minimization step to project the updated image on a piecewise constant space. To solve the $\mathbf{y} = \mathbf{Ax}$ problem in ART step, a randomized Kaczmarz algorithm is used [168]. If $\mathbf{y} = \mathbf{Ax}$ is a linear system of equations and \mathbf{x}_0 is an arbitrary initial approximation to the solution, randomized Kaczmarz applies the following updating step at each iteration

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\mathbf{y}_{r(i)} - \langle a_{r(i)}, \mathbf{x}_k \rangle}{\|a_{r(i)}\|_2^2} a_{r(i)}, \quad (5.25)$$

where $r(i)$ is chosen randomly from the set $\{1, 2, \dots, m\}$, with probability proportional to $\|a_{r(i)}\|_2^2$, a_i is the rows of \mathbf{A} , and $\langle ., . \rangle$ is the inner product of two vectors. Using this algorithm followed by the split Bregman TV minimization (denoted by SBROF in Algorithm 6), the ART-TV based method is described in Algorithm 6. In our simulations the number of inner iterations used in Kaczmarz algorithm is 10 and the number of outer iterations is 50.

Algorithm 6 Pseudo-code of ART-TV based method used in simulations.

Initialize: $r_1 = 0$, inner, outer, λ , tol

```

while  $\frac{\|\mathbf{x}_k - \mathbf{x}_{k-1}\|_2}{\|\mathbf{x}_k\|_2} > \text{tol}$  or  $k < \text{outer}$  do
    1       $r = \mathbf{y} - \mathbf{Ax}_k$ 
    2       $x_{new} = \text{randKaczmarz}(\mathbf{A}, r, \text{inner})$ 
    3       $\mathbf{x}_k = \text{SBROF}(\mathbf{x}_{k-1} + x_{new}, \lambda)$ 
    4       $k \leftarrow k + 1$ 
endwhile
```

A major improvement in the proposed method, compared to the other CS-based reconstruction techniques, is its much reduced computational burden. Figure 5.3 compares the recovery time

of 1) filtered back projection (FBP), 2) the proposed method (FCSA-LEM), and 3) the ART-TV based method described in Algorithm 6. It can be seen that for a 512×512 image the recovery time for the proposed method is approximatively 10-30 sec., and approaches that of FBP when the image size is 1024×1024 . Since the sparsity of the Shepp-Logan phantom increases by increasing its size, the ART-TV and FCAS-LEM algorithms converge faster for larger Shepp-Logan phantom. It should be noted that this figure is based on 200 projections. Increasing the number of projections, increases the size of measurement matrix \mathbf{A} and therefore increases the reconstruction time, while does not affect the proposed method (FCSA-LEM). Figure 5.4

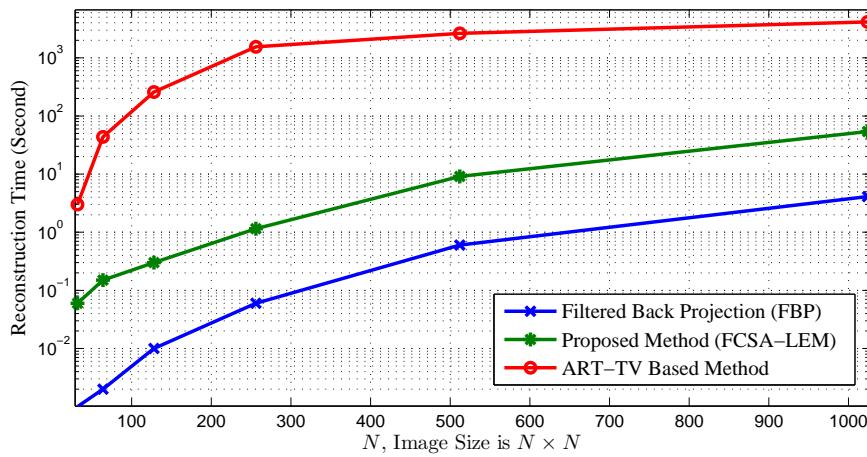


Figure 5.3: Reconstruction time comparison using a standard desktop computer, for: (1) fan beam filtered back projection (FBP) reconstruction, (2) the proposed method (FCSA-LEM), and (3) a fan beam ART-TV based method.

shows the phantom reconstructed by ART-TV described in Algorithm 6 from 200 projections. The normalized reconstruction error $\frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}$ is 2×10^{-2} .

5.4.3 Results and Discussion

Equiangular fan beam projections of the Shepp-Logan phantom were computed on 128 projection views. This data was then rebinned to parallel rays on equally-sloped angles as described by equation (2.9). Figure 5.5 compares the reconstructed 512×512 images with the original image using 1) the inverse pseudo-polar Fourier transform (least squares method), 2) an iterative soft threshold-based method (TwIST) without consideration of the weights [27], and 3) the proposed FCSA-LEM method. Note that the lower row consists of an expanded view of the marked central region, from which it is clear that the FCSA-LEM method yields results very close to that shown in figure 5.5-A. Based on the same phantom, figure 5.6 compares the accuracy of the reconstruction error for all three methods as the number of projections is varied

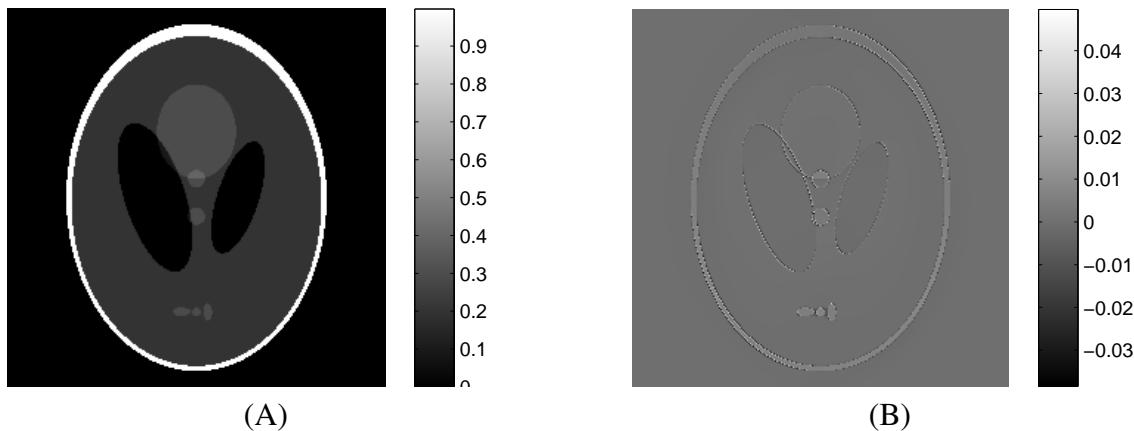


Figure 5.4: ART-TV Reconstruction results: (A) Shep-Logan phantom reconstructed from 200 projections, and (B) Reconstruction Error.

from 50 to 1000. Since the purpose of this particular study was to examine effects of EAW inclusion on rebinning interpolation error, noise was not included in the simulations. Both of these figures show that the recovery accuracy is improved significantly by the inclusion of EAW to correct for rebinning errors. In particular, figure 5.6 shows that the use of more than 300 projections for a 512×512 image does not significantly affect the reconstruction accuracy in absence of noise.

To examine the effects of EAW on both rebinning errors and measurement noise, 100 equiangular projections through the Shepp-Logan phantom were computed and Poisson noise was added to the projections. Figure 5.7 shows the effect of EAW inclusion for different input peak signal-to-noise ratios (PSNR) on the PSNR of the reconstructed images. Images were reconstructed with the proposed method once with including EAW that is calculated by equation (5.5) and once without EAW. As can be seen, the PSNR is improved when the input noise is larger (small input PSNR) and its effect is less when the noise is low (larger input PSNR). The input PSNR was measured from the FBP reconstructed images.

Results obtained from the custom fabricated cardiac plaque phantom are shown in figure 5.8, which provides a comparison of reconstructions using FBP from 900 projections with the results obtained from two datasets: 1) 200 equiangular projections and 2) 450 projections gathered from the white areas shown in figure 5.8-D. In regard to the latter, it should be noted that meanwhile the CT scanners are unable to switch their X-ray sources on and off fast enough to achieve the proposed equiangular simulations. To overcome this problem, the mask used in figure 5.8-D is used that addresses this concern by turning the X-ray source off in the black areas over a range of angles, and then turning it on in the white areas.

The reconstructions computed from the X-ray projections through the Lungman chest phantom are shown in figure 5.9. These images provide a comparison between FBP images obtained

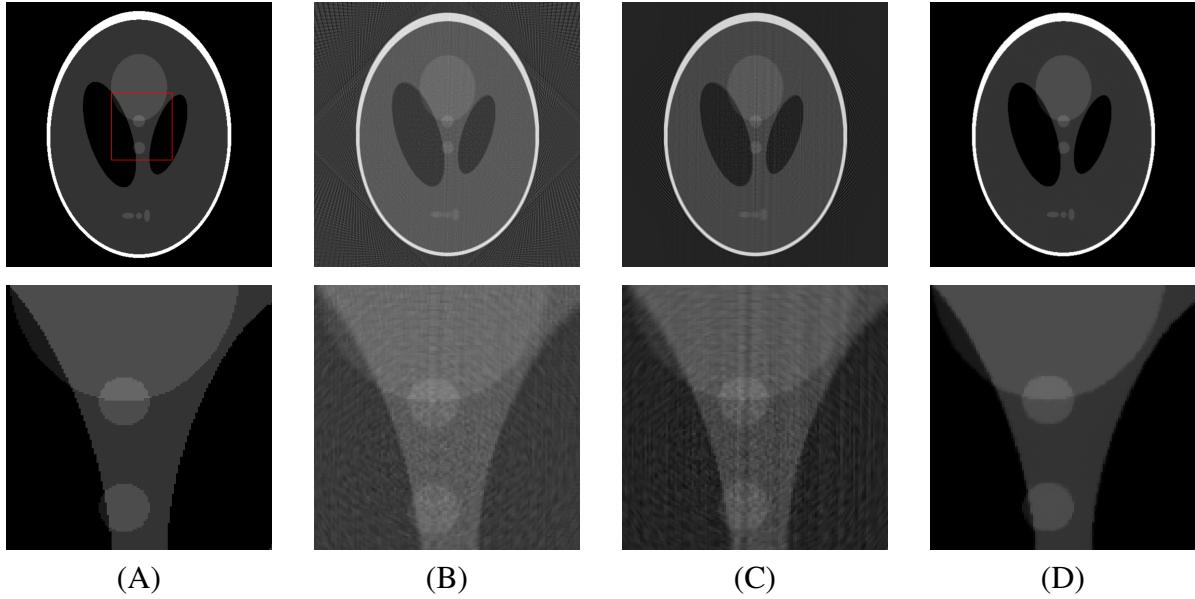


Figure 5.5: Simulation results for a Shepp-Logan phantom. The bottom row shows an expanded view of the region marked in panel (A). (A) Original phantom image. Reconstructions using 128 projections with (B) inverse pseudo-polar Fourier transform using the least squares method (normalized error ≈ 0.9), (C) an iterative soft thresholding based method (TwIST) (normalized error $\approx 10^{-1}$), and (D) the proposed FCSA-LEM method (normalized error $\approx 10^{-2}$). Rebinned parallel rays were used in all three methods to recover the image.

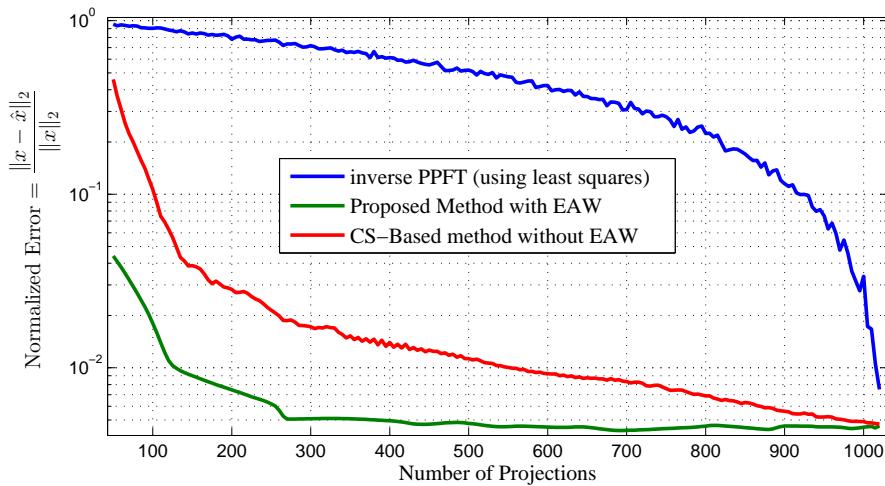


Figure 5.6: Normalized reconstruction error for the simulated Shepp-Logan phantom reconstructed with the inverse pseudo polar Fourier transform (PPFT) using least squares (LS) method, using an iterative soft thresholding based method (TwIST), and the proposed FCSA-LEM method. Rebinned parallel rays were used in all three methods to recover the 512×512 image.

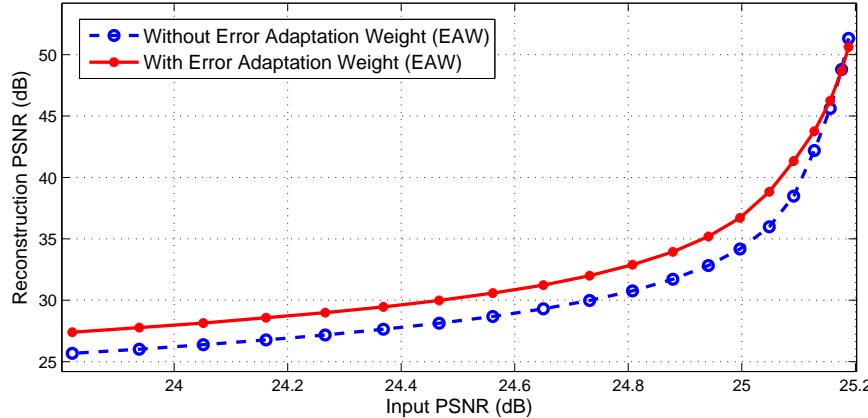


Figure 5.7: Effect of Error Adaptation Weight (EAW) on the PSNR of the reconstructed image. Solid red line shows the results when EAW was included and dashed blue line shows the reconstructed image results when EAW is not included.

from 900 projections with 200 projections using the proposed method. From a qualitative perspective it can be seen that the image quality of the images reconstructed by FCSA-LEM with fewer projections and noise standard deviation of 22HU are better than the FBP reconstructed images from 900 projections with a noise standard deviation of 50HU. By comparison, the middle panel, which was reconstructed from 200 projections using FBP, suffers from a significant streak artifact. Finally, figure 5.10 shows reconstructions of a chest CT scan from a hospital patient using FBP from 900 projections and the proposed method with 200 projections. Again, it is evident that the image reconstructed with the proposed method has almost the same quality as the FBP with all the details preserved and thus, in comparison with FBP the FCSA-LEM method uses about four times less projections, i.e. the radiation dose is decreased by a factor of four. Note that the middle panel (200 projections, FBP) shows serious artifacts. The number of projections are chosen based on the results shown in figure 5.6, in which the reconstruction error of 200 projections is small enough.

Figure 5.11 shows the results for the simulated 3D phantom described earlier in this section, using the helical rebinnning method discussed in section 5.4.1. With reference to figure 2.10, the helix source position was defined as $\psi = [R\cos(\phi), R\sin(\phi), P\frac{\phi}{2\pi}]$ with a pitch factor $P = 0.5$ and the distance of X-ray source from origin was $R = 541\text{cm}$. The scanner was assumed to have 896 detector channels with 32 rows in z -direction. Each sub-image contains 320×320 pixels with a 2 mm pixel size. A total of 3600 projections were taken through the object and 1200 of them were chosen equiangularly and used in the FCSA-LEM algorithm. As can be seen, aside from the start or end of the scan, the reconstruction is almost perfect. However, when the image is close to one of the endpoints the error of rebinnning increases and as a result

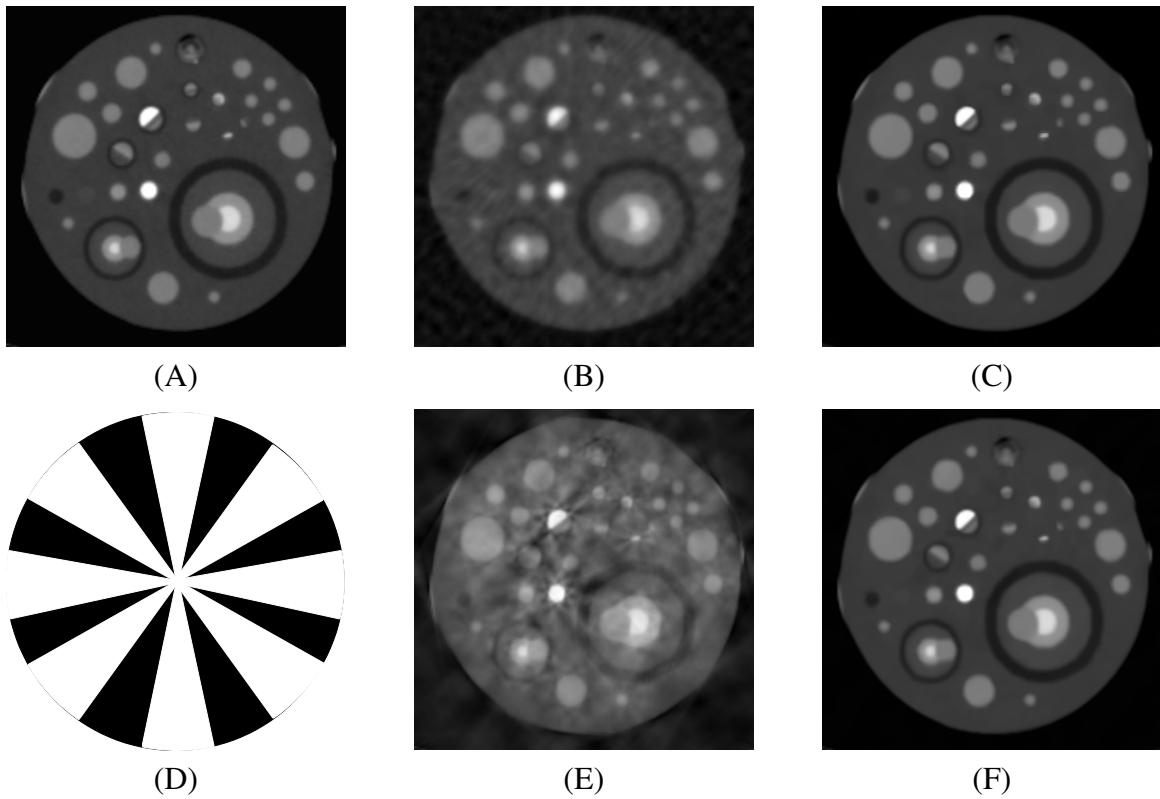


Figure 5.8: Comparison of FBP and the proposed method for a custom built cardiac plaque phantom. (A) Image reconstructed from 900 projections using FBP, (B) Image reconstructed from 200 equiangular projections with FBP, (C) Image reconstructed from 200 equiangular projections with the proposed method, FCSA-LEM, (D) scan protocol used in images (E&F), in which the projections are taken within the white areas, (E) FBP reconstructed image from the 450 projections gathered from the mask shown in image (D), and (F) FCSA-LEM reconstructed image from projections shown in image (D).

the image reconstruction error increases³.

³To calculate the helical projections and the rebinned fan beam data (CB-SSRB rebinning method), the IRT package provided by Fessler et al. (<http://web.eecs.umich.edu/~fessler/irt/irt/>) was used.

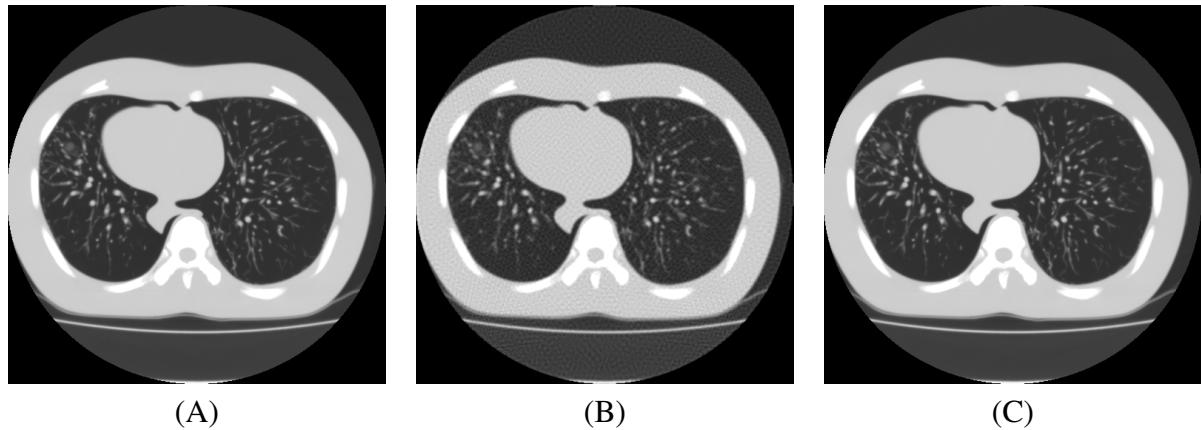


Figure 5.9: Comparison of FBP and the proposed method for a Lungman CT phantom. (A) Image reconstructed from 900 projections with FBP (noise standard deviation 50HU), (B) Image reconstructed from 200 projections with FBP, (C) Image reconstructed from 200 projections with the proposed method, FCSA-LEM (noise standard deviation 22HU).

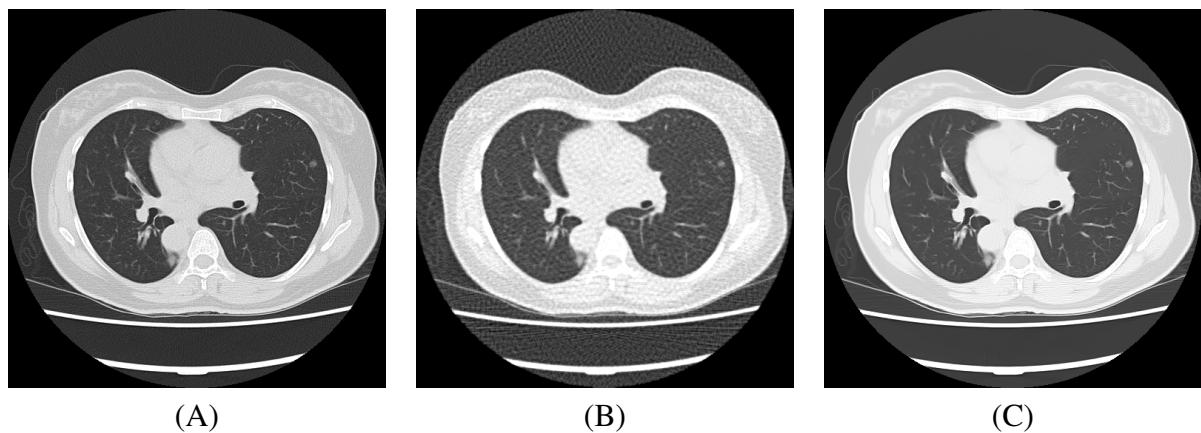


Figure 5.10: Comparison of FBP and the proposed method for the chest CT scan data from a patient. (A) Image reconstructed from 900 projections with FBP, (B) Image reconstructed from 200 projections with FBP, (C) Image reconstructed from 200 projections with the proposed method, FCSA-LEM.

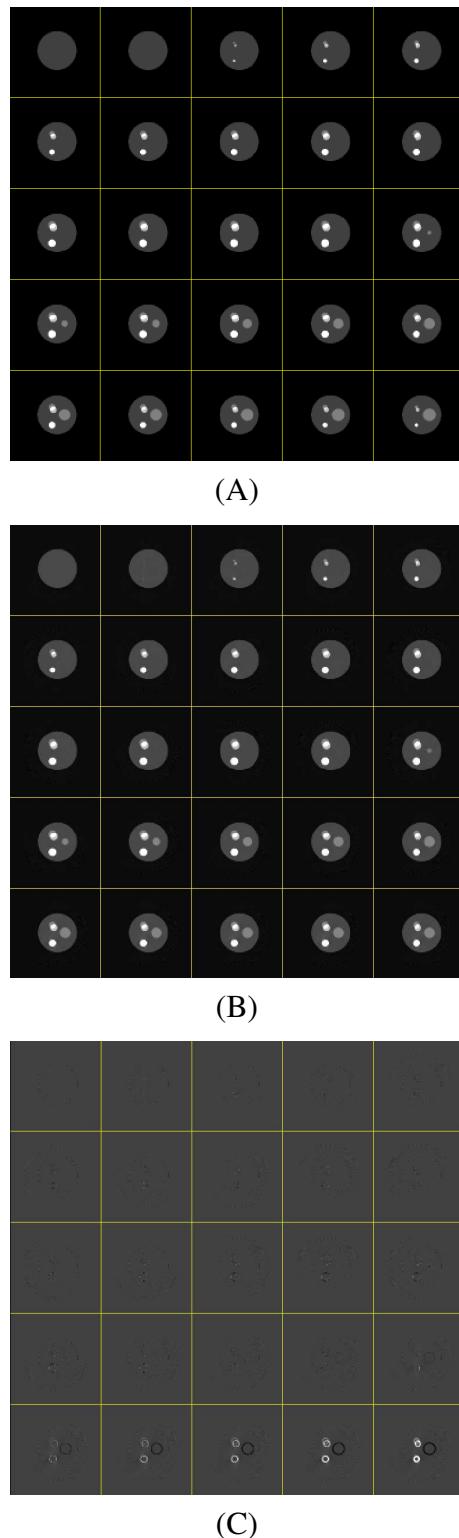


Figure 5.11: Helical scan tested on a simple simulated phantom. (A) The original phantom. (B) Image reconstructed with the proposed method. (C) Difference between the true image and the reconstructed image.

Chapter 6

An Adaptive Method for Updating Regularization Parameters

Improving the image quality is a critical objective in low dose Computed Tomography (CT) and is the focus of CT image denoising. State of the art CT denoising algorithms are mainly based on iterative minimization of an objective function, in which the performance is controlled by a set of regularization parameters. To achieve the best result, these should be chosen correctly. However, these values are typically chosen in an ad-hoc manner, which can cause the algorithms to converge slowly and become trapped in a local minimum. To overcome these issues, we propose an iterative updating algorithm based on a new Noise Confidence Region Evaluation (NCRE) method, which evaluates the denoising results at the end of each iteration and updates the parameters to achieve a better match to the noise statistics. Moreover, it stops the iterations when the statistical characteristics of the residual resemble Gaussian noise.

Using NCRE, we propose a new iterative CT image denoising approach that combines NCRE and the fundamentals of a non-iterative Block Matching 3D (BM3D). The proposed method, called NCRE-BM3D, improves the the performance of the BM3D algorithm, which is a state of the art denoising method, in terms of the mean squared error (MSE) and feature preservation, measured by Structural Similarity Index (SSIM) [169].

However, this is not the only application of NCRE. In this chapter, we start with a general noiseless signal recovery formulated with Lagrangian multipliers. This model can be used in a number of different applications. In addition to NCRE-BM3D, in this chapter the role of NCRE is explained in adaptive regularization of wavelet soft thresholding, Total Variation denoising, and Iterative Soft Thresholding compressed sensing recovery. Simulation results showed that in all these cases NCRE improves the performance of the signal restoration and enables the algorithms to converge in fewer iterations.

6.1 Importance of the Regularization Parameter Selection

Estimating a noiseless signal from noisy measurements appears in diverse applications ranging from communications to biomedical signal analysis. The most common model used for noise is additive white Gaussian noise (AWGN):

$$\mathbf{y} = \mathbf{Ax} + \mathbf{n}, \quad (6.1)$$

where $\mathbf{y} \in \mathbb{R}^m$ is the noisy measurements, $\mathbf{A} \in \mathbb{R}^{m \times N}$ is the measurement matrix, $\mathbf{x} \in \mathbb{R}^N$ is the noiseless signal, and $\mathbf{n} \sim N(0, \sigma_n^2)$ is a zero mean Gaussian noise with variance of σ_n^2 . This model appears in many problems, such as regression, denoising, and compressed sensing (CS) [18, 90]. To estimate the noiseless signal, unconstrained Lagrangian multiplier optimization problem is usually used [18, 90, 99, 170, 171, 95],

$$\hat{\mathbf{x}} = J(\mathbf{y}, \mathbf{A}, \lambda) = \operatorname{argmin}_x \frac{1}{2} \|\mathbf{Ax} - \mathbf{y}\|_2^2 + \lambda h(x), \quad (6.2)$$

in which λ is a regularization parameter that controls the trade-off between data-fidelity and regularization term $h(x)$. This optimization problem is usually solved iteratively and there is a strong dependence of the quality of the result on the regularization parameter [170, 172, 173]. It is a challenging task to find the regularization parameter λ that provides the best balance between signal smoothing and feature preservation. Specifically, if λ is not appropriately adjusted, the optimization is trapped in a local minimum, i.e., if λ is too small, noise is only partially removed and if it is too large, the image will be over-smoothed [30, 172]. Some methods have been proposed to update the regularization parameters iteratively, such as use of the discrepancy principle [174] that seeks for the noise-only residual by comparing the noise and residual variances, Generalized cross-validation [173] that minimizes the prediction error, and L-curve, which is based on the plot of the norm of the regularized solution versus the norm of the corresponding residual [175]. These methods fail in certain situations, are problem specific, and generally increase the computational complexity of the recovery algorithms.

One straight forward approach used in many algorithms is the use of a heuristic value λ combined with a criterion to stop the algorithm before the estimated signal is over-smoothed [176]. Different stopping criteria have been proposed for iterative denoising problems. For instance, switching median filter algorithm [177] stops the algorithm when the number of changed pixels in the denoising iterations is minimum. In [178], statistical properties of high frequency wavelet subbands are used to stop the total variation iterations. A major problem in these stopping conditions arises from considering just two situations: that the signal is either noisy or well denoised. Consequently, they are not able to differentiate an over-smoothed data from

a well denoised data. To avoid over-smoothness the updating steps are usually chosen to be small, which decreases the convergence speed.

In this chapter we propose an algorithm that adaptively updates the regularization parameter of noiseless signal restoration problems, presented in an unconstrained form with one Lagrangian multiplier, similar to (6.2). Our proposed method, NCRE, validates the statistical properties of the denoising residual at the end of each iteration and provides the information that the signal is well denoised, partially denoised, or over-smoothed. This information is used to update the parameters in the next iteration or to stop the algorithm. The role of NCRE in denoising and compressed sensing [90, 179] problems is explained here with few examples: 1) wavelet soft thresholding, 2) Total Variation (TV) denoising [99], 3) parameterless CS recovery using a modified Iterative Soft Thresholding Algorithm (ISTA) [95], and 4) Computed Tomography (CT) image denoising using a new iterative BM3D method.

6.2 NCRE Stopping Criterion and Regularization Parameter Updating Method

If $\hat{\mathbf{x}}_i$ denotes the signal recovered at i^{th} iteration by solving (6.2), the denoising residual at the end of i^{th} iteration can be expressed by $\Delta\mathbf{y}_i = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_i$ which, ideally, should be white Gaussian noise (6.1). Here, we provide a quantitative measure that verifies the similarity between the structure of $\Delta\mathbf{y}_i$ and that of \mathbf{n} . The method stops at the end of the iteration for which the similarity is satisfied. Algorithm 7 shows the proposed iterative λ updating scheme, which is described in the following sections.

6.2.1 Noise Confidence Region Evaluation (NCRE)

In [180] it was shown that the following function of zero mean white Gaussian noise, \mathbf{n} , with length m , and any given scalar value of z ,

$$\mathcal{G}(z, \mathbf{n}) = \frac{1}{m} \sum_{j=1}^m \mathcal{G}(z, \mathbf{n}_j), \quad \mathcal{G}(z, \mathbf{n}_j) = \begin{cases} 1 & \text{if } |\mathbf{n}_j| \leq z \\ 0 & \text{if } |\mathbf{n}_j| > z, \end{cases} \quad (6.3)$$

is equivalent to sorting the absolute value of the noise elements \mathbf{n}_i . The expected value of this function is $E(\mathcal{G}(z, \mathbf{n})) = F(z)$ and its variance is $\text{var}(\mathcal{G}(z, \mathbf{n})) = \frac{1}{m}F(z)(1 - F(z))$, where $F(z) = 2\phi(\frac{z}{\sigma_n}) - 1$ and $\phi(x)$ is the cumulative distribution function (CDF) of Gaussian distri-

Algorithm 7 Proposed Iterative Regularization Parameter Updating.

```

 $\hat{\mathbf{x}}_0 = J(\mathbf{y}, \mathbf{A}, \lambda),$ 
 $\Delta\mathbf{y}_0 = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_0, \mathcal{G}_0 = \text{sort}(|\Delta\mathbf{y}_0|), i = 1,$ 
while  $\mathcal{G}_i \not\subseteq$  Region II and  $i < \text{maxiter}$  do,
    if  $\mathcal{G}_i \subset$  Region I then
        increase  $\lambda$  towards Region II,
         $\hat{\mathbf{x}}_i = J(\mathbf{y}_i, \mathbf{A}, \lambda),$ 
         $\mathbf{y}_i = (1 - \alpha_1)\mathbf{A}^T\mathbf{y} + \beta_1\hat{\mathbf{x}}_i + (\alpha_1 - \beta_1)\hat{\mathbf{x}}_{i-1},$ 
    end if
    if  $\mathcal{G}_i \subset$  Region III then
        decrease  $\lambda$  towards Region II,
         $\hat{\mathbf{x}}_i = J(\mathbf{y}_i, \mathbf{A}, \lambda),$ 
         $\mathbf{y}_i = (1 - \alpha_2)\mathbf{A}^T\mathbf{y} + \beta_2\hat{\mathbf{x}}_i + (\alpha_2 - \beta_2)\hat{\mathbf{x}}_{i-1},$ 
    end if
     $\Delta\mathbf{y}_i = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_i, \mathcal{G}_i = \text{sort}(|\Delta\mathbf{y}_i|), i \leftarrow i + 1,$ 
end while

```

bution. Therefore, it is bounded by the following lower (L_n) and upper (U_n) values,

$$\begin{aligned} L_n(z) &= F(z) - \zeta \sqrt{\frac{1}{m}F(z)(1-F(z))} \\ U_n(z) &= F(z) + \zeta \sqrt{\frac{1}{m}F(z)(1-F(z))} \end{aligned} \quad (6.4)$$

with probability of $\Pr\{|\mathcal{G}(z, \mathbf{n}) - F(z)| \leq \zeta \sqrt{\frac{1}{m}F(z)(1-F(z))}\} \approx 2\phi(\zeta) - 1^1$. If the sorted absolute values of a signal lies between these two boundaries for a large enough ζ , that signal follows a white Gaussian distribution with confidence probability close to one.

As shown in Fig. 6.1, these boundaries divide the $(z, \mathcal{G}(z, \mathbf{n}))$ space into three regions denoted by Region I, Region II, and Region III. At the end of each iteration $\mathcal{G}(z, \Delta\mathbf{y}_i)$ the sorted absolute value of the residual, is calculated. If this sequence falls into Region II², it means that the residual has a Gaussian like structure and denoising stops. On the other hand, if the denoising at the i^{th} iteration has removed not only the noise but also parts of the noiseless data itself, $\Delta\mathbf{y}_i$ will have some of the image information making its samples larger than Gaussian noise. Therefore, for a specific value of z , $\mathcal{G}(z, \Delta\mathbf{y}_i)$ (average number of $\Delta\mathbf{y}_i$ s with absolute values smaller than z) is smaller than $\mathcal{G}(z, n)$ and falls in Region III. This will enforce continuation of the denoising to the $(i+1)^{th}$ step with changing of the regularization parameters λ such that the denoising algorithm extracts less noise in the next iteration, i.e., *decreasing* $\lambda = \lambda/s$, $s > 1$.

¹When needed the Median-Absolute Deviation (MAD) is used to estimate noise variance [181].

²In our approach (Algorithm 7) being a subset of a region is evaluated by having a high fraction of $\mathcal{G}(z, \Delta\mathbf{y}_i)$ in that region. For example, in our simulations this fraction is 90%.

If $\mathcal{G}(z, \Delta\mathbf{y}_i)$ falls in Region I, it indicates that the noise is partially removed. In this case the algorithm continues to the $(i+1)^{th}$ step and changes the regularization parameters λ such that more noise is extracted by the denoising algorithm, i.e. *increasing* $\lambda = s \times \lambda$, $s > 1$.

In summary, at each iteration if $\mathcal{G}(z, \Delta\mathbf{y}_i)$ is in regions I or III, the regularization parameter is updated such that it moves toward region II. Value of s can be tuned for each application, as a fixed or an adaptively changing variable based on the euclidean distance between $\mathcal{G}(z, \Delta\mathbf{y}_i)$ and $F(z)$, $d_i = \|\mathcal{G}(z, \Delta\mathbf{y}_i) - F(z)\|^2$.

6.2.2 Robust Convergence Step

The proposed algorithm uses a memory strategy for recovery of possible lost edges and fine details:

$$\mathbf{y}_i = (1 - \alpha_j)\mathbf{A}^T \mathbf{y} + \beta_j \hat{\mathbf{x}}_i + (\alpha_j - \beta_j) \hat{\mathbf{x}}_{i-1}. \quad (6.5)$$

in which $j \in [1, 2]$, α_j and β_j are positive scalars, chosen based on the conditions given in [27]. This stage of the algorithm is inspired by the second order iterative methods [27, 182] that improve the convergence rate of the iterative methods.

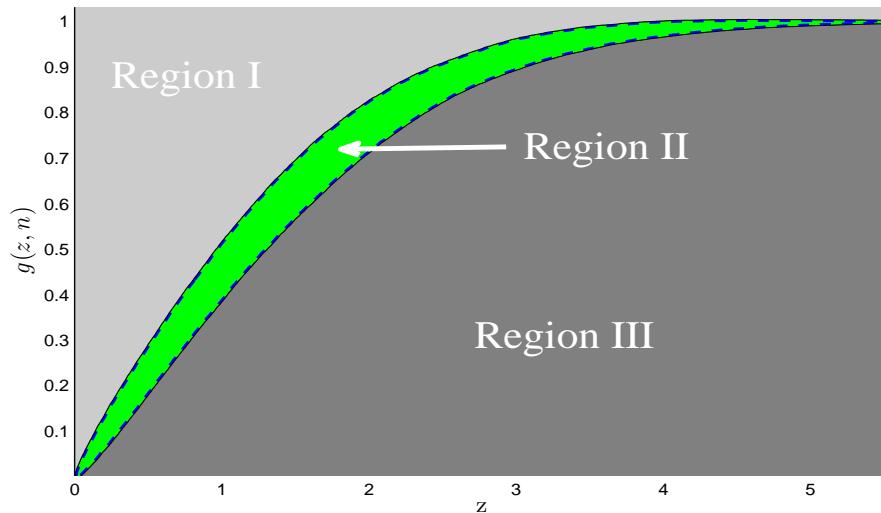


Figure 6.1: Three possible regions for the residual at the end of each iteration. If it lies in Region II (Noise Confidence Region) denoising is stopped.

6.3 NCRE in Denoising and Compressed Sensing

This section explains the role of the NCRE for updating the regularization parameters in denoising and CS recovery. Simulation results are provided in section 6.4. It should be noted that $\mathbf{A} = I_{N \times N}$ is the identity matrix in denoising problems and, in CS problem \mathbf{A} is a measurement matrix with $m \ll N$, which makes (6.2) an ill-conditioned problem.

6.3.1 Soft Thresholding (NCRE-ST)

In the wavelet soft thresholding method, (6.2) is minimized with $h(x) = \|Wx\|_1$, where W is the wavelet transform matrix and $\|x\|_p = (\sum_i |x_i|^p)^{1/p}$. This can be solved by $\hat{x} = S_\lambda(x) = \text{sign}(x) \max(|x| - \lambda, 0)$. Several different algorithms for choosing λ have been proposed [183, 181, 184, 155], but these are suboptimal. NCRE can be used to improve these algorithms by iteratively updating the threshold values. To initialize λ , BayesShrink [181] can be used, for example, with $\lambda = \frac{\sigma_n^2}{\sigma_{\bar{x}}}$, where $\sigma_{\bar{x}} = \sqrt{\sigma_y^2 - \sigma_n^2}$ and σ_y is the standard deviation of the noisy signal.

6.3.2 TV Denoising (NCRE-TV)

NCRE-TV adaptively updates the regularization parameter in TV image denoising [99]. It uses $h(x) = \mathbf{TV}(x) = \sqrt{|\nabla x^h|^2 + |\nabla x^v|^2}$ to model the piece-wise linearity of the images [80], where ∇x^h and ∇x^v are the first order derivatives in horizontal and vertical directions, correspondingly. The split Bregman algorithm (algorithm 4) [99] is used to solve (6.2) and λ is initialized by $\lambda = \|\mathbf{TV}(\mathbf{y})\|_\infty$.

6.3.3 Iterative Soft Thresholding Algorithm (NCRE-ISTA)

Selecting an optimal λ is a challenging task in CS formulation, making parameter-free CS algorithms very desirable [179]. Here we use NCRE to propose a parameter-free ISTA, called NCRE-ISTA. ISTA (algorithm 1) is a simple and effective CS recovery algorithm [95] that uses $\hat{x}^{i+1} = S_\lambda(\hat{x}^i + \mathbf{A}^T(y - \mathbf{A}\hat{x}^i))$ to update the signal in the $(i+1)^{th}$ iteration, a process that is known to converge slowly.

In addition, ISTA is used to solve ℓ_p -minimization problem using a generalized shrinkage operator [185] given by:

$$S_{\lambda,p}(x) = \text{sign}(x) \max\{|x| - \lambda^{2-p}|x|^{p-1}, 0\}. \quad (6.6)$$

Based on the discussion in subsection 3.2.1 and figure 3.3 it can be expected that better recovery performance will be obtained with ℓ_p , $0 \leq p < 1$ compared to ℓ_1 . However, choosing an optimum λ for this optimization problem is challenging. NCRE is used to address this problem.

Our observations show that, when a fixed λ is used, the distance between $\mathcal{G}(z, \Delta\mathbf{y}_i)$ and the center of region II, $F(z)$, gets smaller at each iteration until a point is reached at which the updating steps get very small, and therefore the convergence rate becomes very slow, often stopping before reaching the true answer.

This problem can be solved with NCRE as follows: if the residual is placed in Region III and the difference of the distances between $F(z)$ and $\mathcal{G}(z, \Delta\mathbf{y}_i)$ in two successive updates is small (e.g., 1/100 of the distance in last update), the value of λ is decreased, if it is in Region I, λ is increased, and if it is Region II the algorithm is stopped. The starting λ in NCRE-ISTA is chosen to be optionally large, $\lambda = 0.8\|\mathbf{A}^T(\mathbf{y})\|_\infty$. The proposed method resolves the slow convergence problem of ISTA and the challenging task of finding an appropriate λ .

6.3.4 CT Image Denoising (NCRE-BM3D)

Modeling the noise characteristics of CT images is complex³ [43]. In general, it can be interpreted as a correlated non-stationary Gaussian noise that its variance is the back projection of the projection noise variances. Our studies have shown that noise in local regions with similar gray scale values (CT#) can be modeled as white Gaussian noise with identical variances (Appendix A) [186]. This property is used to denoise the CT images with a new iterative Block Matching 3D (BM3D) approach [187], called NCRE-BM3D.

BM3D [187] is a non-iterative denoising method that still outperforms many newer algorithms [188]. It is composed of two major filtering steps as shown in figure 6.2. In both stages collaborative filtering is utilized, which itself has four stages: 1) grouping similar patches with a reference patch, as shown in figure 6.3, 2) calculation of the 3D wavelet coefficients of each stack of patches, 3) denoising the wavelet coefficients (thresholding in step 1 or Wiener filtering in step 2, and 4) recovering the denoised image by calculating the inverse 3D wavelet transformation. BM3D aims to denoise the patches by Wiener filtering, which is done in step 2. This requires reliable estimate of the noiseless image to find best similar patches in the step 2 and to reliably estimate the Wiener coefficients. Due to the similarity of the patches, the noise variance of CT images in each block is almost identical. The main purpose of the step 1 is to find the best initial estimate of the noiseless image. The input of this step, which is a hard thresholding block, is the 3D noisy wavelet coefficients of the similar patches located by block

³More details about the CT image noise is provided in Appendix A.

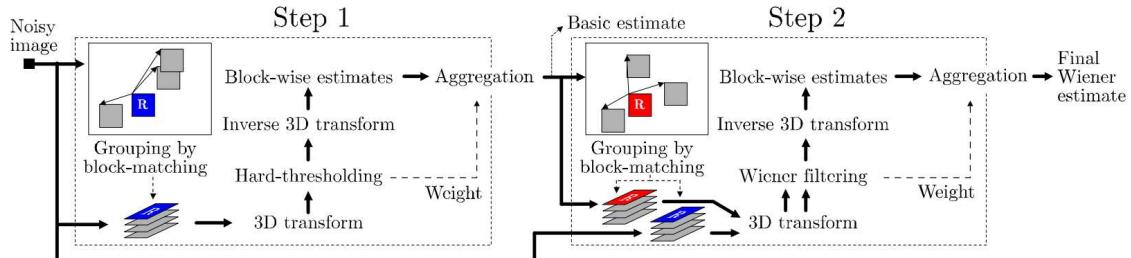


Figure 6.2: Flowchart of BM3D denoising algorithm. The operations surrounded by dashed lines are repeated for each processed block, marked with R . Reproduced, with permission, from [187].

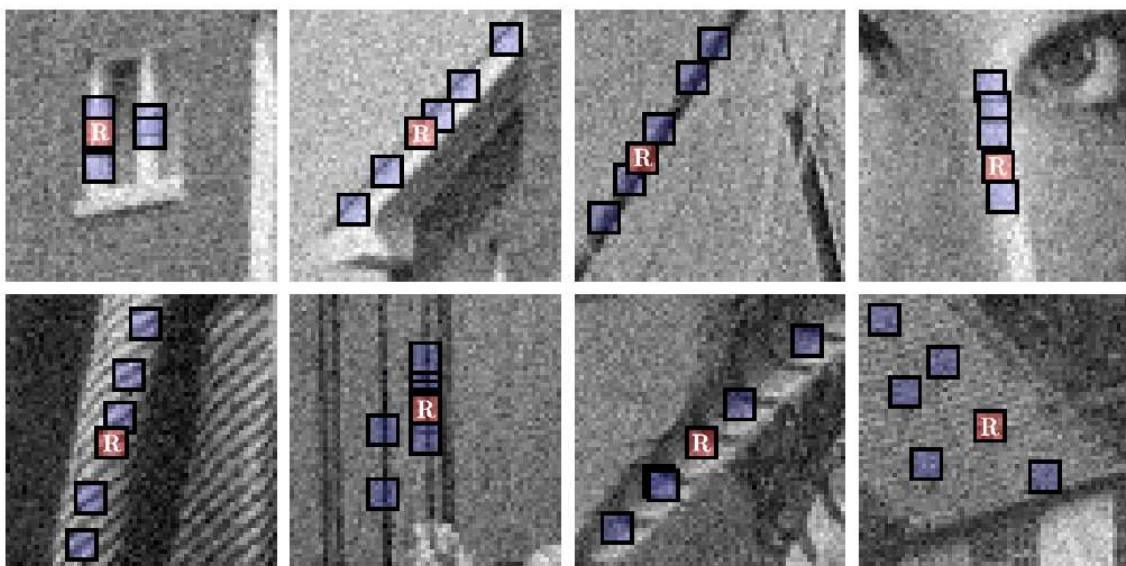


Figure 6.3: Grouping the blocks of noisy images corrupted by a zero mean white Gaussian noise with standard deviation 15. Reproduced, with permission, from [187].

matching applied to the available noisy image. A simple hard threshold with value of $2.7\sigma_n$ is used in the first step, as found heuristically [187]. The resulting denoised coefficients are then transformed back to spatial domain to be used as initial estimates of the noiseless data used for calculating the Wiener filter coefficients.

In the proposed method, a noisy image is treated as an arranged set of small similar blocks and denoising is performed on each block. Noise in the blocks is assumed to be a locally additive Gaussian distribution, so that noise level of different blocks may differ [189]. We use the fundamentals of BM3D to propose an iterative scheme: the output of the Wiener filter is a better estimate of the original image than the input of Wiener filter from the first step. Therefore, this output can be fed into the first step to provide better Wiener coefficients in the second iteration. BM3D is used iteratively in Algorithm 7 to denoise the CT images, where NCRE adjusts the

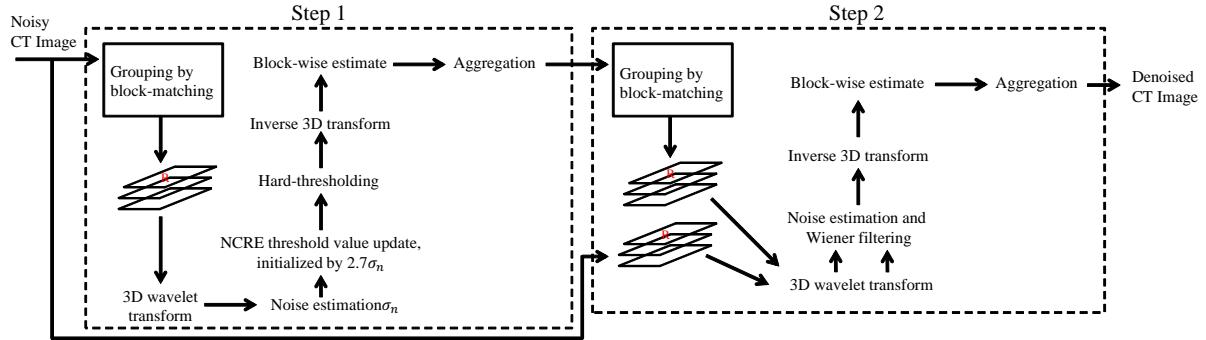


Figure 6.4: Schematic of the proposed CT image denoising scheme (NCRE-BM3D).

threshold values applied on 3D wavelet coefficients of the first step. λ is initialized by $2.7\sigma_n$ with noise variances estimated independently for each 3D stack. In each iteration, if the sorted absolute value of the residual falls into Region I, the threshold value will be increased and, if it falls into Region III, the threshold values is decreased; so that in the next iteration the residual moves towards Region II. A schematic of NCRE-BM3D is depicted in figure 6.4.

6.4 Results

Testing of NCRE-ST was performed using four standard 1D signals (available in MATLAB by ‘wnoise(i,11)’, $i=1,\dots,4$) and four images (Cameraman, Hills, Lena, and Peppers). Table 6.1 compares the performance of the NCRE-ST and BayesShrink for the 1D signals based on signal to noise ratio (SNR) in dB. ‘Haar’ wavelets with 5 levels were used with $\alpha_i = 1$, $\beta_i = s_i$, $s_i = 1 - 2 \exp(d_i^2/10)$, $i \in [1, 2]$. As can be seen, the denoising efficiency has improved significantly. In addition, in average of 50 runs, NCRE-ST improved the SNR of the denoised images by 2.1 dB, compared to BayesShrink. Figure 6.5 C-D compares an image denoised with BayesShrink and NCRE-ST. As can be seen, the image is denoised better and the details are well-preserved by NCRE-ST.

Figure 6.5 E-F compares an image denoised with TV denoising and the NCRE-TV, using $\alpha_i = \beta_i = 1, i \in [1, 2]$, $s = 1.2$. The proposed method not only eliminates the need of manual λ optimization, but also substantially improves the edge preservation.

To test the NCRE-ISTA, a 4096 element 1D signal was used having 3936 zeros and 160 non-zero elements, randomly chosen to be either 1 or -1. A random 1024×4096 measurement matrix, \mathbf{A} , with normalized columns was used to under-sample the signal, which were corrupted by Gaussian noise with $\sigma_n = 10^{-3}$ and $\alpha_i = \beta_i = 1, i \in [1, 2]$, $s = 1.05$. Figure 6.6 compares the signal recovered by ISTA, with a manually optimized $\lambda = 0.05\|\mathbf{A}^T(\mathbf{y})\|_\infty$, and NCRE-ISTA. As can be seen, the recovery error is improved by four order of magnitudes. Moreover, the

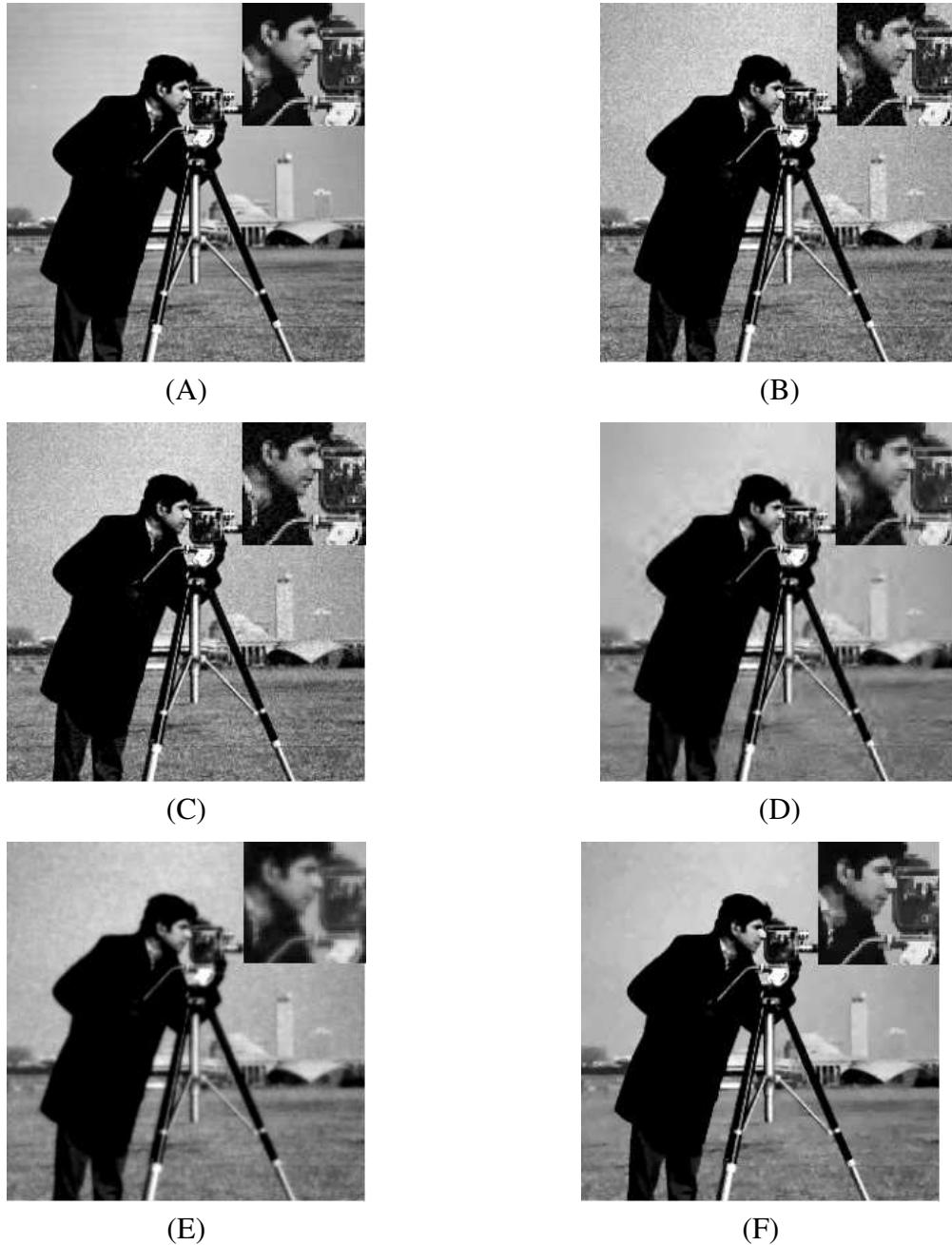


Figure 6.5: Comparison of regular TV denoising and NCRE-TV: (A) Original Image, (B) Noisy image ($\text{SNR} = 19.94 \text{ dB}$), (C) Denoised with BayesShrink ($\text{SNR} = 21.93 \text{ dB}$), (D) Denoised with NCRE-ST ($\text{SNR} = 23.09 \text{ dB}$), (E) Denoised by regular TV denoising ($\text{SNR} = 22.49 \text{ dB}$), and (F) Denoised by NCRE-TV ($\text{SNR} = 25.56 \text{ dB}$).

Table 6.1: Comparison of BayesShrink and NCRE-ST with 1D test signals. SNR values are in dBs and are the average of 50 runs.

σ_n	0.1	0.2	0.3	0.4	0.5	0.6
Input SNR						
Blocks (i=1)	29.46	23.50	19.95	17.50	15.58	14.03
Bumps (i=2)	25.13	19.14	15.67	13.24	11.45	9.98
Heavy Sine (i=3)	29.84	23.78	20.24	17.83	15.95	14.38
Doppler (i=4)	9.82	4.98	2.93	1.88	1.28	0.94
BayesShrink						
Blocks	29.79	24.17	20.95	18.82	17.22	15.98
Bumps	25.49	19.89	16.82	14.76	13.35	12.22
Heavy Sine	30.03	24.20	20.88	18.70	17.03	15.68
Doppler	11.74	8.42	7.47	7.16	6.92	6.84
NCRE-ST						
Blocks	37.09	30.83	26.56	25.37	23.54	21.76
Bumps	27.90	22.83	20.24	18.64	17.24	16.10
Heavy Sine	33.11	28.95	26.90	25.24	24.44	23.58
Doppler	15.24	11.61	10.05	8.81	7.57	6.97

number of iterations in the proposed method is almost 1/10 of the regular method (1000 vs. 120). To compare the performance of ℓ_1 and ℓ_p , $p = 0.2$, the number of non-zero elements of the signal was changed between 100 and 300, while the number of measurements was kept fixed and equal to 1024. Figure 6.7 shows that ℓ_p , $p = 0.2$, recovers the signal with 260 non-zeros, but ℓ_1 fails after 210 non-zero elements.

A CATPHAN phantom (Phantom Laboratory, Greenwich, NY, USA) was used to test NCRE-BM3D. It was scanned using a low dose (50 mAs, 120kVp) and a high dose (300mAs, 120kVp) protocol. Image reconstructions were performed with a Toshiba CT scanner using a lung kernel (FC52) and proprietary iterative reconstruction, Adaptive Iterative Dose Reduction 3D (AIDR3D). Our proposed denoising method was applied to the images reconstructed with a high spatial resolution filter algorithm, FC52, and these are compared to the images reconstructed with AIDR3D. Figure 6.8 shows that the detectability of low contrast objects is improved with our method and outperforms the commercially available iterative algorithm. Figure 6.9 shows that both AIDR3D and the proposed method keep the line-pair resolution unchanged. In addition, clinical low dose (50mAs, 120kVp) lung CT's of a patient were reconstructed using FC52 and were processed by BM3D and BM3D-NCRE with $\alpha_i = 0.9$, $\beta_i = 0.1$, $i = [1, 2]$, $s = 1.2$. As can be seen, NCRE-BM3D removes the noise more effectively compared to BM3D and preserves the fine features which from a diagnostic perspective can be very important.

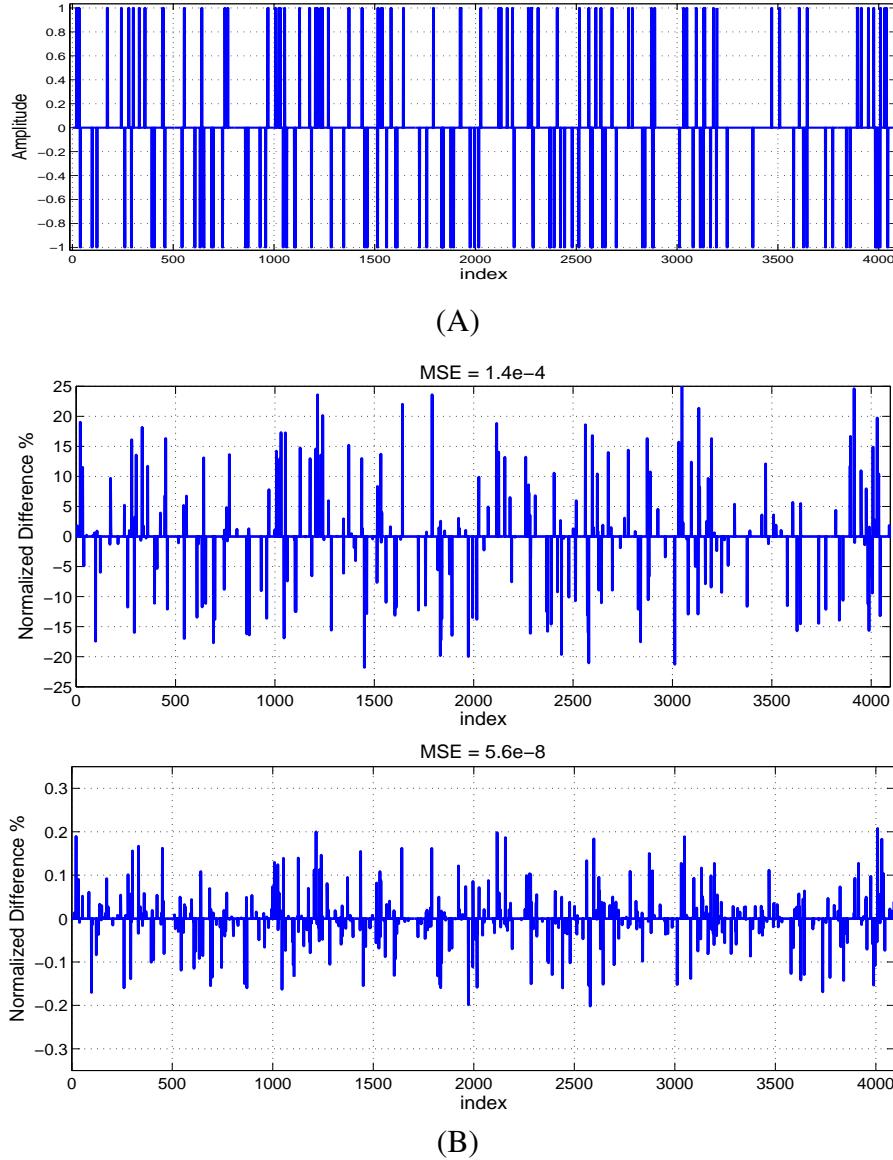


Figure 6.6: (A) Original Signal, and (B) Recovery error of ISTA (Top) and NCRE-ISTA (Bottom).

6.5 Summary

An iterative regularization parameter updating scheme was proposed for signal denoising problems formulated by Lagrangian multipliers. The proposed NCRE method compared the statistical characteristics of the Gaussian noise with denoising residual, quantitatively, to determine if the denoising was effectively, weakly or strongly executed. Based on this information the denoising parameters were adjusted for the next iteration. NCRE was used to improve the following denoising and compressed sensing algorithms: 1) wavelet soft thresholding, 2) TV

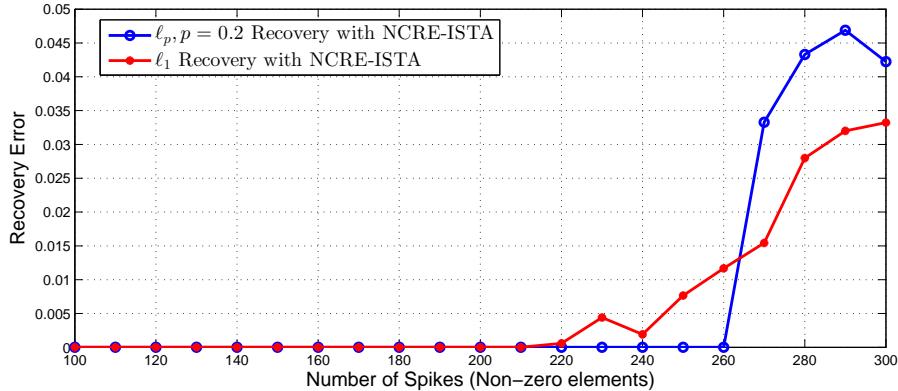


Figure 6.7: Compares the number of measurements needed for exact recovery by ℓ_p and ℓ_1 minimizations, using NCRE-ISTA.

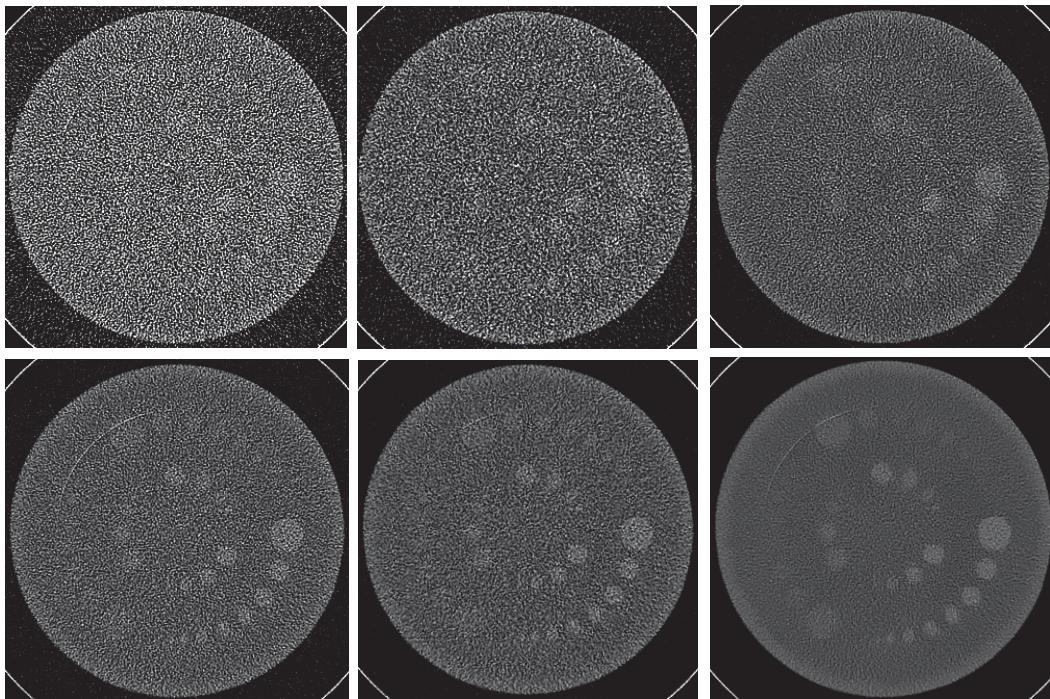


Figure 6.8: Low contrast slice of CATPHAN denoised by the proposed method and Adaptive Iterative Dose Reduction (AIDR3D). Top: scanned with 50mAs/120kVp and Bottom: scanned with 300mAs/120kVp. Left: image reconstructed with FC52, Middle: reconstructed with AIDR3D, and Right: reconstructed with FC52 and denoised by BM3D-NCRE. In all images Window-Width/Window-Level is: WW=100/WL=70.

denoising, 3) parameter-free iterative soft thresholding CS recovery algorithm, and 4) a new iterative BM3D based CT image denoising. It was shown that NCRE not only improved the SNR performance of the recovery algorithms, but also significantly reduced the number of iterations needed to achieve an acceptable range of error in TV minimization and ISTA.

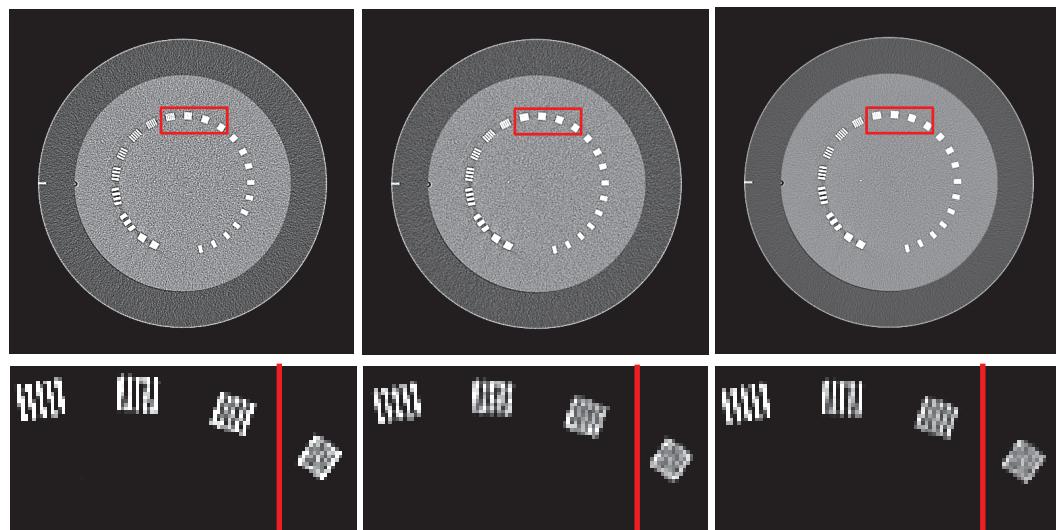


Figure 6.9: Top: Line pair slice of the CATPHAN scanned with 50mAs and 120kVp (WW=400/WL=60). Bottom: Red rectangular ROI of the images, the red lines show the cut off line pair resolution (WW=400/WL=500). Left: Phantom reconstructed with FC52 (noise standard deviation=64HU), Middle: Phantom reconstructed with FC52+AIDR3D (noise standard deviation=31HU), and Right: Phantom reconstructed with FC52 and denoised with the proposed method (noise standard deviation=22HU).

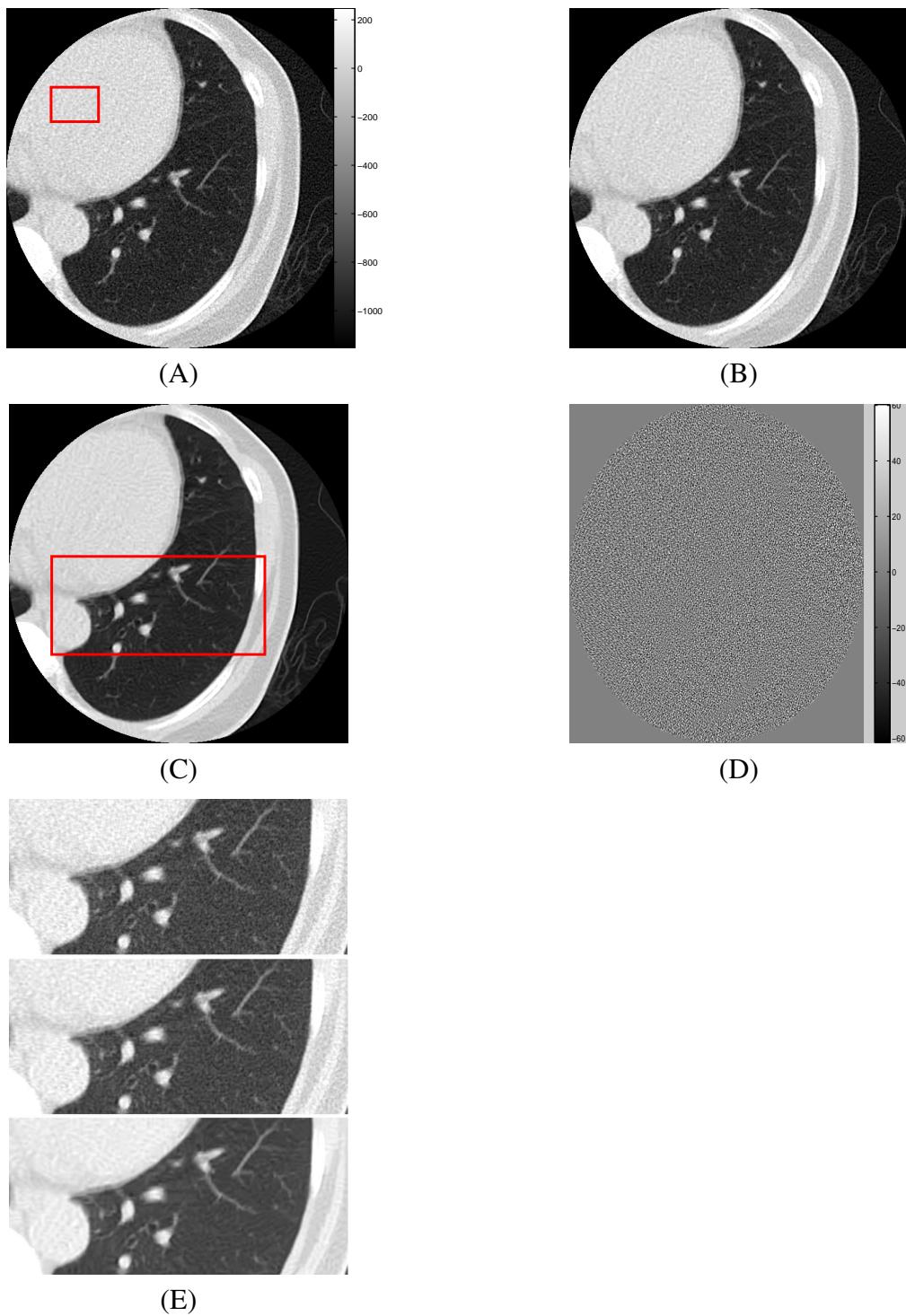


Figure 6.10: (A) CT lung image, noise std in red rectangle is 60HU, (B) Denoised with BM3D, noise 38HU, and (C) Denoised with NCRE-BM3D, noise 15HU. WW/WL = 1400/-450HU, (D) The difference between the noisy and denoised image, and (E) Top: Original, Middle: BM3D, Bottom: NCRE-BM3D, WW/WL = 1500/-600HU.

Chapter 7

Compressed Sensing in Local CT Reconstruction

One potential method for reducing the radiation dose is to limit the exposure area to those regions that need to be examined, as illustrated in figure 7.1. Reconstruction methods based on the FBP technique are non-local because of the use of Fourier transform as its theoretical foundation. This means that the recovery of any pixel in a CT image requires information from all projections thereby exposing the patient to a relatively large amount of radiation even when only a small region needs to be examined.

Several methods for local CT reconstruction have been previously studied. These algorithms are usually based on the localized nature of wavelet transforms [31, 32, 33], iterative reconstruction methods [34], or generalized FBP and truncated Hilbert transform [36, 35]. Here, we focus on wavelet inverse Radon (WiR) based methods to recover the wavelet subbands of a region of interest (ROI). Another important application of wavelet transform in signal processing is its use as a sparsifying transform in Compressed Sensing. However, since CS-based CT reconstruction methods use the discrete Radon matrix, they are not suitable for local CT reconstruction. In this chapter it is shown that CS-based CT reconstruction and wavelet inverse Radon can be combined to achieve an ultra low dose local CT imaging technique. First, WiR is improved by decreasing its computational load. WiR uses back projection (BP) to recover the wavelet subbands. In the proposed method, denoted by accelerated wavelet inverse Radon (AWiR), BP is replaced by Direct Fourier Reconstruction (DFR) [48] on pseudo-polar grids; this significantly reduces the computational complexity. By combining CS with AWiR, we propose a low dose CS-based local CT reconstruction algorithm, denoted by Accelerated Wavelet inverse Radon Subband adaptive Iterative Shrinkage/Thresholding Algorithm (AWiR-SISTA), that takes advantage of both local CT imaging and CS-based CT reconstruction to reconstruct ultra-low dose local CT images. This method, not only reduces the effective radiation, but also

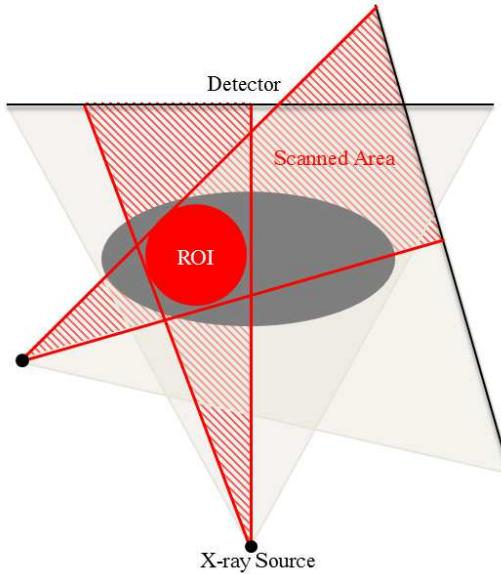


Figure 7.1: Region of Interest scanning method.

decreases the computational complexity of CS recovery and, in addition, by using an innovative combination of Two step Iterative Soft Thresholding (TwIST) [27] and subband dependent CS recovery [97, 190], the convergence rate of iterative soft thresholding algorithm (ISTA) is improved.

7.1 Wavelet Transform and CT image Reconstruction

The problem of low dose local CT imaging is studied in this chapter by proposing a wavelet subband dependent CS-based CT reconstruction algorithm. To achieve high quality low dose CT images, it should be noted that CS-based CT reconstructions recover the images using fewer projections; and that wavelet inverse Radon algorithms recover wavelet subbands of CT images from locally scanned projections. Moreover, it has been shown that subband CS algorithms accelerate the convergence of the CS recovery methods, such as ISTA (described in Algorithm 1). Here, we propose an innovative combination of a newly developed accelerated wavelet inverse Radon transform and non-convex CS formulation to recover the wavelet subbands of CT images from a reduced number of locally scanned X-ray projections. Fast pseudo-polar Fourier transform is used to accelerate wavelet inverse Radon, and as a result, to decrease the computational complexity of CS recovery. Therefore, the proposed method, denoted by AWiR-SISTA, reduces the radiation dose by simultaneously decreasing the X-ray exposure area and the number of projections, decreases the CS computational complexity, and accelerates the CS recovery convergence rate.

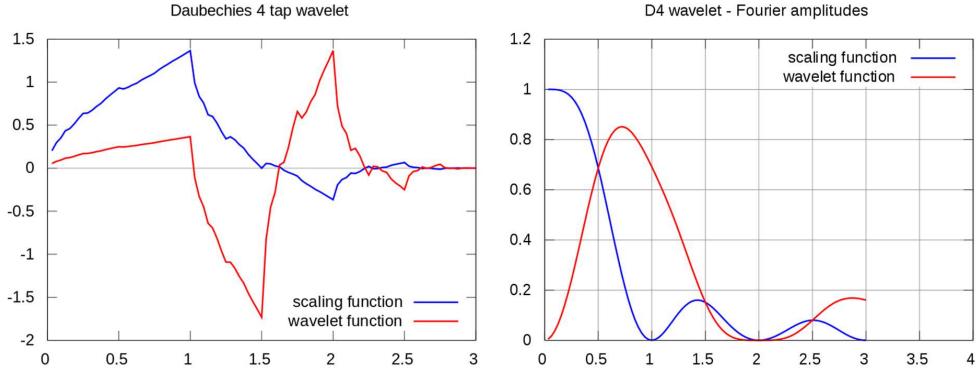


Figure 7.2: Scaling and wavelet functions of Daubechies 4.

7.1.1 Wavelet based Local CT Reconstruction

If ψ is a wavelet with n vanishing moments that generates an orthonormal basis of \mathbb{R}^2 , then it has a support of size larger than or equal to $2n - 1$. It is well known that a Daubechies' wavelet has a minimum-size support equal to $[-n + 1, n]$ and that the support of the corresponding scaling function ϕ is $[0, 2n + 1]$. An example is shown in figure 7.2. Since the support of the Hilbert transform of a compactly supported scaling function is the same as the scaling function itself, the wavelet subbands of CT images can be reconstructed from locally scanned projections [31, 32, 33].

Using this property of wavelet transform, an important class of local CT reconstruction methods [31, 32, 33] makes use of wavelet inverse Radon transform (WiR) to overcome the non-locality of FBP. These algorithms recover the wavelet coefficients of the desired ROI of the CT image from locally measured X-ray projections. A typical WiR algorithm is depicted in figure 7.3-A, where the input X-ray projections are convolved with wavelet ramp filters. This convolution is performed in the Fourier domain by multiplying 1D Fourier coefficients of the X-ray projection with the Fourier transform of the wavelet ramp filters. For example, with a 2 level wavelet transformation, the Fourier transform of the approximate and detail wavelet ramp filters are given by [31]:

$$\begin{aligned} H_{\phi}^A &= |\omega| \mathcal{F}_{1D} \tilde{\phi}_{2j}(\omega \cos \varphi) \mathcal{F}_{1D} \tilde{\phi}_{2j}(\omega \sin \varphi) \mathcal{W}(\omega) \\ H_{\phi}^{D^1} &= |\omega| \mathcal{F}_{1D} \tilde{\phi}_{2j}(\omega \cos \varphi) \mathcal{F}_{1D} \tilde{\psi}_{2j}(\omega \sin \varphi) \mathcal{W}(\omega) \\ H_{\phi}^{D^2} &= |\omega| \mathcal{F}_{1D} \tilde{\psi}_{2j}(\omega \cos \varphi) \mathcal{F}_{1D} \tilde{\phi}_{2j}(\omega \sin \varphi) \mathcal{W}(\omega) \\ H_{\phi}^{D^3} &= |\omega| \mathcal{F}_{1D} \tilde{\psi}_{2j}(\omega \cos \varphi) \mathcal{F}_{1D} \tilde{\psi}_{2j}(\omega \sin \varphi) \mathcal{W}(\omega), \end{aligned} \quad (7.1)$$

where φ is the projection angle, $\mathcal{W}(\omega)$ is a smoothing function, ϕ is the scaling function, ψ is the wavelet function, $\psi_u = u\psi(u \vec{t})$, $\phi_u = u\phi(u \vec{t})$, $\tilde{\psi}(t) = \psi(-t)$, $\tilde{\phi}(t) = \phi(-t)$, and \mathcal{F}_{1D}

represents the Fourier transform. Back projection of the filtered X-ray projections recovers the 2D wavelet subbands of the desired ROI. Applying an inverse 2D wavelet transform, the image is recovered. Consequently, to recover the subbands, four back projections should be applied with complexity of order $O(N^3)$, for an $N \times N$ pixel image. This complexity can be reduced by utilizing central slice theorem (CST) that relates the 1D Fourier coefficients of the X-ray projection to the 2D Fourier coefficients of the image. Ideally, CST enables use of Fast Fourier Transform (FFT) algorithms instead of back projection to reduce the computational load. However, to recover the images, a polar to Cartesian interpolation must be used requiring use of oversampled projections to decrease the interpolation error, which in turn, increases the radiation dose.

7.1.2 Compressed Sensing based CT Reconstruction

It has been shown that CT images are sparse in the wavelet domain, enabling CS to recover images from fewer projections, with a corresponding reduction of the radiation dose. CS-based CT reconstruction uses an optimization functional similar to equation (3.15). Conventional CS-based CT recoveries use the discrete Radon matrix, but suffer from high computational complexity due to the size of the measurement matrix. Using the sparsity of the wavelet transform, an ℓ_1 optimization problem is solved to recover the image from under-sampled measurements. Various algorithms have been proposed to solve this optimization problem, as discussed in section 3.4. Some methods, such as ADM based methods, use a surrogate relaxed reformulation of the optimization problem and the regularization term. This introduces extra regularization parameters that makes the tuning task more difficult [98]. Alternatively, as used here, an iterative shrinkage/thresholding algorithm (ISTA) [95] can be used. This is a very simple and nonparametric method consisting of a Landweber updating step followed by soft thresholding. The pseudo-code of ISTA is provided in Algorithm 1.

A potential difficulty of ISTA is its slow convergence especially when the problem is ill-posed (e.g., few X-ray projections are measured). Different approaches have been proposed to accelerate the convergence: 1) By combining the result of past iterations to speed up convergence, such as two-step iterative shrinkage/thresholding (TwIST) [27] and fast ISTA (FISTA) [159]; 2) By using a wavelet subband dependent CS methods like subband adaptive ISTA (SISTA) [97, 190].

7.2 Proposed method: AWiR-SISTA

As noted earlier, the WiR transform recovers wavelet subbands of local CT images, and CS-based CT reconstruction algorithms use wavelets as a sparsifying transform to recover the images from a reduced number of projections. Moreover, subband CS recovery is known to improve the convergence rate of ISTA. Using an improved version of WiR with CS-based CT reconstruction, we propose a local subband CS-based CT reconstruction algorithm, called Accelerated Wavelet inverse Radon subband adaptive Iterative Shrinkage/Thresholding Algorithm (AWiR-SISTA), which benefits from both techniques to further reduce the effective radiation dose. In addition, the proposed method decreases the complexity of the WiR, reduces the computational load of CS recovery and improves the convergence rate of the ISTA, used in CS recovery.

7.2.1 Accelerated Wavelet inverse Radon (AWiR) Transform

To accelerate the recovery of the subbands, the back projection step is replaced by pseudo-polar Fourier transform (PPFT) based DFR. The proposed accelerated WiR (AWiR) is shown in figure 7.3-B. While the DFR reduces the computational complexity, the polar to Cartesian interpolation makes this approach problematic. Pseudo-polar Fourier transform [191] addresses this problem by proposing a fast recovery algorithm (with complexity of $O(N^2 \log N)$) that directly computes the images from the Fourier samples measured on the pseudo-polar girds, given by equation (2.7).

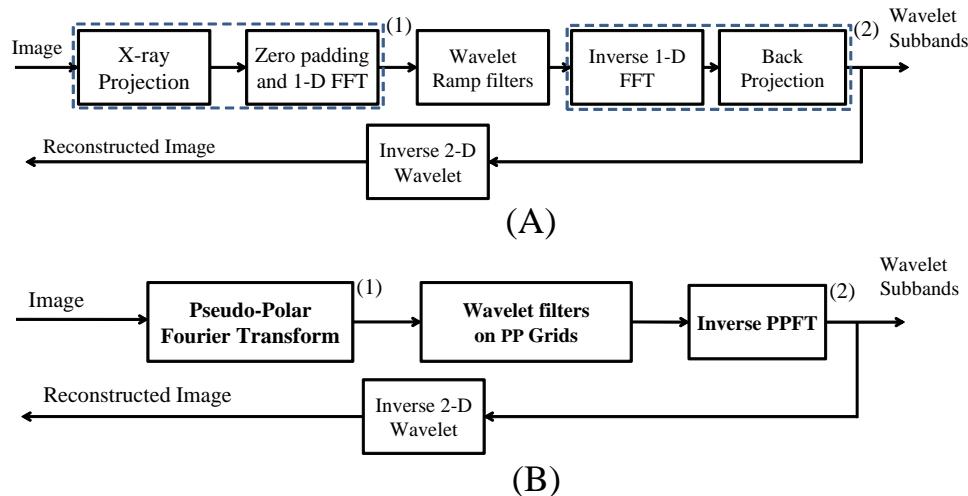


Figure 7.3: Wavelet inverse Radon transform methods. (A) Wavelet inverse Radon transform according to [31], and (B) Accelerated wavelet subbands recovery. The blocks marked by (1) and (2) in (A) are replaced by the corresponding ones in (B).

Because these grids are closer to polar grids the interpolation error is significantly reduced. In addition, because CT scanners are theoretically able to measure the X-ray projections on equally sloped lines, as given by equation (2.9), enabling pseudo-polar Fourier samples to be computed from zero-padded 1D Fourier transforms of the equally sloped samples with negligible error. In AWiR the X-ray projections are measured or interpolated on the pseudo-polar grids and are multiplied by the wavelet filters that are interpolated on pseudo-polar grids with $\mathcal{W}(\omega) = 1$: a process referred to as pseudo-polar wavelet filtering. The wavelet subbands are then recovered using a fast inverse pseudo-polar Fourier transform as proposed in [54].

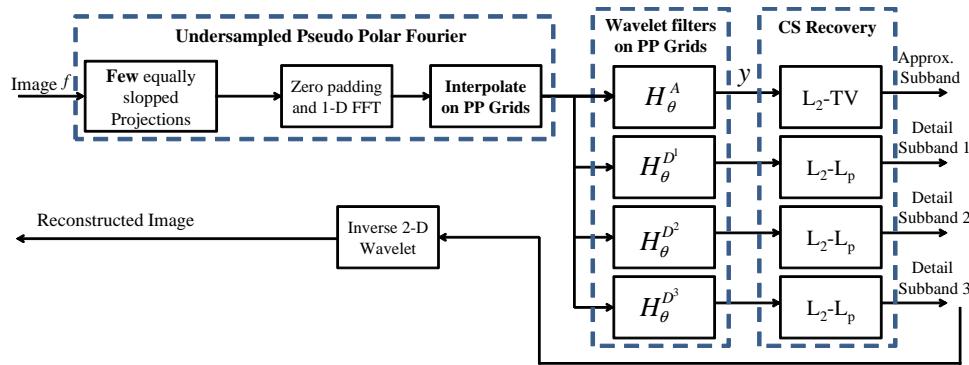


Figure 7.4: Proposed AWiR-SISTA method for local CT image reconstruction.

7.2.2 AWiR-SISTA steps

In the AWiR-SISTA method, illustrated in figure 7.4, the inverse PPFT block of AWiR (see figure 7.3-B) is replaced by a CS recovery enabling the algorithm to reconstruct the wavelet subbands from a reduced number of projections. The computational load of the CS reconstruction is significantly reduced by replacing the huge X-ray projection measurement matrix with a fast PPFT algorithm. This combination enables independent recovery of the wavelet subbands; and the independent recovery of the subbands accelerates the convergence rate of ISTA. In addition, the different statistical properties of the subbands effectively adjust the thresholding algorithm of each subband, as proposed in [155].

To recover the local CT images the following steps are performed in AWiR-SISTA:

Step 1: The CT projections are measured on a few equally sloped radial lines or they are redistributed (rebinned) from other geometries to equally sloped beams [111, 152]. A 1D interpolation is used to map the 1D Fourier transform of the projections to pseudo-polar grids using zero padding to decrease the interpolation error.

Step 2: This data is multiplied with pseudo-polar wavelet filters, resulting in the PPFT of the

wavelet subbands at the corresponding angles¹.

Step 3: CS recoveries are used to reconstruct each wavelet subband separately from the undersampled data calculated in step 2. A fast PPFT is used as the measurement function, A , in the CS recoveries.

Step 4: An inverse 2D wavelet is used to recover the ROI of the CT image from the reconstructed subbands.

Step 5: Interior CT images usually have a bias [31]. Few steepest decent updating steps are applied on the images recovered in step 4 to reduce the recovery error and to decrease this bias. An additional advantage of AWiR-SISTA is that different optimization schemes can be used to recover the subbands, based on their statistical properties. For instance, the approximation subband is not as sparse as the detailed subbands, but is smoother, which makes the use of $\ell_2 - TV$ recovery, similar to equation (3.22), more beneficial. In this application u is the approximation wavelet coefficients, and y denotes the pseudo-polar Fourier coefficients of the projections multiplied by H_φ^A .

In addition, our observations showed that a non-convex $\ell_2 - \ell_p$ minimization:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda \|W^T \mathbf{x}\|_p, \quad (7.2)$$

is more accurate for detailed subband recovery, where $0 \leq p < 1$ is estimated from statistical model of wavelet subband as described in [155], \mathbf{x} denotes the wavelet detailed subbands, and \mathbf{y} is the pseudo-polar Fourier coefficients of the projections multiplied by $H_\varphi^{D^i}, i \in [1, 2, 3]$.

For further convergence acceleration, TwIST [27] was used to solve the ℓ_p - and TV-minimization problems. It applies the following update at $(l+1)^{th}$ iteration:

$$\hat{\mathbf{x}}^{l+1} = (1 - \alpha) \hat{\mathbf{x}}^{l-1} + (\alpha - \beta) \hat{\mathbf{x}}^l + \beta \Gamma(\hat{\mathbf{x}}^l), \quad (7.3)$$

in which α and β are two constant which should be in a range defined in [27], and $\Gamma(\mathbf{x}) = \gamma(\mathbf{x} + \mathbf{A}^T(\mathbf{y} - \mathbf{Ax}))$, where $\gamma(\mathbf{x})$ is a denoising/smoothing function, depending on the optimization problem that TwIST is used for, \mathbf{y} is the 1D Fourier coefficients of the equally sloped projections interpolated on pseudo-polar grids multiplied by the corresponding wavelet filter, and \mathbf{A}^T is the conjugate transpose of PPFT [54]. For detailed wavelet subband recovery ℓ_p -minimization is solved using the shrinkage operator given in equation (6.6). For the approximation subband recovery, $\gamma(x)$ is a TV denoising function that is solved by the split-Bregman method, as proposed in [99] and provided in Algorithm 4.

¹The idea of combining pseudo-polar transform and wavelet transform has been used in Ridgelet transform in a different way [192]

7.3 Results

A simulated 512×512 Shepp-Logan head phantom was used for testing the proposed algorithm. All the tests used Daubechi 4 together with $p = 0.5$ in $\ell_2 - \ell_p$ optimizations for the wavelet detailed subband recoveries. Equally sloped projections were calculated on 1024 lines along the angles given by equation (2.9). Figure 7.5 shows the image and its subbands recovered from 200 projections, selected uniformly from the calculated projections, using WiR and AWiR-SISTA. As can be seen, the proposed method fully recovers the image and its subbands, with a normalized error of $\|x - \hat{x}\|_2 / \|x\|_2 = 1.02 \times 10^{-3}$.

Figure 7.6 shows an ROI at the center of the phantom with $1/2$ of the radius of the full phantom recovered by WiR from a full data set (1024 projections) and by AWiR-SISTA from 200 projections. Since the ROI radius is half, the effective radiation dose is half of that needed for the full phantom. In addition, only $1/5$ of the projections are used to recover the ROI with AWiR-SISTA. Consequently, the effective radiation dose is $1/10$ of the radiation used to recover a full phantom with AWiR, which needs 1024 projections for an exact recovery. The bias of the reconstructed ROI is reduced from 0.4 in WiR to 10^{-2} in AWiR-SISTA. Figure 7.7 shows an ROI from a lung CT image recovered by WiR and AWiR-SISTA from 200 projections (50mAs, 120kVp), which were selected equiangularly from 900 X-ray projections taken using a Toshiba Aquilion ONE[©] scanner (Toronto General Hospital, Canada). Data from the central row of a volumetric scan on one single rotation served as the fan beam data and were rebinned to pseudo-polar beams. AWiR-SISTA decreases the recovery error by almost an order of magnitude (4×10^{-1} to 0.7×10^{-2}). It should be noted that the error in WiR is due to the ROI reconstruction bias and CT# changes caused by the use of few projections.

Finally, figure 7.8 compares the quality of a non-centered ROI recovered by AWiR-SISTA and WiR from only 200 projections. However, since variable collimation is needed in different projection angles of the non-centered ROI imaging, its practical implementation may not be possible with the available hardwares.

7.4 Summary

A subband dependent compressed sensing algorithm was proposed for ultra low dose local CT imaging. The proposed method (AWiR-SISTA) not only reduced the X-ray radiation by reducing the X-ray exposure area and the number of projections, but also decreased the computational complexity as a result of using a fast PPFT algorithm. Independent recovery of the subbands accelerated the convergence of CS recovery. It also provided a possible means for accounting for the different statistical properties of the subbands in CS recovery and, by paral-

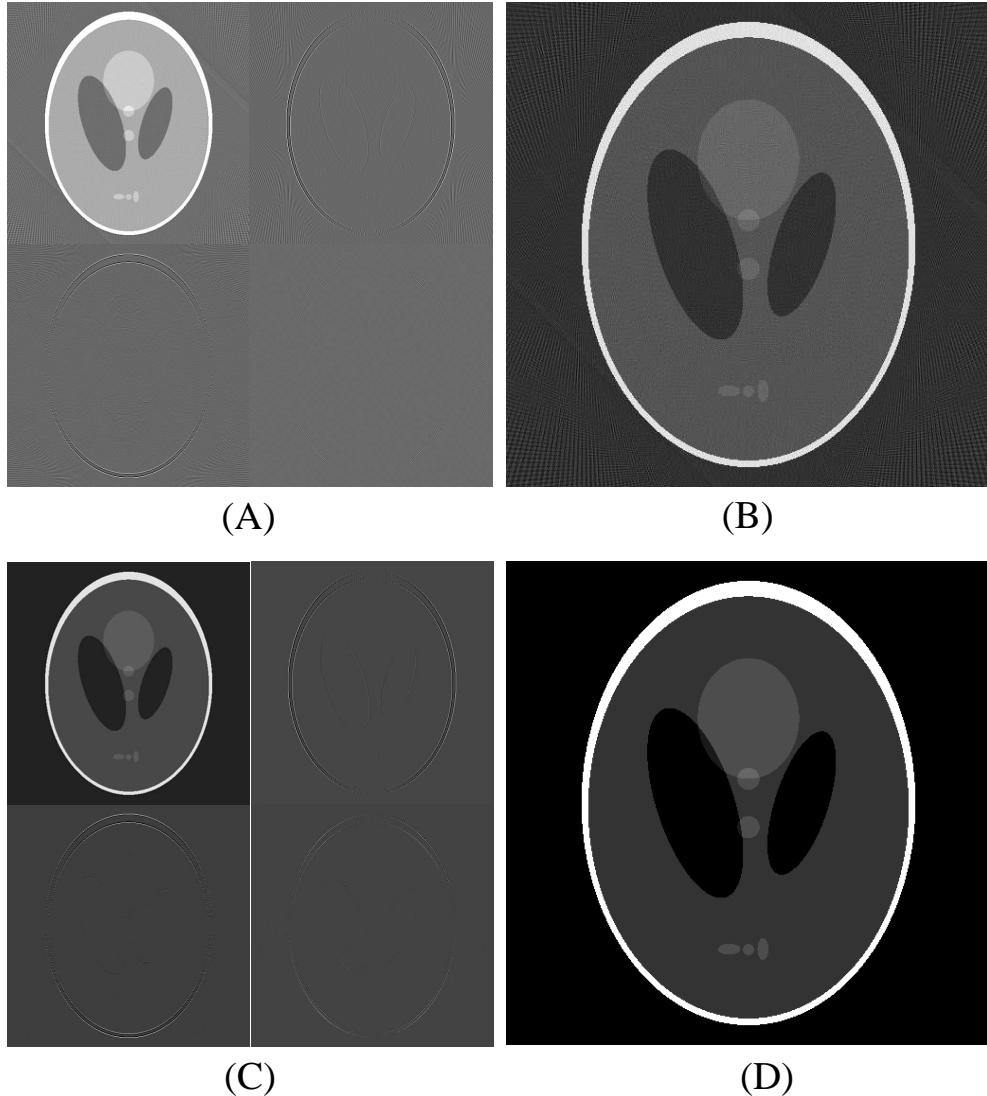


Figure 7.5: Comparison of full Shepp-Logan phantom images reconstructed from 200 projections by WiR and AWiR-SISTA: (A) Subbands recovered by AWiR, (B) image recovered by AWiR with a normalized error of 2×10^{-1} , (C) image subbands reconstructed with AWiR-SISTA, and (D) phantom reconstructed by AWiR-SISTA with a normalized error of 1.02×10^{-3} .

lel reconstruction of the subbands, it increased the overall image reconstruction speed. Results showed that a high quality ROI with half the radius of the phantom can be recovered from 200 projections, leading to a 90% reduction in the radiation dose. As a result, our proposed method provides strong evidence that the cancer risk associated with CT can be substantially reduced without degrading the image quality.

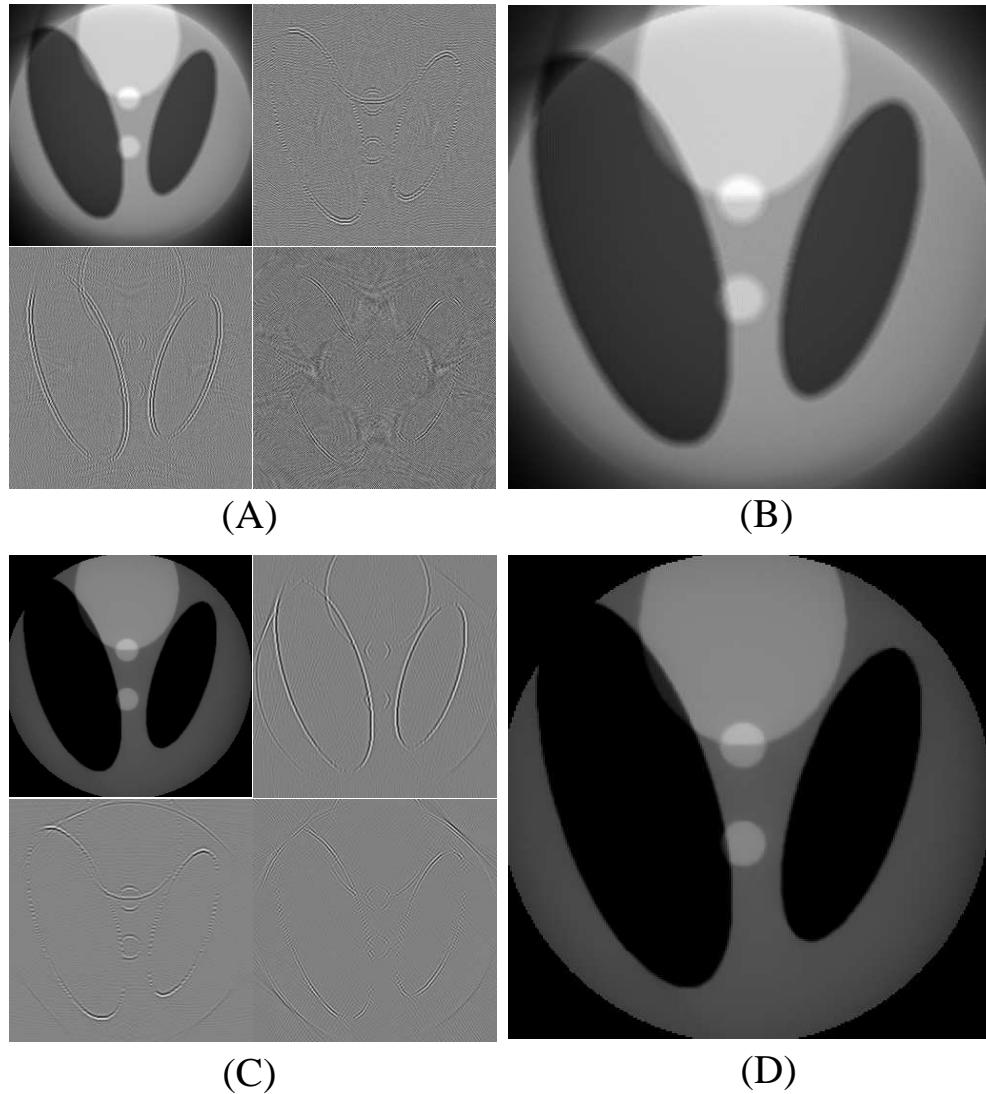


Figure 7.6: An ROI with radius of $1/2$ of the phantom: (A) Subbands recovered by WiR from full data, (B) ROI image recovered by WiR with a normalized error of 2×10^{-1} (C) Subbands recovered by AWiR-SISTA, and (D) ROI image recovered by AWiR-SISTA from 200 projections with a normalized error of 0.8×10^{-2} .

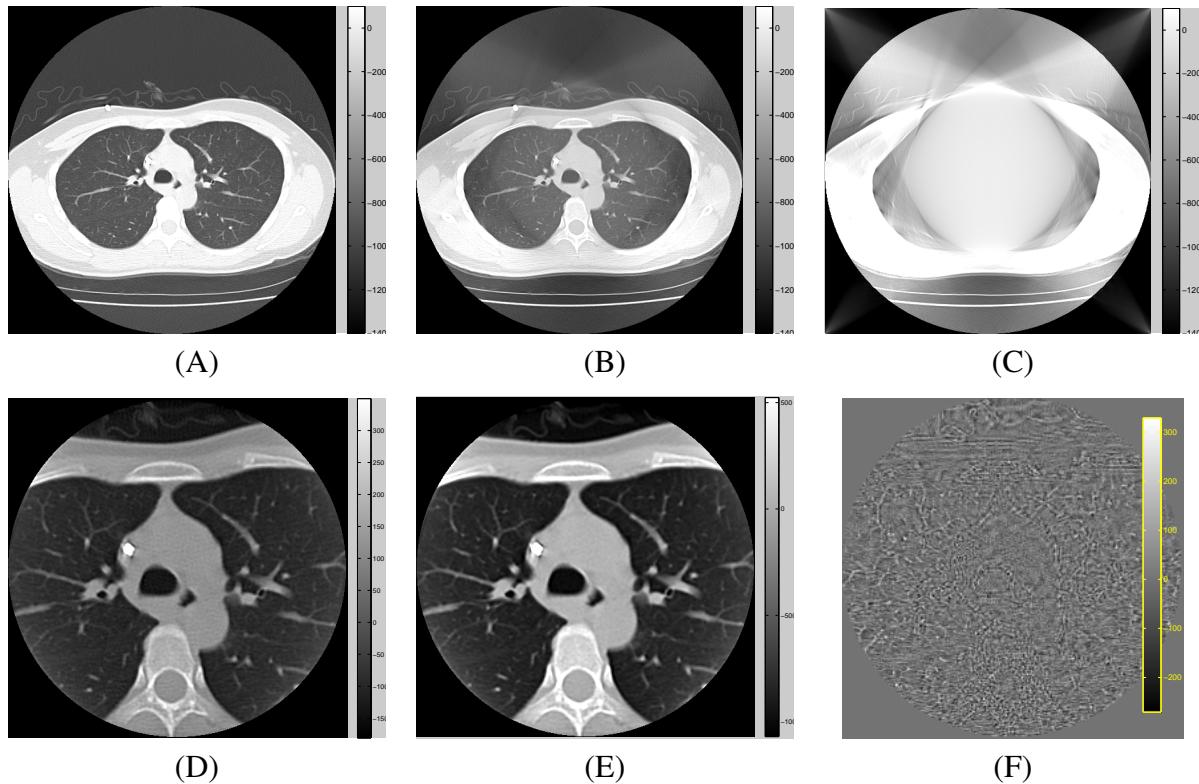


Figure 7.7: Patient ROI with radius of $1/2$ that of a lung CT image recovered by. (A) Original Image, (B) Difference between the ROI image reconstructed by WiR and the original image, (C) Difference between the ROI image reconstructed with the proposed method, (D) WiR with normalized error of 4×10^{-1} , (E) AWiR-SISTA with normalized error of 0.7×10^{-2} , and (F) Difference of the original ROI image and the ROI recovered by AWiR-SISTA.

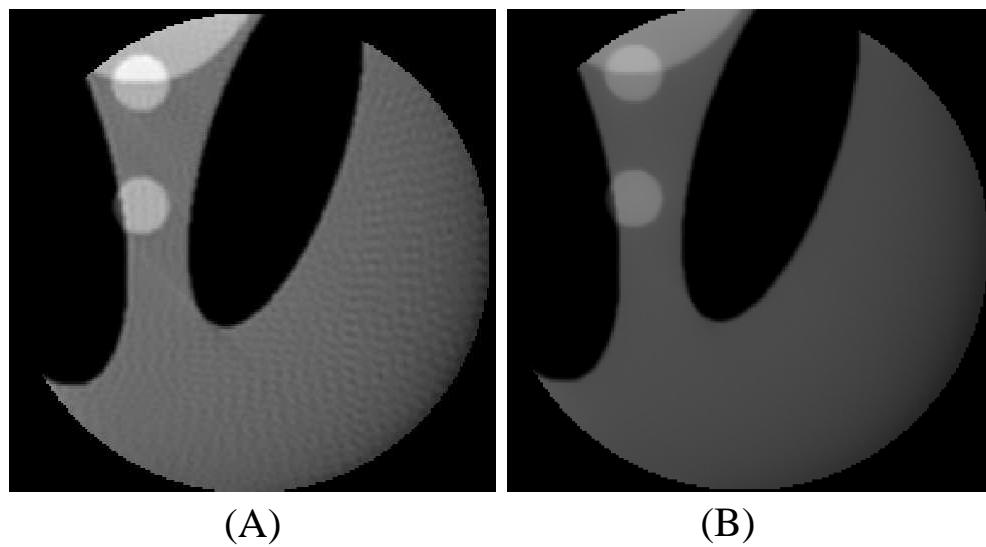


Figure 7.8: A non-centered ROI in a Shepp-Logan phantom with a radius $1/2$ phantom reconstructed from 200 projections by: (A) WiR with a normalized error of 2.5×10^{-1} , and (B) AWiR-SISTA with a normalized error of 1.1×10^{-2} .

Chapter 8

Conclusions, Summary and Future Directions

While the average effective dose of a standard dose chest CT is approximately 8 mSv, different studies have reported the dose of posteroanterior and lateral chest X-ray radiography to be between 0.05 and 0.24 mSv [193, 194, 195]. The goal of low dose CT imaging is to decrease this difference without sacrificing the clinical quality of the images. To provide some perspective on this issue, it should be noted that the low dose FBP protocol used at the Toronto General Hospital uses 50 mAs (current-rotation time product) and 120kVp, corresponding to a radiation dose of about 1mSv.

The purpose of this study was to examine potential methods of achieving a reduction in CT dose without significantly degrading the quality of the CT images. For this purpose, it was assumed that the effective X-ray dose in CT is proportional to the product of the exposure area, the number of projections and the dose for each projection. Based on this simplified approach, four X-ray dose reduction approaches were proposed: each of which controls one of these factors.

Increasing the slice thickness increases the number of photons and therefore decreases the noise, but at the same time it reduces the spatial resolution and the accuracy of the measurements. Considering this relation, in chapter 4 the slice thickness/overlap reconstruction was optimized for optimal nodule detection and characterization during thoracic low dose CT. In addition, effect of AIDR3D, Toshiba's iterative method, was tested on the detectability of the nodules. Two CT slice reconstructions were suggested: 4.0/2.0 mm reconstructions through the entire thorax using a lung kernel and lung window display for nodule detection, and 2.0/2.0 mm reconstructions through the nodule using a soft tissue kernel and display for optimal nodule characterization. If practical considerations require a single protocol for clinical reading of low dose CT images, then a 3.0/1.5 mm slice thickness/overlap reconstruction was a satisfac-

tory compromise for both reconstructions.

The method studied in chapter 5 was to reduce the number of projections. It was shown that CT reconstruction can be statistically modeled as a weighted compressed sensing optimization problem. The proposed model was derived from the MAP model of CT imaging with sparsity and piecewise linearity constraints. To solve the proposed model a fast CS recovery method was proposed in which pseudo polar Fourier transform was used as the measurement function to reduce the computational complexity. Moreover, to be able to reconstruct CT images from fan beam and helical cone beam projections, rebinning to parallel beams was used. To enable the proposed CS recovery method to remove the interpolation error, a weighting approach (EAW) was proposed, in which the weights accounted for the measurement noise and interpolation errors. This enabled CT images to be reconstructed from a reduced number of fan or helical cone beam X-ray projections. It was shown that using EAW improves the reconstruction quality substantially. For instance, a 512×512 Shepp-Logan phantom reconstructed with 128 projections using a conventional CS method had $\sim 10\%$ error. However, using the same data with our proposed method the reconstruction error was as low as $\sim 1\%$, leading to $\sim 78\%$ dose reduction. The proposed weighted CS-based CT reconstruction model was solved with a proposed Expectation Maximization based method, called FCSA-LEM. The low computational complexity of our FCSA-LEM method made fast recovery of the CT images possible. For example, we were able to recover a 512×512 image in less than 30 sec on a desktop computer without numerical optimizations, thus our proposed method may be among the first CS-CT methods whose computational complexity is within the realm of what could be clinically relevant today.

Improving the quality of low dose CT images, reconstructed by FBP, was studied in chapter 6. An iterative denoising scheme was proposed for low dose CT images, which adjusts the denoising parameters at each iteration based on the effect of the denoising method in the previous iteration. Noise confidence region evaluation (NCRE) was used to compare the noise with denoising residual to determine if the denoising was effectively, weakly or strongly executed. Based on this information the denoising parameters were adjusted for the next iteration. BM3D was used in the proposed iterative scheme, which improved the results of the BM3D method in the sense of both MSE and feature preservation, based on qualitative assessments. The phantom study showed that our proposed method improved low contrast detectability. The patient study demonstrated that the image was efficiently denoised and the visibility of small objects was preserved. In addition to the denoising problems, NCRE was used in compressed sensing optimization problems to adaptively adjust the regularization parameters. It improved the performance of ISTA in sense of the recovery error (MSE) and helped it to converge faster. In chapter 7, a subband dependent compressed sensing based algorithm was proposed for ul-

tra low dose local CT imaging. The proposed method (AWiR-SISTA) not only reduced the X-ray radiation by reducing the X-ray exposure area and the number of projections, but also decreased the computational complexity as a result of using a fast PPFT algorithm. Independent recovery of the subbands accelerated the convergence of CS recovery. It also provided a possible means for accounting for the different statistical properties of the subbands in CS recovery and, by parallel reconstruction of the subbands, it increased the overall image reconstruction speed. Results showed that a high quality ROI with half the radius of the phantom can be recovered from 200 projections, leading to a 90% reduction in the radiation dose. As a result, our proposed method provides good evidence that the risks associated with CT can be substantially reduced without degrading the image quality.

8.1 Contributions

The following contributions have been made within this thesis:

- A slice thickness/overlap reconstruction was proposed to optimize the trade-off between the detection and maximum characterization accuracy of the small non-calcified pulmonary nodules. For the first time the effect of slice thickness and overlap were considered together. In addition, the effect of table position was considered and minimized for the first time.
- A new weighted compressed sensing formalization was proposed for CT reconstruction, based on statistical modeling of the reconstruction process. The weights were mathematically calculated as a function of the measurement noise.
- A fast recovery algorithm was proposed to solve the proposed optimization scheme. The computational complexity was significantly reduced by the use of pseudo-polar Fourier transform.
- The model was extended to fan and helical cone beam geometries, by rebinning. To compensate the interpolation error caused by rebinning, the weights of the CS model, denoted by Error Adaptation Weight (EAW), were modified.
- A new method was proposed to adaptively update the regularization parameters in the optimization problems formulated in the form of Lagrangian multiplier. This method, denoted by NCRE, improved different denoising algorithms in the sense of recovery error. In addition, it was shown that this method can improve the performance of iterative soft thresholding algorithms to solve ℓ_p , $0 < p \leq 1$, minimization problems.

- A new iterative scheme was proposed for CT image denoising. A stopping criterion was proposed for iterative denoising algorithms, which can determine if the image is well denoised, partially denoised, or over-smoothed.
- A compressed sensing based local (interior) tomography algorithm was proposed. The proposed algorithm used compressed sensing to directly recover the wavelet subbands, which accelerated the convergence rate of the recovery algorithms and enabled them to consider the statistical properties of each subband independently.

8.2 Suggestions for Future Work

1. To find the exact amount of dose that can be reduced by the proposed weighted CS-based CT model, more accurate studies, similar to the ones done in [196], should be done to compare the standard CT quality metrics of the proposed method with the statistical iterative algorithms. The results should be followed by a more clinically relevant test; and the quality of FBP, SIR, and CS-based reconstructed images should be compared by radiologists. To measure and compare the quality of the images in the context of a clinical task, such as nodule detection in thoracic CT scans, receiver operating characteristic analysis can be used [197].
2. The experimental studies reported in this thesis were performed with a Toshiba scanner that used equiangular scan trajectory. To be consistent with the real data, the simulation results were also based on the equiangular projections. However, the rebinning error could be minimized by optimizing the projection angles. For example, the optimum angles for a fan beam system similar to Toshiba scanner, used in this thesis, are shown in figure 8.1, with solid blue line. As can be seen, these angles are slightly different from the pseudo-polar angles, shown by red dashed line. As a result, more accurate results could be achieved if the projection angle of the scanner could be controlled.
3. The helical cone beam reconstruction method needs to be more carefully investigated. The algorithm described in this thesis could be improved by using better rebinning algorithms, such as the PI based methods [198].
4. The application of the proposed method can be extended to cone beam CT by the use of Grangeat formula. Unlike fan beam projections, which are the same as the 2D Radon transform, cone beam projection is different from the 3D Radon transform. Consequently, 3D central slice theorem cannot be used to accelerate the image reconstruction. The Grangeat formula relates the cone beam projections to 3D Radon transform and

therefore to 3D central slice theorem [199]. Combining the proposed method with the Grangeat formula and including the accrued interpolation errors in EAW, cone beam images can be reconstructed very fast. However, since cone beam data do not satisfy the sufficiency conditions, regular methods may not be able to recover high quality images. In this case 3D patch based recovery approaches could be used.

5. As shown in [200, chapter 4], the interpolation method has a major impact on the quality of the reconstructed images. The influence of this on the proposed weighted CS model could be explored.
6. Non-diagonal elements of D in equation (5.5) can be used to model photon scattering [201]. This could be explored in future studies.
7. The proposed NCRE-BM3D denoising algorithm should be tested for medical cases and be evaluated by radiologists. In addition, more accurate noise models should be considered.
8. The proposed local CS-based CT reconstruction should be combined with the proposed weighted CS formulation to extend its application to fan and helical cone beam geometries and to improve the quality of the images.
9. To date, the choice of regularization weights has not been justified in any publication. Optimal weights are inconsistent across different data sets, and may change slightly for the same data set with different sampling ratio. Therefore, improved CS models, which are insensitive to the regularization parameter, can be expected to apply to routine clinical applications. This problem was studied briefly in chapter 6 with some preliminary results. Further tests, with real data and real applications, could be a future research area [202, 179].

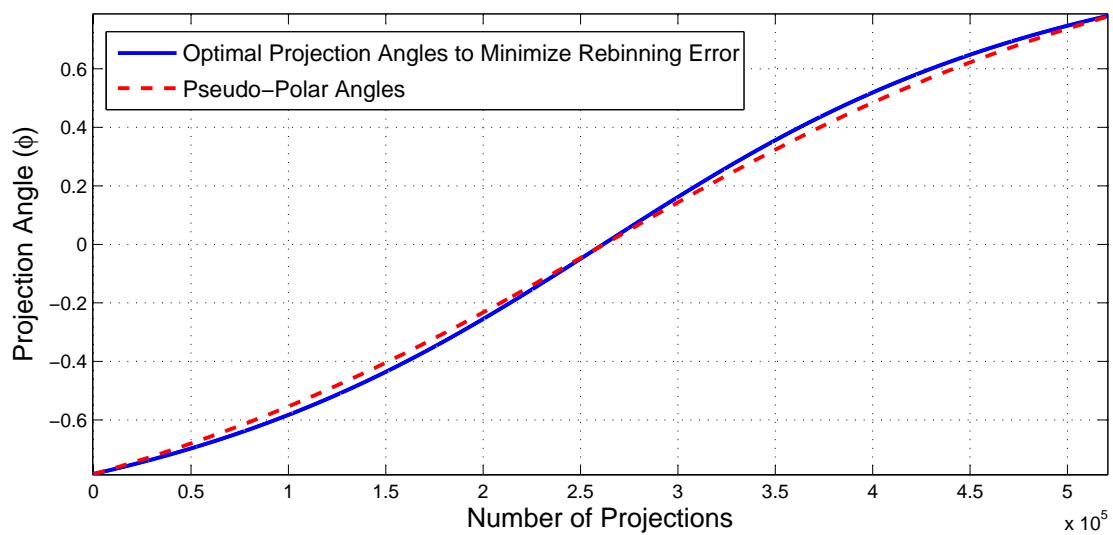


Figure 8.1: The projections angles that minimize the fan beam to pseudo-polar parallel beams interpolation error.

Appendix A

Noise in Computed Tomography Reconstructed Images

A.1 Noise in CT images

X-ray projection measurements in CT scanners are based on counting the photons that hit the detector channels. Although a reasonable statistical model for this case is the independent Poisson distributions [43], it has been shown that the corrected polyenergetic X-Ray projections can be modeled more accurately by Gaussian distribution with the following relationship between its mean and variance:

$$\sigma_{i,n}^2 = \sqrt{(J_i \times \exp(\bar{\xi}(i,n)/s))^2 + (\sigma_e^2)_{i,j}}, \quad (\text{A.1})$$

where $\bar{\xi}$ is the mean and $\sigma_{i,n}^2$ is the variance of the projections at i^{th} projection angle (φ_i) and n^{th} detector bin placed in distance l_n from detectors center, s is a scaling factor, σ_e^2 is electronics noise variance, and J_i is a parameter adaptive to different detector channels [28].

During the reconstruction process noise distribution is changed as a function of the reconstruction algorithm, scan protocols (e.g., using current modulation and bowtie filters), denoising and ramp filters. Due to the complicated dependencies of noise on scan parameters and on spatial position, noise distribution in the reconstructed CT images is usually unknown. Using discrete parallel filtered back projection relation:

$$f(x,y) = \frac{\pi\Delta t}{N_{\pi g}} \sum_{i=1}^{N_{\pi g}} \sum_{n=1}^{N_g} c(x \cos \varphi_i - y \sin \varphi_i - l_n) g_{i,n}, \quad (\text{A.2})$$

the noise in reconstructed images can be described by [203]

$$\sigma_n^2(x, y) = \left(\frac{\pi\Delta t}{N_{\pi g}}\right)^2 \sum_{i=1}^{N_{\pi g}} \sum_{n=1}^{N_g} c^2(x \cos \varphi_i - y \sin \varphi_i - l_n) \sigma_{i,n}^2, \quad (\text{A.3})$$

where $g_{i,n}$ is the parallel projection at i^{th} angle and n^{th} detector bin, $c(.,.)$ is the ramp filter in spatial domain, and $\sigma_{i,n}^2$ is the projection noise (logarithm of the photon count measurements), which depends mainly on the photon count and electrics noise (e.g. $\sigma_{i,n}^2 \approx \frac{(\sigma_e^2)_{i,j} + \bar{\xi}(i,n)}{\bar{\xi}^2(i,n)}$). Equation A.3 can be interpreted as the back projection of the projection noise variances, which makes it non-stationary, object dependent, and correlated. Moreover, since the variance at a certain position is a summation of the variances from many angles, if one of the variances in one direction is significantly larger than the other projection angles, the variances along that direction are more correlated than the other directions [204]. This is observed as streak artifact in the reconstructed images, as shown in figure A.1. In addition, due to the signal processing algorithms used to improve the quality of the images and to reduce the image artifacts, the noise characteristics are changed in the reconstructed image, which could be different in the images acquired and reconstructed by different scanners and different vendors.

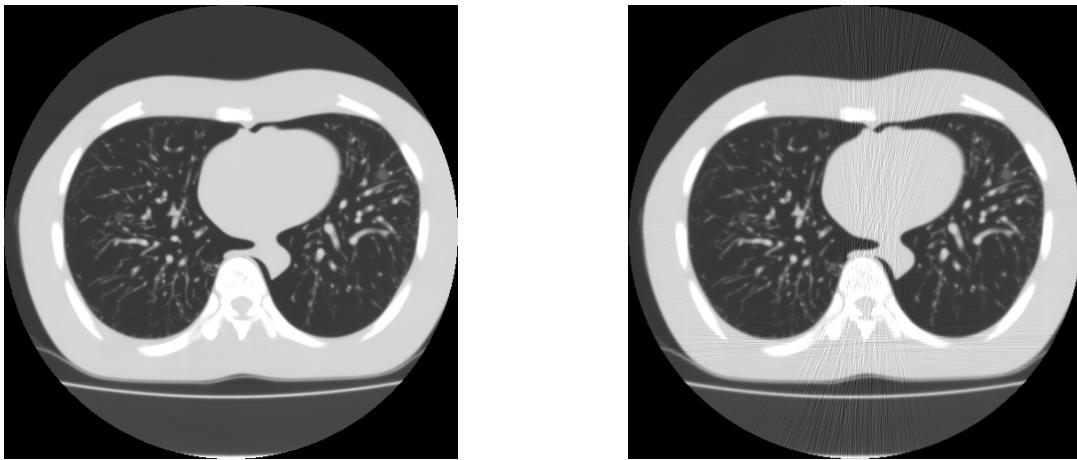


Figure A.1: Streak artifact caused by the presence of high attenuating bone structures in the image. Left: Image taken from Toshiba Aquilion one CT scanner with corrections, and Right: image without the corrections.

Various approaches have been proposed to reduce the effect of noise in CT images, such as projection noise removal prior to the image reconstruction step [28], and statistical iterative CT reconstructions [15], which suppress the noise during the reconstruction by the use of more realistic assumptions compared to FBP. Another common area of research includes the development of algorithms for noise reduction in reconstructed CT images. These methods are required to reduce noise while preserving edges, as well as small structures that might be im-

portant for diagnosis. Partial differential equation (PDE) based methods [29, 7] have been used as the standard edge-preserving methods in image domain denoising.

In many CT image denoising algorithms, noise is modeled as additive white Gaussian noise (AWGN) $f(x, y) = \bar{f}(x, y) + \mathbf{n}(x, y)$, where $(x, y) \in \mathbb{R}^2$ is the coordinate of the image pixels, \bar{f} is the noiseless image, and \mathbf{n} is a zero mean white Gaussian noise with variance of σ_n^2 [186]. This model provides a great deal of flexibility to use state of the art denoising algorithms. Examples of such algorithms are multi-scale transform based denoising methods [205], anisotropic diffusion [29], and sparsity based denoising and dictionary learning methods [206, 207]. However, this model is not accurate, and thereby, the denoising methods based on this assumption are not optimum. Other methods are available, which consider non-stationary, correlated additive Gaussian noise models [28, 203, 204, 205, 208, 209, 189] for better CT image denoising. In this appendix we study the noise statistics of the reconstructed CT images. It is shown that in small neighborhoods with the same CT# the image noise can be modeled by an additive white Gaussian distribution. This result is used in Chapter 6 to denoise the CT images by BM3D-NCRE. Toshiba Aquilion One CT scanner is used to scan the phantom shown in figure A.2. The same study can easily be done by other scanners.

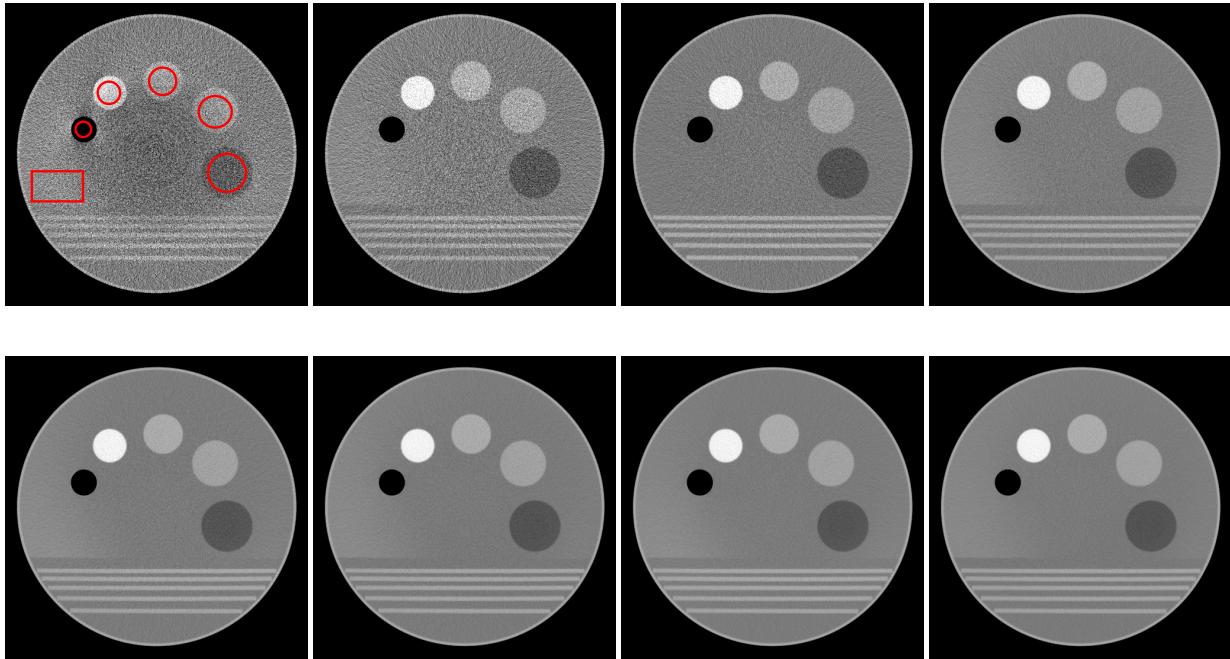


Figure A.2: Phantom scanned with eight different X-ray source currents: top left to right 5, 10, 25, 50 mAs and bottom left to right 100, 150, 200, 250 mAs.

A.2 Notes on Noise Statistics in reconstructed CT images

Six regions are considered in the phantom, a region in background and 5 regions inside the circles with different CT#s. This phantom is scanned with eight different doses controlled with changing the tube current as [5,10,25,50,100,150,200,250](mAs). The distribution of noise in each region is compared with additive white Gaussian noise in figure A.3. As can be seen, the noise distribution matches the white Gaussian noise with a large correlation.

Figure A.4 shows the noise variance changes for each region with different CT#s and figure A.5 shows the dependence of the noise variances on the CT# values. Noise changes, as shown in these figures, are not very significant for tube currents more than 50 mAs. This justifies the application of 50 mAs as the optimum current used in the routine diagnostic low dose protocol in Toronto General Hospital. In addition, these figures show that CT image noise is locally Gaussian with variable variances as a function of the CT# values.

A.3 Noise Considerations in our proposed Algorithm (NCRE-BM3D)

Based on the results of the last section, we use a non-stationary additive Gaussian noise model, in which the noise variance depends on the location and attenuation value of the pixels. To simplify the problem, only the pixels with soft tissue attenuations around the lungs are used in NCRE to validate the noise statistics. These pixels have similar CT# values and since they only include a narrow region around the lungs, have almost the same radial locations, as shown in figure A.6. Consequently, the noise is assumed to be an identical additive Gaussian noise in these regions.

In addition, BM3D-NCRE method uses the collaborative filtering fundamentals of BM3D, which groups the similar patches of the image in 3D stacks and performs the denoising on each 3D block. Therefore, due to the similarity of the patches, the noise variance in each block is almost identical and the denoising of the patches with independently estimated variances is accurate enough. This property is shown in figure A.7.

A.4 Summary

Although the CT image noise is not a pure additive white Gaussian noise, it was shown here that the noise in the soft tissue area around the lung is very close to white Gaussian noise. Moreover, it was shown that in BM3D, due to the use of similar patches selected from a local

region in the CT images, noise characteristics are very close to those of an additive white Gaussian noise.

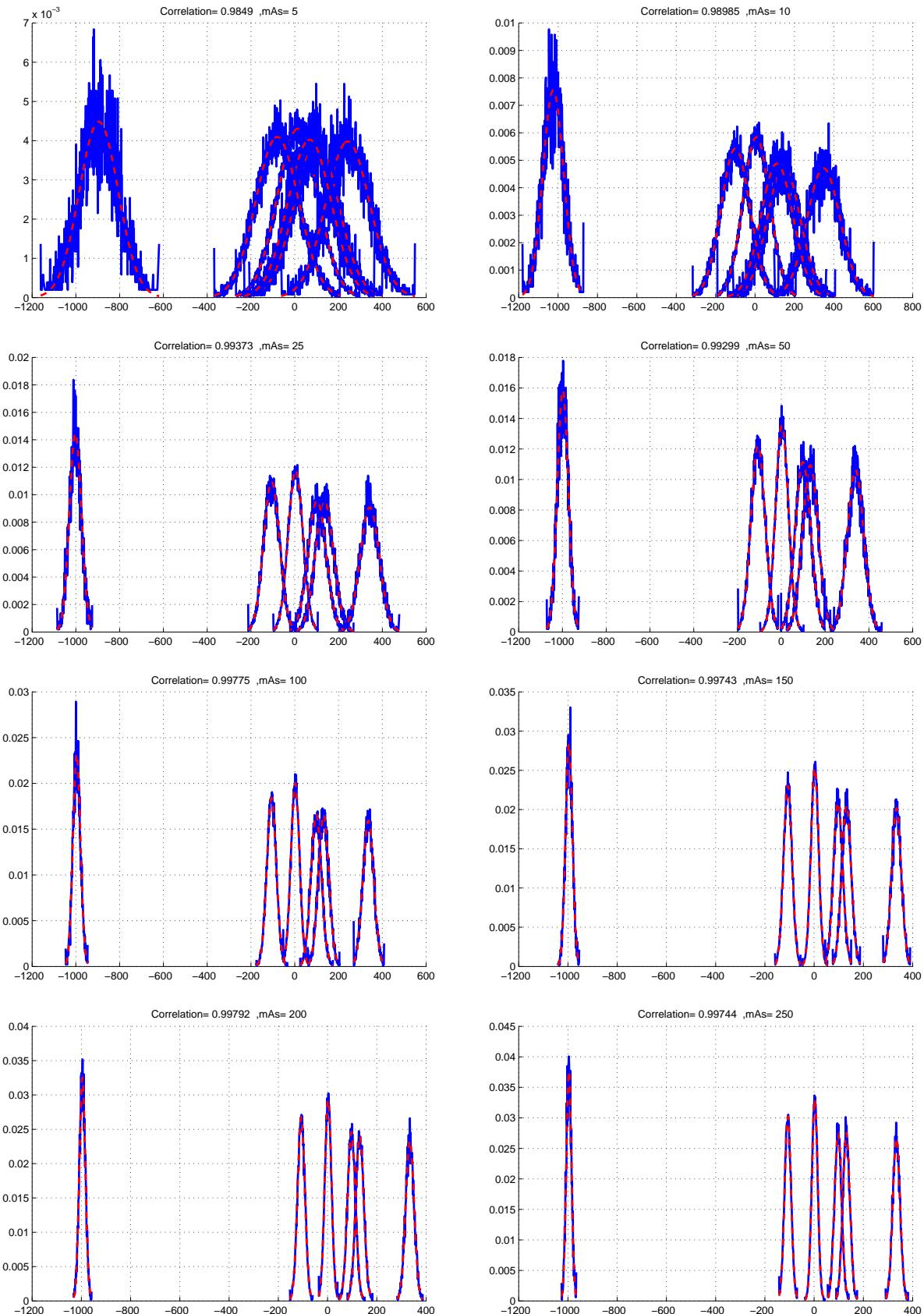


Figure A.3: Compares the noise distribution of the noise in the six regions of the phantom shown in figure A.2 with white Gaussian noise.

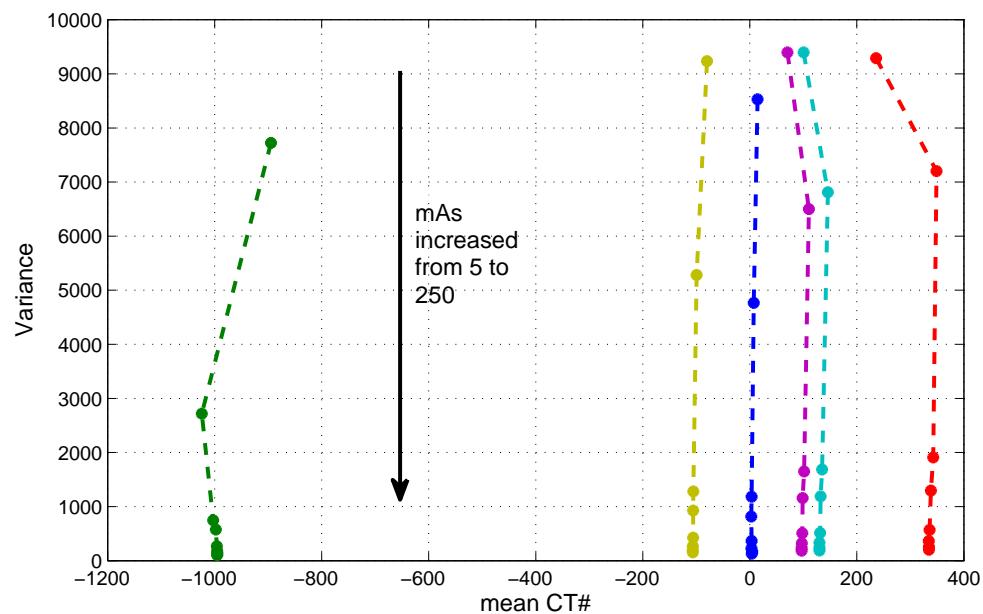


Figure A.4: Shows the noise variance changes for different CT#'s in the six regions of the phantom.

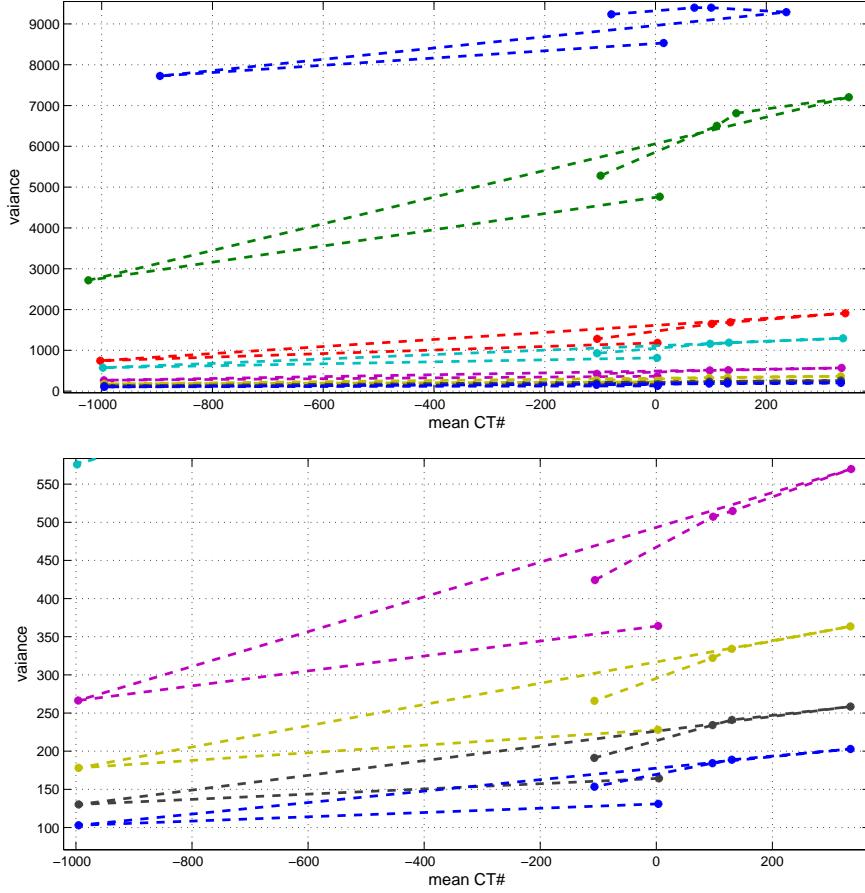


Figure A.5: Shows the noise variances of each region for different scan doses. Each connected line shows the result of a fixed dose (tube current). The figure at the bottom shows the 100, 150, 200, 250 mAs data which are compact in the upper figure. The region with CT# of 0 HU has smaller variances, since this is the CT# associated to water and the scanners are well calibrated for its variations.

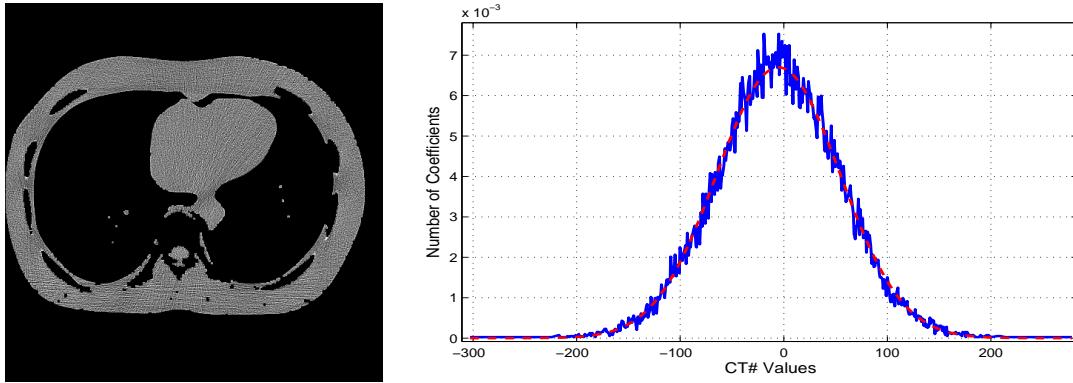


Figure A.6: Soft tissue region around the lung, which is used to evaluate the noise characteristics in the NCRE algorithm. The statistical distribution of the noise in this region is shown, with blue solid line, and is compared with Gaussian distribution with same variance, red dashed line.

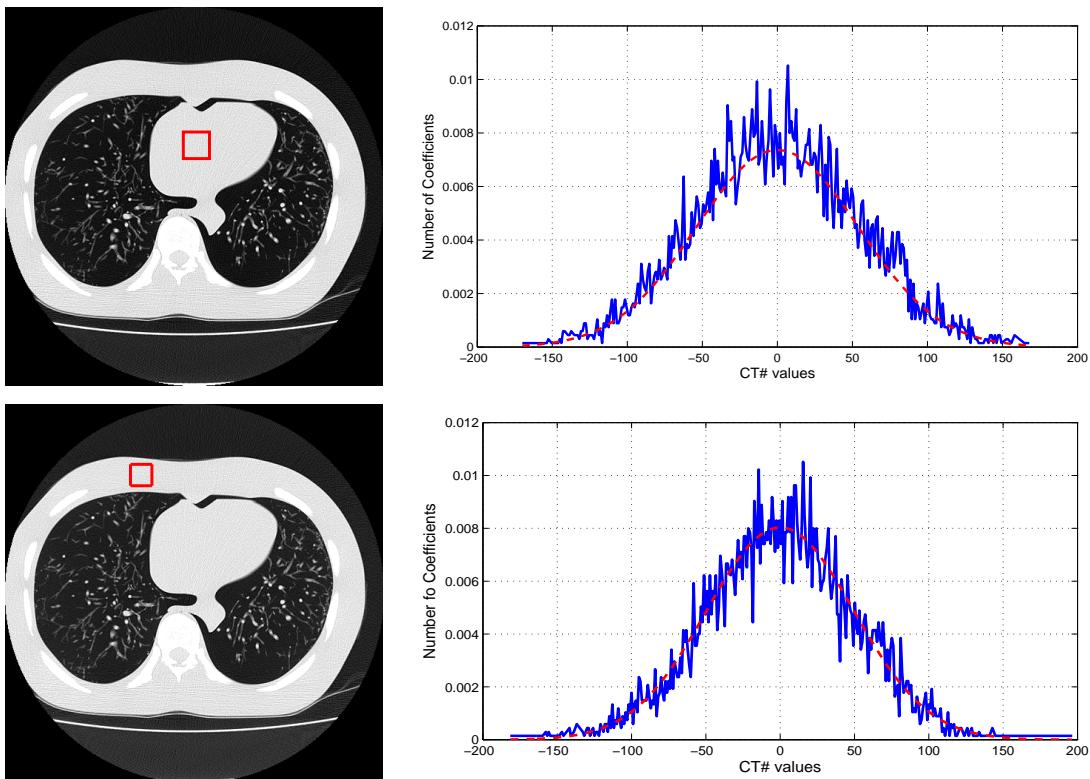


Figure A.7: Depicts the noise statistics of two small regions used in BM3D denoising and compares their noise distribution with that of similar Gaussian noises.

Appendix B

Measured CT number in Small Pulmonary Nodules

In this appendix the effect of slice thickness on the measured CT# of a small pulmonary nodule is studied. It is assumed that the nodule is spherical with a homogeneous attenuation and is surrounded by air with CT# of -1000 HU.

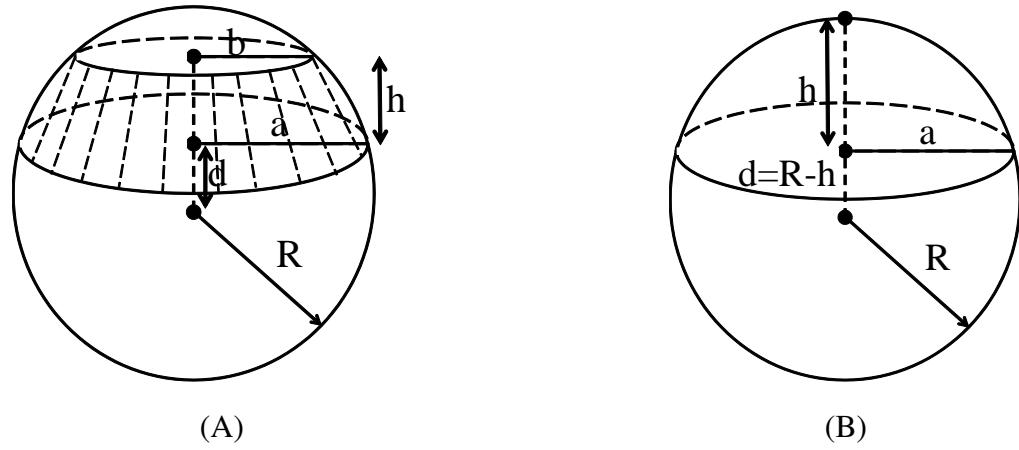


Figure B.1: (A) Spherical Segment, and (B) Spherical Cap.

B.1 Volume of Spherical Segment

The volume of the spherical segment of height h , shown in figure B.1-A, can be calculated from

$$\begin{aligned} V &= \pi \int_d^{d+h} (R^2 - y^2) dy \\ &= \pi h \left[R^2 - d^2 - hd - \frac{1}{3}h^2 \right], \end{aligned} \quad (\text{B.1})$$

where,

$$\begin{aligned} a &= \sqrt{R^2 - d^2} \\ b &= \sqrt{R^2 - (d+h)^2}. \end{aligned} \quad (\text{B.2})$$

So that (B.1) simplifies to

$$V = \frac{1}{6}\pi h (3a^2 + 3b^2 + h^2). \quad (\text{B.3})$$

B.2 Volume of Spherical Cap

When $b = 0$, similar to figure B.1-B, this volume becomes

$$V_{cap} = \frac{1}{3}\pi h^2 (3R - h). \quad (\text{B.4})$$

Therefore, the volume V_{seg} that includes the part of the nodule in the reconstructed slice with thickness of S can be calculated by

$$V_{seg} = \frac{2}{3}\pi R^3 - \frac{1}{3}\pi h^2 (3R - h). \quad (\text{B.5})$$

To measure the CT# of the nodule we choose the image slice with the largest cross section. If the slice thickness is thinner than the nodule radius R , the slice with the largest cross section is one of the cases shown in figure B.2, which includes two smaller volumes

$$V_1 = \frac{2}{3}\pi R^3 - \frac{1}{3}\pi(R-h_1)^2(3R-(R-h_1)) = \frac{1}{3}\pi(3R^2h_1 - h_1^3), \quad (\text{B.6})$$

and

$$V_2 = \frac{2}{3}\pi R^3 - \frac{1}{3}\pi(R-h_2)^2(3R-(R-h_2)) = \frac{1}{3}\pi(3R^2h_2 - h_2^3). \quad (\text{B.7})$$

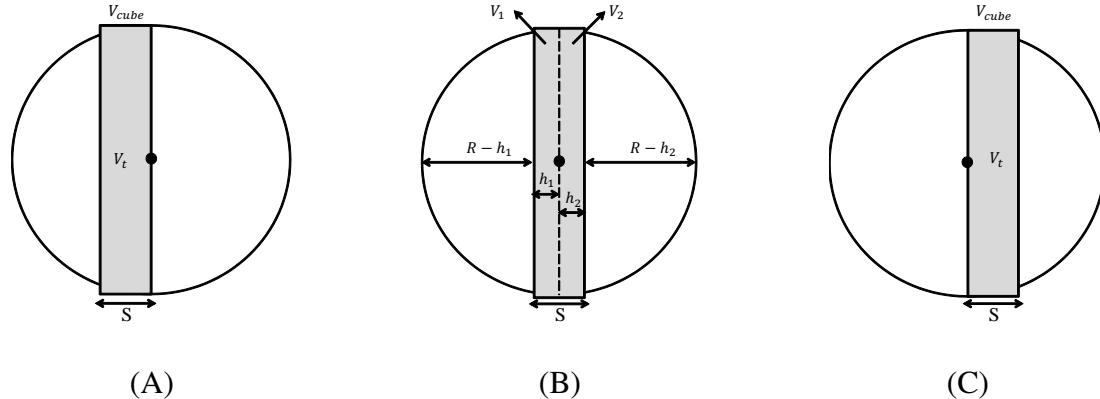


Figure B.2: Three cases with largest cross section of a nodule, when the slice thickness is less than the nodule radius.

Since $h_1 = S - h_2$, the overall volume can be calculated by:

$$\begin{aligned}
V_t &= V_1 + V_2 \\
&= \left(\frac{1}{3} \pi (3R^2(S - h_2) + (S - h_2)^3) \right) + \left(\frac{1}{3} \pi (3R^2h_2 + h_2^3) \right) \\
&= \pi \left(R^2S + S^2h_2 - Sh_2^2 - \frac{1}{3}S^3 \right). \tag{B.8}
\end{aligned}$$

Using equation (B.8), the average CT# inside the measurement slice with volume of $V_{cube} = 4R^2S$ would be:

$$\begin{aligned} CT_{mean} &= \frac{(V_{cube} - V_t)(-1000) + CT_{true}V_t}{V_{cube}} \\ &= -1000 + \frac{\pi}{4} \left(1 + \frac{Sh_2}{R^2} - \frac{h_2^2}{R^2} - \frac{S^2}{3R^2} \right) (CT_{true} + 1000), \end{aligned} \quad (\text{B.9})$$

in which CT_{true} is the nodule CT#. Assuming that h_2 is a random variable, which can be anything between $0 < h_2 < S$ with the same probability, i.e., it has a uniform distribution with probability distribution function of $pdf(h_2) = \frac{1}{S}$, the expected value of the measured CT# would be:

$$\begin{aligned} E(CT_{mean}) &= \int_0^S \frac{1}{S} \left(-1000 + \frac{\pi}{4} \left(1 + \frac{Sh_2}{R^2} - \frac{h_2^2}{R^2} - \frac{S^2}{3R^2} \right) (CT_{true} + 1000) \right) dh_2, \\ &= -1000 + \frac{\pi}{4} \left(1 - \frac{S^2}{6R^2} \right) (CT_{true} + 1000). \end{aligned} \quad (\text{B.10})$$

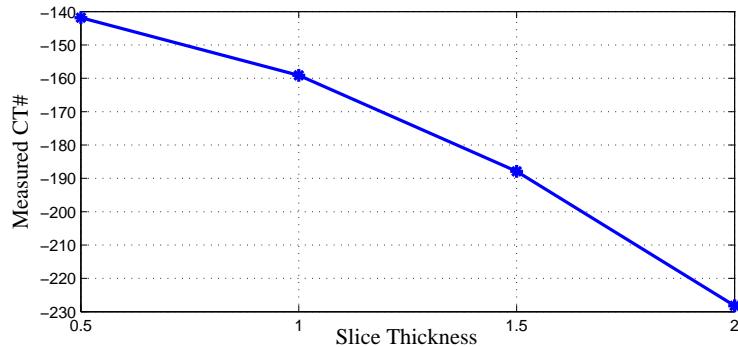


Figure B.3: CT# changes of a nodule with diameter of 5 mm and CT# of 100 HU, as a function of slice thickness.

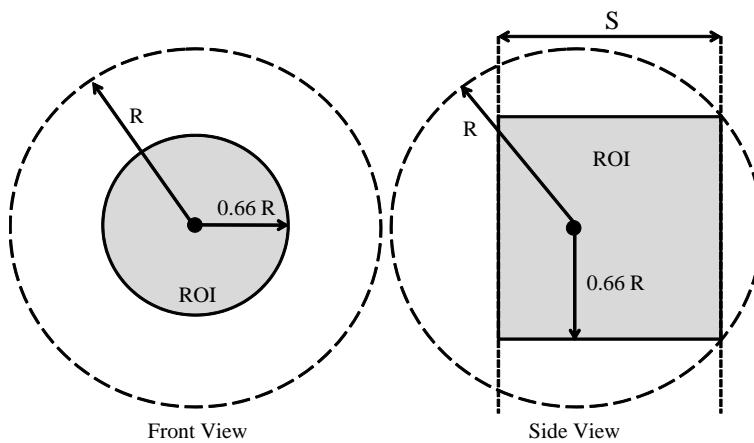


Figure B.4: A 66% ROI at the center of a nodule.

B.3 Importance of ROI in measurements

Figure B.3 shows the average measured CT# in four different slice thicknesses. As can be seen, the measured values are highly affected by slice thickness. To decrease this effect, the CT# is not measured from whole area, but only from a 66% ROI at the center of the nodule. This means that for slice thicknesses with $S < \sqrt{R^2 - (0.66R)^2} = 0.75R$ the averaging error is zero. Consequently, a nodule with diameter of 5 mm can be accurately characterized with a slice thickness smaller than 1.88 mm.

Considering a 66% ROI, as shown in figure B.4, few different cases would happen, which need to be calculated separately. To simplify the calculations, MATLAB simulations are used in chapter 4 to numerically calculate the reconstruction errors as a function of slice thickness.

References

- [1] D.J. Brenner and E.J. Hall. Computed tomography—an increasing source of radiation exposure. *The New England Journal of Medicine*, 357:2277–2284, Nov. 2007.
- [2] M.S. Pearce, J.A. Salott, M.P. Little, K. McHugh, C. Lee, K.P. Kim, N.L. Howe, C.M. Ronckers, P. Rajaraman, L. Parker A.W. Craft, and A.B. de Gonzálezl. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *The Lancet*, 380(9840):499–505, Aug. 2012.
- [3] S. Leng, L. Yu, and Y. Zhang. Correlation between model observer and human observer performance in CT imaging when lesion location is uncertain. *Medical Physics*, 40(8):081908, Aug. 2013.
- [4] K. Iniewski. *Medical Imaging: Principles, Detectors, and Electronics*. John Wiley & Sons, 2009.
- [5] B. Desjardins and E.A. Kazerooni. Review: ECG-gated cardiac CT. *American Journal of Roentgenology*, 128(4):993–1010, Apr. 2004.
- [6] J. Ming and W. Ge. Convergence of the simultaneous algebraic reconstruction technique (SART). *IEEE Transaction on Image Processing*, 12(8):957–961, Aug. 2003.
- [7] E. Sideky, Y. Duchin, X. Pan, and C. Ullberg. A constrained, total-variation minimization algorithm for low-intensity x-ray CT. *Medical Physics*, 38(7):S117–S125, Jul. 2011.
- [8] G.H. Chen, J. Tang, and S. Leng. Prior image constrained compressed sensing (piccs): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Medical Physics*, 35(2):660–663, Feb. 2008.
- [9] H. Lee, L. Xing, R. Davidi, R. Li, J. Qian, and R. Lee. Improved compressed sensing-based cone-beam CT reconstruction using adaptive prior image constraints. *Physics in Medicine and Biology*, 57(8):2287–2307, Feb. 2012.

- [10] F. Zhu, T. Carpenter, D. Rodriguez Gonzalez, M. Atkinson, and J. Wardlaw. Computed tomography perfusion imaging denoising using gaussian process regression. *Physics in Medicine and Biology*, 57(12):183–198, Jun. 2012.
- [11] S.A. Hyder and R. Sukanesh. An efficient algorithm for denoising MR and CT images using digital curvelet transform. *Advances in Experimental Medicine and Biology*, 696:471–480, Sep. 2011.
- [12] T. Strohmer and R. Vershynin. A randomized kaczmarz algorithm with exponential convergence. *Journal of Fourier Analysis and Applications*, 15(2):262–278, Apr. 2009.
- [13] G.T. Herman and L.B. Meyer. Algebraic reconstruction techniques can be made computationally efficient. *IEEE Transaction on Medical Imaging*, 12(3):600–609, Sep. 1993.
- [14] J. Fessler. Statistical image reconstruction methods for transmission tomography. In M. Sonka and J. Michael Fitzpatrick, editors, *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*. SPIE, Bellingham, 2000.
- [15] J. Browne and A.B. de Pierro. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Transaction on Medical Imaging*, 15(5):687–699, Oct. 1996.
- [16] K. Li, J. Tang, and G.H. Chen. Statistical model based iterative reconstruction (MBIR) in clinical CT systems: Experimental assessment of noise performance. *Medical Physics*, 41(4), Apr. 2014.
- [17] E.J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transaction on Information Theory*, 52(2):489–509, Feb. 2006.
- [18] D.L. Donoho. Compressed sensing. *IEEE Transaction on Information Theory*, 52(4):1289–1306, Apr. 2006.
- [19] E.Y. Sidky, C.M. Kao, and X. Pan. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *Journal of X-Ray Science and Technology*, 14(2):119–139, Jan. 2006.
- [20] J. Tang, B.E. Nett, and G.H.G Chen. Performance comparison between total variation TV-based compressed sensing and statistical iterative reconstruction algorithms. *Physics in Medicine and Biology*, 54(19):5781–5804, Oct. 2009.

- [21] K. Choi, J. Wang, L. Zhu, T.S. Suh, S. Boyd, and L. Xing. Compressed sensing based cone-beam computed tomography reconstruction with a first-order method. *Medical physics*, 37(9):5113–5125, Sep. 2010.
- [22] Z. Zhu, K. Wahid, P. Babyn, D. Cooper, I. Pratt, and Y. Carter. Improved compressed sensing-based algorithm for sparse-view CT image reconstruction. *Computational and Mathematical Methods in Medicine*.
- [23] J.C. Park, B. Song, J.S. Kim, S.H. Park, H.K. Kim, Z. Liu, T.S. Suh, and W.Y. Song. Fast compressed sensing-based CBCT reconstruction using barzilai-borwein formulation for application to on-line igt. *Medical physics*, 39(3):1207–1217, Mar. 2012.
- [24] Y. Junfeng, Z. Yin, and Y. Wotao. A fast alternating direction method for TVL1-L2 signal reconstruction from partial fourier data. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):288–297, Apr. 2010.
- [25] B.P. Fahimian, Y. Mao, P. Cloetens, and J. Miao. Low-dose x-ray phase-contrast and absorption CT using equally sloped tomography. *Physics in Medicine and Biology*, 55(18):5383–5400, Aug. 2010.
- [26] M. Jiang, J. Jin, F. Liu, Y. Yu, L. Xia, Y. Wang, and S. Crozier. Sparsity-constrained SENSE reconstruction: an efficient implementation using a fast composite splitting algorithm. *Magnetic Resonance Imaging*, 31(7):1218–1227, Sep. 2013.
- [27] J.M. Bioucas-Dias and M.A.T. Figueiredo. A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transaction on Image Processing*, 16(12):2992–3004, Dec. 2007.
- [28] T. Li, X. Li, J. Wang, J. Wen, H. Lu, J. Hsieh, and Z. Liang. Nonlinear sinogram smoothing for low-dose x-ray CT. *IEEE Transactions on Nuclear Science*, 51(5):2505–2513, Oct. 2004.
- [29] A.M. Mendrik, E.J. Vonken, A. Rutten, M.A. Viergever, and B. Van Ginneken. Noise reduction in computed tomography scans using 3-d anisotropic hybrid diffusion with continuous switch. *IEEE Transactions on Medical Imaging*, 28(10):1585–1594, Oct. 2009.
- [30] Y. Wang and H. Zhou. Total variation wavelet-based medical image denoising. *International Journal of Biomedical Imaging*, 2006.

- [31] F. Rashid-Farrokhi, K.J.R. Liu, C.A. Berenstein, and D. Walnut. Wavelet-based multiresolution local tomography. *IEEE Transactions on Image Processing*, 6(10):1412–1430, Oct. 1997.
- [32] T. Olson and J. DeStefano. Wavelet localization of the Radon transform. *IEEE Transactions on Signal Processing*, 42(8):2055–2067, Aug. 1994.
- [33] M. Holschneider. Inverse Radon transforms through inverse wavelet transforms. *Inverse Problems*, 7(6):853–861, 1991.
- [34] J. Deng, H. Yu, J. Ni, L. Wang, and G. Wang. Parallelism of iterative CT reconstruction based on local reconstruction algorithm. *The Journal of Supercomputing*, 48(1):1–14, Apr. 2009.
- [35] H. Yu, Y. Ye, S. Zhao, and G. Wang. Local ROI reconstruction via generalized FBP and BPF algorithms along more flexible curves. *International Journal of Biomedical Imaging*.
- [36] *International Journal of Biomedical Imaging*, 2007.
- [37] The nobel prize in physiology or medicine 1979. http://www.nobelprize.org/nobel_prizes/medicine/laureates/1979/.
- [38] J. Radon. über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. *Berichte der Sächsischen Akademie der Wissenschaft*, 69:262–277, 1917.
- [39] R. Mohan, A. Singh, and M. Gundappa. Three-dimensional imaging in periodontal diagnosis - utilization of cone beam computed tomography. *Journal of Indian Society of Periodontology*, 15(1):11–17, Jan. 2011.
- [40] W. A. Kalender. *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*. John Wiley & Sons, 2011.
- [41] L.W. Goldman. Principles of CT and CT technology. *Journal of Nuclear Medicine Technology*, 35:115–128, Sep. 2007.
- [42] S.R. Deans. *The Radon Transform and Some of Its Applications*. John Wiley & Sons, 1983.
- [43] J. Hsieh. *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*, volume PM188. SPIE Press Book, Bellingham, WA, 2nd edition, 2009.

- [44] P.E. Danielsson, M.M. Seger, and H. Turbell. The PI-methods for helical cone-beam tomography. 2002.
- [45] H.K. Tuy. An inversion formula for cone-beam reconstruction. *SIAM Journal on Applied Mathematics*, 43(3):546–552, 1983.
- [46] F. Noo, M. Defrise, and R. Clackdoyle. Single-slice rebinning method for helical cone-beam CT. *Physics in Medicine and Biology*, 44(2):561–570, Feb. 1999.
- [47] J. Sunnegårdh. Iterative enhancement of non-exact reconstruction in cone beam CT. Master’s thesis, Linköping University, Sep. 2004.
- [48] D. Gottleib, B. Gustafsson, and P. Forssen. On the direct fourier method for computer tomography. *IEEE Transaction on Medical Imaging*, 19(3):223–232, Mar. 2000.
- [49] H. Stark, J. Woods, I. Paul, and R. Hingorani. Direct fourier reconstruction in computer tomography. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 29(2):237–245, Apr. 1981.
- [50] H. Schomberg and J. Timmer. The gridding method for image reconstruction by fourier transformation. *IEEE Transactions on Medical Imaging*, 14(3):596–607, Sep. 1995.
- [51] Y. Mao, B.P. Fahimian, S.J. Osher, and J. Miao. Development and optimization of regularized tomographic reconstruction algorithms utilizing equally-sloped tomography. *IEEE Transaction on Image Processing*, 19(5):1259–1268, May 2010.
- [52] J. Miao, F. Förster, and O. Levi. Equally sloped tomography with oversampling reconstruction. *Physical Review B*, 72(5):052103, Aug. 2005.
- [53] E. Lee, B.P. Fahimian, C.V. Iancu, C. Suloway, G.E. Murphy, E.R. Wright, D. Castano-Diez, G.J. Jensen, and J. Miao. Radiation dose reduction and image enhancement in biological imaging through equally-sloped tomography. *Journal of Structural Biology*, 164(2):221–227, Nov. 2008.
- [54] A. Averbuch, R.R. Coifmanb, D.L. Donoho, M. Elad, and M. Israeli. Fast and accurate polar fourier transform. *Applied and Computational Harmonic Analysis*, 21:145–167, Sep. 2006.
- [55] G. Besson. CT image reconstruction from fan-parallel data. *Medical Physics*, 26:415–426, Dec. 1999.

- [56] A.K. Hara, R.G. Paden, A.C. Silva, J.L. Kujak, H.J. Lawder, and W. Pavlicek. Iterative reconstruction technique for reducing body radiation dose at CT: feasibility study. *American Journal of Roentgenology*, 193(3):764–771, Sep. 2009.
- [57] H. Dang, A.S. Wang, Z. Zhao, M.S. Sussman, J.H. Siewerdsen, and J.W. Stayman. Joint estimation of deformation and penalized-likelihood CT reconstruction using previously acquired images. In *International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, pages 424–427, 2013.
- [58] I.A. Elbakri and A.J. Fessler. Statistical x-ray-computed tomography image reconstruction with beam-hardening correction. In *Medical Imaging 2001*, pages 1–12. International Society for Optics and Photonics, 2001.
- [59] H.M. Hudson and R.S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609, Dec. 1994.
- [60] D. Hart, D.G. Jones, and B.F. Wall. Estimation of effective dose in diagnostic radiology from entrance surface dose and dose-area product measurements. Technical report, National Radiological Protection Board, Chilton, United Kingdom, 1994.
- [61] C.H. McCollough, A.N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner. Strategies for reducing radiation dose in CT. *Radiologic Clinics of North America*, 47(1):27–40, Jan. 2009.
- [62] D.H. Yang and H.W. Goo. Pediatric 16-slice CT protocols: radiation dose and image quality. *Journal of the Korean Radiological Society*, 59(5):333–347, 2008.
- [63] F. Tatsugami, L. Husmann, B.A. Herzog, N. Burkhard, I. Valenta, O. Gaemperli, and P.A. Kaufmann. Evaluation of a body mass index–adapted protocol for low-dose 64-MDCT coronary angiography with prospective ECG triggering. *American Journal of Roentgenology*, 192(3):635–638, Mar. 2009.
- [64] N.S. Paul, H. Kashani, D. Odedra, A. Ursani, C. Ray, and P. Rogalla. The influence of chest wall tissue composition in determining image noise during cardiac CT. *American Journal of Roentgenology*, 197(6):1328–1334, Dec. 2011.
- [65] M.K. Kalra, M.M. Maher, S.R. Prasad, M.S. Hayat, M.A. Blake, J. Varghese, E.F. Halpern, and S. Saini. Correlation of patient weight and cross-sectional dimensions with subjective image quality at standard dose abdominal CT. *Korean Journal of Radiology*, 4(4):234–238, Oct-Dec. 2003.

- [66] Y.Y. Jung and H.W. Goo. The optimal parameter for radiation dose in pediatric low dose abdominal CT: cross-sectional dimensions versus body weight. *Journal of the Korean Radiological Society*, 58(2):169–175, 2008.
- [67] U. Nyman, T.L. Ahl, M. Kristiansson, L. Nilsson, and S. Wettemark. Patient circumference adapted dose regulation in body computed tomography. a practical and flexible formula. *Acta Radiologica*, 46(4):396–406, Jul. 2005.
- [68] D.J. Brenner and M.A. Georgsson. Mass screening with CT colonography: should the radiation exposure be of concern? *Gastroenterology*, 129(1):328–337, Jul. 2005.
- [69] C.H. McCollough, M.R. Bruesewitz, and J.M. Kofler. CT dose reduction and dose management tools: Overview of available options 1. *Radiographics*, 26(2):503–512, Mar-Apr. 2006.
- [70] H. Greess, H. Wolf, U. Baum, M. Lell, M. Pirkl, W. Kalender, and W.A.Bautz. Dose reduction in computed tomography by attenuation-based on-line modulation of tube current: evaluation of six anatomical regions. *European Radiology*, 10(2):391–394, Jan. 2000.
- [71] H.W. Goo and D.S. Suh. Tube current reduction in pediatric non-ECG-gated heart CT by combined tube current modulation. *Pediatric radiology*, 36(4):344–351, Apr. 2006.
- [72] L. Yu, M.R. Bruesewitz, K.B. Thomas, J.G. Fletcher, J.M. Kofler, and C.H. McCollough. Optimal tube potential for radiation dose reduction in pediatric CT: principles, clinical implementations, and pitfalls. *Radiographics*, 31(3):835–848, 2011.
- [73] W.A. Kalender, S. Buchenau, P. Deak, M. Kellermeier, O. Langner, M. van Straten, S. Vollmar, and S. Wilharm. Technical approaches to the optimisation of CT. *Physica Medica*, 24(2):71–79, Jun. 2008.
- [74] X. Pan, E.Y. Sidky, and M. Vannier. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Problems*, 25(12):123009, Jan. 2009.
- [75] H. Kudo, T. Suzuki, and E.A. Rashed. Image reconstruction for sparse-view CT and interior CT-introduction to compressed sensing and differentiated backprojection. *Quantitative Imaging in Medicine and Surgery*, 3(3):147, Jun. 2013.
- [76] M.A. Davenport, M.F. Duarte, Y.C. Eldar, and G. Kutyniok. *Compressed Sensing: Theory and Applications*, chapter Introduction to Compressed Sensing. Cambridge University Press, 2012.

- [77] A.M. Bruckstein, D.L. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, Feb. 2009.
- [78] J.F. Claerbout and F. Muir. Robust modeling with erratic data. *Geophysics*, 38(5):826–844, Oct. 1973.
- [79] F. Santosa and W.W. Symes. Linear inversion of band-limited reflection seismograms. *SIAM Journal on Scientific and Statistical Computing*, 7(4):1307–1330, Oct. 1986.
- [80] Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, Nov. 1992.
- [81] P. Garrigues and L.E. Ghaoui. An homotopy algorithm for the lasso with online observations. In *Advances in Neural Information Processing Systems*, pages 489–496, 2009.
- [82] D. Angelosante and G.B. Giannakis. RLS-weighted lasso for adaptive estimation of sparse signals. In *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2009*, pages 3245–3248, Apr. 2009.
- [83] Y.C. Eldar, editor. *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.
- [84] L. Dettori and L. Semler. A comparison of wavelet, ridgelet, and curvelet-based texture classification algorithms in computed tomography. *Computers in Biology and Medicine*, 37(4):486–498, Apr. 2007.
- [85] K. Guo and D. Labate. Optimally sparse multidimensional representation using shearlets. *SIAM Journal on Mathematical Analysis*, 39(1):298–318, May 2007.
- [86] R. Coifman, F. Geshwind, and Y. Meyer. Noiselets. *Applied and Computational Harmonic Analysis*, 10:27–44, Jan. 2001.
- [87] X.M. Lu, J.M. Reid, and S.H. Wang. A practical application of ℓ_1 deconvolution for medical ultrasound. In *IEEE Ultrasonics Symposium*, volume 2, pages 1177–1180, Dec. 1991.
- [88] A.Y. Ng. Feature selection, ℓ_1 vs. ℓ_2 regularization, and rotational invariance. In *21 International Conference on Machine Learning, Banff, Canada*, pages 78–86, 2004.
- [89] E.J. Candès and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, Dec. 2005.

- [90] E.J. Candès and M.B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, Mar. 2008.
- [91] E.J. Candès and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse Problems*, 23(3):969–985, Apr. 2007.
- [92] S. Becker, E.J. Candès, and M. Grant. Templates for convex cone problems with applications to sparse signal recovery. *Mathematical Programming Computation*, 3(3), Aug. 2011.
- [93] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>, Mar. 2014.
- [94] M. Asif and J. Romberg. Sparse recovery of streaming signals using ℓ_1 homotopy. *IEEE Transactions on Signal Processing*, 62(16):4209–4223, Aug. 2014.
- [95] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):182–202, Mar. 2009.
- [96] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
- [97] I. Bayram and I.W. Selesnick. A subband adaptive iterative shrinkage/thresholding algorithm. *IEEE Transactions on Signal Processing*, 58(3):1131–1143, Mar. 2010.
- [98] J.F. Yang and Y. Zhang. Alternating direction algorithms for ℓ_1 -problems in compressive sensing. *SIAM Journal on Scientific Computing*, 33(1):250–278, Feb. 2011.
- [99] T. Goldstein and S. Osher. The split bregman method for ℓ_1 -regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, Apr. 2009.
- [100] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for ℓ_1 -minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, 1(1):143–168, 2008.
- [101] S.G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, Dec. 1993.
- [102] T. Blumensath and M.E. Davies. Iterative hard thresholding for compressed sensing. *Applied and Computational Harmonic Analysis*, 27(3):265–274, Nov. 2009.
- [103] T. Blumensath. Accelerated iterative hard thresholding. *Signal Processing*, 92(3):752–756, Mar. 2012.

- [104] J.A. Tropp and A.C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, Dec. 2007.
- [105] D.L. Donoho, Y. Tsaig, I. Drori, and J.L. Starck. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 58(2):1094–1121, Feb. 2012.
- [106] D. Needell and J.A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, May 2009.
- [107] E.J. Candès, M.B. Wakin, and S. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, Dec. 2008.
- [108] R. Chartrand and W. Yin. Iteratively reweighted algorithms for compressive sensing. In *33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 3869–3872, Mar. 2008.
- [109] Keijo Hämäläinen, Lauri Harhanen, Andreas Hauptmann, Aki Kallonen, Esa Niemi, and Samuli Siltanen. Total variation regularization for large-scale X-ray tomography. *International Journal of Tomography & Simulation*, 25(1):1–25, 2014.
- [110] M.A. Figueiredo and R.D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Transaction on Image Processing*, 12(8):906–916, Aug. 2003.
- [111] B.P. Fahimian, Y. Zhao, Z. Huang, R. Fung, Y. Mao, C. Zhu, M. Khatonabadi, J.J. DeMarco, S.J. Osher, M.F. McNitt-Gray, and J. Miao. Radiation dose reduction in medical x-ray CT via fourier-based iterative reconstruction. *Medical Physics*, 40(3):031914–1–031914–10, Mar. 2013.
- [112] J.C. Ramirez-Giraldo, J. Trzasko, S. Leng, L. Yu, A. Manduca, and C.H. McCollough. Nonconvex prior image constrained compressed sensing (NCPICCS): Theory and simulations on perfusion CT. *Medical physics*, 38(4):2157–2167, Apr. 2011.
- [113] J. Zhang, J. Wang, H. Zuo, G. Xu, and J.B. Thibault. Compressed sensing algorithms for fan-beam computed tomography image reconstruction. *Optical Engineering*, 51(7):071402–1, May 2012.
- [114] Y. Kaganovsky, D. Li, A. Holmgren, H. Jeon, K.P. MacCabe, D.G. Politte, J.A. O’Sullivan, L. Carin, and D.J. Brady. Compressed sampling strategies for tomography. *Journal of the Optical Society of America*, 31(7):1369–1394, Jul. 2014.

- [115] I.Y. Chun, B. Adcock, and T. Talavage. Efficient compressed sensing SENSE parallel MRI reconstruction with joint sparsity promotion and mutual incoherence enhancement. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC14)*, pages 1–4, Aug. 2014.
- [116] M. Lustig, D.L. Donoho, J.M. Santos, and J.M. Pauly. Compressed sensing mri. *IEEE Signal Processing Magazine*, 25(2):72–82, Mar. 2008.
- [117] K. Choi and D.J. Brady. Coded aperture computed tomography. In *Procideenig SPIE, Adaptive Coded Aperture Imaging, Non-Imaging, and Unconventional Imaging Sensor Systems*, volume 7468, page 74680B, 2009.
- [118] A. Jemal, R. Siegel, J. Xu, and E. Ward. Cancer statistics, 2010. *CA: A Cancer Journal for Clinicians*, 60(5):277–300, Sep-Oct. 2010.
- [119] C.J. Beadsmoore and N.J. Screamton. Classification, staging and prognosis of lung cancer. *European Journal of Radiology*, 45(1):8–17, Jan. 2003.
- [120] Survival of patients with stage I lung cancer detected on CT screening. *New England Journal of Medicine*, 355(17):1763–1771, Oct. 2006.
- [121] H. MacMahon, J.H. Austin, G. Gamsu C.J. Herold, J.R. Jett, D.P. Naidich, E.F. Patz, and S.J. Swensen. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner society. *Radiology*, 237(2):395–400, Nov. 2005.
- [122] C.I. Henschke, D.I. McCauley, D.F. Yankelevitz, D.P. Naidich, G. McGuinness, O.S. Miettinen, D.M. Libby, M.W. Pasmanier, J. Koizumi, N.K. Altorki, and J.P. Smith. Early lung cancer action project: overall design and findings from baseline screening. *Lancet*, 354(9173):99–105, Jul. 1999.
- [123] Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 365(5):395–409, Aug. 2011.
- [124] M.M. Goodsitt, H.P. Chan, T.W. Way, S.C. Larson, E.G. Christodoulou, and J. Kim. Accuracy of the CT numbers of simulated lung nodules imaged with multi-detector CT scanners. *Medical Physics*, 33(8):3006–3017, Aug. 2006.
- [125] J.G. Ravenel, W.M. Leue, P.J. Nietert, J.V. Miller, K.K. Taylor, and G.A. Silvestri. Pulmonary nodule volume: effects of reconstruction parameters on automated measurements - a phantom study. *Radiology*, 247(2):400–408, May 2008.

- [126] L. Yu, X. Liu, S. Leng, j.M. Kofler, J.C. Ramirez-Giraldo, M. Qu, J. Christner, J.G. Fletcher, and C.H. McCollough. Radiation dose reduction in computed tomography: techniques and future perspective. *Imaginig in Medicine*, 1(1):65–84, Oct. 2009.
- [127] C.I. Henschke, D.F. Yankelevitz, R. Mirtcheva, G. McGuinness, D. McCauley, and O.S. Miettinen. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *American Journal of Roentgenology*, 178(5):1053–1057, May 2002.
- [128] J.M. Goo, C.M. Park, and H.J. Lee. Ground-glass nodules on chest CT as imaging biomarkers in the management of lung adenocarcinoma. *American Journal of Roentgenology*, 196(3):533–543, Mar. 2011.
- [129] D.M. Xu, R.J. van Klaveren, G.H. de Bock, A.L.M. Leusveld, M.D. Dorrius, Yi Zhao, Y. Wang, H.J. de Koning, E.T. Scholten, J. Verschakelen, M. Prokop, and M. Oudkerk. Role of baseline nodule density and changes in density and nodule features in the discrimination between benign and malignant solid indeterminate pulmonary nodules. *European Journal of Radiology*, 70(3):492–498, Jun. 2009.
- [130] F. Fischbach, F. Knollmann, V. Griesshaber, T. Freund, E. Akkol, and R. Felix. Detection of pulmonary nodules by multislice computed tomography: improved detection rate with reduced slice thickness. *European Radiology*, 13(10):2378–2383, Oct. 2003.
- [131] M. Petrou, L.E. Quint B. Nan, and L.H. Baker. Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. *American Journal of Roentgenology*, 188(2):306–312, Feb. 2007.
- [132] S. Diederich, M.G. Lentschig, F. Winter, N. Roos, and G. Bongartz. Detection of pulmonary nodules with overlapping vs non-overlapping image reconstruction at spiral CT. *European Radiology*, 9(2):281–286, 1999.
- [133] M.A. Gavrielides, R. Zeng, K.J. Myers, B. Sahiner, and N. Petrick. Benefit of overlapping reconstruction for improving the quantitative assessment of CT lung nodule volume. *Academic Radiology*, 20(2):173–180, Feb. 2013.
- [134] O. Honda, H. Sumikawa, T. Johkoh, N. Tomiyama, N. Mihara, A. Inoue, M. Tsubamoto, J. Natsag, S. Hamada, and H. Nakamura. Computer-assisted lung nodule volumetry from multi-detector row CT: Influence of image reconstruction parameters. *European Journal of Radiology*, 62(1):106–113, Apr. 2007.
- [135] M. Katsura, I. Matsuda, M. Akahane, J. Sato, H. Akai, K. Yasaka, A. Kunimatsu, and K. Ohtomo. Model-based iterative reconstruction technique for radiation dose reduction

- in chest CT: comparison with the adaptive statistical iterative reconstruction technique. *European Radiology*, 22(8):1613–1623, Aug. 2012.
- [136] P. Therasse, S.G. Arbuck, E.A. Eisenhauer, J. Wanders, R.S. Kaplan, L. Rubinstein, J. Verweij, M. Van Glabbeke, A.T. van Oosterom, M.C. Christian, and S.G. Gwyther. New guidelines to evaluate the response to treatment in solid tumors. European organization for research and treatment of cancer, national cancer institute of the United States, national cancer institute of Canada. *Journal of the National Cancer Institute*, 92(3):205–216, Feb. 2000.
- [137] B.M. Gramer, D. Muenzel, V. Leber, A.K. von Thaden, H. Feussner, A. Schneider, M. Vembar, N. Soni, E.J. Rummeny, and A.M. Huber. Impact of iterative reconstruction on CNR and SNR in dynamic myocardial perfusion imaging in an animal model. *European Radiology*, 22(12):2654–2661, Dec. 2012.
- [138] Q. Li, H. Yu, L. Zhang, L. Fan, and S.Y. Liu. Combining low tube voltage and iterative reconstruction for contrast-enhanced CT imaging of the chest-initial clinical experience. *Clinical Radiology*, 68(5):e249–e253, May 2013.
- [139] M. Yanagawa, Y. Tanaka, M. Kusumoto, S. Watanabe, R. Tsuchiya, O. Honda, H. Sumikawa, A. Inoue, M. Inoue, M. Okumura, N. Tomiyama, and T. Johkoh. Automated assessment of malignant degree of small peripheral adenocarcinomas using volumetric CT data: Correlation with pathologic prognostic factors. *Lung Cancer*, 70(3):286–294, Dec. 2010.
- [140] B.R. Whiting. Signal statistics in X-ray computed tomography. In L.E. Antonuk and M.Joel Yaffe, editors, *Proc. SPIE 4682, Medical Imaging 2002: Physics of Medical Imaging*, volume 4682. SPIE Proceedings, May 2002.
- [141] K. Knight. *Mathematical Statistics*. Chapman and Hall, New York, 2000.
- [142] Multipurpose chest phantom n1. kyoto kagaku co. <http://www.kyotokagaku.com/products/detail03/ph-1.html>.
- [143] N.S. Paul, J. Blobel, E. Prezelj, P. Burey, A. Ursani, R.J. Menezes H. Kashani, and J.H. Siewerdsen. The reduction of image noise and streak artifact in the thoracic inlet during low dose and ultra-low dose thoracic CT. *Physics in Medicine and Biology*, 55(5):1363–1380, Mar. 2010.

- [144] S.T. Schindera, D. Odedra, S.A. Raza, T.K. Kim, H.J. Jang, Z. Szucs-Farkas, and P. Rogalla. Iterative reconstruction algorithm for CT: Can radiation dose be decreased while low-contrast detectability is preserved? *Radiology*, 269(2):511–518, Nov. 2013.
- [145] S. Iwano, N. Makino, M. Ikeda, S. Itoh, M. Tadokoro, H. Satake, and T. Ishigaki. Solitary pulmonary nodules: Optimal slice thickness of high-resolution CT in differentiating malignant from benign. *Clinical Imaging*, 28(5):322–328, Sep-Oct. 2004.
- [146] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang. Low-dose x-ray CT reconstruction via dictionary learning. *IEEE Transactions on Medical Imaging*, 31(9):1682–1697, Sep. 2012.
- [147] S. Ramani and J.A. Fessler. A splitting-based iterative algorithm for accelerated statistical x-ray CT reconstruction. *IEEE Transactions on Medical Imaging*, 31(3):677–688, Mar. 2012.
- [148] J. Tang, B.E. Nett, and G.H. Chen. Performance comparison between total variation (tv)-based compressed sensing and statistical iterative reconstruction algorithms. *Physics in Medicine and Biology*, 54(19):5781, 2009.
- [149] E. J. Candès, M. Wakin, and S. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14:877–905, Oct. 2008.
- [150] M.A. Khajehnejad, X. Weiyu, A.S. Avestimehr, and B. Hassibi. Analyzing weighted ℓ_1 minimization for sparse recovery with nonuniform sparse models. *IEEE Transactions on Signal Processing*, 59(5):1985–2001, May 2011.
- [151] N. Vaswani and L. Wei. Modified-CS: Modifying compressive sensing for problems with partially known support. In *IEEE International Symposium on Information Theory, ISIT 2009.*, pages 488–492, 2009.
- [152] M. Hashemi, S. Beheshti, P.R. Gill, N.S. Paul, and R.S.C. Cobbold. Fast fan/parallel beam CS-based low-dose CT reconstruction. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, May 2013.
- [153] J.B. Thibault, D.K. Sauer, C.A. Bouman, and J. Hsieh. A three-dimensional statistical approach to improved image quality for multislice helical CT. *Medical Physics*, 34(11):4526–4544, Aug. 2007.
- [154] S. Maeda, W. Fukuda, K. Atsunori, and S. Ishii. Maximum a posteriori X-ray computed tomography using graph cuts. In *Journal of Physics: Conference Series, International*

- Workshop on Statistical-Mechanical Informatics 2010*, volume 233, pages 4526–4544, 2010.
- [155] M. Hashemi and S. Beheshti. Adaptive bayesian denoising for general gaussian distributed signals. *IEEE Transactions on Signal Processing*, 62(5):1147–1156, Mar. 2014.
- [156] M. Basseville, A. Benveniste, K.C. Chou, S.A. Golden, R. Nikoukhah, and A.S. Willsky. Modeling and estimation of multiresolution stochastic processes. *IEEE Transaction on Information Theory*, 38(2):766–784, Mar. 1992.
- [157] M. Lassas and S. Siltanen. Can one use total variation prior for edge-preserving bayesian inversion? *Inverse Problems*, 20(5):1537–1563, Aug. 2004.
- [158] T. Knopp, S. Kunis, and D. Potts. A note on the iterative MRI reconstruction from nonuniform k-space data. *International journal of biomedical imaging*, 2007, 2007.
- [159] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, Mar. 2009.
- [160] J. Huang, S. Zhang, H. Li, and D. Metaxas. Composite splitting algorithms for convex optimization. *Computer Vision and Image Understanding*, 115(12):1610–1622, Dec. 2011.
- [161] Q. Kun and A. Dogandzic. Sparse signal reconstruction via ECME hard thresholding. *IEEE Transactions on Signal Processing*, 60(9):4551–4569, Sep. 2012.
- [162] M.A. Figueiredo and R.D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Transaction on Image Processing*, 12(8):906–916, Aug. 2003.
- [163] L. Scharf and C. Demeure. *Statistical signal processing: detection, estimation, and time series analysis*. Addison-Wesley Pub. Co., Reading, MA, 1991.
- [164] M.A.T. Figueiredo and R.D. Nowak. Wavelet-based image estimation: an empirical bayes approach using Jeffrey’s noninformative prior. *IEEE Transaction on Image Processing*, 10(9):1322–1331, Sep. 2001.
- [165] E.Y. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in medicine and biology*, 53(17):4777, 2008.

- [166] L. Ritschl, F. Bergner, and C. Fleischmann and M. Kachelrieß. Improved total variation-based CT image reconstruction applied to clinical data. *Physics in medicine and biology*, 56(6):1545, 2011.
- [167] J. Bian, J.H. Siewerdsen, X. Han, E.Y. Sidky, J.L. Prince, C.A. Pelizzari, and X. Pan. Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. *Physics in medicine and biology*, 55(22):6575, 2010.
- [168] T. Strohmer and R. Vershynin. A randomized kaczmarz algorithm with exponential convergence. *Journal of Fourier Analysis and Applications*, 15(2):262–278, 2009.
- [169] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.
- [170] K. Chen, E.L. Piccolomini, and F. Zama. An automatic regularization parameter selection algorithm in the total variation model for image deblurring. *Numerical Algorithms*, pages 1–20, Oct. 2013.
- [171] J.M. Pauly M. Lustig, D. Donoho. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58:1182–1195, Dec. 2007.
- [172] S. Osher, A. Sole, and L. Vese. Image decomposition and restoration using total variation minimization and the H^1 . *Multiscale Modeling & Simulation*, 1(3):349–370, 2003.
- [173] H. Liao, F. Li, and M.K. Ng. Selection of regularization parameter in total variation image restoration. *Journal of the Optical Society of America A*, 26(11):2311–2320, Nov. 2009.
- [174] G. M. Vainikko. The discrepancy principle for a class of regularization methods. *USSR computational mathematics and mathematical physics*, 22(3):1–19, 1982.
- [175] P.C. Hansen. Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34(4):561–580, 1992.
- [176] E.J. Candès and Y. Plan. Near-ideal model selection by ℓ_1 minimization. *The Annals of Statistics*, 37(5A):2145–2177, 2009.
- [177] S. Akkoul, R. Harba, and R. Lédée. An image dependent stopping method for iterative denoising procedures. *Multidimensional Systems and Signal Processing*, pages 1–10, Jul. 2012.

- [178] Y. Wang and H.M. Zhou. Total variation wavelet based medical image denoising. *International Journal of Biomedical Imaging*.
- [179] A. Mousavi, A. Maleki, and R.G. Baraniuk. Parameterless optimal approximate message passing. *arXiv:1311.0035 [cs.IT]*, 2014.
- [180] S. Beheshti, M. Hashemi, X.P. Zhang, and N. Nikvand. Noise invalidation denoising. *IEEE Transactions on Signal Processing*, 58(12):6007–6016, Dec. 2010.
- [181] S.G. Chang, Y. Bin, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing*, 9(9):1532–1546, Sep. 2000.
- [182] Owe Axelsson. *Iterative Solution Methods*. Cambridge University Press.
- [183] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81(1):425–455, 1994.
- [184] D. L. Donoho and I.M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90, May 1995.
- [185] S. Voronin and R. Chartrand. A new generalized thresholding algorithm for inverse problems with sparsity constraints. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013*, pages 1636–1640, 2013.
- [186] T. Lei and W. Sewchand. Statistical approach to x-ray CT imaging and its applications in image analysis. i. statistical analysis of x-ray CT imaging. *IEEE Transactions on Medical Imaging*, 11(1):53–61, Mar. 1992.
- [187] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, Aug. 2007.
- [188] P. Chatterjee and P. Milanfar. Patch-based near-optimal image denoising. *IEEE Transactions on Image Processing*, 21(4):1635–1649, Apr. 2012.
- [189] D.H. Trinh, M. Luong, J.M. Rocchisani, C.D. Pham, H.D. Pham, and F. Dibos. An optimal weight method for CT image denoising. *JOURNAL OF ELECTRONIC SCIENCE AND TECHNOLOGY*, 10(2):124–129, Jun. 2012.
- [190] M. Guerquin-Kern, M. Haberlin, K.P. Pruessmann, and M. Unser. A fast wavelet-based reconstruction method for magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 30(9):1649–1660, Sep. 2011.

- [191] A. Averbuch, I. Sedelnikov, and Y. Shkolnisky. CT reconstruction from parallel and fan-beam projections by a 2-d discrete radon transform. *IEEE Transaction on Image Processing*, 21(2):733–741, Feb. 2012.
- [192] D.L. Donoho and A.G. Flesia. Digital ridgelet transform based on true ridge functions. In C.K. Chui, P. Monk, and L. Wuytack, editors, *Beyond Wavelets*, volume 10 of *Studies in Computational Mathematics*, pages 1–30. Elsevier, 2003.
- [193] E. Quaia, E. Baratella, S. Cernic, A. Lorusso, F. Casagrande, V. Cioffi, and M.A. Cova. Analysis of the impact of digital tomosynthesis on the radiological investigation of patients with suspected pulmonary lesions on chest radiography. *European Radiology*, 22(9):1912–1922, Sep. 2012.
- [194] F.A. Mettler, W. Huda, T.T. Yoshizumi, and M. Mahesh. Effective doses in radiology and diagnostic nuclear medicine: A catalog. *Radiology*, 248(1):254–263, Jul. 2008.
- [195] K. Bacher, P. Smeets, K. Bonnarens, A. De Hauwere, K. Verstraete, and H. Thierens. Dose reduction in patients undergoing chest imaging: digital amorphous silicon flat-panel detector radiography versus conventional film-screen radiography and phosphor-based computed radiography. *American journal of roentgenology*, 181(4):923–929, Oct. 2003.
- [196] 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. *Medical Physics*, 41(6), Jun. 2014.
- [197] C.M. Florkowski. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: Communicating the performance of diagnostic tests. *The Clinical Biochemist Reviews*, 29:S83–S87, Aug. 2008.
- [198] R. Proksa, T. Kohler, M. Grass, and J. Timmer. The n-pi-method for helical cone-beam CT. *IEEE Transactions on Medical Imaging*, 19(9):848–863, Sep. 2000.
- [199] S.W. Lee, G. Cho, and G. Wang. Artifacts associated with implementation of the grangeat formula. *Medical physics*, 29(12):2871–2880, Dec. 2002.
- [200] Johan Sunnegårdh. *Iterative Enhancement of Non-Exact Reconstruction in Cone Beam CT*. PhD thesis, Department of Electrical Engineering, Computer Vision, Linköping University, The Institute of Technology, 2004.
- [201] P.M. Joseph and R.D. Spital. The effects of scatter in X-ray computed tomography. *Medical physics*, 9(4):464–472, Jul-Aug. 1982.

- [202] J. Miao, F. Huang, and D.L. Wilson. Investigation on compressed sensing regularization parameter using case-pdm. In *Proceeding of the 19 th Annual Meeting of ISMRM. Montréal, Canada*, 2011.
- [203] D. Bartuschat, A. Borsdorf, H. Kostler, R. Rubinstein, and M. Sturmer. A parallel k-svd implementation for CT image denoising. Technical report, Technical Report - University of Erlangen-Nurnberg, 2009.
- [204] J. Hsieh. Adaptive streak artifact reduction in computed tomography resulting from excessive X-ray photon noise. *Medical Physics*, 25(11):2139–2147, Nov. 1998.
- [205] H. Rabbani, R. Nezafat, and S. Gazor. Wavelet-domain medical image denoising using bivariate laplacian mixture model. *IEEE Transactions on Biomedical Engineering*, 56(12):2826–2837, Dec. 2009.
- [206] F. Yu, Y. Chen, and L. Luo. CT image denoising based on sparse representation using global dictionary. In *2013 ICME International Conference on Complex Medical Engineering (CME)*, pages 408–411, May 2013.
- [207] K. Abhari, M. Marsousi, J. Alirezaie, and P. Babyn. Computed tomography image denoising utilizing an efficient sparse coding algorithm. In *11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, pages 259–263, Jul. 2012.
- [208] S. Amjad Ali, S. Vathsal, and K.K. Lal. An efficient denoising technique for CT images using window-based multi-wavelet transformation and thresholding. *European Journal of Scientific Research*, 48(2):315–325, Dec. 2010.
- [209] T. Jingtian, Y. Xiaoli, and P. Meisen. The application of independent component analysis in CT image de-noising. In *3rd International Conference on Bioinformatics and Biomedical Engineering , 2009. ICBBE 2009.*, pages 1–4, Jun. 2009.