Real-Time Volumetric Ultrasound Imaging for Handheld Applications

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

in partial fulfillment

of the requirements for the degree

Doctor of Philosophy

by

Kevin Owen

August

2012

APPROVAL SHEET

The dissertation

is submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

AUTHOR

The dissertation has been read and approved by the examining committee:

Advisor

Accepted for the School of Engineering and Applied Science:

Dean, School of Engineering and Applied Science

August

2012

Acknowledgements

I would like to acknowledge my parents, Catherine and David Owen, for always encouraging me to follow my dreams and not ever telling me that something was impossible or couldn't or shouldn't be done, and also for supporting me during my early academic career, helping to start a long journey that would end in this PhD dissertation. Also, I am eternally grateful to my wonderful wife Kate, both for inspiring me to go back to school with her own positive PhD experience, but perhaps more importantly for her unfailing support, including more recently several periods of effective single-parenthood on her part. I consider myself very fortunate to have the love and support of my family in this important endeavour.

I am also grateful for the funding sources which have enabled this course of study, including from the National Institutes of Health and the Wallace H. Coulter Foundation. Additionally, this work would not have been possible if it were not for the co-operation and support of the two start-up companies based upon the technology in this dissertation, Pocketsonics Inc., and Rivanna Medical LLC.

Special gratitude is reserved for my PhD advisor, Dr. John Hossack, for graciously accepting me into his laboratory, and providing the support, advice and coaxing needed to get me to the finish line. Additionally, I thank Dr. Will Mauldin for his very practical advice and guidance on navigating the final pitfalls of the PhD process. I am also grateful to the other members of my dissertation committee, Drs. Craig Meyer, Scott Acton, Yong Kim and Mohamed Tiouririne, for their helpful advice and insightful comments. Similarly, I acknowledege the important contributions of the co-authors of publications based on this work, including Dr. Will Mauldin, Dr. Michael Fuller, and Sarah Nguyen. Finally, I am very thankful for the support and advice of lab-mates over the years, including Dr. Drake Guenther, Dr. Mike Ellis, Dr. Matt Eames, Dr. Linsey Philips, Dr. Abhay Patil, Ali Dhanaliwala, Joe Kilroy, Dan Lin, Shiying Wang, Adam Dixon and Johnny Chen.

Abstract

Medical ultrasound has many benefits over other imaging modalities, including lack of ionizing radiation, relative portability and low cost. However, the majority of ultrasound scanners image only a 2D plane of tissue, which moves with the probe. This arrangement, along with the remote display screen, requires a high level of skill and experience to operate. Some recently introduced scanners have more intuitive 3D operating modes, but are often bulky and expensive. In this dissertation, several of the fundamental challenges of handheld, intuitive 3D ultrasound imaging systems are addressed, including energy efficient 2D beamforming, enhanced motion tracking using sector-scan probes, and optimal combination of motion estimates from various sensor modalities. In the field of vascular ultrasound imaging, the size and portability of ultrasound devices is limited by the energy cost of beamforming, particularly when transducers with many thousands of elements are involved. A separable approach to 2D beamforming, optimized for complex short-time-sequence signals is developed that reduces the energy cost of beamforming by a factor of 20, enabling real-time imaging in a 150 g device with multi-hour battery life. Imaging of spinal bone anatomy has poor performance for most ultrasound systems due to extremely bright bone reflections and systems optimized for tissue. Recognizing that a mechanically-scanned single piston transducer has intrinsic contrast advantages when imaging bone, techniques are developed to improve motion estimation using sector-scan ultrasound data, so that handheld freehand 3D spinal bone imaging is enabled. Finally, to address anisotropic motion estimation resolution using ultrasound alone, other sensor modalities (camera, accelerometer) are optimally combined to produce a handheld 3D imaging system capable of real-time guidance of epidural anesthesia procedures, with an RMS bone surface localization error of only 2.2 mm. These capabilities are demonstrated in a handheld battery-powered prototype with real-time 3D bone surface display. Initial *in-vivo* 2D and 3D images demonstrate feasibility of the device and imaging methodology.

Contents

List of Figures **v**

List of Tables **xv**

1 INTRODUCTION **1**

2 REAL-TIME VASCULAR IMAGING USING SEPARABLE BEAMFORMING **6**

2.A Abstract 6

2.B Introduction 7

2.C Theory 10

2.C.i Non-separable 2D Array Focusing 10

2.C.ii Separable 2D Array Focusing 11

2.C.iii FFT-Based 2D Array Focusing 16

2.D Materials and Methods 16

2.D.i Focusing Algorithm Implementations 16

2.D.ii Simulation and Experimental Methods 19

2.D.iii Focusing Computational Metrics 21

2.E Results 22

2.E.i Focusing Phase Errors and Delay-and-Sum Interpolation Errors for Separable 2D Array Focusing 22

2.E.ii Simulated Point-Spread Functions and Beamplots 22

2.E.iii Simulated and Experimental Anechoic Cylinder Images . 24

2.E.iv Computational Performance of Separable and Non-Separable 2D Focusing 31

2.F Discussion 31

2.F.i Imaging Performance of Separable 2D Beamforming Algorithms . . . 31
2.F.ii Computational Performance of of Separable 2D Beamforming Algorithms . . . 35
2.G Conclusion . . . 38

3 SECTOR-SCAN MOTION TRACKING FOR FREEHAND 3D IMAGING 39
3.A Abstract . . . 39
3.B Introduction . . . 41
3.C Theory . . . 44
3.C.i Statistical Model for Complex PSF Correlation with Motion 44
3.C.ii Statistics of Normalized Complex Correlation with PSF Transforms . . . 49
3.C.iii Sector Scan Geometric Model and Azimuthal Motion Estimation . . . 50
3.C.iv Sector Scan Elevational Motion Estimation . . . 53
3.C.v Relationship between Normalized Correlations for Complex RF and Detected Data . . . 56
3.C.vi Speckle-Tracking Motion Bias In Sector-Scan Systems . 57
3.D Materials and Methods . . . 61
3.D.i Complex Correlation Statistical Model Validation . . . 61
3.D.ii Joint Azimuthal-Elevational Motion Estimation Evaluation 63
3.D.iii Conventional Azimuth and Elevational Motion Estimation 66
3.D.iv Motion Tracking Validation Using 3D Spine Model . . . 67
3.E Results . . . 70
3.E.i Single A-line Complex Correlation Statistics . . . 70
3.E.ii Joint Azimuthal-Elevational Motion Estimation Evaluation 76
3.E.iii Conventional Azimuth and Elevational Motion Estimation 79
3.E.iv Motion Tracking Validation Using 3D Spine Model . . . 83
3.F Discussion . . . 86
3.F.i Statistics of Normalized Complex Correlation with PSF Transforms . . . 86
3.F.ii Speckle-Tracking Lateral Motion Estimation Bias Prediction 89

3.F.iii Joint Azimuthal-Elevational Motion Estimation Performance 90
3.G **Conclusion** 93
3.H **Appendices** 94
3.H.i Derivation of SVD Components as a Function of Complex PSF Inner Product 94

4 REAL-TIME 3D SPINE IMAGING USING ROBUST MULTIMODAL MOTION TRACKING 97
4.A **Abstract** 97
4.B **Introduction** 98
4.C **Background** 100
4.C.i Multi-Modal Motion Estimation 100
4.C.ii Probabilistic Combination of Multi-Modal Motion Estimates 104
4.C.iii Spine Imaging Assessment Using Receiver Operating Characteristic Analysis 106
4.D **Materials and Methods** 110
4.D.i Handheld Spinal Bone Imaging System 110
4.D.ii Calibration and Testing of Camera and Ultrasound Motion Estimation 113
4.D.iii Freehand 3D Spine Imaging Using Multi-Modal Motion Estimation 116
4.D.iv Receiver Operating Characteristic Analysis of Freehand 3D Scans 118
4.E **Results** 119
4.E.i Combined Ultrasound/Camera Motion Estimation 119
4.E.ii Freehand 3D Spine Imaging Using Multi-Modal Motion Estimation 124
4.F **Discussion** 130
4.F.i Combined Ultrasound/Camera Motion Estimation 130
4.F.ii Freehand 3D Spine Imaging Using Multi-Modal Motion Estimation 132
4.G **Conclusion** 134

5 FINAL CONCLUSIONS AND FUTURE DIRECTIONS **136**
5.A Conclusions 136
5.B Future Directions 138
5.C Publications 138

REFERENCES **140**

List of Figures

2.1 Non-separable 2D array focusing an M x N aperture (smaller white box) is translated across the array (larger white box) to form P beamformed points in the azimuthal direction (gray area), forming a weighted sum of M x N time-delayed (or phase rotated) signals for each of P points (the black square). This is repeated Q times to form P x Q focused outputs. 12

2.2 Separable 2D array focusing a 1 x N aperture (vertical white box) is translated to form P partially beamformed sums in the azimuthal direction, shown as a gray line in (A). The partial sums are each reused M times as an M x 1 aperture (white horizontal box) is translated to form P output points, shown as a gray line in (B). In conjunction these two apertures form an M x N aperture (dotted box). The process of (A) and (B) is repeated Q times to form P x Q focused outputs. 13

2.3 Unweighted (A) and apodization-weighted (B) RMS phase errors for separable focusing algorithm with varying f/# and focal depth = 5mm, 15mm or 25mm. 23

2.4 Simulated beamplots at focal depth of 15mm, f/#=1.4 for non-separable (A) and separable (B) delay-and-sum algorithms using 8-sample Hamming-windowed sinc interpolation (Sinc8) and cubic B-spline interpolation (BSpline). 23

2.5 Simulated 2D elevational-azimuthal point-spread functions for separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming, using f/#=1.4 and focal depth of 15mm (65dB logarithmic display range, 2-cycle transmit for DAS, 4-cycle transmit for DSIQ). 25
2.6 Simulated beamplots for non-separable and separable delay-and-sum and DSIQ beamforming, using f/#=1.4 and a focal depth of 5mm (A,B), 15mm (C,D) and 25mm (E,F). 26
2.7 Point spread functions for the 2D array from Table 1, focused using the NDAS, SDAS and SDF methods, with logarithmic display range of 65 dB. Row A shows PSFs for f/#=0.8, focal depth=5mm, center frequency=5MHz. Row B shows f/#=1.0, focal depth=10mm, center frequency=5MHz. Row C shows f/#=1.4 focal depth=15mm, center frequency=7 MHz. 27
2.8 Simulated 10-mm diameter anechoic cylinder C-mode image slices (first 2 columns) and RMS absolute value profile (taken over y dimension). For each, separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming are shown, using f/#=1.4 and focal depth of 15mm (50 dB logarithmic display range). White rectangles represent areas used to estimate CNR values. 28
2.9 Simulated 10-mm diameter anechoic cylinder B-mode image slices for separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming, using f/#=1.4 (50 dB logarithmic display range). 29
2.10 Experimental 10-mm diameter anechoic cylinder C-mode image slices (first 2 columns) and RMS absolute value profile (taken over y dimension). For each, separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming are shown, using f/#=1.4 and focal depth of 15mm (35 dB logarithmic display range). White rectangles represent areas used to estimate CNR values. 30

2.11 (A) C-Mode imaging frame rates for C-with-SIMD implementations of separable and non-separable delay-and-sum and DSIQ focusing on the OMAP 3530 processor (SDAS, SDF, NDAS, NDF). (B) Frame rates for NDF, SDF and FFT-based focusing implemented in MATLAB on the Intel Core i5 processor. 32

3.1 Sensitivity to scatterers of a single pixel from a focused complex demodulated A-line as a function F(x,y,z) of three dimensional space shown in (A). After transducer motion of vector r, the new scatterer sensitivity is shown in (B), as the function $\mathbf{F}(\mathbf{x}-\mathbf{r_x}, \mathbf{y}-\mathbf{r_y}, \mathbf{z}-\mathbf{r_z})$. 45

3.2 Multiple A-lines from a sector scan, showing the scatter sensitivity function or PSF at the same depth for each. Also shown is the local co-ordinate system for each A-line, with the unit vectors k_n in the A-line direction, j_n orthogonal to k_n in the azimuthal plane, and i_n in the elevational direction. Local co-ordinate axis vectors are defined in the global co-ordinates x, y, z, and centered on the PSF (shown offset for clarity). 51

3.3 Illustration of how tissue detected by a single pixel from a vertical A-line will be detected identically after lateral translation, in the case of identical, non-angled A-lines from a linear array system. This leads to an unbiased, single-pixel correlation peak. 59

3.4 Illustration of how tissue detected by a single pixel from a vertical A-line will be detected by rotated PSFs from other, angled A-lines after lateral translation, in the case of identical, angled A-lines from a sector-scan system. This leads to an offset single-pixel correlation peak, as the PSFs from angled lines at increasing translations rotate out of the tissue area detected by the original PSF. The geometry of rotated PSFs is also shown, where tissue detected by an original A-line at angle α_1 is detected by a PSF with rotation α_2 at a lateral shift of $r_0 + r$. 60

3.5 Magnitude and phase of single A-line normalized complex correlation for 200-pixel window with shifts of varying magnitudes at different angles relative to pure lateral motion (0°). For all results, RF pixel is at focus of piston transducer, at 56 mm depth. Correlation values calculated using a deterministic simulated 3D PSF inner product are shown in A and B. The corresponding measurements for PSF-shift simulations in fully-developed 3D speckle are shown in C and D. Correlation magnitude and phase values from experiments using a tissue-mimicking phantom are shown in C and F. Results from C and D were averaged over 50 independent 3D speckle realizations, and those from E and F were averaged over 50 independent phantom locations. 72
3.6 Standard deviation of single A-line normalized complex correlation magnitude for 200-pixel window with shifts of varying magnitudes at different angles relative to pure lateral motion including 0° (A), 10° (B), 20° (C) and 30° (D). For all results, RF pixel is at focus of piston transducer, at 56 mm depth. Standard deviation values are shown for a simulated 3D PSF inner product model (Mod.), for 50 realizations of a PSF-shift simulation in fully-developed 3D speckle (Sim.), and for experimental data at 50 different phantom locations (Expt.). 73
3.7 Angular standard deviation of single A-line normalized complex correlation angle for 200-pixel window with shifts of varying magnitudes at different angles relative to pure lateral motion including 0° (A), 10° (B), 20° (C) and 30° (D). For all results, RF pixel is at focus of piston transducer, at 56 mm depth. Circular standard deviation values are shown for a simulated 3D PSF inner product model (Mod.), for 50 realizations of a PSF-shift simulation in fully-developed 3D speckle (Sim.), and for experimental data at 50 different phantom locations (Expt.). 74

3.8 Normalized complex correlation values for a 200-pixel window from a single A-line, while undergoing shifts with magnitude 0.1 mm (○), 0.3 mm (×), 0.5 mm (+) and 0.7 mm (△). Panels A, B and C show 50 realizations each for the PSF model, 3D speckle simulation and experimental data with a shift angle of 10°, while D, E and F show the same data for a shift angle of 20°. 75
3.9 Azimuthal motion estimation performance using simulated data. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for gaussian PSF model (mg, sg) and Field II PSF model (mf, sf). 76
3.10 Azimuthal motion estimation performance using experimental data from a tissue mimicking phantom. Panel (A) shows azimuthal estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown. 77
3.11 Elevation motion estimation performance using simulated data from separable gaussian PSF model. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for 'fast' (equation 3.23, m1, s1) and 'accurate' (equation 3.22, m2, s2) estimators. 77
3.12 Elevation motion estimation performance using simulated data from Field II PSF model. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for 'fast' (equation 3.23, m1, s1) and 'accurate' (equation 3.22, m2, s2) estimators. 78

3.13 Elevation motion estimation performance using experimental data from a tissue-mimicking phantom. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for 'fast' (equation 3.23, m1, s1) and 'accurate' (equation 3.22, m2, s2) estimators. 78
3.14 Azimuthal motion estimation performance for sector-scan and speckle-tracking methods using experimental data from a tissue mimicking phantom. Panel (A) shows azimuthal estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for the sector scan method (mss, sss) and speckle-tracking method (mst, sst). 80
3.15 Elevational motion estimation performance for sector-scan and speckle-tracking methods using experimental data from a tissue mimicking phantom. Panel (A) shows elevational estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for the sector scan method (mss, sss) and the decorrelation-based method (mdc, sdc). 80
3.16 Simulated and experimental normalized correlation curves for single, vertically-oriented detected A-line, using 200-pixel window centered at depth of 56 mm. Underlying shifts of 0.5 mm, 1.0 mm and 1.5mm are shown. Top panel (A) shows simulated correlation curves with peaks (○), and actual shifts (+). Experimental results averaged over 50 realizations are shown in panel (C). Bottom panels (B) and (D) are identical to (A) and (C), zoomed to area of interest. 81

3.17 Simulated and experimental bias measurements for lateral motion estimation using samples at various initial A-line sector angles with lateral shifts of 0.5 mm, 1.0 mm and 1.5 mm. Simulated bias (A) is found using a deterministic PSF model to find the peak of the correlation curve with lateral shifts. Experimental bias (B) is calculated for a given initial A-line angle by finding the 200-sample long vertical strip of scan-converted, detected pixels corresponding to that angle, then shifting laterally to find the correlation peak. 82

3.18 True and estimated paths taken during a semi-random walk, with two kinds of motion estimation, sector-scan based joint azimuthal-elevation estimation, and conventional, speckle-tracking azimuthal motion estimation coupled with decorrelation-based elevation motion estimation. 84

3.19 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from 258 sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using ground-truth location for each sector-scan image. 84

3.20 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from 258 sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using joint azimuthal-elevational sector-scan motion estimation to place each sector-scan image. 85

3.21 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from 258 sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using speckle-tracking and decorrelation-based motion estimation to place each sector-scan image in the lateral and elevational axes respectively. 85

4.1 Representative image from fluoroscopically guided epidural procedure, with needle to left of image. 101

4.2 Six degrees of freedom required to uniquely identify the position and orientation of an object. On the left, the BACK-FRONT(x), LEFT-RIGHT (y) and UP-DOWN (z) translational degrees of freedom are indicated. On the right, the rotational degrees of freedom are shown, including rotation about the x axis (ROLL), y axis (PITCH) and z axis (YAW). The gray trapezoid represents a B-mode scan plane. 102

4.3 Example ROC curves, including from a random classifier, and a generic example classifier. Also shown is the optimal operating point on the example curve, defined as the point on the curve with the smallest Euclidean distance to the top-left of the graph. . 109

4.4 Block diagram of the ‘SpineFinder’ experimental handheld spinal bone imaging system. The three different sensor modalities used for motion detection are shown, along with the CPU, mass storage (Secure Digital or SD Card), network conectivity and auxiliary features such as rechargeable battery and user interface components. 111

4.5 Top view of the ‘SpineFinder’ experimental handheld spinal bone imaging system. The user interface components (touchscreen, buttons, optical finger sensor) are visible, along with the hemispherical sector-scan ultrasound transducer at the very bottom of the system. 111

4.6 Bottom view of the ‘SpineFinder’ experimental handheld spinal bone imaging system. The camera circular aperture is clearly visible in the device handle, to the right of the ultrasound transducer. 112

4.7 Side view of the ‘SpineFinder’ experimental handheld spinal bone imaging system. The preferred orientation of the device is shown, with the ultrasound transducer oriented perpendicularly to the skin surface, and the camera aperture parallel to the skin. 112

4.8 Histograms of motion estimation error for ultrasound azimuthal (A) and elevational (B), camera azimuthal (C) and elevational (D), and for optimally combined azimuthal (E) and elevational (F) estimation. 120
4.9 Montage of individual photos taken by camera on semi-random walk over phantom with overlaid porcine tissue (country ham slice, 10 mm). Three runs with different random walk motion are shown in A, B and C. 121
4.10 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using ground-truth location for each sector-scan image. 122
4.11 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using ultrasound-derived location for each sector-scan image. 122
4.12 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using camera-derived location for each sector-scan image. 123
4.13 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using combined ultrasound/camera-derived location for each sector-scan image. 123
4.14 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a bidirectional scan of the spine using ultrasound-derived location for each sector-scan image. 125

4.15 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a bidirectional scan of the spine using camera-derived location for each sector-scan image. 126
4.16 Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a bidirectional scan of the spine using combined ultrasound/camera-derived location for each sector-scan image. 126
4.17 ROC Curves from freehand scan 1, using ultrasound-only, camera-only and combined motion estimation. 127
4.18 Rotational motion in the azimuthal plane (‘roll’) estimated using sector-scan ultrasound data and accelerometer data. 127
4.19 B-mode image of slice of human spine from freehand ultrasound scan, showing the spinous process and transverse process for a single vertebrae. 128
4.20 B-mode image of slice of human spine from freehand ultrasound scan, showing the interlaminar space and vertebral body for a single vertebra. 129
4.21 3D surface rendering of human spine from freehand ultrasound scan, showing the spinous (white, top) and transverse processes (left, right, yellow), and vertebral body (dark, bottom) from a single vertebra. 129
4.22 3D surface rendering of human spine from freehand ultrasound scan, showing one vertebra with spinous (white) and transverse processes (yellow), and with clearly visible epidural gap between this and adjacent vertebra to the top-left. Spinal axis is in top-left to bottom-right direction in this rendering. 130

List of Tables

2.1 Array Parameters for Experimental and Simulated System 17
2.2 Processor Parameters 22

3.1 Transducer/Scan Parameters for Experimental and Simulated System 61
3.2 Complex Correlation Gaussian-Magnitude and Linear Phase Parameter Fits for Model, Simulated and Experimental Data 71
3.3 Experimental Lateral Motion Estimation Bias 83
3.4 Root-Mean-Square Bone Surface Error for Ground-Truth, the Proposed Sector-Scan Tracking Algorithm and for Conventional Speckle-Tracking/Decorrelation-based Motion Estimation 83

4.1 Motion Estimation Capabilites of Different Sensor Modalities . . . 104
4.2 Bone Surface Classification Categories for ROC Analysis 108
4.3 Truth Table for ROC Voxels 118
4.4 Root-Mean-Square Estimation Error for Ultrasound-based, Camera-based and Combined Estimation 124
4.5 Root-Mean-Square Bone Surface Error for Various Motion Estimators Over 3 Random Walks 124
4.6 Root-Mean-Square Bone Surface Error for Various Motion Estimators Over 3 Bidirectional Freehand Spine Scans 125
4.7 ROC Curve Parameters for Freehand Scan 1 128

Chapter 1

INTRODUCTION

Medical ultrasound is a widely used medical imaging modality with benefits that, with respect to most competing modalities, include: high resolution, lack of ionizing radiation, low cost and portability. However, the majority of existing ultrasound scanners are B-mode systems that produce 2D image slices through a 3D volume of tissue. Considerable skill and experience is required to comprehensively acquire and interpret these 2D slices, which can limit clinical utility. For instance, when using conventional B-mode ultrasound to perform spinal anesthesia, it has been demonstrated that success rates are highly dependent on user skill and familiarity with ultrasound (1). Additionally, ultrasound images contain a number of artifacts that manifest differently with changing interrogation angles. Common ultrasound artifacts occur due to reverberation, phase aberration, bright off-axis targets, and shadowing (2). Recently, ultrasound scanners with 3D volume imaging capabilities have been introduced, with the potential for reduced interpretation errors (3), (4) and more intuitive imaging of real tissue. 3D imaging typically requires a 2D array with several thousand elements, resulting in high scanner and transducer complexity and cost (5). Thus, current 3D volume imaging is mostly limited to expensive and bulky freestanding systems.

Recently, handheld ultrasound imaging systems have been developed (6), (7). These systems typically still use the B-mode imaging mode and experience the same B-Mode image interpretation challenges. In addition, current handheld ultrasound systems follow the longstanding paradigm of separating the image display unit from the transducer - further hindering intuitive interpretation of 3D tissue structure. Lower cost handheld ultrasound systems with more intuitive display modes may lead to more pervasive use of ultrasound in medicine and subsequently improved patient outcomes. Increased device mobility will also potentially alleviate some workflow problems associated with bulky, freestanding scanners, i.e. the scanner can be taken to the patient rather than the patient to the scanner.

The combination of 3D ultrasound imaging capabilities with a handheld form-factor and intuitive display has the potential to have profound impact in the medical imaging field. Current handheld ultrasound scanners may be viewed as being essentially a result of a number of cost-performance trade-offs applied to existing cart-based conventional designs. In contrast, if 3D imaging is used in conjunction with non-standard, intuitive display modes, which are tuned for specialized applications, these advances may enable wider adoption by non-specialist users with the result of better clinical outcomes at lower cost to the healthcare system.

In this PhD dissertation, substantive contributions are made toward low-cost, handheld 3D imaging with intuitive display of the target tissue for two specialized applications: imaging of the vasculature for intravenous needle guidance, and imaging of the spinal bone anatomy to aid in central neuroaxial anesthesia procedures. Ultrasound guidance is increasingly used for intravenous (IV) access procedures in emergency and routine situations (8), and has been demonstrated to increase ultimate success rates for peripheral IV access from a baseline rate of 33% up to 97% in patients with difficult venous access (9). Central neuroaxial anesthesia procedures (e.g. epidural and spinal anesthesia) are typically performed at the bedside without guidance (i.e. the blind method) (10), resulting in failure rates of 40%-70% (11), (12), (13) and poor patient outcomes (11). Patients who are difficult to image with high quality using ultrasound (e.g. obese patients) can be imaged using fluoroscopy. However this exposes the patient to ionizing X-ray radiation, recently shown to be responsible for 2% of all cancers and up to 11,000 deaths per year in the U.S. (14), (15). Unfortunately, the patients that are most likely to need fluoroscopic guidance in lumbar procedure (i.e. those with high body mass index) also typically require higher radiation doses during fluoroscopy, compounding the problem. In both of these applications, intuitive 3D imaging is expected to improve outcomes, but there are difficulties to be overcome in producing a functional handheld device with the required low cost, portability, efficacy and battery charge-cycle life.

In chapter 2, a beamforming strategy that leads to 20 x improvements in frame rate or energy efficiency for real-time handheld vascular imaging is presented; this is a significant enabling technology without which the form factor, battery life or imaging quality in this application would be significantly degraded. Additionally, it is shown that the improvements in computational efficiency using this algorithm have a negligible associated performance reduction in most practical conditions. Although targeted at C-mode imaging (16) in extremely small (170 g)(17) devices, this method enables, and has demonstrated handheld volumetric imaging, if reduced frame-rate and/or battery life are permissible.

Chapter 3 introduces a motion-tracking technology that addresses deficiencies in existing algorithms when applied to sector-scan ultrasound systems. Recently published work (18) contributed to by the author indicates that a mechanically-scanned piston transducer has attractive properties when imaging bright specular-reflecting targets such as bone surfaces. A new statistical model for ultrasound motion tracking is introduced in this chapter, and used to both successfully improve motion estimation performance for sector-scan systems, and also make experimentally-validated predictions about bias in existing motion estimation algorithms applied to sector-scan systems.

In chapter 4, the ultrasound-based motion estimation of chapter 3 is optimally combined with information from other motion-sensing modalities (camera, accelerometer) to produce a robust, multi-modality motion estimator. This work is an enabling technology for the 'Spine Finder', the world's first real-time handheld ultrasound imaging system for imaging spinal bone anatomy and guiding spinal access procedures. Additional work performed by the author includes design and implementation of all device hardware and software in the prototype system described in this chapter.

The work presented in this dissertation is organized into self-contained chapters, with each chapter's content largely taken from peer-reviewed journal articles, or from drafts currently submitted for peer-reviewed publication, or for content imminently awaiting submission to a peer-reviewed journal. In the interests of clarity and consistency, minor editing is performed on the described papers/chapters.

Chapter 2

REAL-TIME VASCULAR IMAGING USING SEPARABLE BEAMFORMING[1]

2.A Abstract

Two-dimensional arrays present significant beamforming computational challenges due to the high channel count and data rate. These challenges are even more stringent when incorporating a 2D transducer array into a battery-powered handheld device, placing significant demands on power efficiency. Previous work in sonar and ultrasound indicates that 2D array beamforming can be decomposed into two separable line-array beamforming operations. This has been used in conjunction with frequency-domain phase-based focusing to achieve fast volume imaging. In this chapter, I analyze the imaging and computational performance of approximate near-field separable beamforming for high-quality delay-and-sum (DAS) beamforming and for a low-cost,

[1]Chapter 2 appears in the peer-reviewed publication :
K. Owen, M. I. Fuller and J. A. Hossack, "Application of X-Y Separable 2D Array Beamforming For Increased Frame Rate and Energy Efficiency in Handheld Devices", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, 59 (7), 2012*

phase-rotation only beamforming method known as Direct-Sampled In-Phase Quadrature (DSIQ). I show that when high-quality time-delay interpolation is used, separable DAS focusing introduces no noticeable imaging degradation under practical conditions. Similar results for DSIQ focusing are observed. In addition, a slight modification to the DSIQ focusing method greatly increases imaging contrast, making it comparable to that of DAS, despite a wider main-lobe and higher sidelobes due to the limitations of phase-only time-delay interpolation. Compared to non-separable 2D imaging, up to a twenty-fold increase in frame rate is possible with the separable method. When implemented on a Texas Instruments OMAP 3530 smart phone oriented processor to focus data from a 60 x 60 channel array using a 40 x 40 aperture, the frame rate per C-mode volume slice increases from 16 Hz to 255 Hz for DAS, and from 11 Hz to 193 Hz for DSIQ. Energy usage per frame is similarly reduced from 75 mJ to 4.8 mJ/frame for DAS, and from 107 mJ to 6.3 mJ/frame for DSIQ. I also show that the separable method outperforms 2D FFT-based focusing by a factor of 1.64 at these data sizes. This data indicates that with the optimal design choices, separable 2D beamforming can significantly improve frame rate and battery life for handheld devices with 2D arrays.

2.B Introduction

The majority of current beamformers operate by summing weighted time-delayed signals from all channels in an aperture to form a single beam on receive (19). For 2D arrays producing volume (azimuth, elevation, time/depth) data, beamforming can be considered a spatial filtering or convolution operation

(20). However, the computational and energy cost of beamforming can be prohibitive when applied to 2D transducer arrays with many thousands of channels. For handheld, battery-operated systems, the energy cost is particularly important for multi-hour battery life. Several methods have been proposed to increase the frame-rate of 2D array ultrasound systems, including sparse 2D arrays (21, 22, 23, 24), synthetic aperture approaches (25, 26, 27), transmit-receive coarrays (28, 29), subaperture methods (5, 30, 31, 32, 33, 34), parallel beamforming (35, 36) and using plane wave transmit with limited-diffraction receive beam focusing (37, 38, 39). Phased 2-D subarray focusing, or micro-beamforming has been suggested for diagnostic ultrasound imaging with 2D arrays as an approach to perform partial focusing close to or in, the transducer assembly, to reduce both interconnect complexity and total computational cost (5, 32). The decomposition of a 2D beamforming process into two separable 1D line array beamforming steps has been proposed for computationally efficient volume focusing (40). Computational efficiencies are achieved with this method by re-using the results of each 1D beamformed partial sum multiple times. Various frequency domain beamforming efficiencies are subsequently employed, as suggested by Maranda (41). These include using a 1D FFT in the time dimension to implement delays for narrowband signals (42), and using a 2D FFT in the X-Y plane for SONAR volume imaging (43). The separable approach has also been extended to near-field wide-band SONAR applications using the chirp zeta transform (CZT) and the Fresnel approximation (44). Separable implementations of 3D ultrasound imaging have been developed to run in real-time on a 16-node PC cluster, (45, 46) using a variation on the time-series 1D FFT acceleration method (42). Real-time 3D ultrasound beamforming implementations using clusters of PCs or several FPGAs (47) primarily target system performance. However, the above methods are not capable of practical real-time imaging on a battery powered, handheld system with a fully-sampled 2D array. In this chapter, I analyze the performance, in terms of resolution, contrast, computational time and energy consumption per frame, of practical 2D separable beamformers for volume and C-mode imaging in handheld devices using successive 1D convolutions in the azimuth and elevation directions. I first investigate a separable version of conventional, high quality delay-and-sum beamforming using different time-delay interpolation methods. In addition, I explore the performance of a 2D focusing method developed especially for power-efficient C-mode imaging in handheld devices (Directly-Sampled In-phase Quadrature [DSIQ] beamforming (16, 17)). This method greatly reduces power consumption by not requiring full time-series on each channel, typically requiring only tens of milliwatts of power for an entire 3600-channel analog front-end, rather than tens of milliwatts per channel for each channel in conventional analog front-end systems (17). However, DSIQ has limited time-delay resolution due to the use of phase-rotations as approximations for time delays. Finally, for comparison purposes, we assess the computational performance of non-separable 2D focusing using fast FFT-based convolution. For an M x N 2D focal aperture, separable focusing yields an MN/(M+N) speed increase over non-separable focusing, producing a twenty-fold increase for a typical 40 x 40 element aperture, independent of array size. This level of performance gain is significant for handheld 2D-array ultrasound systems, where any intrinsic frame rate capability above 30 frames per second is recovered as additional battery life.

2.C Theory

2.C.i Non-separable 2D Array Focusing

To form a single beamformed output value in a 2D array ultrasound system, the signals from a M x N receive aperture from a larger array, arranged laterally about the projection of the focal point onto the transducer plane are appropriately delayed and weighted before summing. This is described in equations 2.1 - 2.3 for time-delay focusing at a point at location (X, Y, Z_f) in space in the region under array element (p, q). Here, $x(i)$ and $y(j)$ are the coordinates of aperture element (i, j), k is the wavenumber $2\pi f_{center}/c$, $R_{XY}(i, j)$ is the distance from aperture element (i, j) to the focal point, $\tau_{XY}(i, j)$ is the associated propagation time delay, $A(i, j)$ is the apodization function over the aperture, and $s(i, j, t - \tau_{XY}(i, j))$ is the time signal from aperture element (i, j) delayed by $\tau_{XY}(i, j)$. The summation output $F_{XY}(p, q, t)$ is a time series that is evaluated at t=0, after envelope detection and other steps. For phase-rotation based focusing, a single complex sample is available for each element in the focal aperture, $s(i, j)$. A complex weight $C(i, j)$ from equation 2.4 is then applied incorporating propagation phase and apodization, before summation as in equation 2.5.

$$R_{XY}(i, j) = Z_f + \sqrt{(X - x(i))^2 + (Y - y(j))^2 + {Z_f}^2} \tag{2.1}$$

$$\tau_{XY}(i, j) = R_{XY}(i, j)/c \tag{2.2}$$

$$F_{XY}(p,q,t) = \sum_{i=-(M-1)/2}^{(M-1)/2} \sum_{j=-(N-1)/2}^{(N-1)/2} A(i,j)s(p-i,q-j,t-\tau_{XY}(i,j)) \quad (2.3)$$

$$C(i,j) = A(i,j)e^{-jkR_{XY}(i,j)} \quad (2.4)$$

$$F_{XY}(p,q) = \sum_{i=-(M-1)/2}^{(M-1)/2} \sum_{j=-(N-1)/2}^{(N-1)/2} C(i,j)s(p-i,q-j) \quad (2.5)$$

The non-separable 2D focusing operation is shown schematically in 2.1. In this case, the M x N focal aperture is shown translating to focus at a series of P points in the azimuthal direction. To focus at one point involves MN delay and sum operations, where delay may be a real time delay or phase-rotation based approximation to a time delay. In each case, aperture apodization is also applied. Therefore to form an output C-mode image plane of P x Q focused points requires MNPQ prototypical delay-and-sum operations. The focusing operation can also be interpreted as a spatial filtering or convolution operation (20). Using equation 2.5, phase-rotation based focusing is equivalent to a 2D complex convolution operation.

2.C.ii Separable 2D Array Focusing

Separable 2D array focusing operates by decomposing the propagation distance $R_{XY}(i,j)$ into two components $R_X(i)$ and $R_Y(j)$, such that $R_{XY}(i,j) \sim R_X(i) + R_Y(j)$. Similarly, the apodization weighting $A(i,j)$ is approximated by the product $A_X(i)A_Y(j)$. This reduces the number of unique delay and weight-

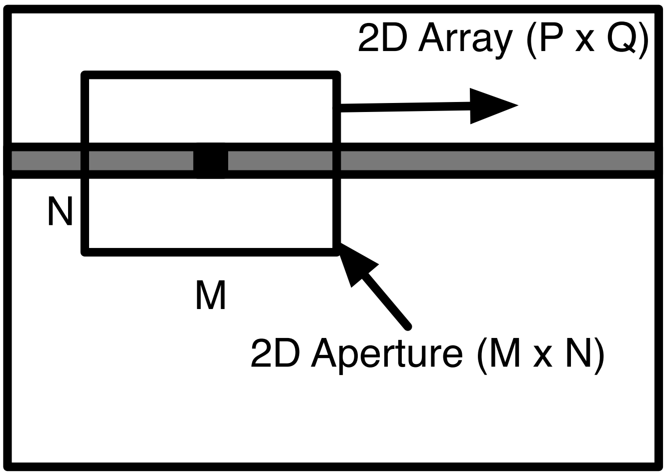

Figure 2.1: Non-separable 2D array focusing an M x N aperture (smaller white box) is translated across the array (larger white box) to form P beam-formed points in the azimuthal direction (gray area), forming a weighted sum of M x N time-delayed (or phase rotated) signals for each of P points (the black square). This is repeated Q times to form P x Q focused outputs.

ing operations for an M x N aperture from MN to M+N, so that each of the M possible unique azimuthal delays and weights for an element can be re-used when the element is at N different elevational positions in an aperture and vice versa. The separable focusing process is shown in Figs. 2.2A and 2.2B as an M x N aperture is used to focus at a line of P points in the azimuthal direction, repeated for each of Q elevational lines (not shown) to form a full P x Q image. In 2.2A, a line of P partially focused outputs are formed by summing over a 1 x N elevational aperture with delays corresponding to $R_Y(j)$ and weights of $B(j)$, as the aperture translates to cover P azimuthal locations. In the second step, shown in 2.2B, an M x 1 aperture with delays corresponding to $R_X(i)$ and weights of $A(i)$ is used to form azimuthal sums as the aperture translates over P azimuthal locations. Each partial sum of N delayed and weighted val-

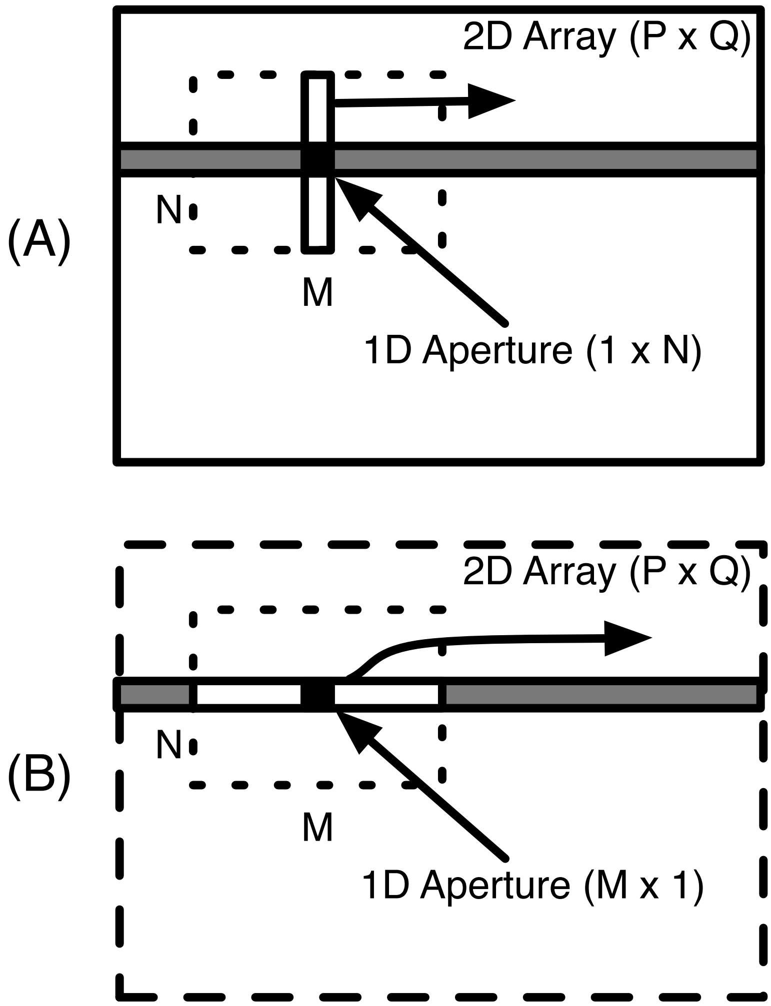

Figure 2.2: Separable 2D array focusing a 1 x N aperture (vertical white box) is translated to form P partially beamformed sums in the azimuthal direction, shown as a gray line in (A). The partial sums are each reused M times as an M x 1 aperture (white horizontal box) is translated to form P output points, shown as a gray line in (B). In conjunction these two apertures form an M x N aperture (dotted box). The process of (A) and (B) is repeated Q times to form P x Q focused outputs.

ues from the first step is re-used M times in the second step. If the process to form a single line of P focused points is repeated for Q similar lines at different elevational positions, the total number of delay, weight and accumulate operations for the two-step separable focusing process is PQ(M+N). This reduces the computational cost by a factor (M+N)/MN compared to the non-separable case.

$$r_{XY} = \sqrt{Z_f^2 + \Delta X^2 + \Delta Y^2} = Z_f \sqrt{1 + \frac{\Delta X^2}{Z_f^2} + \frac{\Delta Y^2}{Z_f^2}} \tag{2.6}$$

$$\sqrt{1+b} = 1 + \frac{1}{2}b - \frac{1}{8}b^2 + \frac{1}{16}b^3 \quad b = \frac{\Delta X^2}{Z_f^2} + \frac{\Delta Y^2}{Z_f^2} \tag{2.7}$$

$$r_{XY} = Z_f \left(1 + \frac{1}{2}\left(\frac{\Delta X^2}{Z_f^2} + \frac{\Delta Y^2}{Z_f^2}\right) - \frac{1}{8}\left(\frac{\Delta X^2}{Z_f^2} + \frac{\Delta Y^2}{Z_f^2}\right)^2 ...\right) \tag{2.8}$$

$$r_X = Z_f \left(1 + \frac{1}{2}\left(\frac{\Delta X^2}{Z_f^2}\right) - \frac{1}{8}\left(\frac{\Delta X^2}{Z_f^2}\right)^2 ...\right) \tag{2.9}$$

$$r_Y = Z_f \left(1 + \frac{1}{2}\left(\frac{\Delta Y^2}{Z_f^2}\right) - \frac{1}{8}\left(\frac{\Delta Y^2}{Z_f^2}\right)^2 ...\right) \tag{2.10}$$

To break the geometric delay $R_{XY}(i,j)$ from equation 2.1 into the separable components $R_X(i)$ and $R_Y(j)$, the rightmost term of $R_{XY}(i,j)$ can be rewritten as r_{XY} in equation 2.6, where the x and y dimension differences are abbreviated to ΔX and ΔY respectively. A Taylor series expansion of equation 2.6 is developed in equations 2.7 and 2.8. The first two terms of this expansion are equivalent to the Fresnel approximation (48, 49). In equations 2.9 and

2.10, r_X and r_Y are the Taylor expansions of equation 2.6 with ΔY and ΔX set to zero respectively. It is clear that a sum of the three-term Taylor expansions of r_X and r_Y is equivalent to the three-term expansion of $r_X Y$ except for an additional constant Z_f and a non-separable X-Y component in the third term of r_{XY}. This suggests that the condition $R_{XY}(i,j) \sim R_X(i) + R_Y(j)$ can be met using the forms of equations 2.11 and 2.12, where the Z_f in $R_Y(j)$ is used to cancel an extra constant that would otherwise appear in the sum. Equation 2.13 describes the resulting azimuthal and elevational propagation time delays, $\tau_X(i)$ and $\tau_Y(j)$, which similarly satisfy $\tau_{XY}(i,j) \sim \tau_X(i) + \tau_Y(j)$.

$$R_X(i) = Z_f + \sqrt{(X - x(i))^2 + Z_f^2} \tag{2.11}$$

$$R_Y(j) = -Z_f + \sqrt{(Y - y(j))^2 + Z_f^2} \tag{2.12}$$

$$\tau_X(i) = R_X(i)/c, \ \ \tau_Y(j) = R_Y(j)/c \tag{2.13}$$

Both of the delay, weight and summation steps of separable focusing are given for a true delay-and-sum implementation in equations 2.14 and 2.15. The summation output $F_{XY}(p,q,t)$ is a time series that is evaluated at $t = 0$ after envelope detection. When phase-rotation based focusing is employed, the two separable focusing steps are as described in equations 2.16 and 2.17 using complex multiplies rather than time delays.

$$F_X(p,q,t) = \sum_{i=-(M-1)/2}^{(M-1)/2} A_X(i) s(p-i, q, t - \tau_X(i)) \tag{2.14}$$

$$F_{XY}(p,q,t) = \sum_{j=-(N-1)/2}^{(N-1)/2} A_Y(j) F_X(p, q-j, t-\tau_Y(j)) \tag{2.15}$$

$$F_X(p,q) = \sum_{i=-(M-1)/2}^{(M-1)/2} A_X(i) e^{-jkR_X(i)} s(p-i, q) \tag{2.16}$$

$$F_{XY}(p,q) = \sum_{j=-(N-1)/2}^{(N-1)/2} A_Y(j) e^{-jkR_Y(j)} F_X(p, q-j) \tag{2.17}$$

2.C.iii FFT-Based 2D Array Focusing

Repeated 2D summation over a focal aperture as it shifts in a 2D plane across the sample data is equivalent to convolution. An N x N sized 2D convolution can be calculated using the convolution theorem and 2D FFTs and IFFTs. 2D FFTs can be calculated in $O(N^2 log(N))$ time as a result of performing N 1D FFTs for each of the X and Y dimensions, taking $O(Nlog(N))$ individually (50). However, zero-padding is required to avoid cyclic convolution issues, and dual domain (time/frequency) data representation increases memory requirements. In addition, if fixed-point arithmetic is used, FFT-based convolution introduces significant rounding errors that increase with FFT length (51, 52).

2.D Materials and Methods

2.D.i Focusing Algorithm Implementations

Unless otherwise stated, experimental and simulated data were focused using the array parameters given in Table 2.1., based on the prototype 2D array of

Eames et al (53) and typical experimental imaging settings for DSIQ and DAS focusing.

Table 2.1: Array Parameters for Experimental and Simulated System

Property	Value	Units
Array Size	60 x 60	Element
Pitch	300	μm
Center Frequency	5	MHz
Cycles (DSIQ)	4	N/A
Cycles (DAS)	2	N/A

Non-separable delay-and-sum (NDAS) and separable delay-and-sum (SDAS) focusing were implemented in MATLAB (Mathworks, Natick, MA). Two different kinds of time-delay interpolation were used: an 8-sample Hamming-windowed sinc function, and a cubic B-spline based method. Cubic B-spline interpolation works by operating a 2-tap IIR filter up, then down each receive channel time series, before application of a 4-tap FIR filter for each individual interpolation step (54, 55). As there are many more time-delay operations than receive channels, in the limit the B-spline method is approximately twice as fast as an 8-tap windowed sinc operation, with an interpolation error reduced by 3.5 dB - 5.5 dB (56). For NDAS focusing, the 2D time delay profile from equation (3) was used to create an N x N x L convolution kernel, with each of the N x N vertical time series implementing a time delay, using windowed sinc (L=8) or B-spline interpolation (L=4),with an integer sample offset. This kernel was then used in a spatially variant 3D convolution with volume data from the 60 x 60 array to produce focused RF output. For SDAS focusing, a 1 x N x L azimuth-focused kernel and an N x 1 x L elevation-focused kernel were similarly created according to equations (11-13) and convolved together to make an effective kernel for focusing as in the NDAS case. Non-separable DSIQ focusing (NDF) and separable DSIQ focusing (SDF) algorithms were also implemented in MATLAB, operating on 4 real samples per channel. These samples were taken at the time intervals $s_1 = t_0$, $s_2 = t_0 + \lambda/4$, $s_3 = t_0 + \lambda$, $s_4 = t_0 + 5\lambda/4$, with $\lambda = f_c/c$ and t_0 the round-trip propagation time from the array to the focal depth. The first two samples per element, separated by a quarter period, are treated as the real and imaginary parts of a first complex sample. The next two samples similarly become the second complex sample. Time delays can then be implemented by weighted phase rotations of the two complex samples per channel. A more detailed treatment of DSIQ sampling is given in (16). The set of first complex samples from each channel were focused separately from the set of second complex samples, and the results added. This permits independent complex focusing kernels for the first and second complex sample data sets, taking into account how close to the geometric waveform center in time the first and second complex samples are. A weighting function was used for each aperture element (i, j) to bias the final output towards the complex sample closest to the ideal time delay. A gaussian function, with full-width half maximum (FWHM) equal to the separation of the two complex samples was chosen to change weighting smoothly while biasing strongly towards the nearest complex sample. This is shown in equation 2.18 where $w_s(i, j)$ is the complex sample weight, t_s is the complex sample time, $\tau(i, j)$ is the geometric target time delay for aperture element (i, j) and k is a constant chosen to set the stated FWHM for the weighting.

$$w_s(i,j) = e^{-k(t_s - \tau(i,j))^2}, \;\; k = -4ln(0.5) f_c^2/\lambda^2 \tag{2.18}$$

For all focusing algorithms an N x N focusing aperture was assumed, based on f-number. In the case of NDF, for each of the two complex samples an N x N array of complex focusing coefficients was calculated for a particular focal depth using equations (2.1, 2.4, 2.5), with radially symmetric apodization and per-element aperture weighting for each of the two arrays according to equation 2.18. The MATLAB function 'conv2' was used to perform 2D complex convolution in-place using double-precision floating point arithmetic. The phase of the non-separable 2D DSIQ focusing aperture is used as a reference phase for calculation of root-mean-square phase error of the separable focusing algorithms. Apodization-weighted RMS phase errors are calculated to give an indication of phase error significance taking into account aperture shading. For SDF focusing, azimuth-focused and elevation-focused focusing vectors were produced, with dimensions 1 x N and N x 1 respectively, according to equations (2.11-2.13) using the same apodization window used for both $A_X(i)$ and $A_Y(j)$. The two 1D focusing vectors were convolved together to form an equivalent N x N outer product convolution kernel, applied independently to the first and second set of complex samples before combination into a final image. For the SDF case, the weighting of equation 2.18 is applied in each of the x- and y- dimensions, producing an N x N product aperture weight as used for NDF.

2.D.ii Simulation and Experimental Methods

All simulations were performed using the Field II program (57), and the parameters of Table 2.1 (unless otherwise stated), with 2 x 2 mathematical elements per physical element and gaussian-windowed transmit pulses with the required bandwidths. In all cases, a sample rate of 128 times the center frequency was employed in Field II to avoid artifacts. The output of the simulation was downsampled to 40 MHz before beamforming to simulate a realistic hardware system. To compare separable beamformer imaging performance to non-separable equivalents, simulated PSFs and beamplots, plus simulated and experimental anechoic cyst images were produced. For the anechoic cyst images, contrast-to-noise ratios (CNRs) were calculated using equation 2.19, where μ and σ represent the log-scale mean and standard deviations of the image in the lesion and background areas as subscripted.

$$CNR = \frac{\mu_{lesion} - \mu_{bgnd}}{\sqrt{\sigma_{lesion} + \sigma_{bgnd}}} \tag{2.19}$$

All experimental data were obtained using a prototype of the Sonic Window handheld C-mode ultrasound scanner (17). This is a fully portable, battery operated system with an integrated 2D transducer array (53) with parameters from Table 2.1, and weighing less than 170 grams. Custom front-end ICs flip-chip attached to the 2D transducer array acquire the received data from every element in parallel after each transmit event. For all experiments a plane wave 4-cycle, 5MHz transmit pulse was used. This was followed by the capture of 4 samples per element at the required instants using a 40MHz sample clock, repeated at different depths to acquire volume data. A tissue mimicking near-field ultrasound phantom, including a 10mm diameter anechoic cylinder at a depth of 15mm was used as a target for the experimental data (Model 050, CIRS, Norfolk, VA).

2.D.iii Focusing Computational Metrics

The computation time required to perform separable and non-separable versions of delay- and-sum (NDAS, SDAS) and DSIQ focusing (NDF, SDF) was measured for a variety of aperture sizes on the Texas Instruments OMAP 3530 smart phone oriented processor, described in Table 2.2. The energy cost of the computations was calculated using an integration of directly measured current consumption of the OMAP processor and associated power management integrated circuit (PMIC), with a fixed supply voltage of 3.3V. All algorithms were implemented in C, using 16-bit signed integer data and compiled with gcc using the -O3 code optimization level, with and without the inner loops optimized to use single-instruction multiple-data (SIMD) assembly instructions. These instructions are capable of performing, for example, 4 multiply-accumulate operations in parallel. In all cases, computation time represents the time to focus a single C-mode slice of a volume-focused image. For the delay-and-sum focusing algorithms using cubic B-spline interpolation, this is the time to perform 4 separate scalar 2D convolutions of a 60 x 60 array with an N x N kernel. For DSIQ-based focusing, computation time is the time to perform 2 complex 2D convolutions of a 60 x 60 array with an N x N kernel. Each timed computation was averaged over 100 runs, alternating between two input data sets to obtain realistic cache usage. In addition, to compare the performance of separable 2D focusing using convolution with FFT-based 2D convolution, both were implemented in MATLAB (using ‘fft2’ and ‘conv2’ built-in functions), operating on double precision complex floating-point data on an Intel Core i5 laptop processor (described in Table 2.2.)

Table 2.2: Processor Parameters

Processor	OMAP 3530	Core i5
Manufacturer	Texas Instruments	Intel
System	Gumstix Overo (www.gumstix.org)	15" Macbook Pro (Apple Inc.)
Family/Variant	Cortex A8	Arrandale (mobile)
Cores	ARM + C64 DSP (unused)	2
Processor Clock	720 MHz	2.4 GHz
L1 memory	16KB(I)+16KB(D)	32KB(I)+32KB(D)/core
L2 memory	256KB	256KB/core
L3 memory	N/A	3MB
Main Memory Access	166MHz/32bit	1066MHz/64bit
Typical Power	1W	35W (Thermal Design Power)

2.E Results

2.E.i Focusing Phase Errors and Delay-and-Sum Interpolation Errors for Separable 2D Array Focusing

Fig. 2.3 shows apodization-weighted and unweighted RMS phase errors for the described simulated array with focal depths of 5mm, 15mm and 25mm, and varying f/# from 0.8 to 3.0. Weighted RMS phase errors are $< 1/32$ cycle for f/# $>$ 1.0, focal depth $<=$ 15 mm. Fig. 2.4 illustrates the effect of interpolation errors in simulated beamplots for non-separable and separable delay and sum algorithms focused at a depth of 15mm, with f/# = 1.4 and using windowed-sinc or B-spline based interpolation methods.

2.E.ii Simulated Point-Spread Functions and Beamplots

Figs. 2.5 and 2.6 illustrate simulated 2D point spread functions (PSFs) and beamplots for non-separable and separable, DSIQ and DAS focusing. Fig.

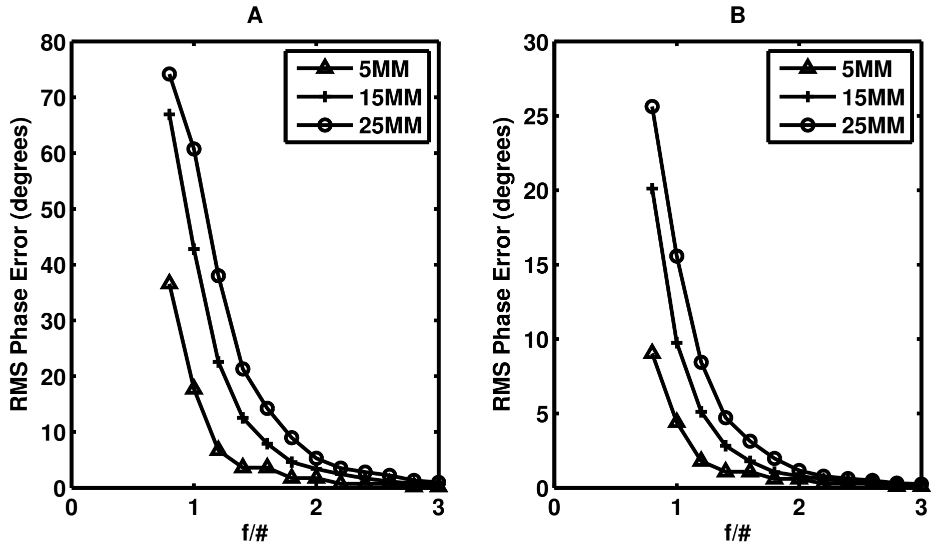

Figure 2.3: Unweighted (A) and apodization-weighted (B) RMS phase errors for separable focusing algorithm with varying f/# and focal depth = 5mm, 15mm or 25mm.

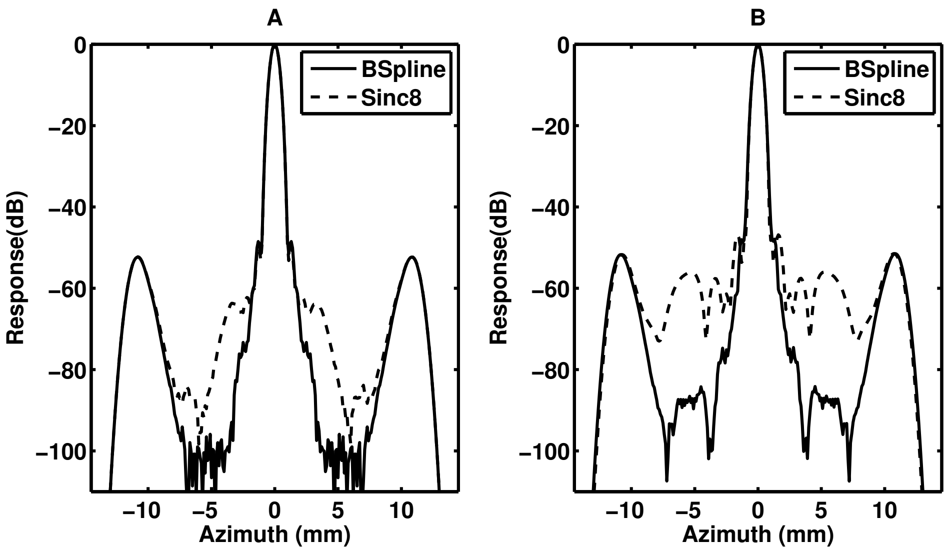

Figure 2.4: Simulated beamplots at focal depth of 15mm, f/#=1.4 for non-separable (A) and separable (B) delay-and-sum algorithms using 8-sample Hamming-windowed sinc interpolation (Sinc8) and cubic B-spline interpolation (BSpline).

2.7 shows simulated 2D PSFs for NDAS, SDAS and SDF focusing methods, in more difficult conditions including low f-number, shallow focal depth and increased excitation pulse frequency.

2.E.iii Simulated and Experimental Anechoic Cylinder Images

Experimental volume data was captured from the Sonic Window 60 x 60 array system, while positioned over a 10mm diameter anechoic cylinder at a depth of 15mm in a CIRS phantom. The phantom was also simulated in Field II and imaged for validation purposes using the NDAS, SDAS, NDF and SDF focusing methods. Fig. 2.8 shows C-mode image slices and lateral plots through the simulated anechoic cyst phantom using all 4 methods. Fig. 2.9 shows a B-mode image slice through the cyst (along the y=0 plane) for the same methods. Fig. 2.10 shows C-mode image slices and lateral plots through the experimental phantom anechoic cyst phantom for all 4 methods, using a 4-cycle transmit pulse.

Using the C-mode slice at the anechoic cylinder center, contrast-to-noise ratio (CNR) was calculated for simulated NDAS, SDAS, SDF and NDF focusing, giving values of 4.06 dB, 4.02 dB, 3.96 dB and 3.91 dB respectively. For the experimental data, the corresponding CNR values were 2.06 dB, 2.03 dB, 2.40 dB and 2.31 dB respectively. Simulated NDAS and SDAS cysts, using a 4-cycle transmit gave CNRs of 3.64dB and 3.53dB, reflecting that DAS has worse contrast than DSIQ when using a non-ideal 4-cycle pulse.

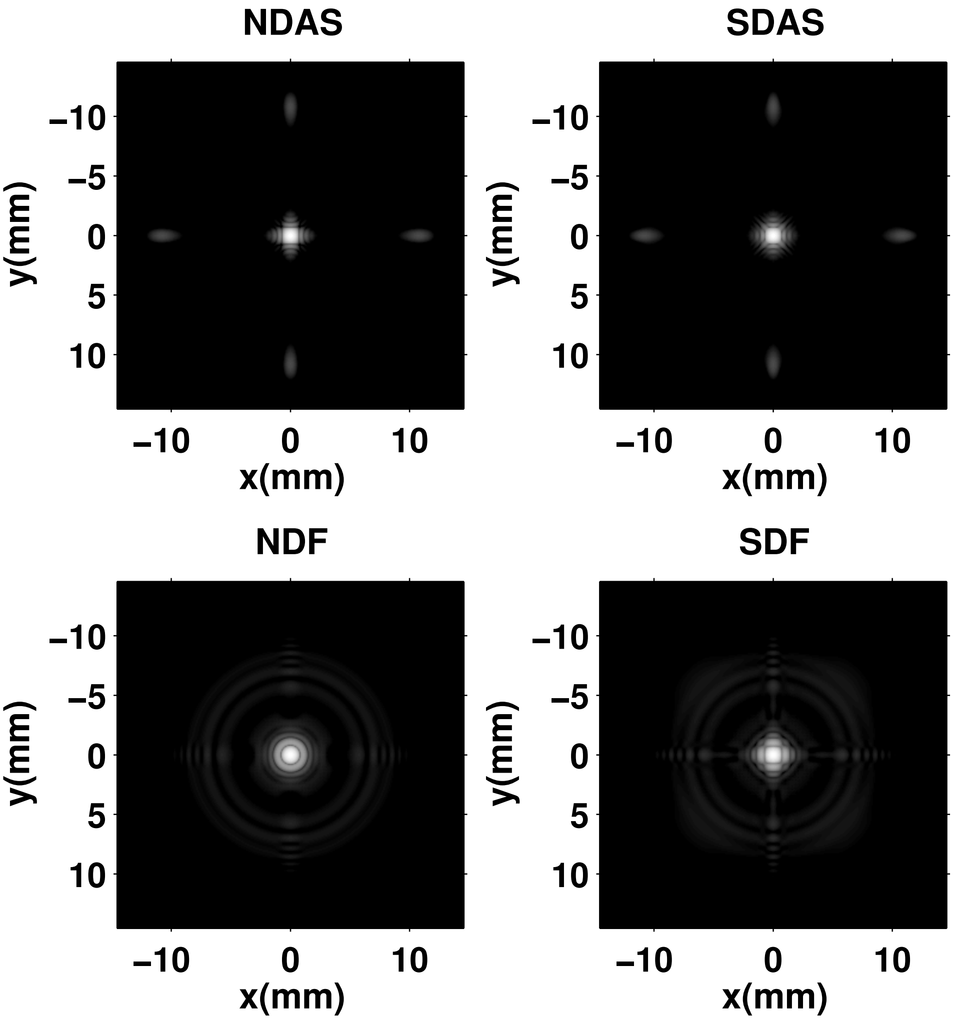

Figure 2.5: Simulated 2D elevational-azimuthal point-spread functions for separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming, using f/#=1.4 and focal depth of 15mm (65dB logarithmic display range, 2-cycle transmit for DAS, 4-cycle transmit for DSIQ).

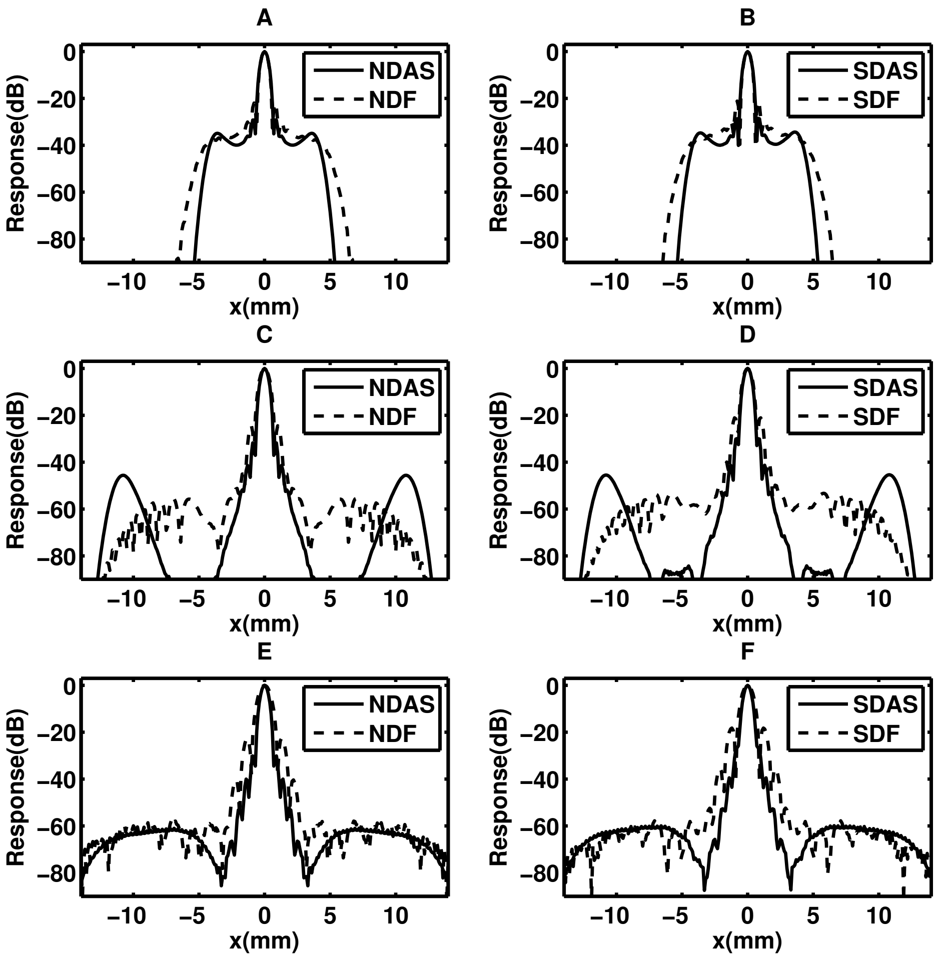

Figure 2.6: Simulated beamplots for non-separable and separable delay-and-sum and DSIQ beamforming, using f/#=1.4 and a focal depth of 5mm (A,B), 15mm (C,D) and 25mm (E,F).

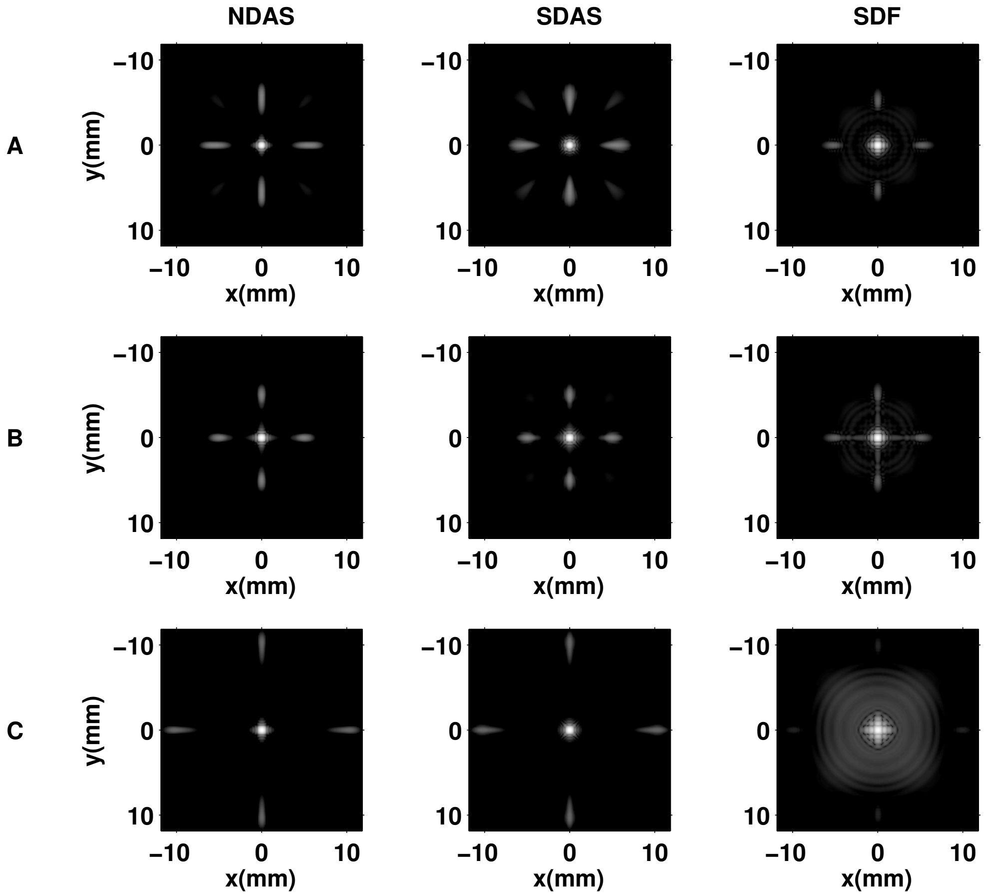

Figure 2.7: Point spread functions for the 2D array from Table 1, focused using the NDAS, SDAS and SDF methods, with logarithmic display range of 65 dB. Row A shows PSFs for f/#=0.8, focal depth=5mm, center frequency=5MHz. Row B shows f/#=1.0, focal depth=10mm, center frequency=5MHz. Row C shows f/#=1.4 focal depth=15mm, center frequency=7 MHz.

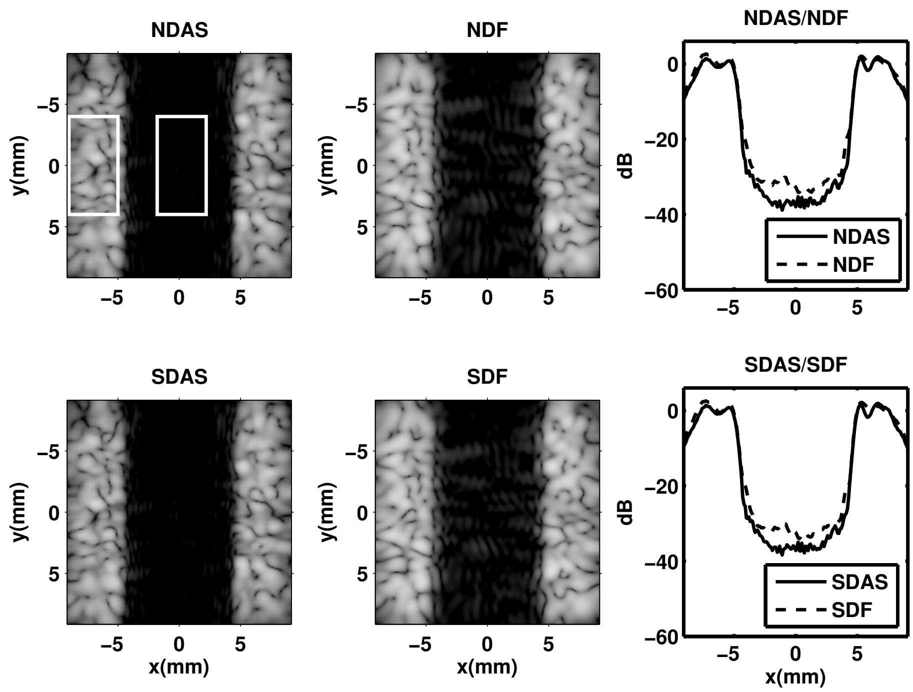

Figure 2.8: Simulated 10-mm diameter anechoic cylinder C-mode image slices (first 2 columns) and RMS absolute value profile (taken over y dimension). For each, separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming are shown, using f/#=1.4 and focal depth of 15mm (50 dB logarithmic display range). White rectangles represent areas used to estimate CNR values.

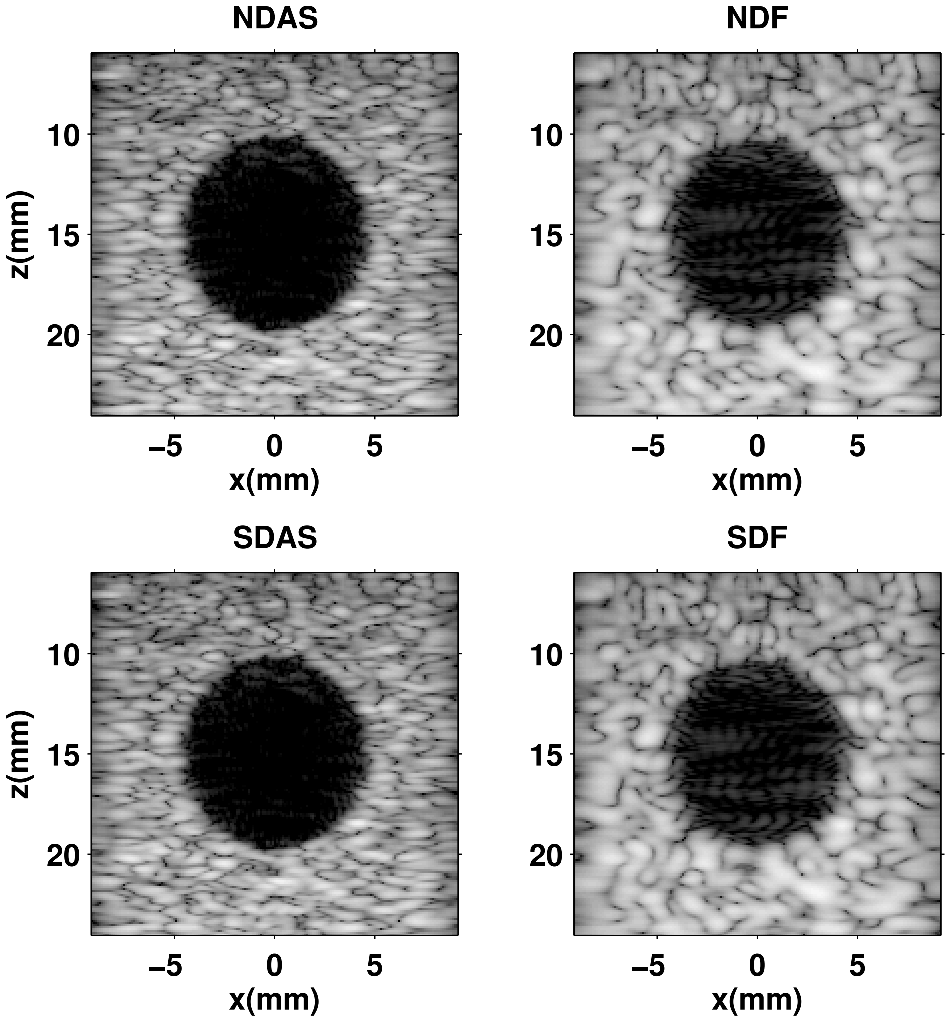

Figure 2.9: Simulated 10-mm diameter anechoic cylinder B-mode image slices for separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming, using f/#=1.4 (50 dB logarithmic display range).

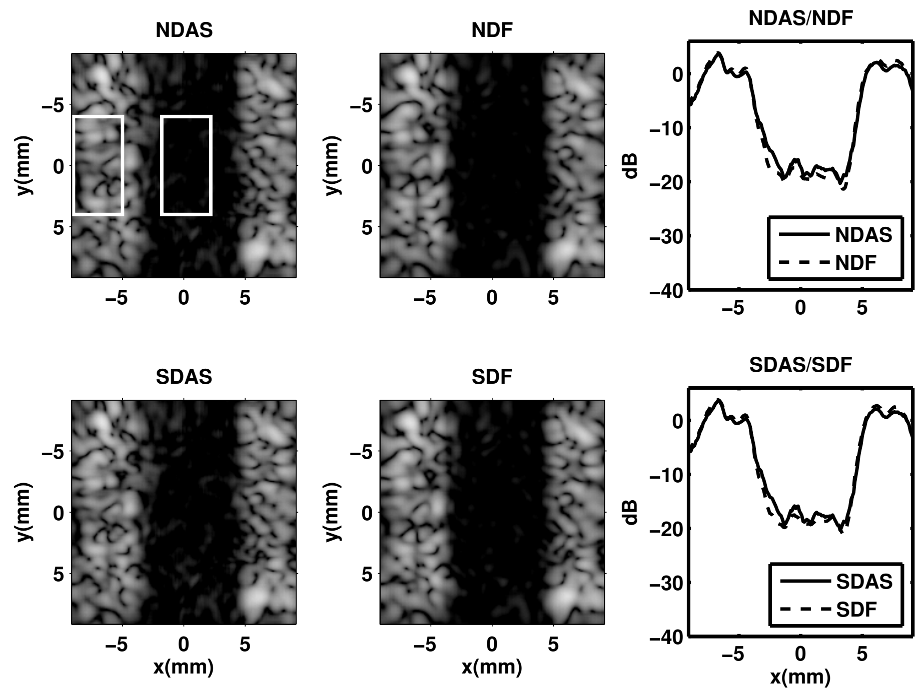

Figure 2.10: Experimental 10-mm diameter anechoic cylinder C-mode image slices (first 2 columns) and RMS absolute value profile (taken over y dimension). For each, separable and non-separable versions of delay-and-sum (NDAS, SDAS) and DSIQ (NDF, SDF) beamforming are shown, using f/#=1.4 and focal depth of 15mm (35 dB logarithmic display range). White rectangles represent areas used to estimate CNR values.

2.E.iv Computational Performance of Separable and Non-Separable 2D Focusing

The execution times of NDAS, SDAS, NDF and SDF focusing were compared on the two different hardware platforms of Table 2. Fig. 2.11-A shows the C-mode imaging frame rates achieved by 'C' implementations of the algorithms with inner-loop SIMD optimizations on the OMAP 3530 processor. SDAS and NDAS rates were 254.8 Hz and 16.3 Hz for 40 x 40 apertures, corresponding to an acceleration factor of 15.6X by using the separable method. For SDF and NDF, frame rates were 192.8 Hz and 11.39 Hz - i.e. an acceleration of 16.9 X by using SDF. Fig. 2.11-B shows the performance of NDF, SDF and 2D FFT-based focusing implemented in MATLAB on the Core i5 processor. Compared to the 2D FFT method for aperture sizes of 20 x 20 and 40 x 40, the SDF algorithm was faster by a factor of 2.12 and 1.64 respectively. The NDAS, SDAS, NDF and SDF algorithms had an energy cost of 75.0 mJ/frame, 4.8 mJ/frame, 107.2 mJ/frame and 6.3 mJ/frame respectively, using a 40 x 40 focusing aperture, and implemented in C with inner-loop SIMD optimizations on the OMAP hardware platform.

2.F Discussion

2.F.i Imaging Performance of Separable 2D Beamforming Algorithms

Separable 2D beamforming is only useful if imaging performance is not significantly degraded. The results from Fig. 2.3 indicate that the RMS phase errors due to separable focusing are much lower when weighted by the aper-

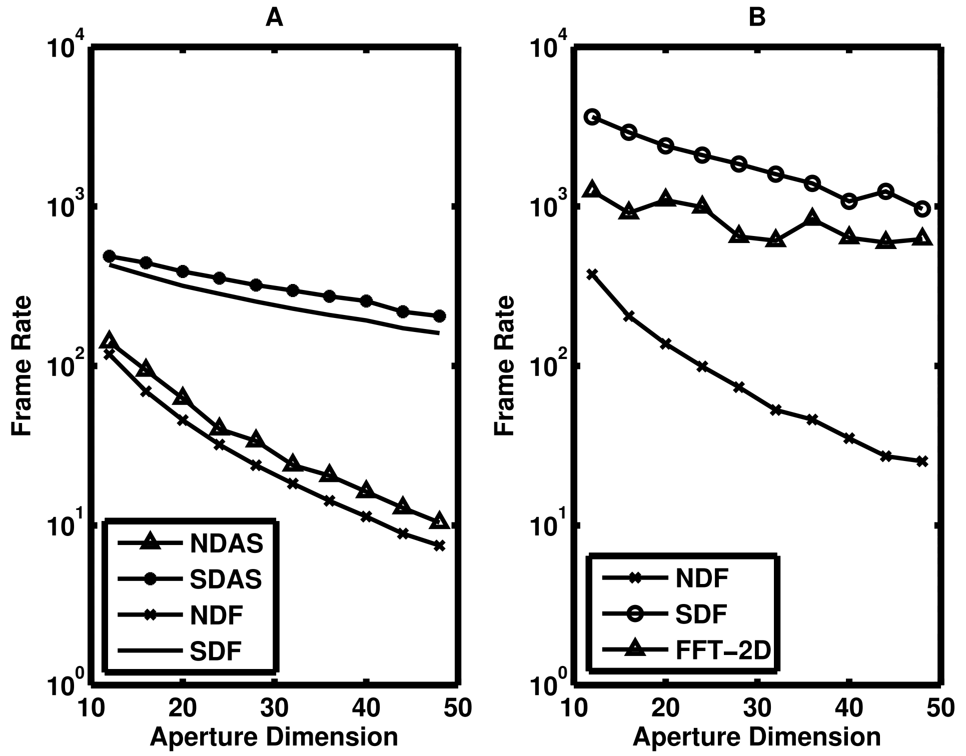

Figure 2.11: (A) C-Mode imaging frame rates for C-with-SIMD implementations of separable and non-separable delay-and-sum and DSIQ focusing on the OMAP 3530 processor (SDAS, SDF, NDAS, NDF). (B) Frame rates for NDF, SDF and FFT-based focusing implemented in MATLAB on the Intel Core i5 processor.

ture apodization function, but increase dramatically for f-numbers lower than about 1.4. In practice, the angular sensitivity of individual elements means that f-numbers below 1.0 are rarely used. For an operating region of f/# $>=$1.4, focal depth $<=$ 25 mm, the weighted RMS phase error is less than 5 degrees, which should not significantly affect focusing quality. It is also important that the delays applied first in the x-direction, then the y-direction, do not introduce cumulative errors. Windowed-sinc time delay interpolation with 8 taps has been shown in Fig. 2.4 to introduce significant beamplot degradation for separable delay-and-sum focusing. In contrast, cubic B-spline interpolation introduces only minimal degradation, and is nominally twice as fast as 8-tap sinc interpolation. The simulated beamplots and PSFs represented in Figs. 2.5 and 2.6 indicate that under typical conditions, separable versions of delay-and-sum and DSIQ focusing suffer minimal degradation over non-separable focusing. In addition, using the separable method, simulated cyst CNR is only reduced from 4.06 dB to 4.02 dB for 2-cycle transmit delay-and-sum, and from 3.96 dB to 3.91 dB for 4-cycle transmit DSIQ. This is a very minor, probably imperceptible, contrast reduction. Additionally, the difference between delay-and-sum and DSIQ contrast is similarly small. This indicates that the application of appropriate weightings to the two DSIQ complex samples significantly reduces the PSF energy outside the mainlobe and side lobes, approaching delay-and-sum contrast, albeit with a wider mainlobe and reduced resolution. Experimental cyst CNR values confirm that there are only marginal differences between NDAS, SDAS, NDF, SDF in this imaging case. The difference in experimental CNR magnitude compared to simulations can be attributed to the presence of distributed phase deficiencies (conservatively estimated at 14 ns RMS delay) across the surface of the prototype array (53). This is probably due to the viscous silver epoxy used for one electrode. Due to prototype limitations, only 4-cycle transmit data was available. This degrades delay-and-sum focusing and CNR relative to DSIQ, verified by delay-and-sum simulation results with a 4-cycle transmit pulse. The separable focusing decomposition is expected to perform worst in conditions of low f-numbers, in the extreme near field (due to increased wavefront curvature), and at operating frequencies where grating lobes are severe, as explored in Fig. 2.7. In all of these conditions, separable delay-and-sum is minimally degraded relative to non-separable delay-and-sum. When compared to delay-and-sum, separable DSIQ focusing only exhibits significant degradation in the high-operating frequency, grating lobe condition, but is remarkably robust in the other cases. Separable focusing performance is governed by the errors in the (separable) Fresnel approximation (48, 49) under typical imaging conditions. A square root expansion of this form converges more quickly when the term b from equation 2.7 is small. The variable b can be related to the f-number used in the system using equation 2.20, which for f/$\#$ $=$ 1.0 is approximately 0.354. For realistic apertures, the series converges rapidly, and the significance of later terms in the expansion falls quickly.

$$b = \left(\sqrt{2} \frac{X^2}{Z_f^2} \right),\ f = \frac{Z_f}{2X} \therefore b = \sqrt{2} \frac{1}{4f^2} = \frac{1}{2\sqrt{2}f^2} \tag{2.20}$$

Furthermore, the non-separable part of the third term in the expansion, $2(\Delta X^2 \Delta Y^2 / Z_f^2)$ is only significant compared to the separable part, $(\Delta X^2/Z_f^2 + \Delta Y^2/Z_f^2)$ in the corners of the aperture. As the corners of the aperture have reduced effective sensitivity due to apodization and element directivity, the approximation error is mitigated. Taking these factors into account, it is logical that the separable focusing approximation causes only minimal imaging performance degradation. In combination, the experimental and simulated results in conjunction with the theoretical justification for separable focusing performance indicate that the separable near-field approximation is functionally equivalent to non-separable in most cases. The subaperture methods suggested by (5, 30, 31, 32, 33, 34) were tested as described in (58), but with plane wave-transmit, receive-only focusing. However, detailed analysis is omitted due to severe grating lobes that degraded imaging performance.

2.F.ii Computational Performance of of Separable 2D Beamforming Algorithms

The decomposition of a 2D beamforming operation into two separable 1-D line array beamforming operations is capable of order-of-magnitude performance increases for near-field wideband 3D ultrasound imaging (44). This method involves applying varying time delays across the azimuthal dimension, followed by the application of further time delays to the azimuthally delayed data, operating across the elevational dimension. When the time delays are simple phase rotations, as in DSIQ focusing, it is trivial to apply the two delays as successive complex multiplications. However, when interpolation operations are used to sample time series at delays of up to tens of samples, this means a full, delayed time series history must be produced by the azimuthal focusing step before elevational focusing. Although the interpolations can be applied using short FIR filters at integer offsets for delay-and-sum focusing, the separable method requires a full time-series to be produced by the first 1D focusing step. For volume focusing, this represents unnecessary oversampling in the axial dimension, detracting from the performance gains from separable decomposition. In effect, this means the separable method can only be used productively with delay-and-sum for volume imaging modes, with reduced flexibility in axial image sampling. In contrast, separable DSIQ can focus volume data with arbitrary axial plane spacing, and form single C-mode slices in isolation. For handheld devices with limited power, DSIQ is a very effective way to use a 2D array for real-time imaging with multi-hour battery life. In addition to energy-efficient beamforming, front-end ASICs using the DSIQ sampling method use very little power due to a low-duty-cycle operating mode. In comparison to typical always-on ultrasound analog front-end integrated circuits, such as the Texas Instruments AFE 5807 (88 mW/channel at 40 MHz, 12-bit ADC) or the Analog Devices AD 9278 (also 88 mW/channel at 40MHz, 12-bit ADC), a typical DSIQ front end only requires 13.8 W per channel at 30 frames/second, or 1.6 mJ per frame to operate all 3600 channels for C-mode imaging (17). This represents less than 1/6000 of the power of the always-on front-ends, and approximately 1/5 of the typical energy cost (7.5 mJ) of the separable DSIQ beamformer processing. Separable 2D focusing has a theoretical computational cost reduction of (M+N)/2 compared to non-separable focusing, where M and N are the focusing aperture dimensions in elements. For typical 40 x 40 apertures a significant speed-up of 20X is predicted. When the separable algorithm was tested on realistic hardware (the OMAP 3530 processor), implemented in C with SIMD optimizations on 16-bit data, actual speed increases ranged between 57%-87%, and 61%-89% of predicted values for delay-and-sum, and DSIQ focusing respectively. In comparison, non-SIMD performance differs from predicted values by just 8% for delay-and sum, and 4% for DSIQ focusing. This indicates that when SIMD instructions are used, giving a 2-3X speed increase, for smaller apertures loop overhead becomes a performance bottleneck. An implementation of focusing using FFT-based 2D convolution on double precision data in MATLAB was 2.12 times slower compared to SDF for 20 x 20 apertures and 1.64 times slower for 40 x 40 apertures. The comparative analysis of (41) shows that for short (32-64 element) 1D apertures, FFT-based 1D convolution is comparable to non-FFT 1D convolution in computational cost. For 2D convolution, the FFT method computational cost increases to $O(N^2 log(N)$, while separable 2D focusing cost increases from $O(N^2)$ to $O((N+N)N2)$ or $O(2N^3)$. The zero-padding and significantly higher memory usage required for FFT-based focusing (compared to SDF focusing) plausibly explains the performance advantage of the separable method for typical data sizes. This effect is expected to be more pronounced in processors with smaller L1 cache memories, such as those likely to be used in handheld, battery operated devices. Although alternate approaches optimized for FFT, such as using FPGAs, specialized DSPs or ASICs are possible, the theoretical performance increases suggested by FFT-based 2D convolution are not necessarily achievable on low-power processors suitable for handheld devices. For an aperture size of 40 x 40, the separable algorithm increased frame rates from approximately 16 to 254 (delay-and-sum), and 11 to 193 (DSIQ) on the OMAP 3530. These are significant increases on the OMAP platform, where the 7.5 mJ/frame energy cost of C-mode 2D DSIQ array beamforming enables real-time, portable operation with multi-hour battery life.

2.G Conclusion

X-Y Separable 2D beamforming is capable of order-of-magnitude improvements in computation time and energy consumption. Our analysis of imaging performance using this method in practical conditions using delay and sum (DAS) and DSIQ beamforming indicates the separable approximation has minimal effect on imaging quality. Simulated and experimental cyst CNR values were reduced by negligible amounts in each example. A high quality interpolator is required in separable DAS beamforming, to prevent cumulative interpolation errors from degrading imaging performance. DSIQ focusing is capable of achieving contrast levels approaching those of DAS, if two complex sample planes are captured and weighted appropriately. Although mainlobe width and sidelobe levels with DSIQ are worse than DAS, separable DSIQ can be used to form C-mode images or volume images with arbitrary axial sampling. In contrast, separable delay-and-sum can only achieve large performance improvements when forming volume images, with additional axial sampling constraints due to the two-step focusing process. The measured performance gains of 15.6X and 16.9X for separable delay-and-sum and DSIQ focusing are very attractive, translating to frame-rates of 255 and 193 C-mode slices per second respectively. This enables real-time but low volume imaging frame rates in each case. Although DSIQ focusing has slightly reduced imaging performance relative to DAS, using the separable method it is capable of processing data from thousands of 2D array channels in real-time, with a low energy cost of 7.5 mJ/frame, enabling multi-hour operation in a handheld, battery-powered device.

Chapter 3

SECTOR-SCAN MOTION TRACKING FOR FREEHAND 3D IMAGING[1]

3.A Abstract

Ultrasound data motion tracking is widely used to both estimate tissue motion for diagnostic purposes, and to estimate bulk transducer motion relative to tissue, for example in freehand imaging (59), where successive 2D ultrasound scan planes are registered in a 3D volume. Speckle-tracking (60) and decorrelation-based methods (61) are used to estimate motion in the azimuthal and elevational planes. However, the performance of speckle-tracking is significantly degraded in sector-scan systems due to PSF rotation with lateral motion (62). Sector-scan systems are used for many applications, including cardiac, abdominal and obstetric imaging, typically using a electronically scanned beam. Mechanical sector-scans have also been shown to be advantageous

[1]Chapter 3 is being edited for submission as the journal article :
K. Owen, F. W. Mauldin, Jr., S. Nguyen, M. Tiouririne and J. A. Hossack, "Improved Elevational and Azimuthal Motion Tracking Using Sector Scans", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*

in bone imaging (18). When there is both azimuthal and elevational motion, the accuracy of estimates in both planes are reduced. In this chapter I develop a new method for joint azimuthal-elevational motion estimation based on the complex correlation of individual IQ-demodulated sector-scan A-lines arising from tissue motion in 3D space. Using a new statistical model, I show that the phase of per-line decorrelation is linearly related to the dot-product of the tissue motion vector and each A-line's direction unit vector, with a normal distribution in realistic conditions. A simple least-squares fit is used to find the maximum-likelihood azimuthal-plane tissue motion vector, given the correlation phases for all A-lines and transducer geometry. Next, the magnitude of each A-line correlation is used in conjunction with the calculated azimuthal motion to estimate elevational motion, using optimal and suboptimal, but fast methods. I show that our method has performance benefits over both speckle-tracking and decorrelation-based tracking for azimuthal and elevational motion estimation in sector-scan systems, with particular benefits when there is both elevational and azimuthal motion. Motion-tracking efficacy is further demonstrated by improved freehand imaging of a known target (anatomically accurate 3D-printed lumbar spine model) in a tissue-mimicking phantom, with a RMS surface distance error of 1.2 mm, compared to 2.43 mm for conventional methods. This data indicates that the new algorithm is capable of improved tracking performance for sector scan systems, enabling effective freehand 3D scanning.

3.B Introduction

Azimuthal and elevational motion of tissue relative to an ultrasound transducer can be estimated during scanning using various motion-tracking algorithms. In addition to diagnostic applications, (e.g. cardiac and vasculature imaging), transducer/tissue relative motion can be used to construct extended field-of-view images from individual scan-frames (63), (64), (65). When applied to conventional B-mode scanners, this approach enables freehand 3D imaging (59), or the production of volumetric composite images from many individual 2D image planes. Traditional B-mode ultrasound systems require substantial training and experience for proper image interpretation, due to the influence of transducer orientation on instantaneous images. In contrast, volumetric images can be more intuitively interpreted (3), (4) with consequent widespread use of high-end 2D phased array systems for volumetric imaging in cardiac and obstetric applications. However, for some applications, such as low-cost, portable imaging, freehand 3D imaging is a viable alternative to 2D phased array technology for intuitive volume imaging.

In-plane motion can be estimated using a variety of speckle-tracking methods, such as normalized cross-correlation (60), (66), sum-of-absolute differences, and variations on these techniques, which allow sub-sample estimation through curve fitting (67). When motion is entirely within the imaging plane, speckle-tracking motion estimation has low jitter, however out-of-plane motion, tissue compression or any rotational motion introduce decorrelation (62), leading to increased estimation error (68). The decorrelation related to out-of-plane motion can also be used to estimate motion magnitude, using inter-frame speckle decorrelation measurements fitted to the two-way point-spread-function (PSF) autocorrelation at a given shift, calibrated using a test phantom(61). This method has been extended to account for partially developed speckle (69), varying elevational motion with depth due to rotational motion, and decorrelation estimation using regression (70) or probabilistic models (71). Decorrelation cannot, however be used to estimate out-of-plane motion direction due to correlation function symmetry in the elevational dimension.

For sector-scan ultrasound systems, azimuthal motion tracking is degraded by rotation of the point-spread function with translation, with rotations of just 2-10 degrees causing rapid decorrelation (62). Sector scans can use either electronic or mechanical scanning and are prevalent in applications where large fields of view are required, and/or the physiological access window is limited, such as abdominal, cardiac and obstetric imaging. Although electronically scanned (phased-array) systems are more prevalent, mechanically scanned systems have some advantages which cannot easily be replicated in electronic systems, including improved elevational focusing (due to piston transducer), complete lack of grating lobes at any angle, and low complexity electronics. In some applications, such as spinal bone imaging (18) and extremely portable systems (e.g. SeeMore USB probes - Interson, Pleasanton, CA), these advantages make mechanical systems preferable.

In this chapter I develop a method to jointly estimate azimuthal and elevational motion using complex decorrelation data from individual A-lines of a sector scan. This technique outperforms speckle-tracking and decorrelation-based tracking when motion has non-zero azimuthal and elevational components, has lower bias for azimuthal motion estimation, and is more tolerant of specular features in the field of view. A statistical model for the complex correlation of a focused, IQ-demodulated RF pixel indicates that the phase of the complex decorrelation of an A-line is normally distributed, with mean linearly related to the projection of tissue motion onto the A-line direction unit vector. Using the property that each A-line makes a different angle with tissue motion, I construct an accurate, maximum-likelihood azimuthal motion estimator using only phase data. I further show that the magnitude of decorrelation for each A-line is related to the magnitude of motion orthogonal to the A-line direction, and using a separable 3D gaussian-magnitude PSF model, develop one maximum-likelihood elevation motion estimator, and two suboptimal but fast elevational motion estimators.

The statistical correlation model is validated using stochastic simulations and comparison to experimental data using a mechanically-scanned piston transducer. Performance of the conventional (speckle-tracking plus decorrelation elevational tracking) and proposed tracking algorithm is compared over a range of azimuthal/elevation translations using a motion stage and tissue-mimicking phantom (72). The statistical model is also used to successfully predict azimuthal speckle-tracking motion estimation bias. Finally, I test motion estimation performance using freehand scanning of a known 3D target (anatomically accurate 3D-printed lumbar spine model embedded in tissue-mimicking phantom), evaluated using root-mean-square bone surface distance error after bone surface detection and 3D registration.

3.C Theory

3.C.i Statistical Model for Complex PSF Correlation with Motion

A single pixel from a focused, IQ-demodulated A-line has a sensitivity to scatterers that is a complex function of 3D space, shown as $F(x, y, z)$ in Fig. 3.1A. The origin of this co-ordinate system is at the center of the function F, corresponding to the location of the pixel in space, the z-axis is in the direction of propagation of the A-line, and the x-axis and y-axis are orthogonal to the z-axis, and arranged to form a right-handed set. The azimuthal plane corresponds to the y-z plane in this system, while elevational motion is in the x-axis direction. If there is subsequently transducer motion relative to tissue, with vector $\mathbf{r}$, the new scatterer sensitivity for the same complex pixel in the A-line is now given by $G(x, y, z) = F(x - r_x, y - r_y, z - r_z)$, a shifted version of the original sensitivity, function, shown in Fig. 3.1B.

The product of the two pixel values is the raw correlation value for the pixel as a function of motion, or $C(\mathbf{r})$. The scatterer function $a(x, y, z)$ describes the scatterer amplitudes in space, a set of scalar, independent, identically distributed (i.i.d) variables with zero mean and a normal distribution. Using this, an expression for the raw complex correlation $C(\mathbf{r})$ can be developed using integrals over 3D space as in equation 3.1. This corresponds to the inner product of the two complex PSFs F and G. A simpler representation of $C(\mathbf{r})$ can be produced using linear algebra (equation 3.2), where the 3D functions F and G are represented by column vectors, with each element corresponding to a single (x, y, z) permutation. The linear algebra approach is simpler, as the 3D integrals become complex inner products, and the six-dimensional integration

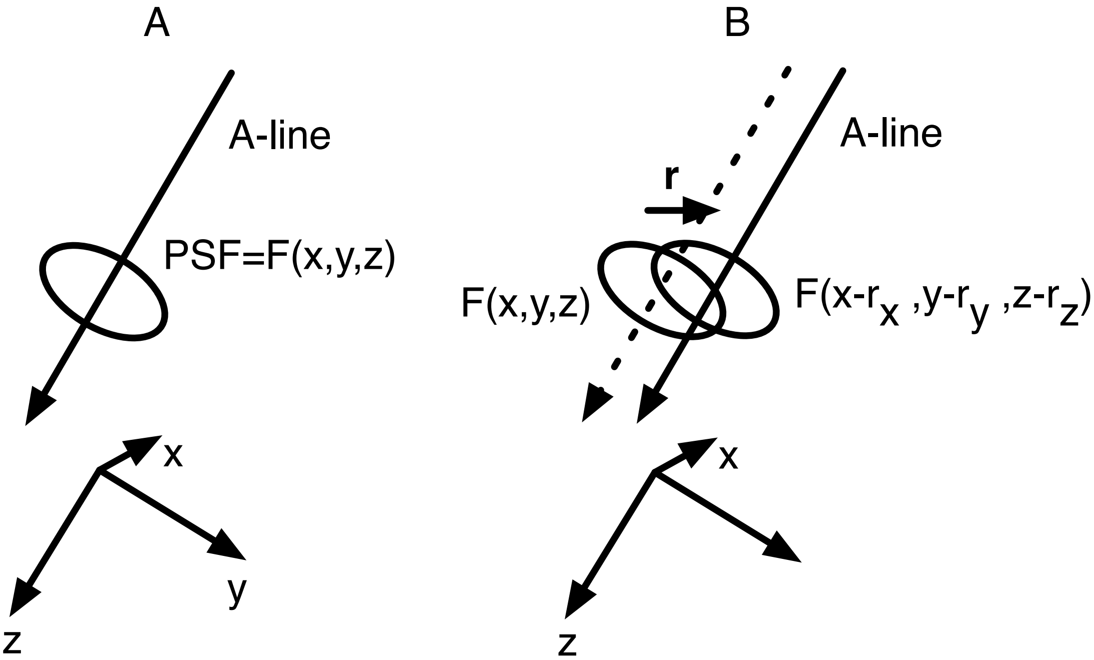

Figure 3.1: Sensitivity to scatterers of a single pixel from a focused complex demodulated A-line as a function F(x,y,z) of three dimensional space shown in (A). After transducer motion of vector r, the new scatterer sensitivity is shown in (B), as the function $\mathbf{F(x - r_x, y - r_y, z - r_z)}$.

required for stochastic analysis of equation 3.1 is avoided. In the limit, if 3D space is discretized finely enough, equations 3.1 and 3.2 are equivalent, but more importantly, matrix factorizations can be used to reduce the problem to the minimum equivalent representation.

$$\begin{aligned} C(\mathbf{r}) = \int\int\int F(x,y,z)a(x,y,z)\,dx\,dy\,dz \\ \int\int\int G^*(x,y,z,\mathbf{r})a^*(x,y,z)\,dx\,dy\,dz \end{aligned} \tag{3.1}$$

$$C(\mathbf{r}) = F^T a a^H G^*(\mathbf{r}) \tag{3.2}$$

The column vectors F and G (of size N x 1, for N discretized scatterers), can be transposed, conjugated and concatenated to form P, an N x 2 matrix as in equation 3.3, giving expressions f and g for the complex pixel value before and after translation. Following this, $P^H a a^H P$ is a 2 x 2 matrix with $C(\mathbf{r})$ as the top-right element and $C(\mathbf{r})^*$ in the bottom-left element, as in equation 3.4. The matrix P can be factorized using the singular value decomposition (SVD), and expressed as $P = USV^H$, or $P^H = VSU^H$, where U is an N x 2 matrix with orthonormal columns, S is a 2 x 2 real matrix with non-zero elements only on the diagonal, and V is a 2 x 2 unitary matrix. Now, the term $P^H a$ can be replaced with $VSU^H a$, where U^H is a complex matrix with orthonormal rows that picks up a, with the distribution $a \sim \mathcal{N}(\mathbf{0}, \sigma_a^2 \mathbf{I})$. Therefore, if $U^H a$ can be replaced with a new, 2 x 1 random variable, B with the same statistical properties as $U^H a$, then in addition to a drastic reduction in problem dimensions, the 2 x 2 matrix VS and the statistics of B will be sufficient to completely describe correlation statistics.

$$P = \begin{pmatrix} F^* & G^*(\mathbf{r}) \end{pmatrix}, \begin{pmatrix} f \\ g \end{pmatrix} = P^H a = \begin{pmatrix} F & G(\mathbf{r}) \end{pmatrix}^T a \tag{3.3}$$

$$P^H a a^H P = \begin{pmatrix} f \\ g \end{pmatrix} \begin{pmatrix} f^* & g^* \end{pmatrix} \begin{pmatrix} ff^* & C(\mathbf{r}) \\ C(\mathbf{r})^* & gg^* \end{pmatrix} \tag{3.4}$$

The term $U^H a$ can be decomposed into real and imaginary components, $U^H a = U_R^H a - jU_I^H a$. As the scatterer distribution is given by $a \sim \mathcal{N}(\mathbf{0}, \sigma_a^2 \mathbf{I})$, and the complex demodulation process ensures that U_R and U_I are orthogonal with equal magnitude, it is clear that $U^H a$ is the sum of independent real and imaginary 2 x 1 real, zero-mean, i.i.d. normally distributed variables. Therefore $U^H a$ can be substituted by the random variable B as in equation 3.5. The properties of the matrices V (unitary), and S (real, diagonal), properties of the complex PSFs F and G, and relations between VS and the complex PSF inner products $F^H F$, $G^H G$ and $F^H G$ can be used to calculate V and S in terms of a phase angle ϕ and the scalar values S_1 and S_2 (equation 3.6, detailed derivation in Appendix A). This result enables the effect of PSF transformations on the complex correlation to be examined analytically in the complex plane, and the small size of the VS matrix (2 x 2) makes stochastic simulations with large numbers of realizations computationally feasible.

$$U^H a \equiv B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, \frac{1}{2}\sigma_a^2 \mathbf{I}) + j\mathcal{N}(\mathbf{0}, \frac{1}{2}\sigma_a^2 \mathbf{I}) \tag{3.5}$$

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ e^{-j\phi} & -e^{-j\phi} \end{pmatrix}, \;\; S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}$$
$$\phi = -arg\left(\frac{F^H G}{F^H F}\right), \quad \frac{S_2}{S_1} = \sqrt{\frac{1 - |\frac{F^H G}{F^H F}|}{1 + |\frac{F^H G}{F^H F}|}} \tag{3.6}$$

The term fg^* from equation 3.4 is the correlation between the complex signals picked up by the two PSFs F and G, with G shifted or otherwise transformed relative to F. Equation 3.7 shows fg^* extracted from the expansion of $VSBB^HSV^H$. The first term in this relation is the weighted sum of two real, chi-squared distributed variables, $S_1^2(b_1b_1^*) \sim S_1^2\chi_2^2$ and $S_2^2(b_2b_2^*) \sim S_2^2\chi_2^2$, rotated by the angle (ϕ) in the complex plane. The second term can be simplified, if the term $(b_2b_1^* - b_1b_2^*)$ is rewritten $|b_1||b_2|(e^{j\theta} - e^{-j\theta})$, then using the Euler relations as $2j|b_1||b_2|sin(\theta)$, where θ is the uniformly distributed angle between independent complex normally distributed variables b_1 and b_2. Equation 3.8 shows the simplified form, where the rightmost term is a modified Bessel distribution of the second kind, altered by the $sin(\theta)$ factor, and is orthogonal to the first, chi-squared term in the complex plane.

$$fg^* = \left(S_1^2 b_1 b_1^* + S_2^2 b_2 b_2^*\right) e^{-j\phi} + S_1 S_2 \left(b_2 b_1^* - b_1 b_2 *\right) e^{-j\phi} \tag{3.7}$$

$$fg^* = e^{-j\phi} \left[\left(S_1^2 b_1 b_1^* + S_2^2 b_2 b_2^*\right) + 2jS_1S_2|b_1||b_2|sin(\theta)\right] \tag{3.8}$$

3.C.ii Statistics of Normalized Complex Correlation with PSF Transforms

Correlation data for motion detection is typically taken from many pairs of focused RF pixel data, mean-reduced then normalized to reduce the effect of localized intensity variations, using equation 3.9. Normalization effectively averages many individual RF pixel correlations $f_i g_i^*$, assumed to be identical, aside from local intensity variations. The numerator of equation 3.9 is ideally the sum of many independent realizations of equation 3.8. In practice, the individual RF pixel correlations from equation 3.9 will be partially correlated with nearby pixels due to PSF overlap, and the PSF will change over a large enough correlation kernel area. However, even with a relatively small number of independent RF pixels, both the chi-squared and modified Bessel (of the 2nd kind) terms from equation 3.8 will approach normal distributions due to the central limit theorem (CLT) (73).

$$C(\mathbf{r}) = \frac{\sum_i (f_i - \overline{f})(g_i(\mathbf{r}) - \overline{g})^*}{\sqrt{\sum_i (f_i - \overline{f})(f_i - \overline{f})^* \sum_i (g_i(\mathbf{r}) - \overline{g})(g_i(\mathbf{r}) - \overline{g})^*}} \tag{3.9}$$

In practice, normalized correlations are not normally distributed, due to the non-linear, non-independent effects of the denominator of equation 3.9. In the absence of a theoretical correlation distribution model, other work has used maximum entropy methods (71), (74) and more general experimentally determined distributions (70) in motion estimation applications. However, as the normalizing denominator is scalar, it can only affect the complex correlation magnitude, but not the phase of a complex correlation, which due to equation 3.8, the CLT and small angle approximation is effectively normally distributed. Therefore, a motion estimator using just the phase of complex correlations can implement maximum likelihood motion estimation using very simple, least-squares techniques. In contrast, normalized correlation magnitude has a complicated distribution with no known analytical model.

3.C.iii Sector Scan Geometric Model and Azimuthal Motion Estimation

In sector scan systems, each A-line typically has a PSF at a given depth that is identical to those of the other A-lines, apart from a rotation in the azimuthal plane. It is convenient to express the PSF function for an individual A-line as a function of a local co-ordinate system, described by the unit vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$, where $\hat{\mathbf{k}}$ is in the A-line direction, $\hat{\mathbf{j}}$ is orthogonal in the azimuthal plane, and $\hat{\mathbf{i}}$ is in the elevation direction. The local co-ordinate systems for multiple A-lines in a sector-scan system are shown in Fig. 3.2 The scatterer sensitivity of the n-th A-line at a displacement given by vector $\vec{\mathbf{p}}$ from the PSF center can be written as $F(\vec{\mathbf{p}} \cdot \hat{\mathbf{i}}, \vec{\mathbf{p}} \cdot \hat{\mathbf{j}}, \vec{\mathbf{p}} \cdot \hat{\mathbf{k}})$. If the tissue is then displaced by the vector $\vec{\mathbf{r}}$, the new scatterer sensitivity is $F((\vec{\mathbf{p}} + \vec{\mathbf{r}}) \cdot \hat{\mathbf{i}}, (\vec{\mathbf{p}} + \vec{\mathbf{r}}) \cdot \hat{\mathbf{j}}, (\vec{\mathbf{p}} + \vec{\mathbf{r}}) \cdot \hat{\mathbf{k}})$. Equation 3.6 can be used to calculate the phase and other parameters resulting from a given displacement vector $\vec{\mathbf{r}}$. A separable analytical representation for the function F as a product of spatial gaussians in the lateral and elevational directions, and a gaussian-windowed complex sinusoid in the axial direction is shown in equation 3.10, parameterized by α, β, γ and λ.

$$F(x, y, z) = e^{-\alpha x^2} e^{-\beta y^2} e^{-\gamma z^2 + j\frac{4\pi}{\lambda} z} = e^{-(\alpha x^2 + \beta y^2 + \gamma z^2 - j\frac{4\pi}{\lambda} z)} \quad (3.10)$$

This allows an analytical evaluation of the $F^H G$ term from equation 3.6

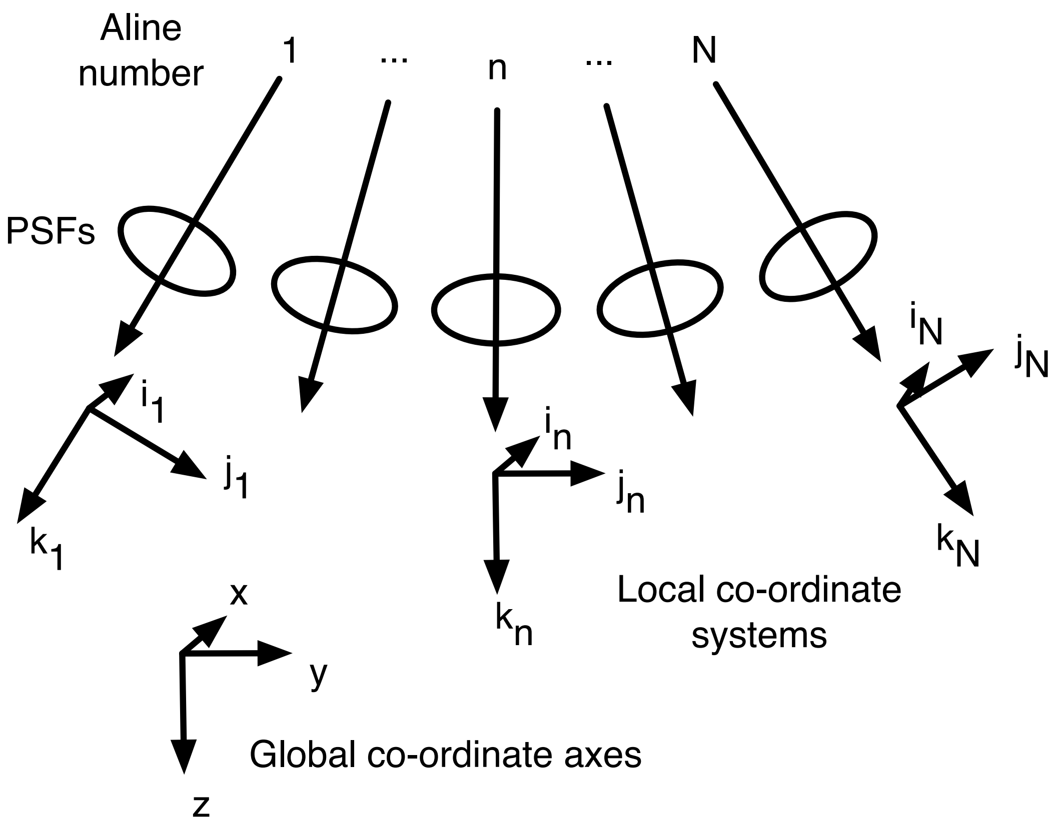

Figure 3.2: Multiple A-lines from a sector scan, showing the scatter sensitivity function or PSF at the same depth for each. Also shown is the local co-ordinate system for each A-line, with the unit vectors k_n in the A-line direction, j_n orthogonal to k_n in the azimuthal plane, and i_n in the elevational direction. Local co-ordinate axis vectors are defined in the global co-ordinates x, y, z, and centered on the PSF (shown offset for clarity).

using integrals, with a solution that is a product of a phase and magnitude term, the magnitude term itself also a product of functions of motion in the $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ directions (equation 3.11). This equation forms the basis for a maximum likelihood estimator of azimuthal-plane motion using phase alone.

$$
\begin{aligned}
\langle F, G(\mathbf{r}) \rangle = \int\int\int_V & e^{-\alpha[(\mathbf{p}\cdot\mathbf{i})^2+((\mathbf{p}-\mathbf{r})\cdot\mathbf{i})^2]} e^{-\beta[(\mathbf{p}\cdot\mathbf{j})^2+((\mathbf{p}-\mathbf{r})\cdot\mathbf{j})^2]} \\
& e^{-\gamma[(\mathbf{p}\cdot\mathbf{k})^2+((\mathbf{p}-\mathbf{r})\cdot\mathbf{k})^2]} e^{j\frac{4\pi}{\lambda}(\mathbf{p}\cdot\mathbf{k}-(\mathbf{p}-\mathbf{r})\cdot\mathbf{k})}\, d\mathbf{p} \\
= & e^{j\frac{4\pi}{\lambda}(\mathbf{r}\cdot\mathbf{k})} \sqrt{\frac{\pi^3}{8\alpha\beta\gamma}} e^{-\frac{1}{2}[\alpha(\mathbf{r}\cdot\mathbf{i})+\beta(\mathbf{r}\cdot\mathbf{j})+\gamma(\mathbf{r}\cdot\mathbf{k})]}
\end{aligned}
\tag{3.11}
$$

To estimate azimuthal translation using the N x 1 vector of complex correlation phases $\mathbf{\Phi}$, a multidimensional normally distributed phase is assumed, fully described by the covariance C_Φ. The phase vector resulting from a given azimuthal translation $\mathbf{r}_{az}$ is given by $\mathbf{\Phi}(\mathbf{r}_{az})$ in equation 3.12, where A is a 2 x N matrix containing the y and z components of the vector k_n for each A-line. This can be used to express the inverse-covariance weighted sum of squared errors between the observed phase vector $\mathbf{\Phi}$ and the phase vector resulting from $\mathbf{r}_{az}$, as in equation 3.13, with scalar distance r and phase Φ related by $r = \frac{\lambda}{4\pi}\Phi$, where λ is effective wavelength, and 4π is used instead of 2π due to round-trip wave propagation. The maximum likelihood azimuthal motion estimate, given the geometric model and correlation phase statistics is found by setting the differential of the error with respect to r_{az} to zero, giving the expression of equation 3.14, with the resulting azimuthal motion estimate covariance of equation 3.15. Due to the multidimensional normal distribution of the phase noise vector, this estimator is not only maximum likelihood, but is also an *effi-*

cient estimator in that it attains the CRLB and hence is the minimun-variance unbiased estimator (75). The complex correlation phase covariance C_Φ has a deterministic relationship with the magnitude of complex correlation for each A-line. This can be used to estimate phase covariance for each A-line, or the covariance can be set to identity for computational simplicity.

$$\Phi(\mathbf{r}_{az}) = A\mathbf{r}_{az} = \begin{pmatrix} k_{y1} & k_{y2} & \cdots & k_{yN} \\ k_{z1} & k_{z2} & \cdots & k_{zN} \end{pmatrix} \begin{pmatrix} r_y \\ r_z \end{pmatrix}_{az} \tag{3.12}$$

$$E = \left(\mathbf{r}_{az}^T A - \frac{\lambda\Phi^T}{4\pi}\right) (\frac{4\pi}{\lambda})^2 C_\Phi^{-1} \left(A^T \mathbf{r}_{az} - \frac{\lambda\Phi}{4\pi}\right) \tag{3.13}$$

$$\mathbf{r}_{az(opt)} = \left(AC_\Phi^{-1}A^T\right)^{-1} AC_\Phi^{-1}\frac{\lambda\Phi}{4\pi} \tag{3.14}$$

$$cov(\mathbf{r}_{az}) = (\frac{\lambda}{4\pi})^2 (AC_\Phi^{-1}A^T)^{-1} \tag{3.15}$$

3.C.iv Sector Scan Elevational Motion Estimation

Once the phase of the complex correlations from all A-lines has been used to estimate the azimuthal motion vector, this can be used in conjunction with the per-A-line correlation magnitude $|c|$ and Bayes' theorem (75) to estimate the non-azimuthal-plane motion component for each A-line as in equation 3.16. Here, R_j and R_k are the projections of global azimuthal motion onto an A-line's $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ unit vectors (from Fig. 3.2), r_i is the A-line's non-azimuthal motion component and $P(r_i)$ and $P(R_j, R_k, |c|)$ are assumed to be constant (flat prior distributions, although incorporation of other prior distribution is easily possible). In equation 3.17 we introduce the actual underlying PSF correlation c_0, and integrate over all possible values weighted by $P(|c|\big||c_0|)$ a 2D probability distribution which can be characterized stochastically using equations 3.5 and 3.6. This is needed to deterministically relate r_i, r_j and r_k as in equation 3.18. In equation 3.19, the left hand side fixes r_i, and in the inner integral r_j and r_k are fixed, leading to an expression for $P(R_j, R_j, |c_0||r_i)$ with fixed r_j that is the product of two gaussian distributions in r_j and r_k, with standard deviations $\sigma_j(|c_0|)$, $\sigma_k(|c_0|)$ derived using equation 3.15. In equation 3.19, $P(r_i\big|R_j, R_k, |c|)$ is proportional to, rather than equal to the right hand side due to the constant prior terms, which do not affect the probability maximization. The variable r_j is needed in the in the inner integral to enable r_k to be found using equation 3.18, but is integrated away, as r_i is the only variable of interest. The value $\hat{r_i}$ from equation 3.20 is the value of r_i which maximizes $P(r_i|R_j, R_k, |c|)$, and is the maximum likelihood non-azimuthal motion for this A-line.

$$P(r_i\big|R_j, R_k, |c|) = \frac{P(R_j, R_k, |c|\big|r_i)P(r_i)}{P(R_j, R_k, |c|)} \tag{3.16}$$

$$P(R_j, R_k, |c|\big|r_i) = \int_0^1 P(|c|\big||c_0|)P(R_j, R_k, |c_0|\big|r_i)d|c_0| \tag{3.17}$$

$$|c_0| = e^{-\frac{1}{2}(\alpha r_i^2 + \beta r_j^2 + \gamma r_k^2)}, \quad r_k\big||c_0|, r_j, r_i = \sqrt{\frac{-2ln(|c_0|) - \alpha r_i^2 - \beta r_j^2}{\gamma}} \tag{3.18}$$

$$P(r_i\big|R_j, R_k, |c|) \propto \int_0^1 P(|c|\big||c_0|) \int_{-\infty}^{\infty} e^{-\frac{(R_j - r_j)^2}{2\sigma_j^2(|c_0|)}} e^{-\frac{(R_k - r_k)^2}{2\sigma_k^2(|c_0|)}} dr_j \; d|c_0| \tag{3.19}$$

$$\hat{r_i} = \underset{r_i}{\mathrm{argmax}} \int_0^1 P(|c|\big||c_0|) \int_{-\infty}^{\infty} e^{-\frac{(R_j - r_j)^2}{2\sigma_j^2(|c_0|)}} e^{-\frac{(R_k - r_k)^2}{2\sigma_k^2(|c_0|)}} dr_j \ d|c_0| \tag{3.20}$$

The maximum-likelihood estimator described here is computationally complex, involving a 2D integration repeated for a range of r_i values. Iterative maximization methods to find $\hat{r_i}$ from equation 3.20 exist, including the Newton-Raphson method (76), and the Expectation-Maximization algorithm (77). However, knowledge of derivatives (76) or some manual data partitioning (77) is required, and convergence is not always guaranteed. In contrast, with some assumptions, simpler and faster suboptimal estimators can easily be constructed. If we assume that the measured correlation magnitude is equal to the underlying correlation magnitude, or $P(|c|\big||c_0|) = \delta(|c|)$, then the 2D integration of 3.19 becomes a 1D integration, as in equations 3.21, 3.22. This can be further simplified, by additionally assuming highly accurate azimuthal estimation, or that $r_j = R_j$ and $r_k = R_k$, which, using equation 3.18 leads to an expression for $\hat{r_i}$, shown in equation 3.23.

$$P(r_i\big|R_j, R_k, |c|) = \int_{-\infty}^{\infty} e^{-\frac{(R_j - r_j)^2}{2\sigma_j^2(|c|)}} e^{-\frac{(R_k - r_k)^2}{2\sigma_k^2(|c|)}} dr_j \tag{3.21}$$

$$\hat{r_i} = \underset{r_i}{\mathrm{argmax}} \int_{-\infty}^{\infty} e^{-\frac{(R_j - r_j)^2}{2\sigma_j^2(|c|)}} e^{-\frac{(R_k - r_k)^2}{2\sigma_k^2(|c|)}} dr_j \tag{3.22}$$

$$\hat{r_i} = \sqrt{\frac{-2ln(|c|) - \beta R_j^2 - \gamma R_k^2}{\alpha}} \tag{3.23}$$

3.C.v Relationship between Normalized Correlations for Complex RF and Detected Data

The correlation statistics for a single pixel of focused complex RF data undergoing a PSF transform (shift, rotation etc) are given in equation 3.8 to be the sum of two parts that are distributed along two orthogonal directions in the complex plane. For before and after complex PSFs with an inner product of the form $Ae^{-j\Phi}$, the first part has a chi-squared distribution along the direction of $e^{-j\Phi}$, while the second part is oriented along the $je^{-j\Phi}$ direction with a distribution very similar to a modified Bessel distribution (2nd kind), with an extra sine factor. When the complex correlation is normalized, the CLT ensures that the expected value of the correlation is equal to the inner product of the two complex PSFs. When the correlation is detected before normalization, this is not necessarily true. However, some conclusions can be made about detected correlations based upon the underlying complex correlation statistics :

1. Changing the angle Φ does not affect the detected correlation, or normalized detected correlation in any way, as it does not affect the magnitude of any detected parameter.

2. Even with a PSF transformation that gives a real-only complex PSF inner product, the uncorrelated part of the before/after PSFs will introduce imaginary components to the complex correlation corresponding to the orthogonal modified Bessel-like part of the distribution (equation 3.8). This illustrates the intrinsic complex stochastic nature of complex PSF correlations in all cases.

3. For a given underlying complex correlation (complex PSF inner product) and equal normalized complex correlation expected value, there is a corresponding expected value for the normalized detected correlation that can be expressed purely as a function of complex correlation magnitude. This follows from (1), and can be characterized by stochastic simulations using the SVD-based model of equation 3.6.

Using the deterministic relation between expected value of normalized-complex and normalized-detected correlations, the effect of various PSF transformations on normalized detected correlation can be easily determined using stochastic simulations.

3.C.vi Speckle-Tracking Motion Bias In Sector-Scan Systems

Speckle-tracking is typically used for motion detection in the azimuthal plane by calculating the normalized correlation between focused, envelope detected A-line pixels in a 2D kernel region as it undergoes various discrete 2D shifts across a larger 2D window region (60), (62). If the PSF at each pixel in each A-line is identical, then the correlation with shifts is a highly averaged, sampled version of the underlying 2D continuous correlation function. Fig. 3.3 shows the PSF of a single pixel from an A-line in, for example a linear array system where after a tissue shift, the correlation peak with lateral shift will fall on or between adjacent A-lines with identical PSFs, without fundamental bias. Expanding to use a window of many pixels simply averages the correlation curve, and interpolation can be used for sub-sample motion estimation (78).

However, in sector scan systems, tracking of azimuthal motion is degraded by rotation of the PSF with translation, with rotations of only 2-10° causing rapid decorrelation (62). After a lateral tissue shift in a sector-scan system, the tissue detected by a PSF from a single pixel from an A-line will now be detected by several other A-lines, each with rotated PSFs, compared to the original, as illustrated in Fig. 3.4. The rotation has two effects the PSF can rotate out of the original tissue area, and as the rotation increases, the original and rotated PSF become more spatially orthogonal with respect to the underlying complex representation, with a deterministic relation to the detected data in correlation. These effects lead to a correlation peak that underestimates lateral motion, shown qualitatively in Fig. 3.4. Additionally the height of the correlation peak will be lower due to rotation than with non-sector scanned systems, leading to larger variance in motion estimates (68). The model and observations of previous sections enable qualitative and quantitative predictions of the magnitude of these effects.

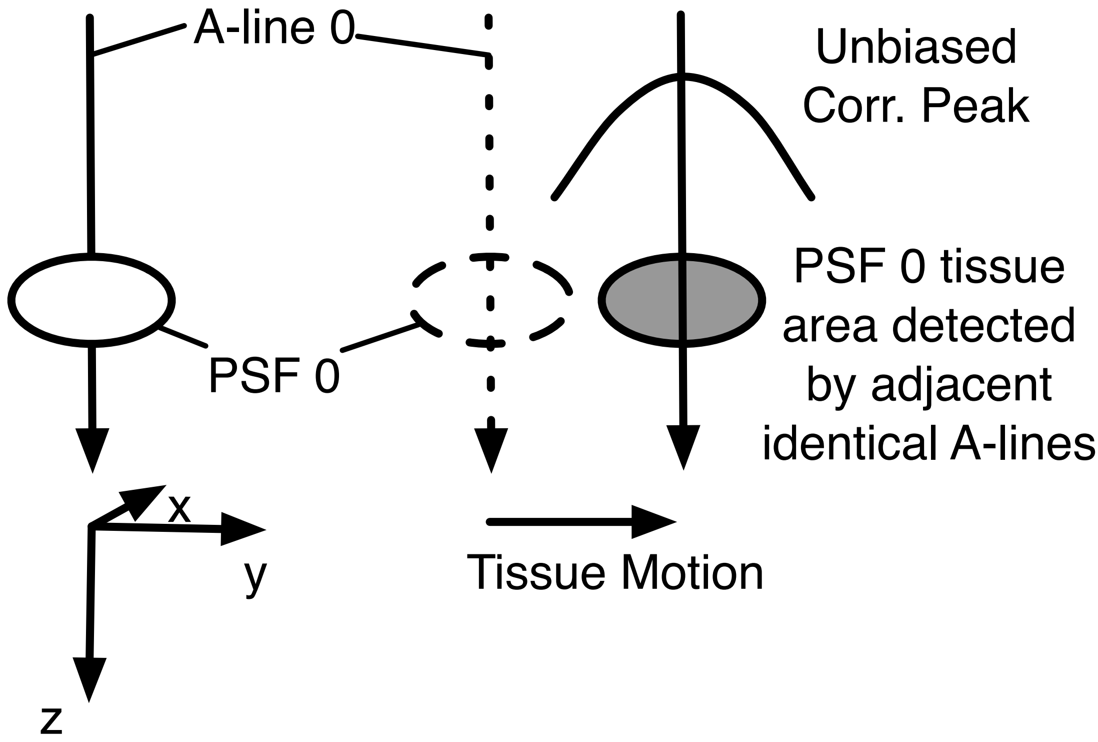

Figure 3.3: Illustration of how tissue detected by a single pixel from a vertical A-line will be detected identically after lateral translation, in the case of identical, non-angled A-lines from a linear array system. This leads to an unbiased, single-pixel correlation peak.

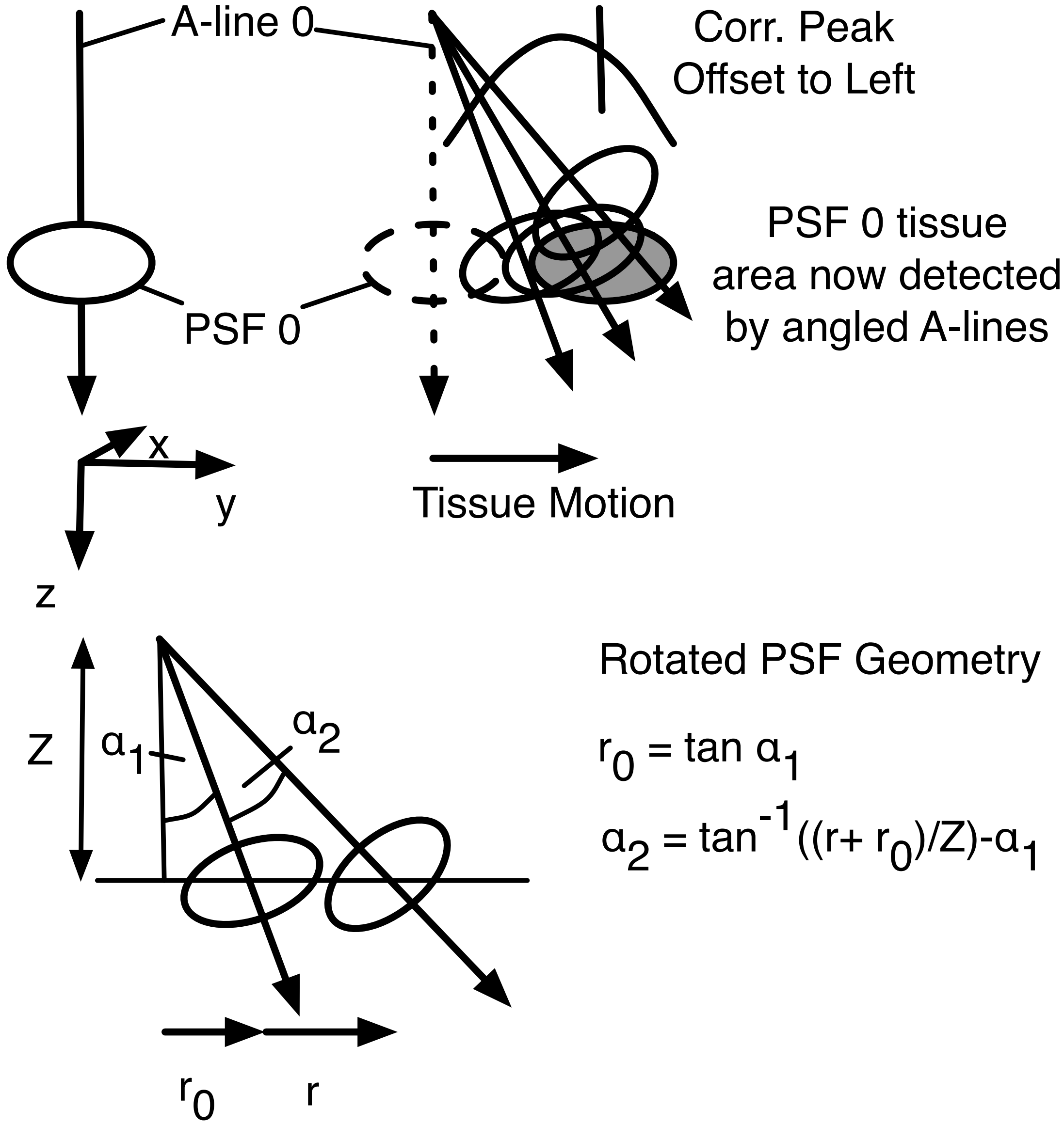

Figure 3.4: Illustration of how tissue detected by a single pixel from a vertical A-line will be detected by rotated PSFs from other, angled A-lines after lateral translation, in the case of identical, angled A-lines from a sector-scan system. This leads to an offset single-pixel correlation peak, as the PSFs from angled lines at increasing translations rotate out of the tissue area detected by the original PSF. The geometry of rotated PSFs is also shown, where tissue detected by an original A-line at angle α_1 is detected by a PSF with rotation α_2 at a lateral shift of $r_0 + r$.

3.D Materials and Methods

3.D.i Complex Correlation Statistical Model Validation

To validate the theoretical stochastic correlation model, the statistics of PSF inner products due to a shift were simulated and compared to full simulations including 3D speckle, and experimental results. Experiments used a commercial mechanical sector-scanned transducer (Interson, Pleasanton, CA), and a custom transmit/receive board we constructed, capable of real-time image display, motion tracking and data capture, with the parameters of table 3.1. The same sector-scan system was also simulated, using the settings from table 3.1 in the Field II (57) package, with an internal sampling rate of 500 MHz, down-sampled to 25 MHz for data readout. All simulations use the transmit pulse captured by a hydrophone at the transducer focus, after appropriate bandpass filtering.

Table 3.1: Transducer/Scan Parameters for Experimental and Simulated System

Property	Value	Units
Sector Angle	60	Degrees
Piston Diameter	12	mm
Focal Depth	56	mm
Number of Lines	256	N/A
Center Frequency	4	MHz
Cycles	4	N/A
Sampling Frequency	25	MHz
Correlation Window Length	200	samples

Simulations of the described system with a stationary, single A-line oriented in the positive direction of the z-axis (i.e. directly down into the target) were used to characterize the spatial sensitivity of a single complex focused

RF pixel to scatterers in 3D space, with 100 micron spacing in the x- and y-directions, and 50 microns in the z-direction. To form complex signals, a 13-tap, 4MHz center frequency, 70% bandwidth finite impulse response (FIR) bandpass filter was created using the 'fir1' command in Matlab (Mathworks, Natick, MA), then Hilbert-transformed to yield separate I and Q bandpass filters.

For each scatterer location, the receive time series was sampled at the time instant corresponding to the complex focused RF pixel using geometric round-trip propagation. The 3D complex focused RF pixel sensitivity is equivalent to the discrete-sampled functions or vector F from equations 3.1, 3.2, 3.3 and 3.6. A version of the sensitivity function shifted by the vector $\mathbf{r}$ is equivalent to $G(\mathbf{r})$, so that the inner-product of F and $G(\mathbf{r})$ can be easily calculated. Equation 3.6 can then be used to build the 2 x 2 matrix VS, which is then multiplied by a 2 x N i.i.d normal random complex matrix B, representing N similar focused RF pixels picking up independent speckle realizations. This corresponds to f and g (equations 3.4, 3.7, 3.8) for N pixels in a single A-line according to the theoretical model, and is used to produce representative statistics for normalized complex correlation as per equation 3.9 for a variety of PSF shifts over 10,000 realizations. To account for the axial correlation between RF samples, an axial covariance matrix based on the pulse parameters is used.

To verify the theoretical model for PSF shifts, correlation statistical properties are compared to normalized correlation for a simulated line of complex focused RF pixels, over many realizations of fully-developed 3D speckle while undergoing the same shifts. In addition, analogous experimental data were captured using the system of table 3.1, while undergoing translations in the azimuthal plane over a tissue-mimicking gelatin phantom (72). For comparison with the simulations, particular A-lines from the experimental data were chosen with angles corresponding to those the simulations as closely as physically possible. To assess statistics from the PSF inner product model, 3D speckle simulations and experimental data, the standard deviation of normalized complex correlation magnitude, and angular deviation calculated from normalized complex correlation angle were used. Angular deviation is a phase-wrap tolerant measure of angular spread (79), (80), calculated using equation 3.24, where r_i is a unit vector in the direction of the i_{th} complex sample, $\theta_i = arg(r_i)$, $\bar{r}$ is the mean resultant vector, R is the l^2-norm of $\bar{r}$, and σ_{ANGLE} is the angular deviation.

$$r_i = \begin{pmatrix} cos\theta_i \\ sin\theta_i \end{pmatrix}, \ \bar{r} = \Sigma_i r_i, \\ R = ||\bar{r}||_2, \ \sigma_{ANGLE} = \sqrt{2(1-R)} \tag{3.24}$$

3.D.ii Joint Azimuthal-Elevational Motion Estimation Evaluation

Two kinds of simulated data were used, in addition to experimental data to assess the performance of the described joint azimuthal-elevational motion estimation algorithm. Simulations were used to calculate the complex PSF inner product over many possible tissue translation vectors, along with the described factorization methods (equations 3.5 and 3.6) to acquire many independent realizations to assess motion estimation statistics. The first simulation method uses the separable PSF model from equation 3.10, and the resulting inner-product from equation 3.11 to assess algorithm performance for an ideal PSF. Parameters for the separable PSF were calculated from experimental data, as detailed in the parameterization steps below.

The second simulation method uses the complex focused RF pixel 3D scatterer sensitivity function calculated using Field II (57), as described in the previous section to calculate the complex inner product for many combinations of A-line angle and azimuthal translation on a regular grid (0 to 30 degrees in 5 degree steps, and 0 to 2.9 mm shifts in 0.1 mm steps). Next, two-dimensional cubic-spline based interpolation is used to find the corresponding normalized complex correlation for a given A-line orientation during an azimuthal/elevational shift. To account for the axial and lateral correlation between adjacent RF pixels, a spatially-varying 2D-separable covariance function based on 3.11 was used to introduce spatial correlation into the random complex B variables from equation 3.5. For both methods, for every realization of 256 sampled A-lines, 256 normalized complex correlations were calculated and used as input to motion estimation algorithms.

For azimuthal motion, the algorithm described by equations 3.12-3.23 was used to produce a maximum likelihood motion estimate using complex correlation phase alone. In this calculation, a diagonal phase covariance matrix was used, with each correlation's variance derived using a deterministic (degree-5 polynomial) relationship to correlation magnitude characterized from simulations. To avoid noise-related phase unwrapping errors, the raw complex correlation data was filtered using a 9-tap boxcar filter across all the A-lines, before determining phase from the complex values. To estimate elevational motion given the azimuthal estimate, the method of maximizing r_i using the 2D integration from equation 3.19 proved computationally intractable, if sampled sufficiently finely in $|c_0|$, r_i and r_j. Instead, the methods of equations 3.22 and 3.23 were both used to form elevational estimates. For the method of equation 3.22, the mean of the most-likely elevational motion for every 8th A-line was used as a global estimate. Due to the form of equation 3.23, the (RMS) of the most-likely r_i value for each A-line was used as the overall elevational motion estimate. For both elevation estimation algorithms, the variable r_i was permitted to be negative, so that a positive elevational motion bias could be avoided.

Simulations were repeated 50 times for each shift, using a window of 200 axial pixels, corresponding to approximately 6.2 mm of depth range. Experimental data was collected using the parameters of table 3.1, while the system imaged a tissue-mimicking gelatin phantom (72) and was translated repeatedly through a range of azimuthal/elevational combinations across different regions of the phantom. All experimental data was partitioned into training and testing segments, acquired on different physical regions of the phantom in order to use one segment for a least-squares fit to the parameters of equation 3.10, the second segment for performance evaluation. Experimental data was evaluated for motion estimation using RF data from depths corresponding to those from simulations, using identical parameters. To find the parameters of equation 3.10 for experimental data at a particular depth, the following steps were followed :

1. For 50 realizations of a range of lateral shifts from 0-1.0 mm, find the parameter $\frac{4\pi}{\lambda}$ from equation 3.11 that is the least-squares fit to the correlation phase from each A-line in the ensemble.

2. For 50 realizations of a range of elevational shifts from 0.1-1.0 mm, find the parameter α from equation 3.11 that is the least-squares fit to the correlation magnitude from each A-line in the ensemble.

3. With fixed α, and $\beta = \alpha$, for 50 realizations of a range of lateral shifts from 0-1.0 mm, find the parameter γ from equation 3.11 that is the least-squares fit to the correlation magnitude from each A-line in the ensemble.

3.D.iii Conventional Azimuth and Elevational Motion Estimation

For comparison, conventional azimuthal-plane and elevation motion estimation was performed on experimental data, using normalized 2D correlation (60), (66) with cosine-fitting and fitted-Gaussian decorrelation-based estimation respectively (61). Normalized correlation has been shown to outperform most other time-delay estimators (67), alongside the sum-squared-differences (SSD) method, and is a good choice for comparison as it forms the basis of the azimuthal sector scan estimator. For subsample estimation, raised-cosine fitting was used as it is amongst the better computationally-efficient interpolation schemes (78). For both methods, RF data was envelope detected and scan converted on a 75 micron (lateral) x 30.8 micron (axial, corresponds to 25 MHz sampling) grid using successive linear interpolation in each rectilinear axis before estimation.

For azimuthal-plane motion estimation, a 2D kernel 200 pixels deep, and as wide as the sector scan, less a margin of 30 pixels on each side was used, with a window region 30 pixels larger in each direction. For elevation, decorrelation-based motion detection, a training set of experimental data from a tissue-mimicking phantom (72) was used with 50 realizations of lateral shifts from 0.1 mm - 1.0 mm to perform least-squares fitting to a parametric Gaussian decorrelation model (61). Conventional azimuthal-plane and elevation motion estimation was not performed on simulated data due to the impractical computational load involved in simulating multiple realizations of 3D speckle over many different shifts.

However, to further test the stochastic complex correlation model, the rotation of PSFs with lateral translation was simulated using the previously described geometric model and separable PSF model of equation 3.10. Stochastic simulations using equation 3.6 were used to form a degree-5 polynomial characterizing the deterministic relationship between the magnitude of the expected value of complex correlation, and the expected value of detected correlations for any kind of PSF shift. Using this model, the bias in normalized detected A-line correlation was simulated for 200-pixel windows and compared to experimental data for A-line correlation and scan-converted speckle-tracking results for lateral shifts.

3.D.iv Motion Tracking Validation Using 3D Spine Model

To evaluate the performance of the sector-scan motion tracking algorithm in 3D volume imaging, a tissue-mimicking gelatin phantom (72) was constructed with an embedded human spine analog, formed by 3D-printing a commercial CAD model of the L3-L5 lumbar spine region (3D Systems, Rock Hill, SC). The spine model has a known shape and dimensions, and can be used as a gold standard for comparison to experimental 3D ultrasound data using 3D registration methods. Ultrasound data were captured during a semi-random translation sequence so that the effect of the motion estimation method can be quantified. Sector-scan azimuthal image planes were captured orthogonal to the spine axis, with translation using a MM3000 motion stage (Newport, Irvine, CA). The semi-random translations consisted of uniformly randomly distributed elevational steps, with a mean of 0.35 mm, and extent of $\pm$ 0.25 mm, along with normally-distributed azimuthal steps ($\mu = 0$, $\sigma = 0.25$ mm). The azimuthal motion steps were low-pass-filtered using a 9-tap boxcar filter to prevent unrealistic rapid motion changes, and to ensure that the spine remained in the field of view, the sequence was further constrained by discarding any sequence that had a maximum absolute azimuthal displacement of > 5 mm.

To convert ultrasound data to surfaces in 3D space, every A-line from each azimuthal image was envelope-detected before edge detection using a 21-tap filter designed to be matched to the bright bone reflection followed by below-average intensity dark shadow. To achieve this, observing that most bone surfaces have a depth of about 7 samples, the filter has 7 unity-value taps followed by 14 taps with the value $-\frac{1}{2}$, with a simple threshold giving a binary edge determination. Following the edge detection filter, scan conversion produces a series of edge points within a rectilinear plane. Each plane of edge points was then translated in 3D space using either known or estimated position, before conversion to polygonal surfaces for comparison to the polygonal 3D spine model. Position estimates were calculated using the joint azimuthal-elevational sector-scan method, with the ‘fast’ elevational estimator variant (equation 3.23), and also using the conventional, speckle-tracking plus elevation decorrelation-based estimator.

The VTK (Visualization Toolkit, Kitware, Clifton Park, NY) software package was used to import and represent the 3D ultrasound image data and 3D spine model as a set of points associated with polygonal (triangular) 3D surfaces. These polygonal surfaces were created from the 'vtkContourFilter' function, which generated an isocontour from the volume of ultrasound image edge data. The isocontour with the smallest value that yielded essentially no speckle signal i.e. only isocontours within the locality of the bone surfaces - was used to produce a root-mean square (RMS) distance between the ultrasound surface points and 3D spine model surface points. Surfaces from both 3D spine model and ultrasound data that cannot be physically interrogated with ultrasound (due to hidden surfaces, limited data extent, etc) were excluded by retaining only the polygonal surfaces and associated points possessing normal vectors directed toward the transducer face within a 180 ° span - i.e. only the top surfaces were retained while the undersides of the VTK spine model or 3D ultrasound representations were excluded. Using the ITK (Insight Toolkit, Kitware, Clifton Park, NY) software package, points associated with the final set of spine model or ultrasound surfaces were iteratively affine-transformed until a similarity metric, the Euclidian distance between points on the ultrasound surface and points on the 3D surface model, was minimized. The transformation was applied using the VTK functions 'RotateX', 'RotateY', 'RotateZ', and 'Translate', which rotate the ultrasound surface about the X, Y, and Z axes and translate the ultrasound surface, respectively. The resulting aligned ultrasound surface and 3D spine model were then superposed using VTK surface rendering tools for visualization. After alignment, the RMS surface distance error was computed, with each distance measurement consisting of the distance between a point in the 3D ultrasound image data and the closest point on the 3D spine model.

3.E Results

3.E.i Single A-line Complex Correlation Statistics

A comparison of single A-line complex correlation magnitude and phase is presented in Figure 3.5, for the deterministic 3D PSF model, simulated PSF shifts in multiple realizations of 3D speckle, and experimental shifts. Normalized correlation phase and magnitude are shown for a range of shift magnitudes (0.1 mm - 2.5 mm , and angles relative to pure lateral motion (0°), using a 200-sample window about the transducer focal depth. For each magnitude curve, a least-squares fit was made to find the parameter α from the Gaussian form $M = e^{-\frac{1}{2}\alpha x^2}$, with M the correlation magnitude and x the shift magnitude. Similarly, a least-squares fit was made to find the phase slope C from $P = Cx$ where P is phase in radians. Results for all cases are presented in table 3.2. The deterministic PSF model α and phase slope (C) parameters exhibit an RMS error relative to the speckle-simulation values of only 9.98 % (α) and 2.46 % (C), and relative to the experimental data, 5.84 % (α) and 10.38 % (C).

Second-order statistics produced by the deterministic 3D PSF complex correlation shift model are show in Figs. 3.6 and 3.7. Standard deviation of the magnitude of complex correlation (Fig. 3.6) and angular standard deviation of the complex correlation angle (Fig. 3.7) from PSF model data are compared to the same parameters from 50 realizations of simulated 3D speckle data, and 50 sets of experimental data from different locations over a tissue-mimicking phantom. To illustrate the distribution of the normalized complex correlation value in the complex plane, Fig. 3.8 shows 50 normalized complex correlation values from each data set, at a range of shift magnitudes, and for two shift angles (10° and 20°).

Table 3.2: Complex Correlation Gaussian-Magnitude and Linear Phase Parameter Fits for Model, Simulated and Experimental Data

Parameter/Angle	0°	10°	20°	30°
Model α ($\times 10^6$)	0.8866	1.2333	1.7152	2.2291
Simulated α ($\times 10^6$)	0.8957	1.3566	1.6427	2.5403
Experimental α ($\times 10^6$)	0.9858	1.2613	1.6396	2.3735
Model C ($\times 10^4$ radians/m)	0.0000	0.5281	1.0390	1.4822
Simulated C ($\times 10^4$ radians/m)	-0.0057	0.5302	1.0749	1.4536
Experimental C ($\times 10^4$ radians/m)	0.0130	0.5675	1.1231	1.6787

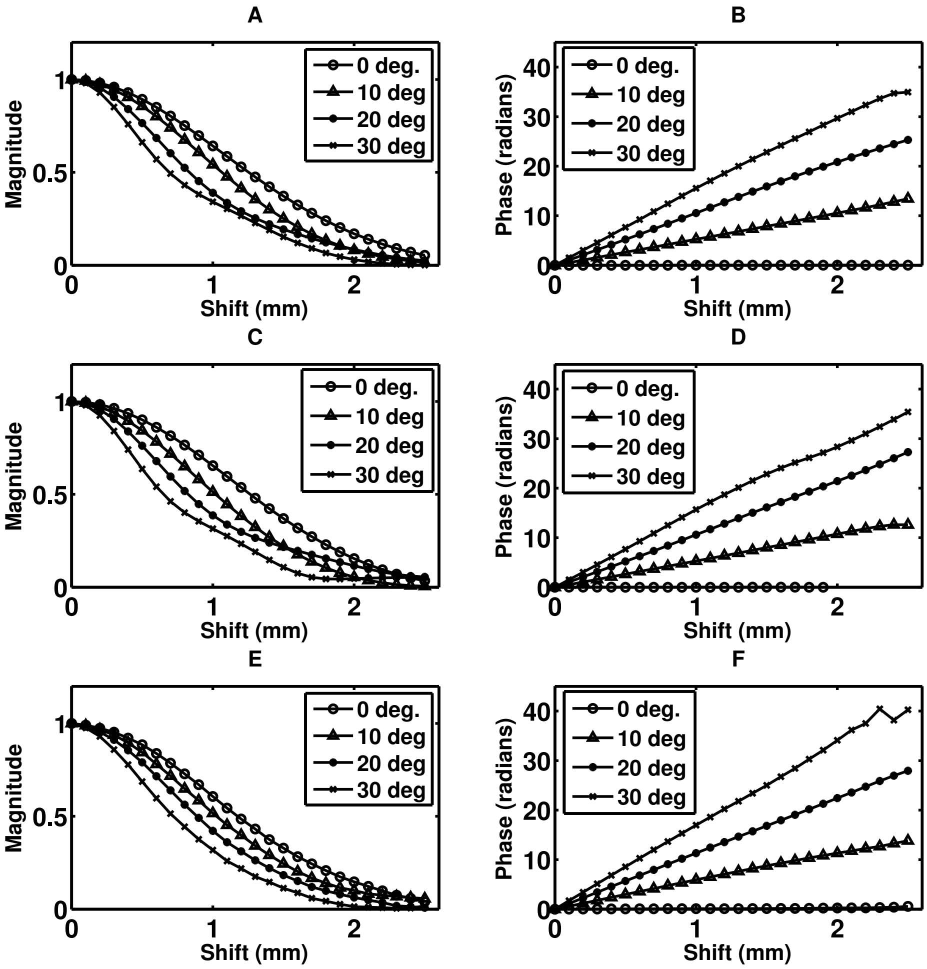

Figure 3.5: Magnitude and phase of single A-line normalized complex correlation for 200-pixel window with shifts of varying magnitudes at different angles relative to pure lateral motion (0°). For all results, RF pixel is at focus of piston transducer, at 56 mm depth. Correlation values calculated using a deterministic simulated 3D PSF inner product are shown in A and B. The corresponding measurements for PSF-shift simulations in fully-developed 3D speckle are shown in C and D. Correlation magnitude and phase values from experiments using a tissue-mimicking phantom are shown in C and F. Results from C and D were averaged over 50 independent 3D speckle realizations, and those from E and F were averaged over 50 independent phantom locations.

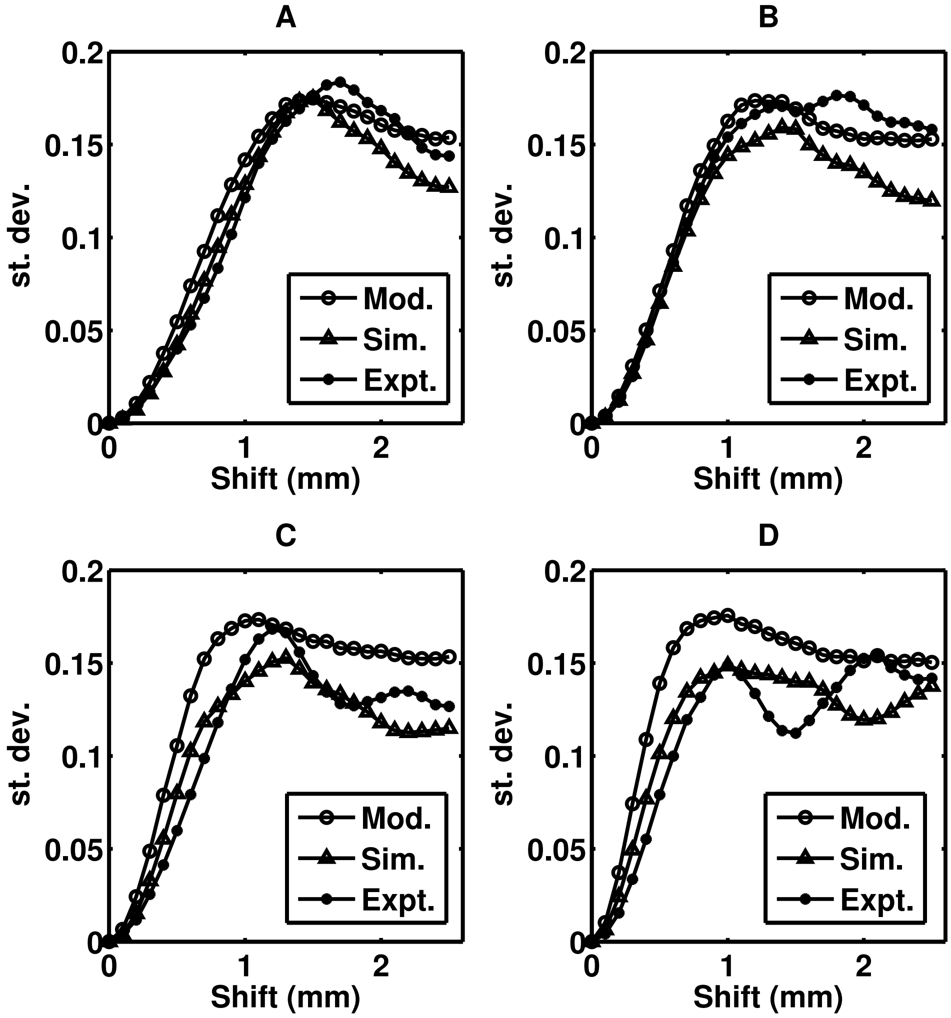

Figure 3.6: Standard deviation of single A-line normalized complex correlation magnitude for 200-pixel window with shifts of varying magnitudes at different angles relative to pure lateral motion including 0° (A), 10° (B), 20° (C) and 30° (D). For all results, RF pixel is at focus of piston transducer, at 56 mm depth. Standard deviation values are shown for a simulated 3D PSF inner product model (Mod.), for 50 realizations of a PSF-shift simulation in fully-developed 3D speckle (Sim.), and for experimental data at 50 different phantom locations (Expt.).

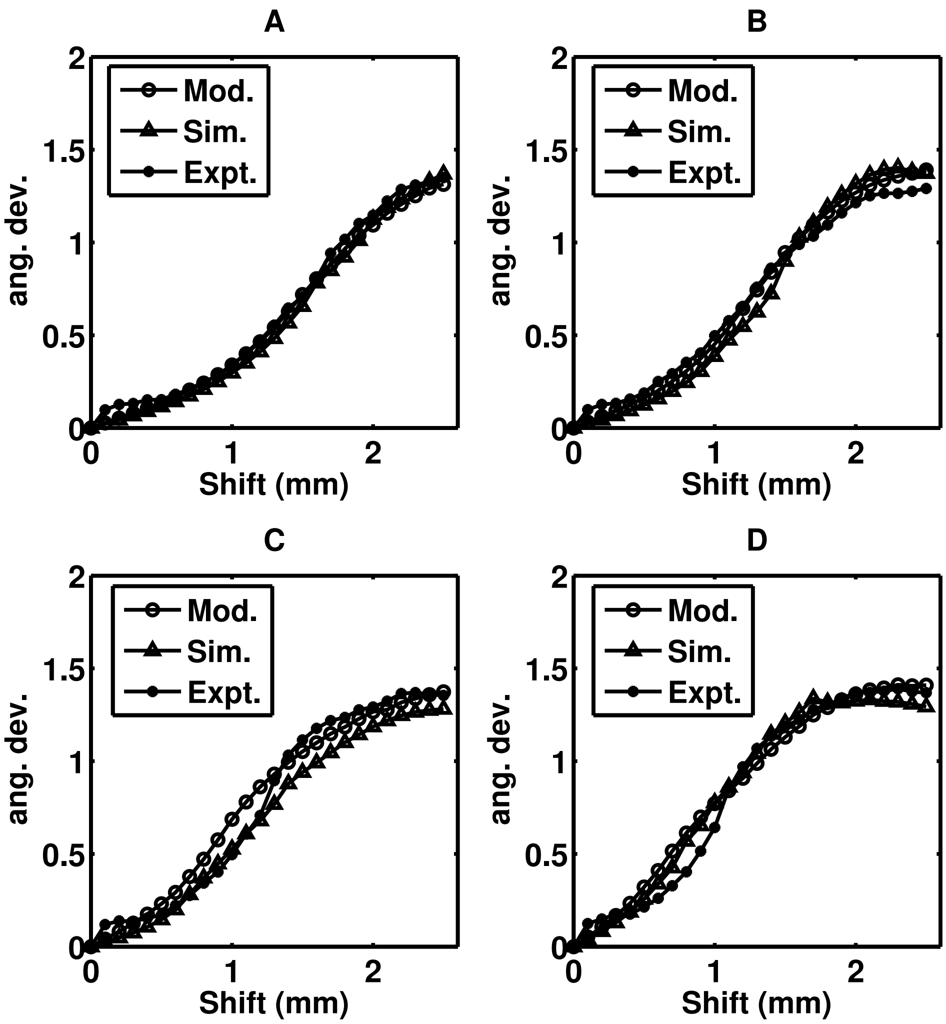

Figure 3.7: Angular standard deviation of single A-line normalized complex correlation angle for 200-pixel window with shifts of varying magnitudes at different angles relative to pure lateral motion including 0° (A), 10° (B), 20° (C) and 30° (D). For all results, RF pixel is at focus of piston transducer, at 56 mm depth. Circular standard deviation values are shown for a simulated 3D PSF inner product model (Mod.), for 50 realizations of a PSF-shift simulation in fully-developed 3D speckle (Sim.), and for experimental data at 50 different phantom locations (Expt.).

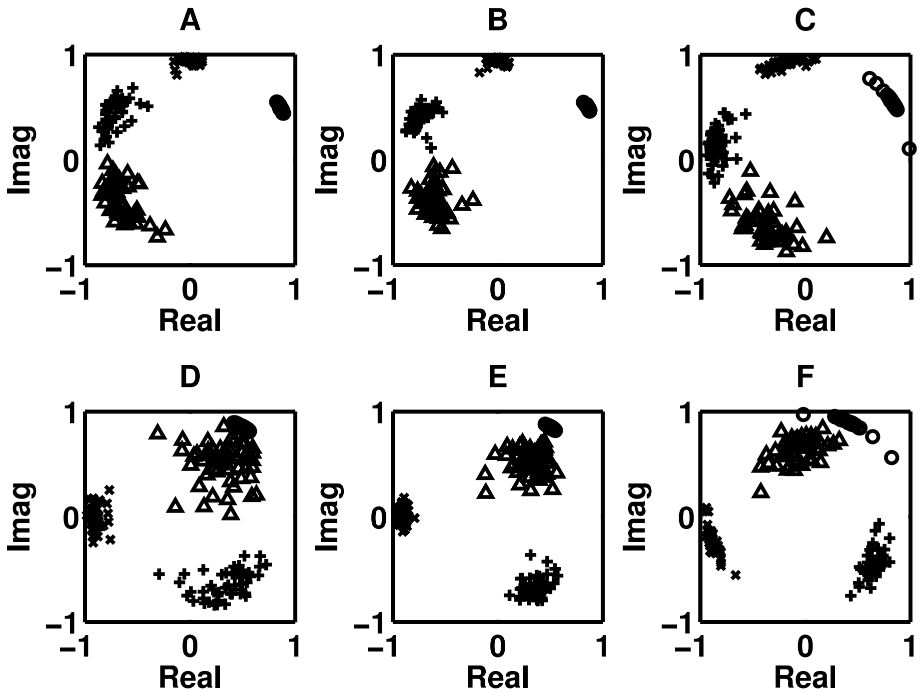

Figure 3.8: Normalized complex correlation values for a 200-pixel window from a single A-line, while undergoing shifts with magnitude 0.1 mm (○), 0.3 mm (×), 0.5 mm (+) and 0.7 mm (△). Panels A, B and C show 50 realizations each for the PSF model, 3D speckle simulation and experimental data with a shift angle of 10°, while D, E and F show the same data for a shift angle of 20°.

3.E.ii Joint Azimuthal-Elevational Motion Estimation Evaluation

The performance of the joint azimuthal-elevational motion estimation algorithm is shown for azimuthal motion estimation for the cases of : azimuthal-only motion, elevation-only motion, and both azimuthal and elevational motion, with two kinds of simulated data (gaussian 3D PSF and Field-II based PSF) in Fig. 3.9, and experimental data in Fig. 3.10. In Figs. 3.11 and 3.12, elevation estimation performance is shown for simulations with the 3 motion cases, using the gaussian 3D PSF model and Field-II based PSF models respectively. Experimental elevation estimation performance is shown in Fig. 3.13 for the same motion cases.

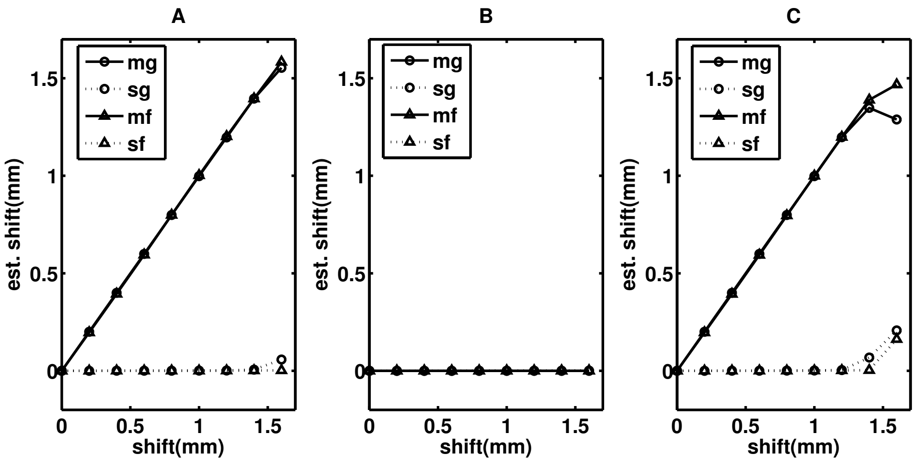

Figure 3.9: Azimuthal motion estimation performance using simulated data. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for gaussian PSF model (mg, sg) and Field II PSF model (mf, sf).

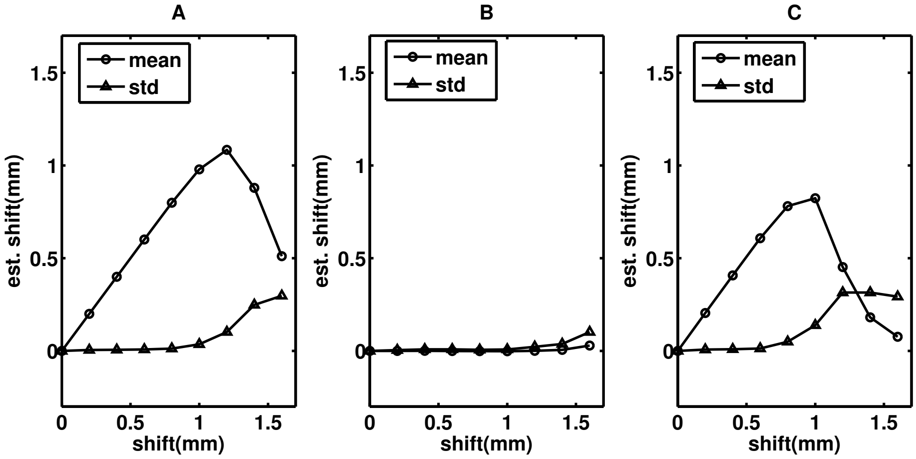

Figure 3.10: Azimuthal motion estimation performance using experimental data from a tissue mimicking phantom. Panel (A) shows azimuthal estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown.

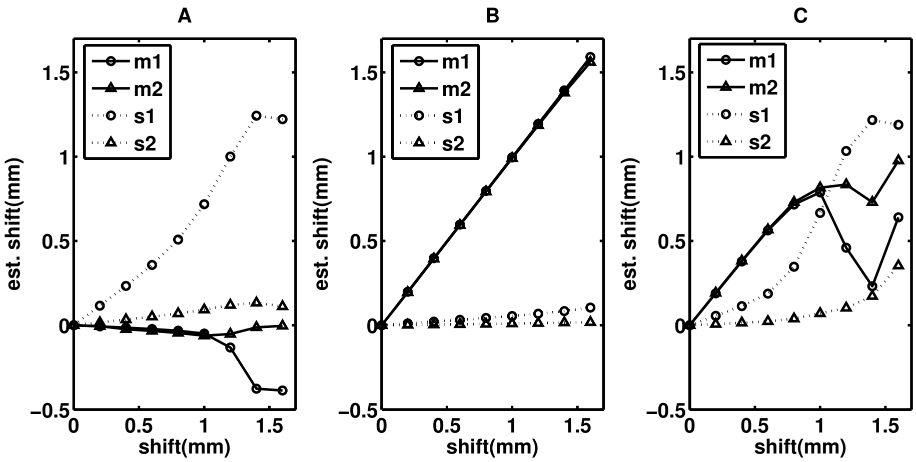

Figure 3.11: Elevation motion estimation performance using simulated data from separable gaussian PSF model. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for ‘fast’ (equation 3.23, m1, s1) and ‘accurate’ (equation 3.22, m2, s2) estimators.

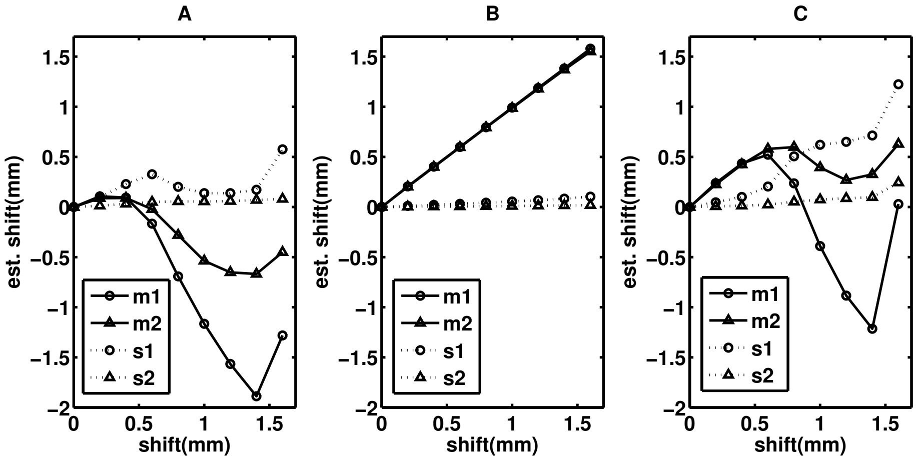

Figure 3.12: Elevation motion estimation performance using simulated data from Field II PSF model. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for 'fast' (equation 3.23, m1, s1) and 'accurate' (equation 3.22, m2, s2) estimators.

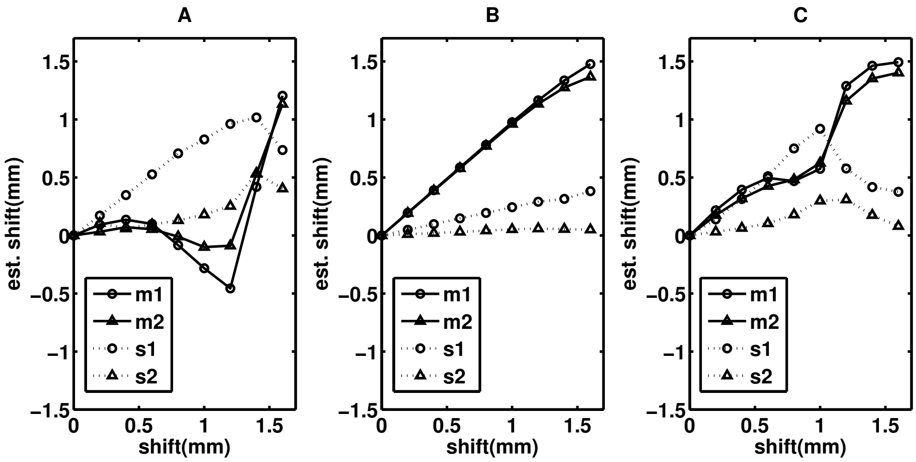

Figure 3.13: Elevation motion estimation performance using experimental data from a tissue-mimicking phantom. Panel (A) shows elevation estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for 'fast' (equation 3.23, m1, s1) and 'accurate' (equation 3.22, m2, s2) estimators.

3.E.iii Conventional Azimuth and Elevational Motion Estimation

A comparison of azimuthal motion estimation performance using the joint azimuthal-elevation sector-scan estimator and conventional azimuthal (speckle-tracking) estimation is shown in Fig. 3.14, for a variety of azimuthal/elevational motion combinations. A similar comparison of elevational motion estimation, using the joint azimuthal-elevation sector-scan method, and conventional elevation (decorrelation-based) estimation is shown in Fig. 3.15, for the same motion combinations.

Figure 3.16 illustrates the shift in the location of the correlation peak for a vertical sector-scan A-line with lateral shifts of 0.5 mm, 1.0 mm and 1.5 mm, both predicted by stochastic simulations using the 3D PSF correlation model, and from experimental data over many realizations. The correlation peak shift from experimental data for the 0.5 mm, 1.0 mm and 1.5 mm shifts, over a range of A-line angles is presented in Fig. 3.17. The mean peak shift, or bias, for the 0.5 mm, 1.0 mm and 1.5 mm shifts, over all sector scan angles is shown in table 3.3, along with the corresponding speckle-tracking lateral shift estimation bias. These values match within $\sim 15\%$ for all three shifts.

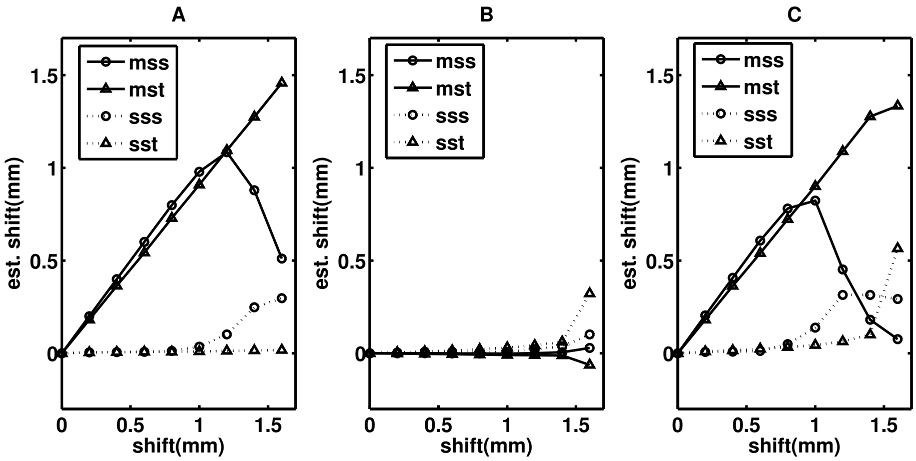

Figure 3.14: Azimuthal motion estimation performance for sector-scan and speckle-tracking methods using experimental data from a tissue mimicking phantom. Panel (A) shows azimuthal estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for the sector scan method (mss, sss) and speckle-tracking method (mst, sst).

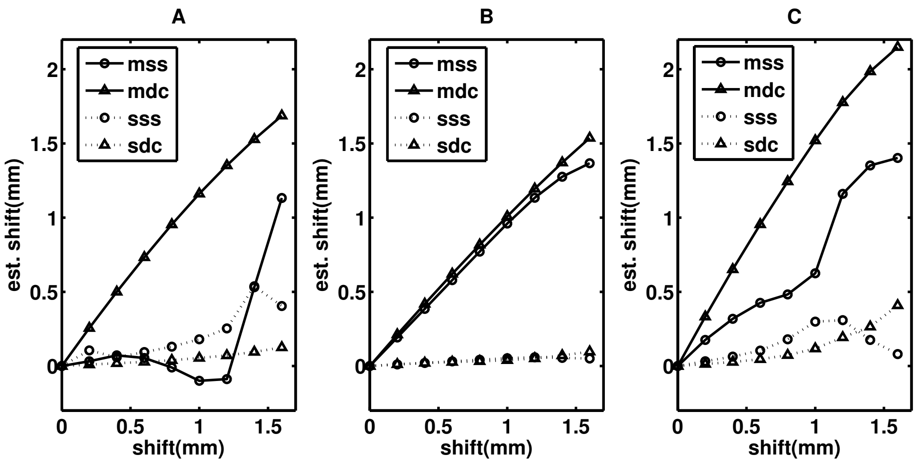

Figure 3.15: Elevational motion estimation performance for sector-scan and speckle-tracking methods using experimental data from a tissue mimicking phantom. Panel (A) shows elevational estimator performance for azimuthal-only shifts, (B) shows elevation-only shifts, and (C) shows diagonal (elevation=azimuthal) shifts. In each panel, estimator mean and standard deviation are shown for the sector scan method (mss, sss) and the decorrelation-based method (mdc, sdc).

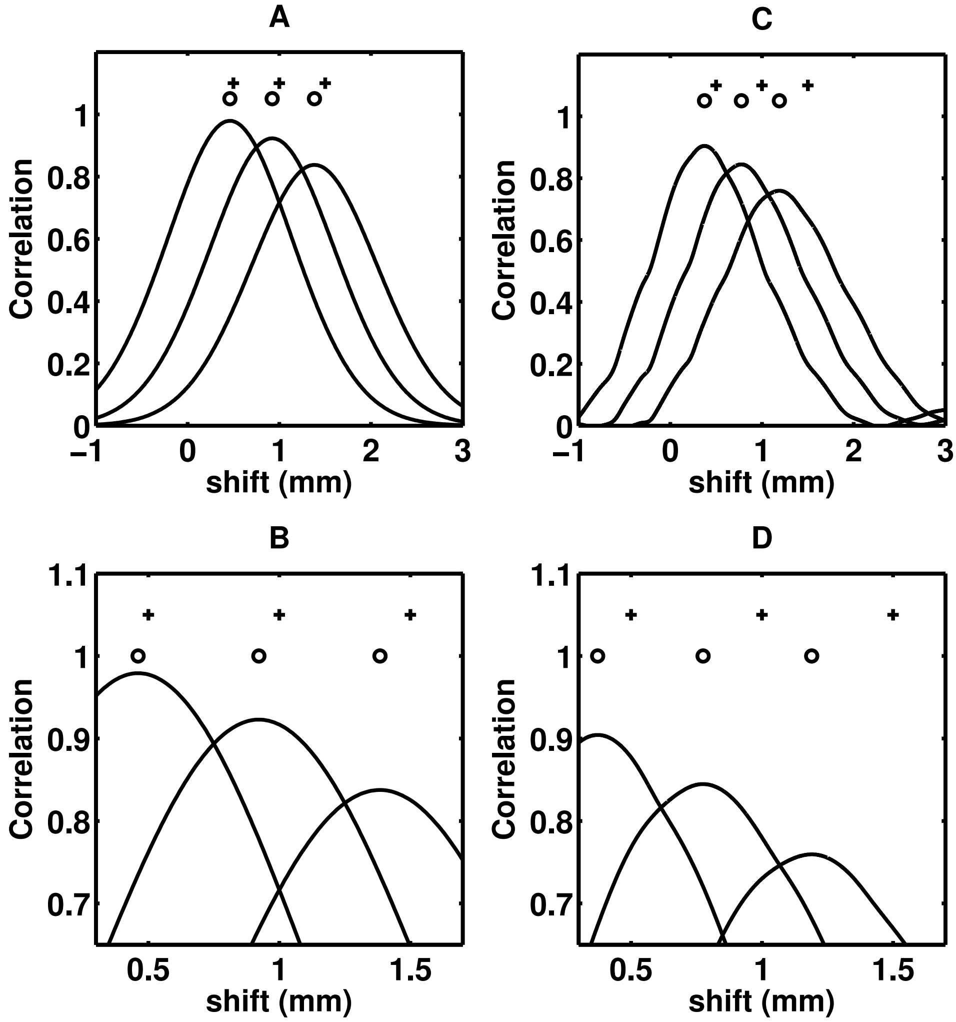

Figure 3.16: Simulated and experimental normalized correlation curves for single, vertically-oriented detected A-line, using 200-pixel window centered at depth of 56 mm. Underlying shifts of 0.5 mm, 1.0 mm and 1.5mm are shown. Top panel (A) shows simulated correlation curves with peaks (○), and actual shifts (+). Experimental results averaged over 50 realizations are shown in panel (C). Bottom panels (B) and (D) are identical to (A) and (C), zoomed to area of interest.

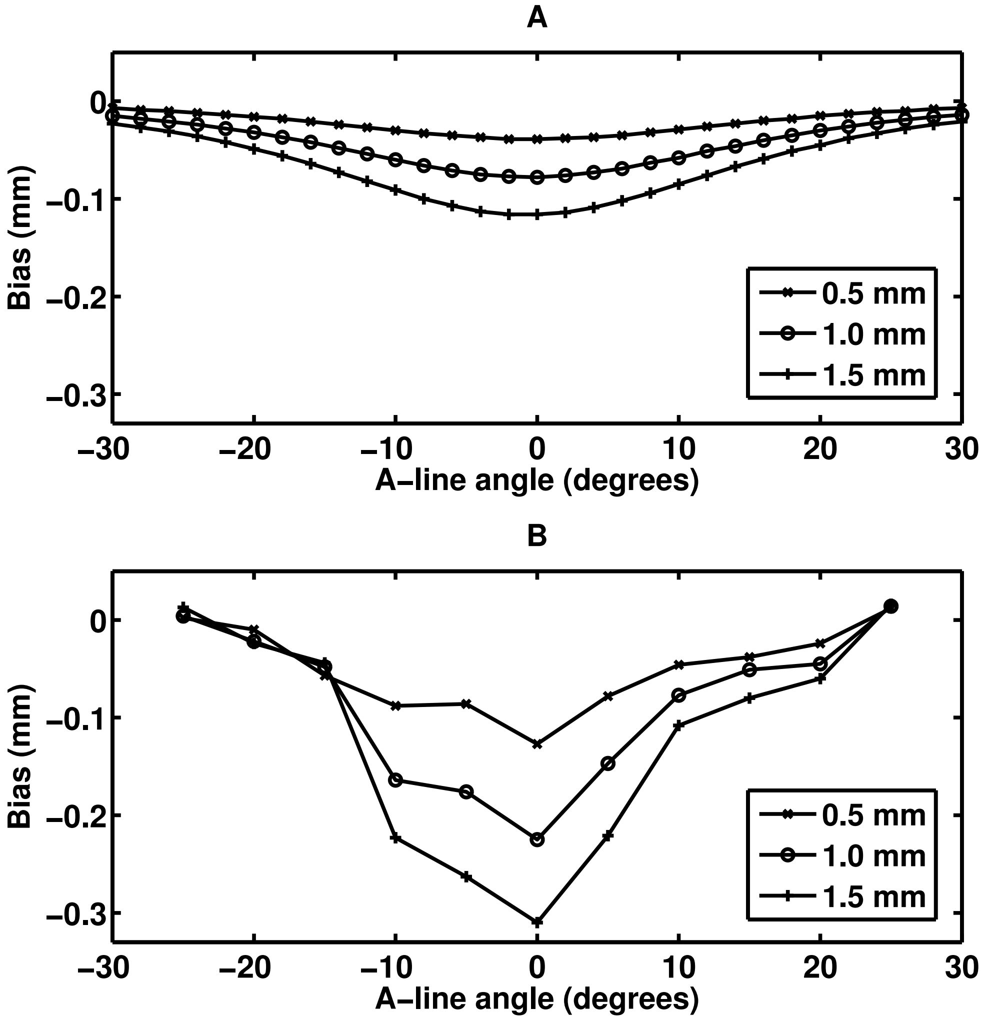

Figure 3.17: Simulated and experimental bias measurements for lateral motion estimation using samples at various initial A-line sector angles with lateral shifts of 0.5 mm, 1.0 mm and 1.5 mm. Simulated bias (A) is found using a deterministic PSF model to find the peak of the correlation curve with lateral shifts. Experimental bias (B) is calculated for a given initial A-line angle by finding the 200-sample long vertical strip of scan-converted, detected pixels corresponding to that angle, then shifting laterally to find the correlation peak.

Table 3.3: Experimental Lateral Motion Estimation Bias

Lateral Shift (mm)	Mean A-line peak Bias (mm)	Mean Speckle-Tracking Bias
5	-0.0415	-0.0491
10	-0.0915	-0.0913
15	-0.1381	0.1351

3.E.iv Motion Tracking Validation Using 3D Spine Model

Fig. 3.18 shows the true path and two estimated paths for a semi-random walk (performed using a motion stage) over the surface of a tissue-mimicking phantom containing a 3D human spine model (L3-L5 vertebrae). Table 3.4 shows the root-mean-square (RMS) error between the 3D spine model surfaces and bone surfaces detected by a sector-scan transducer during the semi-random walk. Figs. 3.19, 3.20 and 3.21 show 3D surface renderings of the 3D spine model with the ultrasound-detected bone surfaces superimposed, for the 'true' motion, motion estimation using the joint azimuthal-elevational method, and conventional motion tracking respectively. For illustrative purposes, to more easily visualize the surface alignment, in figs. 3.19, 3.20 and 3.21, ultrasound data was raised 7mm and the isocontour threshold was set to a value smaller than the value used to calculate the RMS surface error.

Table 3.4: Root-Mean-Square Bone Surface Error for Ground-Truth, the Proposed Sector-Scan Tracking Algorithm and for Conventional Speckle-Tracking/Decorrelation-based Motion Estimation

Estimation Method	RMS Bone Surface Error (mm)
Ground-Truth (motion stage displacement)	1.082
Sector-Scan Joint Azimuthal/Elevation Estimation	1.204
Speckle-Tracking/Decorrelation-based Estimation	2.425

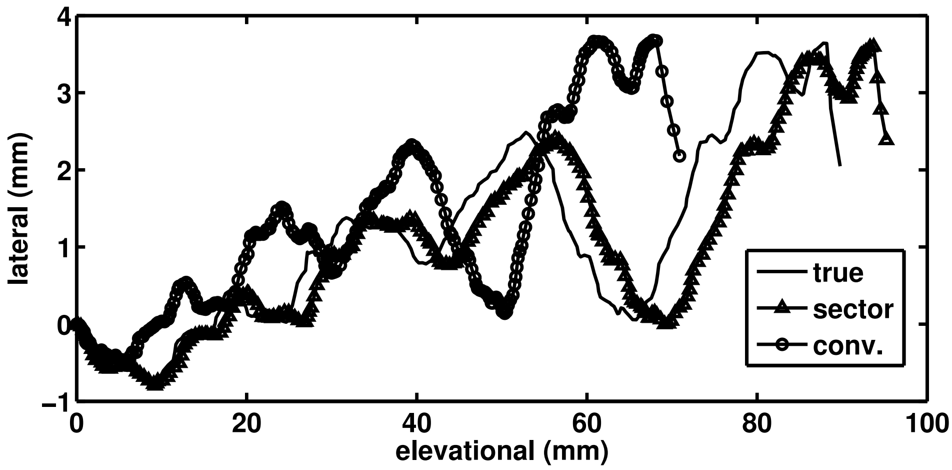

Figure 3.18: True and estimated paths taken during a semi-random walk, with two kinds of motion estimation, sector-scan based joint azimuthal-elevation estimation, and conventional, speckle-tracking azimuthal motion estimation coupled with decorrelation-based elevation motion estimation.

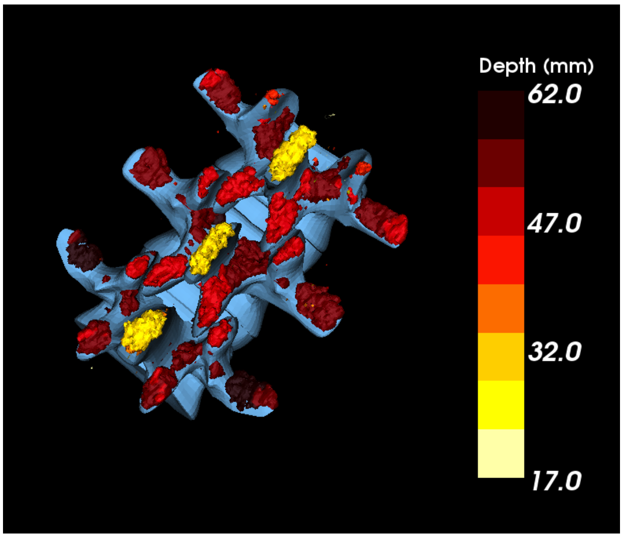

Figure 3.19: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from 258 sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using ground-truth location for each sector-scan image.

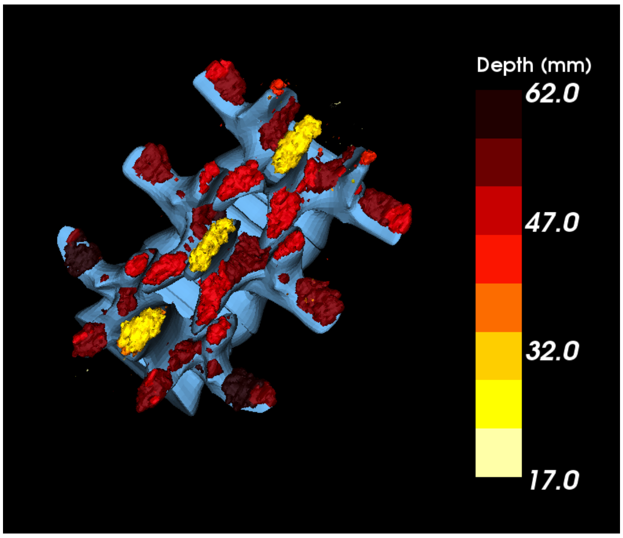

Figure 3.20: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from 258 sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using joint azimuthal-elevational sector-scan motion estimation to place each sector-scan image.

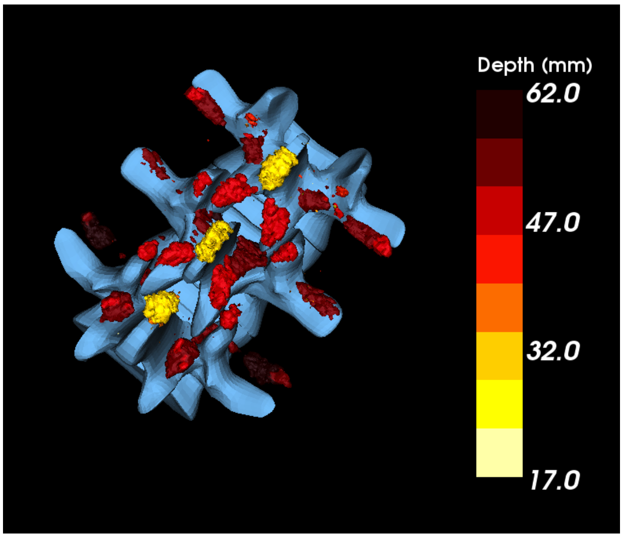

Figure 3.21: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from 258 sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using speckle-tracking and decorrelation-based motion estimation to place each sector-scan image in the lateral and elevational axes respectively.

3.F Discussion

3.F.i Statistics of Normalized Complex Correlation with PSF Transforms

The theoretical background of second-order speckle statistics is well-known (81), (82), (83) and is successfully used as the basis of methods to detect out-of-plane motion magnitude using a parametric fit to the speckle decorrelation curve (61). Motion detection in the azimuthal plane using speckle tracking (60), (66) also effectively uses the speckle correlation curve, in finding the peak correlation corresponding to a subsample shift. However, when there is simultaneous motion in the azimuthal (scan-plane) and elevational (out-of-plane) directions, the decorrelation due to out-of-plane motion appears as a noise source to azimuthal motion detection using speckle-tracking (68). Similarly, in-plane lateral motion also causes decorrelation, which will be spuriously interpreted as out-of-plane motion by the decorrelation-based estimator.

Speckle-tracking motion estimation has significantly higher resolution in the axial dimension, due to higher spatial frequency, smaller speckle size for RF data in this dimension (81), (82), and the ability to use estimators that approach the Cramer-Rao lower bound (CRLB) (84). However, as in-plane motion is predominantly in the lateral dimension, normalized correlation with detected data is typically used for speckle-tracking. No known analytical model exists for the probability distribution of normalized speckle correlation with translations, although some efforts have been made to fit the probability distribution of normalized correlation from experimental data to a general statistical model (71), (70) for improved out-of-plane motion estimation.

Sector-scan motion azimuthal motion estimation is degraded by rapid decorrelation due to PSF rotations with lateral shifts (62). However, as sector-scan A-lines span a range of angles, lateral azimuthal-plane motion will have an axial component relative to some of the A-lines, which can be estimated with higher accuracy using RF data. Complex correlation data (from IQ demodulation) can be used to estimate the axial motion for each A-line (and by geometry, the azimuthal-plane motion vector), independently of any elevation-plane motion. As the azimuthal motion estimation is highly accurate, it can be used to account for the azimuthal-motion component of normalized correlation magnitude for each A-line, so that elevational motion can be estimated using the remaining decorrelation. The separation of motion estimation into two parts - azimuthally using highly-accurate axial complex correlation phase, and elevationally using using the remaining correlation magnitude after accounting for azimuthal motion, should reduce motion estimation errors when both motion components are present.

I have presented a statistical model for complex PSF correlation for a single RF pixel detecting speckle (equation 3.1), and by discretizing the problem, formed a linear algebraic equivalent (equation 3.2). The linear algebra formulation can be used to factorize the single-pixel complex PSF correlation expression into a much smaller and simpler form (equations 3.3, 3.4). This is used to replace the very large scalar, normally distributed scatterer vector with a 2 x 1 complex normally distributed vector (equation 3.5), and fully characterize the PSF correlation statistics using a simple 2 x 2 matrix derived from complex PSF correlation magnitude and phase (equation 3.6). The form of equation 3.6 indicates that the single RF pixel complex correlation is the sum of two parts, a chi-squared distribution in the direction of the underlying PSF complex correlation, and a Bessel-like distribution in the orthogonal direction in the complex plane (equations 3.7, 3.8). If many identically distributed RF-pixels from an A-line are used in normalized correlation (equation 3.9), the central limit theorem (73) and small-angle approximation can be used to show that the phase of the normalized correlation is approximately normally distributed. Using the commonly-used separable gaussian 3D PSF model ((81), equation 3.10), we relate the complex correlation phase and magnitude to the motion vector relative to individual A-lines (equation 3.11). Along with the correlation phase statistics, this is used to form a maximum-likelihood azimuthal motion estimator using normalized correlation phase alone (equations 3.12 - 3.15). In equations 3.16 - 3.20, Bayes' Theorem is used to develop a maximum-likelihood elevational motion estimator, given the azimuthal motion estimate and observed normalized correlation magnitudes. However, for adequate sampling this estimator requires probability evaluation over at least 10^8 variable permutations for each estimation, and so is impractical. Varying levels of simplification are used to derive two suboptimal but fast and practical estimators, in equations 3.22 and 3.23.

The statistical model with factorization was tested against fully-developed 3D speckle simulations and experimental data in Figs. 3.5 - 3.8 and table 3.2. The expected-value of complex correlation phase and magnitude are shown to match closely between model, simulation and experimental data in Fig. 3.5, with similar magnitude standard deviation (Fig. 3.6) and angular deviation (Fig. 3.7). In table 3.2, parametric fits to phase slope and gaussian magnitude are shown, with differences of $<=$10.4$\%$. Examples of the distribution of complex correlation in the complex plane are shown in Fig. 3.8, again with a close qualitative match. These results indicate that the statistical model with factorization is equivalent to a full 3D speckle simulation, and closely matches experimental data, within the limits of how well the modeled and actual PSFs match.

3.F.ii Speckle-Tracking Lateral Motion Estimation Bias Prediction

The rotation of PSFs with lateral translation in sector scan systems (62) is known to cause rapid decorrelation and degrade speckle-tracking motion estimation. The decorrelation acts to increase the ‘noise’ seen by the motion estimator when searching for correlation peaks (68), however we show qualitatively in Figs. 3.3 and 3.4 that in sector scans the PSF rotation causes lateral motion underestimation. The statistical complex correlation PSF model can be used to model correlation of detected data by a deterministic transform applied to complex correlation magnitude. Using this method, the model was used to quantitatively predict lateral motion underestimation by modeling the PSF rotation for individual A-lines and finding the location of detected correlation magnitude peaks for different shifts (as in Fig. 3.4). The shifted correlation peaks arising from PSF rotation during various shifts are shown in Fig. 3.16 for simulations and experimental data.

Qualitatively the correlation peaks are very similar, with a consistent left-shift (underestimation). The experimental correlations have lower peak magnitudes and larger shifts, this is likely due to inaccuracies in the separable 3D gaussian PSF model used in simulations relative to the real experimental PSF. The estimator bias for various A-line angles and shifts is shown in Fig. 3.17 for simulated and experimental data. Again, the forms of the bias graphs for the simulated and experimental data are similar, with the main difference greater bias for experimental A-lines with angles close to vertical. When averaged over all A-line angles, the average underestimaton of lateral motion (Table 3.3) ranges from $\sim 9\ \%$ (for 1.0 mm and 1.5 mm shifts) to $\sim 17\%$ (for 0.5 mm shift). This is consistent with the magnitude of lateral motion estimation bias from speckle-tracking performed on scan-converted, detected data (also table 3.3), which has underestimation of 9 to 10 $\%$. The statistical complex PSF correlation model has been shown to be accurate and applicable to detected correlation modeling, with the capability to improve motion estimation by iterative analysis. For example, from Fig. 3.17, it can be seen that motion estimator bias changes with A-line angle, and is largest for vertically angled A-lines, therefore scan-converted speckle tracking can possibly be improved by an angle-weighting scheme favoring outer sector-scan A-lines.

3.F.iii Joint Azimuthal-Elevational Motion Estimation Performance

Figure 3.9 shows azimuthal estimation performance for the joint estimation method, using the complex correlation statistical model with 3D gaussian separable PSFs, and PSFs simulated in Field II (57). Three cases are shown, pure azimuthal motion, pure elevational motion, and diagonal (equal azimuth/elevational) motion. In comparison, azimuthal estimation performance for experimental data using the joint estimation method is shown in Figs. 3.10 and 3.14, and using speckle-tracking on scan-converted data is shown in Fig. 3.14. Both simulation techniques indicate very accurate azimuthal motion estimation, with low variance, for displacements of up to 1.2 mm. Using the joint estimation method on experimental data performs similarly up to 1.0 mm displacements for azimuthal-only and elevation-only displacements, and works well up to 0.8 mm displacements for diagonal motion. Speckle-tracking motion estimation (Fig. 3.14) is effective for larger displacements, but exhibits a negative bias (underestimation) of 9 to 10 $\%$. Joint estimation method variance is approximately equal to the speckle-tracking variance, for motion up to 1 mm for azimuthal-only motion, or up to 0.8 mm for diagonal motion. These results indicate that the joint estimation method is more accurate than speckle tracking, with similar jitter for displacements with a small azimuthal component.

Elevational performance of the joint estimation method is shown for two kinds of simulated data (3D gaussian PSF, Field II PSF) in Figs. 3.11 and 3.12 respectively, and for experimental data in Fig. 3.13. Equivalent performance for conventional (i.e. decorrelation-only estimation on detected, scan-converted data) is shown in Fig. 3.15. The simulated cases include elevation motion estimation using the ‘accurate’ and ‘fast’ methods of equations 3.22 and 3.23 respectively. The ‘accurate’ method has much lower variance than the ‘fast’ method. However, this comes at the cost of increased computational complexity. For gaussian-PSF simulated azimuth-only motion of $<=$ 1.0 mm, both estimation methods give an elevation estimate error of $<$ 0.1 mm. For elevation-only motion, estimation is accurate, with low variance. In the case of diagonal motion, both estimation methods work well for shifts of up to 0.8 mm before failing. For equivalent simulations using the ‘Field’-generated PSF, performance is degraded, with rapidly increasing errors for displacements above 0.6 mm for azimuthal-only or diagonal motion. Elevation motion estimation performance for the joint (‘accurate’ algorithm) and conventional methods is compared in Fig. 3.15. For azimuthal-only motion, the conventional method has a large error that increases with azimuthal motion, due to lateral PSF decorrelation, while the joint method has an absolute elevational motion estimate of $<$ 0.1 mm for azimuthal shifts up to 1.0 mm. For diagonal motion, the conventional method overestimates elevational motion by on average 52 $\%$, while the joint method underestimates by 25 $\%$ on average.

A comparison of motion tracking performance during a semi-random walk over a spine analog target in a tissue-mimicking phantom is presented in Figs. 3.18, 3.19, 3.20, 3.21. The first figure shows the lateral-elevational path taken by the semi-random walk, and the path estimated by the joint estimation method and conventional method. This indicates that both the joint method and conventional method track lateral motion reaonably well, with a maximum lateral position error of 0.35 mm for conventional, and 0.11 mm for joint estimation.

However, for elevational estimation, the joint method significantly outperformed the conventional method, with a maximum elevational position error of only 5.88 mm, compared to 20.21 mm for the conventional method. When the motion estimates were used to form volumetric ultrasound data sets, and registration was performed against the known 3D spine model (Figs. 3.19-3.21), the improved performance of the joint motion estimation method was reflected in an RMS surface distance error of only 1.2 mm, compared to 2.43 mm for the conventional estimator, and 1.08 mm for the actual motion (with motion stage). The elevational underestimation of the conventional estimation method is surprising, as previous results indicated an overestimation was to be expected, due to azimuthal shifts causing decorrelation which is interpreted as elevational shifts. However, when conventional motion estimation was tested in semi-random walks over pure speckle phantoms (i.e. containing no spine analog), elevation estimates were consistently too high as expected (by $>=$20 $\%$). This indicates that the presence of the bright specular reflecting spine may be causing artificially high correlation between adjacent frames in the elevational dimension, leading to elevational underestimation. The joint estimation method is less affected by this, as the bright spots only affect a small number of A-lines. Therefore, in addition to significantly outperforming the conventional motion estimation for realistic freehand scan motion, the joint estimation method is also more immune to localized areas with non-fully-developed speckle. For large displacements (e.g. $>$ 1.2mm per frame), the sector-scan method fails in the azimuthal (y-z) directions, while conventional tracking methods continues to work successfully. However, with a typical sector-scan frame-rate of 20, only azimuthal translations at speeds above 24 m/s will suffer difficulties, so this should not be a problem in practice. Additionally, the sector-scan method can be extended to correlations between adjacent lines to support faster displacements. In total, the joint method has proven to have large advantages over the conventional method for diagonal motion, with resulting greatly-improved freehand scan performance.

3.G Conclusion

The statistical model for complex correlation with a PSF transformation introduced in this work is capable of accurately modeling the statistics of correlation-based motion estimators, including those using detected data. This has been verified using simulated and experimental data involving PSF shifts, and also by making theoretical predictions about estimator bias, subsequently verified in experiments. Using the model, we have established that the phase of normalized complex correlation for RF pixels from a window on a sector-scan A-line is effectively normally distributed with a mean given by a geometric vector relation. This forms the basis of a computationally simple, highly accurate maximium-likelihood azimuthal motion estimator using indidividual A-lines' complex correlation phase. Furthermore, given the azimuthal motion estimate, we have developed a maximum-likelihood elevational estimator and two faster but suboptimal elevational estimators. Used together, the joint azimuthal/elevational sector-scan motion estimator is shown to outperform conventional (speckle-tracking azimuthal/decorrelation elevational) motion estimation for small displacements, with smaller bias for azimuthal motion, and much smaller error for diagonal (azimuthal/elevational) motion. Experimental data indicates that the benefits of the joint sector-scan estimation method lead to significantly better freehand motion estimation, where diagonal motion is likely to be present. For applications where sector-scan systems are preferable (such as handheld bone imaging (18)), this work represents a significant enabling step for 3D freehand scanning.

3.H Appendices

3.H.i Derivation of SVD Components as a Function of Complex PSF Inner Product

If the matrix $P = (F^* G^*)$ is broken down by the SVD into $P = VSU^H$, then the matrix $VSSV^H$ can be related to the inner products between complex PSFs F and G, as in equation 3.25. As the matrix V is unitary, inner products between any rows or columns of V are by definition zero, leading to the constraints in equation 3.26. If F and G are identical PSFs except for an affine transform, then we can assert that $F^H F = G^H G$, and $F^H G = (G^H F)^*$, leading to equation 3.27. The constraints of equations 3.26 and 3.27 can only be satisfied if the elements of V have the form of equation 3.28, with the (diagonal, real) form of S also shown for convenience. If $VSSV^H$ is expanded using the form of 3.28, and held to be equal to the right hand side of 3.25, four relations, shown in 3.29-3.32 must be satisfied.

$$P^H P = VSSV^H = \begin{pmatrix} F^H F & G^H F \\ F^H G & G^H G \end{pmatrix} \tag{3.25}$$

$$V_{11}V_{12}^* + V_{21}V_{22}^* = 0, \;\; V_{11}V_{21}^* + V_{12}V_{22}^* = 0 \tag{3.26}$$

$$V_{11}V_{11}^* + V_{21}V_{21}^* = 0, \;\; V_{12}V_{12}^* + V_{22}V_{22}^* = 0 \tag{3.27}$$

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{ja} & e^{jb} \\ e^{jc} & e^{jd} \end{pmatrix}, \;\; S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix} \tag{3.28}$$

$$S_1^2 + S_2^2 = F^H F \tag{3.29}$$

$$S_1^2 + S_2^2 = G^H G \tag{3.30}$$

$$S_1^2 e^{j(a-c)} + S_2^2 e^{j(b-d)} = G^H F = |G^H F| e^{j\Phi}, \Phi = arg(G^H F) \tag{3.31}$$

$$S_1^2 e^{-j(a-c)} + S_2^2 e^{-j(b-d)} = F^H G = |G^H F| e^{-j\Phi}, \Phi = arg(G^H F) \tag{3.32}$$

Equations 3.31 and 3.32, along with the requirement that $|G^H F| <= |F^H F|$ can only be satisfied if $(a - c) = \Phi$, and $(b - d) = \pi + (a - c) = \pi + \Phi$. Noting that in 3.31 and 3.32, only the relative values $(a - c)$ and $(b - d)$ affect the right hand sides, we can set $a = b = 0$, leaving $c = -\Phi$, and $d = -\pi - \Phi$, with the resultant form of V shown in equation 3.33. This leads to equation 3.34 relating the variables S_1 and S_2 to the inner product of the complex PSFs F and G. Equations 3.29 and 3.34 together form a simultaneous equation in two unknowns, that can be solved to give expressions for S_1,S_2 and S_2/S_1, shown in equations 3.35 and 3.36.

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ e^{-j\phi} & -e^{-j\phi} \end{pmatrix} \tag{3.33}$$

$$S_1^2 - S_2^2 = |F^H G| \tag{3.34}$$

$$S_1 = \sqrt{\frac{1}{2}(F^H F + |F^H G|)},\ \ S_2 = \sqrt{\frac{1}{2}(F^H F - |F^H G|)} \tag{3.35}$$

$$\frac{S_2}{S_1} = \sqrt{\frac{1 - |\frac{F^H G}{F^H F}|}{1 + |\frac{F^H G}{F^H F}|}} \tag{3.36}$$

Chapter 4

REAL-TIME 3D SPINE IMAGING USING ROBUST MULTIMODAL MOTION TRACKING[1]

4.A Abstract

Spinal bone imaging is an important ultrasound application, as in patients with a high body-mass-index (BMI), epidural procedures often fail (11) due to an inability to locate the relevant lumbar spine anatomy. Fluoroscopy is another alternative for epidural procedure guidance, however this is undesirable due to the ionizing radiation involved. Conventional 2D (B-Mode) ultrasound has been used successfully in spinal bone imaging (1), however most systems have difficulties imaging bone (18), and require a high level of skill and experience to operate. Freehand 3D ultrasound systems are more intuitive to use, as they build up a 3D image from the motion of a 2D (B-mode) transducer (59), however no handheld freehand 3D ultrasound systems currently exist.

[1]Chapter 4 will be submitted as the following journal article :
K. Owen, F. W. Mauldin, Jr., S. Nguyen, M. Tiouririne and J. A. Hossack, "Lumbar Spinal Bone Anatomy Imaging using Ultrasound and Robust Multimodal Motion Tracking", *Computerized Medical Imaging and Graphics.*

In this chapter, the ability of a handheld freehand 3D ultrasound system (the 'SpineFinder') to accurately localize bone surfaces using both ultrasound and camera-based motion tracking is tested. A theoretical model for optimal combination of multi-modal sensor estimates is developed and calibrated using a realistic lumbar spine phantom, for translations with and without rotational motion. Bone localization efficacy is tested using a novel 3D registration and visualization methodology, complemented by a receiver operating characteristic (ROC) analysis of bone/speckle classification performance. Without rotational motion, RMS bone surface accuracy of 1.12 mm is achieved, rising to 2.2 mm with realistic amounts of rotation. This indicates that performance is adequate without accounting for all rotational motion, but that doing so will improve performance.

4.B Introduction

Epidural and spinal anesthesia procedures require precise location of the epidural gap in the L3-L5 lumbar region of the spine. The current standard of care is the 'blind' method (10), where manual palpation alone is used to locate the spinous processes in the lumbar region, and by visual interpolation the correct entry point and angle for the needle is identified. However in the growing obese patient population (85), this method is increasingly difficult, contributing to failure rates of 40% to 70% (11), (12), (13) and poor patient outcomes (11). Fluoroscopic procedure guidance uses high quality real-time images (as in Fig. 4.1), but exposes the patient to ionizing X-ray radiation, recently shown to be responsible for 2% of all cancers and up to 11,000 deaths per year in the U.S.

(14), (15). Recently, ultrasound has begun to be used for epidural procedure guidance, with improved success rates (1) and no radiation risk.

However, conventional B-mode ultrasound requires considerable skill and experience to mentally register the instantaneous scan plane position relative to the target area, in contrast to 3D ultrasound imaging which requires much less training and skill (4). Current 3D ultrasound systems are typically bulky and expensive, due to the degree of transducer and beamforming complexity associated with large 2D transducer arrays. Simpler and cheaper 'freehand' 3D ultrasound systems have been suggested and implemented, using motion tracking technology to register many 2D (B-mode) image planes into a volume (59), (69), (64), (65). Motion tracking within the azimuthal ultrasound plane is straightforward and high quality using speckle-tracking (60), however out-of-plane motion detection using ultrasound has diminished accuracy (82), and can only be used to find motion magnitude not direction (61). Various methods have been proposed to overcome the elevational accuracy and directional ambiguity (86), (87), including modified transducers for improved elevational tracking (63), (88), magnetic sensors (65), combined inertial/optical sensors (64) and constrained mechanical motion (4).

The existing and described freehand 3D imaging systems appear to uniformly add extra motion tracking capabilities (or algorithms) to existing B-mode ultrasound systems. Although the resulting 3D images are more intuitive to understand, mental registration of the image to the probe position is still required, and the addition of extra motion tracking sensors likely hinders portability. In this chapter, I introduce the 'SpineFinder', an intuitive handheld freehand 3D imaging system with collocated display and transducer for intuitive imaging, specialized sector-scan transducer optimized for bone imaging (18) and an inexpensive camera module used to enhance motion tracking performance.

Previous work indicates that cameras can be used to improve tracking accuracy in combination with ultrasound-based tracking (89). The 'SpineFinder' similarly exhibits improved overall motion estimation when optimally combining motion from both sensor modalities, after calibration. To assess freehand 3D spine imaging performance, motion estimation using ultrasound, camera and optimally combined data is performed using a custom phantom with embedded human spine analog and aberrating tissue layer. A known 3D CAD model of the embedded spine is used, along with sophisticated registration algorithms to quantify root-mean-square (RMS) bone surface error between the 3D ultrasound data and the spine model. Additionally, the problem of discriminating between bone targets and locally bright speckle is analyzed using the receiver operating characteristic (ROC) formalism. Finally, using a high-quality measurement of in-plane rotation using sector-scan geometery, in combination with the 3D registration metrics, the relative importance of rotational motion estimation is analyzed, along with the possibility that using a 3-axis accelerometer for rotational measurements could improve overall estimation.

4.C Background

4.C.i Multi-Modal Motion Estimation

Motion estimates formed using different sensor modalities have different capabilities and properties. Sensor data from an ultrasound transducer, camera or accelerometer can be analyzed to yield a subset of the entire motion parameter

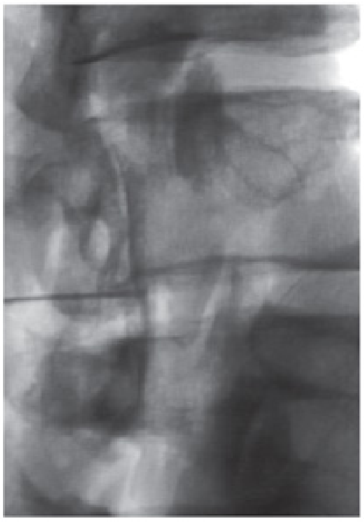

Figure 4.1: Representative image from fluoroscopically guided epidural procedure, with needle to left of image.

set, but not the entire parameter set. A three-dimensional object has six degrees of freedom, three translational (front-back, left-right and up-down), and three rotational (roll, pitch, yaw). These parameters are indicated pictorially in Fig. 4.2, relative to a B-mode scan plane.

Camera-based motion estimation involves acquiring a series of images at sequential points in time, and performing image registration between successive frames to estimate the particular transform occuring between the frames. There are many methods to perform image registration (90), (91), broadly broken down into the categories of feature-based, and area-based detection. Images containing artifical components (e.g. maps, street-views, radar images) are often amenable to feature extraction before registration. However, medical images typically do not have easily identifiable features without expert intervention (91), and so for these images area-based registration is preferable in automatic systems. In general, image transforms occuring between frames

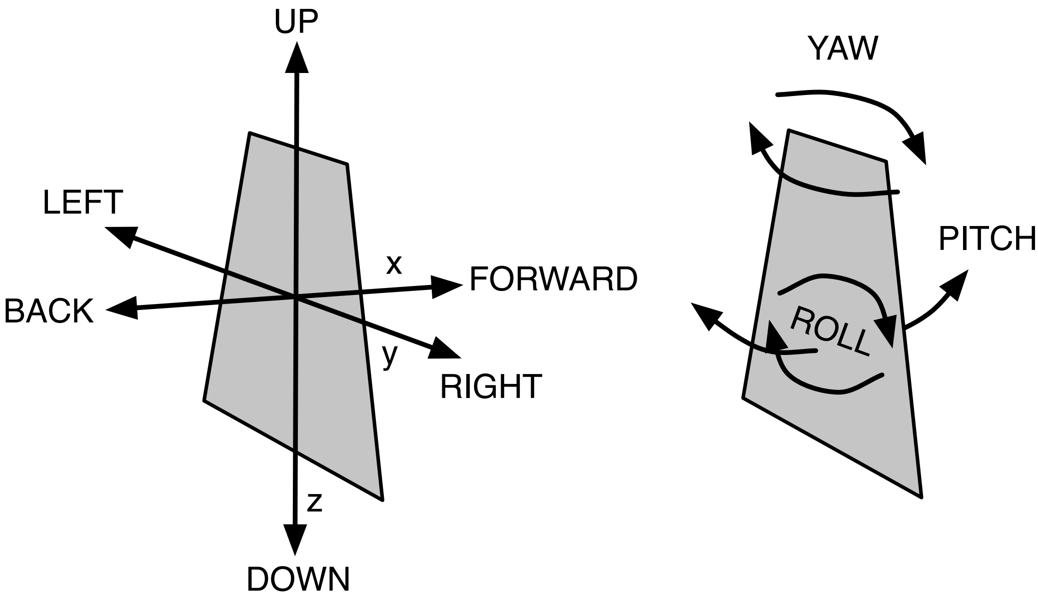

Figure 4.2: Six degrees of freedom required to uniquely identify the position and orientation of an object. On the left, the BACK-FRONT(x), LEFT-RIGHT (y) and UP-DOWN (z) translational degrees of freedom are indicated. On the right, the rotational degrees of freedom are shown, including rotation about the x axis (ROLL), y axis (PITCH) and z axis (YAW). The gray trapezoid represents a B-mode scan plane.

may include global transforms, for example including affine, projective, perspective and polynomial transforms, and local transforms which are spatially variant image changes. For the purposes of this work, only global transforms will be considered, with the further assumption that the image sensor will be approximately parallel to the skin, so that only image translations and rotations are possible.

A B-mode ultrasound transducer images a two-dimensional plane that is typically perpendicular to the skin surface. Unlike the camera, the resolution cell of this modality is a three-dimensional volume, corresponding to the point-spread-function of the system. This allows a larger subset of the six degrees-of-freedom to be estimated. However, ultrasound-based motion estimation is anisotropic in resolution and accuracy due to the narrow interrogated volume in the out-of-plane direction (82), and can only estimate magnitude (not direction) of out-of-plane motion components (61). Various methods are used to detect motion in the azimuthal (y-z) plane and elevational (x) direction, including speckle-tracking (60), decorrelation-based methods (61), and the joint azimuthal-elevational estimation technique for sector-scan systems from chapter 3. As mechanically scanned sector systems using a single piston transducer are efficacious in bone imaging (18), and the joint estimation method has been show to outperform the other methods, it will be exclusively used here.

Inexpensive micromachined sensors for measuring acceleration and other inertial properties have become widely available in recent years (92). Although it is theoretically possible to integrate displacement from acceleration, using methods such as Verlet integration (93), errors and drift quickly accumulate in practice. This means accelerometers can only be used for displacement estimation over very short distances and time periods, even when special, optimal calibration methods are employed (94). However, low-cost accelerometers are effective at measuring the projection of vector acceleration due to gravity $\hat{g}$ onto three accelerometer axes, providing some orientation information.

The estimation capabilities of the three modalities are summarized in table 4.1. It is clear that for a full determination (i.e. including magnitude and direction for each parameter) of the position and orientation of the system, information from all three modalities is required.

Table 4.1: Motion Estimation Capabilites of Different Sensor Modalities

Modality/Parameter	X	Y	Z	Roll	Pitch	Yaw
Camera	Yes	Yes	No	No	No	Yes
Ultrasound	Magnitude	Yes	Yes	Yes	Magnitude	Magnitude
3-Axis Accelerometer	No	No	No	Yes	Yes	No

4.C.ii Probabilistic Combination of Multi-Modal Motion Estimates

Motion estimates from different sensing modalities can be combined in a number of ways. However, one simple and effective way to do so is to find the most likely motion given the various estimates from different modalities. This requires a probabilistic model for the distribution of each modality's estimate, given the true motion, shown for ultrasound and camera-based estimators in equations 4.1 and 4.2, where r is the true motion, r_U and r_C are the ultrasound and camera motion estimates, b_U and b_C are the estimator biases, σ_U^2 and σ_C^2 are the estimator variances, and independent normal distributions are assumed. Using normal distributions for the estimators is an approximation.

However this is typically justified as the central limit theorem (CLT) (73) dictates estimators with a lot of averaging converge to a normal distribution, and this is a common practice for unknown probability distrutions whose first two moments are known (75). Qualitatively, the estimators used in this work have unimodal distributions, further justifying the approximation. If the different modality estimators have independent normal distributions, this further simplifies probability maximization, as scalar factored probability expressions can be used. In this case, although the ultrasound estimator is affected by the aberrating layer, it is reasonable to assume that the surface texture imaged by the camera is uncorrelated with the ultrasound aberration properties.

Equation 4.3 shows the probability of a motion r given the ultrasound and camera estimates, and assumed bias and variance from a normally distributed PDF (probability distribution function) for both modalities, using Bayes' Theorem (75). In equation 4.4 this is further developed to show the factorization of the individual estimator PDFs, and in equation 4.5 is the final equation for most-likely combined estimate $\hat{r}$, given the individual modality estimates r_U and r_C, and the stated assumptions. Differentiation of equation 4.5, and setting the derivative to zero yields a closed-form solution for most-likely motion $\hat{r}$, shown in equation 4.6. Although equations are shown here for combination of estimations from ultrasound and camera-based sensor data, this method is generally applicable to combining several, independent motion estimates.

$$P(r_U|r) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma_U^2}(r_U-(r+b_U))^2} \tag{4.1}$$

$$P(r_C|r) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma_C^2}(r_C-(r+b_C))^2} \tag{4.2}$$

$$P(r|r_U, r_C) = \frac{P(r_U, r_C|r)P(r)}{P(r_U, r_C)} \tag{4.3}$$

$$P(r|r_U, r_C) \propto e^{-\frac{1}{2\sigma_U^2}(r-(r_U-b_U))^2} e^{-\frac{1}{2\sigma_C^2}(r-(r_C-b_C))^2} \tag{4.4}$$

$$\hat{r} = \underset{r}{\operatorname{argmax}}\, e^{-\frac{1}{2\sigma_U^2}(r-(r_U-b_U))^2} e^{-\frac{1}{2\sigma_C^2}(r-(r_C-b_C))^2} \tag{4.5}$$

$$\hat{r} = \big(\frac{1}{\sigma_U^2} + \frac{1}{\sigma_C^2}\big)^{-1}\Big(\frac{r_U - b_U}{\sigma_U^2} + \frac{r_C - b_C}{\sigma_C^2}\Big) \tag{4.6}$$

4.C.iii Spine Imaging Assessment Using Receiver Operating Characteristic Analysis

In this application the lower lumbar spinal bone anatomy is imaged using ultrasound, which has an intrinsically bright background ‘noise’ signal in the form of speckle, with a theoretical signal-to-noise ratio of approximately 1.91(81). Although bone reflections are up to 35 dB brighter than background tissue (i.e. speckle) (95), various effects including oblique beam incidence on a specular reflector, the presence of a phase-aberrating subcutaneous fat layer (96), (97) and intervening ligaments, nerves or vascular elements act to reduce bone-speckle contrast. Qualitatively, spinal bone anatomy imaging performance can be assessed based on the ability to both correctly identify spinal bone surface locations, and also discriminate between which ultrasound reflections are bone rather than locally bright speckle regions. Receiver operating characteristic (ROC) analysis is a formal framework for performing this type of quantitative performance analysis (98). ROC analysis was first used to measure the discriminatory performance of systems in binary identification (friend-or-foe) of military radar targets. More recently, this form of analysis has been used extensively to evaluate the efficacy of medical diagnostic systems (99), including diagnostic medical imaging (100).

For spinal bone imaging, ROC analysis is useful in numerically assessing bone detection sensitivity (what fraction of the bone surfaces are detected by ultrasound), and specificity (the fraction of non-bone that is correctly identified as non-bone). More formally, in the classification of detected surfaces, four categories are possible, shown in table 4.2. Using the true positive (TP), false positive (FP), true negative (TN) and false negative (FN) rates, two very useful metrics can be constructed. Sensitivity, defined in equation 4.7, measures how successful the system is at detecting all the bone surfaces. Specificity, defined in equation 4.8, is a measure of how well the system identifies non-bone surfaces. A perfect system will have a sensitivity of unity (detect all possible bone surfaces), and specificity of unity (identify all non-bone surfaces correctly). In a real system, if a threshold is used to detect bone surfaces, it is easily possible to achieve high sensitivity by setting a low threshold, but significant amounts of bright, non-bone speckle signal will be detected as bone (FP), leading to poor specificity. Similarly, with a high threshold, specificity will be increased (less misidentification of bright speckle as bone surface), but some less-bright bone reflections will be classified as non-bone, therefore reducing sensitivity.

Table 4.2: Bone Surface Classification Categories for ROC Analysis

Category	Decription
Positive (P)	Bone surface
Negative (N)	Non-bone surface
True Positive (TP)	Bone surface classified as bone surface
False Positive (FP)	Non-bone surface classified as bone surface
True Negative (TN)	Non-bone surface classified as non-bone surface
False Negative (FN)	Bone surface classified as non-bone surface

$$sensitivity = \frac{TP}{P} = \frac{TP}{TP + FN} \tag{4.7}$$

$$specificity = \frac{TN}{N} = \frac{TN}{FP + TN} \tag{4.8}$$

As system parameters, such as the bone classification threshold are varied, different sensitivity and specificity values arise. If the sensitivity and specificity values are plotted on a 2D graph as the active parameter is varied, a locus is traversed, as in the example graph of Fig. 4.3. Here, the baseline performance given by a random classifier is shown, which will operate at a point along the diagonal line ranging from perfect sensitivity, zero specificity to zero sensitivity, perfect specificity. A more realistic example ROC curve is also shown in Fig. 4.3, which is always at or to the upper-left of the random ROC curve. The optimal operating point for ROC curves depends on the relative importance of sensitivity and specificity in a particular application, however one general-purpose optimal operating point is the point on the curve that is closest to the top-left of the graph in a Euclidean sense, representing both good sensitivity and specificity.

In a realistic bone-detection system, the amount of bright speckle signal will vary with individual patients, so that the parametric variable (bone detection threshold) will be varied by the operator during use. Therefore, the area under the ROC curve is a way to measure classifier performance over all possible parametric variable values. The ROC curve area provides a measure of how well the system in general classifies, so that the effect of changing ultrasound parameters such as transmit pulse, bone detection filter characteristics, imaging mode and beamforming variables can be assessed.

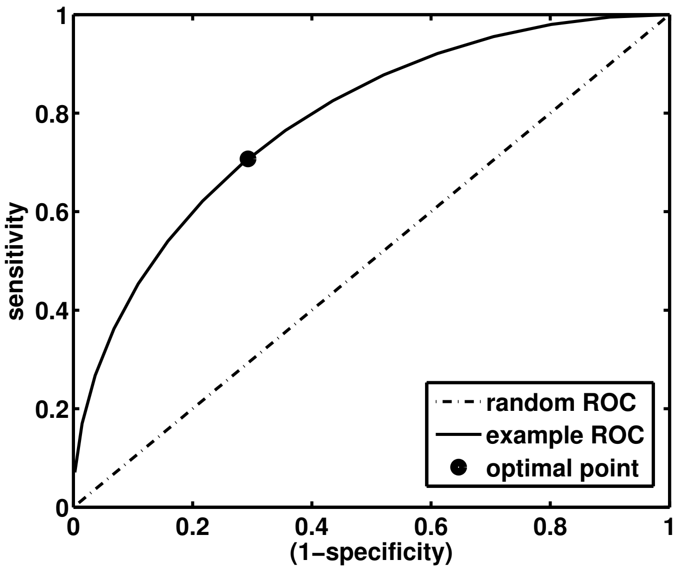

Figure 4.3: Example ROC curves, including from a random classifier, and a generic example classifier. Also shown is the optimal operating point on the example curve, defined as the point on the curve with the smallest Euclidean distance to the top-left of the graph.

4.D Materials and Methods

4.D.i Handheld Spinal Bone Imaging System

A custom ultrasound device was designed and constructed to evaluate spinal bone imaging using multi-modal motion estimation (The 'SpineFinder'). The system includes an ARM (Advanced RISC Machine)-based microprocessor, along with hardware to interface to an ADCM 2650 CMOS camera module (Agilent, Santa Clara, CA), an ADXL 345 three-axis accelerometer (Analog Devices, Norwood, MA), and an ultrasound transmit-receive board. A block diagram of the system is shown in Fig. 4.4. The ultrasound electronics are capable of high-voltage transmit and receive on two channels simultaneously, and are optimized to drive a mechanically scanned transducer, such as the single-piston, 60° sector scan transducer (Interson, Pleasanton, CA) with parameters given in table 3.1. The camera module has up to VGA (640 x 480) resolution, a variable focus, high quality F/2.6 lens and auto-exposure and gain features for operation in a range of lighting conditions. The SpineFinder device is capable of both independent real-time display of 2D/3D ultrasound echoes, and real-time capture of multimodal sensor readings to a storage device (Secure Digital card) for later download and offline analysis.

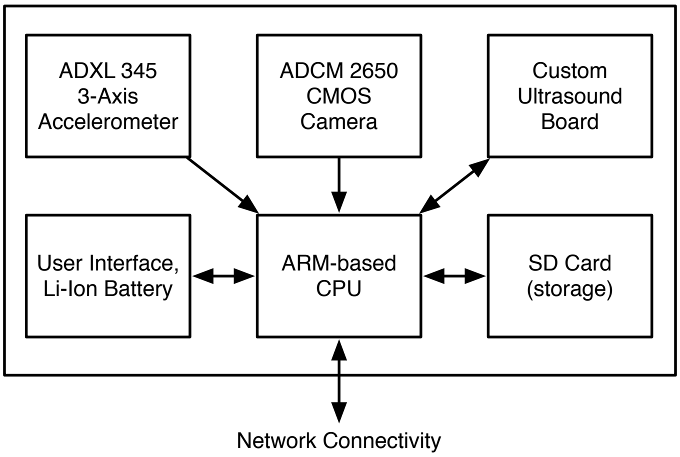

Figure 4.4: Block diagram of the 'SpineFinder' experimental handheld spinal bone imaging system. The three different sensor modalities used for motion detection are shown, along with the CPU, mass storage (Secure Digital or SD Card), network conectivity and auxiliary features such as rechargeable battery and user interface components.

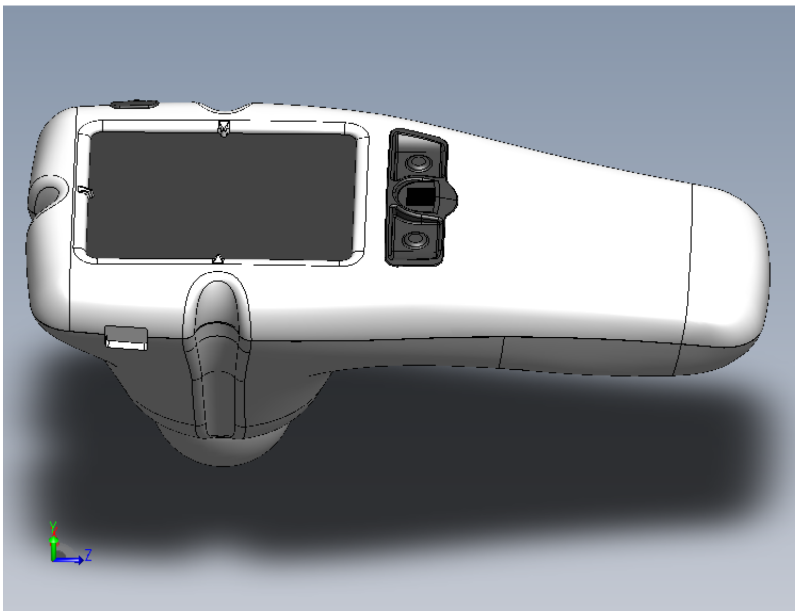

Figure 4.5: Top view of the 'SpineFinder' experimental handheld spinal bone imaging system. The user interface components (touchscreen, buttons, optical finger sensor) are visible, along with the hemispherical sector-scan ultrasound transducer at the very bottom of the system.

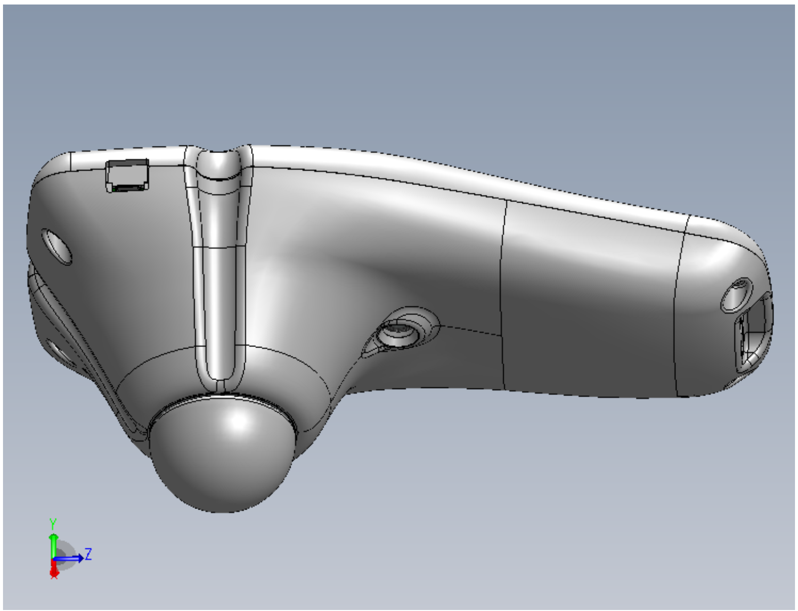

Figure 4.6: Bottom view of the 'SpineFinder' experimental handheld spinal bone imaging system. The camera circular aperture is clearly visible in the device handle, to the right of the ultrasound transducer.

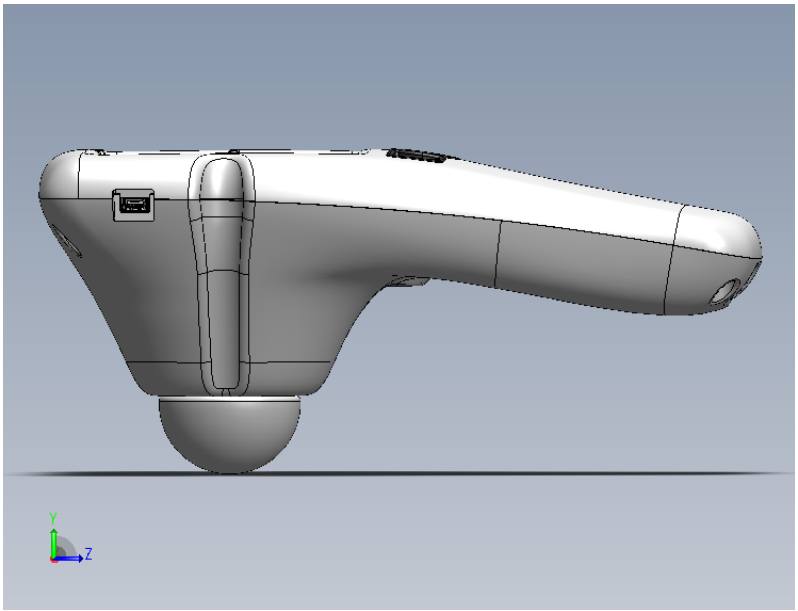

Figure 4.7: Side view of the 'SpineFinder' experimental handheld spinal bone imaging system. The preferred orientation of the device is shown, with the ultrasound transducer oriented perpendicularly to the skin surface, and the camera aperture parallel to the skin.

The SpineFinder system is illustrated in Figs. 4.5, showing display and user controls, 4.6 showing the relative locations of the ultrasound transsducer and camera, and 4.7 with the side-view, indicating the preferred orientation of the device over and parallel to the skin surface. The accelerometer is located on the CPU board inside the device, close to the mounting holes as per the device manufacturer's recommendations (101).

Acoustic output of the device was measured under various conditions in accordance with the Food and Drug Administration (FDA) guidelines for medical ultrasound devices (102), and was found to have maximum values for mechanical index (MI) and derated intensity, spatial-peak temporal-average ($I_{\mathrm{SPTA.3}}$) of 0.27 and 3.7 mW/cm^2 respectively, significantly under the corresponding FDA maximum-permissible values of 1.9 and 720 mW/cm^2.

4.D.ii Calibration and Testing of Camera and Ultrasound Motion Estimation

To calibrate the performance of camera and ultrasound-based translational motion estimation, controlled translations of the device were performed using an MM3000 motion stage (Newport, Irvine, CA), relative to a special tissue mimicking phantom with embedded spine analog and an overlaid 10-12 mm thick layer of porcine tissue (country ham). The spine analog is a 3D-printed CAD model of the human L3-L5 vertebrae, as described in chapter 3, in a gelatin/graphite speckle-producing background(72). The porcine tissue layer has two functions - to provide an optical target with surface texture broadly similar to human skin, and also to mimic the speed-of-sound variations and other phenomena that cause ultrasound phase aberration in human tissue (96), (97).

For ultrasound-based motion estimation, the joint azimuthal-elevational estimation technique described in chapter 3 was used, with the algorithm associated with equation 3.23 used to calculate elevational motion. A recalibration of the decorrelation curve parameters α and γ from chapter 3 was performed, using the methodology of chapter 3, sub-section D.ii, as it is known that the inhomogeneities in real tissue (i.e. the porcine layer) affect the PSF and decorrelation statistics (96),(69).

The ADCM 2650 camera on the system was operated in QVGA portrait (240 x 320), 8-bit grayscale mode with automatic exposure and gain controls enabled. The variable focus was manually adjusted to bring the porcine layer into sharp focus when the device is in position for ultrasound scanning (as in Fig. 4.7). In this orientation, the camera sensor is parallel to the porcine layer, and each image pixel corresponds to the illumination of an area on the porcine tissue with the same aspect ratio as the sensor pixel, dimensionally scaled only by a factor due to the lens and lens-tissue distance. That is, there is no distortion due to pitch or roll motion of the device while attached to the motion stage. To estimate translational motion between camera frames, 2D normalized cross-correlation (60), (66),(90), (91), with cosine-fitting was used with a 240 x 320 window from one frame, and a 180 x 260 kernel from the next frame, giving a maximum 30-pixel interframe displacement, corresponding to approximately 3 mm, or 3 cm/s motion with a 10 frame/s imaging rate. To qualitatively assess camera image motion estimation and image quality, all camera images taken during a single walk were combined at the appropriate positions into a larger composite image.

Calibration of the ultrasound and camera translational motion estimation parameters was performed using semi-random walks over the phantom, as described in sub-section D.iv of chapter 3. For each semi-random walk, camera and ultrasound data captures were performed after each motion step. Accelerometer data was not captured for this experimental sequence, as there is no rotational motion involved, and the motion stage moves very slowly. The translations in each dimension corresponding to a single-pixel shift in the camera image were found by a least-squares fit to the actual translations over two training random walks.

Following this, three independently generated semi-random walks over the phantom were performed, with ultrasound and camera data used to estimate translational motion for each step. Data from the three walks was used to calculate the bias, variance and error distribution for both modalities, following equations 4.1-4.2, and using equations 4.3-4.6, combined motion estimates for each step were formed. To assess the performance of camera, ultrasound and combined motion estimates on spinal bone surface localization accuracy, the ultrasound B-mode 'slices' acquired during a walk were processed to detect bone surfaces as described in chapter 3, sub-section D.iv, before being placed in a rectilinear volume using motion estimates and imported into the VTK software package for further processing. Polygonal surfaces corresponding to the ultrasound data and 3D CA model were created in VTK, and iteratively registered using translations and rotations as previously described. For each walk and motion estimation method, this process produces an RMS distance error between the ultrasound surfaces and the 3D CAD model surface, that can be used to assess motion estimation efficacy.

4.D.iii Freehand 3D Spine Imaging Using Multi-Modal Motion Estimation

For freehand 3D imaging assessment, the system was specially programmed to image for 10 seconds (to allow a good start point to be found), followed by 25 seconds of imaging with synchronized capture of ultrasound, camera and accelerometer data to on-device storage (SD card). This results in 250 multi-modal data records, which are downloaded from the device for offline processing. Due to limited storage space and bandwidth to the storage device during real-time imaging, the ultrasound data was processed on the device, including edge detection and scan-conversion as described in chapter 3, producing a single 160 x 120 ultrasound image per frame. This reduces the ultrasound data frame size from 1.57 MB to 38.4 KB. As the full ultrasound data set is no longer available for off-line processing, the 256 individual A-line complex correlations using in the sector-scan motion estimation algorithm of chapter 3 are calculated on the device and stored with each frame for later analysis and use in motion tracking.

Three freehand scans were recorded, in each case the same phantom target (spine in tissue-mimicking background with porcine tissue) as the previous section was used. In the first 10 seconds, the start of the 3D embedded spine was found using the live 2D image on the device display, following this a scan was made for half of the recording time along the length of 3 spine vertebrae, followed by a direction change and return to the start position. For a freehand scan, despite every attempt to move the device evenly without any rotation or deviation from a straight line, rotation and a meandering path is inevitable in practice. It is unclear to what extent rotational information is needed, if at all, in forming 3D freehand spine images with sufficient quality to guide lumbar puncture procedures. With a sector scan system, high-quality measurements of the 'roll' rotation can be made, using a 1D cross-correlation with cosine fitting for shifts across the differently-angled A-lines. For comparison, the device orientation was also determined using the inverse tangent of the ratio of the projection of the gravity vector onto the x and z axes of the devce. This data was collected for each frame to quantify the degree of angular rotation occuring during a freehand scan with no intentional rotation.

The ultrasound and camera tracking estimates were combined for each freehand scan, using equations 4.3-4.6 as in the previous section. As ultrasound tracking alone does not provide elevational motion direction, the sign of camera elevational motion was found by thresholding low pass filtered (10-tap boxcar) camera elevational motion. This allows fair comparison of the magnitude of ultrasound estimates. However, for freehand scans, no 'ground truth' motion is available. Instead, the ultrasound, camera and combined motion estimates are used to form bone surfaces in a 3D volume as in the previous section, which are similarly used to register with the 3D spine model and quantify bone surface RMS error.

Finally, a freehand scan was performed on the spine of a human subject, BMI = 27.1, with motion estimation identical to the ex-vivo freehand scans. This data was analyzed and processed to produce 2D and 3D images, however as the shape and size of this spine is unknown, only qualitative analysis of these data will be performed. This experiment exposes the subject to no health risks, as the measured acoustic output of the device falls well under FDA limits for mechanical index (M.I.) and acoustic intensity.

4.D.iv Receiver Operating Characteristic Analysis of Freehand 3D Scans

The receiver operating characteristis (ROC) curve was calculated for one of the 3 freehand scans, using ultrasound-only, camera-only and combined ultrasound-camera motion tracking. To find the ROC curve, 3D surface data from the ultrasound scan with B-mode slices placed using motion estimation was used, with the 'iso-contour' threshold as the parametric variable. The iso-contour value is a threshold, below which edges from B-mode slices do not get translated into polygons in the VTK toolkit. The iso-contour threshold was varied over a range of values, and for each the true-positive, false-positive, true negative and false-negative counts were taken, as defined in table 4.2. As the 3D spine model and the ultrasound data points are not sampled on a regular grid, a series of regularly spaced 4 mm x 4 mm x 4 mm voxels were each classified as TP, FP, TN or FN. To perform this classification, for each voxel a determination was made whether any point from an ultrasound surface or from the 3D spine model was within a 4 mm radius, then a simple truth table (table 4.3) was used to choose the correct ROC category. Using the ROC curve for each estimation method, the ROC curve area and optimal sensitivity/specificity values were determined, the latter corresponding to the point on the curve with the smallest Euclidean distance to the top-left of the graph.

Table 4.3: Truth Table for ROC Voxels

Ultrasound Surface $<=$ 4 mm ?	Model Surface $<=$ 4 mm ?	ROC Classification
No	No	True Negative
No	Yes	False Negative
Yes	Yes	True Positive
Yes	No	False Positive

4.E Results

4.E.i Combined Ultrasound/Camera Motion Estimation

The performance of the ultrasound-only, camera-only and combined ultrasound-camera motion estimates over a series of 3 semi-random walks over a tissue-mimicking phantom with embedded spine analog and overlaid porcine tissue layer are shown in table 4.4. Azimuthal and elevational RMS estimation error, normalized correlation (to known motion steps) and bias are shown separately for each estimation variety. Fig. 4.8 shows the distributions of ultrasound, camera and combined elevational and azimuthal motion estimates over the 3 walks. Composite images of the porcine tissue (country ham) produced by combining multiple camera images at the positions indicated by camera motion estimates are shown in Fig. 4.9. RMS bone surface distance errors for registration of the ultrasound volume images to the 3D spine model, when using the different motion estimation methods are shown in table 4.5, for all individual random walks, plus average values across all 3 walks. Three-dimensional superimposed surface renderings of the ultrasound-detected surfaces and 3D spine model are shown in Figs. 4.10, 4.11, 4.12 and 4.13 for ground-truth motion, ultrasound-derived motion, camera-derived motion and combined ultrasound/camera derived motion respectively, for the second random walk.

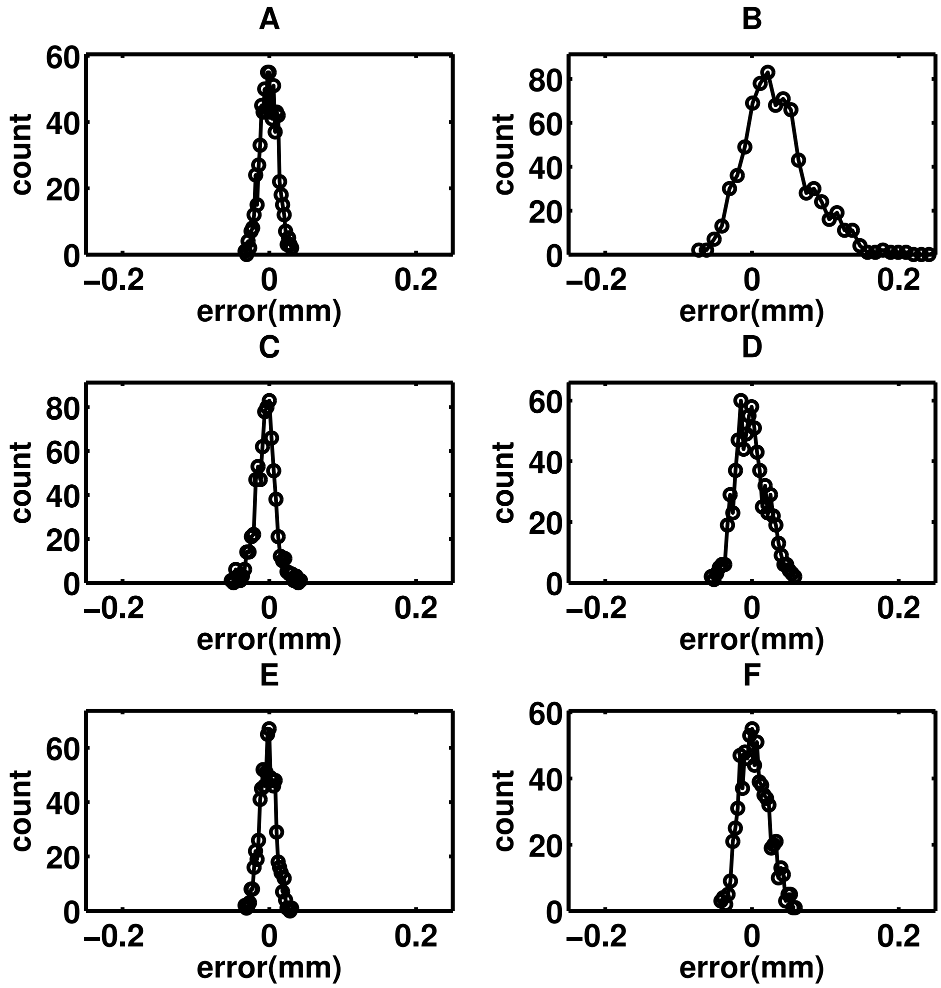

Figure 4.8: Histograms of motion estimation error for ultrasound azimuthal (A) and elevational (B), camera azimuthal (C) and elevational (D), and for optimally combined azimuthal (E) and elevational (F) estimation.

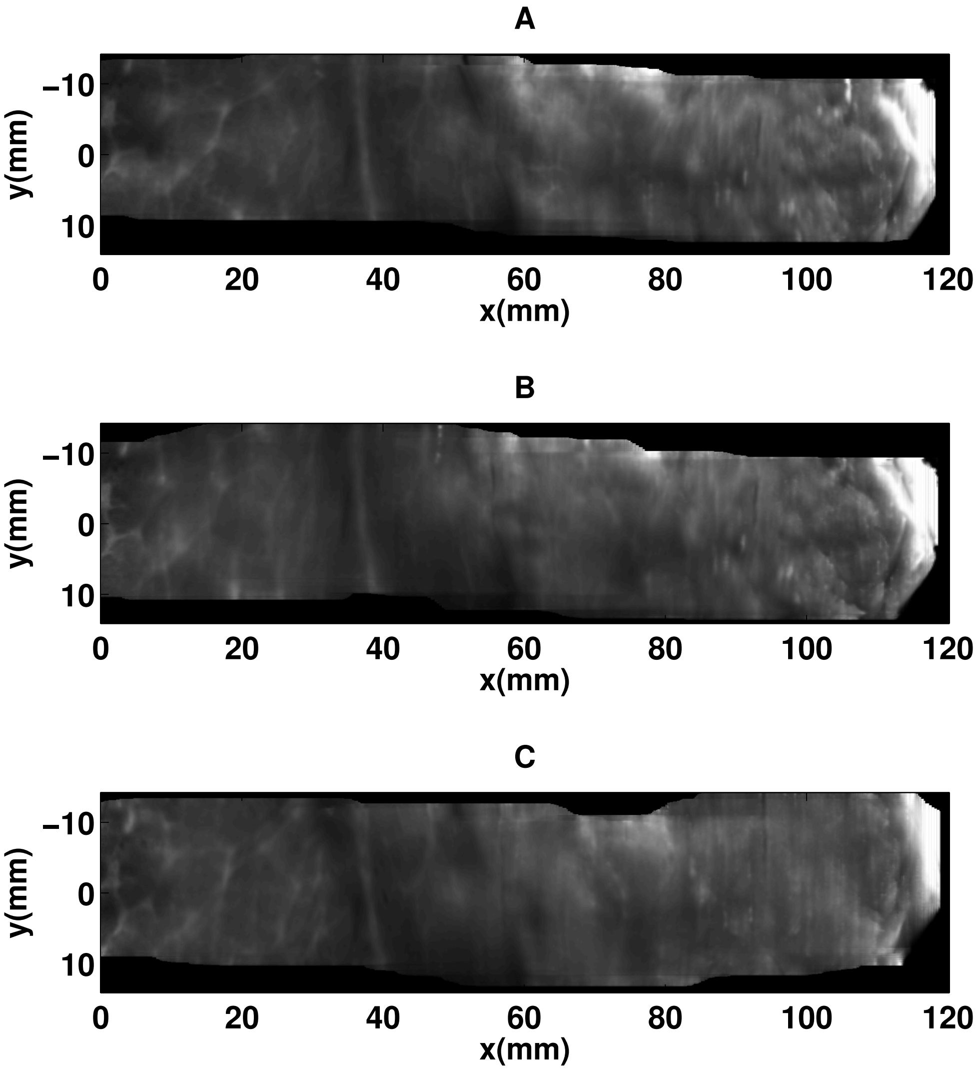

Figure 4.9: Montage of individual photos taken by camera on semi-random walk over phantom with overlaid porcine tissue (country ham slice, 10 mm). Three runs with different random walk motion are shown in A, B and C.

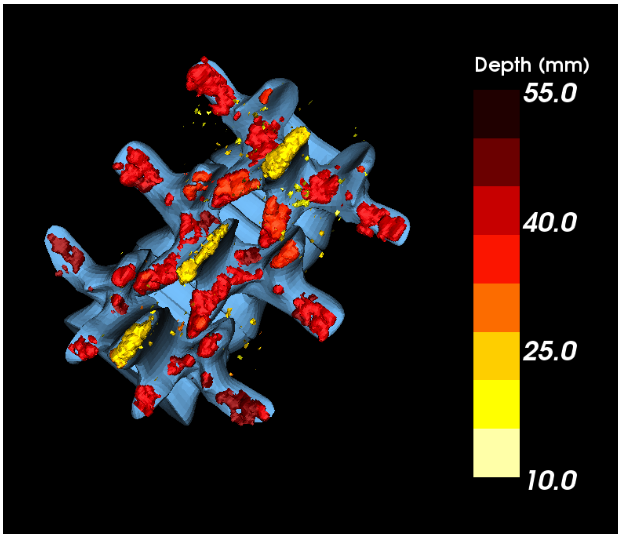

Figure 4.10: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using ground-truth location for each sector-scan image.

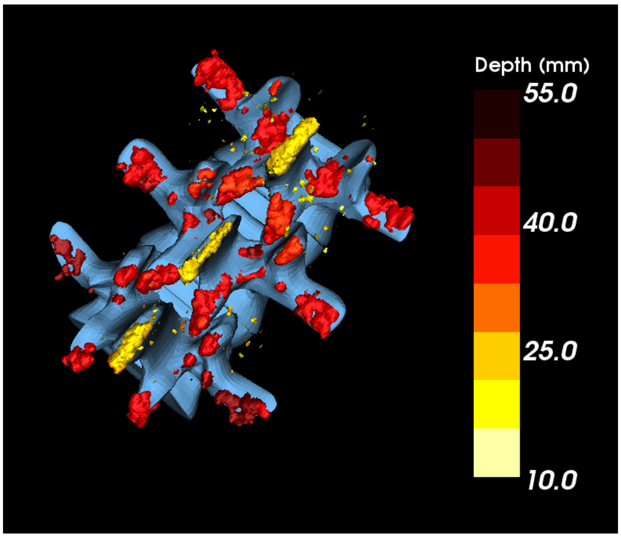

Figure 4.11: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using ultrasound-derived location for each sector-scan image.

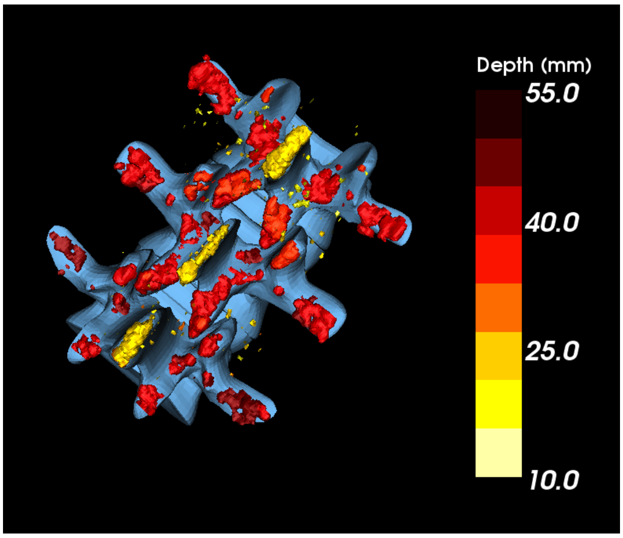

Figure 4.12: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using camera-derived location for each sector-scan image.

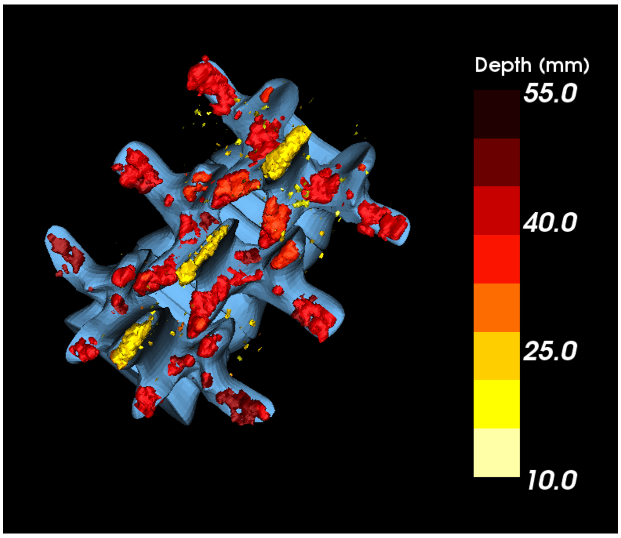

Figure 4.13: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a semi-random walk in the lateral-elevational plane, using combined ultrasound/camera-derived location for each sector-scan image.

Table 4.4: Root-Mean-Square Estimation Error for Ultrasound-based, Camera-based and Combined Estimation

Estimation	Bias (mm)	RMS Error (mm)	Norm. Corr.
Ultrasound (azimuth)	-0.0005	0.0115	0.9946
Ultrasound (elevation)	0.0350	0.0570	0.9511
Camera (azimuth)	-0.00515	0.0144	0.9877
Camera (elevation)	-0.0009	0.0209	0.9898
Combined (azimuth)	0.0023	0.0108	0.9943
Combined (elevation)	-0.0002	0.0194	0.9913

Table 4.5: Root-Mean-Square Bone Surface Error for Various Motion Estimators Over 3 Random Walks

Estimation/Index	1	2	3	Mean
True Motion	1.046	1.048	1.243	1.112
Ultrasound	1.137	1.187	1.394	1.239
Camera	1.058	1.054	1.252	1.121
Combined	1.051	1.048	1.254	1.118

4.E.ii Freehand 3D Spine Imaging Using Multi-Modal Motion Estimation

For three separate freehand scans over the tissue-mimicking spine phantom with overlaid porcine tissue, the RMS bone surface errors are shown in table 4.6, for ultrasound, camera and combined motion estimation along with the mean values over all three scans. The corresponding 3D renderings (for run 1) showing the ultrasound surface data superposed with the 3D spine CAD model, are shown in Figs. 4.14, 4.15 and 4.16 for ultrasound-only, camera-only and combined ultrasound/camera motion estimation algorithms. The ROC curves corresponding to each case are shown in Fig. 4.17, with ROC curve area, and sensitivity/specificity values at the optimal operating point shown in table 4.7. The rotation estimate in the 'roll' orientation is shown in Fig. 4.18, for run 1 of the freehand scans, calculated using ultrasound data, and also using the projection of the gravity vector on to the accelerometer y and z axes. The two measurements are not very well correlated, with an average correlation value of 0.635 over all 3 runs.

Table 4.6: Root-Mean-Square Bone Surface Error for Various Motion Estimators Over 3 Bidirectional Freehand Spine Scans

Estimation/Index	1	2	3	Mean
Ultrasound	2.384	2.023	2.265	2.224
Camera	2.212	2.171	2.253	2.212
Combined	2.205	2.077	2.222	2.168

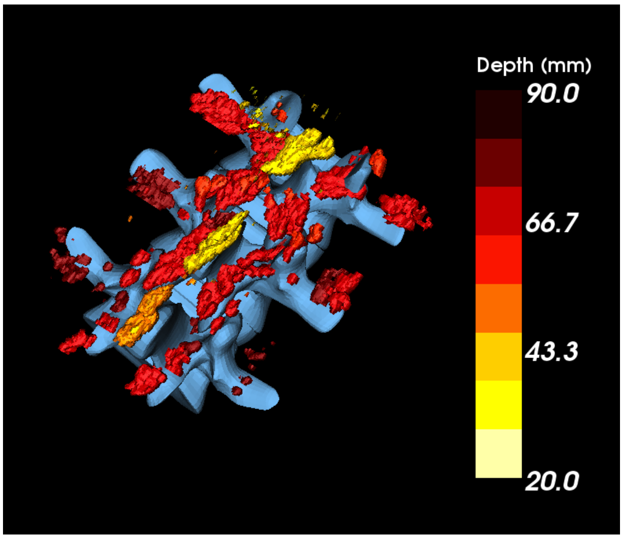

Figure 4.14: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a bidirectional scan of the spine using ultrasound-derived location for each sector-scan image.

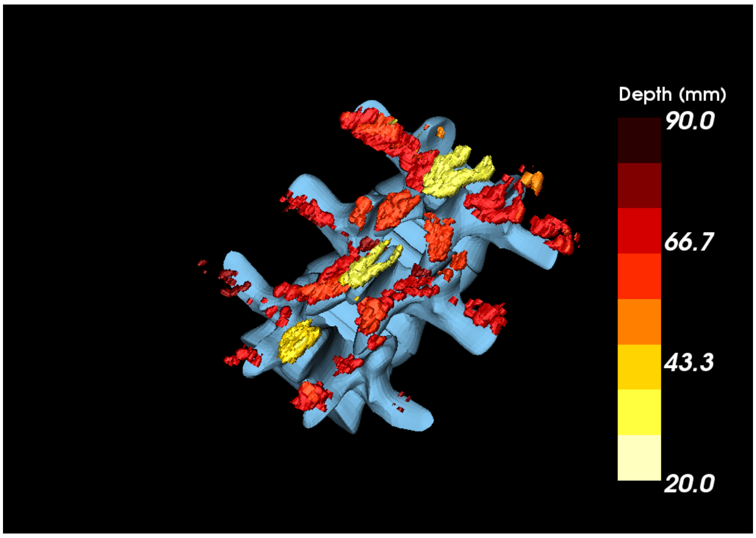

Figure 4.15: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a bidirectional scan of the spine using camera-derived location for each sector-scan image.

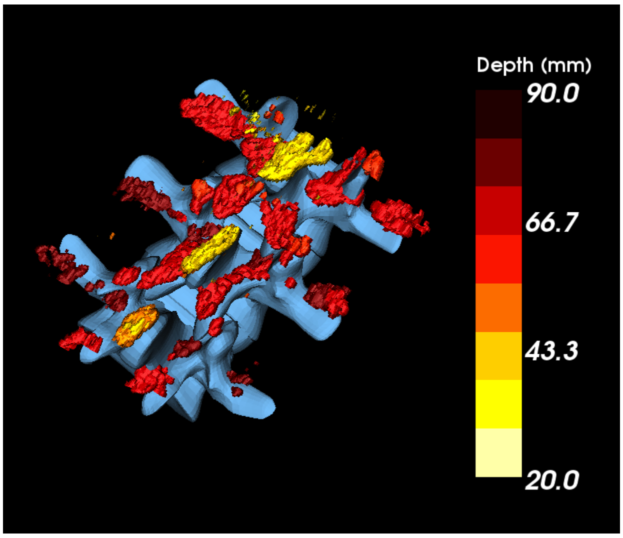

Figure 4.16: Superimposed visualization of 3D CAD spine model (blue-gray) and spinal bone surfaces (white-yellow-red) from sector-scan azimuthal plane images taken during a bidirectional scan of the spine using combined ultrasound/camera-derived location for each sector-scan image.

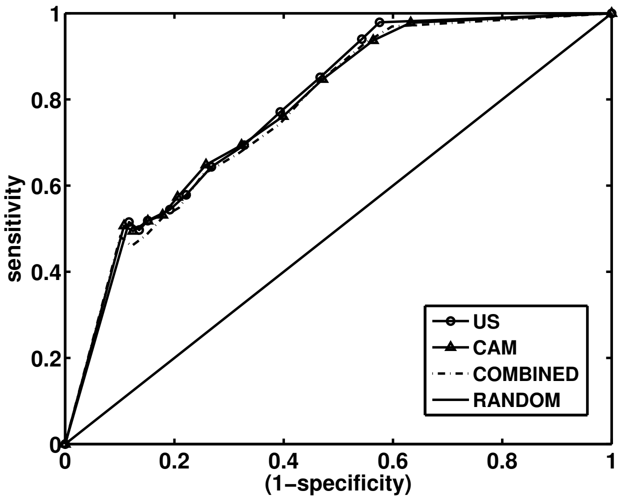

Figure 4.17: ROC Curves from freehand scan 1, using ultrasound-only, camera-only and combined motion estimation.

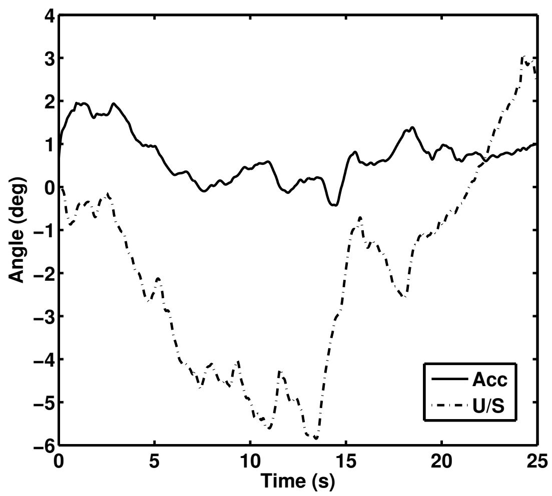

Figure 4.18: Rotational motion in the azimuthal plane ('roll') estimated using sector-scan ultrasound data and accelerometer data.

Table 4.7: ROC Curve Parameters for Freehand Scan 1

Estimation	ROC Area	Sensitivity (Opt.)	Specificity (Opt).
Ultrasound	0.7812	0.6436	0.7322
Camera	0.7793	0.6484	0.7421
Combined	0.7736	0.6256	0.7481

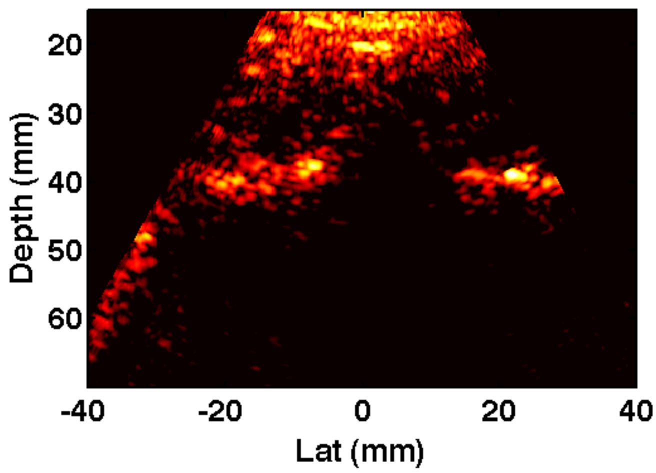

Figure 4.19: B-mode image of slice of human spine from freehand ultrasound scan, showing the spinous process and transverse process for a single vertebrae.

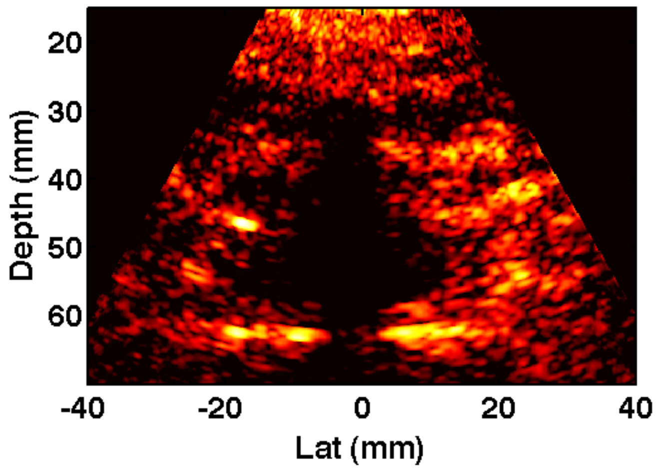

Figure 4.20: B-mode image of slice of human spine from freehand ultrasound scan, showing the interlaminar space and vertebral body for a single vertebra.

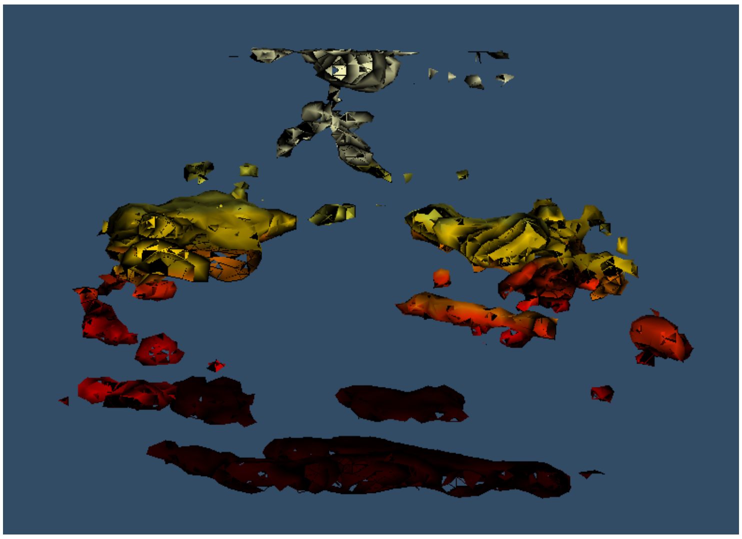

Figure 4.21: 3D surface rendering of human spine from freehand ultrasound scan, showing the spinous (white, top) and transverse processes (left, right, yellow), and vertebral body (dark, bottom) from a single vertebra.

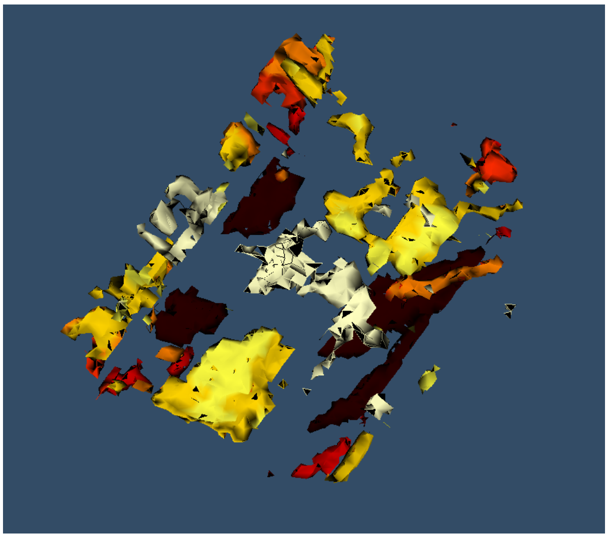

Figure 4.22: 3D surface rendering of human spine from freehand ultrasound scan, showing one vertebra with spinous (white) and transverse processes (yellow), and with clearly visible epidural gap between this and adjacent vertebra to the top-left. Spinal axis is in top-left to bottom-right direction in this rendering.

4.F Discussion

4.F.i Combined Ultrasound/Camera Motion Estimation

The results from experiments involving known translations but no rotational motion indicate that both ultrasound and camera-based translational motion estimation perform reasonably well in ideal conditions (table 4.4). As predicted by theory (82), ultrasound motion estimation in the elevational direction has an approximately 5 x larger RMS error, at 0.057 mm, compared to 0.0115 mm for azimuthal motion, averaged over all steps in 3 separate random walks. Camera azimuthal RMS estimation error is slightly worse than for ultrasound, at 0.0144 mm, but still highly accurate. In the elevational direction, surprisingly camera motion estimation RMS error was slightly worse, at 0.0209 mm. Possible reasons for this include anisotropic surface texture or lighting characteristics, or could simply be due to the different distribution of elevational and azimuthal random walk steps. Even though the errors here are all very small, it should be considered that for many small steps the errors add up rapidly, although it is possible to intentionally skip some frames to reduce error accumulation (103). All the measurements had very low bias, with the exception of ultrasound elevational estimates, with a bias of 0.035 mm. This can be attribute to a system PSF that does not match the ideal gaussian shape from (61), possibly exacerbated by the effects of an aberrating tissue layer on the PSF (96). In addition, electronic noise in the ultrasound system produces further decorrelation, which contributes to a positive bias in elevational estimates.

Using the optimal combination of multimodal motion estimation measurements detailed in section C.ii, combined motion estimates were in all cases better than either modality alone, with similarly improved normalized correlation with the true motion steps. In Fig. 4.8, the distributions of azimuthal and elevational estimate error are clearly unimodal for both camera and ultrasound modalities, with reduced distributoon width due to combination of measurements. This indicates that the distribution assumptions of equations 4.1-4.6 are justified. A qualitative visual confirmation of camera motion tracking accuracy, consistency and imaging quality is provided in the extended field-of-view images of the porcine tissue shown in Fig. 4.9.

The 3D surface renderings of Figs. 4.10-4.13 indicate good performance for all tracking methods, with RMS bone surface error in the 1.1 to 1.2 mm range (details in table 4.5). As the epidural gap is approximately 10 mm wide, these results indicate that without rotational motion, both ultrasound and camera motion tracking perform very well, with small improvements from combining measurements. The very close match between ultrasound surfaces and bone surface from the 3D spine model verifies that this registration method can be used to assess motion tracking even in the absence of a known set of motion steps.

4.F.ii Freehand 3D Spine Imaging Using Multi-Modal Motion Estimation

The freehand scans of Figs. 4.14 - 4.16 show a markedly worse match to the 3D spine model, with RMS bone surface error of approximately 2.2 mm on average. The surface distance error is exaggerated by the intentional use of an out-and-back scan motion, and in some 3D surface renderings a second ‘shadow’ image is seen for some parts of spinal bone anatomy. However, all motion estimates are still accurate enough for location of spinal features such as the epidural gap. The visibly inferior surface renderings are due to some combination of rotation introducing motion estimation errors, and rotation reducing bone/speckle contrast due to changing angle of incidence. From the summary of table 4.1, it is clear that using camera and ultrasound data, the direction of ‘pitch’ rotation cannot be determined. However, the accelerometer is capable of calculating both pitch and roll using the projection of the gravity vector onto the accelerometer axes and geometry. The accelerometer should make measurements of pitch and roll with equal accuracy, enabling a comparison of accelerometer ‘roll’ measurement to the high-quality roll measurement from the sector-scan ultrasound data, shown in Fig. 4.18. The two rotation measurements over time have similarities but only have a correlation of 0.635 when averaged over 3 runs. This indicates that the accelerometer used cannot easily be used to find ‘roll’ and ‘pitch‘ angles.

As the RMS surface distance errors for all motion estimators using the freehand out-and-back scans are quite similar, it is not surprising that the ROC curves for each estimator (Fig. 4.17 are similar along with ROC area and sensitivity/specificity values (4.7). However, the ROC values and images are quite reasonable. This indicates that the presence of rotation has a larger effect on bone surface localization and identification than differences between different motion estimation methods.

The human 2D spine images of figures 4.19 and 4.20 clearly show the spinous processes, transverse processes, vertebral body and interlaminar space, and appear to be superiour than ultrasound spine images from the literature. This is qualitative verification of the conclusions of (18) regarding the qualities of piston transducers for bone imaging. However, the system is designed to image spinal bone anatomy in high BMI patients (> 35), whereas the test subject has a BMI of only 27.1. Therefore, the focal depth of the transducer (5.6 cm) is ideal for imaging deeper spines, rather than the 1-1.5 cm deep spine of the test subject. This depth mismatch means that the resolution and contrast of the imaging system is degraded, due to imaging outside the focal zone, near-field effects and reverberation. As shallow spines are easily found using manual palpation, non-high BMI patients would not normally require ultrasound guidance for epidural procedures. However, different methods can be used to improve shallow imaging, including synthetic aperture methods (104), tissue harmonic imaging (105), and transducer optimizations to reduce reverb.

4.G Conclusion

In the absence of rotational motion, translational motion estimates from both ultrasound and camera data are highly accurate, and capable of producing 3D spinal bone surface renderings with RMS surface errors of around 1 mm. As the epidural gap, through which epidural procedures are performed, is approximately 10 mm in diameter, a handheld device with this level of bone surface localization capabilities will be easily capable of guiding such procedures. However, as the freehand scans demonstrated, uncompensated rotational motion degrades bone localization capabilities, but in realistic conditions (the freehand out-and-back scans), the system is still capable of adequate positional accuracy ($\sim$2.2 mm) for epidural procedure guidance. Clearly, in addition to the 'roll' rotation (which is already accounted for), use of existing ultrasound and camera data to estimate 'pitch' and 'yaw' rotations will greatly increase practical bone imaging performance. Theoretically, a 3-axis accelerometer is capable of providing rotational information that the camera and ultrasound system lack (direction of 'pitch' for example). However, accelerometer estimates of a well-known quantity ('roll') proved that the accelerometer used is not accurate enough to usefully estimate rotational parameters usefully. It may be possible to use this data crudely i.e. just to estimate sign of rotational variables, but it is more reasonable to investigate accelerometers with higher sensitivity and lower noise. Taken together, the motion tracking and 3D bone surface registration results indicate that it is feasible to perform epidural procedure guidance using the described device, however, the quality of rotational estimates has a significant effect on bone surface localization performance. The 3D images of figures 4.22 and 4.21, obtained without sigificant system optimization on a non-target patient, indicate the feasibility of the system for 3D freehand imaging of spinal bone anatomy.

Chapter 5

FINAL CONCLUSIONS AND FUTURE DIRECTIONS

5.A Conclusions

In this dissertation I have introduced several key technologies, enabling intuitive handheld 3D imaging, optimized for two specialized applications. In chapter 2, a separable factorization of the 2D beamforming process when applied to short-time-sequence, complex or approximated complex signals is shown to result in a 20 X increase in computational efficiency. Depending on the mode of operation, with a low beamforming duty cycle, low energy real-time operation can result, or with a high duty-cycle, very high frame rates can be achieved. I have additionally demonstrated that the separable factorization approach introduces negligible imaging degradation under all practical conditions. This work enables the smallest and lightest (170 g) real-time ultrasound system in the world today (to my knowledge), capable of 30 frames/second operation, with multi-hour battery life, optimized for vascular needle guidance.

In chapter 3, a specialized motion estimation algorithm is developed, that significantly improves estimation accuracy and robustness in sector-scan systems. This is an important application because previous work has shown that mechanically-scanned single-element transducers have particular benefits when imaging bone, specifically the intrinsic lack of grating lobes, and radially symmetric resolution increases bone contrast and visibility (18). I have developed a detailed, novel statistical model for complex point-spread-function correlation with translation. The model has been successfully used to gain insight into the statistics of normalized complex correlation, and to develop a maximum-likelihood sector-scan azimuthal motion estimator, and three different elevational motion estimators with varying degrees of optimality versus execution speed. The joint azimuthal-elevational motion estimator designed in this chapter has been proven to outperform conventional motion tracking for sector scans in typical conditions. In addition, the statistical model is further validated by correctly modeling and predicting bias in speckle-tracking motion estimation for sector-scan systems.

Finally, in chapter 4, the instrinsic directional ambiguity in some directions for ultrasound motion estimation using B-mode scan planes is addressed with the introduction of other sensor modalities for optimal combined motion estimation. Extensive experimental validation of the combined motion estimation is performed, indicating that the inclusion of a camera module can increase the robustness of motion estimation, particularly in the elevational direction. The peformance of the system in imaging 3D spinal bone anatomy is well-classified, using a realistic spine phantom and novel 3D registration methodology, along with receiver operating characteristic analysis. The experimental results indicate that in realistic conditions, even without accounting for all rotational degrees of freedom, the system is accurate enough to guide epidural anesthesia procedures. Initial *in-vivo* 2D and 3D images further indicate the feasibility of freehand 3D imaging of spinal bone anatomy.

Taken together, these contributions represent a significant contribution to the field of specialized, handheld intuitive 3D imaging systems.

5.B Future Directions

There are a number of obvious future directions for these research areas. For separable beamforming, varying the number of time samples and optimizing focusing weights taking into account the separable factorization is a simple way to further improved performance. Sector-scan motion tracking is very important in obstetric and cardiac applications, where a wide field of view is required through a small window into the body. The methods developed here could easily be applied to cardiac motion analysis. For robust 3D spinal bone anatomy imaging, the next steps involve exhaustively optimizing imaging and motion tracking parameters in animal and human models. Institutional review board approval is pending for a comparative spine imaging study relative to MRI and CT scans. Another area for improvement is in imaging well at shallow depths, whereas the current device is optimized for deep imaging in obese patients.

5.C Publications

All refereed publications related to this dissertation, or other medical ultrasound research during my academic studies at the University of Virginia are listed below.

K. Owen, M. I. Fuller, J. A. Hossack. Application of X-Y Separable 2D Array Beamforming for Increased Frame Rate and Energy Efficiency in Handheld Devices. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control. Accepted, to be published in July 2012 edition.

F. W. Mauldin, Jr., K. Owen, J. A. Hossack. The Effects of Medical Ultrasound Transducer Geometry on Image Artifacts Prevalent in Bone Imaging. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 59, pp. 1101-1114, 2012.

K. Owen, D. A. Guenther, W. F. Walker. Beamformer Enhancement by Post-Processing for Improved Spatial Resolution and Signal-to-Noise Ratio, Proceedings of SPIE, vol. 6920, 692002 (2008).

F. Viola, R. L. Coe, K. Owen, D. A. Guenther, W. F. Walker. Multi-Dimensional Spline-Based Estimator (MUSE) for Motion Estimation : Algorithm Development and Initial Results. Annals of Biomedical Engineering, vol. 36, no. 12, pp. 1942-1960, 2008.

REFERENCES

[1] T. Grau, E. Bartusseck, R. Conradi, E. Martin, and J. Motsch, "Ultrasound imaging improves learning curves in obstetric epidural anesthesia: a preliminary study." *Can J. Anesth*, vol. 50, no. 10, 2003. 1, 97, 99

[2] S. et al, "Artifacts and pitfall errors associated with ultrasound-guided regional anesthesia. Part I. understanding the basic principles of ultrasound physics and machine operations,," *Regional Anesthesia and Pain Medicine*, vol. 32, no. 5, pp. 412–418, 2007. 1

[3] T. R. Nelson and T. T. Elvins, "Visualization of 3D Ultrasound Data," *Computer Graphics and Applications*, vol. 13, pp. 50–57, 1993. 1, 41

[4] A. Fenster, D. B. Downey, and H. N. Cardinal, "Three-Dimension Ultrasound Imaging," *Phys. Med. Biol.*, vol. 46, pp. 67–99, 2001. 1, 41, 99

[5] B. Savord and R. Solomon, "Fully sampled matrix transducer for real time 3D ultrasonic imaging," in *Proceedings of IEEE Ultrasonics Symposium*, 2003, pp. 945–953. 2, 8, 35

[6] J. Hwang J, J. Quistgaard, J. Souquet, and L. A. Crum, "Portable Ultrasound Device for Battlefield Trauma," in *Proceedings of IEEE International Ultrasonics Symposium*, vol. 2, 1998, pp. 1663–1667. 2

[7] A. M. Chang, P. P. Chang, and S. R. Broadstone, "P-C Based Imaging System in a Probe," in *Proceedings of IEEE International Ultrasonics Symposium*, vol. 2, 2000, pp. 1255–1260. 2

[8] C. Abboud P-A and J. L. Kendall, “Ultrasound Guidance for Vascular Access,” *Emerg. Med. Clin. N. Am*, vol. 22, pp. 749–773, 2004. 3

[9] T. G. Constantino, A. K. Parikh, W. A. Satz, and J. P. Fojtik, “Ultrasonography-Guided Peripheral Intravenous Access Versus Traditional Approaches in Patients With Difficult Intravenous Access,” *Annals of Emergency Medicine*, vol. 46, no. 5, pp. 456–461, 2005. 3

[10] C. Arzola, S. Davies, A. Rofaeel, and J. C. A. Carvalho, “Ultrasound using the transverse approach to the lumbar spine provides reliable landmarks for labor epidurals,” *Anesthesia and Analgesia*, vol. 104, pp. 1188–1192, 2007. 3, 98

[11] M. Balki, Y. Lee, S. Halpern, and J. C. A. Carvalho, “Ultrasound Imaging of the Lumbar Spine in the Transverse Plane: The Correlation Between Estimated and Actual Depth to the Epidural Space in Obese Parturients,” *Anesthesia and Analgesia*, vol. 108, pp. 1876–1881, 2009. 3, 97, 98

[12] D. D. Hood and D. M. Dewan, “Anesthetic and obstetric outcome in morbidly obese parturients,” *Anesthesiology*, vol. 79, pp. 1210–1218, 1993. 3, 98

[13] K. J. Chin, A. Perlas, V. Chan, D. Brown-Shreves, A. Koshkin, and V. Vaishnav, “Ultrasound Imaging Facilitates Spinal Anesthesia in Adults with Difficult Surface Anatomic Landmarks,” *Anesthesiology*, vol. 115, pp. 94–101, 2011. 3, 98

[14] D. J. Brenner and E. J. Hall, “Current concepts - Computed tomography - An increasing source of radiation exposure,” *New England Journal of Medicine*, vol. 357, pp. 2277–2284, 2007. 3, 99

[15] A. C. Society, “Cancer Facts & Figures 2010,” American Cancer Society. 3, 99

[16] K. Ranganathan, M. K. Santy, T. N. Blalock, J. A. Hossack, and W. F. Walker, “Direct sampled I/Q beamforming for compact and very low-cost

ultrasound Imaging," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 51, pp. 1082–1094, 2004. 4, 9, 18

[17] M. I. Fuller, K. Owen, T. N. Blalock, J. A. Hossack, and W. F. Walker, "Real Time Imaging with the Sonic Window: A Pocket-Sized, C-Scan, Medical Ultrasound Device," in *Proceedings of IEEE Ultrasonics Symposium*, vol. 1, 2009, pp. 196–199. 4, 9, 20, 36

[18] F. W. J. Mauldin, K. Owen, M. Tiouririne, and J. A. Hossack, "The effects of transducer geometry on artifacts common to diagnostic bone imaging with conventional medical ultrasound," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 59, pp. 1101–1114, 2012. 4, 40, 42, 94, 97, 100, 103, 133, 137

[19] K. E. Thomenius, "Evolution of Beamformers," in *Proceedings of 1996 IEEE Ultrasonics Symposium*, 1996, pp. 1615–1622. 7

[20] B. D. Van Veen and K. M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, vol. 5, pp. 4–24, 1988. 8, 11

[21] G. R. Lockwood, J. R. Talman, and S. S. Brunke, "Real-Time 3-D Ultrasound Imaging Using Sparse Synthetic Aperture Beamforming," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 45, pp. 980–988, 1998. 8

[22] J. T. Yen, J. P. Steinberg, and S. W. Smith, "Sparse 2-D Array Design for Real-Time Rectilinear Volumetric Imaging," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 47, pp. 93–110, 2000. 8

[23] K. S. Kim and T. K. Song, "High Volume Rate 3-D Ultrasound Imaging Using Cross Array Based on Synthetic Transmit Focusing," in *Proceedings of IEEE Ultrasonics Symposium*, 2004, pp. 1409–1412. 8

[24] K. Kim and T. Song, "Method and Apparatus for Forming Ultrasonic Three-Dimensional Ultrasound Images Using Cross Arrays," U.S. Patent 5 901 708, 1999. 8

[25] L. F. Nock and G. E. Trahey, "Synthetic Receive Aperture Imaging with Phase Correction for Motion and for Tissue Inhomogeneities," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 39, pp. 489–195, 1992. 8

[26] M. Karaman, P. C. Li, and M. O'Donnell, "Synthetic Aperture Imaging for Small Scale Systems," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 42, pp. 429–442, 1995. 8

[27] N. M. Daher and J. T. Yen, "2-D Array for 3-D Ultrasound Imaging Using Synthetic Aperture Techniques," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 53, pp. 912–924, 2006. 8

[28] M. Karaman, I. O. Wygant, O. Oralkan, and B. T. Khuri-Yakub, "Minimally Redundant 2-D Array Designs for 3-D Medical Ultrasound Imaging," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 28, pp. 1051–1061, 2009. 8

[29] R. Y. Chiao and L. Thomas, "Aperture Formation on Reduced Channel Arrays Using the Transmit-Receive Apodization Matrix," in *Proceedings of IEEE Ultrasonics Symposium*, 1996, pp. 1567–1571. 8

[30] J. Johnson, M. Karaman, and B. Khuri-Yakub, "Coherent-Array Imaging Using Phased Subarrays. Part 1: Basic Principles," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 52, pp. 37–50, 2005. 8, 35

[31] M. Karaman and M. O'Donnell, "Subaperture Processing for Ultrasonic Images," *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 45, pp. 126–135, 1998. 8, 35

[32] B. J. Savord, "Beamforming Methods and Apparatus for Three-Dimensional Ultrasound Imaging Using Two-Dimensional Transducer Array," U.S. Patent 6 013 032, 2000. 8, 35

[33] J. Petrofsky, S. H. Maslak, and C. R. Cole, "Ultrasonic Receive Beamformer with Phased Subarrays," U.S. Patent 5 573 001, 1996. 8, 35

[34] J. G. Petrofsky, S. H. Maslak, and C. R. Cole, “Ultrasonic Receive Beamformer With Phased Subarrays,” U.S. Patent 5 676 147, 1997. 8, 35

[35] S. W. Smith, H. G. Pavy, and O. T. von Ramm, “High Speed Ultrasound Volumetric Imaging System - Part 1: Transducer Design and Beam Steering,” *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 38, pp. 100–108, 1991. 8

[36] O. T. von Ramm, S. W. Smith, and H. G. Pavy, “High Speed Ultrasound Volumetric Imaging System - Part II: Parallel Processing and Image Display,” *I*, vol. 38, pp. 109–115, 1991. 8

[37] J.-y. Lu, “2D and 3D High Frame Rate Imaging With Limited Diffraction Beams,” *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 44, pp. 839–856, 1997. 8

[38] J.-Y. Lu, “Experimental Study of High Frame Rate Imaging With Limited Diffraction Beams,” *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 45, pp. 84–97, 1998. 8

[39] J. Cheng and J.-y. Lu, “Extended High Frame Rate Imaging Method with Limited-Diffraction Beams,” *IEEE Transactions on Ultrasonics, Ferrorelectrics & Frequency Control*, vol. 53, pp. 880–899, 2006. 8

[40] S. Stergiopoulos, Ed., *Advanced Signal Processing Handbook: Theory and Implementation for Radar*. Boca Raton, FL: CRC, 2000. 8

[41] B. Maranda, “Efficient Digital Beamforming In the Frequency Domain,” *J. Acoust. Soc. Am.*, vol. 86, pp. 1813–1819, 1989. 8, 37

[42] A. Tawfik, “A Generic Processing Structure Decomposing the Beamforming Process of 2D and 3-D Arrays of Sensors into Sub-Sets of Coherent Process,” in *Proceedings of the 16th National Radio Science Conference, NRSC*, 1999. 8

[43] V. Murino and A. Trucco, “Underwater 3D Imaging By FFT Dynamic Focusing Beamforming,” in *IEEE International Workshop on Imaging Systems and Techniques*, 2007. 8

[44] M. Palmese and A. Trucco, "Digital Near Field Beamforming for Efficient 3-D Underwater Acoustic Image Generation," in *IEEE International Workshop on Imaging Systems and Techniques*, 2007. 8, 35

[45] F. Zhang, A. Bilas, A. Dhanantwari, K. N. Plataniotis, R. Abiprojo, and S. Stergiopoulos, "Parallelization and Performance of 3D Ultrasound Imaging Beamforming Algorithms on Modern Clusters," in *Proceedings of the 16th International Conference on Supercomputing*, 2002. 8

[46] A. C. Dhanantwari, S. Stergiopoulos, L. Song, C. Parodi, F. Bertora, P. Pellegretti, and A. Questa, "An Efficient 3-D Beamformer Implementation for Real-Time 4-D Ultrasound Systems Deploying Planar Probes," in *Proceedings of IEEE Ultrasonics Symposium*, 2004, pp. 1421–1424. 8

[47] K. Wall and G. R. Lockwood, "Development of a Versatile Signal Processing Board for Real-Time 3-D Beamforming," in *Proceedings of IEEE Ultrasonics Symposium*, 2007, pp. 1526–1528. 9

[48] T. L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*. Burlington , MA: Elsevier Academic Press, 2004. 14, 34

[49] J. W. Goodman, *Fourier Optics*. Englewood, CO: Roberts & Co, 2005. 14, 34

[50] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Mathematics of Computation*, vol. 19, pp. 297–301, 1965. 16

[51] P. D. Welch, "A Fixed-Point Fast Fourier Transform Error Analysis," *IEEE Transactions on Audio and Electroacoustics*, vol. 17, no. 2, pp. 151–157, 1969. 16

[52] W. R. Knight and R. Kaiser, "A Simple Fixed-Point Error Bound for the Fast Fourier Transform," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, no. 6, pp. 615–620, 1979. 16

[53] M. Eames and J. A. Hossack, “Fabrication and Evaluation of Fully-Sampled, Two Dimensional Transducer Array for ’Sonic Window’ Imaging System,” *Ultrasonics*, vol. 48, pp. 376–383, 2008. 17, 20, 34

[54] M. Unser, “Splines A Perfect Fit for Signal and Image Processing,” *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, 1999. 17

[55] M. Unser, A. Aldroubi, and M. Elden, “Fast B-Spline Transforms for Continuous Image Representation and Interpolation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 277–285, 1991. 17

[56] P. Thevenaz, T. Blu, and M. Unser, “Interpolation Revisited,” *IEEE Transactions on Medical Imaging*, vol. 19, pp. 739–758, 1991. 17

[57] J. A. Jensen and N. B. Svendsen, “Calculation of Pressure Fields from Arbitrarily Shaped, Apodized and Excited Ultrasound Transducers,” *IEEE Transactions on Ultrasonics Ferroelectrics & Frequency Control*, vol. 39, pp. 262–267, 1992. 19, 61, 64, 90

[58] Y. Li, T. N. Blalock, W. F. Walker, and J. A. Hossack, “Phased Subarrays for Low-Cost C-Scan Applications,” in *Proceedings of IEEE Ultrasonics Symposium*, 2003, pp. 1408–1411. 35

[59] A. Gee, R. Prager, G. Treece, and L. Berman, “Engineering a Freehand 3D Ultrasound System,” *Pattern Recognition Letters*, vol. 24, pp. 757–777, 2003. 39, 41, 97, 99

[60] G. Trahey, J. W. Allison, and O. T. Von Ramm, “Angle Independent Ultrasonic Detection of Blood Flow,” *IEEE Transactions on Biomedical Engineering*, vol. BME-34, pp. 965–967, 1987. 39, 41, 57, 66, 86, 99, 103, 114

[61] F. Chen J, J. B. Foulkes, P. L. Carson, and J. M. Rubin, “Determination of Scan-Plane Motion Using Speckle Decorrelation : Theoretical Considerations and Initial Test,” *Intl. Journal of Imaging Systems and Technology*, vol. 8, pp. 38–44, 1997. 39, 42, 66, 67, 86, 99, 103, 131

[62] J. Meunier and M. Bertrand, "Ultrasonic Texture Motion Analysis : Theory and Simulation," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 14, pp. 293–300, 1995. 39, 41, 42, 57, 58, 87, 89

[63] J. A. Hossack, T. S. Sumanaweera, S. Napel, and J. S. Ha, "Quantitative 3D Diagnostic Ultrasound Imaging Using a Modified Transducer Array and an Automated Tracking Technique," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 49, pp. 1029–1038, 2002. 41, 99

[64] A. M. Goldsmith and T. L. Szabo, "An Inertial-Optical Tracking System for Portable, Quantitative, 3D Ultrasound," *Proc. IEEE Ultrasonics Symposium*, pp. 45–49, 2002. 41, 99

[65] D. F. Leotta, P. R. Detmer, O. H. Gilja, and E. Al, "Three-Dimensional Ultrasound Imaging using Multiple Magnetic Tracking Systems and Miniature Magnetic Sensors," in *Proceedings of IEEE International Ultrasonics Symposium*, vol. 2, 1995, pp. 1415–1418. 41, 99

[66] B. H. Friemel, L. N. Bohs, and G. E. Trahey, "Relative performance of two-dimensional speckle-tracking techniques: normalized correlation, non-normalized correlation and sum-absolute-difference," in *Proceedings of IEEE International Ultrasonics Symposium*, vol. 2, 1995, pp. 1481–1484. 41, 66, 86, 114

[67] F. Viola and W. F. Walker, "A Comparison of the Performance of Time-Delay Estimators in Medical Ultrasound," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 50, pp. 392–401, 2003. 41, 66

[68] I. Cespedes, J. Ophir, and S. K. Alam, "The combined effect of Signal Decorrelation and Random Noise on the Variance of Time Delay Estimation," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 44, pp. 220–225, 1997. 41, 58, 86, 89

[69] A. Gee, R. J. Housden, P. Hassenpflug, G. Treece, and R. W. Prager, "Sensorless Freehand 3D Ultrasound in Real Tissue: Speckle Decorrelation without Fully Developed Speckle," *Medical Image Analysis*, vol. 10, pp. 137–149, 2006. 42, 99, 114

[70] R. W. Prager, A. H. Gee, G. M. Treece, C. J. C. Cash, and L. H. Berman, "Sensorless Freehand 3D Ultrasound Using Regression of the Intensity," *Ultrasound in Medicine and Biology.*, vol. 29, pp. 437–446, 2003. 42, 49, 86

[71] C. Laporte and T. Arbel, "Probabilistic Speckle Decorrelation for 3D Ultrasound," *Medical Image Computing and Computer-Assisted Intervention*, vol. 4791, pp. 925–932, 2007. 42, 49, 86

[72] E. L. Madsen, J. A. Zagzeg, R. A. Banjavie, and R. E. Jutila, "Tissue Mimicking Materials for Ultrasound Phantoms," *Medical Physics*, vol. 5, pp. 391–394, 1978. 43, 63, 65, 67, 113

[73] W. Feller, "The Fundamental Limit Theorems in Probability," *Bull. Amer. Math Soc.*, vol. 51, pp. 800–832, 1945. 49, 88, 105

[74] R. Baker, "Probability Estimation and Information Principles," *Struct. Saf.*, vol. 9, pp. 97–116, 1990. 49

[75] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1:Estimation Theory*. Prentice-Hall, 1993. 53, 105

[76] R. I. Jennrich and P. F. Sampson, "Newton-Raphson and Related Algorithms for Maximum-Likelihood Variance Component Estimation," *Technometrics*, vol. 18, pp. 11–17, 1976. 55

[77] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, vol. 39, pp. 1–38, 1977. 55

[78] I. Cespedes, Y. Huang, and O. S. Spratt, "Methods for Estimation of Subsample Time Delays of Digitized Echo Signals," *Ultrasonic Imaging*, vol. 17, pp. 142–171, 2002. 57, 66

[79] P. Berens, "Circstat: A matlab toolbox for circular statistics," *Journal of Statistical Software*, vol. 31, no. 10, pp. 1–21, 9 2009. [Online]. Available: http://www.jstatsoft.org/v31/i10 63

[80] J. Zar, *Biostatistical Analysis*, 4th ed. Prentice Hall, 1999. 63

[81] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of Speckle in Ultrasound B-Scans," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 30, pp. 156–163, 1983. 86, 88, 106

[82] B. S. Ramamurthy and G. E. Trahey, "Potential and Limitations of Angle-Independent Flow Detection Algorithms using Radio-Frequency and Detected Signals," *Ultrasonic Imaging*, vol. 13, pp. 252–268, 1991. 86, 99, 103, 130

[83] J. G. Abbott and F. L. Thurstone, "Acoustic Speckle : Theory and Experimental Analysis," *Ultrasonic Imaging*, vol. 1, pp. 303–324, 1979. 86

[84] W. F. Walker and G. E. Trahey, "A Fundamental Limit on Delay Estimation Using Partially Correlated Speckle Signals," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 42, pp. 301–308, 1995. 86

[85] Y. Wang, M. A. Beydoun, L. Liang, B. Caballero, and K. Kumanyika, "Will All Americans Become Overweight or Obese? Estimating the Progression and Cost of the US Obesity Epidemic," *Obesity*, vol. 16, pp. 2323–2330, 2008. 98

[86] O. V. Solberg, F. Lindseth, H. Torp, R. E. Blake, and T. A. N. Hernes, "FREEHAND 3D ULTRASOUND RECONSTRUCTION ALGORITHMS A REVIEW," *Ultrasonic Imaging*, vol. 33, pp. 991–1009, 2007. 99

[87] L. Mercier, T. Lango, F. Lindseth, and L. D. Collins, "A REVIEW OF CALIBRATION TECHNIQUES FOR FREEHAND 3-D ULTRASOUND SYSTEMS," *Ultrasound in Medicine and Biology.*, vol. 31, pp. 143–165, 2005. 99

[88] C. D. Garson, Y. Li, and J. A. Hossack, "Quantitative 3D Diagnostic Ultrasound Imaging Using a Modified Transducer Array and an Automated Tracking Technique," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 54, pp. 966–977, 2007. 99

[89] K. Owen, F. W. Mauldin, and J. A. Hossack, "Transducer Motion Estimation Using Combined Ultrasound Signal Decorrelation and Optical Sensor Data for Low-cost Ultrasound Systems with Increased Field of View," in *Proceedings of IEEE International Ultrasonics Symposium*, 2011. 100

[90] L. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24, pp. 325–376, 1992. 101, 114

[91] B. Zitova and J. Flusser, "Image Registration Methods : A Survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003. 101, 114

[92] N. Yazdi, F. Ayazi, and K. Najafi, "Micromachined Inertial Sensors," *Proceedings of the IEEE*, vol. 86, pp. 249–265, 1998. 103

[93] L. Verlet, "Computer 'Experiments' on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules," *Physical Review*, vol. 159, pp. 98–103, 1967. 103

[94] G. Pang and H. Liu, "Evaluation of a Low-cost MEMS Accelerometer for Distance Measurement," *Journal of Intelligent and Robotic Systems*, vol. 30, pp. 249–265, 2001. 104

[95] P. Hoskins, A. Thrush, K. Martin, and T. Whittingham, *Diagnostic ultrasound: physics and equipment*. Cambridge University Press, 2003. 106

[96] G. E. Trahey and S. W. Smith, "Properties of acoustical speckle in the presence of phase aberration part I: First order statistics," *Ultrasonic Imaging*, vol. 10, pp. 12–28, 1988. 106, 113, 114, 131

[97] G. E. Trahey, "Phase Aberration Measurement and Correlation in Medical Ultrasound," in *Proceedings of IEEE Eng. Med. Biol. Society International Conference*, 1989. 106, 113

[98] T. Fawcett, “An Introduction to ROC Analysis,” *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006. 107

[99] K. P. Adlassnig and W. Scheithauer, “Performance Evaluation of Medical Expert Systems Using ROC Curves,” *Computers and Biomedical Research*, vol. 22, pp. 297–313, 1989. 107

[100] J. A. Swets, “ROC analysis applied to the evaluation of medical imaging techniques,” *Investigative Radiology*, vol. 14, pp. 109–121, 1979. 107

[101] “ADXL345 data sheet,” Analog Devices, Norwood, MA. 113

[102] Food and D. Administration. (2008) Guidance for industry and fda staff - information for manufacturers seeking marketing clearance of diagnostic ultrasound systems and transducers. [Online]. Available: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm070856.htm 113

[103] J. S. Ha, W. F. Walker, and J. A. Hossack, “Determination of an Optimal Image Frame Interval for Frame-to-Frame Ultrasound Image Motion Tracking,” *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 52, pp. 386–396, 2005. 131

[104] M.-L. Li, W.-J. Guan, and P.-C. Li, “Improved Synthetic Aperture Focusing Technique with Applications in High-Frequency Ultrasound Imaging,” *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 51, pp. 63–70, 2004. 133

[105] F. Tranquart, N. Gernier, V. Eder, and L. Pourcelot, “Clinical Use of Ultrasound Tissue Harmonic Imaging,” *Ultrasound in Medicine and Biology*, vol. 25, pp. 889–894, 1999. 133