

Collagen Alignment Imaging and Analysis for Breast Cancer Classification

By

Jeremy S. Bredfeldt

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Medical Physics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2014

Date of final oral examination: 1/10/2014

The dissertation is approved by the following members of the Final Oral Committee:

Thomas R. Mackie, Professor Emeritus, Medical Physics

Kevin W. Eliceiri, Associate Scientist, Medical Physics

Edward F. Jackson, Professor, Medical Physics

Patricia J. Keely, Professor, Regenerative Biology

Paul J. Campagnola, Professor, Medical Physics

Sean B. Fain, Associate Professor, Medical Physics

UMI Number: 3609840

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3609840

Published by ProQuest LLC (2014). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Acknowledgements

I am sincerely grateful to all those who have made possible the writing of this dissertation.

To Kevin Eliceiri for his incredible enthusiasm and for including me in the LOCI family.

To Rock Mackie for his amazing optimism, creativity, and support.

To the members of the Keely, LOCI and Medical Engineering groups for their insightful assistance and valuable discussions.

To my parents, Thomas and Susan Bredfeldt, and Mark and Pamela Seppelt, for encouraging me to follow my passions and for their tremendous love and support throughout the past four years.

To my children Oliver, Hazel and Rose whose enthusiasm, creativity, passion, sense of humor and resilience inspire and re-energize me every day.

To my wife, Bobbie Bredfeldt, for her unbelievable love, patience, and humility throughout this adventure.

This work was financially supported by NIH R01 Grants CA114462 and CA136590, the UW Radiological Sciences Training Grant T32 CA009206, and the Morgridge Institute for Research.

Table of Contents

List of Tables	v
List of Figures	vii
Abstract.....	xiv
Chapter 1 Introduction	1
1.1 Significance.....	2
1.2 Optical Biopsy.....	2
1.3 Intrinsic Contrast Imaging with Multiphoton Microscopy	4
1.4 Collagen's Role in Cancer Progression	5
1.5 Collagen alignment imaging	6
1.6 SHG Imaging Microscopy	7
1.7 3D SHG Imaging.....	8
1.8 Automated histopathology imaging and analysis	10
1.9 Collagen alignment analysis.....	11
1.10 Patient classification.....	12
1.11 Innovation.....	12
Chapter 2 Curvelet Image Analysis	14
2.1 Introduction	14
2.2 Materials and Methods.....	20
2.2.1 The Curvelet Transform	20
2.2.2 CurveAlign	21
2.2.3 Validation Analysis	24
2.2.4 SHG Microcopy.....	25
2.2.5 Mouse Mammary Tissues and Tumor Explants	25
2.3 Results and Discussion.....	26
2.4 Conclusion.....	36
Chapter 3 Collagen Fiber Extraction	38
3.1 Introduction	38
3.2 Materials and Methods.....	42
3.2.1 Sample Preparation.....	43
3.2.2 SHG Imaging	44
3.2.3 FIRE Algorithm.....	44
3.2.4 Preprocessing Algorithms.....	45
3.2.5 Algorithm integration and evaluation.....	47

3.3 Results	49
3.4 Discussion	54
3.5 Conclusion.....	58
Chapter 4 Electronic Tumor Association Collagen Signatures	60
4.1 Background	60
4.2 Methods	62
4.2.1 Human breast carcinoma tissue microarray.....	64
4.2.2 Imaging System	65
4.2.3 TACS Model.....	68
4.2.4 Classification and Survival Analysis	75
4.3 Results	75
4.4 Discussion and Conclusions.....	82
Chapter 5 Three Dimensional SHG Imaging.....	87
5.1 Introduction	87
5.2 Materials and methods	89
5.2.1 Collagen Gel Preparation.....	89
5.2.2 SHG Imaging System	90
5.2.3 Image Processing.....	91
5.3 Results	91
5.3.1 Elevation angle sensitivity of backward-propagating SHG	92
5.3.2 Elevation angle sensitivity of forward-propagating SHG	93
5.3.3 NA dependence on axial detection with forward-propagating SHG	95
5.4 Discussion	96
5.5 Conclusion.....	99
Chapter 6 Sequential Erosive Tissue Imaging	100
6.1 Introduction	100
6.2 Materials and Methods	103
6.2.1 Sample Preparation.....	103
6.2.2 Imaging and Erosion Instrumentation	104
6.2.3 Image Reconstruction	105
6.3 Results and Discussion.....	109
Chapter 7 Overall Conclusions and Future Outlook.....	113
Appendix 1 Hardware Bill of Materials.....	115
Appendix 2 SETI Protocols	118

Bibliography	121
--------------------	-----

List of Tables

Table 1. Descriptions of tumor-related collagen I structures[91].....	16
Table 2. Results of Kolmogorov-Smirnov testing between manual measurements and CurveAlign measurements on six images of mouse mammary tumors. This table presents the mean, standard deviation and Kolmogorov-Smirnov test results between three images of TACS-2(A-C) mouse mammary tumors and three images of TACS-3(A-C) mouse mammary tumors. The result $h = 0$ indicates that the measured and calculated angle distributions were similar based on a 95% confidence level.	28
Table 3. Results validating the CurveAlign measurements of fiber angles in 5 simulated test images. M = mean angle in degrees, SD = standard deviation of the distribution of angles in degrees, KS = the result of the Kolmogorov-Smirnov statistical test validating the null hypothesis that the ground truth and measured distributions are statistically similar at the 95% confidence level.	29
Table 4. Features used in epithelial cell segmentation and TACS-3 fiber classification tasks.	71
Table 5. Correlation between manual and the eTACS approach. There is statistically significant positive correlation between the eTACS method and all manual scoring approaches (N=196).	78
Table 6. Univariate Cox Proportional Hazard Analysis results for various eTACS feature combinations. This table shows that the eTACS scoring method that includes both fiber and cellular information produces the more significantly prognostic scores compared to fiber information alone (N=196).	78
Table 7. Feature ranking based on SVM feature weight w for a 16 patient (8 TACS-3 positive and 8 TACS- 3 negative) training set including the average feature value difference between the positive and negative training cases $fp - fn$. These values show the trend of the feature values between TACS-3 positive (N=112) and TACS-3 negative (N=84) images.	81

List of Figures

Figure 1. SHG image of collagen fibers in fresh, unstained mouse mammary tissue.....	7
Figure 2. Standard optical layout of an epi-detecting SHG microscope. Images are captured in the en-face plane by scanning the focus of the beam in two dimensions. 3D imaging is performed by translating the objective or sample in the third dimension. Red is the excitation light and green is the emission light.	8
Figure 3. Geometry of 3D fiber imaging with an SHG microscope. The incident light is polarized along the x-axis and propagates in the negative z-direction (dotted line). Example fibers are oriented in the x-y plane (imaging plane) and in the z-direction.	10
Figure 4. Largest 0.4% of the curvelets at the second finest scale produced by the Fast Discrete Curvelet Transform overlaid in green on a synthetic test image. Green lines indicate the orientation of each curvelet while the green circles indicate their center position.....	27
Figure 5. CurveAlign results from synthetic images of crisscrossing fibers (A-C) and parallel fibers (D-F). A,D: Individual curvelets overlaid on the original images. B,E: alignment maps overlaid on the original images where green indicates low probability of alignment, yellow intermediate alignment, and red high alignment. C,F: Histograms of the detected curvelet angles within each image. Although the histograms of these images are quite similar, the alignment maps show areas of localized alignment in only the parallel fiber image.	30
Figure 6. CurveAlign output examples. (A) 3D collagen gel image with boundary (yellow), analyzed curvelets (green) and discarded curvelets (red), overlaid on the original image. (B) Resulting histogram of fiber angles with respect to the boundary. (C inset) A zoomed version of A showing the boundary (yellow), curvelets (green), and the associations between the curvelets and the boundary (blue). (D) The alignment map overlaid on the original figure showing regions within	

the image where there are high (red), medium (yellow) and low (green) concentrations of perpendicularly aligned fibers 31

Figure 7. Influence of curvelet coefficient threshold on measurements. Curvelets overlaid on original image representing 0.1% (A), 0.25% (B), and 1.0% (C) of the top curvelet coefficients with corresponding angle distributions (D-E respectively). The coefficient of alignment was 0.972, 0.976, and 0.973 for case A-C respectively 32

Figure 8. Collagen alignment measurements made in thick excised human breast cancer tissue samples. Fiber angle distributions relative to the epithelial stromal boundary for an adjacent normal terminal lobular end-bud unit are shown in (A,C) while the collagen alignment relative to the invasion front of a group of invasive ductal carcinoma cells are shown in (B,D). In the invasive case, the angle distribution is weighted more heavily toward perpendicular angles. E-Cadherin stained epithelial cells are shown in red while collagen fibers are shown in green. Scale bar = 50 μ m 34

Figure 9. Representative collagen patterns observed in human breast cancer tissue sections demonstrating the heterogeneous nature of collagen structure. Wavy (A) and straight (B). High (C) and low density (D). Thick bundles (E) and thin strands (F). Discontinuous (G) and continuous (H). Crossing (I) and parallel (J). Scale bar is 10 microns 39

Figure 10. Fibers extracted by the FIRE algorithm alone without preprocessing. A and D are the original images, B and E show manual segmentations of the fibers, C and F show the automatic fiber segmentations that are extracted by the FIRE algorithm and show many falsely segmented fibers. Scale bar is 25 microns 41

Figure 11. Diagram of the approach for quantitative collagen analysis showing the iterative process for optimizing the performance of a single image-processing filter for fiber tracking. The raw image is processed by the image filter using an initial normalization parameter, the result of which is

sent to the FIRE fiber-tracking algorithm. Automated fiber extractions are compared against manually performed fiber extractions. Several normalization parameters are evaluated and one optimal parameter is selected for each filter based on the fiber evaluation result.....	43
Figure 12. Output of the image processing techniques (row 1) and output of the fiber tracking algorithm (row 2) for a single test case. The first column is without a filter, column 2: GF, column 3: SPTV filter, column 4: TF, and column 5: CT. Scale bar is 25 microns.	50
Figure 13. Two test cases (A&B), showing different processing methods in each column. The original image (column 1) is shown overlaid with a manual segmentation (column 2), GF (column 3), SPTV (column 4), TF (column 5), and CT filter (column 6) where each filter is followed by FIRE fiber extraction. Scale bar is 25 microns.....	51
Figure 14. F-measure, recall, and precision results comparing the automated segmentation techniques to the manual segmentations of three independent raters, for 25 test cases, representing a total of 9290 fiber evaluations. The error bars indicate the standard deviation between average F-measure, recall and precision scores of each of the raters. Recall is the fraction of relevant fibers that were found. Precision is the fraction of fibers found that were relevant. F-measure is the harmonic sum of recall and precision.....	52
Figure 15. Distribution of angles (top row) and lengths (bottom row) of all fibers in all simulated test cases. Ground truth data is on the left and the results of the automated CT+FIRE algorithm are shown on the right.....	53
Figure 16. Demonstration of automated fiber segmentation feature extraction in an <i>in-vivo</i> mouse model for breast cancer. A mammary window was placed immediately superficial to a palpable mammary tumor and the collagen microenvironment was imaged 8 and 12 weeks of age. Automated fiber extractions are shown overlaid on representative images from the 8 (A) and 12 (B) week time points. The bar graph (C) shows the ratio of the number of wavy fibers to total	

fibers found in the image. Fibers are labeled wavy if the distance along the fiber divided by the distance between fiber endpoints is greater than 1.08. Error bars indicate one standard deviation of the computed average wavy-fractions among the 9 images analyzed for each time point. Scale bar is 25 microns. 54

Figure 17. Block diagram of the eTACS system. The red steps are performed with WiscScan, the blue steps are performed with ImageJ/FIJI, and the green steps are performed with MATLAB based tools. The left side shows the steps performed on the SHG images, while the right side shows the steps performed on the bright field images. The middle column combines information from both modalities. 64

Figure 18. Photo of the CAMM imaging system (A) and optical block diagram panel (B). PMT = photomultiplier tube, QCam = QImaging RGB camera, LED = bright field lamp, DBS = dichroic beamsplitter, TL = tube lens, CL = condenser lens, L = lens, M = mirror, BE = beam expander, Z = z-direction translation, XY = xy translation, XYG = xy galvonometer driven mirrors, PC = Pockel's cell, RAMM = Rapid Automated Modular Microscope (ASI), λ_2 , λ_4 = half and quarter waveplates. 66

Figure 19. Integrated fiber angle and epithelial boundary feature algorithms. Panel A shows a single fiber (green) and epithelial region boundary (red) with one highlighted fiber endpoint v . Zoomed versions of panel A are shown in panels B, C and D where individual image pixels are represented as filled circles. The nearest distance from v to the boundary is indicated in panel B, intersection between the endpoint extension line and the boundary is shown in panel C and the quadratic curve fit to the boundary at the intersection point and tangent line are shown in panel D. 74

Figure 20. Micro array SHG overlay on H&E image with zoomed breakout panel. Scale bar = 25 μm 75

Figure 21. Examples of TACS-3 positive regions. Collagen fibers (yellow) registered and overlaid onto bright field images of H&E stained epithelial cells. These images illustrate the features of the TACS-3 signature, particularly, straightened aligned collagen fibers that terminate at steep angles relative to epithelial region boundaries. Scale bar = 25 μm	76
Figure 22. Sample fiber extraction results. The resulting image after curvelet denoising (A) shows likely fiber pixels in white and likely background pixels in grey. The extracted fiber network is overlaid on the original SHG image (B) showing many appropriate segmentations and a few under- and over-segmentations (green arrows). Scale bar = 50 μm	77
Figure 23. Sample epithelial cell segmentation results. The raw probability map produced by the Trainable Weka Segmentation ImageJ plugin (A) is filtered by a 2D Gaussian filter matched to the average cell cluster width (B) and thresholded (C) to produce epithelial region boundaries which are overlaid onto original bright field images to validate the segmentation (D). Scale bar = 100 μm	79
Figure 24. KM curves for DFS and DSS showing the prognostic classification produced by our eTACS approach. For TACS-3 Positive N=112 and for TACS-3 Negative N = 84.....	80
Figure 25. Bar graph showing the mean feature difference for the top 18 features. The features that had larger values for TACS-3 positive are towards the top of the graph and those with smaller values for TACS-3 positive fall towards the bottom of the graph. Green lines indicate the $\sigma/4$ away from the mean.....	82
Figure 26. A focused beam (red) traveling in the negative z direction with an arbitrary polarization angle β , incident on a collagen fiber (yellow) oriented at an arbitrary elevation angle δ	88
Figure 27. X-Y (A) and X-Z (B) projections of a 50X50X50 μm volume captured with MPFE (red) and backward SHG (green). The X-Y projection is a single section from a z-stack while the X-Z projection is a 5 μm thick maximum intensity projection in the Y-dimension. Fiber extraction was	

performed on both the MPFE and SHG image stacks. Comparing panel A and B immediately reveals that the excitation volume is longer in the Z-dimension compared to the X and Y-dimensions. The extraction and reconstruction for the SHG channel is shown in C while the distribution of elevation angles is shown in D. Scale bar = 10 microns.....93

Figure 28. X-Y (A) and X-Z (B) projections of a 50X50X50 μm volume captured with MPFE (red) and forward SHG (green). The X-Y projection is a single section from a z-stack while the X-Z projection is a 5 μm thick maximum intensity projection in the Y-dimension. Fiber extraction was performed on both the MPFE and SHG image stacks. The extraction and reconstruction for the SHG channel is shown in C while the distribution of elevation angles for both MPFE and SHG is shown in D. Scale bar = 10 microns.95

Figure 29. X-Y (A-C) and X-Z (D-F) projections of a 25X25X25 μm volume of an unstained collagen gel captured with forward SHG using 1.0 (A,D), 0.8 (B,E) and 0.5 (C,F) NA. The X-Y projection is a single section from a z-stack while the X-Z projection is a 5 μm thick maximum intensity projection in the Y-dimension. Arrows highlight the presence of axially oriented collagen fibers visible even down to NA 0.5. Scale bar = 5 microns.96

Figure 30. A cleared mouse mammary tissue sample (A) compared to a similar but un-cleared tissue (B). The cleared tissue was incubated in the ScaleA2 formulation for 48 hours, while the control tissue was incubated in phosphate buffered saline for 48 hours. The effective multiphoton imaging depth for the cleared sample was only about 20 microns deeper compared to the un-cleared sample.....101

Figure 31. Block diagram of the SETI hardware (A) and a photograph of the prototype system (B) with inset zoomed in on the resin embedded sample. FCam = Fluorescence camera, BFCam = Bright field camera, WL = White light lamp, LED = light emitting diode lamp, S = sample.....104

Figure 32. Standard light microscope PSF (A) and MTF (B) demonstrating the missing cone problem for a 0.2 NA objective lens. The vertical direction in these figures is the optical axis of the microscope. There is only a sliver of Fourier space that is represented by the data captured by a standard light microscope. The SETI PSF (B) and MTF (D) are modified by the physical sectioning feature of the system. Physical sectioning adds high frequency data along the optical axis of the microscope, improving the missing cone problem.....	107
Figure 33. Model of SETI image formation in a clear sample. The simulated object (A) is a model of a low concentration mixture of subresolution fluorescent beads suspended in resin. The raw SETI image (B) is produced by convolving the PSF with the simulated object and decimating in the z-dimension to the z-step size of the microscope. The z-dimension blurring is corrected in C by a simple subtraction algorithm that subtracts image i from the next image i+1. The inset in C shows the artifact left by this simple algorithm.	108
Figure 34. Simulated fluorescent beads in plastic. The simulated raw data (A) is compared to a simple subtraction based reconstruction algorithm (B) and a filtered subtraction algorithm (C).....	110
Figure 35. 6 micrometer diameter fluorescent beads embedded in resin and imaged with SETI. The raw data (A) is processed with the simple subtraction filter (B) and the filtered subtraction algorithm (C). These experimental results are qualitatively similar to our simulation results presented above.	110
Figure 36. A 3D reconstruction of the bony structures within the regenerated limb of an axolotl (A) and the folds of a fetal mouse bladder with the developing prostate on the lower right (B).	111

Abstract

Mortality in cancer patients is directly attributable to the ability of cancer cells to metastasize to distant sites from the primary tumor. This migration of tumor cells begins with a remodeling of the local tumor microenvironment including changes to the extracellular matrix. In breast cancer, it has been proposed that the alignment of extracellular collagen fibers surrounding tumor epithelial cells can serve as a quantitative image-based biomarker for survival of invasive ductal carcinoma patients. This dissertation reports on the design and validation of a set of optical and computational tools to enable large-scale investigations of the interaction between collagen and tumor cells.

We have built a novel imaging system which captures large field of view images of collagen and cells within tissue sections on microscope slides. The system uses second harmonic generation (SHG) imaging to capture collagen information and bright field imaging to capture information about cellular structures. We have named the system the compact automated multiphoton microscope (CAMM).

We have developed computational tools that semi-automatically score collagen interactions with tumor cells via a variety of metrics, a method we call Electronic Tumor Associated Collagen Signatures (eTACS). The eTACS system produced classifications that had statistically significant correlation with manual TACS classifications and significantly prognostic correlation with disease specific survival. The eTACS classifications accurately predicted breast cancer patient recurrence producing a disease free survival hazard ratio of 2.59. Feature rank analysis revealed that TACS positive fibers were more well aligned with each other, generally lower density, and terminated within or near groups of epithelial cells at steeper angles of interaction.

In addition, we have investigated the limitations to 3D collagen imaging with SHG and have observed that forward SHG detection is required for accurate 3D imaging. We have also developed a prototype

system using single photon fluorescence optical imaging combined with physical sectioning to capture three dimensional images of whole mount tissues. We call the system SETI (Sequential Erosive Tissue Imaging). SETI is being developed to study contrast agent localization, prostate cancer progression mechanisms, and breast cancer extracellular matrix properties.

Chapter 1 Introduction

Mortality for cancer patients depends significantly on the propensity of cancer cells to metastasize. Since breast cancer, like many other cancers, is a diverse and heterogeneous disease, it can be difficult to predict the metastatic potential of a specific breast cancer tumor. This variability suggests that treatments should be patient and even tumor specific. The development of new laboratory assays and digital pathology scoring systems have helped to classify patients into high and low recurrence risk groups and have helped to identify patients as candidates for some specific treatments. However, challenges still remain related to the clinical adoption and overall effectiveness of these new techniques. Clearly, physicians need better tools for stratifying breast cancer patients based on metastatic potential. The aim of this dissertation is to develop semi-automated, clinically relevant optical imaging and analysis tools for determining whether collagen alignment can predict breast cancer patient outcome. The resulting tools may be used to further develop an emerging group of collagen alignment related candidate biomarkers. These biomarkers have the potential to aid in rapid and precise stratification of cancer patients and may allow for individualized decision making by patients and physicians, ultimately leading to improved clinical outcomes. The specific aims of this dissertation are as follows:

Specific Aim 1. **Develop optical imaging techniques for collagen alignment assessment.**

Specific Aim 2. **Develop scalable analysis methods for collagen alignment quantification.**

Specific Aim 3. **Validate quantification measures by correlation with clinical endpoints.**

Recent studies have shown that collagen fiber alignment plays an important role in cancer progression [1-3], potentially predicting patient outcome. With these specific aims, we build upon these previous studies by designing the technology needed to perform collagen alignment imaging and analysis on significantly larger cohorts. These tools are critical components in the effort to validate collagen alignment as more than just a candidate breast cancer biomarker. We assess collagen fiber alignment

with a combination of second harmonic generation (SHG) imaging and wide-field microscopic imaging. The SHG technique offers the unique ability to capture sub-micron resolution, high contrast images of collagen fibers in either fresh, unstained tissue or prepared histopathology slides. We also develop a unique surface-imaging platform for capturing 3D images of collagen and other biological structures in whole mount tissues.

1.1 Significance

Breast cancer is a significant disease worldwide, with approximately 1.38 million women diagnosed with the disease in 2008 [4]. Progress has been made in diagnosing and staging breast cancer. However, physicians still need better tools for selecting optimally beneficial treatments. Outcome predictions are important for making treatment decisions, for example whether systemic chemotherapy should be administered in early stage breast cancer. If outcome is likely to be worse, then more aggressive treatment options may be selected. Whereas if outcome is likely to be better, then aggressive treatment may be forgone in order to avoid potentially dangerous side effects. The following sections outline the literature relevant to the development of tools for predicting breast cancer patient outcome based on collagen alignment images.

1.2 Optical Biopsy

Diagnostic imaging in clinical medicine is dominated by X-ray, CT, MRI, PET, SPECT, and ultrasound. These techniques are the workhorses of clinical medicine, feeding information into the diagnostic and treatment decision making process. Each of these systems benefit from being highly non-invasive. However, especially in the case of cancer, a definitive diagnosis and treatment plan is not made until a surgical specimen has been removed, sectioned, stained and observed under a light microscope by a pathologist. Although the former clinical imaging systems provide extremely useful clinical information over large fields of view, none of the non-invasive techniques provide cellular resolution or the ability to

directly visualize the important features of the extracellular matrix that are critical to disease diagnosis and staging. These details can most readily be observed using the light microscopy techniques used by pathologists for centuries [5]. The standard histopathology workflow starts by excising a small piece of tissue from the patient. This tissue is embedded in paraffin wax, sectioned at 5 μm thickness and placed onto microscope slides. The tissue is then stained, permanently mounted under a coverslip and viewed with a wide field, white light microscope. This technique provides incredible amounts of information to the trained pathologist about the health of the cells in the tissues of interest. The main disadvantage to histopathology is the highly invasive nature of the process. Tissue must be surgically removed from the patient adding to the cost and risk of the procedure. Sample bias becomes a potential issue, since it is not feasible to section, stain and observe entire tissues. Usually only a few regions are selected from the gross surgical specimen and only a few sections from each region are observed under the microscope. In addition, traditional histopathology processing and analysis are time consuming, causing potentially significant delays in the clinical decision making process, and require highly trained individuals to perform each part of the process. In an ideal situation, such exquisite detail could be extracted from tissues without the need for tissue excision. This ideal technique is often described as an optical biopsy [6-8], which infers histopathology level diagnostics without the need for excision. If excision was not a requirement, tissue could be imaged longitudinally over time allowing the acquisition of more samples and to assess disease progression or treatment response. In addition, patients would not need to suffer the side effects associated with the surgical removal of precious tissue. Tissue could also be assessed prior to, during or following surgical removal, effectively guiding surgery without the need to wait for histopathology processing. With its high resolution and deep optical sectioning capability, multiphoton microscopy (MPM) promises to be a key modality in the quest to develop optical biopsy systems and may ultimately contribute to overcoming the need for investigative surgical removal of tissue. Although more work remains, significant progress has been made in the development of instrumentation for

MPM based optical biopsy such as the work of Webb and colleagues who have developed a miniature MPM endoscope [9]. One of the main challenges associated with in-vivo or in-situ MPM is that of contrast. In standard wide field microscopy, contrast is generated by colorimetric or fluorescent stains. The same is generally true for optical sectioning techniques such as MPM. Unfortunately, administration of stains or fluorescently conjugated labels brings risks for deleterious side effects and generally reduces the value of these methods for routine screening. Therefore, if MPM is to be used for in-vivo optical biopsy, then imaging without stains or labels is desirable. Many naturally occurring biological molecules with intrinsic fluorescent signatures are currently being investigated for their potential as surrogate stains for in-vivo or in-situ MPM.

1.3 Intrinsic Contrast Imaging with Multiphoton Microscopy

Although standard fluorescent labeling approaches using injectable fluorophores or green fluorescent proteins have been the method of choice for most multiphoton imaging studies to date, intrinsic fluorescence multiphoton imaging allows for essentially non-invasive inspection of tissue, a key requirement for optical biopsy. This technique uses naturally present fluorophores such as NADH, elastin, collagen or flavins. Multiphoton induced intrinsic fluorescence was studied early on by Xu et al. [10], but was considered mainly for reasons of rejecting "undesirable" autofluorescence. This work mapped the two-photon excitation (TPE) spectra for metabolic molecules NADH, NADPH and flavins FMN and FAD. Other early work by Masters et al. used signals from autofluorescent metabolites to estimate redox ratios in tissues [11, 12]. Zipfel et al. [13] later studied the properties of many intracellular and extracellular autofluorescent molecules and used them to extract important physiologic and anatomic information from a wide range of tissues. Skala et al. used intrinsic fluorescence multiphoton imaging to estimate in-vivo NADH and FAD redox states and showed that redox ratios changed significantly between undifferentiated and mature basal epithelial cells [14]. Vergen et. al. used the fact that multiphoton excited fluorescence lifetimes differ between free and enzyme bound NADH

to estimate metabolic state [15]. In addition to these intracellular metabolites, a variety of extracellular structures also exhibit intrinsic contrast. Collagen is an important and abundant extracellular matrix component that generates intrinsic contrast via second harmonic generation (SHG), a special case of TPE [16]. SHG has been used to visualize collagen structures in many disease states and is a key imaging method used throughout this dissertation. We will discuss the role of collagen in cancer progression below, then discuss methods for collagen imaging in subsequent sections.

1.4 Collagen's Role in Cancer Progression

Collagen plays an important role in breast cancer morphogenesis and invasion [17-20] and may be a key factor in the link between breast density and breast cancer risk [21, 22]. In a mouse model, early tumor cell clusters were observed to be encapsulated by tangential, circumferential collagen matrices [23]. At later stages, the circumferential pattern was observed to be replaced by a radial pattern. This radial pattern, in addition to providing less encapsulation, may also provide a scaffolding on which cells with metastatic potential can climb away from the primary tumor. The angular orientation of collagen fibers with respect to cellular boundaries, as observed with SHG imaging, has been shown to correlate with cancer invasiveness [2, 3, 24-34] and clinical outcome for patients [1]. The collagen patterns observed have been termed "tumor-associated collagen signatures" or TACS [3]. In particular, the third TACS signature, so-called TACS3, corresponds to straightened bundles of radially aligned collagen around the tumor. The presence of the TACS3 phenotype has been shown to be a strong prognostic indicator for breast cancer patients [1]. In addition, TACS3 was shown to be prognostic independent from other well established biomarkers such as human epidermal growth factor receptor 2 (HER2) and estrogen receptor (ER), making it a potentially valuable classifier for further stratifying patients within breast cancer subgroups. This is potentially significant because sub-classification could allow for more specifically tailored treatment decisions. Although TACS3 was shown to be a powerful, independent prognostic factor in a single clinical study to date, additional studies are necessary to further validate

TACS3. Furthermore, significant technical challenges remain to enable clinical trials and exploit the potential of TACS3 and other collagen alignment changes as a research and clinical tool. These challenges and technical innovations in response are addressed by the aims of this dissertation.

1.5 Collagen alignment imaging

Collagen is naturally eosinophilic, therefore can be visualized in routine histological preparations such as hematoxylin and eosin (H&E) stained slides. Unfortunately, eosin has an affinity for many other proteins such as elastin or proteoglycans which makes visual separation of collagen from other extracellular structures difficult in H&E stained slides. Picosirius red or Masson's trichrome stains combined with polarization microscopy provide more specific staining of collagen, allowing for individual fibers to be readily observed with high contrast [35, 36]. Transmission and scanning electron microscopy have also been used for imaging collagen at nanometer resolution [3, 37]. These techniques all require extensive tissue processing which causes tearing and dehydration artifacts and severely limits the observation of three dimensional (3D) collagen architecture. 3D volumes have been created from histological sections [38, 39]. However, the process is extremely labor intensive and encounters difficulty with section alignment. Alternatively, optical sectioning microscopy methods have been used to great success for observing collagen architectures including reflectance confocal imaging [40, 41] and most notably SHG imaging microscopy [13, 42-44]. An example of an SHG image of the collagen surrounding the mouse mammary fat pad is shown in Figure 1. This image demonstrates the specificity to collagen and the resolution that SHG imaging affords. All of the thread like structures are type I collagen fibers while cell clusters occupy the empty regions between fibers. In this image, we are able to resolve fibers with diameters less than 1 μm .

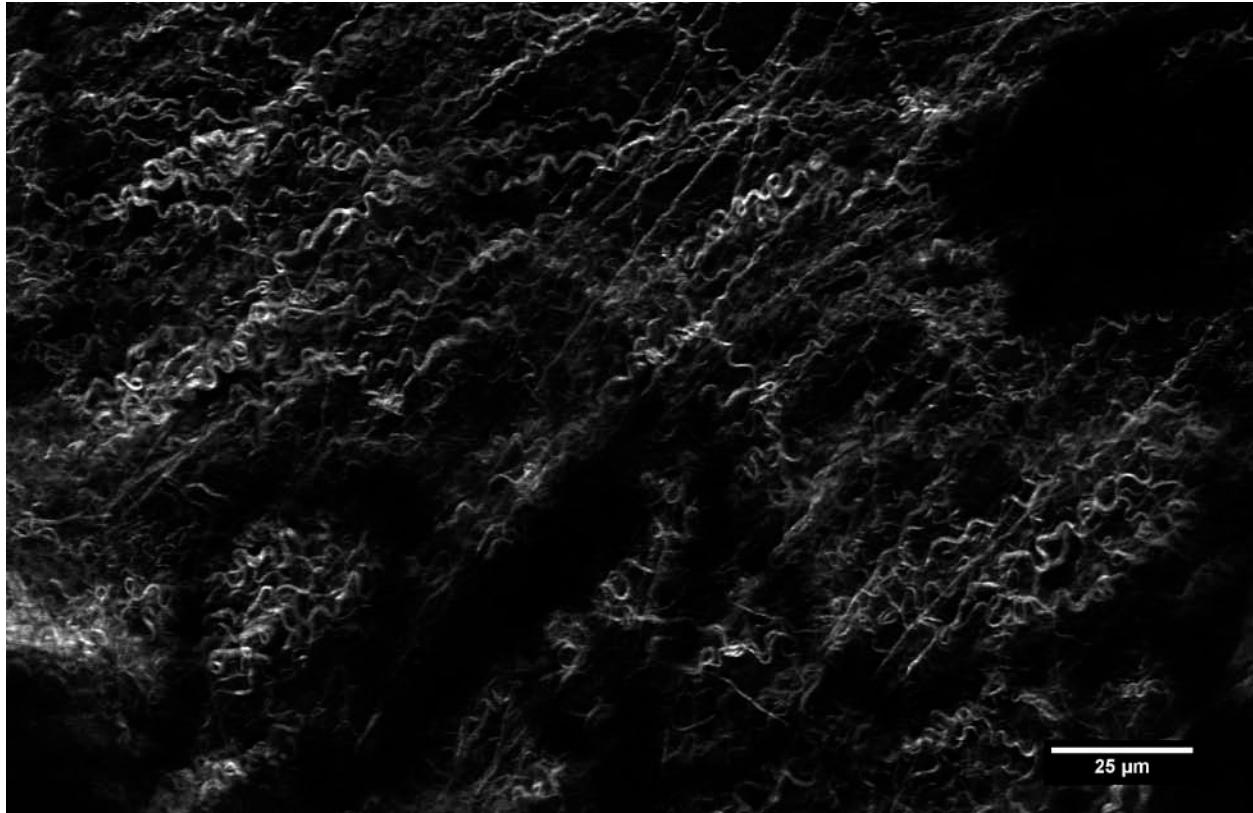


Figure 1. SHG image of collagen fibers in fresh, unstained mouse mammary tissue.

In circumstances where in-vitro collagen samples are of interest, collagen monomers can be labeled with fluorescent probes and visualized via confocal or multiphoton microscopy [45]. Although not microscopic in nature, ultrasound [46, 47], MR [48, 49] and CT [50, 51] have all been used to probe collagen architectures. This dissertation will focus on the development of SHG imaging microscopy for capturing images of collagen alignment due to its ability to capture high resolution/high contrast images of collagen in samples of fresh tissue or histopathology slides.

1.6 SHG Imaging Microscopy

A major historical challenge for thick sample microscopy has been the elimination of out of focus light, a challenge commonly referred to as the "missing cone" problem [52]. The "missing cone" corresponds to the empty regions of the modulation transfer function along the optical axis of an imaging system. This problem was addressed by the confocal microscope using a pinhole to spatially filter light from above

and below the focal plane [53], effectively filling in the missing cone in Fourier space. A two-photon excited fluorescence (TPEF) microscope achieves similar out of focus light elimination, with improved light collection efficiency, improved fluorescence stability and reduced photo toxicity [54]. The fluorescence excitation volume in TPEF is reduced to the region where the likelihood of simultaneous absorption of two photons is appreciable. A SHG microscope benefits from all of the advantages of multiphoton microscopy, but uses second harmonic generation rather than TPEF [55, 56]. The standard layout of an epi-detecting SHG microscope is shown in Figure 2.

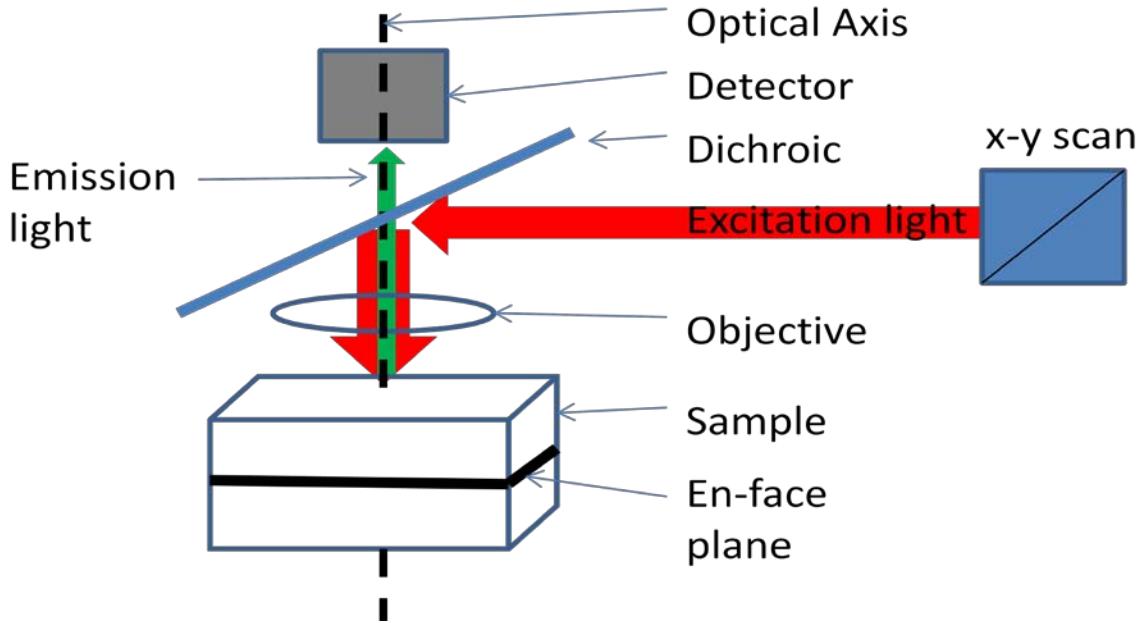


Figure 2. Standard optical layout of an epi-detecting SHG microscope. Images are captured in the en-face plane by scanning the focus of the beam in two dimensions. 3D imaging is performed by translating the objective or sample in the third dimension. Red is the excitation light and green is the emission light.

Second harmonic light is only produced in materials that lack a center of symmetry, such as type I collagen. This allows for excellent image contrast for type I collagen fibers. SHG microscopy has been used to the study the role of collagen in a wide variety of tissues and diseases [1, 3, 13, 26, 34, 42, 43, 57-64].

1.7 3D SHG Imaging

In second harmonic generation imaging, the high-intensity, incident electric field mixes with the non-linear susceptibility tensor of the sample to induce a second order, non-linear polarization in the sample. This occurs most efficiently if the polarization vector of the incident light is oriented parallel with the charge transfer axis of the harmonophore [65]. Each polarized subunit of the sample material can be modeled as a dipole antenna and Green's function [66] or phased array approaches [67] can be used to solve for the far field radiation pattern. The polarization sensitivity of SHG has many advantages. For example, by measuring the response of a sample to varying linear polarizations and by measuring the polarization of the SHG emission light, information about molecular pitch angle [44, 68], fiber orientation [69] and susceptibility tensor information [70] can be extracted from the sample. Each of these pieces of information are currently being used to study disease states in collagen related diseases. Although the coherent nature of SHG has advantages, it makes 3D imaging of fiber networks potentially difficult. Since light is a transverse wave, the only dipole orientation that can be easily polarized is in the plane perpendicular to the direction of travel of the light through the microscope (the optical axis). The imaging geometry showing this principle is sketched in Figure 3.

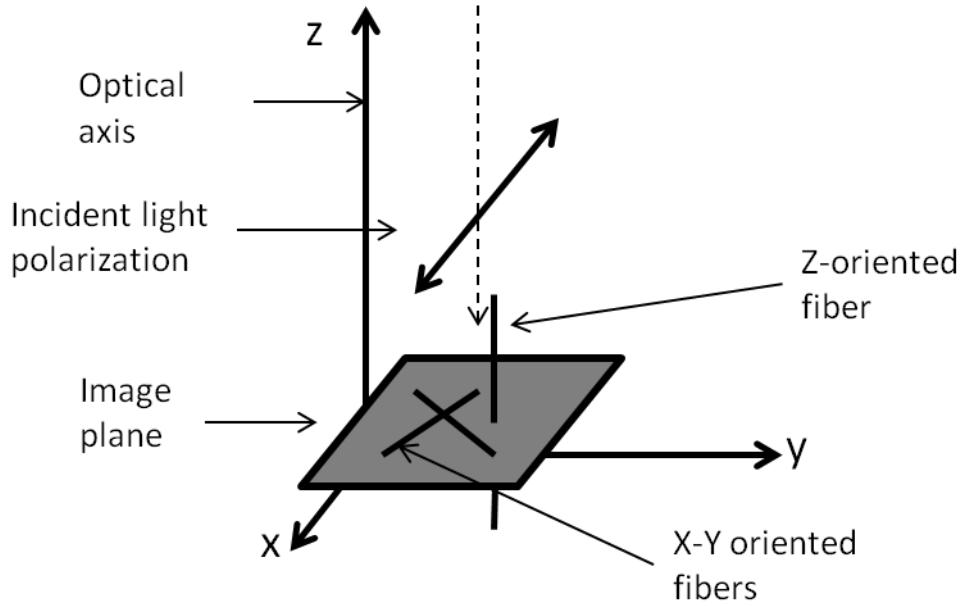


Figure 3. Geometry of 3D fiber imaging with an SHG microscope. The incident light is polarized along the x-axis and propagates in the negative z-direction (dotted line). Example fibers are oriented in the x-y plane (imaging plane) and in the z-direction.

Fortunately, the molecular dipole moments within a type I collagen have been observed to be anisotropic [69], meaning that they can be stochastically oriented with respect to the fiber axis, therefore fibers of all orientations are observed to be able to create SHG. However, the SHG light generated in axially oriented fibers is more likely to scatter in the forward direction compared to the backward direction. This has been established both theoretically and experimentally and is due to the combination of the Gouy phase anomaly, the phase matching requirements for SHG creation, and the size scale of individual type I collagen fibers. Although SHG imaging has proven useful in numerous research studies, fairly little work has been done to automate the process of SHG image capture and analysis. Many of the advances in automated histopathology imaging and analysis could be applied to further improve the value of SHG imaging and enable large scale screening for TACS as described above.

1.8 Automated histopathology imaging and analysis

Automated analysis of microscope images and, in particular, histological slides is a growing industry. Evidence for this can be found in the new digital pathology business units formed at GE, Philips and Leica focused on automated microscopy and computer aided image interpretation. There are a growing collection of commercially available imaging systems and software packages available for imaging and analyzing sectioned and stained samples of surgically resected tissue. For example, software by Aperio, GE Healthcare, and Philips all tackle this difficult challenge. These systems attempt to aid the pathologist in making diagnoses. Unfortunately, these commercial software packages are costly, inflexible and algorithms are kept confidential. None of the systems provide the imaging or analysis tools for capturing collagen alignment information. In addition none of the commercial systems can be tailored to meet the specific needs of collagen alignment scoring. The automated registration and integrated analysis of both SHG and H&E images in search of TACS is a complicated and unique task that require the flexibility and extensibility of open source image processing tools such as FIJI [71] or CellProfiler [72, 73]. Here we report on the development of a novel technique for multimodality collagen alignment scoring that uses open source publicly available software. This will allow others to easily access our code, use what we have developed and tailor our code to their needs. With an open source toolset, the collagen alignment assay will be available for investigation by other labs.

1.9 Collagen alignment analysis

Collagen alignment image analysis has been performed both manually and with a variety of automated algorithms in both open and commercial image analysis software. In the recent study where collagen alignment was associated with breast cancer patient survival, aggressive patterns of collagen alignment were identified by three panelists who were asked to look at images and find the aggressive patterns in a manual fashion [1]. In other studies, the collagen orientation was observed by tracing fibers and measuring fiber angles by hand [3, 74]. These methods are subjective and result in both inter and intra-observer errors. For these reasons, there have been a number of algorithms developed for automatically

evaluating fiber orientation in images including Fourier domain methods [75-77], fiber-tracking [40, 45, 78], vector field transforms [79], and curvelet analysis [80, 81]. Each of these algorithms have advantages, but none of them effectively brings together all information needed to analyze collagen alignment for breast cancer patient classification. This challenge is addressed in the subsequent chapters of this dissertation.

1.10 Patient classification

Breast cancer patients can be specifically classified according to numerous metrics available in health care today including tumor size, nodal involvement, estrogen receptor, progesterone receptor and HER2 status. Depending on the sub-type of cancer, patients may benefit from certain treatments and not others. For example, HER2 positive patients will benefit from targeted antiHER2 therapy [82]. A vast array of new biomarkers for breast cancer are being developed causing some to call for biomarker registries [83] and new approaches to assess potential biomarker efficacy [84]. Some of the new biomarker techniques include methods of immunohistochemistry, gene constitution, gene and protein expression, and single nucleotide polymorphism [85]. Despite these significant efforts, no assays are available that allow for the rapid analysis of fresh tissue in an operating room setting and few allow for the comprehensive evaluation of large tissue volumes. TACS-3 evaluation using SHG microscopy has the potential to solve both of these problems.

1.11 Innovation

If collagen alignment is to be clinically adopted as a new diagnostic tool, it must be shown to be a powerful prognostic or predictive indicator in many trials. It must be tested on a large patient population and shown to be able to clearly prognosticate patient survival and/or predict response to treatment. In addition, the imaging and analysis tools must be compatible with clinical workflows. These are common challenges for any new medical test or device. Indeed, translating technologies from the

lab bench to the bed side is a major challenge for biomedical science in general [86]. Collagen alignment is no exception, and thus faces significant translational challenges. For example, in our recent collagen alignment study [1], significant time and expertise was required for operation of imaging instruments and interpretation of the resulting images. Manual operation of the instrument and manual interpretation of the images introduced human error into the process. Not only would automating the imaging and interpretation process allow for higher throughput, it would eliminate subjectivity, allowing for improved precision. Automated imaging and analysis would also help make collagen alignment more clinically compatible by reducing training and labor costs. Chapter 2, Chapter 3, and Chapter 4 of this dissertation report on novel techniques to address some of these challenges, attempting to lower the barriers to collagen alignment analysis and therefore allowing this new assay to be investigated at the large scales necessary for clinical translation or application to other research problems.

In addition to aiding the translation of collagen alignment, there is a need to characterize collagen fibers in three dimensions (3D), since collagen and cells interact in a 3D environment [29, 87]. To date, collagen alignment has only been observed with respect to breast cancer invasion in stacks of 2D images of biological tissues [2]. Even though these 2D image stacks can be used to form a 3D image, the alignment of collagen in the third dimension has not been observed. We report in Chapter 5 new experimental evidence supporting models for SHG image contrast and data that can help to interpret 3D collagen fiber images. Chapter 6 summarizes the protocols, design decisions, and early results from the Sequential Erosive Tissue Imaging (SETI) system.

Chapter 2 Curvelet Image Analysis

2.1 Introduction

The extracellular matrix (ECM) is a complex meshwork of proteins (e.g., collagens, proteoglycans, laminin, fibronectin, entactin, etc.) and polysaccharides (e.g., glycosaminoglycans) secreted by cells. Components of the ECM can bind to one another and to adhesion receptors on the cell surface, providing both physical scaffolding to support tissues and mechanical and biochemical cues that regulate cellular behavior [88]. For example, in the mammary gland, the ECM impacts cell phenotype, behavior and complex cellular functions [89]. In particular, collagen I is known to have a significant effect on both mammary morphogenesis and tumor progression [89–93], playing an important role in maintaining normal cell behavior [89, 94]. Moreover, increased density and organization of the collagen matrix is one of the hallmarks of the transition from carcinoma *in situ* to invasive carcinoma [95, 96]. An important facet of this relationship is the local reorganization of collagen fibers at the tumor-stroma boundary that facilitates invasion of transformed epithelial cells into the surrounding ECM[91, 97], with tumor cells migrating directly along collagen fibers [98].

In this chapter we present a technique for semi-automatically extracting quantitative image features that describe the collagen fiber reorganization that occurs at the tumor-stromal boundary. While the techniques described here are widely applicable to imaging data obtained using multiple modalities, nonlinear optical imaging techniques, such as multiphoton laser-scanning microscopy (MPLSM [99]) are particularly useful for studying the interaction between the collagen matrix and cell behavior [100]. MPLSM is an optical sectioning technique that can simultaneously produce multiphoton excitation (MPE) and second harmonic generation (SHG). MPE occurs when two or more low-energy photons excite a fluorophore, which then emits a single photon with higher energy than the individual incident photons [99, 101]. This method restricts fluorophore excitation to the plane of focus (i.e. optical

sectioning; [99]), thus reducing phototoxic effects [100, 102] while increasing the effective imaging depth in comparison to conventional confocal microscopy [103]. SHG depends on the non-absorptive interaction of the incident photons with a non-centrosymmetric structure, such as fibrillar collagen [65, 104]. The result of this interaction is a non-linear, second order polarization, producing a coherent wave traveling at exactly one-half the incident wavelength [57, 104], allowing straightforward separation of MPE and SHG emission signals using band-pass optical filters or through measurement of the fluorescence lifetime [100]. SHG requires no staining, is specific to certain biological molecules like collagen I and myosin, and can provide high-resolution 3D data sets from thick tissue sections [105]. The combination of MPE induced cell fluorescence, whether extrinsic such as GFP labeled cells[106] or intrinsic such as metabolic co-factors[107], along with direct collagen detection provides a powerful tool for studying the interaction between normal and cancer cells and changes in collagen density and alignment.

SHG imaging has been used to study many disease states[105]. For example, in breast cancer, SHG has enabled the detection of specific and consistent changes in stromal collagen organization at various stages of tumor progression[91, 97]. Of particular relevance, two of these signatures differentiate non-invasive from invasive disease and as such correlate with metastatic potential and patient survival [1]. Specifically, non-invasive regions are contained by collagen fibers oriented parallel to the tumor boundary while regions of local invasion possess areas where collagen has been realigned perpendicular to the tumor boundary to facilitate local invasion[91, 92, 97]. Therefore, it has been hypothesized that the angle of collagen fibers relative to the tumor boundary may be used as a predictor of imminent invasion and metastasis[108]. These collagen changes, known as Tumor Associated Collagen Signatures

(TACS)[91], fall into three categories and will be used herein as a practical example of the utility of quantifying collagen architecture (Table 1):

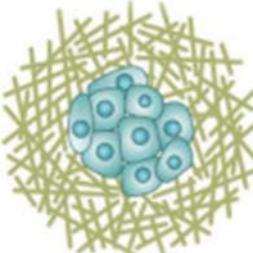
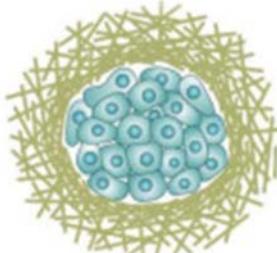
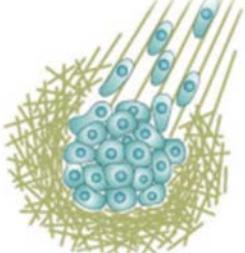
TACS-1	TACS-2	TACS-3
		
<ul style="list-style-type: none"> Region of dense collagen I surrounding tumor Collagen fibers have no specific alignment Tumors may be pre-palpable 	<ul style="list-style-type: none"> Collagen I fibers wrapped around tumor Fibers appear aligned and stretched across a relatively smooth tumor boundary Tangential orientation of fibers (approximately 0° to tumor boundary) predominant 	<ul style="list-style-type: none"> Collagen I fibers aligned normal to tumor boundary Tumor boundary is irregular in shape indicative of local invasion Fibers aligned in the direction of cell invasion (approximately 90° to tumor boundary)

Table 1. Descriptions of tumor-related collagen I structures[91].

Despite the sophisticated imaging techniques available to examine the relationship between changes in collagen I and disease state - dermal wounds [109-111], breast cancer [91, 92, 112-114], ovarian [115], prostate cancer [116], asthma [117, 118] – there is still a dearth of robust computational methods for characterizing these changes. As a result, these measurements are still often performed by hand. This is extremely time-consuming, can be sensitive to the subjectivity of the individual performing the measurements, and may be prone to inter-observer variation. Furthermore, it may take an hour or more to measure a single image which, when combined with the variability of manual measurement, makes quantitative analysis of large datasets difficult and impractical. The user variability can arise from

differing individual abilities to precisely detect minor changes in collagen alignment, which can be very subtle especially when imaging samples with high collagen density. Indeed the high detection sensitivity and excellent signal to noise ratio provided by SHG imaging allows for the detection of subtle collagen changes at the single cell level and often reveals highly complex fiber networks. However, the complex data sets produced by SHG imaging are often extremely difficult to analyze manually or with conventional image processing tools. The challenge for automated methods is to detect meaningful, significant changes to alignment within the noise produced by high levels of collagen heterogeneity observed in tissue. The ability to accurately and robustly quantify fiber angle and alignment in a highly automated fashion is a key requirement for an experimental tool to be effectively applied to collagen alignment based biomarker development. In addition, in order to improve signal to noise, collagen alignment measurements must be limited to the most important biological regions and must allow for measurements of relative collagen alignment with respect to biologically relevant interfaces, such as the tumor-stromal boundary or vascular walls. The novel tools we describe in this chapter address both of these image analysis requirements.

Previous efforts to apply signal-processing techniques to the measurement of biological structures such as collagen have been largely based on the Fourier transform (e.g.[119-123]), which works particularly well with periodic structures. Several studies have also combined SHG imaging and subsequent Fourier transform analysis to characterize collagen alignment in tissues such as skeletal muscle [124], cornea [125], cardiac muscle [126] as well as collagen gels[127, 128]. In particular Bayan, et al. [127] described the use and improvement of Fourier based methods to automatically analyze and classify collagen alignment in thick collagen gels imaged by SHG. Their approach has advantages over previously reported collagen alignment studies in that their analysis is based on a combination of the Fourier and Hough transforms, to examine both the overall alignment of collagen fibers in a gel and the local orientation of collagen fibers. Their fully automated approach begins with a novel adaptive thresholding technique

followed by Fourier transform analysis to determine the degree of collagen fiber alignment along a dominant direction in the image. The image is then decomposed into a series of small tiles comprising the whole of the image, and the Hough transform applied to each tile to determine the direction and magnitude of the dominant line in the tile. This approach gave superior results over previous studies in accuracy and reproducibility. However, this approach, like other Fourier-based methods, is still not particularly effective at detecting edges in images with overlapping fibers or curves, which are a major feature in images of the ECM. Other approaches such as filter kernel based methods have been successfully applied to finding alignment in, for example, acto-myosin fibers [129]. As an alternative, the curvelet transform, developed by Candès and Donoho [130], is specifically designed to produce a sparse representation of edges in images even in the presence of complex geometries such as those associated with stromal collagen.

Multiscale image transformations parsimoniously describe images in a sparse manner while retaining information such as position, orientation and scale. The most salient features of an image f can be represented in terms of its m -term approximation, the linear combination of the m basis functions corresponding to the largest coefficients of the transform (i.e., the m largest terms in the decomposition). A sparse representation results in fewer coefficient values of importance, resulting in better m -term approximations when compared with dense representations. While Fourier analysis works well on periodic structures, and wavelet analysis works well on singularities, neither is well suited for the task of sparsely representing edges, nor is wavelet analysis particularly suitable for determining orientation [131]. There are other edge detection methods, such as Gabor filters[132] and the Canny algorithm[133] , that are suitable for finding the orientation of edges in images. Gebäck and Koumoutsakos compared these two methods to the curvelet transform for edge detection in microscopy images[134]. They found that, in comparison to the curvelet transform, the other two approaches were much more sensitive to noise and did not perform as well as the curvelet transform in

finding the main features of an image[134]. In addition, the greater flexibility in parameter choice when using Gabor filters (as opposed to the fixed, discrete partitioning of the frequency domain used by the curvelet transform) can make the filters more complicated to use due to difficulty finding an appropriate parameter set[134].

Curvelets are a class of directional wavelet-like functions and, like wavelets in general, have been shown to be good at compression and de-noising. However, unlike classical wavelets, curvelets are very good at detecting and representing edge orientation and alignment. There exists a closely related class of directional wavelet-like functions[135], called contourlets, that are also sensitive to edge directionality. For the purposes of our studies we are using the curvelet transform, due to the existence of the curvelet transform software library, its ability to give very accurate measurements of angle distributions for defined regions of interest and the demonstrated performance of curvelets in other applications including seismic data processing [136-138], edge detection and denoising of images [134, 139, 140], and feature extraction [141]. Similar results may be obtained with contourlets due to the very similar function of the two transforms.

Here, we demonstrate application of the curvelet transform to quantify collagen architecture and present a new software tool, CurveAlign, for measuring collagen alignment. The authors of the curvelet transform^{41, 61} have made CurveLab toolbox freely available as a library, www.curvelet.org. This library is a MATLAB (Natick, Massachusetts: The MathWorks Inc.) code base designed to allow developers to interface to their specification application. The CurveLab toolbox is the algorithmic engine for our CurveAlign application, which features a graphical user interface to allow for user-guided, semi-automated measurement of collagen angles using the curvelet library. While curvelets have been used previously in image enhancement and feature extraction [142, 143], MRI image denoising of breast cancer images[139] and image-based breast cancer diagnosis[143], we present below to our knowledge

the first application of the curvelet transform to the problem of collagen alignment in cell resolution image data, particularly in SHG images.

2.2 Materials and Methods

2.2.1 The Curvelet Transform

The curvelet transform, which was originally proposed in [130], is an overcomplete representation in terms of local, wavelet-like functions, each associated with a specific scale, orientation and position. It is shown in [130] that the m -term approximation using a curvelet representation outperforms both the Fourier and wavelet representations, in the sense that fewer curvelets, than sinusoids or wavelets, are needed to accurately represent image structure. Conceptually, the curvelet transform is a multiscale pyramid with many orientations and positions at each length scale, and needle shaped elements at fine scales. Curvelets obey a parabolic scaling relation which says that at scale 2^{-j} each element has an envelope that is aligned along a "ridge" of length $2^{-j/2}$ and width 2^{-j} [144]. A brief overview of the mathematical framework from Candès, et al.[144] is now presented to give the reader a formal representation of curvelets. Consider the case where our spatial variable x lies in \mathbb{R}^2 . Start with a windowing of the space both radially and angularly and call these windows $W(r)$ and $V(t)$ respectively, where $r \in (\frac{1}{2}, 2)$ and $t \in [-1, 1]$. Now, consider the frequency window U_j defined in the Fourier domain by:

$$U_{j(r,\theta)} = 2^{-3j/4} W(2^{-j}r) V\left(\frac{2^{\lfloor j/2 \rfloor} \theta}{2\pi}\right)$$

Where $\lfloor j/2 \rfloor$ is the integer part of $j/2$. Therefore, the support of U_j is a polar "wedge" defined by the support of W and V . The equispaced sequence of rotation angles are defined as $\theta_l = 2\pi \cdot 2^{\lfloor -\frac{j}{2} \rfloor} \cdot l$

with $\mathbf{l} = \mathbf{0}, \mathbf{1}, \mathbf{2}, \dots$ such that $0 \leq \theta_l < 2\pi$, with translation parameters $\mathbf{k} = (\mathbf{k}_1, \mathbf{k}_2) \in \mathbb{Z}^2$. The waveform $\varphi_j(\mathbf{x})$ is defined by its Fourier transform, $\widehat{\varphi}_j(\omega) = U_j(\omega)$. Curvelets at scale $2^{-j/2}$, orientation θ_l and position $x_k^{j,l} = R_{\theta_l}^{-1}(\mathbf{k}_1 \cdot 2^{-j}, \mathbf{k}_2 \cdot 2^{-j/2})$ where R_θ is the rotation by θ radians, can now be defined as:

$$\varphi_{j,k,l}(x) = \varphi_j(R_{\theta_l}(x - x_k^{j,l}))$$

A curvelet coefficient is then the inner product between an element $f \in L^2(\mathbb{R}^2)$ and a curvelet $\varphi_{j,k,l}$:

$$c(j, k, l) := \langle f, \varphi_{j,k,l} \rangle = \int_{\mathbb{R}^2} f(x) \overline{\varphi_{j,k,l}(x)} dx$$

Our CurveAlign program uses an implementation of the curvelet transform called the Discrete Curvelet Transform via Wrapping. The details of the wrapping implementation of the algorithm have been previously described[144].

2.2.2 CurveAlign

CurveAlign allows a user to easily open their image or images (for comparison) and analyze collagen alignment. The program has several options for how metrics can be presented both as a graphic representations and as comma separated value (csv) files that may be imported into another program for subsequent analysis. The tool is available freely at <http://loci.wisc.edu/software/curvelet-based-alignment-analysis>, with full install, usage instructions and example data.

There are two different measurement schemes built into this program: a boundary-free measurement scheme that measures the distribution of fiber alignment in an image with respect to an absolute

reference, and a scheme that measures the distribution of fiber alignment in an image relative to a user-defined boundary.

The basic procedure is as follows:

1. Perform the 2D FDCT on the input image
2. Discard all coefficients except those in the second-finest scale
3. Threshold the remaining coefficients based on a user defined threshold (generally keeping only the largest .1- 1%)
4. Find the center and spatial orientation of each curvelet corresponding to the remaining coefficients
5. Application-specific analytics are then performed using the measured angles and locations

In the boundary-free measurement scheme, the default reference is the horizontal axis and the list of angles generated in step 5 above is the list of absolute angle values. The boundary measurement scheme requires the user to draw a line on the image that corresponds with a biologically relevant interface by clicking points along that boundary. In order to compute a relative angle between a curvelet and a boundary, there must be an association made between each curvelet and a single position on the boundary. This is performed by first identifying if the line created by the curvelet orientation and center point intersect with the boundary. If an intersection point is found, then this point is associated with the curvelet. However, if no intersection point is found, then the nearest point on the boundary is found using a nearest neighbor search algorithm[145] and that point is associated with the curvelet. When an association point is found on the boundary, then the angle of the boundary at that location is determined by computing the angle of the tangent line to the boundary at that point. The relative

angles between each curvelet and each associated point on the boundary line are then calculated and returned.

Once the angles have been determined further statistics are computed. For example, the mean, median, variance, standard deviation, skewness and kurtosis of the distribution are calculated and written out to a CSV file. The directional statistics methods published by [146] are used for all statistical analysis of the fiber orientation information. The overall alignment of the fibers is determined by calculating the resultant vector length of all orientation vectors. This yields a unitless number between 0 and 1 that indicates how well the distribution is grouped about the median angle, with 0 being completely random and 1 being completely aligned in the direction of the median angle.

In both of the measurement schemes described above, there is no need for any pre-processing such as thresholding or denoising. Both are accomplished by the selection of the appropriate curvelet coefficients for the analysis. The very highest and lowest scales of the curvelet transform, which contain the high frequency noise and the low frequency background in the image, are discarded. The image is further thresholded/denoised by the user's selection of a coefficient threshold – the very largest coefficients correspond to the highest frequency edges in the image, which, in the case of a collagen SHG image, are the very strongest and most defined fibers. The larger the threshold value, the greater the number of fibers that will be analyzed (see Figure 4).

The available outputs from CurveAlign are:

1. Overlay Image - This allows the user to see where curvelets were found within the image by overlaying center point and orientation of each curvelet on the original image.
2. Local Orientation Map - Indicates regions of dense local alignment of collagen fibers within an

image.

3. Reconstructed Image – this is an image reconstructed from the thresholded curvelets. It shows all of the edges in the image that were measured.
4. Histogram – a bar histogram of the measured angles, either with respect to the boundary or to the horizontal axis if no boundary is used.
5. Compass Plot – an angular histogram of the measured angles
6. Values – the values of the measured angles as well as basic descriptive Statistics

The algorithm for generating the local orientation map is as follows. First, an image is created where the grey level at each point containing a curvelet is assigned a value that corresponds to the absolute or relative orientation of the curvelet in that pixel. All pixels without a curvelet are assigned a value of zero. This image is then processed using a 2-D max filter followed by a 2-D Gaussian filter. The max filter and the Gaussian filter widths are defined by default as one 64th of the width of the image, but should be adjusted depending on the pixel size in the image.

The curvelet threshold, boundary (if used) and the choice of outputs will be applied to all of the input images. Output is saved automatically. Image outputs are saved in tiff format. Histograms, compass plots, values and statistics are saved in CSV files. Histograms are saved as a row of counts above a row of bins, compass plots are saved as the x and y components of the curvelet vectors.

2.2.3 Validation Analysis

We validated CurveAlign against manual collagen angle measurements in six real SHG images of mouse mammary tumors and five computer generated test cases (results in Tables 2 and 3). Among the six real SHG images were three examples of TACS-2 collagen alignment and three examples of TACS-3 collagen alignment. Manual collagen angle measurements were performed using the region of interest selection tools in ImageJ. The computer generated test cases contained simulated collagen fibers that attempted to recapitulate the fiber length and angle distributions observed in collagen gels of 1.0 mg/mL concentration. The resulting manual and automated distributions of measurements were compared using a two-sample Kolmogorov-Smirnov test.

2.2.4 SHG Microcopy

For all second harmonic imaging, a custom multiphoton workstation at the University of Wisconsin Laboratory for Optical and Computational Instrumentation (LOCI) was used. This system and SHG approaches have been reviewed extensively[91, 97, 108, 147, 148]. The samples were typically imaged using a TE300 inverted microscope (Nikon, Tokyo, Japan) equipped with a Plan APO VC 20X air immersion (N.A. 0.75, WD 1mm; Nikon Instruments, Tokyo, Japan) objective lens by using a mode-locked Ti:Sapphire laser (Spectra Physics Mai Tai, Mountain View, CA). Tuning the excitation wavelength to 890 nm, a 445 nm \pm 20 nm narrow band pass emission filter (Chroma, Burlington, VT) was used to detect the SHG signal of collagen in the backscattered mode using a H7422P-40P GaAsP photon counting photomultiplier tube (Hamamatsu Photonics KK, Shizuoka, Japan). All images (either 512X512 or 1024X1024 pixels) were acquired using WiscScan (<http://www.loci.wisc.edu/software/wiscscan/>).

2.2.5 Mouse Mammary Tissues and Tumor Explants

Mouse mammary images used for validation in this study were originally created and acquired for Provenzano, et al.[23] and methods are fully described therein. The human samples described in Figure 8 were collected with Institutional Review Board approval at the University of Wisconsin-Madison. They were immediately snap frozen following excision, transported to the imaging location, thawed on ice,

fixed in 4% paraformaldehyde, embedded in 1% agarose gel, sectioned to 300 microns thick, and finally stained for E-cadherin using anti-body conjugated DyLight 488.

2.3 Results and Discussion

The fundamental advantage of the curvelet transform for collagen alignment analysis is the ability of the transform to retain orientation information from the image. This results in the ability to examine all prominent edges at a particular orientation and a particular scale. When applied to the collagen alignment analysis problem, the curvelet transform becomes a powerful tool for detecting the presence of filamentous structures, their location, scale, and orientation. By obtaining accurate quantitative data regarding collagen amount, morphology, and organization/orientation, biologically relevant data can be derived.

CurveAlign is not intended to precisely follow individual fibers, but rather to determine the overall trend in fiber alignment in an image. The measured angles are binned according to the θ_l divisions of the frequency domain (described in the experimental section). In addition to the binning that is inherent to the transform, the measurements from an image are weighted with the thickest, longest or brightest fibers being associated with the highest number of large curvelet coefficients. In Figure 4, the location and orientation of curvelets corresponding to large coefficients are represented as green circles and green lines respectively on a synthetic image of crisscrossing lines. This example demonstrates the ability of the curvelet transform to highlight prominent edges in an image and maintain accuracy even at locations where fibers may be cross-linked.

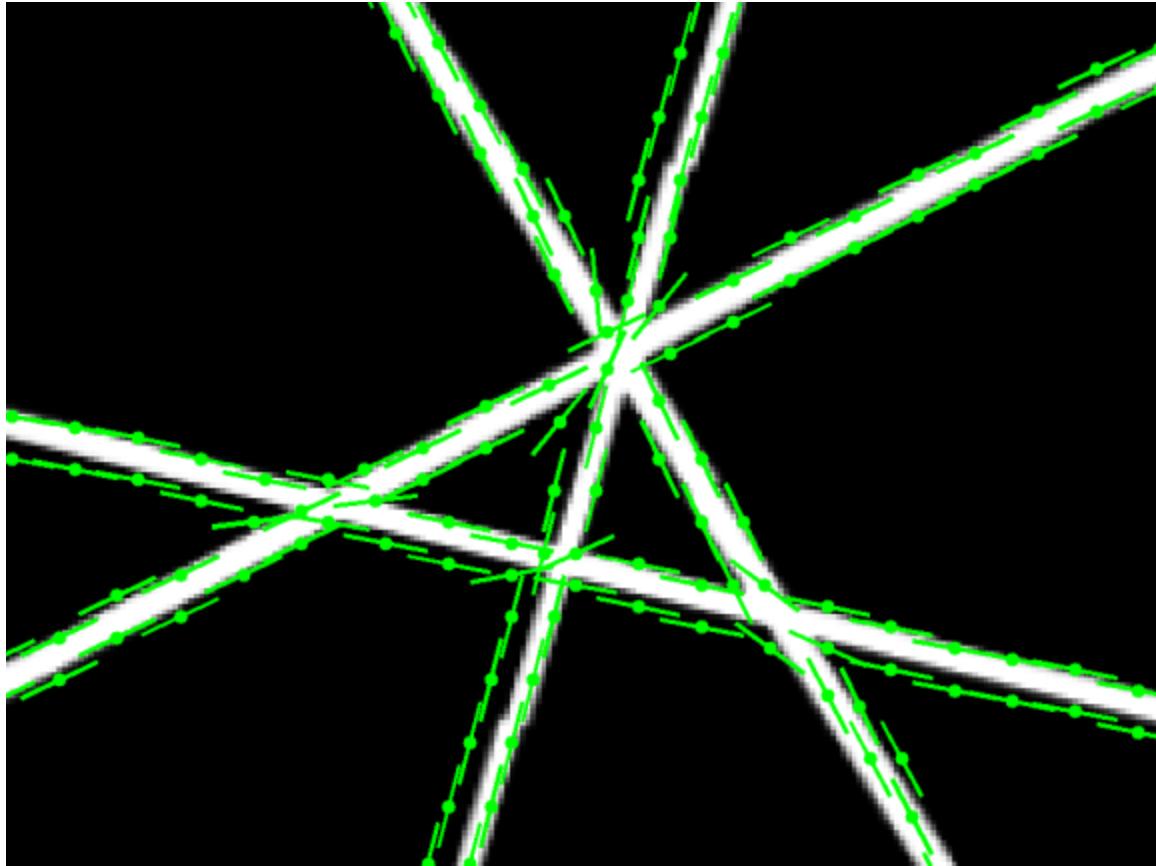


Figure 4. Largest 0.4% of the curvelets at the second finest scale produced by the Fast Discrete Curvelet Transform overlaid in green on a synthetic test image. Green lines indicate the orientation of each curvelet while the green circles indicate their center position.

We validated CurveAlign's performance against manual angle measurements in six real SHG images of excised tissue from a polyoma-middleT mouse breast cancer model [23] and against ground truth in a series of five simulated fiber images. Our results comparing CurveAlign to manual analysis in real images is shown in Table 2. Three examples of TACS-2 collagen architecture and three examples of TACS-3 collagen architecture were evaluated. Both CurveAlign and manual analysis produced angle distributions with similar means and standard deviations for all test cases. In addition, the mean angles for each produced the expected low relative angle for the TACS-2 cases and high relative angle for the TACS-3 cases.

		Ground Truth	CurveAlign
TACS-2A	M	34.1	32.9
	SD	24.1	22.1
	KS	$h = 0$	
TACS-2B	M	30.6	27.7
	SD	24.2	23.7
	KS	$h = 0$	
TACS-2C	M	24.8	22.0
	SD	16.2	12.2
	KS	$h = 0$	
TACS-3A	M	66.1	63.6
	SD	18.0	26.2
	KS	$h = 0$	
TACS-3B	M	68.1	68.3
	SD	15.8	16.4
	KS	$h = 0$	
TACS-3C	M	62.1	61.4
	SD	21.1	19.4
	KS	$h = 0$	

Table 2. Results of Kolmogorov-Smirnov testing between manual measurements and CurveAlign measurements on six images of mouse mammary tumors. This table presents the mean, standard deviation and Kolmogorov-Smirnov test results between three images of TACS-2(A-C) mouse mammary tumors and three images of TACS-3(A-C) mouse mammary tumors. The result $h = 0$ indicates that the measured and calculated angle distributions were similar based on a 95% confidence level.

The Kolmogorov-Smirnov (KS) test result reveals that to within a 95% confidence level, there were no statistically significant differences between the angles obtained via manual measurement and those generated by the CurveAlign program were drawn from the same distribution. The results of our validation against simulated test cases are shown in Table 3.

		Ground Truth	CurveAlign
Image 1	M	68.5	60.1
	SD	46.0	39.0
	KS	$h = 0$	
Image 2	M	93.9	92.3
	SD	51.5	54.8
	KS	$h = 0$	
Image 3	M	73.2	71.8
	SD	45.8	42.6
	KS	$h = 0$	
Image 4	M	59.7	68.1
	SD	44.9	43.1
	KS	$h = 0$	
Image 5	M	91.1	97.6
	SD	47.5	45.1
	KS	$h = 0$	

Table 3. Results validating the CurveAlign measurements of fiber angles in 5 simulated test images. M = mean angle in degrees, SD = standard deviation of the distribution of angles in degrees, KS = the result of the Kolmogorov-Smirnov statistical test validating the null hypothesis that the ground truth and measured distributions are statistically similar at the 95% confidence level.

Similar to the real SHG image test cases, the mean, standard deviation and KS test results indicate close agreement between CurveAlign measurements and ground truth. Furthermore, the CurveAlign analysis was significantly faster than the manual approach. The time required for manual analysis in ImageJ, or a similar program, can approach 1 hour or more per image depending on the number of features in the image and their complexity, while computational analysis using CurveAlign is on the order of seconds and provide consistent and accurate measurements that minimize cross-user variability.

With the results summarized in Tables 2 and 3, we have shown that the curvelet transform is able to generate accurate absolute angle distributions from SHG images and CurveAlign can use those angle distributions to accurately measure angle statistics relative to a boundary. Although the angle distribution can be a powerful tool for alignment measurements, spatial information is often lost. To help solve this issue, CurveAlign has been designed to indicate areas of local alignment and display these as an overlay on the original image. The usefulness of this feature is illustrated by the results shown in Figure 5.

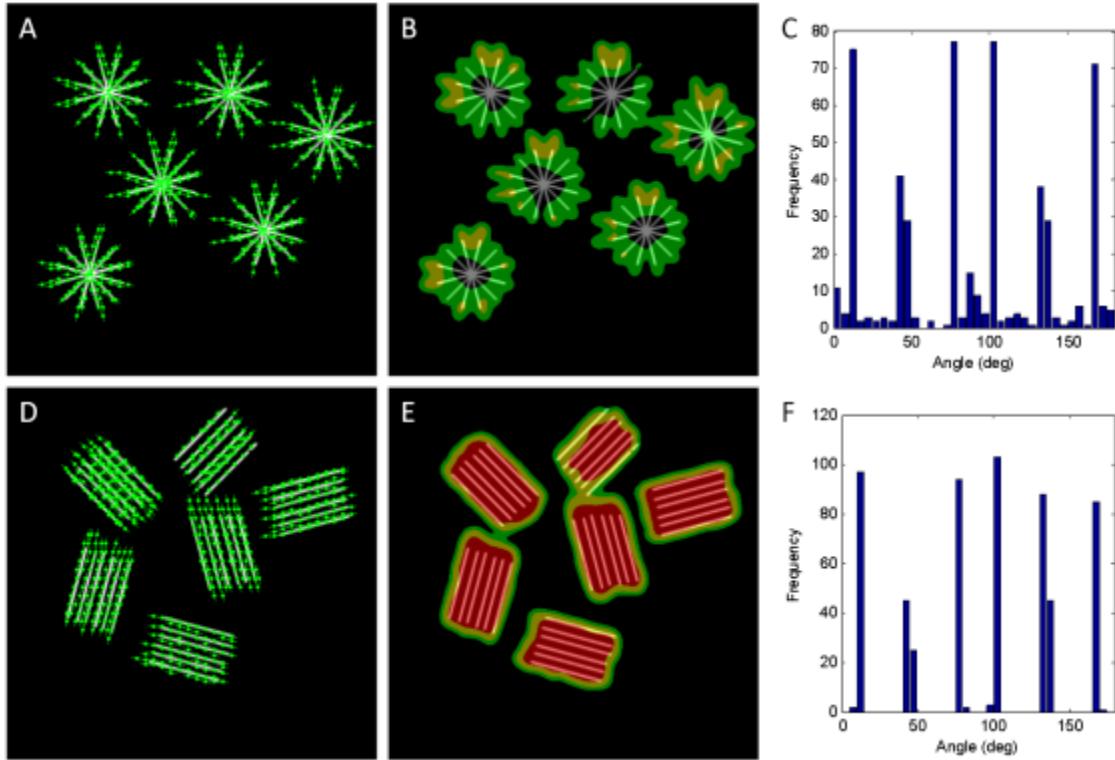


Figure 5. CurveAlign results from synthetic images of crisscrossing fibers (A-C) and parallel fibers (D-F). A,D: Individual curvelets overlaid on the original images. B,E: alignment maps overlaid on the original images where green indicates low probability of alignment, yellow intermediate alignment, and red high alignment. C,F: Histograms of the detected curvelet angles within each image. Although the histograms of these images are quite similar, the alignment maps show areas of localized alignment in only the parallel fiber image.

CurveAlign is used to analyze two synthetic images containing the same number of fibers at similar angles, but with different spatial distributions. For the first image (Figure 5A-C), there are fibers at 15, 45, 75, 105, 135, and 165 degrees configured in star patterns. For the second image, (Figure 5D-F), there are fibers at the same angles, but configured in aligned bundles. Indeed the histograms for these two images (Figure 5C and F) are very similar, showing peaks at the same locations. The local alignment maps shown in Figure 5B and E, on the other hand, are able to clearly differentiate between the two patterns.

One of the main functions of CurveAlign is the measurement of collagen fiber angle relative to a user-selected boundary. This is particularly useful for TACS-2 and TACS-3 measurements where collagen alignment relative to the tumor-stromal boundary is of primary interest. We illustrate some of the details of this function on an SHG image of a collagen gel shown in Figure 6.

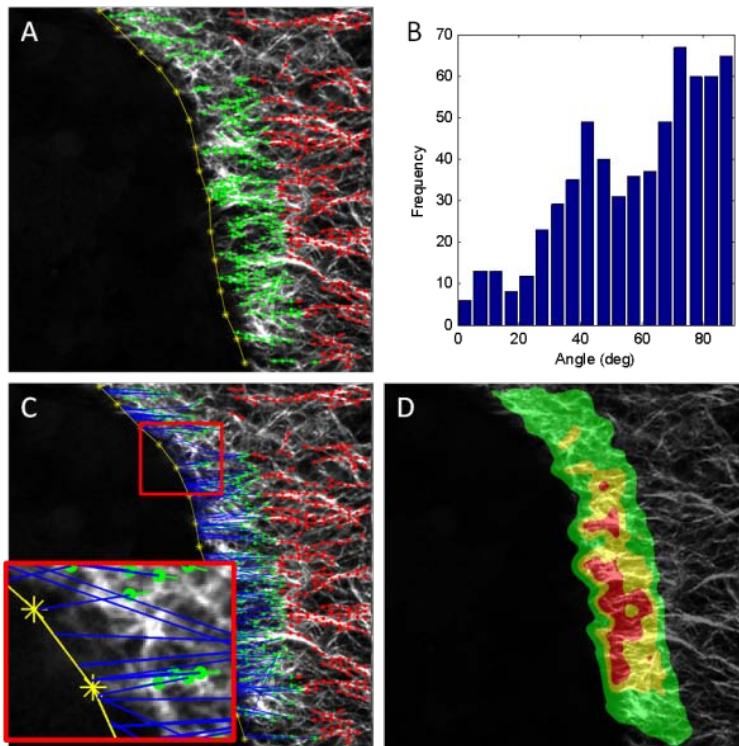


Figure 6. CurveAlign output examples. (A) 3D collagen gel image with boundary (yellow), analyzed curvelets (green) and discarded curvelets (red), overlaid on the original image. (B) Resulting histogram of fiber angles with respect to the boundary. (C inset) A zoomed version of A showing the boundary (yellow), curvelets (green), and the associations between the curvelets and the boundary (blue). (D) The alignment map overlaid on the original figure showing regions within the image where there are high (red), medium (yellow) and low (green) concentrations of perpendicularly aligned fibers.

The manually entered boundary is the yellow line in Figure 6A. CurveAlign allows the user to evaluate collagen alignment within a user-defined distance away from the boundary. In Figure 6A, the curvelets that are included in the analysis are shown in green, while those discarded from the analysis are shown in red. The relative angle distribution for the green curvelets is shown in Figure 6B. In order to measure

the relative angle between the curvelet and the boundary, a point on the boundary must be selected to compare with a given curvelet. The result of our association algorithm is illustrated in Figure 6C where blue lines connect each curvelet to the point on the boundary where the relative angle measure is made. The local alignment map shown in Figure 6D indicates where in the image there are high (red), medium (yellow), and low (green) density regions of fibers aligned perpendicular to the boundary. CurveAlign requires that the user make a selection as to what percent of the largest curvelet coefficients will be used in an analysis. The selection of this threshold can have an impact on the resulting measurements. However, as shown in Figure 7, there is often little effect on the shape of the resulting fiber angle distribution as the threshold is adjusted between 0.1% (Figure 7A), 0.25% (Figure 7B), to 1.0% (Figure 7C). The key difference between these trials is that the number of curvelets found in each bin is increased as the threshold increases.

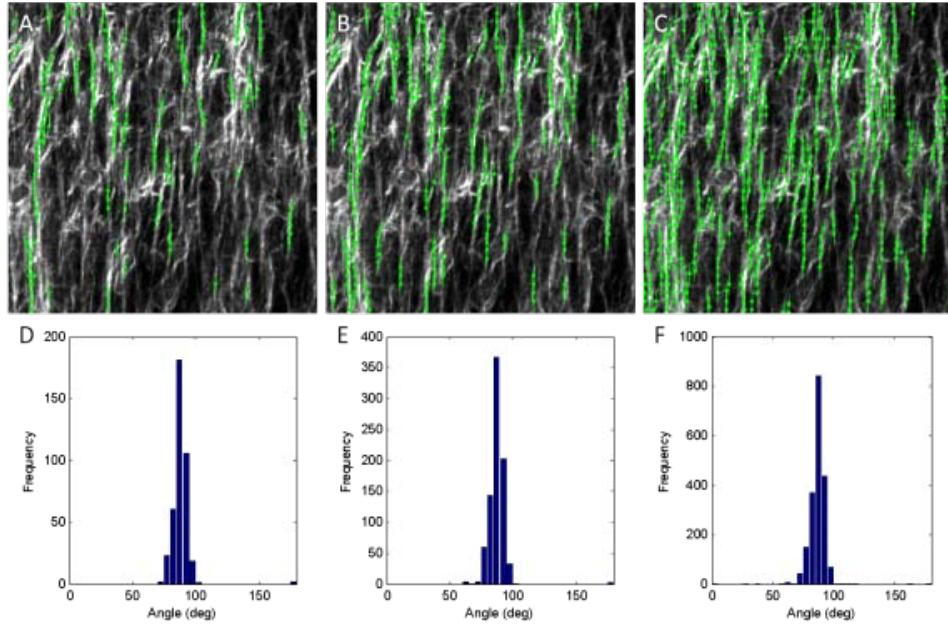


Figure 7. Influence of curvelet coefficient threshold on measurements. Curvelets overlaid on original image representing 0.1% (A), 0.25% (B), and 1.0% (C) of the top curvelet coefficients with corresponding angle distributions (D-E respectively). The coefficient of alignment was 0.972, 0.976, and 0.973 for case A-C respectively.

This program is now actively being used in our own breast cancer research and also in the labs of several collaborators who are investigating collagen alignment in cancer[149, 150], cervical tissue[151], and cardiac disease[152], as well as other filamentous structures such as microtubules[153]. For example, in breast cancer invasion and progression studies CurveAlign was used to quantitatively assess the effect of cancer cells on the remodeling of collagen during cancer onset and progression. A xenograft of MDA-231 breast cancer cells were implanted after 3 weeks in an adult mouse (using methods similar to those described in [27]) and CurveAlign was used to quantify the difference in alignment at the graft boundary. CurveAlign has also been used to analyzed collage fibers surrounding normal murine mammary gland (NMuMG) cells cultured in a high-density 3D collagen matrix[154] where regions along the leading edge of an invading cell cluster were abutted against collagen fibers distributed more closely around 90 degrees when compared to the alignment of fibers surrounding the distal and lateral edges of the cell cluster. In addition, CurveAlign was recently used in cardiac research to quantitatively assess the remodeling of the TissueMend[®] (TEI Biosciences, Boston, MA, USA) matrix by mesenchymal stem cells (MSCs) after being applied to mouse myocardium *in-vivo*[152]. The combined approach of SHG and curvelet analysis showed a significant decrease in the coefficient of alignment of fibers in the presence of MSCs compared with the matrix alone, and the MSCs were capable of migrating out of the matrix and into the host tissue[152]. Finally, CurveAlign has recently been used to quantify the alignment of collagen in excised human breast tissue as illustrated in Figure 8.

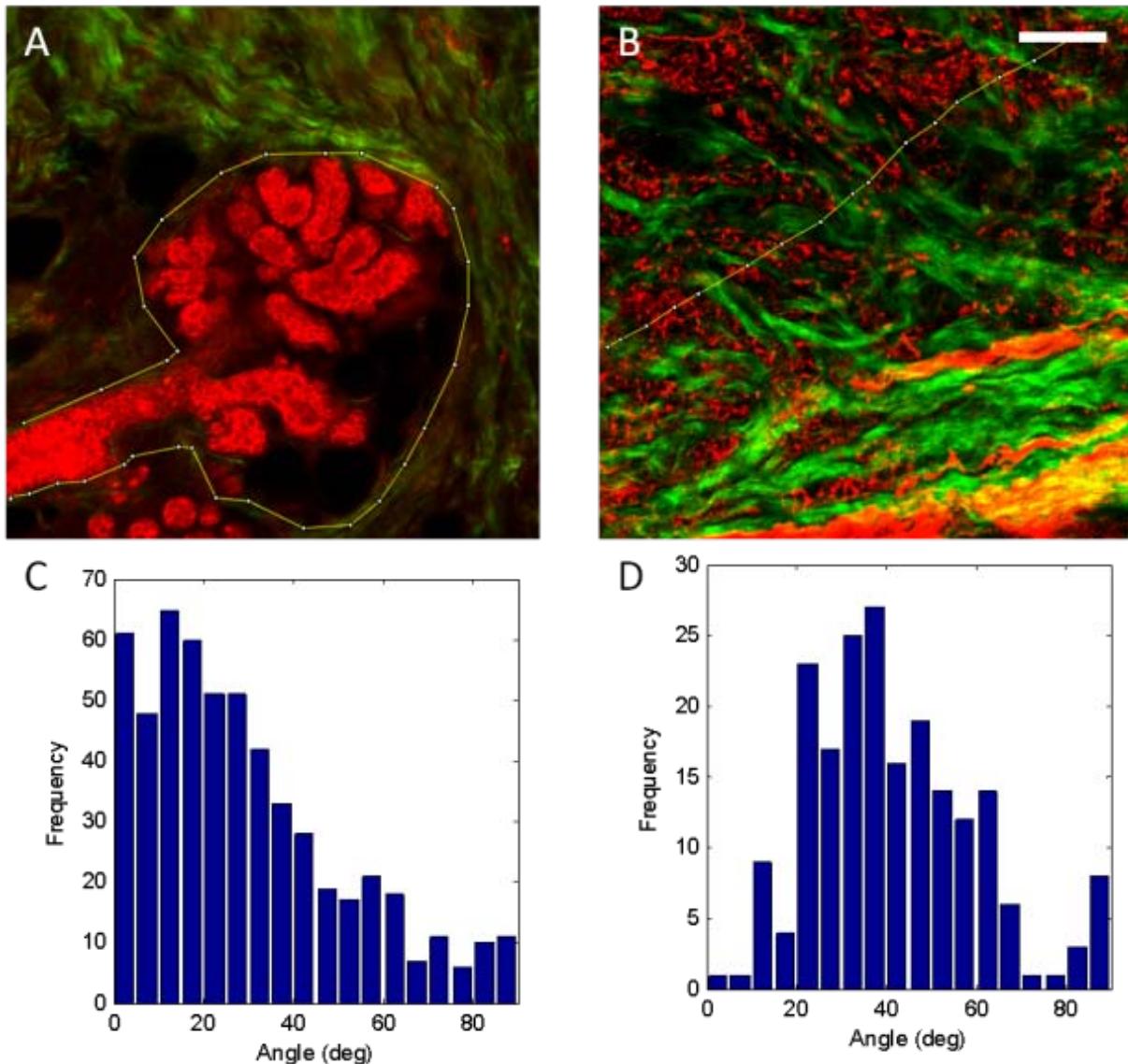


Figure 8. Collagen alignment measurements made in thick excised human breast cancer tissue samples. Fiber angle distributions relative to the epithelial stromal boundary for an adjacent normal terminal lobular end-bud unit are shown in (A,C) while the collagen alignment relative to the invasion front of a group of invasive ductal carcinoma cells are shown in (B,D). In the invasive case, the angle distribution is weighted more heavily toward perpendicular angles. E-Cadherin stained epithelial cells are shown in red while collagen fibers are shown in green. Scale bar = 50 μ m.

Whole mount E-cadherin epithelial cell staining was performed on 300 micron thick tissue sections followed by MPLSM and SHG imaging. Epithelial cells in Figure 8A and B are pseudocolored red while the collagen fibers are green. A boundary was drawn around the normal terminal lobular end unit is shown in Figure 8A and a boundary is drawn perpendicular to the invasion front of the invasive cancer cells in

Figure 8B. The angle distributions in Figure 8C and D illustrate that fiber angles in the invasive case are more heavily weighted toward higher angles compared to the collagen angle distribution surrounding normal epithelial cells.

The rapid increase in applications for CurveAlign is due to the growing interest in collagen alignment quantification by our group and others. Along with these new applications comes the need for new CurveAlign functionality. While the MATLAB library does provide great flexibility, we realize the value of making CurveAlign more widely available to biologists that either don't utilize MATLAB or require the functionality of an image visualization package. We plan on making CurveAlign available as an ImageJ plugin as part of our ongoing ImageJ development (www.imagejdev.org). This CurveAlign plugin would still utilize the CurveLab toolbox, but would do so without relying on MATLAB. We also plan on utilizing our Bio-Formats library[155] (www.loci.wisc.edu/bioformats) to allow CurveAlign to open additional image formats. CurveAlign can currently open standard, single channel TIFF files but currently cannot open the proprietary, often multichannel datasets used on many microscopes. Bio-Formats will enable that functionality and provide the appropriate dimensionality to CurveAlign. Additionally, we plan on interfacing CurveAlign to machine learning approaches such as Cell Profiler (www.cellprofiler.org) to allow for automatic detection of dense collagen areas for subsequent curvelet analysis.

Advances in the field of SHG imaging of collagen drive further opportunities for curvelet-based alignment analysis. SHG signal propagates in both the forwards and backwards directions and can be simultaneously collected with the appropriate detection hardware [57]. The ratio of forwards and backwards directions has been shown to not only uniquely characterize and discriminate collagen types[156] but also characterize remodeling, such as that occurring in cancer invasion and progression. There is also strong evidence for the importance of polarization in collagen alignment. Future studies using curvelet-based alignment analysis may find it useful to be able to take into account these added

dimensions. More challenging is the growing need to look at changes in collagen alignment that happen in three dimensions. Currently, the vast majority of studies have looked at collagen alignment in two dimensions. When researchers refer to three dimensional analysis, they often only quantify changes in collagen in individual z planes. However, to our knowledge, there is not yet any conclusive evidence of specific collagen fiber alignment changes along the z dimension. There is a 3D version of the curvelet transform[157] which we have not investigated to date not only due to the lack of compelling collagen alignment datasets in 3D but also due to the substantial increase in computational resources required to execute collagen alignment analysis in 3D. As future studies begin to demonstrate changes in collagen alignment in 3D, there will be a need for tools that can measure collagen fiber changes occurring between focal planes.

Another challenge for the CurveAlign program is sensing the difference between randomly oriented, short, straight fibers and long curvy fibers. The angle distributions produced by CurveAlign for these two scenarios would likely be very similar. The differentiation between these structures would be useful in at least one published study of ovarian cancer[115] which showed that normal ovarian tissue contains more straightened aligned collagen fibers, while collagen in diseased tissue is more curvy. The tissue in an intermediate “at risk” state has random collagen with a complete lack of organization. This randomness and presence of multiple alignment types intersecting with each other make quantification difficult. The fiber tracking techniques described in Chapter 3 hold promise for solving problems such as these. Given the rising importance of collagen architecture as a biomarker, and the significance of ovarian cancer, strategies for analyzing these complex changes in collagen organization will be useful for future studies.

2.4 Conclusion

Here we have presented the utility of the curvelet transform using a semi-automated computational tool for the quantification of collagen fiber alignment, both relative to a tumor boundary and in tissues or gels lacking a defined boundary structure. By taking advantage of several key features of the curvelet transform, we are able to determine the orientation and alignment of collagen fibers in an SHG image without any pre-processing steps. We have shown that the curvelet transform, and our practical software implementation of this, CurveAlign, is effective at measuring and displaying changes in collagen alignment in an array of collagen samples from collagen gels to mouse mammary tissue. This tool is available for download at: <http://loci.wisc.edu/software/curvelet-based-alignment-analysis>.

Chapter 3 Collagen Fiber Extraction

3.1 Introduction

The extracellular collagen matrix (ECM) has been found to promote the progression of many types of cancer. However, the underlying mechanism behind this relationship is not fully understood, and is the subject of intense biomedical research. Much of this research has benefited from the capabilities of laser scanning microscopy techniques, in particular second harmonic generation (SHG) imaging [55], to capture high-resolution, high-contrast images of individual collagen fibers in tissue and *in-vitro* tissue models [13, 24, 30, 32, 158]. For example, Conklin et. al. [1] showed that patterns in SHG images of collagen can predict breast cancer patient outcome. Raub et. al. [159] showed that SHG image characteristics can be used to predict bulk mechanical properties of collagen hydrogels, a common *in-vitro* tissue model for studying cancer cell motility. Nadiarnykh et. al. [160] and Watson et. al. [161] found that SHG image characteristics in ovarian tissue provide quantitative discrimination between tumor and benign tissue. Although SHG has been used successfully in these and many other studies, quantification of collagen fiber shape changes remains a difficult challenge, in part due to large heterogeneities in the patterns observed in SHG images of tissue. For example, in the breast tissue images shown in Figure 9, collagen fibers can be described as wavy or straight (Figure 9A and B), high or low density (Figure 9C and D), thick bundles or thin strands (Figure 9E and F).

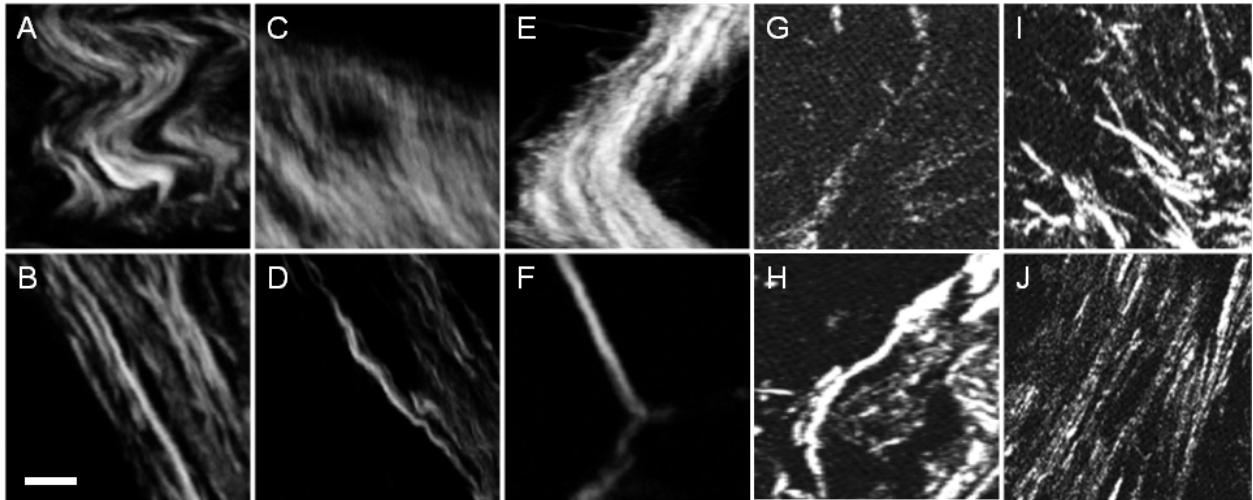


Figure 9. Representative collagen patterns observed in human breast cancer tissue sections demonstrating the heterogeneous nature of collagen structure. Wavy (A) and straight (B). High (C) and low density (D). Thick bundles (E) and thin strands (F). Discontinuous (G) and continuous (H). Crossing (I) and parallel (J). Scale bar is 10 microns.

These descriptions are consistent with previously published observations of common collagen structures in tissues [1, 162, 163]. In addition, depending on imaging parameters such as depth within the tissue, images can have low signal to noise (SNR) and potentially low dynamic range (Figure 9G, H, I, and J). Quantitative analysis techniques for SHG images of collagen need to provide robust and informative features within this heterogeneous collection of patterns and image qualities. Also, in order to elucidate interactions between cells and individual collagen fibers, effective quantitative analysis techniques should be able to extract information about individual fibers, such as fiber number, length, angle and curvature. The work reported here is motivated by these two requirements, the need for robust performance and the need for fiber-level information in SHG image analysis of collagen.

While these two requirements may be met by manual analysis [23], inter-observer and intra-observer variance can be significant and time requirements can often approach tens of minutes per image. Computer assisted image feature extraction is poised to help meet these requirements in an automated fashion. Transform or filter based methods have been used for SHG collagen analysis such as the Fourier transform method published by Falzon [164], the combined Fourier and Hough transform

method by Bayan [76], the Curvelet transform method by Pehlke [81], the Fourier and fractal based method reported by Frisch [165], the directional gradient technique suggested by Altendorf [166] , and the grey level co-occurrence method published by Hu [167]. These techniques can be highly robust, often able to detect important features in a diversity of image settings. However, since these techniques do not extract individual fibers, they lack the ability to identify fiber-level information. For example, transform based methods can provide general information about fiber size and direction at each point in an image, but cannot determine the actual fiber number, nor the length or curvature of the fibers. Since pixels are not grouped into individual fibers, these methods may not sense the difference between many short, randomly-oriented, straight fibers and long curvy fibers, features that may help to classify patients into high and low risk groups for ovarian [160] and other cancers. In addition, angle distributions generated by these algorithms would generally produce bias toward longer and potentially thicker or brighter fibers, since more pixels are present in longer and thicker fibers and distributions are based on pixels and not fibers. Avoidance of bias such as this is necessary to accurately pursue biological hypotheses that are based on high resolution microscopy image data, for example, the hypothesis that fiber angle distribution may help predict metastatic potential of cancer cells [1, 2, 30, 32].

Fortunately, fiber tracking and extraction methods, such as those published by Wu [40, 78] and Stein[45], have been developed to extract fiber-level information from images of *in-vitro* collagen matrices. These methods can enable the automated measurement of important fiber-level parameters, such as fiber length, number and curvature, and have been used to estimate collagen gel mechanical properties based on confocal images of stained gels. However, they have not been applied to SHG images of collagen *in-situ*. While these approaches are powerful, perhaps they have not been applied *in-situ* because they often fail to properly segment fibers in the dense or low signal-to-noise (SNR) situations commonly encountered in SHG images of tissue. Examples of two SHG images are shown in Figure 10A and D with corresponding manual fiber extractions shown in Figure 10B and E. The Fiber

Extraction (FIRE) algorithm, developed and made available by Stein [45], produces the overly complex fiber network shown in Figure 10C, an erroneous star pattern in Figure 10F and in both cases fails to identify many of the fibers extracted by the human observer.

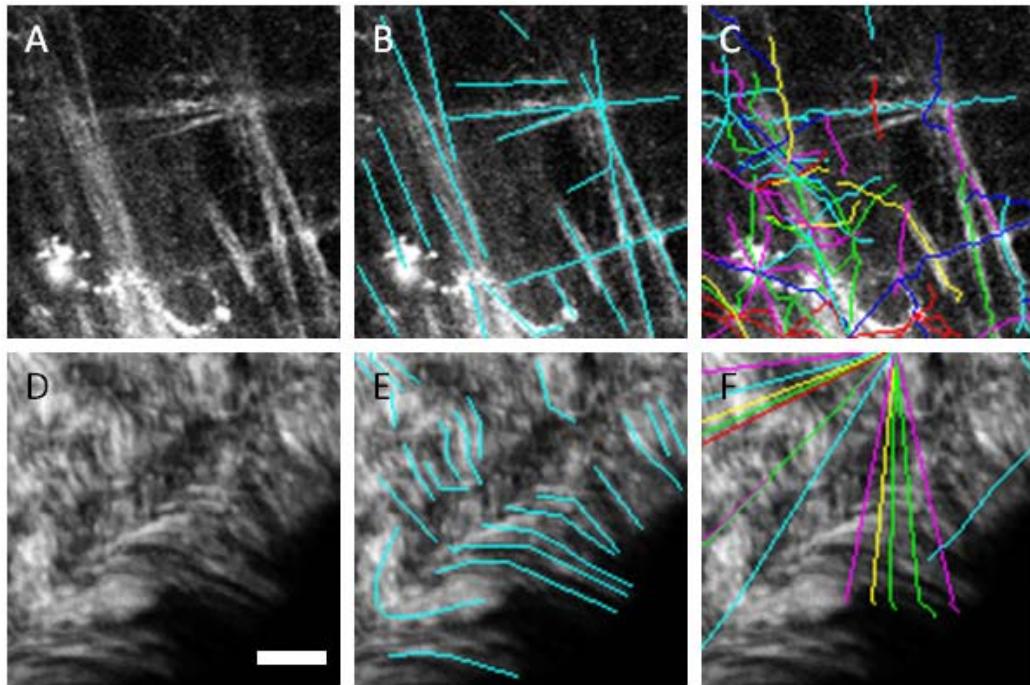


Figure 10. Fibers extracted by the FIRE algorithm alone without preprocessing. A and D are the original images, B and E show manual segmentations of the fibers, C and F show the automatic fiber segmentations that are extracted by the FIRE algorithm and show many falsely segmented fibers. Scale bar is 25 microns.

Instead of using transform based methods or fiber extraction methods alone, a more strategic approach would be to combine complimentary methods and use transform/filter based methods as a preprocessing step to fiber extraction. This combined approach has potential for robust performance in a wide range of challenging imaging situations commonly seen in cancer imaging while simultaneously allowing for fiber-level information to be extracted from images. In this study, we present an approach for integrating transform/filter based preprocessing techniques with fiber extraction and evaluate the performance of our combined approach with the ultimate goal of improving the fiber extraction accuracy of an algorithm such as FIRE. Our hypothesis is that the application of appropriate pre and post

image processing algorithms can significantly improve fiber extraction in SHG images. This will enable, for example, more accurate fiber angle distributions, thus allowing for increased sensitivity to detect collagen alignment changes related to cancer progression.

We evaluate four candidate preprocessing techniques including the simple Gaussian filter, the SPIRAL-TV filter [168], the tubeness filter [169], and a curvelet transform based denoising filter [170, 171]. Other than the Gaussian filter, these filters were chosen based on their published ability to highlight edge information in images while simultaneously suppressing spatially uniform structures and noise. We have chosen to use the FIRE algorithm based on evidence of its ability to extract fibers from *in-vitro* collagen gel networks and its availability [172], but other fiber extraction tools may be substituted for the FIRE algorithm. We focus our analysis on 2D SHG images because the effective nonlinear susceptibility declines sharply when fibers are tipped out of the imaging plane [70, 173, 174]. However, our methods may naturally extend to 3D without significant alteration.

3.2 Materials and Methods

Our experimental approach for the evaluation of quantitative collagen fiber extraction is illustrated in Figure 11. Twenty-five images were annotated by three human observers who traced each fiber within each image to create a surrogate for ground truth. The same images were then filtered with one of four image filters, which are described in the following sections. Fiber tracking and extraction was then performed using the FIRE algorithm. We evaluated the automated fiber extraction accuracy by comparing the length, angle and position of each fiber extracted by the FIRE algorithm with each fiber that was manually extracted to determine if fibers can be considered detected, missed, or falsely detected. F-measure scores were created based on these rates and a single parameter in each image filter was adjusted to optimize the F-measure score for each algorithm. The details of each step in this process are given in the following sections. Following algorithm evaluation, we applied the best

performing algorithm to quantify collagen fiber shape changes during tumor progression in an *in-vivo* mouse model of breast cancer.

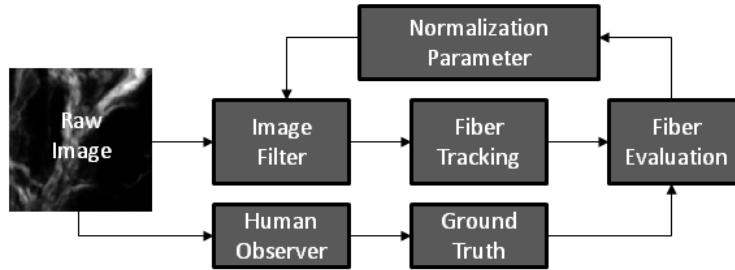


Figure 11. Diagram of the approach for quantitative collagen analysis showing the iterative process for optimizing the performance of a single image-processing filter for fiber tracking. The raw image is processed by the image filter using an initial normalization parameter, the result of which is sent to the FIRE fiber-tracking algorithm. Automated fiber extractions are compared against manually performed fiber extractions. Several normalization parameters are evaluated and one optimal parameter is selected for each filter based on the fiber evaluation result.

3.2.1 Sample Preparation

To evaluate algorithm accuracy, we chose to use images of both human and mouse tissues as they are both routinely used by our group and others for studying stromal interactions during breast cancer development. Human ductal carcinoma *in-situ* biopsy samples were obtained from two completely de-identified patients, paraffin processed, sectioned to 5 microns, hematoxylin and eosin (H&E) stained and coverslipped. In addition, tumor bearing mouse mammary glands from 8 week old MMTV-polyoma middle-T [23] mice were harvested and imaged fresh, intact, and unstained. All human tissue images were collected following Institutional Review Board approval at the University of Wisconsin at Madison. Mouse tissue images were obtained using protocols approved by the University of Wisconsin at Madison Institutional Animal Use and Care Committee. We demonstrate an application of our method by measuring fiber curvature in an *in-vivo* tumor model. *In-vivo* images were captured through a glass intravital imaging window that was surgically placed immediately superficial to palpable tumors within the mammary gland [175] in live 8 week old MMTV-polyoma middle-T (PyMT) mice. Animals were anesthetized while imaging was performed at 8 weeks and 12 weeks of age.

3.2.2 SHG Imaging

SHG images were captured with an excitation wavelength of 890 nm, a pulse length of approximately 100 fs, and an emission filter centered at 445 nm with a 20 nm bandwidth (Semrock, Rochester, NY). Excitation light was focused onto the sample using a 10X (N.A. 0.5) objective. Pixel size was approximately 0.75 microns and pixel dwell time was approximately 10 microseconds per pixel. Average laser power at the sample was adjusted to approximately 15 mW for slide imaging and 30 mW for intact and *in-vivo* tissue imaging using a Pockel's cell and a polarizer. Forward SHG was used to image slides and backward SHG was used to image intact and *in-vivo* mouse tissue. Emission light was detected with a 7422-40P (Hamamatsu, Japan) photomultiplier tube in both forward and backward SHG cases. All SHG images of collagen were captured in regions adjacent to mammary ductal epithelium verified by white light images of H&E in the slides and by cellular autofluorescence from FAD in the intact tissue and *in-vivo* cases [23]. Single images were captured of the slides and z-series were captured for the intact tissue and *in-vivo* imaging experiments. For intact tissue imaging, representative images were selected from each z-series for quantitative analysis. For *in-vivo* imaging, z-series at 3 imaging locations were captured at each time point. Three images at depths of approximately 5, 10 and 15 microns below the imaging window were selected for analysis at each location for a total of 9 images per time point.

3.2.3 FIRE Algorithm

We briefly review the FIRE process here. A more detailed description of the algorithm can be found in Chapter 4 and reference [45]. FIRE is an automated tracking method that can extract the geometric structure of three dimensional collagen images and is capable of generating information about the number, length and curvature of the collagen fibers in an image. The first step of this method is to apply a threshold to form a binary image such that foreground pixels represent potential fibers and background pixels represent pixels where no fiber is present. Next, the distance transform on the binary image is performed to yield the distance from each foreground pixel to the nearest background pixel. Then the maximal ridges of the smoothed image formed by the distance transform are searched to

create a list of nucleation points. Branches are formed by extending the fiber from each nucleation point based on fiber trajectory. Short fiber branches are then pruned and closely associated fibers are finally linked based on the fiber length, fiber direction and the distance between adjacent fibers. In the associated FIRE software [172], there are about 20 adjustable parameters initialized with default values. There are usually only a few parameters that need to be adjusted, such as those impacting the binary image generation, the search for nucleation points and fiber linkage. To our knowledge, the FIRE method has only been tested on confocal reflectance and confocal fluorescence images of *in-vitro* collagen gels, but has not been applied to extract collagen fibers from SHG images of tissue.

Each preprocessing technique described in this chapter was followed by nearly identical implementations of the FIRE algorithm. The only difference is in the threshold used for creating the initial binary image. This threshold was hand optimized to produce the highest quality fiber extractions across all test cases for each algorithm. Our criteria for quality was the F-measure score which is described below.

3.2.4 Preprocessing Algorithms

The four preprocessing algorithms evaluated here are described briefly below. More detailed background information on the advanced filters can be found in their respective references.

3.2.4.1 Gaussian filter

A simple 2-D Gaussian filter (GF) was used as a baseline for comparison against the other more advanced filters. The standard deviation of the simple GF was optimized to produce fiber extractions that most closely matched the human observers using the iterative approach diagrammed in Figure 11.

3.2.4.2 SPIRAL-TV filter

The SPIRAL-TV (SPTV) algorithm, by Harmany et. al. [168] was developed to accurately extract features from images where Poisson noise dominates, a common occurrence in SHG imaging of collagen in tissue

or *in-vitro* collagen gels due to the low signal levels often encountered in such imaging experiments [176]. This algorithm has applications in compressed sensing, nuclear medicine tomographic reconstruction and super-resolution reconstruction in astronomy. The algorithm iteratively approximates a solution to the constrained optimization problem given by

$$f^{k+1} = \arg \min_{f \in \mathbb{R}^n} F^k(f) + \tau * pen(f)$$

subject to $f \geq 0$

where f is the approximation to the image of interest, $F^k(f)$ is the negative Poisson log-likelihood function at iteration k , and $pen(f) = \|f\|_{TV}$ is the total variation seminorm penalty scheme [177]. The scalar parameter τ was optimized to produce the highest F-measure score when comparing human and automated fiber extractions. SPTV was shown to perform well at highlighting strong edges in images and smooth noise in low gradient areas [168]. The designers of this algorithm have tested it on noisy computed tomography reconstruction data, but it has not been heretofore applied to preprocessing for fiber extraction from SHG images.

3.2.4.3 Tubeness filter

The tubeness filter (TF) is an ImageJ plugin implemented by Longair, Preibisch and Schindelin [178] and is based on the work published by Sato et. al. [169]. The algorithm highlights fiber-like structures in images while attenuating homogeneous or noisy regions and has found application in processing images of neurons and blood vessels [169, 179]. This filter was used to enhance fiber structures by first applying a 2-D Gaussian filter with the standard deviation optimized to produce the highest F-measure score. Next, the Hessian is computed at each point in the image and the eigenvalues, λ_1 and λ_2 for the 2-D case, of the Hessian matrix are found. The resulting pixel value is given by the following rule:

$$\lambda_g = \begin{cases} |\lambda_2|, & \lambda_2 < 0 \\ 0, & otherwise \end{cases}$$

To our knowledge, this filter has not been evaluated for its ability to highlight collagen fibers in SHG images of tissue.

3.2.4.4 Curvelet filter

We have also implemented a denoising filter based on the 2-D curvelet transform (CT). The CT was developed by Candes and Donoho [180] to overcome the missing ability of the conventional wavelet transform to highlight lines and edges. Our group has recently reported on the successful use of the CT for finding fiber alignment information in SHG images of collagen [81]. Here we report on the use of the CT as a preprocessing step to fiber extraction. Briefly, the CT represents images as superpositions of elements that are constant along ridgelines and wavelets in the orthogonal direction. Curvelet lengths and widths vary with scale and obey the rule $\text{width} \approx \text{length}^2$. Simple curvelet coefficient thresholding has been shown to be an improvement over advanced denoising techniques based on wavelets such as decimated or undecimated wavelet transforms [170]. Our denoising implementation uses the frequency wrapping version of the Fast Discrete Curvelet Transform [171] and reconstructs images using the top $x\%$ of the curvelet coefficients from the intermediate scales 4, 5, and 6 out of 7 total scales in our test cases. The parameter x was optimized to produce the best overall results as indicated in the block diagram in Figure 11. Scale selection may vary with different applications. We chose to remove the finest scale (7th scale) due to the high noise content present at this scale. The coarser scales (scales 1-3) did not represent the size of the fibers in our images and were therefore discarded.

3.2.5 Algorithm integration and evaluation

As shown in Figure 11, in order to yield optimal fiber evaluation results, each filter was optimized in an iterative manner to achieve the highest F-measure score. The FIRE parameters could have also been iteratively optimized in this fashion. However, we decided to hand optimize the FIRE parameters based on a visual evaluation of the results from a small training set of 3 unprocessed images. We then kept the FIRE parameters fixed for each of the preprocessing algorithms, except for the initial threshold that

separates fiber pixels from background pixels. This threshold was adjusted for each algorithm, again by visual inspection on a small training set of 3 images. Then the threshold was fixed across images within a preprocessing group. The method for evaluating the fiber segmentation was as follows: three human observers were asked to manually segment all fibers in each of the test images into regions of interest (ROI). The images were annotated using the ImageJ ROI Manager. The ROIs for each of the test cases were saved for each of the 3 observers. These ROIs were then read into MATLAB using the Miji toolbox [181]. The fibers extracted by each automated algorithm were then compared with the manually extracted fibers for each test case and each observer. Fiber angle agreement, fiber length agreement, and distance between manual and automatically extracted fibers were used to score the accuracy of the automated segmentation. The average angle of a fiber was computed by finding the absolute angle of the line connecting the end points of the fiber. Fiber length was computed as the euclidean distance traveled along the fiber. Distance between manual and automatically extracted fibers was computed using a k-nearest neighbor search algorithm [145]. Consider a set M of i manually segmented fibers each with j points and a set A of n automatically segmented fibers each with m points. The function $KNN(M_{i,j}, A_{n,m})$ produces $D_{i,j,n}$ where D is the Euclidean distance from point j on fiber i of set M to the nearest neighbor point on the n th fiber of set A . The metric for the distance between manually segmented fiber i and automatically segmented fiber n is then $C_{i,n} = \sum_j D_{i,j,n}/L_i$, where L_i is the distance along the path of the i th manual fiber. A manually segmented fiber was associated with an automated fiber, and vice versa, if the two had similar average angles, similar lengths, and similar positions. The number of true positive fibers (TP), false positive (FP), and false negative fibers (FN) were then found by counting the number of associated manual fibers, unassociated automated fibers, and unassociated manual fibers respectively for each test case. Precision (also called positive predictive value) and recall (also called sensitivity, hit rate, or true positive rate) were computed as $Precision =$

$TP/(TP + FP)$ and $Recall = TP/(TP + FN)$, and the harmonic sum of the two was computed as follows

$$Fmeasure = 2 * \frac{Precision * Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN}$$

The $Fmeasure$ result for each of the preprocessing algorithms was averaged over all test cases for a given observer, producing $Fmeasure_n$, where n represents observer number. Then, the $Fmeasure_n$ result was averaged over all observers and the standard deviation between observers was computed.

3.3 Results

Comparing the four image-processing techniques to each other, as shown in row 1 of Figure 12, reveals that edge-preserving filters such as SPTV, although effective for denoising without loss of edge information, do not lend themselves well to improving the fiber tracking results. On the other hand, the TF and CT create ridges along fiber centers (Fig. 4, row 1), helping to ease the difficulty of threshold selection and helping the fiber tracking algorithm to follow the centers of thick or noisy fibers.

Examination of fiber tracking results in Figure 12 row 2, shows many completely erroneous fiber tracks for the unprocessed, GF and SPTV filtered cases (red arrows), whereas the TF and CT filtered results show several properly segmented fibers (green arrows). Each of the images in Figure 12 are representative 128 by 128 pixel regions cropped out of larger images.

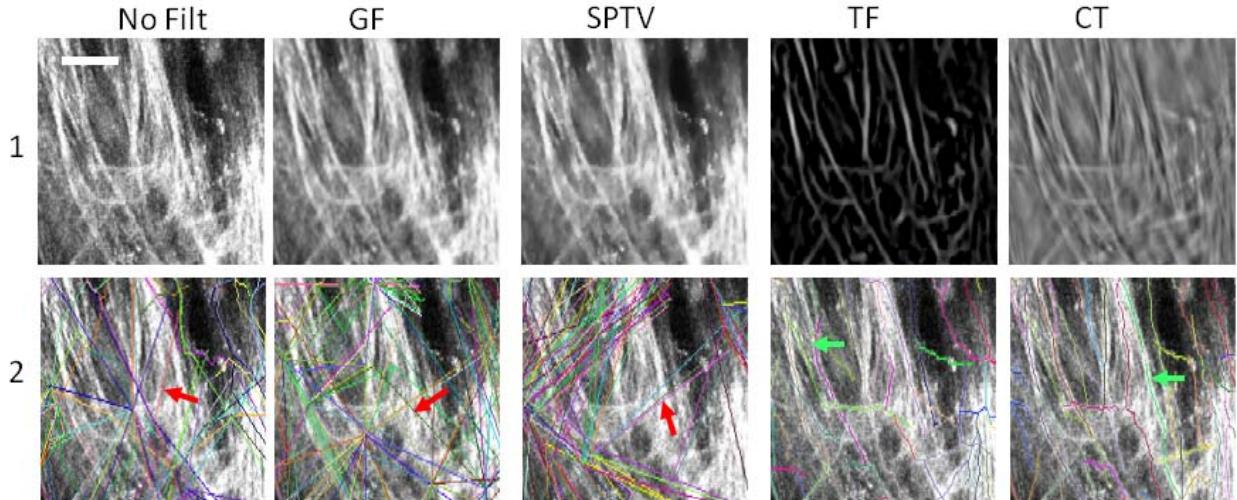


Figure 12. Output of the image processing techniques (row 1) and output of the fiber tracking algorithm (row 2) for a single test case. The first column is without a filter, column 2: GF, column 3: SPTV filter, column 4: TF, and column 5: CT. Scale bar is 25 microns.

To further test the effect of preprocessing on collagen fiber segmentation, the manual segmentation results were compared to the automated fiber extraction results for two representative test cases, shown in Figure 13. Each row in the figure is a different test case, while each column represents a different method of fiber segmentation. Column 1 shows the original images with no overlaid segmentations. Columns 2 through 6 show the original image with overlays of the manual, GF, SPTV filter, TF, and CT filter segmentations respectively, where FIRE was performed following each of the filter preprocessing steps. Although we had three observers manually segment each of the test cases, the manual segmentations shown in column 2 represent the segmentations of a single observer. Each tile in Figure 13 is a 128 by 128 pixel crop of a larger image. The quality of extraction of individual fibers under the various processing conditions was compared (arrows in Figure 13A and B) where it was found that improperly segmented fibers as a result of GF or SPTV filtering were accurately segmented using the TF or CT filters.

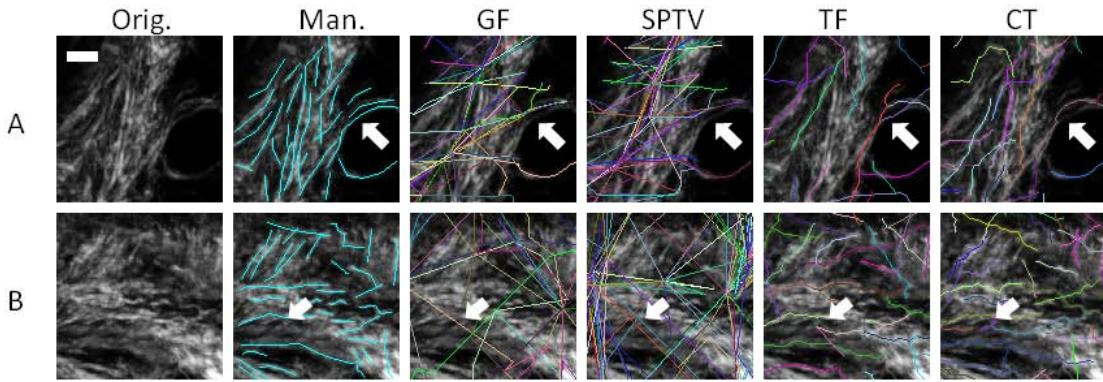


Figure 13. Two test cases (A&B), showing different processing methods in each column. The original image (column 1) is shown overlaid with a manual segmentation (column 2), GF (column 3), SPTV (column 4), TF (column 5), and CT filter (column 6) where each filter is followed by FIRE fiber extraction. Scale bar is 25 microns.

In order to assess the accuracy of collagen fiber segmentation, the results of each of the fiber extraction preprocessing algorithms were compared against each of the three segmentations performed by the independent observers using a collection of custom scripts written in MATLAB (MathWorks, Natick, MA). If a fiber segmented by the automated process had a similar angle, close proximity, and similar length to a manually segmented fiber, then an association was made between the automated and manual fibers, indicating a true positive. The thresholds for angle, proximity and length were chosen to optimize the sum of the F-measure scores for all algorithms. After all fibers were evaluated, all remaining unassociated manual fibers were counted as false negatives (misses) and all remaining unassociated automatic fibers were counted as false positives (false hits). Precision, recall and their harmonic sum (F-measure) were computed and compiled for all test cases and all observers. Overall average F-measure scores for each of the preprocessing algorithms are shown in Figure 14. The average F-measure score for the CT filter was the highest followed by the TF, SPTV, and GF. The error bars indicate the standard deviation between the F-measure scores from each of the 3 observers and show that the scores between observers were very similar, meaning that the CT filter result was the closest match to all 3 observers. Comparing the computation time for each algorithm on identical images resulted in respective times of 0.22, 225.05, 0.83, and 4.65 seconds for the Gaussian, SpiralTV,

Tubeness, and CT processing. Focusing on the CT denoising filter, we further validate its performance combined with FIRE using a collection of computationally generated images of collagen fibers that were designed to mimic the length and curvature characteristics found in collagen gels of approximately 1.0 mg/mL[45]. We processed these images with the CT-FIRE algorithm to extract length and angle information about the individual fibers. The results of this test are shown in Figure 15 and show that the CT-FIRE algorithm produces accurate length and angle distribution measures in known synthetic test cases.

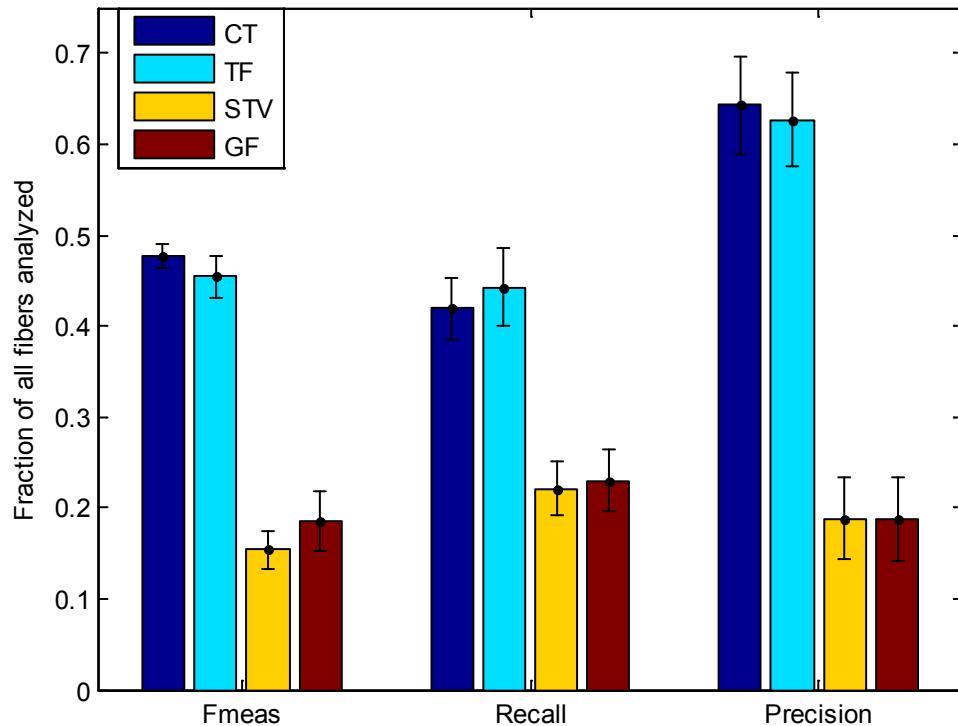


Figure 14. F-measure, recall, and precision results comparing the automated segmentation techniques to the manual segmentations of three independent raters, for 25 test cases, representing a total of 9290 fiber evaluations. The error bars indicate the standard deviation between average F-measure, recall and precision scores of each of the raters. Recall is the fraction of relevant fibers that were found. Precision is the fraction of fibers found that were relevant. F-measure is the harmonic sum of recall and precision.

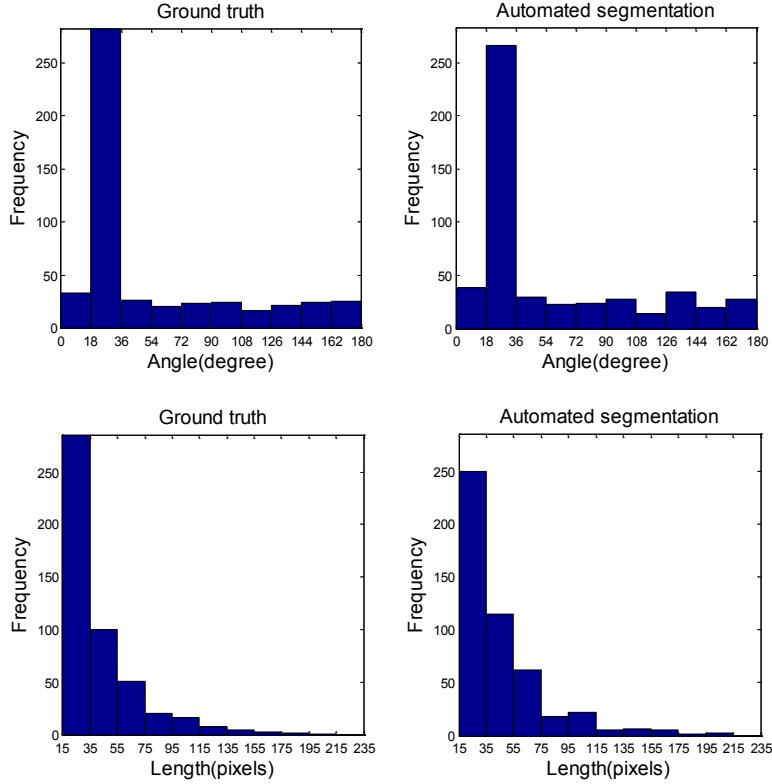


Figure 15. Distribution of angles (top row) and lengths (bottom row) of all fibers in all simulated test cases. Ground truth data is on the left and the results of the automated CT+FIRE algorithm are shown on the right.

After identifying the CT filter as a top performing preprocessing algorithm, we applied the combined CT-FIRE algorithm to the measurement of collagen fiber shape changes in an *in-vivo* mouse model for breast cancer. The results of this study are shown in Figure 16. Representative images show clear differences in waviness of fibers between the early (Figure 16A) and late (Figure 16B) time points. The colored lines overlaid on the images are the automated fiber segmentations produced by CT-FIRE. These overlays qualitatively illustrate the high fiber segmentation quality that can be expected from the CT-FIRE algorithm. Fiber waviness (W) was quantified for each extracted fiber by dividing the distance along the path of the fiber (d_0) by the distance between the end points of the fiber (d_n).

$$W = \frac{d_0}{d_n}$$

Thus for perfectly straight fibers $W = 1.0$ and more wavy fibers $W > 1.0$. We labeled a fiber as wavy if W was greater than a threshold value of 1.08. Then, to compute the wavy-fraction per image, the number of wavy fibers was divided by the total number of fibers found in each image.

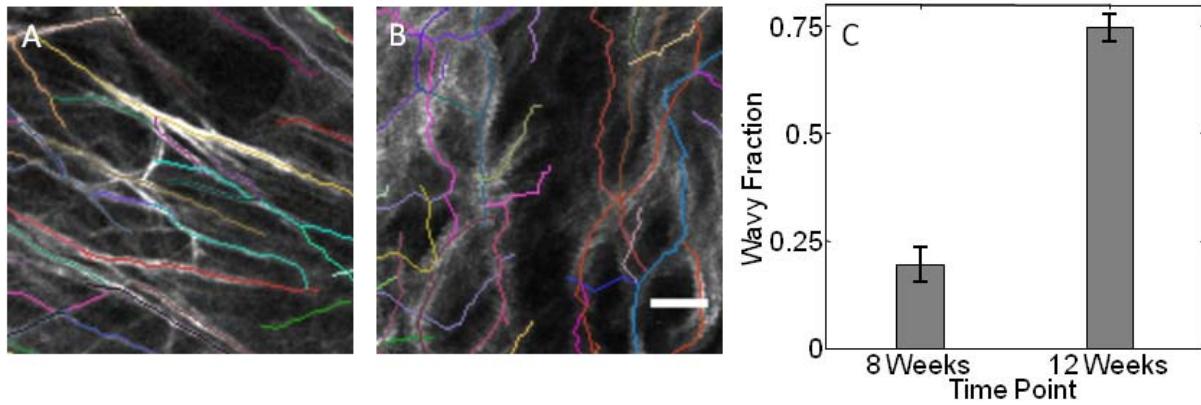


Figure 16. Demonstration of automated fiber segmentation feature extraction in an *in-vivo* mouse model for breast cancer. A mammary window was placed immediately superficial to a palpable mammary tumor and the collagen microenvironment was imaged 8 and 12 weeks of age. Automated fiber extractions are shown overlaid on representative images from the 8 (A) and 12 (B) week time points. The bar graph (C) shows the ratio of the number of wavy fibers to total fibers found in the image. Fibers are labeled wavy if the distance along the fiber divided by the distance between fiber endpoints is greater than 1.08. Error bars indicate one standard deviation of the computed average wavy-fractions among the 9 images analyzed for each time point. Scale bar is 25 microns.

The resulting wavy-fraction values were averaged over all images at each time point and plotted in the bar graph shown in Fig. 8C. We observe that the fraction of curvy fibers at the 8 week time point was approximately 0.19 ± 0.04 and 0.75 ± 0.03 at the 12 week time point. The error terms given here and error bars in Fig. 8C represent one standard deviation around the average wavy-fractions for the 9 images analyzed at each time point and indicate that there was close agreement between all images within a given time point.

3.4 Discussion

In the present study we compare preprocessing approaches prior to the application of the FIRE fiber extraction algorithm to identify fiber-level collagen characteristics in a series of SHG images of collagen in mammary tissue. Fiber extraction facilitates automated analysis of collagen features such as fiber number, length, and curvature. These features are important to researchers studying the role of the extracellular matrix in cancer progression. Computer assisted interpretation of these fiber-level collagen patterns has the potential to generate more reliable and reproducible results compared to manual or transform/filter-based quantification methods. Furthermore, an algorithm that identifies collagen fiber characteristics in tissue samples may enable large scale studies of tumor associated collagen signatures supporting the manual analysis performed previously [1].

To our knowledge, FIRE has not been applied to SHG images of collagen in tissue. According to our testing, FIRE works well in some situations without any preprocessing or pre-filtering. However, the algorithm fails when collagen fibers are densely packed or image quality is degraded, both common occurrences while imaging collagen in tissue. Our work aims to extend FIRE's application sphere to include complicated SHG images in tissue and to quantitatively compare the performance of a selection of preprocessing algorithms. Our results show that both the CT and the TF approaches are very promising and improve the fiber extraction accuracy achieved by the FIRE algorithm in many key situations. In addition, we demonstrate the application of CT-FIRE to extract *in-vivo* fiber curvature changes during the development of a mouse mammary tumor. Although FIRE is used in our study for fiber extraction, other effective approaches that have been developed for vessel segmentation or neural diffusion mapping such as statistical tracking [182-184] may be effective in SHG image analysis. We believe the CT and TF methods would generally improve these algorithms as well.

A recent review [131] suggested that there might be advantages of the CT method being applied in combination with other approaches for image processing such as fiber extraction, as we have now

demonstrated is the case. By selecting and thresholding the most representative scales, the CT based method shows good performance for both denoising the image and enhancing edge information producing a better fiber extraction among the proposed preprocessing algorithms discussed in this chapter. Evidence supporting this claim is presented in Figure 14 where the overall F-measure result was slightly higher for the CT method and notably a threefold improvement over a standard GF when considering all 25 test cases analyzed in this study. In addition, the CT based method simplifies the often-difficult choice of selecting a threshold to binarize the image early in the FIRE process. Image thresholding can be difficult in low SNR and non-stationary images but may be alleviated through the application of more complicated thresholding techniques [185] or via the grey level distance threshold [186]. In our case, the inverse CT makes grey level threshold selection simple by placing the background on the negative side of zero and the foreground on the positive side of zero, allowing the threshold to always remain at zero. Although grey level thresholding is simplified, CT denoising adds a threshold selection step. We chose a hard thresholding approach since it had robust performance for all cases we tested, though other soft thresholding or scale-adaptive thresholding techniques may be adopted to finely adjust the CT-reconstruction. In addition, to take full advantage of the multiscale analysis of CT based approaches, an optimal scale combination must be obtained according to the features of the images to take into account different fiber width, length and dynamic intensity changes. The CurveAlign software [187] previously developed in our group may be used to show the curvelet centers and directions of the fiber edges at a specified scale, which may be helpful for choosing the optimal scales and threshold of the curvelet coefficients.

The TF method produces slightly lower overall fiber segmentation accuracy compared to the CT method as shown by the lower optimized F-measure score in Figure 14. However, the optimized recall score of the TF method was higher than the CT method. This indicates the use of the TF method if recall is most critical, in other words if priority is placed on not missing real fibers. In this study, we decided to

balance recall and precision equally, therefore a missing fiber and a false alarm fiber were considered equally important (see F-measure calculation). When this is taken into account, although close, the CT method produced higher accuracy segmentations on average compared to the TF method.

The GF and SPTV methods produce similarly inferior segmentations compared to the CT and TF methods. One reason for this is that the GF and SPTV methods lack the ability to normalize the fibers in the images, such that bright or dim fibers, thick or thin fibers do not generate the same signal level in the output image. This lack of image normalization in the GF and SPTV methods cause difficulty in the threshold selection step. In addition, these filters do not enhance the ridges along the centers of the fibers, which is an attractive feature of both the CT and TF methods. The GF method is able to attenuate high frequency noise, but does not preserve edges. The SPTV method filters high frequency noise and preserves edge information, but allows plateaus of high signal level to remain in the image, such as those seen at the center of bright, thick fibers. For these reasons, the GF and SPTV methods are not ideal for preprocessing SHG images of collagen prior to fiber tracking.

Following our evaluation and comparison of these image processing algorithms, we applied a top performing algorithmic approach, termed CT-FIRE, to quantify collagen fiber shape changes over time in an *in-vivo* mouse model for breast cancer. We observed a significantly larger fraction of highly curved fibers at the late time point compared to the early time point, indicating a quantifiable collagen matrix reorganization in the vicinity of the developing mammary tumor. Although the precise mechanisms underlying this observed matrix reorganization are currently unknown, we have demonstrated the power of a tool like CT-FIRE to quantify key aspects of these dynamic processes and others like it, enabling further studies of ECM remodeling.

This chapter focuses on image processing as a technique for quantifying structural information about collagen fibers in SHG images. However, it should be noted that there are a number of related

techniques that use information about the polarization or directionality of the SHG signal to make inferences about collagen fiber orientation or estimates of the non-linear susceptibility tensor [16, 44, 65, 69, 70, 188, 189]. These techniques have the potential drawback of adding costly components to the imaging system and often requiring multiple images to be captured per imaging frame, limiting their usefulness for applications such as *in-vivo* imaging. Our goal here was to establish a robust technique for quantitative collagen architecture analysis of images captured with standard SHG imaging techniques.

It is worth mentioning also that although the CT and TF preprocessing methods can improve the results of the FIRE algorithm to some degree, they may do little about some intrinsic limitations of FIRE, such as the ability to properly segment crossing or cross-linked fibers, extremely curvy fibers, or fibers with gaps due to the fibers that travel in and out of the focal plane as we observed in our testing. However, with the improvements provided by the combined approach of CT-FIRE, we anticipate being able to more accurately measure collagen fiber angle distributions in a highly automated fashion, thereby leading to better understanding of the interactions between cells and collagen fibers. In order to link collagen architecture to cellular features, SHG imaging and CT-FIRE may be combined with complementary imaging techniques such as multiphoton excited fluorescence imaging [190] and fluorescence lifetime imaging [191] which allow imaging of both extrinsic and intrinsic fluorescence of tumor and stromal cells. In the future, accurate assessment of tumor-stromal interactions will help reveal prognosis or treatment response in diseases such as breast cancer.

3.5 Conclusion

We demonstrate here an integrated approach for quantitative SHG collagen image analysis and algorithm evaluation. We show the application of curvelet transform denoising as a preprocessing step for FIRE fiber extraction, a process we call CT-FIRE, performs more accurate fiber segmentations, compared to other techniques we investigated in a variety of collagen images of human breast and

mouse mammary tissue. We then demonstrate that CT-FIRE can automatically sense changes in collagen fiber curvature from images captured in an *in-vivo* breast cancer mouse model. Our current work uses both MATLAB and Fiji [178] image processing tools in combination. To make these approaches more widely accessible, we plan to develop a single Fiji plug-in to perform the CT-FIRE process to produce 2D and 3D collagen fiber network extractions. Other future efforts will include the evaluation of multiple fiber tracking algorithms applied to collagen fiber tracking in SHG images. Although this study focused solely on breast cancer, the use of these fiber quantification techniques should be easily adapted to SHG images of other collagen related diseases. A MATLAB implementation of the CT-FIRE algorithm is available at <http://loci.wisc.edu/software/ctfire>.

Chapter 4 Electronic Tumor Association Collagen Signatures

4.1 Background

Breast cancer diagnosis and staging have been revolutionized by new molecular screening assays based on immunohistochemistry [192], fluorescence in-situ hybridization[193], and reverse transcription polymerase chain reaction[194] which are all used to personalize care. These tools are helping patients live longer and receive better treatment than ever before. However, there remains a significant group of breast cancer patients for whom these new techniques ultimately fail, due to several factors including varying patient genotype and primary or acquired resistance to drugs such as the HER2 targeting drug Trastuzumab (trade name Herceptin) [195]. In addition, molecular screens are confounded by the high degree of intra-tumor genetic diversity and often require extra tissue sections to be cut, stained and evaluated on top of the standard hematoxylin and eosin (H&E) preparation. New assays that predict patient outcome and response to treatment are therefore critically needed if we are to continue improving breast cancer treatment and prevention. One promising area of development is image based assays which leverage high content imaging hardware and image analysis software to classify biological samples [196-198]. In many cases, image based analysis does not require more than the standard histopathology H&E stained slides prepared as part of the normal clinical workflow. In this chapter, we demonstrate the use of a new image based assay for predicting patient outcome using information about tumor-stromal interactions from standard H&E stained histopathology specimens.

Aberrant tumor-stromal interactions have been shown to accelerate tumorigenesis in breast cancer [199-201]. The importance of stromal collagen in breast cancer is highlighted by the link between breast cancer, breast density, and the increased deposition of stromal collagen [19, 21, 22, 202, 203]. Interestingly, although mammographic density, which is attributable to collagen content, is one of the largest risk factors for the development of breast tumors, there is currently no clinical intervention

based on mammographic density alone. This is due in part to the lack of a clear correlation observed between increased mammographic density and patient outcome. Most of the work to date [204-207] has defined mammographic density as a etiological factor and not as a prognostic factor. Recently Cil et al [208] explored mammographic density as predictor of local breast cancer recurrence. They reported that women with intermediate and high breast density had a significantly elevated risk to develop a local breast cancer recurrence. However follow-up clinical trials that incorporate additional risk factors such as obesity, are needed to examine the possible prognostic value of mammographic density in large and diverse patient cohorts before using density as a possible clinical target. As well recently there has been an effort to investigate the underlying contributor to mammographic density, focusing on one of the largest components, collagen. Several studies have shown a link between collagen remodeling and the invasion and progression of mammary cancer in mouse models [23, 24, 163]. Furthermore, there was a link observed between collagen morphology, particularly collagen alignment, and breast cancer patient outcome [1]. Provenzano et al. [23] first introduced the so called Tumor Associated Collagen Signature (TACS) nomenclature to describe collagen alignment patterns. Phenotypes are classified into three groups as described in Chapter 2, Table 1. Conklin et al. [1] qualitatively searched for these patterns in human breast cancer samples and found that the presence of the TACS-3 alignment phenotype was a prognostic indicator for disease free and disease specific survival for invasive breast cancer patients. Our quantitative study, presented here, builds on this previous work by defining an algorithmic model for TACS-3. We then apply this model to score for the presence of TACS-3 in a cohort of 196 patients and correlate the TACS-3 score with breast cancer recurrence and survival.

Previous collagen alignment studies have largely been facilitated by the development of second harmonic generation (SHG) microscopy techniques which have the ability to capture high contrast images of the collagen fiber extracellular matrix without the need for exogenous stains [13, 44, 57, 190]. The application of SHG imaging in cancer research is growing rapidly. For example, changes in the ratio

of the forward to backward propagating SHG signal have been recently linked to breast tumor progression [209] and positive lymph node status [210]. SHG directionality was also used by Ajeti et al. to quantify the collagen composition in breast cancer models [43] while Ambekar et al. used Fourier transform and polarization-resolved SHG imaging to differentiate malignant from benign tissues in breast biopsies [211].

In addition, many new computational techniques are being developed to quantify patterns observed in SHG images. For example, a directional gradient method developed by Altendorf et al. [166] provides 3D orientation and radius information about fibers in SHG images. Due to the fibrous nature of the collagen matrix, SHG images are particularly well suited for the curvelet transform, which is a multiscale, orientation sensitive version of the wavelet transform. The curvelet transform [81] and combined fiber tracking methods [212] have been applied to extract fiber orientation, length, curvature and radius from SHG images of collagen. One key feature that is missing from all of the available image analysis techniques is the ability to incorporate cellular information into the analysis. The interaction between tumor cells and collagen fibers cannot be fully assessed without integration of information about cellular morphology and associated collagen morphology. As well this information is critical for finding regions of interest with TACS, an essential task for any type of high-throughput screening where manual user searching is not practical. Herein, we describe a computational protocol that achieves this goal by integrating information about collagen fibers from SHG images with information about cells captured through bright field imaging of standard H&E stained slides to perform highly automated, prognostic TACS scoring. We call this novel quantitative platform Electronic Tumor Associated Collagen Signatures (eTACS).

4.2 Methods

In order for TACS to become a useful and fully validated biomarker, it must be screened for in several large studies containing many patients and diverse populations. As well, besides screening in heterogeneous populations it ideally needs to be screened in diverse sample types to account for possible subtle differences in surgery, pathology or sample preparation that could negatively impact sample consistency. This ability to rapidly screen in many sample types of large diverse populations would also open the door for TACS to be explored in other cancer types such as pancreatic and renal cancer. Heretofore, there has not been a method that automates enough of the process to enable such large scale adaptation. In our original studies, collagen fiber angles were measured by hand, one at a time, using ImageJ region of interest (ROI) marking tools [23, 26]. These experiments used information gathered *a priori* or from autofluorescence to identify tumor-stromal boundaries. In addition, imaging locations were chosen manually. Conklin et al. manually captured each individual image, used bright field images to manually identify tumor-stromal boundaries, and manually estimated collagen fiber angles [1]. In each of these cases, many subjective decisions were made while identifying which areas to image, which fibers to measure and what should be considered a tumor-stromal boundary. There has been progress made by our group and others in automating the fiber angle analysis steps of this task [76, 81, 164, 212, 213]. However, none of these methods bring automation to all four steps of the TACS analysis process, which are: **1.** Image capture **2.** fiber angle measurement, **3.** tumor-stromal boundary identification, and **4.** relative angle measurement between fiber and boundary. In this chapter, we describe our eTACS system which uses image analysis and supervised learning techniques to enable the automation of each of these tasks. The block diagram of our eTACS protocol is shown in Figure 17.

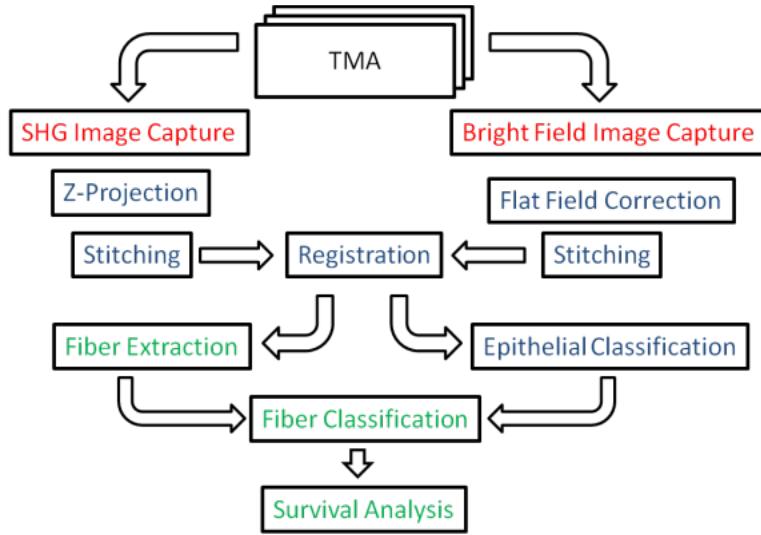


Figure 17. Block diagram of the eTACS system. The red steps are performed with WiscScan, the blue steps are performed with ImageJ/FIJI, and the green steps are performed with MATLAB based tools. The left side shows the steps performed on the SHG images, while the right side shows the steps performed on the bright field images. The middle column combines information from both modalities.

Starting with the tissue microarray (TMA) slides, we capture registered, whole-slide SHG and bright field images, extract fibers from the SHG images, tumor-stromal boundaries from the bright field images, and measure relative angles, all in a scripted pipelined process. We believe that the eTACS method will allow significantly larger scale studies to be performed in order to validate TACS as a prognostic biomarker in breast cancer and potentially other cancer types and investigate if TACS can be used to predict patient response to targeted therapies.

4.2.1 Human breast carcinoma tissue microarray

The TMA used here is the same as that used by Conklin et al.[1] for manual collagen alignment analysis.

The clinical profiles of all patients whose tissue is included in this TMA have been described in a previous study[214]. All tissue and patient information used in this study was acquired following institutional review board approval. Tumor tissues from 353 patients diagnosed with invasive carcinoma were resected by the same surgeon between 1981 and 1995. Pieces of each resected tumor were embedded in paraffin according to standard histopathology protocols. After excluding tumors smaller than 5 mm and severely damaged samples, 196 patients remained for analysis. Sections of 4 μm thickness were cut

from TMA blocks containing 1.0 mm diameter tissue cores, placed on glass slides, stained with H&E and mounted under a glass coverslip. Patients were followed for a median of 6.2 years, ranging from 1 month to 18.6 years.

4.2.2 Imaging System

TACS analysis requires the simultaneous analysis of information about epithelial cells and extracellular collagen. The interactions between collagen and cells can only be assessed computationally if the cellular information is carefully registered with images of the collagen. We have therefore optimized our imaging system for highly automated capture of large field of view, registered SHG and bright field images of stained microscope slides with the purpose of analyzing collagen angle with respect to cell cluster boundaries. For this study, we originally planned to use the same SHG and bright field images captured by Conklin et al. [1] since these were already manually annotated. Unfortunately, these images contained artifacts which, although trivial for the human visual perception system to overcome, are extremely difficult for computational systems to handle effectively. For example, SHG images were originally captured in the backwards direction with linearly polarized light, causing two artifacts. The first was simply a low signal to noise level due to few SHG photons traveling in the backward direction from the thin tissue sections [158, 215]. The second artifact was observed as a larger relative SHG signal from fibers in the direction parallel to the laser polarization [65]. Artifacts in the bright field images included significant vignetting at field edges and low signal to noise due to short exposure times. These artifacts were easily hurdled by the human observers making TACS assessments in our original study [1]. However, these artifacts are particularly difficult to handle by a computer vision based approach. We therefore decided to develop an imaging system and protocol that would fix many of these artifacts and allow for more consistent automated imaging. Similar image quality and consistency can be achieved with other SHG microscopes including commercial ones with the appropriate hardware but our analysis protocol did identify a necessary rigorous acquisition protocol that is best achieved with our new

automated SHG microscope described below. In general, the system should allow for forward SHG and bright field imaging with field of view as large and flat as possible, numerical aperture of at least 0.75, automated xyz motion control with appropriate position logging, circular polarization at the sample for SHG imaging, autofocus and automated switching between SHG and bright field imaging.

All samples in this study were imaged with the custom built integrated forward SHG/bright field imaging system shown in Figure 18. We call this system the Compact Automated Multiphoton Microscope (CAMM).

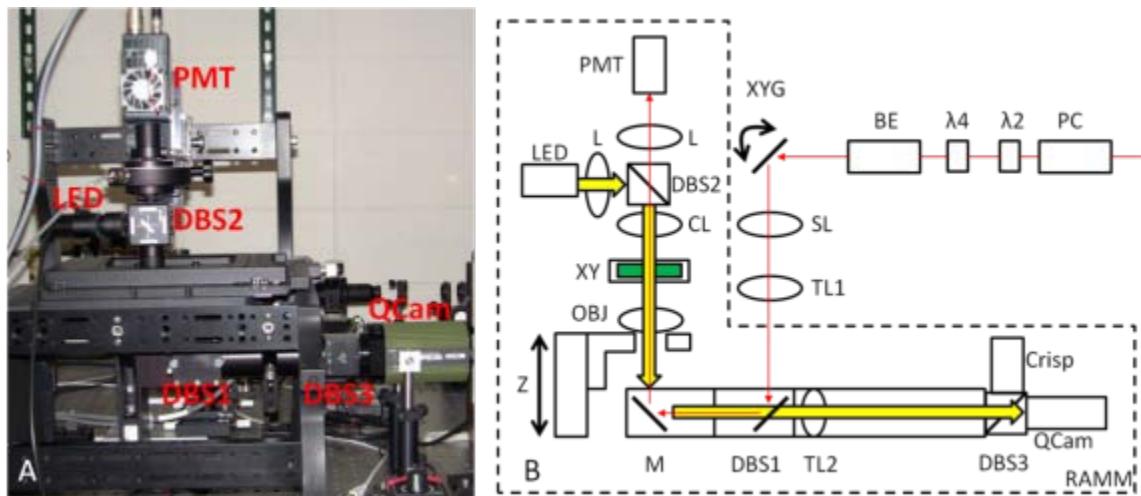


Figure 18. Photo of the CAMM imaging system (A) and optical block diagram panel (B). PMT = photomultiplier tube, QCam = QImaging RGB camera, LED = bright field lamp, DBS = dichroic beamsplitter, TL = tube lens, CL = condenser lens, L = lens, M = mirror, BE = beam expander, Z = z-direction translation, XY = xy translation, XYG = xy galvanometer driven mirrors, PC = Pockel's cell, RAMM = Rapid Automated Modular Microscope (ASI), λ_2 , λ_4 = half and quarter waveplates.

A MIRA 900 tuned to 780 nm, with a pulse length of approximately 100 fs, was directed through a Pockel's cell (ConOptics, Danbury, CT), half and quarter waveplates (ThorLabs, Newton, NJ), beam expander (ThorLabs), a 3mm galvo pair (Cambridge, Bedford, MA), a scan/tube lens pair (ThorLabs), through a dichroic beam splitter (Semrock, Rochester, NY) and focused by a 20X/0.75NA objective (Nikon, Melville, NY). SHG light was collected in the forward direction with a 0.54NA condenser (ThorLabs) and filtered with an interference filter centered at 390 nm with a full width at half maximum

bandwidth of 22.4 nm (Semrock). The back aperture of the condenser lens was imaged onto the 5mm aperture of a 7422-40P photomultiplier tube (Hamamatsu, Hamamatsu, Japan) the signal from which was amplified with a C7319 integrating amplifier (Hamamatsu) and sampled with an A/D converter (Innovative Integration, Simi Valley, CA). Timing between the galvo scanners, signal acquisition, and motorized stage positioning was achieved using our custom WiscScan software [216]. The Rapid Automated Modular Microscope system (Applied Scientific Instrumentation, Eugene, OR) served as our microscope base and we used ASI motorized translation stages for x, y and z motion control. The SHG light source was verified to be circularly polarized at the sample using the protocol of Chen et. al. [44]. SHG images were captured as stacks of 3 images spaced 3 μm apart, then z-projected to improve field flatness. Bright field images were captured with the same system using a MCWHL2 white LED lamp (ThorLabs) set up for Kohler illumination. White light from this lamp was separated from SHG light traveling through the condenser assembly using a short pass dichroic mirror with a cutoff at 670 nm (Semrock). An RGB camera (QImaging, Surrey, BC, Canada) was used to capture bright field images through WiscScan to allow for acquisition within a single application. Both SHG and white light images were tiled with 10% overlap using automation provided by WiscScan. Stage positions for individual images and pixel size data were stored in Bio-Formats image metadata [217] and this was then used by the Grid/Collection stitching ImageJ plugin [71] to reassemble a high resolution large field of view image of the entire TMA. When capturing large field of view images, the sample plane often walks out of the in-focus imaging plane as the stage is translated over large distances in x or y. We alleviated this issue using the CRISP autofocus system (Applied Scientific Instrumentation) which maintained an accurate distance between the coverslip and the objective throughout the whole slide stitched image capture. This allowed for a single bright field image to be captured at each location rather than a z-stack, improving capture speed, reconstruction speed, and production of unnecessary data. After SHG and white light images were captured and stitched, the two modalities were registered with the Landmark

Correspondences ImageJ plugin using 5 control points per image. The image of the entire TMA was registered in a single step, then each individual TMA core was cropped out of the full TMA image, producing 196 images. The resulting TMA core images were each 2048X2048 pixels, consisting of 4, 8-bit channels. The first three channels represented the red, green and blue planes of the white light image, while the fourth channel contained the SHG information.

4.2.3 TACS Model

Our TACS model is based on published observations by our group and others and relates information about collagen structure to information about epithelial region shape. The first step in the TACS scoring process is the identification of groups of straightened, aligned collagen fibers. The second step is to determine if those fibers terminate at or near regions of epithelial cells at steep angles. If a fiber meets these criteria, it is considered TACS-3 positive. If one or more TACS-3 positive fibers are found in a sub region of an image, that region is scored as TACS-3 positive. The number of regions with TACS-3 positive scores are then used to score the entire image. There are many details in these steps and defining parameters to account for each would produce a potentially fragile model. Instead, we have implemented a supervised learning approach that allows the data to most appropriately define the model. We perform this task computationally using a series of cascaded classifiers. The first classifier is trained to find epithelial regions in the images using a small set of annotated regions of interest (ROIs). The resultant epithelial cell model is then used to segment epithelial cell regions within the entire cohort of images. Features describing the epithelial regions are then combined with features derived from our fiber extraction algorithm and are fed into a second classifier which is trained to score each image as being TACS positive or negative based on a training set of annotated images. TACS scores are then fed into a Cox proportional hazard model to regress to censored survival data.

Our entire image cohort can be represented by a set S of registered SHG and bright field (RGB) images.

Each image $I_i(\vec{u}, g) \in S$ is composed of pixels $\vec{u} = (x_i, y_i) \in R^2, 0 < x_i, y_i < N$ and the function $g(\vec{u})$

which maps each pixel to a quartet of intensity values corresponding to the R, G, B and SHG intensity channels of the image. The first step of our protocol involves the extraction of collagen fiber objects from the SHG channel.

4.2.3.1 Fiber Extraction

We applied a technique called CT-FIRE[212] to the SHG images to enhance, trace and extract a network of collagen fibers for each SHG image $I(\vec{u}, g_{shg})$. CT-FIRE combines the advantage of the curvelet transform[171] for denoising the image and enhancing the fiber ridge features with the advantage of a fiber tracing algorithm[45] for automatic fiber extraction, being capable of extracting fiber geometric information such as length, angle, width, and curvature of each fiber. We apply the fast discrete curvelet transform (FDCT) to capture a collection of coefficients C^D in curvelet space which are defined as the inner product of the input SHG image channel with each of the curvelet basis functions

$$C^D(j, l, k) := \sum_{0 < x_i, y_i < n} I(\vec{u}, g_{shg}) \gamma_{jkl}^D(\vec{u}),$$

where $\gamma_{jkl}^D(\vec{u})$ is the digital curvelet waveform and jkl represent the scale, orientation, and location indices respectively. We use the open source FDCT MATLAB [171] library and specifically the "wrapping" version of the FDCT due to its simplicity. To denoise the image, we set all curvelet coefficients to zero that fall below a user defined threshold T as shown below

$$C_T^D(j, l, k) = \begin{cases} C^D(j, l, k), & |C^D(j, l, k)| > T \\ 0, & \text{otherwise} \end{cases}$$

This threshold is determined empirically on a small subset of SHG images to determine the appropriate level of noise reduction. The inverse FDCT is then applied to reconstruct an edge enhanced, noise reduced version of the image (Figure 22A). After reconstruction, CT-FIRE traces fibers, using the method of Stein et al. [45], by first finding local maxima in the result of the smoothed distance transform. Fiber

branches are formed by creating regions surrounding each local maxima, the size of which are defined by the result of the distance transform at the location of the local maximum point. The edges of this region are then searched for further local maxima. This process is repeated until no new local maxima are found indicating the end of a fiber branch. Short branches are then pruned from the network and closely spaced, similarly oriented fibers are merged. Fiber width (FW) was quantified for each extracted fiber by averaging the fiber widths ($2R_i$) at n points that were used to form the fiber

$$FW = \frac{1}{n} \sum_{i=1}^n 2R_i$$

where, R_i is the fiber radius at i^{th} point, estimated by the result of the distance transform at that location. Fiber straightness (FS) was quantified for each extracted fiber by dividing the distance between the end points of the fiber (d_n) by the distance along the path of the fiber (d_0)

$$FS = \frac{d_n}{d_0}$$

Thus for perfectly straight fibers $FS = 1.0$ and wavy fibers $FS < 1.0$. After fiber objects have been extracted from each of the images, we next segment epithelial cell regions.

4.2.3.2 Epithelial Cell Segmentation

The TACS-3 phenotype consists of straightened aligned collagen fibers that terminate near regions of epithelial cells such that the angle of the collagen fibers appear perpendicular to the epithelial stromal boundary. Detecting this TACS-3 phenotype requires knowledge of the locations of epithelial cells within the sample. We must then identify regions of epithelial cell clusters and identify a boundary between the epithelial cells and surrounding stroma. This task is performed in two steps. Step 1 uses the Trainable Weka Segmentation ImageJ plugin [178] to find epithelial cell nuclei and step 2 applies a cascaded matched filter, threshold operation to identify clusters and boundaries. A training set of 15

cropped 256X256 pixel images denoted as $t_i(\vec{u}_t, g) \in S$ is created that contains representative features from 5 classes: epithelial cell nuclei, other cell nuclei (including lymphocytes and fibroblasts), cytoplasm, collagen, and background. A further subgroup of pixels within the training images $\vec{u}_a \in \vec{u}_t$ were annotated as belonging to each class $w_k(\vec{u}_a)$, $k \in (1, \dots, 5)$. A feature vector $p_i(\vec{u}_t, g)$ is computed for each pixel and each channel of the training image where $i \in (1 \dots d)$ is the feature index and d is the dimensionality of the feature subspace. The feature set we used is listed in Table 4 and incorporates features at 5 scales for a total of 80 feature planes for each image channel.

Task	Feature Description	Total number of features
Epithelial Cell Classification	Gaussian Blur	5
	Sobel Filter	5
	Hessian	40
	Difference of Gaussians	24
	Membrane Projections	6
TACS-3 Fiber Classification	Fiber Curvature	1
	Fiber Width	1
	Fiber Length	2
	Fiber Density	9
	Fiber Alignment	9
	Epithelial Proximity	3
	Relative Epithelial Angle	2

Table 4. Features used in epithelial cell segmentation and TACS-3 fiber classification tasks.

The detailed implementations for each of these features are given in the online documentation for the Weka Segmentation plugin [218]. We then use a multithreaded implementation of the random forests classifier [219, 220] with a forest of 200 trees and 2 random features per node to build a model based on \vec{u}_a and w_k . The trained model was then applied to every pixel \vec{u} in the cohort producing a probability map for each class and each image using a scripted version of the plugin. The probability map for an example image is shown in Figure 23A. The epithelial class probability map is then filtered with a Gaussian filter matched to the average width of the epithelial cell nuclei (3 microns) and thresholded such that the top 80 percent of resulting pixels are retained. The resulting image is then filtered with a Gaussian filter matched to the width of the average sized epithelial cell cluster (25 microns) (Figure 23B),

then finally thresholded such that, again, the top 80 percent of resulting pixels are retained (Figure 23C). Following the final threshold step, regions smaller than 50 pixels in area are discarded and a mask is generated with epithelial cell clusters in the foreground and all else in the background. Epithelial mask pixels are represented here as e_i while epithelial region boundary pixels are created using a 8-connected neighborhood and are denoted as b_i . Mask images are saved as tiff files and read, along with the extracted fiber data, into our CurveAlign software for fiber/epithelial region feature extraction. Outlines of the resulting mask files are overlaid onto the original white light image to qualitatively validate the segmentation accuracy of the applied epithelial region model (Figure 23D).

4.2.3.3 Combined Fiber-Epithelial Features and Fiber Classification

In the sections above, we described our methods for epithelial cluster segmentation and collagen fiber extraction. With these two pieces of information we associate fibers with epithelial cell clusters and calculate measures of orientation and proximity between fibers and cell clusters. This task is performed by our open source, MATLAB based tool called CurveAlign [81]. This tool starts by reading in a fiber database file (generated by CT-FIRE) and an epithelial mask file (generated by our epithelial segmentation script). A feature vector p_i is then built for each fiber endpoint $v_i \in R^2$ in the image. The feature vector is populated directly with features derived above in the fiber extraction section including fiber length, curvature, radius and grey level. Both endpoints are given the same values for these single fiber derived features. The rest of the features are unique to each fiber end point. All features used in TACS-3 fiber classification are listed in Table 4. Many of the features in this section rely heavily on the nearest neighbor search routine which is formulated here as

$$\varphi^n(X, Q) = \underset{X \in D}{\operatorname{argmin}} \rho(X, Q)$$

where $D = \{X_1, \dots, X_n\}$ is a set of vectors in R^2 , Q is a query vector, $\rho(X, Q)$ is the Euclidean norm $\|X - Q\|_2$, and $\varphi^n(X, Q)$ is a vector of n points in X that are nearest to each point in Q . In addition,

some of the features compute a metric of alignment using vector addition according to the following algorithm

$$\sigma(\theta_X) = \frac{1}{n} \left| \sum_X \exp(i * \theta) \right|$$

where $\theta \in \{0, \dots, \pi\}$ is a vector of n orientations associated with the vector of n positions in X . The factor of 2 is included since we are using fiber orientations rather than 2π directions. In words, the alignment metric is calculated as the normalized vector sum of orientation vectors. Fiber density features are computed as the average distance from the current fiber endpoint to the $n = 2, 4, 8$, and 16 nearest neighbors. The density features for fiber endpoint v are therefore given by

$$fd_{n,v} = \frac{1}{n} \sum_j \rho(v - \varphi_j^n(v_i, v)_j)$$

Fiber alignment features are computed as the absolute values of the vector sum of the n nearest neighbor fiber endpoints and are given by $fa_{n,v} = \sigma([\theta_v, \theta_{\varphi_j^n(v_i, v)}])$ for fiber endpoint v where $[\cdot, \cdot]$ indicates vector concatenation and $n = 2, 4, 8$, and 16. Features that incorporate epithelial cell information include distance to nearest epithelial mask point $de_v = \rho(v - \varphi^1(e_i, v))$ and distance to nearest epithelial region boundary $db_v = \rho(v - \varphi^1(b_i, v))$. These features have the same value if the fiber end point is outside an epithelial cell region, but they are different if the end point is colocal with an epithelial region. Next we extract relative angle features. Angle with respect to nearest epithelial region boundary point is computed as $ab_v = \sigma([\theta_v, \theta_{\varphi(b_i, v)}])$, and angle with respect to nearest "extension boundary intersection point" is given by $ae_v^r = \sigma([\theta_v, \theta_{\varphi(l_i^r, v)}])$, where the set of points in l^r is computed by taking the intersection of all epithelial boundary points b_i and a line of length $2r$ extending from the fiber endpoint at an angle θ and is formulated below

$$l^r = b_i \cap bres(v - q, v + q)$$

where *bres* indicates a modified Bresenham algorithm [221] which is used to find all pixels along a line between the 2 points $v - q$ and $v + q$. The term $q = [r \exp(i * \theta_v)]$ is the offset from v in the x and y directions. For this last feature, three values of r (50, 100 and 200 μm) are calculated. These three lengths correspond to 5, 10 and 20 times the diameter of a typical epithelial cell and were chosen based on estimates of intercellular signaling distances [222]. If no intersection is found, then the ae_v^r feature value is set to zero. The angle of the tumor stromal boundary line $\theta_{\varphi(l_i^r, v)}$ is estimated by fitting a quadratic to 9 contiguous points on the boundary surrounding the intersection point (or nearest boundary point in the previous feature) and computing the tangent angle of the line fit at the midpoint.

The steps in the process of relative angle feature extraction are diagrammed in Figure 19.

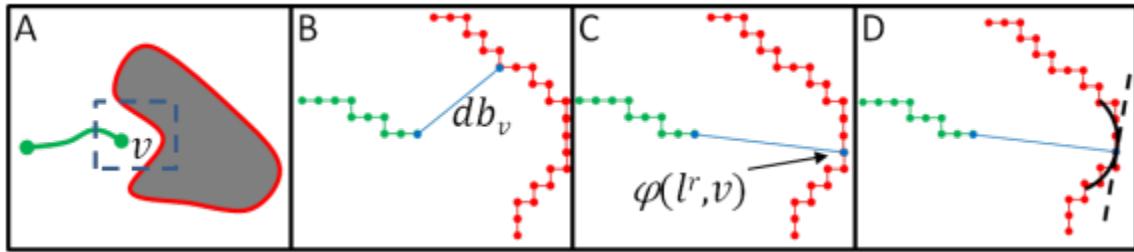


Figure 19. Integrated fiber angle and epithelial boundary feature algorithms. Panel A shows a single fiber (green) and epithelial region boundary (red) with one highlighted fiber endpoint v . Zoomed versions of panel A are shown in panels B, C and D where individual image pixels are represented as filled circles. The nearest distance from v to the boundary is indicated in panel B, intersection between the endpoint extension line and the boundary is shown in panel C and the quadratic curve fit to the boundary at the intersection point and tangent line are shown in panel D.

Each of these fiber level features p_i are calculated for every fiber endpoint v_i in the cohort. Fiber level features are then averaged among all fibers in a given image and training is performed with a subset of 16 images $I_t \in I_i$ that had been manually annotated as being TACS-3 positive or negative. A linear support vector machine is used to build a model which is then applied to all images in the cohort for classifying each image as being TACS-3 positive or negative.

4.2.4 Classification and Survival Analysis

The TACS-3 scores are correlated with disease free survival (DFS) and disease specific survival (DSS) data using the Cox-proportional-hazards regression method [223]. DFS is defined as the time from date of diagnosis to the first date of recurrence and DSS is defined as the time from diagnosis to death from breast cancer or date of last follow-up evaluation. In both cases, all other events are censored. The Kaplan-Meier method is used to compare DFS and DSS between TACS-3 negative and TACS-3 positive patients. Hazard ratios are computed using a log-rank test. Correlations between manual and computationally generated TACS-3 scores are made using the Pearson's linear correlation coefficient.

4.3 Results

Registered SHG and bright field images of a subsample of the TMA are shown in Figure 20 along with two zoomed versions of regions within the image. SHG information is added as an alpha channel on top of the raw RGB bright field image and pseudo colored yellow. The zoomed panel shows the detail available in the full resolution images captured with the 20X, 0.75NA lens and shows a region with a positive TACS-3 signature.

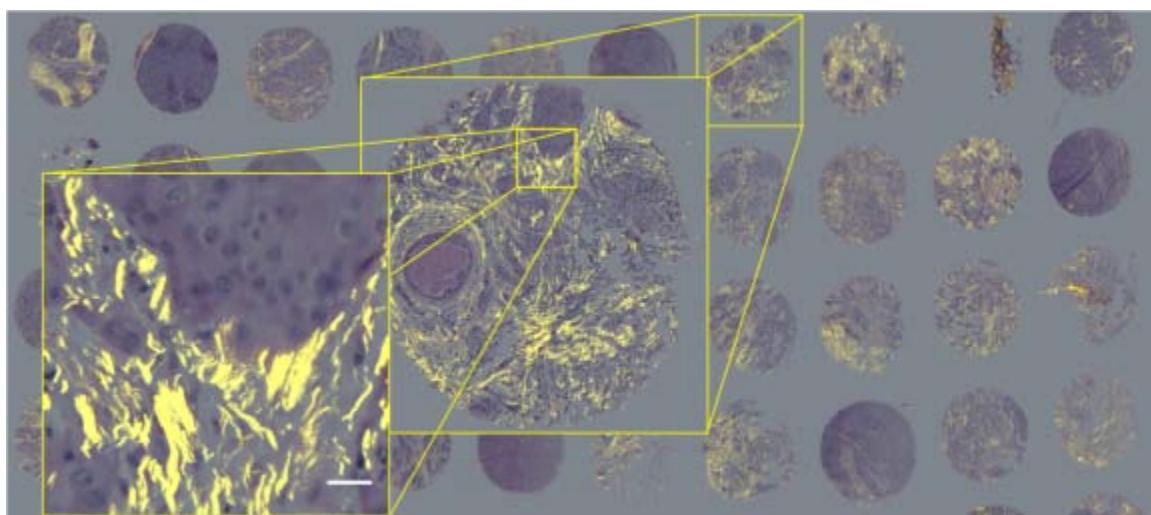


Figure 20. Micro array SHG overlay on H&E image with zoomed breakout panel. Scale bar = 25 μm .

A collection of 6 more TACS-3 positive regions were cropped out of the TMA images and shown in Figure 21. These images illustrate the features that are common to the TACS-3 signature including straightened, aligned fibers terminating in or near regions of epithelial cells at near perpendicular angles with respect to the epithelial region boarder.

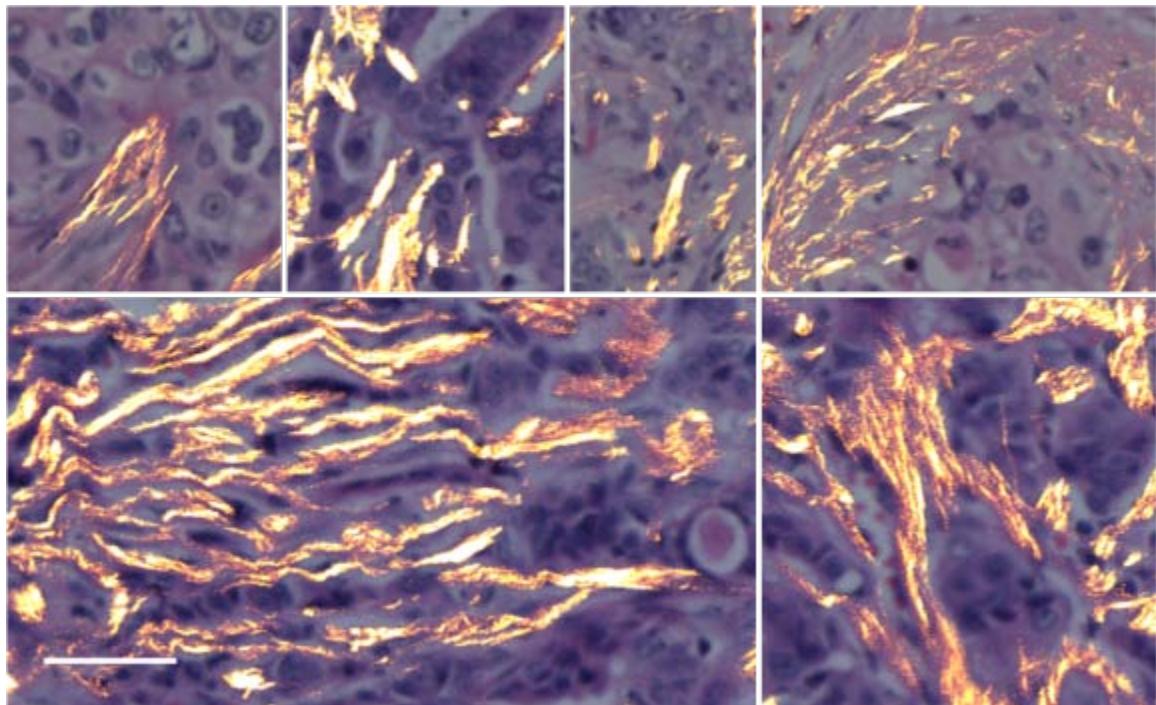


Figure 21. Examples of TACS-3 positive regions. Collagen fibers (yellow) registered and overlaid onto bright field images of H&E stained epithelial cells. These images illustrate the features of the TACS-3 signature, particularly, straightened aligned collagen fibers that terminate at steep angles relative to epithelial region boundaries. Scale bar = 25 μm .

A sample of our fiber extraction and epithelial region segmentation results are shown in Figure 22 and Figure 23, respectively. The absolute accuracy of each of these steps was assumed to be validated by the overall TACS-3 hazard analysis. Therefore, only qualitative evaluation of the outcome from these individual steps was performed. In both cases, epithelial region segmentation and fiber extraction were observed to accurately represent the data. The orientations of the epithelial cell region boundaries were compared to collagen fiber angles derived from the results of our fiber object extraction algorithm CT-FIRE which has been shown to perform well in comparison to other techniques [212]. A representative

sample of the results produced by this algorithm are shown in Figure 22. The intermediate product after the CT denoising step is shown in Figure 22A while the extracted fiber network is shown overlaid on the original SHG image in Figure 22B. Although some fibers are over- or under-segmented (annotated by green arrows), most of the extracted fibers properly represent the data. Figure 23D clearly demonstrates the ability of our epithelial cell segmentation algorithm to properly classify many of the regions of epithelial cells as positive. However, a few small regions of stromal fibroblasts and endothelial cells are included in the epithelial cell regions (annotated by green arrows). Although these errors occurred occasionally throughout the cohort, the noise they generated did not overcome the TACS-3 signal. Another feature evident in Figure 23D is the smoothness of the epithelial region boundaries. The boundary smoothness was dependent on the selection of our filter widths and binary mask thresholds. These parameters were selected to accurately represent the boundary orientation at the spatial scale of the epithelial cell regions.

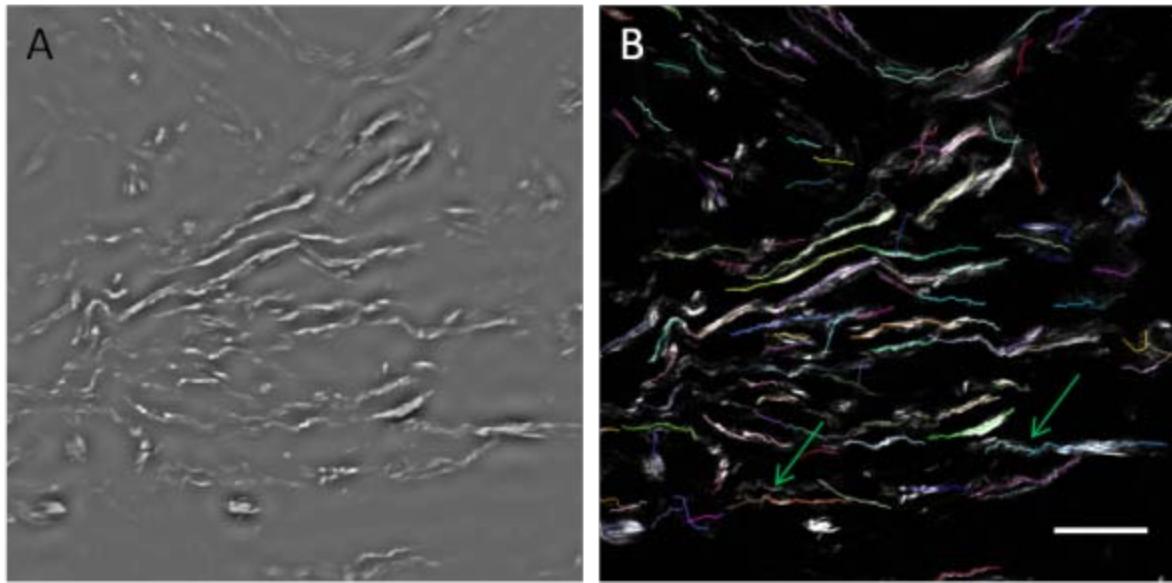


Figure 22. Sample fiber extraction results. The resulting image after curvelet denoising (A) shows likely fiber pixels in white and likely background pixels in grey. The extracted fiber network is overlaid on the original SHG image (B) showing many appropriate segmentations and a few under- and over-segmentations (green arrows). Scale bar = 50 μm .

Although correlation with survival is our ultimate goal, automated TACS-3 scores were also correlated with manual scores for each of the images. The Pearson linear correlation coefficient was used to determine this correlation, the results of which are in Table 5. The manual analysis performed by Conklin et al. produced 3 scores. Score 1 was the number of TACS-3 positive regions divided by the total number of regions analyzed, score 2 was the average number of TACS-3 positive votes per region among 3 observers, and score 3 indicated if one or more region received a TACS-3 positive rating. Table 5 shows positive correlation between all manual scoring methods and our eTACS scoring system presented here, with the highest correlation observed to be with manual score 2.

	Correlation Coeff.	p-value
Manual Score 1	0.295	2.7E-5
Manual Score 2	0.311	0.9E-5
Manual Score 3	0.271	1.2E-5

Table 5. Correlation between manual and the eTACS approach. There is statistically significant positive correlation between the eTACS method and all manual scoring approaches (N=196).

	Hazard Ratio (DFS)	p-value (DFS)	Hazard Ratio (DSS)	p-value (DSS)
Fiber Curvature	1.432	0.179	1.657	0.077
Fiber Density	2.195	0.003	1.831	0.032
Fiber Alignment	1.958	0.011	1.588	0.100
TACS-3 Score	2.588	0.002	2.250	0.008

Table 6. Univariate Cox Proportional Hazard Analysis results for various eTACS feature combinations. This table shows that the eTACS scoring method that includes both fiber and cellular information produces the more significantly prognostic scores compared to fiber information alone (N=196).

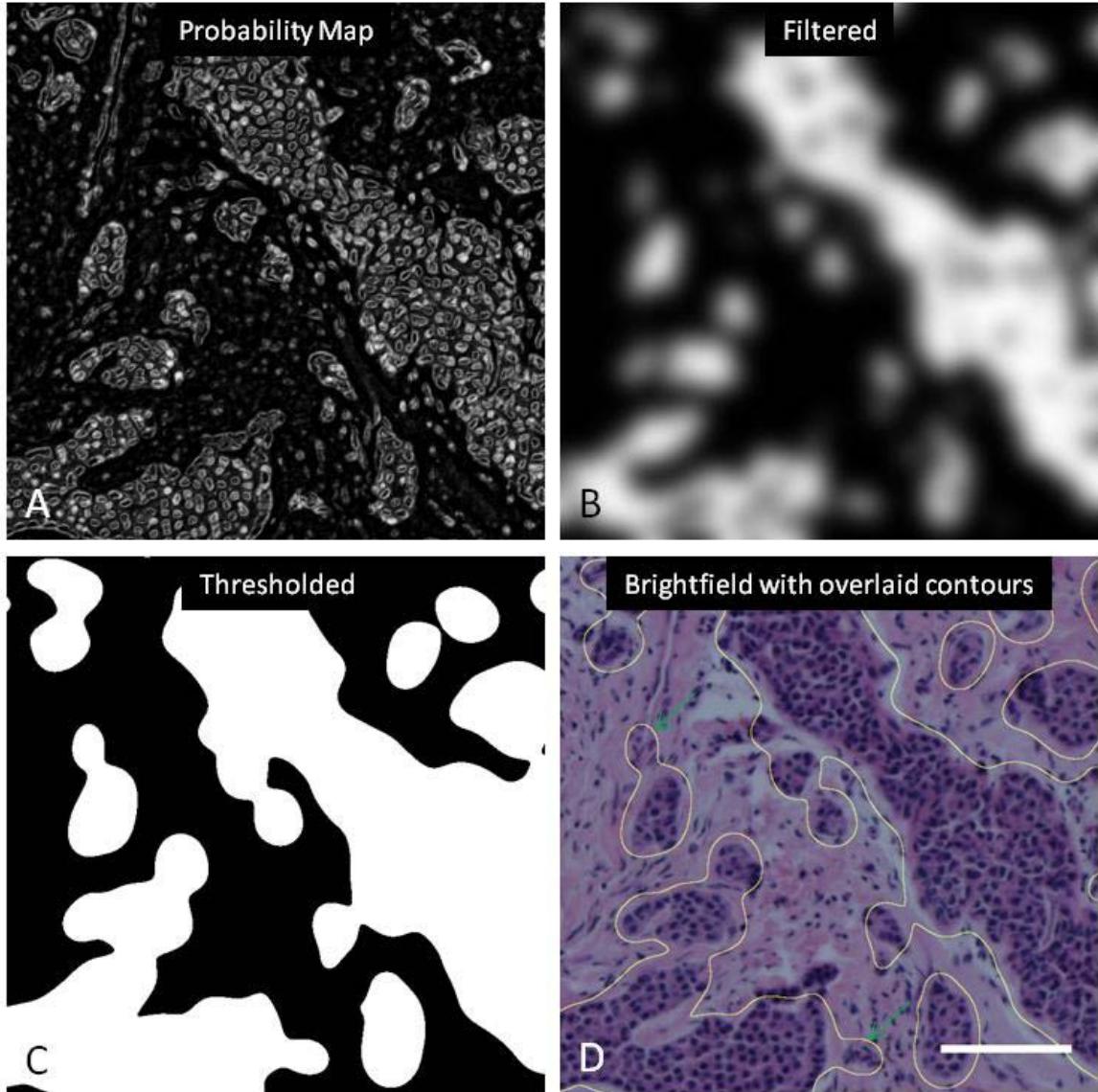


Figure 23. Sample epithelial cell segmentation results. The raw probability map produced by the Trainable Weka Segmentation ImageJ plugin (A) is filtered by a 2D Gaussian filter matched to the average cell cluster width (B) and thresholded (C) to produce epithelial region boundaries which are overlaid onto original bright field images to validate the segmentation (D). Scale bar = 100 μm .

The Kaplan-Meier curve in Figure 24 demonstrates the prognostic potential of our TACS-3 scoring system. TACS-3 negative patients showed significantly better disease-free and disease-specific survival compared to TACS-3 positive patients. In addition, Cox proportional hazards regression showed significant correlation between our computationally generated TACS-3 scores and survival as listed in Table 6. We also correlated scores created by individual fiber feature metrics with survival. Although

fiber features alone were in many cases correlated with survival, the highest correlation was observed when integrated fiber/epithelial features were considered in the analysis.

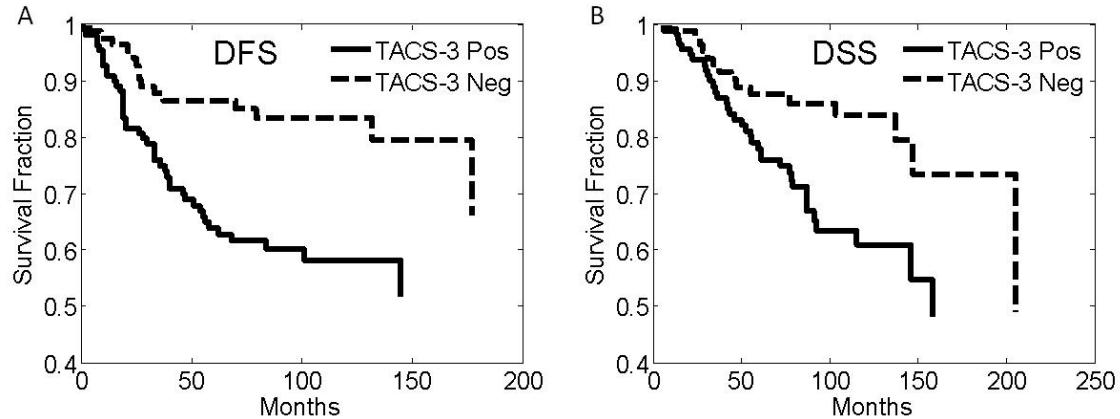


Figure 24. KM curves for DFS and DSS showing the prognostic classification produced by our eTACS approach. For TACS-3 Positive N=112 and for TACS-3 Negative N = 84.

Table 7 lists the 14 most informative features in the TACS-3 scoring process ranked according to their weight produced by the linear SVM algorithm. The SVM weight is used to assess which features were more or less informative in the classification. Of particular interest are the features labeled as "nearest distance to boundary" and "inside epithelial region". These features indicate the proximity between fibers and epithelial cell regions and were highly important in the TACS-3 classification. In addition, the difference in mean feature scores $d_f = f_p - f_n$ for the training set is shown in Table 7 for each of the ranked features. If d_f is greater than zero, then the TACS-3 positive images had larger values for those features and if d_f is less than zero, then the TACS-3 negative images had larger values. For example, the density features resulted in lower d_f values in the TACS-3 positive cases indicating that the TACS-3 positive images had lower density collagen fibers. On the other hand, d_f was positive for the alignment features indicating that TACS-3 positive images tended to have more aligned fibers. Interestingly, relative boundary angle was not as highly informative as many other features. However, it still was ranked within the top 14 of 27 features.

Feature Name	w	$f_p - f_n$	Feature Description
std nearest align	0.716	-0.004	Standard deviation of all alignment features
nearest dist to bound	0.491	-0.086	Shortest distance from fiber end point to epithelial cell boundary
inside epi region	0.393	0.572	If fiber end point is inside epithelial cell region, $f = 1$, else $f = 0$
std nearest dist	0.348	0.211	Standard Deviation among all nearest distance measures
mean nearest dist	0.238	0.188	Average among all nearest distance measures
curvature	0.219	0.010	Fiber curvature
box density 128	0.190	-0.311	Number of fiber endpoints within a 128 pixel radius
width	0.183	-0.050	Fiber Width
box alignment 64	0.155	0.136	Alignment of all fibers within a 64 pixel radius
box alignment 32	0.150	0.079	Alignment of all fibers within a 32 pixel radius
box alignment 128	0.111	0.145	Alignment of all fibers within a 128 pixel radius
box density 64	0.109	-0.249	Number of fiber endpoints within a 64 pixel radius
nearest relative boundary angle	0.087	0.004	Angle of collagen with respect to boundary at the nearest point
total length	0.079	0.004	Fiber length

Table 7. Feature ranking based on SVM feature weight w for a 16 patient (8 TACS-3 positive and 8 TACS-3 negative) training set including the average feature value difference between the positive and negative training cases $f_p - f_n$. These values show the trend of the feature values between TACS-3 positive ($N=112$) and TACS-3 negative ($N=84$) images.

The mean feature difference for the top 18 features is plotted as a bar graph in Figure 25. This figure helps to illustrate the most important features contained in the TACS-3 model and indicates whether the features tend to have larger or smaller values depending on the classification. The "inside epi region" feature reveals that TACS-3 positive fibers tend to terminate inside epithelial regions more often than TACS-3 negative fibers. The "box alignment 128" feature indicates that TACS-3 fibers are more well aligned with one another. The "extension point angle" indicates that the fiber angle relative to the epithelial cell boundary is greater, or steeper, for TACS-3 positive fibers and the "box density 128" feature reveals that TACS-3 positive fibers tended to be lower density compared to TACS-3 negative. The green error bars only indicate $\sigma/4$ away from the mean, illustrating that the individual features are not strong classifiers. Only by putting together the most informative features, are we able to boost the strength of the overall classifier and produce significantly prognostic results.

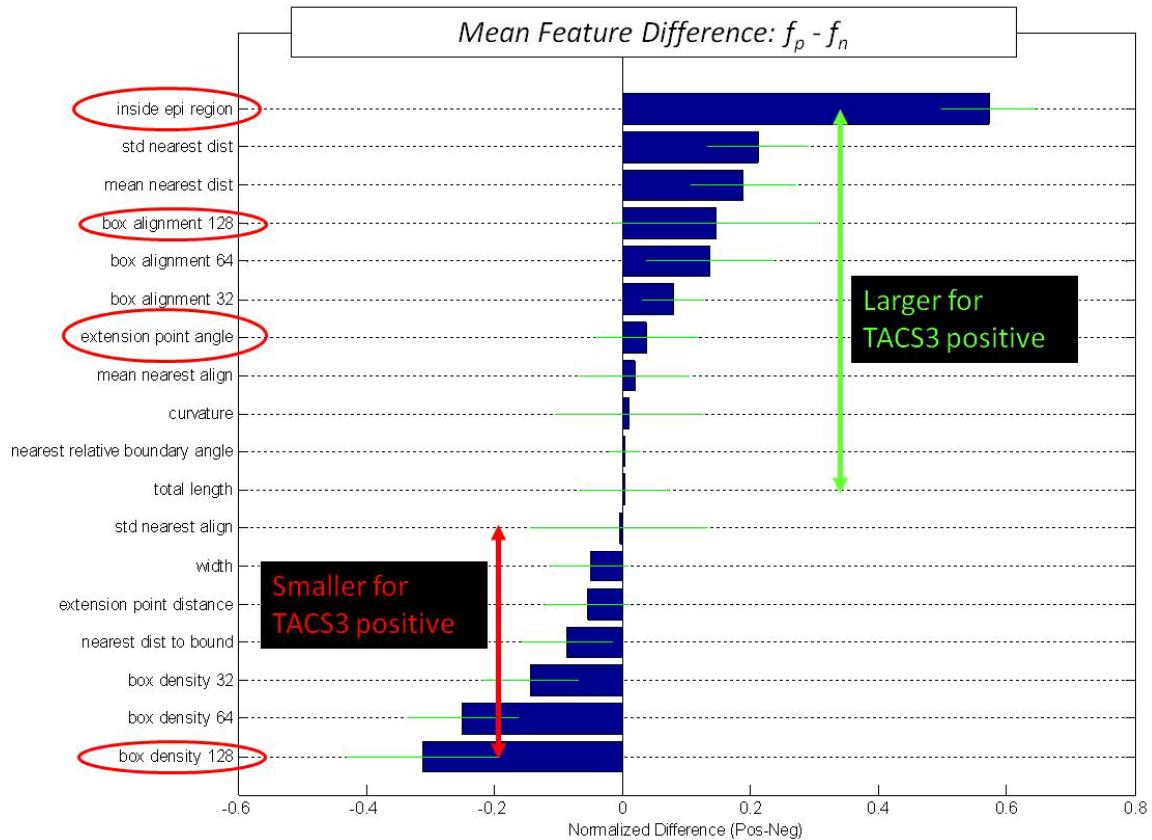


Figure 25. Bar graph showing the mean feature difference for the top 18 features. The features that had larger values for TACS-3 positive are towards the top of the graph and those with smaller values for TACS-3 positive fall towards the bottom of the graph. Green lines indicate the $\sigma/4$ away from the mean.

4.4 Discussion and Conclusions

The search for new prognostic and predictive breast cancer biomarkers is motivated by the need to improve patient outcome. A significant number of patients present with none of the currently available markers. In addition, survival and treatment response is often heterogeneous among patients within current biomarker classifications. The discovery and validation of new biomarkers will help to further improve breast cancer diagnosis and treatment planning. These new biomarkers need to be quantifiable, scalable and ideally correlate with both disease outcome and treatment specific response. The candidate biomarker we are focused on in this study (TACS-3) measures collagen characteristics

relative to tumor-stromal boundaries and has been associated with progression in mouse models and has been shown to predict disease recurrence and survival in human patients. Here, we demonstrate a protocol for using large field of view imaging techniques, image analysis and supervised learning to automate and quantify all of the steps in the process of TACS-3 scoring. These advances provide the tools for increasing the scale of TACS-3 investigations and applying TACS-3 scoring to cancers in other tissues such as ovarian [160], and pancreatic cancer [224, 225] where collagen fiber characteristics are predicted to correlate with prognosis. These techniques could also be used to characterize other TACS both current and to be identified to see if they have research value in animal models or prognostic value in clinical specimens.

Our eTACS system of imaging and analysis to produce prognostic TACS scores uses standard histopathology H&E slide preparations. The technique is therefore completely compatible with routine clinical protocols and is intended to augment currently available diagnostic tests. The current process requires no changes to current clinical protocol and the sample is returned to the clinician unmodified. We present a system that uses SHG imaging to capture collagen fiber images. Alternatively, wide field polarization sensitive techniques [226] such as LC-PolScope [227] or Picosirius red staining [35, 228] may be used to capture images of collagen fibers. One advantage of using SHG is that it does not require additional stains and can capture 3D fiber information in thick, unstained tissue samples. Unfortunately, when imaging in thick unstained tissue, the identification of epithelial regions can be difficult. Fortunately, techniques using autofluorescence and fluorescent lifetime imaging have been shown to be capable of this task [28, 191]. As implemented here, our eTACS algorithm is necessarily two dimensional, since we are relying on H&E stained slides for our epithelial cell information. However, fiber extraction, epithelial region segmentation and relative angle measurements can be extended to 3D without significant alteration of our general protocol. In addition, although our current eTACS protocol is able to

process standard H&E stains, staining for epithelial cells, with, for example, pan-cytokeratine conjugated stains, may simplify and improve epithelial cell segmentation.

Collagen alignment related image features are interesting not only because they have been shown to be prognostic, but because they have been shown to be directly linked to cancer biology. Our group and others have found that cells are more likely to invade along parallel, aligned collagen fibers [17, 27], features that are directly being measured by our system. This is not always the case for other machine vision systems developed for biological image classification such as WNDCHRM [196], where informative image features do not necessarily relate to the biology at hand. In the case of eTACS, biological observations have driven the image analysis model, therefore features are more easily linked back to biological functions.

The TMA used in this study was split between two slides. For the combination of SHG and bright field images, each slide required approximately 8 hours of scope time and stitching required approximately 2 hours of computer time on a 1.5 GHz Xeon processor (Intel, Santa Clara, CA). The imaging and stitching processes were not optimized for speed in this study. Multi-focal scanning [229] or wide field multiphoton techniques [230] could be employed for faster SHG imaging. In addition, bright field slide scanning speed could be improved through time-delay-integration microscopy techniques [231]. Stitching time could be eliminated by stitching the data while it is being collected. The eTACS system was able to successfully analyze 196 out of a possible 206 tissue cores in our TMA. Ten of the samples were discarded due to significant folding or tearing in the tissue. These ten samples were discarded while cropping each of the TMA core images out of the whole slide scanned images. Though they could have been automated, these two steps, 1) cropping and 2) discarding damaged cores, only required approximately 20 minutes to manually process all 206 samples. Fiber extraction required approximately 4 minutes per image, epithelial cell segmentation required 3 minutes per image, and the final eTACS

feature extraction process required approximately 1 minute per TMA image. Each of these steps could be performed in parallel, since each image is independent from one another, therefore the total computational processing time for all samples could be reduced to approximately 8 minutes with appropriate parallelization. In our case, we processed images on an Intel Corei7 machine with six processor cores, allowing for six images at a time to be processed in parallel.

High mammographic density is one of the largest risk factors for the development of breast cancer and has been associated with increased epithelial cellularity and increased collagen density [21, 22, 207]. Increased collagen density has been observed to promote tumor progression in a mouse tumor model [26] and in node positive breast cancer [210] leading one to potentially conclude that collagen density causes elevated risk. However, Maller et al. [232] observed that high density, non-fibrillar collagen protected against tumor progression and alternatively, that linearized collagen fibers induced invasive cellular behavior. In agreement with these recent findings, we observe here that TACS-3 fibers are more commonly present in regions of lower fiber density and are more likely to be thinner, more linearized fibers. Thick, curvy, and more dense collagen fibers are unlikely to contain TACS-3 fibers and are observed to be associated with a better prognosis. These observations support the hypothesis that collagen fiber shape and organization is a key aspect of the invasive ECM phenotype.

In conclusion, we present an imaging and analysis protocol that uses high content imaging techniques coupled with supervised learning to perform semi-automated TACS-3 scoring of slide mounted biopsy samples. We apply our eTACS technique to a previously annotated TMA containing tissue from 207 patients with invasive breast cancer. The resulting scores are shown to positively correlate with manual annotations and to predict patient outcome with good statistical significance. Future work will attempt to validate this technique on larger cohorts of breast cancer patients, to study ECM targeted drug responses in animal models, and to study collagen alignment in other cancers. As well future work will

focus on improving the clinical application of these techniques so they can be run by non-trained clinical personnel and be run automated at the time of acquisition to find regions of interest and TACS within those regions automatically. Together with more automation, eTACS screening has great potential as a clinical diagnostic tool to be performed on large population cohorts.

Chapter 5 Three Dimensional SHG Imaging

5.1 Introduction

Collagen fiber orientation is emerging as an important factor in many biological processes ranging from osteogenesis imperfecta [42] and birth defects [233] to tendon remodeling [234] and ovarian cancer [160]. In breast cancer in particular, there is growing evidence that the collagen, which is a major component of the extracellular matrix (ECM), plays an important role in cancer invasion and progression. Non-linear microscopy is the key tool used for observing interactions between cancer cells and collagen fibers in each of these studies. The closely linked non-linear microscopy modalities of Multiphoton Fluorescence Excitation (MPFE) and Second Harmonic Generation (SHG) microscopy optically section a sample using focused ultrafast pulses of light at near infrared wavelengths. MPFE microscopy is a fluorescence based imaging method that can excite intrinsic fluorophores such as nicotinamide adenine dinucleotide, flavin adenine dinucleotide, elastin [15] and extrinsic fluorophores such as green fluorescent protein [235] deep within intact tissue [236] and even live animals [236, 237]. Since fluorophore excited state lifetimes are usually longer than the stochastic motions of fluorophore molecules, MPFE signal creation is not particularly sensitive to changes in excitation polarization. On the other hand, in SHG microscopy, the excitation light deterministically interacts with the nonlinear susceptibility tensor of the sample to produce the coherent second harmonic of the excitation frequency. The frequency doubled emission light can be detected in both the backwards and forwards directions and can be detected and differentiated simultaneously with MPFE signals [13]. Unlike MPFE, SHG signal generation is highly dependent on the polarization of the excitation light with respect to the sample orientation, which effects the three-dimensional (3D) imaging capability of this technique. Although, both of these methods have been reported to be capable of producing 3D images, SHG imaging of collagen fibers should only be considered to be fully 3D if all 4π fiber orientations are

represented in the captured images. Most observers have suggested that the SHG signal from collagen is strongest when the excitation polarization is parallel to the collagen fiber axis [13, 70, 188]. It has also been noted [44] that if the polarization of the excitation light is circular at the focus, then fibers in the x-y plane at any azimuth angle β , as shown in Figure 26, should be visible in both the forward and backward directions. On the other hand, the literature is somewhat conflicted about how to best capture SHG signals when the elevation angle δ of the fiber is non-zero, that is, when the fiber tips out of the x-y plane such as the fiber depicted in Figure 26.

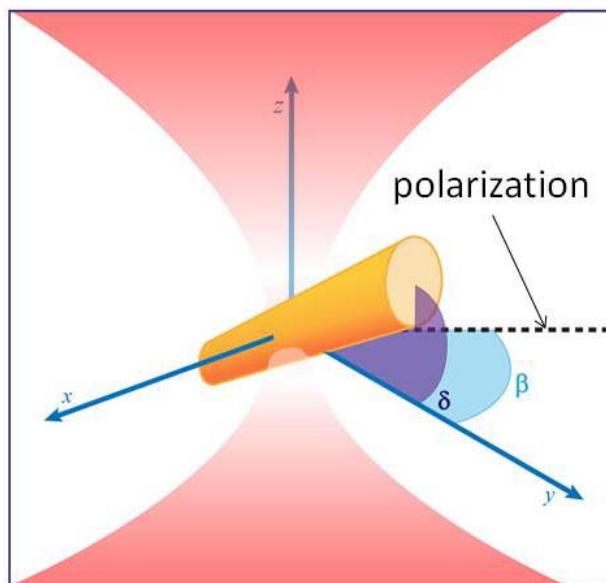


Figure 26. A focused beam (red) traveling in the negative z direction with an arbitrary polarization angle β , incident on a collagen fiber (yellow) oriented at an arbitrary elevation angle δ .

Some research has suggested that fibers oriented with a large δ angle are only visible in the forward direction [13, 70, 158, 238], while others seem to suggest that 3D SHG imaging is possible in the backward direction as well [166, 239]. Still others suggest that 3D fiber orientation are ideally probed by radially polarized beams using a high numerical aperture (NA) lens and condenser [173, 174, 240].

There are many challenges to 3D imaging with the SHG microscope including the non-spherical excitation volume, attenuation due to scattering in the sample, and the sensitivity of SHG detection to

collagen fiber orientation. In this chapter, we focus on fiber orientation sensitivity and set out to experimentally measure the ability of SHG microscopy to detect axially oriented fibers. We first image a stained collagen gel in the backward direction using MPFE and SHG using a high numerical aperture lens. After tracing the 3D fibers in both the MPFE and SHG images, we are able to measure the elevation angle at which collagen fibers tend to disappear from the backward SHG images. Our second experiment images and analyzes a stained collagen gel in the forward direction with both MPFE and SHG showing that all collagen fiber angles are represented in the SHG image. Our third experiment shows that axially oriented collagen fibers are visible in the forward direction at NAs as low as 0.5. These three experiments reveal interesting features about the SHG contrast mechanism that could help to inform studies focused on 3D collagen imaging and may be important in studies that attempt to estimate mechanical properties of collagen networks based on SHG images.

5.2 Materials and methods

5.2.1 Collagen Gel Preparation

Unstained collagen gels of 1 mg/mL were prepared by mixing 0.2 mL 10X PBS, 0.016 mL 0.1 M NaOH, and 1.45 mL distilled water on ice. Then 0.33 mL Nutrigen (5010-D, Advanced Biomatrix, Poway, CA) was added and mixed gently with a pipette. This solution was then poured into a glass bottom dish (FD35-100, World Precision Instruments, Sarasota, FL) and incubated at 23 degrees Celsius for 24 hours.

Stained collagen was prepared by following the protocol of Baici et al. [241]. Briefly, 2 mL of Nutrigen was dialyzed in 0.4M acetic acid for 3 hours, 0.25M sodium bicarbonate for 3 hours, fluorescein isothiocyanate (FITC) for 12 hours (0.2 mg/mL in sodium bicarbonate), then 0.4M acetic acid again for 3 hours. Stained gels of 1 mg/mL were then poured using the procedure listed above for unstained gels.

Gels were made at low concentration, using low pH and low incubation temperature to achieve thicker fibers with larger pore sizes [238] to allow for easier identification of individual collagen fibers and

reduce multiple scattering of forward propagating light. The volume of solution combined with the dish well diameter produced a gel that was approximately 500 microns thick. 3D gel imaging on our upright microscope was challenging due to the gels' tendency to wobble during z-stack imaging. We found if care was taken to not release the gel from the glass dish and if imaging was performed near the glass gel interface, then motion artifact free image stacks could be captured.

5.2.2 SHG Imaging System

Our forward/backward SHG imaging system was developed around an Olympus microscope stand as reported by Hall et al. [242]. The laser source was a Coherent Chameleon centered at 890 nm. Laser power was modulated by a motor controlled half waveplate and a polarizing beamsplitter. Power at the focus was set to 10 mW during the stained collagen gel imaging experiments because the stained gel tended to denature quickly at higher powers. 50 to 70 mW were typical at the focus for unstained gel imaging. Approximate circular polarization was achieved at the focus with a half and quarter waveplate placed into the beam following the procedure of Chen et al. [44]. Scanning and data acquisition were performed using our custom open source Wiscscan software package [216]. MPFE and SHG signals were differentiated with bandpass filters. Our SHG filters were centered at 445 nm with a 20 nm bandwidth (Semrock, Rochester, NY) and our fluorescene filters were centered at 520 nm with a 35 nm bandwidth (Semrock, Rochester, NY). Identical volumes were captured with each filter individually. The MPFE image stack was captured first, the filter was changed, then the SHG image stack was captured. The objective used in this study was the Olympus 60X 1.0 NA LUM Plan FL N water immersion lense. Numerical aperture was adjusted between 1.0, 0.8 and 0.5 by placing different sized apertures at the back aperture of the objective. Identical volumes were captured with all three NAs one after another by capturing a z-stack at one NA then carefully changing the NA such that the sample was not disturbed and repeating the z-stack capture. Laser power was set to 40, 45 and 70 mW for the 1.0, 0.8 and 0.5 NA images respectively to compensate for photon density reductions at the focus. The NA was verified by imaging a

0.2 μm diameter YG fluorescent bead (Molecular Probes, Grand Island, NY) and measuring the full width at half max of the center of the bead in the X-Y plane. The measurements were then substituted into the equations given by Zipfel et al. [190] to estimate the actual NA. Measured NAs were found to be within 5% of the estimated NAs. We set the condenser NA to 0.9 for all forward SHG images. Stacks were captured with an axial extent of 50 μm with a 0.5 μm step size for the stained gel images and 25 μm extent with a 0.25 μm step size for the unstained gel images. Pixel dwell time was set to 3.8 microseconds for all imaging experiments. Our detectors in the forward and backward directions were identical 7422-40P photomultiplier tubes (Hamamatsu, Japan).

5.2.3 Image Processing

The images presented here were first window leveled using the default automatic Window/Level function provided by FIJI/ImageJ [178]. All images in each stack were then filtered using a 3-D Gaussian kernel with a standard deviation of 0.2 μm . Lateral projections were created with the 3D Projection ImageJ plugin using the brightest point projection method and interpolation selected.

3D fibers were extracted from image stacks using a 3D skeletonization tool named FIRE [45] which was written in MATLAB (Mathworks, Natick, NY). Briefly, this algorithm finds local maximima of the 3D distance transform to produce fiber seed points. These points are then connected to form fibers based on their proximity and local trajectory. The fiber network is then pruned to remove branches and short fibers. All parameters used in this algorithm were empirically tuned to generate acceptable fiber extractions. The 3D fiber tracking capability of FIRE was characterized and verified using both synthetic 3D fiber images and stained collagen gel images captured with a confocal microscope [45]. The tool was only slightly modified to output fiber elevation angle distributions where equal length fiber segments of approximately 2 μm counted as 1 observation in the histograms reported below.

5.3 Results

5.3.1 Elevation angle sensitivity of backward-propagating SHG

To experimentally determine the angular selectivity of the backward SHG imaging configuration, similar volumes of FITC stained collagen gels were imaged with both MPFE and SHG using a 1.0 NA water immersion objective. Although the FITC staining produced fibers that were significantly more photosensitive, the fibers retained the ability to produce SHG signals. X-Y projections of the merged MPFE and SHG images reveal that both MPFE and SHG are able to detect all fibers in the X-Y plane, but, as described previously [238], MPFE images contain many additional circular, punctate regions that are not present in the backward SHG images (Figure 2). Using the X-Z projections, these circular regions are confirmed to be fibers oriented out of the imaging plane. If we allow the MPFE images to represent ground truth, then fiber extraction and elevation angle quantification show that as the fiber elevation angle reaches approximately 30 degrees, backward SHG imaging misses about 85% of the fibers that are present in the volume, as shown in figure 2D.

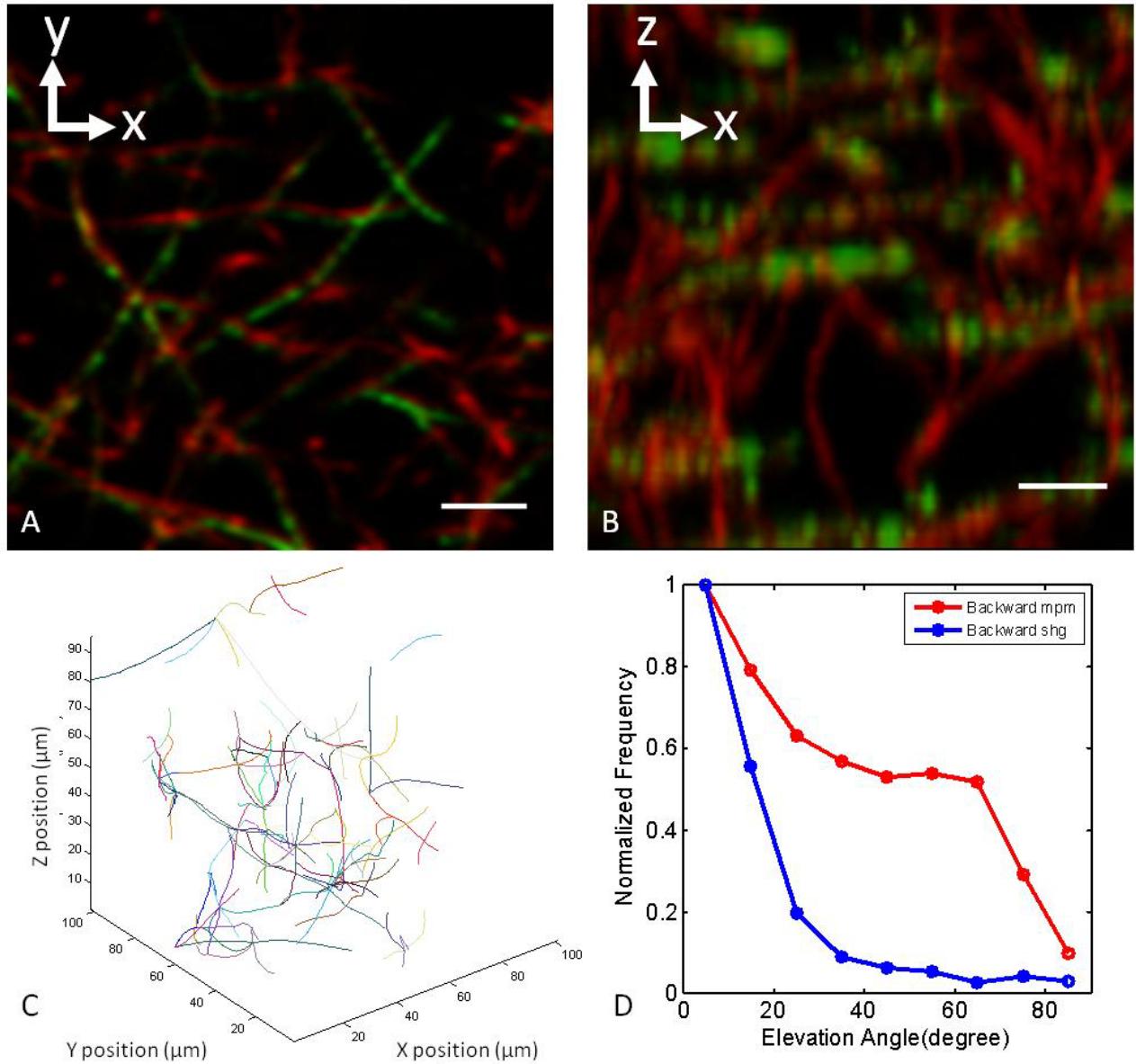


Figure 27. X-Y (A) and X-Z (B) projections of a 50X50X50 μm volume captured with MPFE (red) and backward SHG (green). The X-Y projection is a single section from a z-stack while the X-Z projection is a 5 μm thick maximum intensity projection in the Y-dimension. Fiber extraction was performed on both the MPFE and SHG image stacks. Comparing panel A and B immediately reveals that the excitation volume is longer in the Z-dimension compared to the X and Y-dimensions. The extraction and reconstruction for the SHG channel is shown in C while the distribution of elevation angles is shown in D. Scale bar = 10 microns.

5.3.2 Elevation angle sensitivity of forward-propagating SHG

Although detection of fibers oriented with a large elevation angle is a significant challenge for backward SHG, this is not the case for forward SHG. This was verified by imaging nearly identical volumes of a stained collagen gel with both MPFE and forward SHG (Figure 2). All fibers that were visible in the MPFE

image were colocalized in the forward SHG images, in both the X-Y and X-Z projections. Further, fiber extraction and elevation angle quantification reveal that all fiber angles present in the MPFE images were well represented in the forward SHG images with the largest deviation being approximately 30% at an angle of 35 degrees. At all other angles, the fiber distributions matched to within approximately 20% of each other. These variations are not observed in the image data and are attributed here to slight errors in the fiber extraction algorithm.

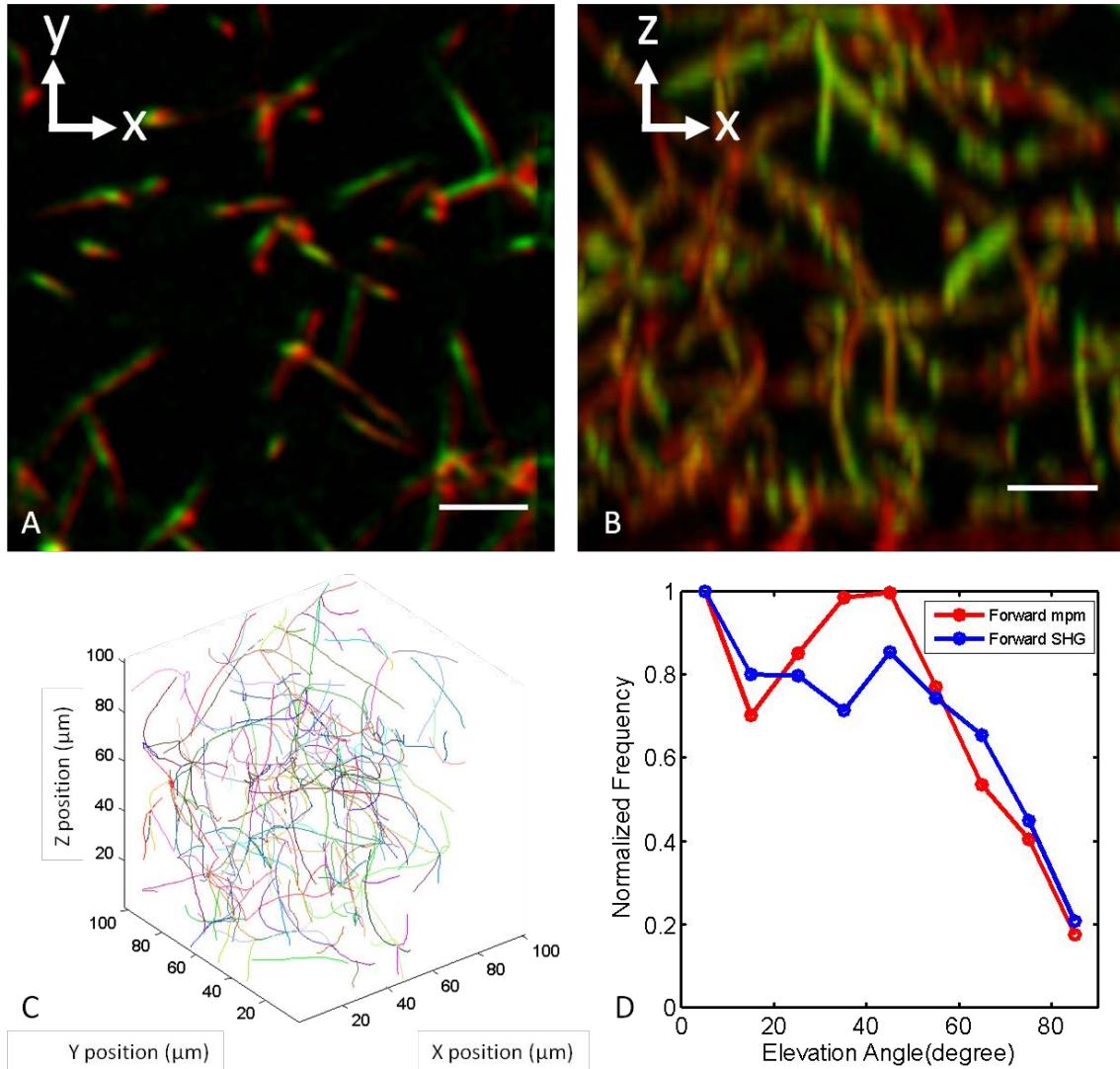


Figure 28. X-Y (A) and X-Z (B) projections of a 50X50X50 μm volume captured with MPFE (red) and forward SHG (green). The X-Y projection is a single section from a z-stack while the X-Z projection is a 5 μm thick maximum intensity projection in the Y-dimension. Fiber extraction was performed on both the MPFE and SHG image stacks. The extraction and reconstruction for the SHG channel is shown in C while the distribution of elevation angles for both MPFE and SHG is shown in D. Scale bar = 10 microns.

5.3.3 NA dependence on axial detection with forward-propagating SHG

Theoretical reports have suggested that axially oriented fiber detection is maximized with high NA

excitation and collection objective lenses [66, 173] where efficient production of axially oriented fields and therefore material polarizations can occur. Based on estimates of Yew and Sheppard [173], for a linearly polarized beam, approximately 30% of the field is tipped into the Z-dimension for an NA of 1.4.

Here, we attempt to determine at which NA axially oriented fibers disappear from forward SHG images by imaging unstained gels using 3 different NAs (Figure 29). Lower NAs resulted in more blurred images in all three dimensions, and, as expected, the Z-dimension was more heavily blurred compared to the X-Y directions. Surprisingly, even though there was significant Z-axis blurring, axially oriented fibers were clearly visible at both 0.8 and 0.5 NAs in all cases we observed. The images shown in Figure 29 are representative of 7 image volumes captured in different locations of the collagen gel.

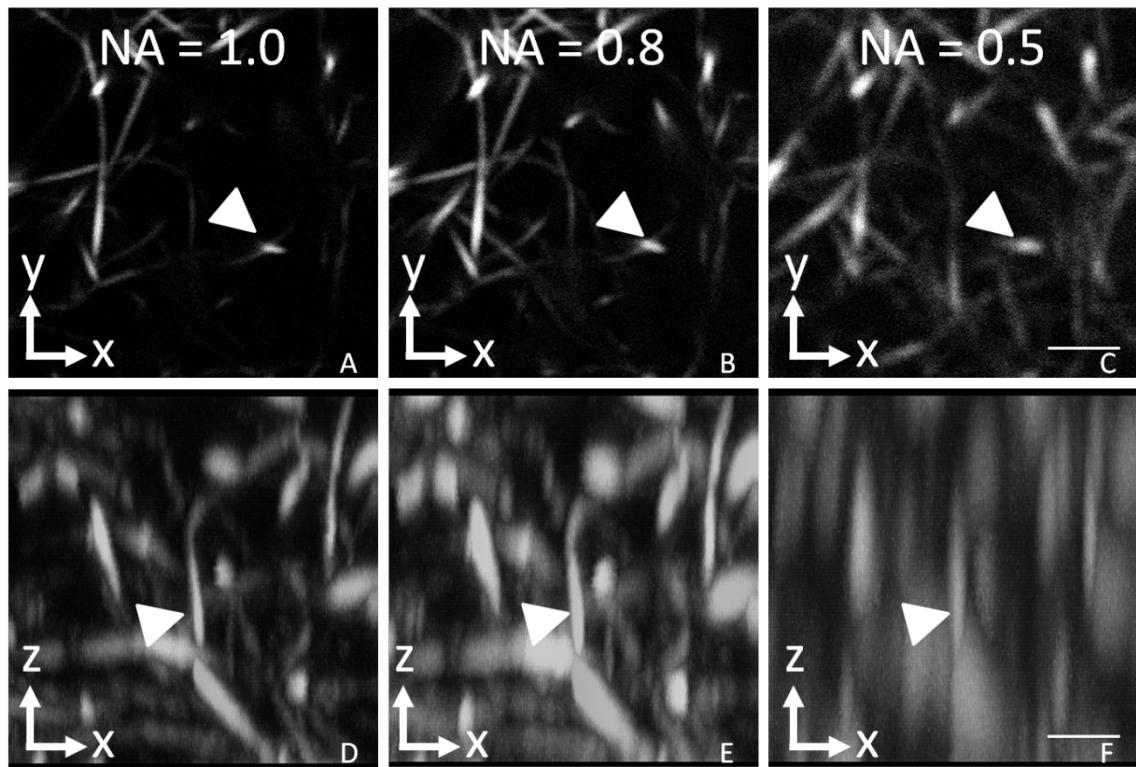


Figure 29. X-Y (A-C) and X-Z (D-F) projections of a 25X25X25 μm volume of an unstained collagen gel captured with forward SHG using 1.0 (A,D), 0.8 (B,E) and 0.5 (C,F) NA. The X-Y projection is a single section from a z-stack while the X-Z projection is a 5 μm thick maximum intensity projection in the Y-dimension. Arrows highlight the presence of axially oriented collagen fibers visible even down to NA 0.5. Scale bar = 5 microns.

5.4 Discussion

Images of collagen fiber networks captured with SHG microscopy depend on many factors including the detection geometry, the orientation of the molecular dipole moments in the fibers themselves and the numerical aperture of the objective lens used for excitation. Experimental measurements have not been performed on how SHG signals change in the forward and backward direction as a function of fiber elevation angle. In addition, how numerical aperture affects detection of axially oriented fibers remains an experimentally untested question. This study is the first, to our knowledge, to use MPFE and SHG in both the forward and backward direction to quantify the fiber detection capabilities of the SHG microscope as a function of fiber angle. Our study shows that fibers with elevation angles above 30 degrees are generally not visible in backward SHG images. On the other hand, forward SHG is capable of detecting fibers of all elevation angles, even down to NA 0.5. We conclude that forward SHG should be used to capture the most accurate representations of collagen gel fiber networks and that even low NA objectives may be used to capture information about axially oriented fibers. These experimental observations are important to consider when designing experiments where biomechanical properties are to be estimated from SHG image data or when assessing 3D collagen alignment differences between tissues.

Our result showing the lack of sensitivity to axially oriented collagen fibers with backward SHG is consistent with previous reports suggesting significant backward SHG axial orientation dependence [13, 70, 158, 188, 238]. Accurate 3D imaging of stained collagen gels has been demonstrated in the epi direction using a confocal microscope [45]. SHG is attractive over this method due to the intrinsic nature of the contrast mechanism. It has also been shown that collagen fibers can produce low levels of intrinsic fluorescence which allows for 3D imaging using MPFE [159, 238]. Unfortunately, generating sufficient fluorescence signal from collagen fibers often requires potentially damaging levels of excitation power. In this study we stained our collagen gels with FITC to allow for high SNR images to be captured with both MPFE and SHG at an excitation power of 10 mW.

Our use of the MPFE imaging as the ground truth for comparison against SHG imaging is supported by reports showing accurate fiber network mechanical property estimates based on fluorescence images of stained collagen gels [45, 243]. Indeed, MPFE of the stained gels was observed here to have an isotropic elevation angle dependence, allowing for equal detection of all 4π fiber orientations. In contrast, backward SHG was unable to detect the majority of the fiber elevation angles present in the gels, therefore should not be relied upon for mechanical property estimates. On the other hand, similar to MPFE, forward SHG signal detection was observed to be independent of elevation angle, therefore could readily be used for assessing details about the 3D collagen microstructure in gels.

The 1 mg/mL gel concentration used in this study was chosen to reduce multiple scattering of forward directed SHG photons in order to elucidate the directionality of the generated SHG signal in our measurements. However, thick tissue imaging studies will observe significant backscattered forward SHG light. It may therefore be expected that some forward directed SHG light from axially oriented fibers would be present in the backward SHG detector due to multiple scattering in thick tissue. In addition, collagen fibers in tissue, particularly in tendon or highly fibrous tissue such as the cervix, often assemble into much larger diameter bundles than the individual fibers captured in the gels used in this study. Large vertically oriented fiber bundles found in tissues may allow for SHG signals to be observed in the backward direction due to differing scattering cross sections.

Previous studies have suggested that high NA excitation and collection is required to capture SHG signals from axially oriented fibers [173, 238]. These studies suggest that high NA is required to introduce axial field components at the focus of the objective. At low NA, such as NA=0.5, the field at the focus can be approximated fairly accurately as a plane wave with no appreciable axial field component [70]. Our observation here that axially oriented fibers are visible with 0.5 NA indicates that the molecular dipole moment of the collagen fibers in our gels must not be oriented along the fiber axis, but at some angle

with respect to the fiber axis. This observation is consistent with recent polarization sensitive SHG measurements made by Tuer et al. [69] who found that dipole moment orientations were measured to be at angles relative to the fiber axis. The observations made here implicate that the lack of backward SHG signal from axially oriented fibers is not due to dipole orientation, but is caused by coherent integration of SHG signals from appropriately spaced collagen molecules in the axial direction only.

5.5 Conclusion

In this chapter we demonstrate the capabilities of the forward SHG microscope for capturing 3D images of collagen fibers. We show that even at a low NA of 0.5, all fiber orientations are represented in forward SHG images. In contrast, we show that collagen fibers disappear from backward SHG images at an elevation angle of 30 degrees and above. Our experimental findings corroborate the theoretical findings of others who predict that axially oriented fibers are not readily detected in the backward direction. However, the low NA detection of vertical fibers raises questions about the need for high numerical aperture for axial fiber detection reported by others. We conclude that forward SHG should be used for the most accurate 3D fiber imaging and analysis in collagen gels. These conclusions should apply to 3D SHG imaging in thin or thick tissue sections as well.

Many recent studies have successfully linked transverse collagen alignment changes to disease progression and survival. Full 3D collagen alignment analysis holds promise for improving the sensitivity and accuracy of these studies by potentially providing more comprehensive information. In trying to extend this work to 3D, it is important to understand the limitations and most appropriate configurations for SHG imaging experiments. The conclusions we draw in this chapter help to guide these future studies.

Chapter 6 Sequential Erosive Tissue Imaging

6.1 Introduction

In general, microscopy is limited to the observation of thin two-dimensional (2D) sections. This is most often accomplished by embedding tissue in wax or tissue freezing media and cutting thin (usually 5 µm) sections from the tissue block using a microtome or cryostat. These sections are then mounted on glass slides and stained with a wide variety of stains but most commonly hematoxylin (which stains nuclei) and eosin (which stains proteins). These stains absorb different bands of the white light spectrum that is transmitted through the thin sections in a standard wide field microscope in order to generate image contrast. This 2D method is the gold standard used in clinical pathology. However, since tissues and organs are inherently three-dimensional (3D) structures, they can be difficult to interpret when viewed as 2D slices. For example, the number and volume of the developing mouse prostate buds required 3D analysis in this study by Timms et. al. [244]. Three dimensional structures can be roughly characterized by manual analysis of serial sections, but this requires significant manual effort in the production of these serially sectioned slides and in the mental registration of information from multiple misaligned 2D images [245]. Computer aided serial section reconstruction has been used for producing cell resolution 3D images [38, 246, 247]. However the process by which tissues are sectioned and mounted on slides can damage tissue architecture, result in section loss and cause section misalignments making automated registration difficult. Alternatively, confocal [248] and MPM [102, 103] are used routinely to capture 3D images of microscopic tissue structures. Both confocal and multiphoton techniques solve the optically section by limiting the depth over which fluorescence detection is made (confocal) or fluorescence excitation is created. These techniques are said to optically sectioning tissue rather than physically section tissue in order to construct 3D images from stacks of 2D images. Unlike serial section reconstruction, these techniques are limited in the depth over which they are able to image in scattering

tissue. For example, Centonze and White [249] demonstrated that the effective imaging depths for fluorescently stained kidney tissues were approximately 30 μm for confocal and 60 μm for multiphoton. These depths can be significantly increased if tissue is cleared with modern clearing agent formulations [250]. Unfortunately, depending on the tissue type, even the best clearing agents are not always effective for deep multiphoton imaging (see Figure 30) and are essentially terminal experiments since normal tissue properties often cannot be recovered.

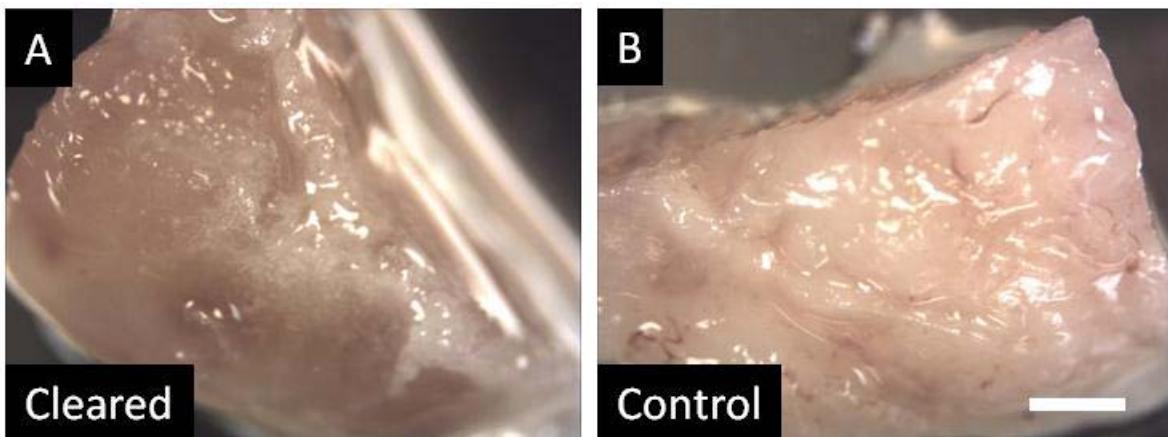


Figure 30. A cleared mouse mammary tissue sample (A) compared to a similar but un-cleared tissue (B). The cleared tissue was incubated in the ScaleA2 formulation for 48 hours, while the control tissue was incubated in phosphate buffered saline for 48 hours. The effective multiphoton imaging depth for the cleared sample was only about 20 microns deeper compared to the un-cleared sample.

Other imaging techniques such as microCT [251], microMR [252], ultrasound biomicroscopy [253], optical projection tomography (OPT) [254], and optical coherence tomography (OCT) [255] have been developed to capture high resolution images of tissue without the need for physical sectioning. Unfortunately, all of these techniques are limited by high cost, low resolution or both when compared to light microscopy.

In this chapter, we focus on a combined light microscopy and physical sectioning technique known as block face imaging. Block face imaging has evolved over several decades from simple point and shoot cameras bolted onto microtomes , to fluorescence microscopy integrated with sledge microtomes [256,

257], to MPM [258] or OCT [259] systems integrated into vibratomes. Perhaps the most famous application for block face imaging is the visible human project [260] which has generated enormous amounts of useful data about human anatomy for educational, diagnostic, treatment planning, artistic and industrial uses. Block face imaging has been applied to many other imaging problems including brain atlas creation [261], embryonic anatomy and development [262], cardiac structural visualization [263], bone microstructure analysis [264] and whole mouse imaging for MR or CT validation [265].

One of the key hurdles to cell resolution block face imaging is that of out of plane blurring, which occurs when light scatters from structures below the surface of the block, reducing the resolution and contrast of the image of the block face. Indeed, out of plane blurring is one of the reasons for sectioning in microscopy in the first place. This issue in microscopy is often called the missing cone problem. If we take the 3D Fourier transform of the microscope point spread function, we see that there is a cone of missing information along the optical axis of the microscope. Confocal, MPM, OPT, structured illumination microscopy (SIM), and selective plane illumination microscopy (SPIM) have all been developed as clever solutions to the missing cone problem. Many computational techniques have also been developed to help solve the missing cone problem, most notably deconvolution microscopy, which attempts to solve the microscopy imaging inverse problem [266] based on a priori knowledge of the point spread function. Block face imaging solves half of the missing cone problem because one half of the point spread function is always outside of the sample. The other half of the problem has been solved in block face imaging by trying to make the sample and block opaque to the excitation light [256, 257] and through post processing deconvolution techniques [264, 267]. Here, we report on a system that combines two approaches to removing out of plane blur from block face images. First, we use a block and tissue opacifier to limit excitation light penetration and second we use deconvolution postprocessing to remove out of plane blurring and reconstruct the 3D volume. We call our technique sequential erosive tissue imaging (SETI).

6.2 Materials and Methods

The SETI system is composed of a set of protocols for sample preparation, instrumentation hardware for sectioning and imaging, and a set of algorithms for image reconstruction and analysis. The materials and methods used in each of these subsystems are described below.

6.2.1 Sample Preparation

SETI provides color bright-field or molecular fluorescence imaging of biological samples. Tissues must therefore be prepared such that contrast for the cells or structures of interest is available. Here we use techniques of in-situ hybridization and immunohistochemistry to label cells with fluorescent dye conjugated antibodies. Tissues are then dehydrated in a graded alcohol bath, embedded in hard resin, and mounted onto a standardized mold platform. The details of each of these sample preparation protocols is given in the sections below.

6.2.1.1 Whole Mount Immunohistochemistry Staining

The protocol used here is based on Keil et. al. [268] and allows for the visualization of epithelial cells in tissue samples up to 1 mm^3 in volume. The details of this protocol are listed in Appendix 2.

6.2.1.2 Embedding Tissue Samples in JB4 Resin

JB4 is a commonly used electron microscopy resin that has a low viscosity allowing for better infiltrations into thicker tissues. The JB4 resin is also water-soluble, therefore can be infiltrated into tissues without the need to dehydrate to absolute alcohol. JB4 is usually clear after polymerization, but we have developed a protocol that mixes india ink into the resin prior to polymerization allowing for the block to turn out opaque. The details of our JB4 embedding protocol are given in Appendix 2.

6.2.1.3 Embedding Tissue Samples in LR White Resin

LR White is also a convenient single part electron microscopy resin that forms hard clear plastic matrices. We started out using LR White, but have transitioned to using JB4 due to its lower viscosity

and improved infiltration qualities. The details of our LR White resin embedding protocol are given in Appendix 2. We have also developed a protocol for adding zinc oxide to LR to create opaque tissue blocks. One of the main challenges to using LR White is the need to keep oxygen away from the block during polymerization.

6.2.2 Imaging and Erosion Instrumentation

The prototype SETI imaging and erosion system consists of an integrated three axis mill, wide field fluorescence microscope, and oblique reflectance macroscope. The block diagram of this system is shown in Figure 31A and a photograph of the prototype is shown in Figure 31B.

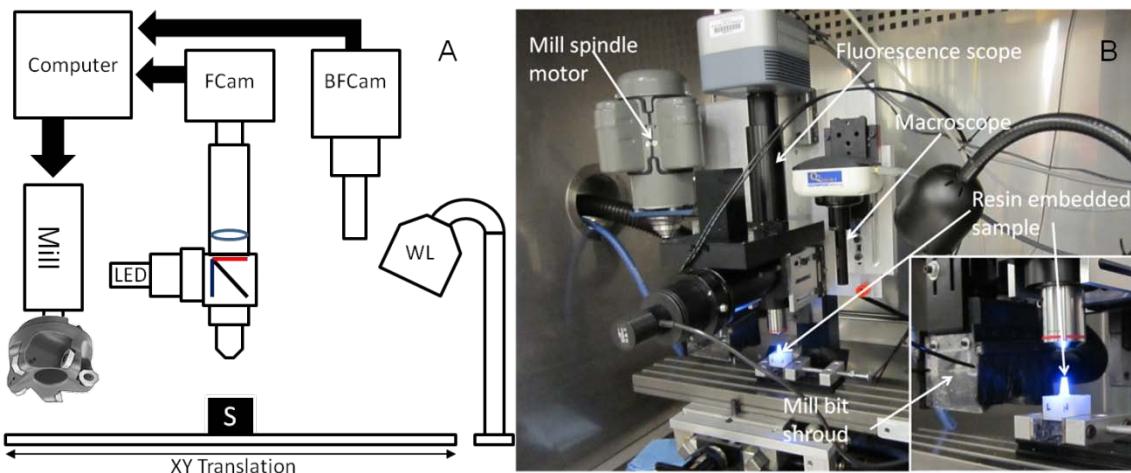


Figure 31. Block diagram of the SETI hardware (A) and a photograph of the prototype system (B) with inset zoomed in on the resin embedded sample. FCam = Fluorescence camera, BFCam = Bright field camera, WL = White light lamp, LED = light emitting diode lamp, S = sample.

We use a modified MicroMill DSLS 3000 three-axis mill (Microproto Systems, Chandler, AZ) for motion control. The mill spindle is equipped with a replaceable insert face mill (McMaster-Carr, Elmhurst, IL) with 5 octagonal inserts. A HEPA vacuum and compressed air system are designed into a 3D printed plastic shroud around the face mill that serves to remove dust from the surface of the sample throughout the cutting process. The fluorescence imaging system consists of a Modular Infinity Microscope (MIM, Applied Scientific Instrumentation, Eugene, OR) through which images are captured with an Andor Clara interline CCD camera (Andor, Belfast, UK). We use a 4X, 0.2NA infinity corrected

objective (Nikon, Japan) as the primary imaging lens and the tube lens is a proprietary 200 mm focal length achromat from Nikon. For imaging fluorophores with excitation and emission spectra similar to fluoresceine such as Alexa 488 and Dylight 488, we use a ThorLabs 470nm collimated LED as the source (ThorLabs, Newton, NJ), and a 49002 filter set (Chroma, Bellows Falls, VT). Our macroscopic imaging system consists of a 94 mm working distance 1.00X magnification Infinistix lens (Infinity-USA, Boulder, CO), a QImaging RGB camera (QImaging, Surrey, BC, Canada) and a white light halogen lamp. The spindle, spindle motor, MIM, and macroscope are all mounted to the vertical axis of the MicroMill using a custom aluminum mounting plate. The imaging systems are focused at the vertical position of the mill cutter by first removing a layer from the sample block, moving the sample block into the field of view of the imaging system, then translating the imaging system relative to the mill spindle using manual translation stages. After this alignment is performed one time, the sample surface then remains in focus after each subsequent layer is removed. The three-axis positioning system and spindle motor are controlled via Mach3 CNC software. Images are captured from the Andor Clara through MicroManager and images from the QCamera are captured using the QCapture software. Automated sectioning and image captured is achieved using the freeware AutoIT Windows GUI scripting tool. This tool allows the user to script repetitive Windows GUI tasks and was used as a simple way to integrate and automate the three executables. The next iteration of the hardware and software should eliminate Windows GUI scripting and integrate all functionality into a single, cross platform executable.

6.2.3 Image Reconstruction

The missing cone problem causes out of plane light to obscure the fine details of the structures at the surface of the block face and causes objects below the surface of the block face to appear as if they are at the surface. Some of the out of plane light is reduced by our block staining protocol. The effects of the rest must be removed computationally. Here we describe a computational method for reducing out of plane blurring and demonstrate its performance on simulated data. The intensity of the excitation light

of a fluorescence microscope as a function of depth z and position x can be modeled as a radially symmetric Gaussian beam given by

$$p(x, z) = \frac{I_0}{w(z)^2} \exp\left(\frac{-2x^2}{w^2(z)}\right)$$

where $w(z)$ is

$$w(z) = w_0 \sqrt{1 + \left(\frac{z}{z_r}\right)^2}$$

and represents the beam radius as a function of z where w_0 is the narrowest point on the beam and $z_r = \frac{\pi w_0^2}{\lambda}$ corresponds to the Rayleigh range. Given the numerical aperture NA of the focusing lens, the beam waist w_0 can be approximated by the following.

$$w_0 = \frac{0.6\lambda}{NA}$$

Since the point spread function (PSF) $p(x, z)$ is radially symmetric, we can model the system in 2D without loss of generality. The 2D modulation transfer function (MTF) of the standard microscope is given by the Fourier transform below.

$$P(k_x, k_z) = \iint p(x, z) \exp(-i(k_x x + k_z z)) dx dz$$

Models of the PSF and the MTF for the standard light microscope are shown in Figure 32A and Figure 32B for a 0.2 NA objective as is currently used on the SETI prototype. This figure reveals that the PSF is significantly taller along the optical axis than in the lateral direction. This causes the MTF to be shaped like a disc in the imaging plane, with all high frequency information along the optical axis being attenuated. This is the cause of the out of plane blurring characteristic of a standard light microscope. The PSF for SETI is a modified version of $p(x, z)$ due to physical sectioning and is given by the following

$$p_S(x, y) = \begin{cases} p(x, y) & z \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Physical sectioning as performed by SETI provides a slightly improved PSF and MTF as shown in Figure 32C and Figure 32D.

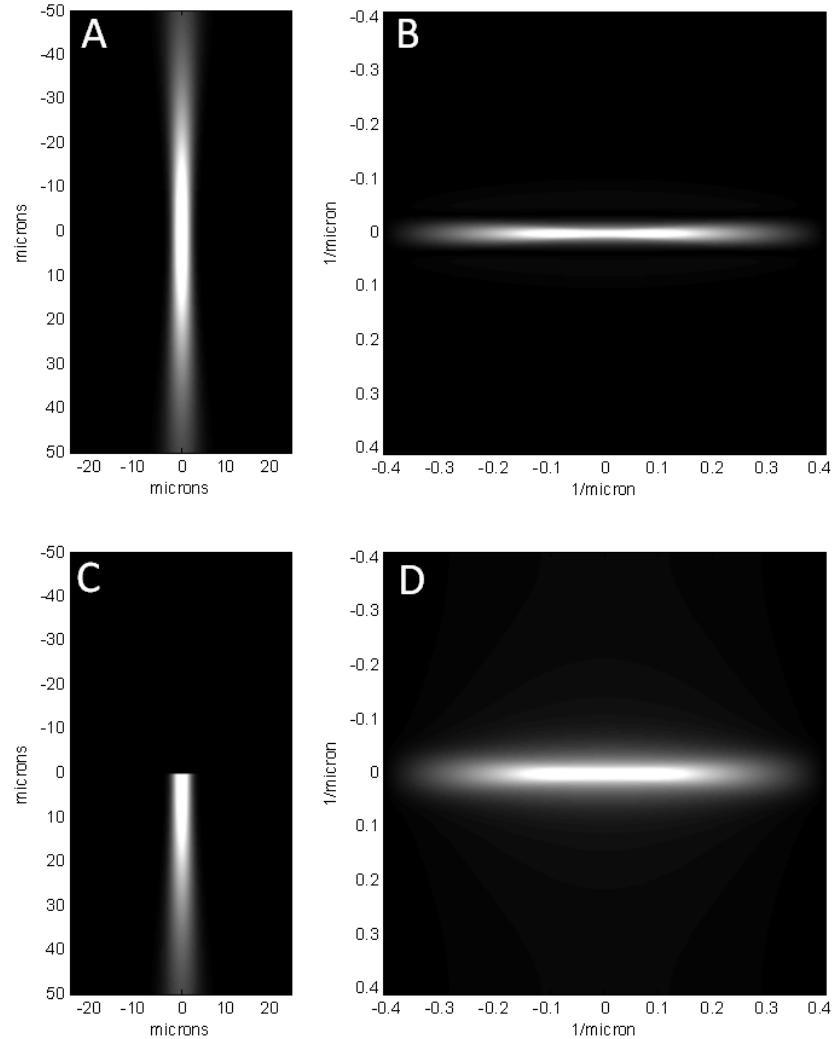


Figure 32. Standard light microscope PSF (A) and MTF (B) demonstrating the missing cone problem for a 0.2 NA objective lens. The vertical direction in these figures is the optical axis of the microscope. There is only a sliver of Fourier space that is represented by the data captured by a standard light microscope. The SETI PSF (B) and MTF (D) are modified by the physical sectioning feature of the system. Physical sectioning adds high frequency data along the optical axis of the microscope, improving the missing cone problem.

The standard wide field fluorescence microscope image $k(x, z)$ can be modeled by convolving $p(x, z)$ with the object $s(x, z)$.

$$k(x, z) = p(x, z) \otimes s(x, z) = \iint p(\tau_1, \tau_2)s(x - \tau_1, z - \tau_2)d\tau_1d\tau_2$$

For the SETI imaging forward problem, we replace $p(x, z)$ in the above equation with $p_S(x, y)$. We model our object as a low concentration mixture of subresolution fluorescent beads randomly positioned within a block of resin. The object and corresponding image are shown in panels A and B of Figure 33.

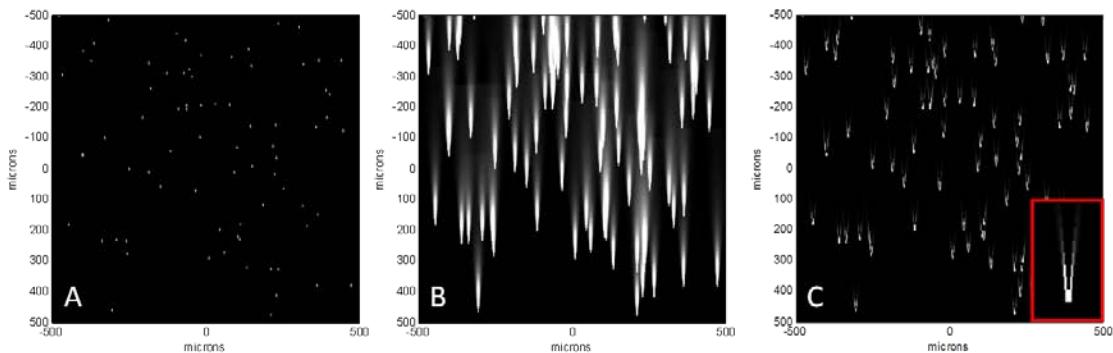


Figure 33. Model of SETI image formation in a clear sample. The simulated object (A) is a model of a low concentration mixture of subresolution fluorescent beads suspended in resin. The raw SETI image (B) is produced by convolving the PSF with the simulated object and decimating in the z-dimension to the z-step size of the microscope. The z-dimension blurring is corrected in C by a simple subtraction algorithm that subtracts image i from the next image $i+1$. The inset in C shows the artifact left by this simple algorithm.

With this model, we have demonstrated that out of plane blurring is clearly a major problem for SETI, since we can expect the raw data from the microscope to look like Figure 33B. We can reduce out of plane blurring by simply subtracting the current image from the next image. Given image I_i where $i \in \{1, z_r\}$ and z_r is the number of images in the stack, subtracting each image from the next subsequent image we have

$$I_i = I_i - I_{i+1} \forall i \in \{1, z_r\}$$

This operation has been performed and the result is shown in Figure 33C. The subtraction effectively performs a high pass filter in the z-direction, attempting to highlight the high frequency step present in the SETI PSF. Although the majority of the contribution from subsequent layers is removed by this operation, there remains a residual chevron shaped artifact above each simulated fluorescent bead as shown in the inset to Figure 33C. This is due to the increased lateral width of the PSF away from the focal plane. In order to fix this problem, we must first filter image I_{i+1} in the lateral direction prior to subtraction. This operation is given below

$$I_i = I_i - (h(w_0) \otimes I_{i+1}) \forall i \in \{1, z_r\}$$

where $h(w_0)$ is a Gaussian filter of width equal to the beam waste governed by the NA of the objective lens.

6.3 Results and Discussion

Sequential erosive tissue imaging (SETI) is able to image deep within tissue while maintaining high sensitivity and high resolution by physically sectioning while simultaneously imaging each newly revealed surface. Physical sectioning allows the imaging system to overcome depth related attenuation and scattering and provides a high frequency step in the PSF of the imaging system. This high frequency component in the PSF provides the opportunity for simple deconvolution techniques to be applied to reduce out of plane blurring. In the previous section, we have applied a simple subtraction based algorithm and compared it to the raw unprocessed data. In this section, we compare the simple subtraction to the more complex algorithm that filters image I_{i+1} prior to subtraction described in the previous section. Our simulated results are shown in Figure 34, where panel A shows the raw data, B shows the result of our simple subtraction algorithm, while panel C shows the results of the filtered subtraction algorithm. The filtered subtraction technique reduces the chevron artifacts above and to the

side of each bead, but produces periodic artifacts directly above the beads. Overall, the filtered subtraction technique is considered to be an improvement over the simple subtraction technique based on a visual inspection of the results shown in Figure 34.

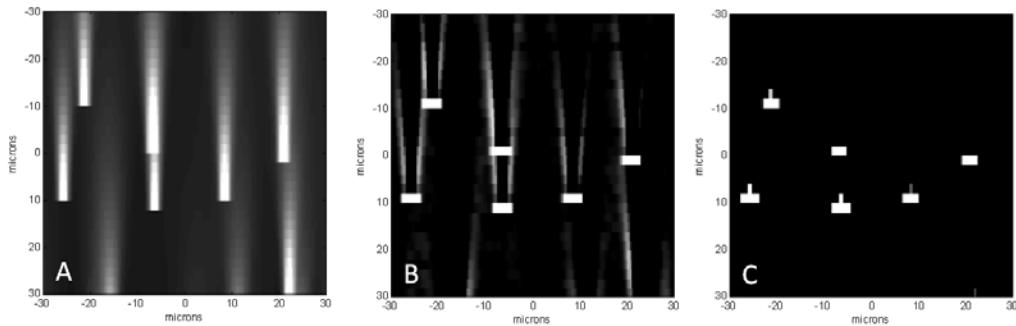


Figure 34. Simulated fluorescent beads in plastic. The simulated raw data (A) is compared to a simple subtraction based reconstruction algorithm (B) and a filtered subtraction algorithm (C).

To experimentally verify our imaging system and reconstruction algorithms, we have embedded fluorescent beads in plastic resin dyed with india ink. The resulting x-z projection of the sample before and after our two reconstruction techniques are shown in Figure 35.

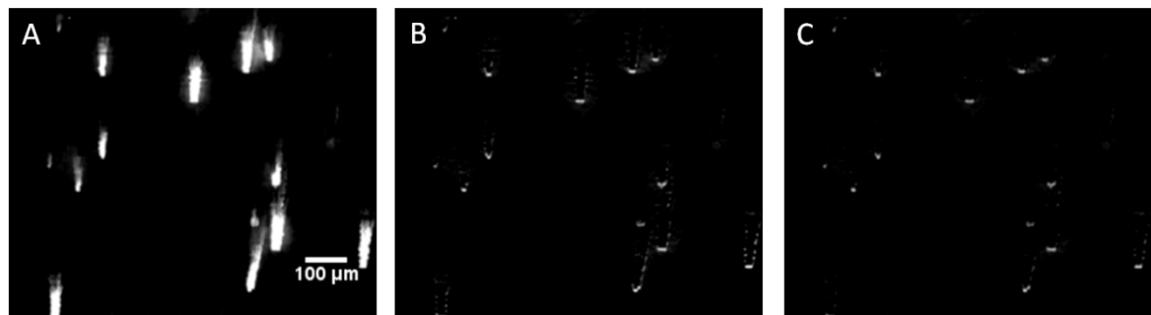


Figure 35. 6 micrometer diameter fluorescent beads embedded in resin and imaged with SETI. The raw data (A) is processed with the simple subtraction filter (B) and the filtered subtraction algorithm (C). These experimental results are qualitatively similar to our simulation results presented above.

We have tested our prototype on a small number of biological samples with the goal of reconstructing the 3D architecture of the vasculature in the regenerating limb of the axolotl and quantifying the number and volume of the developing prostate buds in a fetal mouse model. Preliminary data on these

two projects are shown in Figure 36A and Figure 36B. The axolotl limb image shows the bony structures in the limb since no contrast was introduced into the vasculature during this experiment. We are currently working on developing the protocol for adding vascular contrast. The fetal mouse bladder and prostate epithelial cells shown in Figure 36B where stained via the IHC protocol given in a previous section of this chapter and imaged with SETI. Both images were reconstructed with the filtered subtraction algorithm given above.

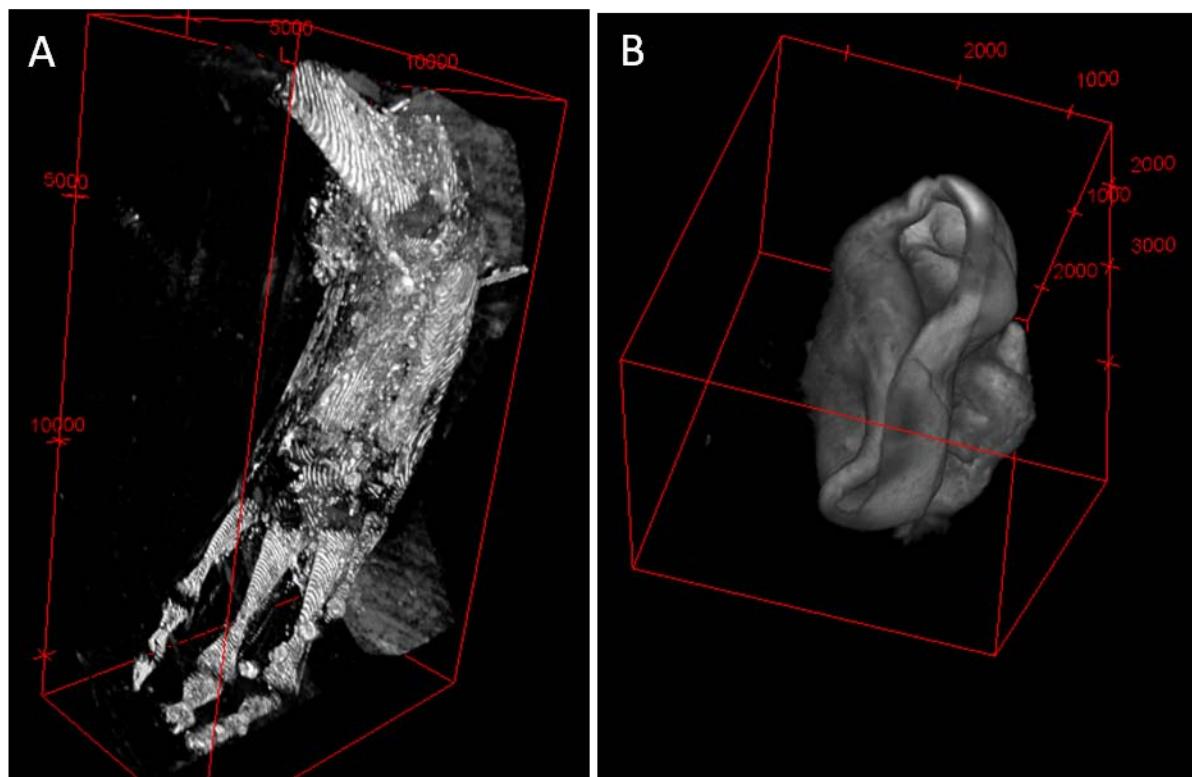


Figure 36. A 3D reconstruction of the bony structures within the regenerated limb of an axolotl (A) and the folds of a fetal mouse bladder with the developing prostate on the lower right (B).

SETI has great potential for providing a low cost 3D imaging solution that fits in the resolution and field of view gap between microscopy and small animal imaging. There are still many hardware design and algorithm development challenges that must be overcome before this system is fully functional. The work presented here represents a significant step toward this goal.

Chapter 7 Overall Conclusions and Future Outlook

Optical imaging provides a vast arsenal of tools and techniques for investigating biological systems, healthy and diseased. From the standard light microscope for visualizing tissue on slides to confocal and multiphoton microscopy for visualizing whole mount tissues in 3D, from serial section reconstruction to sequential erosive tissue imaging, each of these techniques provide cell resolution representations of tissues that allow for key structural information to be gathered. Perhaps as important as the hardware itself is the integration of this hardware with open access computational tools directed at mining and applying the collected image information. In this dissertation, we have presented a number of key hardware and computational developments that are focused on 1) extending the field of view of optical imaging allowing for the more thorough and consistent collection of data and 2) mining the results of these large data collections for patterns that are relevant to disease outcome and treatment response.

In Chapter 2, we demonstrated that the curvelet transform forms a convenient foundation to many useful tools for analyzing SHG images of collagen. The curvelet transform combined with fiber tracking algorithms was shown in Chapter 3 to more accurately segment fibers from SHG images of collagen in comparison to other algorithms. Chapter 4 applied our curvelet transform tools in combination with cell segmentation techniques showing a correlation between automatically extracted collagen alignment features and breast cancer patient survival. In Chapter 5, we focused on extending this work to 3D and experimentally tested some of limiting physical properties of SHG imaging. And finally, in Chapter 6, we presented our preliminary work on SETI, a low cost 3D optical imaging system that removes the depth limitations present in traditional microscopy techniques. Each of these techniques have been successfully applied to solve specific problems. The test of their ultimate utility will be in the transferability of the tools to other scientists performing similar studies. Future effort will therefore be placed in the convenient packaging and presentation of these tools and techniques for more wide dissemination among the community. For our CurveAlign and CT-FIRE software, these tools must be

developed to communicate more seamlessly with one another and with FIJI/ImageJ. They both have users manuals and software posted online, however screencasts and tutorials for generating data similar to that presented in this dissertation will be helpful to users trying to use these tools in the future. The hardware platforms developed as part of this dissertation, CAMM and SETI, should continue development as well, particularly on improving the usability of these instruments. This will involve further streamlining imaging protocols and simplifying user interfaces.

Appendix 1 Hardware Bill of Materials

CAMM Parts List			
Description	Supplier	Part No	Unit Price
Base			
Scope Base	ASI	RAMM	~
XY Stage	ASI	MS2000	~
Z Motor	ASI	MIM-RA-FOCUS-K	~
Piezo Z Drive	ASI	Piezo Z	~
Tube For White Light	ASI	C-60-Tube-B	~
Motor Controller	ASI	RM2000	~
Cube Holder	ASI	MIM-Cube-II-K	~
Excitation			
Laser	Coherent	Mira 900	~
Laser Shutter	Uniblitz	VS14S2ZM1	550
Laser Shutter controller	Uniblitz	VMM D4	1865
Pockel's Cell	Con Optics	350 80-LA	4200
Pockel's Cell Mount	Con Optics	102	700
Pockel's Cell Alingment Tool	Con Optics	103	85
Pockel's Cell Controller	Con Optics	302rm	3475
XY Galvos	Cambridge	6210	~
XY Galvo Mounting Block	Cambridge	61021003R	~
Galvo Servos	Cambridge	671	~
80 inch Galvo Cables	Cambridge	6010-22-080 F-LF	50
Beam Expander	Thor	BE-02-05-B	1100
Beam Expaner Mount	Thor	KS2-02	125
Achromatic 1/2 wave plate	Thor	AHWP05M-980	783
Achromatic 1/4 wave plate	Thor	AQWP05M-980	783
Rotation mount	Thor	CRM1	88
Tube Lense	Thor	ITL200	450
Scan Lense	Thor	LA1145-B	39.8
Dichroic	Semrock	IX 71 Style, FF720-SDi01-25X36	965
Dichroic Holder	Semrock	U-MF2	550
20X Objective	Nikon	CFI Plan Apo Lambda 20X	~
Detection			
Detector shutter	Uniblitz	VS25S2T0-106	765
Detector	Hamamatsu	7422-40P	4309
Detector control module	B&H	DCC100	2842

Detector cable	B&H	C300 DCC H74XL	295
Integrating Amplifier	Hamamatsu	C7319	~
PC	Super Logics	SL-4U-SBC-Q57-HA	2164
Rack	Starcase	RF 1625-19" Wide, 25" Deep, 16RU High Rack Kit	238.6
Rack Shelf	Starcase	RF Shelf19	30.85
Casters for new rack	Starcase	RFCASTERS-3"	30.85
DSP Board	Innovative	M62/67	~
D/A	Innovative	DAC40	~
A/D	Innovative	A4D1	~
Attenuator for adjusting D/A signal	John Peterson	Placed order through Bock system	~
Hardware for PC control of Pockel's Cell	NI	USB-6008	295
DSP Breakout 1	Newark	90F7874	83.02
Din Rail	Newark	19C7295	4.31
15 pin breakout	Newark	90F7883	68.23
15 pin cable	Newark	23T5431	32.5
Asphere Condenser lens	Thor	ACL2520-A	25.6
Dichroic holder Complex	Thor	DFM	295
SM1 (ext) to C (int)	Thor	SM1A10	19.75
C (ext) to SM1 (int)	Thor	SM1A9	19.75
4 inch lens tube	Thor	SM1L40	44
2 inch lens tube	Thor	SM1L20	16.5
0.5 inch lens tube	Thor	SM1L05	12.59
0.3 inch lens tube	Thor	SM1L03	12.16
1 inch lens tube	Thor	SM1L10	14.25
2 inch lens tube spacer	Thor	SM1S20	14
Graduated Iris	Thor	SM1D12C	92
SM1 Retaining ring	Thor	SM1RR	4.5
Frosted Condenser	Thor	ACL2520-DG6-A	25.6
Adjustable lens tube	Thor	SM1V05	29.6
SM1 Coupler 0.5"	Thor	SM1T2	19.5
Cold White LED	Thor	MCWHL2	187.5
LED Driver	Thor	LEDD1B	284
LED Driver Power Supply	Thor	TPS001	25
Short pass dichroic	Chroma	FF670-SDi01-25x36	549
Servo Control Box			
Anti ripple filter	Vicor	V1-Ram-E1	49

Galvo Project Box	Digikey	HM999-ND	164.47
Power switch and plug	Digikey	Q340-ND (plugtype: IEC320C13)	8.11
Heat sink	Digikey	345-1052-ND	14.4
12V DC-Power Supply	Allied	744-0336	34.94
Fans	Digikey	1053-1239-ND	10.26
BNC connectors	Digikey	ARF1064-ND	4.45
Terminal block	Digikey	WM5765-ND	4.75
Terminal block jumpers	Digikey	3-1437644-7-ND	0.79
LED	Digikey	67-1061-ND	0.53
LED holder	Digikey	67-1330-ND	0.376
8 pin connectors	Digikey	WM2006-ND	0.77
2 pin connectors (input)	Digikey	WM 18874-ND	0.62
Connector pins	Digikey	WM17406-ND	0.308
2 pin connector (output)	Digikey	WM18873-ND	0.5
Cambridge connector kit	Cambridge	671CK	~
Galvo Power Supply	Vicor	FlatPac VI-PULL-EX Output 1: 28V-75W, Output 2: 28V-75W	350
Tools Etc.			
IR Viewer	Find-r-scope	84499A-5	1,795.0 0
Power meter	Newport	1917-R	890
Power meter head	Newport	818P Thermopile	901
basic x-y stage	Thor Labs	PT1M	264
Metric screw kit	Thor Labs	HW-KIT2/M - M6 Cap Screw and Hardware Kit	106.1
Metric ball driver set	Thor Labs	BD-Kit/M	26.92
Long Metric driver	Thor Labs	BD-5ML	8.2
Metric tool caddy	Thor Labs	TC4/M	65.2
Standard tool caddy	Thor Labs	TC4	70.4
XY Stage for Galvo Positioning	Thor	DT25	128
XY Stage Riser Plate	Thor	DT25A	21
Galvo vert stage	Thor	mVS005	595

Appendix 2 SETI Protocols

Whole Mount Immunohistochemistry Staining Protocol

Sample Fixation

1. Fix Samples overnight at 4C in 4% buffered paraformaldehyde (US Biologicals #10043-1LT).
2. Dehydrate at 25C in MeOH/TBSTx (0.1% Triton-X 100) see recipe:
 - 10 min in 25% MeOH
 - 10 min in 50% MeOH
 - 10 min in 75% MeOH
 - 10 min in 100% MeOH (Repeat 2X and store at -20(C)

Day 1 IHC

1. Incubate 15 min at 4C **on orbital shaker** in 100% MeOH containing 6% H₂O₂. Vigorous rocking necessary to cause samples to move within microfuge tube.
2. Rehydrate at 25C: Vigorous rocking.
 - 10 min in 100% MeOH
 - 10 min in 75% MeOH + 25% TBSTx
 - 10 min in 50% MeOH + 50% TBSTx
 - 10 min in 25% MeOH + 75 % TBSTx
 - Rinse 2X 5 min in TBSTx
3. Incubate 20 min **at -20C** in **(pre-chilled)** 100% acetone (Sigma 534064-4L).
4. Rinse 2X 5 min at 25C in TBSTx.
5. Incubate 25 min at 25C in **TBSTx** containing *Clostridium histolyticum* collagenase (1mg/ml final collagenase conc., diluted from Sigma #C9891-100mg). **On orbital shaker (speed setting 3.5). Samples should be swishing about within tube and not be at the base of the tube.**
6. Rinse quickly in TBSTx.
7. Rinse 3X 5 min at 25C in TBSTx.
8. Incubate 20 min at 25C in 4% buffered paraformaldehyde.
9. Rinse 2X 5 min at 25C in TBSTx.
10. Block tissues minimum 1h at 25C in RGBTx (**containing 0.5% Triton X and 1% DMSO**)
11. Incubate for a minimum of 24h at 4C in RGBTx (containing **0.1% TritonX and no DMSO**) with primary antibody.

Day 3 IHC

12. Rinse 5X 60 min at **4C on shaker** in TBSTx.
13. Incubate overnight at 4C in RGBTx (0.1% Triton X with **no** DMSO) containing secondary antibody containing biotinylated secondary antibody or fluorophore conjugated secondary antibody

Day 4 IHC

14. rinse 7X 25 min at **4 C on shaker** in TBSTx
For biotinylated secondary antibodies, skip to step 18.
For fluorophore-conjugated secondary antibodies:

15. Incubate 10 min in DAPI solution (300 µM aliquots are available in -20°C freezer, diluted from Invitrogen #D3571). Skip to step # 25

For biotinylated antibodies:

16. Prepare streptavidin conjugate. We have two forms. Select A or B below (both work well).

- a. Prepare, during step #16, 1X avidin biotin complex (ABC) reagent (Vectastain Elite ABC Kit (standard), Vector labs Cat. No. PK6100):
 - 1ml TBSTx
 - 8ul Reagent A
 - 8ul Reagent B
 - Place ABC reagent in a tube and incubate for 45 min at 25°C on rocker.
- b. OR prepare, during step #16, diluted streptavidin-HRP. Add 0.25µl streptavidin-HRP (Jackson Immunoresearch Cat. No. 016-030-084) per 1 mL TBSTx). Unlike ABC reagent above, no incubation period is required for the diluted streptavidin-HRP.

17. Incubate 45 min at 25°C in ABC reagent or streptavidin-HRP.

18. Rinse 5X 15 min at 25°C in TBSTx.

19. Prepare, during the last wash, 1X DAB solution from the DAB substrate kit for peroxidase (vector labs SK-4100):

- a. 1 ml TBSTx
- b. + 8ul Buffer Stock Solution and Mix
- c. + 8ul DAB stock solution and Mix

20. Incubate tissues in 1X DAB solution for 10 min with shaking.

21. Add 8 µl Hydrogen Peroxide Solution (from the DAB kit) per 1 mL 1X DAB solution.

Monitor tissues under microscope for color development (appr. 1-10 min).

22. Rinse 5X 10 min at 25°C in TBSTx.

JB4 Embedding Protocol

1. Make embedding solution in 2 or 5 mL amounts depending on the number of samples or the size of the samples needed.
 - a. For 2 mL, use 2 mL of Solution A + 25 mg catalyst
 - b. For 5 mL, use 5 mL of Solution A and 62.5 mg catalyst
2. Combine Solution A and catalyst in a conical tube
3. Mix on rotating plate for 20 min or until catalyst is completely dissolved
4. Add 30 uL india ink/1000 uL of solution
5. Mix well (quickly invert a few times)
6. If adding fluorescent beads, add at this time (e.g. 20 uL of 1/100 dilution 6 um diameter Y-G beads)
7. Add 50 uL accelerator/1000 uL solution
8. Mix well (quickly invert a few times)
9. Add 1000 uL to small rectangular molds and allow to set for 30 s - 1 min
10. Cover with plastic caps
11. Fill in openings around the edges with paraffin wax
12. Refrigerate for at least 2 hours or overnight to allow to harden

LR White Embedding Protocol

1. For mouse prostate sample use a total Volume of 4ml.
2. Put on rotating plate to mix for the designated time in parenthesis next to the solution.
3. After the designated time change the solution to the next one in the order given below.
4. 50% ethanol in PBS (15 mins)
5. 70% ethanol in PBS (15 mins)
6. 90% ethanol in PBS (15 mins)
7. 100% ethanol (15 mins)
8. 50% LR White Resin in Ethanol (1 hr)
9. 2x 100% LR White Resin
 - a. First time 30 mins, second time 15 mins
 - b. *When changing solutions either mix the new solutions before adding the tissue sample, or add PBS first then ethanol to the sample (or Ethanol first in the LR White Resin/Ethanol mix).
10. Remove as much excess LR white from the sample as possible.
11. Add a few drops of LR white + Zinc Oxide to a 0.5mL eppendorf tube.
12. Add the sample and flick to combine LR White and LR White + Zinc Oxide
13. Fill to the tube to 0.2mL with LR White + Zinc Oxide and add 2µL of LR White activator, and Flick a few times to mix.
14. Fill the 0.5mL tube the rest of the way with LR White + Zinc Oxide so that it is completely filled (including the well on the cap). Add 3µL of LR white activator. If any LR white +Zinc Oxide has spilled, refill it to minimize any air bubbles and close the top. Use a paper towel to wipe off any excess Resin from the outside of the tube.
15. Invert the tube a few times to mix the activator into the solution.
16. Put on ice and allow the resin to harden.

LR White Resin + Zinc Oxide mixture

- 0.8g Zinc Oxide
 - 10ml LR White Resin
1. Combine Zinc Oxide and LR White Resin in a 15mL conical tube.
 2. To mix thoroughly, use a combination of vortexer and sonicator.
 3. Mixing should take about 5 minutes.
 4. Excess LR White + Zinc Oxide can be kept in a refrigerator, but must be mixed again before the next use.

Bibliography

- [1] M. W. Conklin, J. C. Eickhoff, K. M. Riching, C. A. Pehlke, K. W. Eliceiri, P. P. Provenzano, *et al.*, "Aligned collagen is a prognostic signature for survival in human breast carcinoma.," *Am J Pathol*, vol. 178, pp. 1221-32, Mar 2011.
- [2] P. P. Provenzano, K. W. Eliceiri, and P. J. Keely, "Shining new light on 3D cell motility and the metastatic process.," *Trends Cell Biol*, vol. 19, pp. 638-48, Nov 2009.
- [3] P. P. Provenzano, K. W. Eliceiri, J. M. Campbell, D. R. Inman, J. G. White, and P. J. Keely, "Collagen reorganization at the tumor-stromal interface facilitates local invasion.," *BMC Med*, vol. 4, p. 38, 2006.
- [4] J. Ferlay, H. Shin, F. Bray, D. Forman, C. Mathers, and D. Parkin. (2010, 04/10/2012). *Globocan 2008 v1.2, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10*. Available: <http://globocan.iarc.fr>
- [5] J. G. van den Tweel and C. R. Taylor, "A brief history of pathology," *Virchows Archiv*, vol. 457, pp. 3-10, Jul 2010.
- [6] D. C. Gritz, I. J. Bigio, T. R. Loree, T. Shimada, R. D. Glickman, and K. Storyheld, "THE OPTICAL BIOPSY SYSTEM - A SPECTRAL-ANALYSIS OF THE OCULAR SURFACE INVIVO," *Investigative Ophthalmology & Visual Science*, vol. 34, pp. 942-942, Mar 15 1993.
- [7] G. J. Tearney, M. E. Brezinski, B. E. Bouma, S. A. Boppart, C. Pitris, J. F. Southern, *et al.*, "In vivo endoscopic optical biopsy with optical coherence tomography," *Science*, vol. 276, pp. 2037-2039, Jun 27 1997.
- [8] T. D. Wang and J. Van Dam, "Optical Biopsy: A New Frontier in Endoscopic Detection and Diagnosis," *Clinical Gastroenterology and Hepatology*, vol. 2, pp. 744-753, Sep 2004.
- [9] D. G. Ouzounov, D. R. Rivera, W. O. Williams, J. A. Stupinski, T. L. Southard, K. H. Hume, *et al.*, "Dual modality endomicroscope with optical zoom capability," *Biomedical Optics Express*, vol. 4, pp. 1494-1503, Sep 1 2013.
- [10] C. Xu, W. Zipfel, J. B. Shear, R. M. Williams, and W. W. Webb, "Multiphoton fluorescence excitation: New spectral windows for biological nonlinear microscopy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, pp. 10763-10768, Oct 1 1996.
- [11] B. R. Masters, D. W. Piston, and W. W. Webb, "3-dimensional nad(p)h redox imaging of the insitu cornea with 2 photon excitation laser scanning microscopy," *Investigative Ophthalmology & Visual Science*, vol. 34, pp. 1402-1402, Mar 15 1993.
- [12] D. W. Piston, B. R. Masters, and W. W. Webb, "3-dimensionally resolved nad(p)h cellular metabolic redox imaging of the in-situ cornea with 2-photon excitation laser-scanning microscopy," *Journal of Microscopy-Oxford*, vol. 178, pp. 20-27, Apr 1995.
- [13] W. R. Zipfel, R. M. Williams, R. Christie, A. Y. Nikitin, B. T. Hyman, and W. W. Webb, "Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation.," *Proc Natl Acad Sci U S A*, vol. 100, pp. 7075-80, Jun 2003.
- [14] M. C. Skala, K. M. Riching, A. Gendron-Fitzpatrick, J. Eickhoff, K. W. Eliceiri, J. G. White, *et al.*, "In vivo multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, pp. 19494-19499, Dec 4 2007.
- [15] J. Vergen, C. Hecht, L. V. Zholudeva, M. M. Marquardt, R. Hallworth, and M. G. Nichols, "Metabolic Imaging Using Two-Photon Excited NADH Intensity and Fluorescence Lifetime Imaging," *Microscopy and Microanalysis*, vol. 18, pp. 761-770, Aug 2012.

- [16] I. Freund, M. Deutsch, and A. Sprecher, "Connective-tissue polarity - Optical 2nd-harmonic microscopy, crossed-beam summation and small angle scattering in rat-tail tendon," *Biophysical Journal*, vol. 50, pp. 693-712, Oct 1986.
- [17] W. Wang, J. B. Wyckoff, S. Goswami, Y. Wang, M. Sidani, J. E. Segall, *et al.*, "Coordinated regulation of pathways for enhanced cell motility and chemotaxis is conserved in rat and mouse mammary tumors.," *Cancer Res*, vol. 67, pp. 3505-11, Apr 2007.
- [18] M. Egeblad, M. G. Rasch, and V. M. Weaver, "Dynamic interplay between the collagen scaffold and tumor evolution.," *Curr Opin Cell Biol*, vol. 22, pp. 697-706, Oct 2010.
- [19] N. F. Boyd, L. J. Martin, M. J. Yaffe, and S. Minkin, "Mammographic density and breast cancer risk: current understanding and future prospects.," *Breast Cancer Res*, vol. 13, p. 223, Nov 2011.
- [20] P. J. Keely, "Mechanisms by which the extracellular matrix and integrin signaling act to regulate the switch between tumor suppression and tumor promotion.," *J Mammary Gland Biol Neoplasia*, vol. 16, pp. 205-19, Sep 2011.
- [21] Y. P. Guo, L. J. Martin, W. Hanna, D. Banerjee, N. Miller, E. Fishell, *et al.*, "Growth factors and stromal matrix proteins associated with mammographic densities," *Cancer Epidemiology Biomarkers & Prevention*, vol. 10, pp. 243-248, Mar 2001.
- [22] N. F. Boyd, H. Guo, L. J. Martin, L. Sun, J. Stone, E. Fishell, *et al.*, "Mammographic density and the risk and detection of breast cancer," *New England Journal of Medicine*, vol. 356, pp. 227-236, Jan 18 2007.
- [23] P. P. Provenzano, K. W. Eliceiri, J. M. Campbell, D. R. Inman, J. G. White, and P. J. Keely, "Collagen reorganization at the tumor-stromal interface facilitates local invasion.," *BMC Med*, vol. 4, pp. 1-16, 2006.
- [24] P. P. Provenzano, K. W. Eliceiri, L. Yan, A. Ada-Nguema, M. W. Conklin, D. R. Inman, *et al.*, "Nonlinear optical imaging of cellular processes in breast cancer.," *Microsc Microanal*, vol. 14, pp. 532-48, Dec 2008.
- [25] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, H. E. Beggs, and P. J. Keely, "Mammary epithelial-specific disruption of focal adhesion kinase retards tumor formation and metastasis in a transgenic mouse model of human breast cancer.," *Am J Pathol*, vol. 173, pp. 1551-65, Nov 2008.
- [26] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, J. G. Knittel, L. Yan, C. T. Rueden, *et al.*, "Collagen density promotes mammary tumor initiation and progression.," *BMC Med*, vol. 6, p. 11, 2008.
- [27] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, S. M. Trier, and P. J. Keely, "Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization.," *Biophys J*, vol. 95, pp. 5374-84, Dec 2008.
- [28] C. T. Rueden, M. W. Conklin, P. P. Provenzano, P. J. Keely, and K. W. Eliceiri, "Nonlinear optical microscopy and computational analysis of intrinsic signatures in breast cancer.," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2009, pp. 4077-80, 2009.
- [29] P. P. Provenzano, K. W. Eliceiri, D. R. Inman, and P. J. Keely, "Engineering three-dimensional collagen matrices to provide contact guidance during 3D cell migration.," *Curr Protoc Cell Biol*, vol. Chapter 10, p. Unit 10.17, Jun 2010.
- [30] J. Condeelis and J. E. Segall, "Intravital imaging of cell movement in tumours.," *Nat Rev Cancer*, vol. 3, pp. 921-30, Dec 2003.
- [31] E. Sahai, J. Wyckoff, U. Philipp, J. E. Segall, F. Gertler, and J. Condeelis, "Simultaneous imaging of GFP, CFP and collagen in tumors in vivo using multiphoton microscopy.," *BMC Biotechnol*, vol. 5, p. 14, 2005.
- [32] M. Sidani, J. Wyckoff, C. Xue, J. E. Segall, and J. Condeelis, "Probing the microenvironment of mammary tumors using multiphoton microscopy.," *J Mammary Gland Biol Neoplasia*, vol. 11, pp. 151-63, Apr 2006.

- [33] J. Wyckoff, B. Gligorijevic, D. Entenberg, J. Segall, and J. Condeelis, "High-resolution multiphoton imaging of tumors *in vivo*," *Cold Spring Harb Protoc*, vol. 2011, pp. 1167-84, Oct 2011.
- [34] W. Zheng, D. Li, S. Li, Y. Zeng, Y. Yang, and J. Y. Qu, "Diagnostic value of nonlinear optical signals from collagen matrix in the detection of epithelial precancer.," *Opt Lett*, vol. 36, pp. 3620-2, Sep 2011.
- [35] L. C. U. Junqueira, G. Bignolas, and R. R. Brentani, "Picosirius staining plus polarization microscopy, a specific method for collagen detection in tissue-sections," *Histochemical Journal*, vol. 11, pp. 447-455, 1979.
- [36] D. G. Young, "Improvements to collagen and nuclear staining of the Masson's trichrome," *Journal of Histotechnology*, vol. 24, pp. 271-273, Dec 2001.
- [37] A. S. Craig and D. A. D. Parry, "Correlation between electron-microscope and x-ray data from collagen fibrils," *New Zealand Journal of Science*, vol. 25, pp. 94-95, 1982.
- [38] J. Streicher, W. J. Weninger, and G. B. Müller, "External marker-based automatic congruencing: a new method of 3D reconstruction from serial sections.," *Anat Rec*, vol. 248, pp. 583-602, Aug 1997.
- [39] S. Saatchi, N. Wanchoo, J. Azuma, S. J. Smith, P. S. Tsao, P. G. Yock, et al., "The Use of Immunofluorescent Array Tomography to Study the Three-Dimensional Microstructure of Murine Blood Vessels," *Cellular and Molecular Bioengineering*, vol. 4, pp. 311-323, Jun 2011.
- [40] J. Wu, B. Rajwa, D. L. Filmer, C. M. Hoffmann, B. Yuan, C. Chiang, et al., "Automated quantification and reconstruction of collagen matrix from 3D confocal datasets.," *J Microsc*, vol. 210, pp. 158-65, May 2003.
- [41] K. Wolf, S. Alexander, V. Schacht, L. M. Coussens, U. H. von Andrian, J. van Rheenen, et al., "Collagen-based cell migration models *in vitro* and *in vivo*," *Seminars in Cell & Developmental Biology*, vol. 20, pp. 931-941, Oct 2009.
- [42] D. A. Redford-Badwal, M. Stover, P. J. Campagnola, A. Millard, and D. W. Rowe, "Visualization of osteogenesis imperfecta collagen fibers using second harmonic generation," *Journal of Bone and Mineral Research*, vol. 16, pp. S442-S442, Sep 2001.
- [43] V. Ajeti, O. Nadiarnykh, S. M. Ponik, P. J. Keely, K. W. Eliceiri, and P. J. Campagnola, "Structural changes in mixed Col I/Col V collagen gels probed by SHG microscopy: implications for probing stromal alterations in human breast cancer.," *Biomed Opt Express*, vol. 2, pp. 2307-16, Aug 2011.
- [44] X. Y. Chen, O. Nadiarynkha, S. Plotnikov, and P. J. Campagnola, "Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure," *Nature Protocols*, vol. 7, pp. 654-669, Apr 2012.
- [45] A. M. Stein, D. A. Vader, L. M. Jawerth, D. A. Weitz, and L. M. Sander, "An algorithm for extracting the network geometry of three-dimensional collagen gels.," *J Microsc*, vol. 232, pp. 463-75, Dec 2008.
- [46] H. Feltovich, L. Reusch, L. Carlson, K. Eliceiri, and T. Hall, "Detection of cervical collagen with quantitative ultrasound," *American Journal of Obstetrics and Gynecology*, vol. 204, pp. S138-S138, Jan 2011.
- [47] S. Kreitz, G. Dohmen, S. Hasken, T. Schmitz-Rode, P. Mela, and S. Jockenhoevel, "Nondestructive Method to Evaluate the Collagen Content of Fibrin-Based Tissue Engineered Structures Via Ultrasound," *Tissue Engineering Part C-Methods*, vol. 17, pp. 1021-1026, Oct 2011.
- [48] C.-T. Chen, M.-F. Lin, W. P. Chan, and W.-F. T. Lai, "MRI and histology of collagen template implantation and regeneration in rabbit knee meniscus," *European Journal of Orthopaedic Surgery and Traumatology*, vol. 21, pp. 231-234, May 2011.
- [49] K. M. Mountain, T. A. Bjarnason, J. F. Dunn, and J. R. Matyas, "The Functional Microstructure of Tendon Collagen Revealed by High-Field MRI," *Magnetic Resonance in Medicine*, vol. 66, pp. 520-527, Aug 2011.

- [50] S. Yunoki, T. Ikoma, A. Monkawa, K. Ohta, M. Kikuchi, and J. Tanaka, "Evaluation of pore architecture in hydroxyapatite/collagen scaffold using micro computed tomography," *Key Engineering Materials*, vol. 309, pp. 1091-1094, 2006.
- [51] H. T. Kokkonen, J. Makela, K. A. M. Kulmala, L. Rieppo, J. S. Jurvelin, V. Tiitu, *et al.*, "Computed tomography detects changes in contrast agent diffusion after collagen cross-linking typical to natural aging of articular cartilage," *Osteoarthritis and Cartilage*, vol. 19, pp. 1190-1198, Oct 2011.
- [52] J. B. Pawley, *Handbook of biological confocal microscopy*, 3rd ed. New York: Springer, 2006.
- [53] C. J. R. Sheppard and T. Wilson, "Image-formation in confocal scanning microscopes," *Optik*, vol. 55, pp. 331-342, 1980.
- [54] W. Denk, J. H. Strickler, and W. W. Webb, "2-photon laser scanning fluorescence microscopy," *Science*, vol. 248, pp. 73-76, Apr 1990.
- [55] P. J. Campagnola, M. D. Wei, A. Lewis, and L. M. Loew, "High-resolution nonlinear optical imaging of live cells by second harmonic generation," *Biophysical Journal*, vol. 77, pp. 3341-3349, Dec 1999.
- [56] P. J. Campagnola, M. D. Wei, and L. M. Loew, "Second harmonic generation imaging of living cells," *Biophysical Journal*, vol. 76, pp. A95-A95, Jan 1999.
- [57] R. M. Williams, W. R. Zipfel, and W. W. Webb, "Interpreting second-harmonic generation images of collagen I fibrils," *Biophys J*, vol. 88, pp. 1377-86, Feb 2005.
- [58] S. V. Plotnikov, A. C. Millard, P. J. Campagnola, and W. A. Mohler, "Characterization of the myosin-based source for second-harmonic generation from muscle sarcomeres.,," *Biophys J*, vol. 90, pp. 693-703, Jan 2006.
- [59] M. Strupler, A. M. Pena, M. Hernest, P. L. Tharaux, J. L. Martin, E. Beaurepaire, *et al.*, "Second harmonic imaging and scoring of collagen in fibrotic tissues.,," *Opt Express*, vol. 15, pp. 4054-65, Apr 2007.
- [60] G. F. a. S. P. a. R. Murison, "Analysis of collagen fibre shape changes in breast cancer," *Physics in Medicine and Biology*, vol. 53, p. 6641, 2008.
- [61] C. Odin, T. Guilbert, A. Alkilani, O. P. Boryskina, V. Fleury, and Y. Le Grand, "Collagen and myosin characterization by orientation field second harmonic microscopy," *Opt Express*, vol. 16, pp. 16151-65, Sep 29 2008.
- [62] P. P. Provenzano, C. T. Rueden, S. M. Trier, L. Yan, S. M. Ponik, D. R. Inman, *et al.*, "Nonlinear optical imaging and spectral-lifetime computational analysis of endogenous and exogenous fluorophores in breast cancer.,," *J Biomed Opt*, vol. 13, p. 031220, 2008 May-Jun 2008.
- [63] P. P. Provenzano, K. W. Eliceiri, and P. J. Keely, "Multiphoton microscopy and fluorescence lifetime imaging microscopy (FLIM) to monitor metastasis and the tumor microenvironment.,," *Clin Exp Metastasis*, vol. 26, pp. 357-70, 2009.
- [64] H. Bao, A. Boussioutas, R. Jeremy, S. Russell, and M. Gu, "Second harmonic generation imaging via nonlinear endomicroscopy," *Opt Express*, vol. 18, pp. 1255-60, Jan 18 2010.
- [65] P. Stoller, B. M. Kim, A. M. Rubenchik, K. M. Reiser, and L. B. Da Silva, "Polarization-dependent optical second-harmonic imaging of a rat-tail tendon," *J Biomed Opt*, vol. 7, pp. 205-14, Apr 2002.
- [66] J.-X. Cheng and X. S. Xie, "Green's function formulation for third-harmonic generation microscopy," *J. Opt. Soc. Am. B*, vol. 19, pp. 1604-1610, 2002.
- [67] L. Moreaux, O. Sandre, S. Charpak, M. Blanchard-Desce, and J. Mertz, "Coherent scattering in multi-harmonic light microscopy.,," *Biophys J*, vol. 80, pp. 1568-74, Mar 2001.
- [68] S. V. Plotnikov, A. C. Millard, P. J. Campagnola, and W. A. Mohler, "Characterization of the myosin-based source for second-harmonic generation from muscle sarcomeres," *Biophysical Journal*, vol. 90, pp. 693-703, Jan 2006.

- [69] A. E. Tuer, M. K. Akens, S. Kruglov, D. Sandkuijl, B. C. Wilson, C. M. Whyne, *et al.*, "Hierarchical Model of Fibrillar Collagen Organization for Interpreting the Second-Order Susceptibility Tensors in Biological Tissue," *Biophysical Journal*, vol. 103, pp. 2093-2105, Nov 2012.
- [70] A. Erikson, J. Ortegren, T. Hompland, C. de Lange Davies, and M. Lindgren, "Quantification of the second-order nonlinear susceptibility of collagen I using a laser scanning microscope," *J Biomed Opt*, vol. 12, pp. 1-10, Jul-Aug 2007.
- [71] S. Preibisch, S. Saalfeld, and P. Tomancak, "Globally optimal stitching of tiled 3D microscopic image acquisitions," *Bioinformatics*, vol. 25, pp. 1463-1465, Jun 2009.
- [72] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, *et al.*, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes.,," *Genome Biol*, vol. 7, p. R100, 2006.
- [73] M. R. Lamprecht, D. M. Sabatini, and A. E. Carpenter, "CellProfiler: free, versatile software for automated biological image analysis.,," *Biotechniques*, vol. 42, pp. 71-5, Jan 2007.
- [74] P. L. Chandran and V. H. Barocas, "Affine versus non-affine fibril kinematics in collagen networks: theoretical studies of network behavior.,," *J Biomech Eng*, vol. 128, pp. 259-70, Apr 2006.
- [75] S. Chaudhuri, H. Nguyen, R. M. Rangayyan, S. Walsh, and C. B. Frank, "A Fourier domain directional filtering method for analysis of collagen alignment in ligaments.,," *IEEE Trans Biomed Eng*, vol. 34, pp. 509-18, Jul 1987.
- [76] C. Bayan, J. M. Levitt, E. Miller, D. Kaplan, and I. Georgakoudi, "Fully automated, quantitative, noninvasive assessment of collagen fiber content and organization in thick collagen gels," *Journal of Applied Physics*, vol. 105, pp. 102042-11, 2009.
- [77] M. J. Pablo, "Fourier analysis and automated measurement of cell and fiber angular orientation distributions," *International Journal of Solids and Structures*, vol. 43, pp. 6413-6423, 2006.
- [78] J. Wu, B. Rajwa, D. L. Filmer, C. M. Hoffmann, B. Yuan, C. S. Chiang, *et al.*, "Analysis of orientations of collagen fibers by novel fiber-tracking software.,," *Microsc Microanal*, vol. 9, pp. 574-80, Dec 2003.
- [79] Y. Xia and K. Elder, "Quantification of the graphical details of collagen fibrils in transmission electron micrographs.,," *J Microsc*, vol. 204, pp. 3-16, Oct 2001.
- [80] J. L. Starck, F. Murtagh, E. J. Candès, and D. L. Donoho, "Gray and color image contrast enhancement by the curvelet transform.,," *IEEE Trans Image Process*, vol. 12, pp. 706-17, 2003.
- [81] C. Pehlke, J. S. Bredfeldt, J. Doot, K. E. Sung, P. Provenzano, K. Riching, *et al.*, "Quantification of Collagen Architecture using the Curvelet Transform," *Integrative Biology*, In Review, January 2014.
- [82] E. Matos, "Treatment of HER2 positive breast cancer patients," *Zdravniški Vestnik-Slovenian Medical Journal*, vol. 81, pp. 236-245, Mar 2012.
- [83] F. Andre, L. M. McShane, S. Michiels, D. F. Ransohoff, D. G. Altman, J. S. Reis-Filho, *et al.*, "Biomarker studies: a call for a comprehensive biomarker study registry," *Nature Reviews Clinical Oncology*, vol. 8, pp. 171-176, Mar 2011.
- [84] A. K. Paul and R. B. Schwab, "Efficacy and pharmacogenomic biomarkers in breast cancer," *Biomarkers in Medicine*, vol. 6, pp. 211-221, Apr 2012.
- [85] N. Saijo, "Critical comments for roles of biomarkers in the diagnosis and treatment of cancer," *Cancer Treatment Reviews*, vol. 38, pp. 63-67, Feb 2012.
- [86] S. H. Woolf, "The meaning of translational research and why it matters.,," *JAMA*, vol. 299, pp. 211-3, Jan 2008.
- [87] I. Eke and N. Cordes, "Radiobiology goes 3D: how ECM and cell morphology impact on cell survival after irradiation.,," *Radiother Oncol*, vol. 99, pp. 271-8, Jun 2011.

- [88] C. M. Ghajar and M. J. Bissell, "Extracellular matrix control of mammary gland morphogenesis and tumorigenesis: insights from imaging," *Histochemistry and Cell Biology*, vol. 130, pp. 1105-1118, 2008.
- [89] C. D. Roskelley, A. Srebrow, and M. J. Bissell, "A hierarchy of ECM-mediated signalling regulates tissue-specific gene expression," *Current Opinion in Cell Biology*, vol. 7, pp. 736-747, 1995.
- [90] P. J. Keely, J. E. Wu, and S. A. Santoro, "The spatial and temporal expression of the [alpha]2[beta]1 integrin and its ligands, collagen I, collagen IV, and laminin, suggest important roles in mouse mammary morphogenesis," *Differentiation*, vol. 59, pp. 1-13, 1995.
- [91] P. P. Provenzano, K. W. Eliceiri, J. M. Campbell, D. R. Inman, J. G. White, and P. J. Keely, "Collagen reorganization at the tumor-stromal interface facilitates local invasion," *BMC Med*, vol. 4, p. 38, 2006.
- [92] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, J. G. Knittel, L. Yan, C. T. Rueden, et al., "Collagen density promotes mammary tumor initiation and progression," *BMC Med*, vol. 6, p. 11, 2008.
- [93] M. Simian, Y. Hirai, M. Navre, Z. Werb, A. Lochter, and M. J. Bissell, "The interplay of matrix metalloproteinases, morphogens and growth factors is necessary for branching of mammary epithelial cells" *Development*, vol. 128, pp. 3117-31, 2001.
- [94] P. Keely, A. Fong, M. Zutter, and S. Santoro, "Alteration of collagen-dependent adhesion, motility, and morphogenesis by the expression of antisense alpha 2 integrin mRNA in mammary cells," *J Cell Sci*, vol. 108, pp. 595-607, February 1, 1995.
- [95] H. Burstein, K. Polyak, J. Wong, S. Lester, and C. Kaelin, "Ductal carcinoma in situ of the breast.," *New England Journal of Medicine*, vol. 350, pp. 1430-41, April 1, 2004.
- [96] D. T. Butcher, T. Alliston, and V. M. Weaver, "A tense situation: forcing tumour progression," *Nat Rev Cancer*, vol. 9, pp. 108-122, 2009.
- [97] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, S. M. Trier, and P. J. Keely, "Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization," *Biophys J*, vol. 95, pp. 5374-84, Dec 2008.
- [98] W. Wang, J. B. Wyckoff, V. C. Frohlich, Y. Oleynikov, S. Huttelmaier, J. Zavadil, et al., "Single cell behavior in metastatic primary mammary tumors correlated with gene expression patterns revealed by molecular profiling," *Cancer Res*, vol. 62, pp. 6278-88, Nov 1 2002.
- [99] W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning fluorescence microscopy," *Science*, vol. 248, pp. 73-6, Apr 6 1990.
- [100] P. P. Provenzano, K. W. Eliceiri, and P. J. Keely, "Multiphoton microscopy and fluorescence lifetime imaging microscopy (FLIM) to monitor metastasis and the tumor microenvironment," *Clin Exp Metastasis*, Sep 3 2008.
- [101] A. Diaspro and C. J. R. Sheppard, "Two-Photon Excitation Fluorescence Microscopy," in *Confocal and Two-Photon Microscopy: Foundations, Applications, and Advances*, A. Diaspro, Ed., ed New York: Wiley-Liss, Inc., 2002, pp. 39-73.
- [102] J. M. Squirrell, D. L. Wokosin, J. G. White, and B. D. Bavister, "Long-term two-photon fluorescence imaging of mammalian embryos without compromising viability," *Nat Biotechnol*, vol. 17, pp. 763-7, Aug 1999.
- [103] V. E. Centonze and J. G. White, "Multiphoton excitation provides optical sections from deeper within scattering specimens than confocal imaging," *Biophys J*, vol. 75, pp. 2015-24, Oct 1998.
- [104] W. Mohler, A. C. Millard, and P. J. Campagnola, "Second harmonic generation imaging of endogenous structural proteins," *Methods*, vol. 29, pp. 97-109, Jan 2003.
- [105] P. Campagnola, "Second Harmonic Generation Imaging Microscopy: Applications to Diseases Diagnostics," *Analytical Chemistry*, vol. 83, pp. 3224-3231, 2011.

- [106] W. Wang, J. B. Wyckoff, V. C. Frohlich, Y. Oleynikov, S. Hottelmaier, J. Zavadil, *et al.*, "Single Cell Behavior in Metastatic Primary Mammary Tumors Correlated with Gene Expression Patterns Revealed by Molecular Profiling," *Cancer Research*, vol. 62, pp. 6278-6288, November 1, 2002.
- [107] D. K. Bird, L. Yan, K. M. Vrotsos, K. W. Eliceiri, E. M. Vaughan, P. J. Keely, *et al.*, "Metabolic Mapping of MCF10A Human Breast Cells via Multiphoton Fluorescence Lifetime Imaging of the Coenzyme NADH," *Cancer Research*, vol. 65, pp. 8766-8773, October 1, 2005.
- [108] P. P. Provenzano, K. W. Eliceiri, and P. J. Keely, "Shining new light on 3D cell motility and the metastatic process," *Trends in Cell Biology*, vol. 19, pp. 638-648, 2009.
- [109] B. A. Torkian, A. T. Yeh, R. Engel, C.-H. Sun, B. J. Tromberg, and B. J. F. Wong, "Modeling Aberrant Wound Healing Using Tissue-Engineered Skin Constructs and Multiphoton Microscopy," *Archives of Facial Plastic Surgery*, vol. 6, pp. 180-187, May/Jun 2004 2004.
- [110] A. T. Yeh, B. Kao, W. G. Jung, Z. Chen, J. S. Nelson, and B. J. Tromberg, "Imaging wound healing using optical coherence tomography and multiphoton microscopy in an in vitro skin-equivalent tissue model," *Journal of Biomedical Optics*, vol. 9, pp. 248-253, 2004.
- [111] S. Zhuo, J. Chen, X. Jiang, X. Cheng, and S. Xie, "Visualizing extracellular matrix and sensing fibroblasts metabolism in human dermis by nonlinear spectral imaging," *Skin Research and Technology*, vol. 13, pp. 406-411, 2007.
- [112] M. W. Conklin, J. C. Eickhoff, K. M. Riching, C. A. Pehlke, K. W. Eliceiri, P. P. Provenzano, *et al.*, "Aligned Collagen Is a Prognostic Signature for Survival in Human Breast Carcinoma," *The American Journal of Pathology*, vol. 178, pp. 1221-1232.
- [113] J. W. Pollard, "Macrophages define the invasive microenvironment in breast cancer," *Journal of Leukocyte Biology*, vol. 84, pp. 623-630, September 1, 2008 2008.
- [114] A. R. Round and et al., "A preliminary study of breast cancer diagnosis using laboratory based small angle x-ray scattering," *Physics in Medicine and Biology*, vol. 50, p. 4159, 2005.
- [115] O. Nadiarnykh, R. B. LaComb, M. A. Brewer, and P. J. Campagnola, "Alterations of the extracellular matrix in ovarian cancer studied by Second Harmonic Generation imaging microscopy," *Bmc Cancer*, vol. 10, Mar 2010.
- [116] J. A. Tuxhorn, G. E. Ayala, M. J. Smith, V. C. Smith, T. D. Dang, and D. R. Rowley, "Reactive Stroma in Human Prostate Cancer," *Clinical Cancer Research*, vol. 8, pp. 2912-2923, September 1, 2002.
- [117] B. B. Araujo, M. Dolhnikoff, L. F. F. Silva, J. Elliot, J. H. N. Lindeman, D. S. Ferreira, *et al.*, "Extracellular matrix components and regulators in the airway smooth muscle in asthma," *European Respiratory Journal*, vol. 32, pp. 61-69, July 1, 2008 2008.
- [118] C. Bergeron, W. Al-Ramli, and Q. Hamid, "Remodeling in Asthma," *Proc Am Thorac Soc*, vol. 6, pp. 301-305, May 1, 2009 2009.
- [119] S. Chaudhuri, H. Nguyen, R. M. Rangayyan, S. Walsh, and C. B. Frank, "A Fourier domain directional filtering method for analysis of collagen alignment in ligaments," *IEEE Trans Biomed Eng*, vol. 34, pp. 509-18, Jul 1987.
- [120] J. P. Marquez, "Fourier analysis and automated measurement of cell and fiber angular orientation distributions," *International Journal of Solids and Structures*, vol. 43, pp. 6413-6423, 2006.
- [121] C. P. Ng, B. Hinz, and M. A. Swartz, "Interstitial fluid flow induces myofibroblast differentiation and collagen alignment in vitro," *J Cell Sci*, vol. 118, pp. 4731-9, Oct 15 2005.
- [122] C. P. Ng and M. A. Swartz, "Mechanisms of interstitial flow-induced remodeling of fibroblast-collagen cultures," *Ann Biomed Eng*, vol. 34, pp. 446-54, Mar 2006.
- [123] Y. Xia and K. Elder, "Quantification of the graphical details of collagen fibrils in transmission electron micrographs," *J Microsc*, vol. 204, pp. 3-16, Oct 2001.

- [124] S. Plotnikov, V. Juneja, A. B. Isaacson, W. A. Mohler, and P. J. Campagnola, "Optical Clearing for Improved Contrast in Second Harmonic Generation Imaging of Skeletal Muscle," *Biophysical Journal*, vol. 90, pp. 328-339, 2006.
- [125] P. Matteini, F. Ratto, F. Rossi, R. Cicchi, C. Stringari, D. Kapsokalyvas, *et al.*, "Photothermally-induced disordered patterns of corneal collagen revealed by SHG imaging," *Opt. Express*, vol. 17, pp. 4868-4878, 2009.
- [126] Y.-W. Chiu, M. T. Lo, M.-R. Tsai, Y.-C. Chang, R.-B. Hsu, H.-Y. Yu, *et al.*, "Applying Harmonic Optical Microscopy for Spatial Alignment of Atrial Collagen Fibers," *Plos One*, vol. 5, p. e13917.
- [127] C. Bayan, J. M. Levitt, E. Miller, D. Kaplan, and I. Georgakoudi, "Fully automated, quantitative, noninvasive assessment of collagen fiber content and organization in thick collagen gels," *Journal of Applied Physics*, vol. 105, May 2009.
- [128] G. Falzon and et al., "Analysis of collagen fibre shape changes in breast cancer," *Physics in Medicine and Biology*, vol. 53, p. 6641, 2008.
- [129] A. Zemel, F. Rehfeldt, A. E. X. Brown, D. E. Discher, and S. A. Safran, "Optimal matrix rigidity for stress-fibre polarization in stem cells," *Nature Physics*, vol. 6, pp. 468-473, Jun 2010.
- [130] E. Candès and D. Donoho, "Curvelets - a surprisingly effective nonadaptive representation for objects with edges," presented at the Curves and Surface Fitting: Saint-Malo 1999, Saint-Malo, France, 1999.
- [131] J. Ma and G. Plonka, "The Curvelet Transform," *IEEE Signal Processing Magazine*, vol. 27, pp. 118-133, Mar 2010.
- [132] R. Mehrotra, K. R. Namuduri, and N. Ranganathan, "Gabor filter-based edge detection," *Pattern Recognition*, vol. 25, pp. 1479-1494, 1992.
- [133] J. Canny, "A Computational Approach to Edge Detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, pp. 679-698, 1986.
- [134] T. Geback and P. Koumoutsakos, "Edge detection in microscopy images using curvelets," *BMC Bioinformatics*, vol. 10, p. 75, 2009.
- [135] F. J. Herrmann, "Curvelet imaging and processing: an overview," in *CSEG National Convention*, 2004.
- [136] H. Chauris and T. Nguyen, "Seismic demigration/migration in the curvelet domain," *Geophysics*, vol. 73, pp. S35-46, 2008.
- [137] G. Hennenfent, L. Fenelon, and F. J. Herrmann, "Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach," *Geophysics*, vol. 75, pp. WB203-210, November 1, 2010.
- [138] F. J. Herrmann, P. Moghaddam, and C. C. Stolk, "Sparsity- and continuity-promoting seismic image recovery with curvelet frames," *Applied and Computational Harmonic Analysis*, vol. 24, pp. 150-173, 2008.
- [139] H. R. R. Arabnia, Q.-N. Tran, S. A. Hyder, and R. Sukanesh, "An Efficient Algorithm for Denoising MR and CT Images Using Digital Curvelet Transform," in *Software Tools and Algorithms for Biological Systems*. vol. 696, ed: Springer New York, pp. 471-480.
- [140] A. Wang, Y. Zhang, S. Meng, and M. Yang, "Image denoising method based on curvelet transform," in *Industrial Electronics and Applications, 2008. ICIEA. 3rd IEEE Conference on*, 2008, pp. 571-574.
- [141] M. Kamel, A. Campilho, T. Mandal, A. Majumdar, and Q. Wu, "Face Recognition by Curvelet Based Feature Extraction," in *Image Analysis and Recognition*. vol. 4633, ed: Springer Berlin / Heidelberg, 2007, pp. 806-817.
- [142] D. Cherifi, A. Beghdadi, and A. H. Belbachir, "A multiscale and multi-orientation contrast enhancement method for digital mammography," in *Communications, Control and Signal Processing, ISCCSP. 3rd International Symposium on*, 2008, pp. 312-315.

- [143] M. M. Eltoukhy, I. Faye, and B. B. Samir, "Breast cancer diagnosis in digital mammogram using multiscale curvelet transform," *Computerized Medical Imaging and Graphics*, vol. 34, pp. 269-276, 2010.
- [144] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast Discrete Curvelet Transforms," *Multiscale Modeling & Simulation*, vol. 5, p. 861, 2007.
- [145] J. H. Friedman, J. L. Bentley, and R. Ari Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Transactions on Mathematical Software*, vol. 3, pp. 209-26, Sept. 1977.
- [146] P. Berens, "CircStat: A MATLAB Toolbox for Circular Statistics," *Journal of Statistical Software*, vol. 31, pp. 1-21, Sep 2009.
- [147] P. P. Provenzano, K. W. Eliceiri, and P. J. Keely, "Multiphoton microscopy and fluorescence lifetime imaging microscopy (FLIM) to monitor metastasis and the tumor microenvironment," *Clinical & Experimental Metastasis*, vol. 26, pp. 357-370, Apr 2009.
- [148] D. L. Wokosin, J. M. Squirrell, K. W. Eliceiri, and J. G. White, "Optical workstation with concurrent, independent multiphoton imaging and experimental laser microbeam capabilities," *Review of Scientific Instruments*, vol. 74, pp. 193-201, 2003.
- [149] J. G. Goetz, S. Minguet, I. Navarro-LÈrida, J. J. Lazcano, R. Samaniego, E. Calvo, et al., "Biomechanical Remodeling of the Microenvironment by Stromal Caveolin-1 Favors Tumor Invasion and Metastasis," *Cell*, vol. 146, pp. 148-163, 2011.
- [150] K. Zhang, C. A. Corsa, S. M. Ponik, J. L. Prior, D. Piwnica-Worms, K. W. Eliceiri, et al., "The collagen receptor discoidin domain receptor 2 stabilizes SNAIL1 to facilitate breast cancer metastasis," *Nature Cell Biology*, vol. 15, p. 677, Jun 2013.
- [151] L. M. Reusch, J. Anderson, L. C. Carlson, H. A. Simmons, C. Retake, K. Eliceiri, et al., "Detecting cervical microstructure via ultrasound and optical microscopy," in *Ultrasonics Symposium (IUS), 2010 IEEE*, pp. 724-727.
- [152] N. A. Kouris, J. M. Squirrell, J. P. Jung, C. A. Pehlke, T. Hacker, K. W. Eliceiri, et al., "A nondenatured, noncrosslinked collagen matrix to deliver stem cells to the heart," *Regenerative Medicine*, vol. 6, pp. 569-582, 2011.
- [153] J. N. Heck, M. G. Garcia-Mendoza, S. M. Ponik, C. A. Pehlke, D. R. Inman, K. W. Eliceiri, et al., "Microtubules Regulate Gef-H1 in Response to Extracellular Matrix Stiffness," *Molecular Biology of the Cell*, vol. 23, pp. 2583-2592, 2012.
- [154] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, and P. J. Keely, "Matrix density-induced mechanoregulation of breast cell phenotype, signaling and gene expression through a FAK-ERK linkage.," *Oncogene*, vol. 28, pp. 4326-43, Dec 2009.
- [155] M. Linkert, C. T. Rueden, C. Allan, J.-M. Burel, W. Moore, A. Patterson, et al., "Metadata matters: access to image data in the real world," *The Journal of Cell Biology*, vol. 189, pp. 777-782, May 31, 2010.
- [156] V. Ajeti, O. Nadiarnykh, S. M. Ponik, P. J. Keely, K. W. Eliceiri, and P. J. Campagnola, "Structural changes in mixed Col I/Col V collagen gels probed by SHG microscopy: implications for probing stromal alterations in human breast cancer," *Biomed. Opt. Express*, vol. 2, pp. 2307-2316, 2011.
- [157] L. Ying, L. Demanet, and E. Candes, "3D discrete curvelet transform," in *Wavelets XI*, San Diego, CA, USA, 2005, pp. 591413-11.
- [158] G. Cox, E. Kable, A. Jones, I. K. Fraser, F. Manconi, and M. D. Gorrell, "3-dimensional imaging of collagen using second harmonic generation," *Journal of Structural Biology*, vol. 141, pp. 53-62, Jan 2003.
- [159] C. B. Raub, V. Suresh, T. Krasieva, J. Lyubovitsky, J. D. Mih, A. J. Putnam, et al., "Noninvasive assessment of collagen gel microstructure and mechanics using multiphoton microscopy.," *Biophys J*, vol. 92, pp. 2212-22, Mar 2007.

- [160] O. Nadiarnykh, R. B. LaComb, M. A. Brewer, and P. J. Campagnola, "Alterations of the extracellular matrix in ovarian cancer studied by Second Harmonic Generation imaging microscopy," *BMC Cancer*, vol. 10, pp. 1-14, Mar 11 2010.
- [161] J. M. Watson, P. F. Rice, S. L. Marion, M. A. Brewer, J. R. Davis, J. J. Rodriguez, *et al.*, "Analysis of second-harmonic-generation microscopy in a mouse model of ovarian carcinoma," *J. Biomed. Opt.*, vol. 17, pp. 1-9, 2012.
- [162] R. Rezakhaniha, A. Agianniotis, J. T. C. Schrauwen, A. Griffa, D. Sage, C. V. C. Bouten, *et al.*, "Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy," *Biomechanics and Modeling in Mechanobiology*, vol. 11, pp. 461-473, Mar 2012.
- [163] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, J. G. Knittel, L. Yan, C. T. Rueden, *et al.*, "Collagen density promotes mammary tumor initiation and progression," *BMC Med*, vol. 6, pp. 1-15, Apr 28 2008.
- [164] G. Falzon, S. Pearson, and R. Murison, "Analysis of collagen fibre shape changes in breast cancer," *Phys Med Biol*, vol. 53, pp. 6641-52, Dec 7 2008.
- [165] K. E. Frisch, S. E. Duenwald-Kuehl, H. Kobayashi, C. S. Chamberlain, R. S. Lakes, and R. Vanderby, Jr., "Quantification of collagen organization using fractal dimensions and Fourier transforms," *Acta Histochemica*, vol. 114, pp. 140-144, 2012.
- [166] H. Altendorf, E. Decenciere, D. Jeulin, P. D. S. Peixoto, A. Deniset-Besseau, E. Angelini, *et al.*, "Imaging and 3D morphological analysis of collagen fibrils," *Journal of Microscopy*, vol. 247, pp. 161-175, Aug 2012.
- [167] W. Y. Hu, H. Li, C. Y. Wang, S. M. Gou, and L. Fu, "Characterization of collagen fibers by means of texture analysis of second harmonic generation images using orientation-dependent gray level co-occurrence matrix method," *J. Biomed. Opt.*, vol. 17, pp. 1-9, Feb 2012.
- [168] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms--theory and practice," *IEEE Trans Image Process*, vol. 21, pp. 1084-96, Mar 2012.
- [169] Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, *et al.*, "Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," *Medical image analysis*, vol. 2, pp. 143-68, 1998.
- [170] J. L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans Image Process*, vol. 11, pp. 670-84, 2002.
- [171] E. Candes, L. Demanet, D. Donoho, and L. X. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling & Simulation*, vol. 5, pp. 861-899, 2006.
- [172] A. Stein. (February 15, 2012). *FIRE FlbeR Extraction*. Available: <http://www.ima.umn.edu/~astein/Andrew%20Stein/Software.html>
- [173] E. Y. S. Yew and C. J. R. Sheppard, "Effects of axial field components on second harmonic generation microscopy," *Optics Express*, vol. 14, pp. 1167-1174, Feb 6 2006.
- [174] E. Y. S. Yew and C. J. R. Sheppard, "Second harmonic generation polarization microscopy with tightly focused linearly and radially polarized beams," *Optics Communications*, vol. 275, pp. 453-457, Jul 15 2007.
- [175] D. Kedrin, B. Gligorijevic, J. Wyckoff, V. V. Verkhusha, J. Condeelis, J. E. Segall, *et al.*, "Intravital imaging of metastatic behavior through a mammary imaging window," *Nature Methods*, vol. 5, pp. 1019-1021, Dec 2008.
- [176] U. Bal, "Dual tree complex wavelet transform based denoising of optical microscopy images," *Biomedical optics express*, vol. 3, 2012.
- [177] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans Image Process*, vol. 18, pp. 2419-34, Nov 2009.

- [178] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, *et al.*, "Fiji: an open-source platform for biological-image analysis," *Nature Methods*, vol. 9, pp. 676-682, Jul 2012.
- [179] M. H. Longair, D. A. Baker, and J. D. Armstrong, "Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes," *Bioinformatics*, vol. 27, pp. 2453-2454, Sep 2011.
- [180] E. J. Candes and F. Guo, "New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction," *Signal Processing*, vol. 82, pp. 1519-1543, Nov 2002.
- [181] D. Sage, D. Prodanov, C. Ortiz, J.-Y. Tivenez, J. Pecreaux, and J. Schindelin. (Accessed: Apr 24, 2012). *A Java package for bi-directional communication and data exchange from Matlab to ImageJ/Fiji*. Available: <http://bigwww.epfl.ch/sage/soft/mij/>
- [182] X. Wang, T. Heimann, P. Lo, M. Sumkauskaitė, M. Puderbach, M. de Bruijne, *et al.*, "Statistical tracking of tree-like tubular structures with efficient branching detection in 3D medical image data," *Phys Med Biol*, vol. 57, pp. 5325-42, Aug 21 2012.
- [183] V. Mohan, G. Sundaramoorthi, and A. Tannenbaum, "Tubular Surface Segmentation for Extracting Anatomical Structures From Medical Imagery," *Ieee Transactions on Medical Imaging*, vol. 29, pp. 1945-1958, Dec 2010.
- [184] O. Friman, M. Hindennach, C. Kuhnel, and H. O. Peitgen, "Multiple hypothesis template tracking of small 3D vessel structures," *Med Image Anal*, vol. 14, pp. 160-71, Apr 2010.
- [185] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, pp. 1277-1294, Sep 1993.
- [186] I. S. Molchanov and P. Teran, "Distance transforms for real-valued functions," *Journal of Mathematical Analysis and Applications*, vol. 278, pp. 472-484, Feb 15 2003.
- [187] C. Pehlke, J. Doot, K. E. Sung, P. Provenzano, K. Riching, D. Inman, *et al.*, "Quantification of Collagen Architecture using the Curvelet Transform," ed, In Press.
- [188] P. Stoller, K. M. Reiser, P. M. Celliers, and A. M. Rubenchik, "Polarization-modulated second harmonic generation in collagen," *Biophysical Journal*, vol. 82, pp. 3330-3342, Jun 2002.
- [189] J. Duboisset, D. Ait-Belkacem, M. Roche, H. Rigneault, and S. Brasselet, "Generic model of the molecular orientational distribution probed by polarization-resolved second-harmonic generation," *Physical Review A*, vol. 85, pp. 1-9, Apr 19 2012.
- [190] W. R. Zipfel, R. M. Williams, and W. W. Webb, "Nonlinear magic: Multiphoton microscopy in the biosciences," *Nature Biotechnology*, vol. 21, pp. 1369-1377, November 2003.
- [191] M. W. Conklin, P. P. Provenzano, K. W. Eliceiri, R. Sullivan, and P. J. Keely, "Fluorescence lifetime imaging of endogenous fluorophores in histopathology sections reveals differences between normal and tumor epithelium in carcinoma in situ of the breast.," *Cell Biochem Biophys*, vol. 53, pp. 145-57, 2009.
- [192] D. C. Allred, J. M. Harvey, M. Berardo, and G. M. Clark, "Prognostic and predictive factors in breast cancer by immunohistochemical analysis," *Modern Pathology*, vol. 11, pp. 155-168, Feb 1998.
- [193] M. Press, D. Slamon, M. Cobleigh, C. Vogel, J. Y. Zhou, S. Anderson, *et al.*, "Improved clinical outcomes for Herceptin (R)-treated patients selected by fluorescence in situ hybridization (FISH)," *Modern Pathology*, vol. 15, pp. 47A-47A, Jan 2002.
- [194] L. A. Habel, S. Shak, M. K. Jacobs, A. Capra, C. Alexander, M. Pho, *et al.*, "A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients," *Breast Cancer Research*, vol. 8, 2006.
- [195] R. Nahta and F. J. Esteva, "HER2 therapy - Molecular mechanisms of trastuzumab resistance," *Breast Cancer Research*, vol. 8, 2006.

- [196] L. Shamir, J. D. Delaney, N. Orlov, D. M. Eckley, and I. G. Goldberg, "Pattern Recognition Software and Techniques for Biological Image Analysis," *Plos Computational Biology*, vol. 6, Nov 2010.
- [197] A. Madabhushi, S. Agner, A. Basavanhally, S. Doyle, and G. Lee, "Computer-aided prognosis: Predicting patient and disease outcome via quantitative fusion of multi-scale, multi-modal data," *Computerized Medical Imaging and Graphics*, vol. 35, Oct-Dec 2011.
- [198] G. Myers, "Why bioimage informatics matters," *Nature Methods*, vol. 9, pp. 659-660, Jul 2012.
- [199] L. Ronnovjessen, O. W. Petersen, V. E. Koteliansky, and M. J. Bissell, "The Origin of the Myofibroblasts in Breast-Cancer - Recapitulation of Tumor Environment in Culture Unravels Diversity and Implicates Converted Fibroblasts and Recruited Smooth Muscle Cells," *Journal of Clinical Investigation*, vol. 95, Feb 1995.
- [200] B. Elenbaas, L. Spirio, F. Koerner, M. D. Fleming, D. B. Zimonjic, J. L. Donaher, *et al.*, "Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells," *Genes & Development*, vol. 15, Jan 1 2001.
- [201] T. D. Tlsty and P. W. Hein, "Know thy neighbor: stromal cells can contribute oncogenic signals," *Current Opinion in Genetics & Development*, vol. 11, Feb 2001.
- [202] N. F. Boyd, L. J. Martin, L. Sun, H. Guo, A. Chiarelli, G. Hislop, *et al.*, "Body size, mammographic density, and breast cancer risk," *Cancer Epidemiology Biomarkers & Prevention*, vol. 15, pp. 2086-2092, Nov 2006.
- [203] N. F. Boyd, L. J. Martin, M. Bronskill, M. J. Yaffe, N. Duric, and S. Minkin, "Breast Tissue Composition and Susceptibility to Breast Cancer," *Journal of the National Cancer Institute*, vol. 102, pp. 1224-1237, Aug 18 2010.
- [204] E. Thurfjell, "Perspective - Breast density and the risk of breast cancer," *New England Journal of Medicine*, vol. 347, pp. 866-866, Sep 19 2002.
- [205] L. A. Habel, J. J. Dignam, S. R. Land, M. Salane, A. M. Capra, and T. B. Julian, "Mammographic density and breast cancer after ductal carcinoma in situ," *Journal of the National Cancer Institute*, vol. 96, pp. 1467-1472, Oct 6 2004.
- [206] N. F. Boyd, J. M. Rommens, K. Vogt, V. Lee, J. L. Hopper, M. J. Yaffe, *et al.*, "Mammographic breast density as an intermediate phenotype for breast cancer," *Lancet Oncology*, vol. 6, pp. 798-808, Oct 2005.
- [207] V. A. McCormack and I. D. S. Silva, "Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis," *Cancer Epidemiology Biomarkers & Prevention*, vol. 15, pp. 1159-1169, Jun 2006.
- [208] T. Cil, E. Fishell, W. Hanna, P. Sun, E. Rawlinson, S. A. Narod, *et al.*, "Mammographic Density and the Risk of Breast Cancer Recurrence After Breast-Conserving Surgery," *Cancer*, vol. 115, pp. 5780-5787, Dec 15 2009.
- [209] K. Burke, P. Tang, and E. Brown, "Second harmonic generation reveals matrix alterations during breast tumor progression," *Journal of Biomedical Optics*, vol. 18, Mar 2013.
- [210] S. M. Kakkad, M. Solaiyappan, P. Argani, S. Sukumar, L. K. Jacobs, D. Leibfritz, *et al.*, "Collagen I fiber density increases in lymph node positive breast cancers: pilot study," *Journal of Biomedical Optics*, vol. 17, Nov 2012.
- [211] R. Ambekar, T.-Y. Lau, M. Walsh, R. Bhargava, and K. C. Toussaint, Jr., "Quantifying collagen structure in breast biopsies using second-harmonic generation imaging," *Biomedical Optics Express*, vol. 3, pp. 2021-2035, Sep 1 2012.
- [212] J. S. Bredfeldt, Y. Liu, M. W. Conklin, R. D. Nowak, P. J. Keely, K. W. Eliceiri, *et al.*, "Computational Segmentation of Collagen Fibers from Second Harmonic Generation Images of Breast Cancer," *J Biomed Opt*, vol. 19, p. 016007, 2014.

- [213] M. Rubbens, A. Driessen-Mol, R. Boerboom, M. Koppert, H. van Assen, B. TerHaar Romeny, *et al.*, "Quantification of the Temporal Evolution of Collagen Orientation in Mechanically Conditioned Engineered Cardiovascular Tissues," *Annals of Biomedical Engineering*, vol. 37, pp. 1263-1272, 2009.
- [214] F. Baba, K. Swartz, R. van Buren, J. Eickhoff, Y. Zhang, W. Wolberg, *et al.*, "Syndecan-1 and syndecan-4 are overexpressed in an estrogen receptor-negative, highly proliferative breast carcinoma subtype," *Breast Cancer Research and Treatment*, vol. 98, pp. 91-98, Jul 2006.
- [215] R. LaComb, O. Nadiarnykh, S. S. Townsend, and P. J. Campagnola, "Phase matching considerations in second harmonic generation from tissues: Effects on emission directionality, conversion efficiency and observed morphology," *Optics Communications*, vol. 281, pp. 1823-1832, Apr 1 2008.
- [216] K. Eliceiri and M. Nazir. (2012, 04/19/2012). *Wiscscan*. Available: <http://loci.wisc.edu/software/wiscscan>
- [217] M. Linkert, C. T. Rueden, C. Allan, J.-M. Burel, W. Moore, A. Patterson, *et al.*, "Metadata matters: access to image data in the real world," *Journal of Cell Biology*, vol. 189, pp. 777-782, May 31 2010.
- [218] A. C. Ignacio Arganda-Carreras, Verena Kaynig, Johannes Schindelin. (October 25, 2013). *Trainable Weka Segmentation*. Available: http://fiji.sc/Trainable_Weka_Segmentation
- [219] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5-32, Oct 2001.
- [220] A. Criminisi, J. Shotton, and E. Konukoglu, "Decision forests: a unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning," *Foundations and Trends in Computer Graphics and Vision*, vol. 7, pp. 81-227, 2011.
- [221] J. E. Bresenham, "Algorithm for Computer Control of a Digital Plotter," *Ibm Systems Journal*, vol. 4, pp. 25-30, 1965.
- [222] K. Francis and B. O. Palsson, "Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, pp. 12258-12262, Nov 11 1997.
- [223] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society Series B-Statistical Methodology*, vol. 34, p. 187, 1972.
- [224] C. R. Drifka, K. W. Eliceiri, S. M. Weber, and W. J. Kao, "A bioengineered heterotypic stroma-cancer microenvironment model to study pancreatic ductal adenocarcinoma," *Lab on a Chip*, vol. 13, pp. 3965-3975, 2013.
- [225] W. Hu, G. Zhao, C. Wang, J. Zhang, and L. Fu, "Nonlinear Optical Microscopy for Histology of Fresh Normal and Cancerous Pancreatic Tissues," *Plos One*, vol. 7, May 24 2012.
- [226] D. S. Kliger, Lewis, J. W., Randall, C. E., *Polarized Light in Optics and Spectroscopy*. New York: Adademic Press, 1990.
- [227] R. Oldenbourg, *Polarization microscopy with the LC-PolScope*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 2005.
- [228] P. Whittaker, R. A. Kloner, D. R. Boughner, and J. G. Pickering, "QUANTITATIVE ASSESSMENT OF MYOCARDIAL COLLAGEN WITH PICROSIRIUS RED STAINING AND CIRCULARLY-POLARIZED LIGHT," *Basic Research in Cardiology*, vol. 89, pp. 397-410, Sep-Oct 1994.
- [229] L. Sacconi, E. Froner, R. Antolini, M. R. Taghizadeh, A. Choudhury, and F. S. Pavone, "Multiphoton multifocal microscopy exploiting a diffractive optical element," *Optics Letters*, vol. 28, pp. 1918-1920, Oct 15 2003.
- [230] H. Choi, E. Y. S. Yew, B. Hallacoglu, S. Fantini, C. J. R. Sheppard, and P. T. C. So, "Improvement of axial resolution and contrast in temporally focused widefield two-photon microscopy with structured light illumination," *Biomedical Optics Express*, vol. 4, pp. 995-1005, Jul 1 2013.

- [231] M. E. Bravo-Zanoguera, C. A. Laris, L. K. Nguyen, M. Oliva, and J. H. Price, "Dynamic autofocus for continuous-scanning time-delay-and-integration image acquisition in automated microscopy," *Journal of Biomedical Optics*, vol. 12, May-Jun 2007.
- [232] O. Maller, K. C. Hansen, T. R. Lyons, I. Acerbi, V. M. Weaver, R. Prekeris, *et al.*, "Collagen architecture in pregnancy-induced protection from breast cancer," *Journal of Cell Science*, vol. 126, pp. 4108-4120, Sep 15 2013.
- [233] G. Tao, A. K. Levay, J. D. Peacock, D. J. Huk, S. N. Both, N. H. Purcell, *et al.*, "Collagen XIV is important for growth and structural integrity of the myocardium," *Journal of Molecular and Cellular Cardiology*, vol. 53, pp. 626-638, Nov 2012.
- [234] F. Sodersten, K. Hultenby, D. Heinegard, C. Johnston, and S. Ekman, "Immunolocalization of Collagens (I and III) and Cartilage Oligomeric Matrix Protein in the Normal and Injured Equine Superficial Digital Flexor Tendon," *Connective Tissue Research*, vol. 54, pp. 62-69, 2013.
- [235] A. Uchugonova, M. Zhao, M. Weinigel, Y. Zhang, M. Bouvet, R. M. Hoffman, *et al.*, "Multiphoton tomography visualizes collagen fibers in the tumor microenvironment that maintain cancer-cell anchorage and shape," *Journal of Cellular Biochemistry*, vol. 114, pp. 99-102, Jan 2013.
- [236] N. Russek-Blum, H. Nabel-Rosen, and G. Levkowitz, "Two-photon-based photoactivation in live zebrafish embryos," *Journal of visualized experiments : JoVE*, 2010.
- [237] A. C. Zenclussen, D. N. Olivieri, M. L. Dustin, and C. E. Tadokoro, "In vivo Multiphoton Microscopy Technique to Reveal the Physiology of the Mouse Uterus," *American journal of reproductive immunology (New York, N.Y. : 1989)*, vol. 69, pp. 281-9, 2013.
- [238] C. B. Raub, J. Unruh, V. Suresh, T. Krasieva, T. Lindmo, E. Gratton, *et al.*, "Image correlation spectroscopy of multiphoton images correlates with collagen mechanical properties., " *Biophys J*, vol. 94, pp. 2361-73, Mar 2008.
- [239] T. Y. Lau, R. Ambekar, and K. C. Toussaint, Jr., "Quantification of collagen fiber organization using three-dimensional Fourier transform-second-harmonic generation imaging," *Optics Express*, vol. 20, pp. 21821-21832, Sep 10 2012.
- [240] K. Yoshiki, K. Ryosuke, M. Hashimoto, T. Araki, and N. Hashimoto, "Second-harmonic-generation microscope using eight-segment polarization-mode converter to observe three-dimensional molecular orientation," *Optics Letters*, vol. 32, pp. 1680-1682, Jun 2007.
- [241] A. Baici, G. Cohen, K. Fehr, and A. Böni, "A handy assay for collagenase using reconstituted fluorescein-labeled collagen fibrils., " *Anal Biochem*, vol. 108, pp. 230-2, Nov 1980.
- [242] G. Hall, K. W. Eliceiri, and P. J. Campagnola, "Simultaneous determination of the Second Harmonic Generation emission directionality and reduced scattering coefficient from 3D imaging of thick tissues," *Journal of Biomedical Optics*, Submitted 2013.
- [243] A. M. Stein, D. A. Vader, D. A. Weitz, and L. M. Sander, "The micromechanics of three-dimensional collagen-I gels," *Complexity*, vol. 16, pp. 22-28, 2011.
- [244] B. G. Timms, K. L. Howdeshell, L. Barton, S. Bradley, C. A. Richter, and F. S. Vom Saal, "Estrogenic chemicals in plastic and oral contraceptives disrupt development of the fetal mouse prostate and urethra," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 7014-7019, May 10 2005.
- [245] E. Gibson, M. Gaed, J. A. Gomez, M. Moussa, C. Romagnoli, S. Pautler, *et al.*, "3D prostate histology reconstruction: An evaluation of image-based and fiducial-based algorithms," *Medical Physics*, vol. 40, Sep 2013.
- [246] W. J. Weninger, J. Streicher, and G. B. Müller, "[3-dimensional reconstruction of histological serial sections using a computer]." *Wien Klin Wochenschr*, vol. 108, pp. 515-20, 1996.
- [247] K. T. Mai, H. M. Yazdi, B. F. Burns, and D. G. Perkins, "Pattern of distribution of intraductal and infiltrating ductal carcinoma: a three-dimensional study using serial coronal giant sections of the breast., " *Hum Pathol*, vol. 31, pp. 464-74, Apr 2000.

- [248] J. G. White, W. B. Amos, and M. Fordham, "An evaluation of confocal versus conventional imaging of biological structures by fluorescence light-microscopy," *Journal of Cell Biology*, vol. 105, pp. 41-48, Jul 1987.
- [249] V. E. Centonze and J. G. White, "Multiphoton excitation provides optical sections from deeper within scattering specimens than confocal imaging," *Biophysical Journal*, vol. 75, pp. 2015-2024, Oct 1998.
- [250] H. Hama, H. Kurokawa, H. Kawano, R. Ando, T. Shimogori, H. Noda, *et al.*, "Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain," *Nature Neuroscience*, vol. 14, pp. 1481-U166, Nov 2011.
- [251] B. D. Metscher, "MicroCT for developmental biology: a versatile tool for high-contrast 3D imaging at histological resolutions.,," *Dev Dyn*, vol. 238, pp. 632-40, Mar 2009.
- [252] G. Pieles, S. H. Geyer, D. Szumska, J. Schneider, S. Neubauer, K. Clarke, *et al.*, "microMRI-HREM pipeline for high-throughput, high-resolution phenotyping of murine embryos.,," *J Anat*, vol. 211, pp. 132-7, Jul 2007.
- [253] C. K. Phoon, "Imaging tools for the developmental biologist: ultrasound biomicroscopy of mouse embryonic development.,," *Pediatr Res*, vol. 60, pp. 14-21, Jul 2006.
- [254] J. Sharpe, U. Ahlgren, P. Perry, B. Hill, A. Ross, J. Hecksher-Sorensen, *et al.*, "Optical projection tomography as a tool for 3D microscopy and gene expression studies," *Science*, vol. 296, pp. 541-545, Apr 19 2002.
- [255] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, *et al.*, "Optical Coherence Tomography," *Science*, vol. 254, pp. 1178-1181, Nov 22 1991.
- [256] A. J. Ewald, H. McBride, M. Reddington, S. E. Fraser, and R. Kerschmann, "Surface imaging microscopy, an automated method for visualizing whole embryo samples in three dimensions at high resolution.,," *Dev Dyn*, vol. 225, pp. 369-75, Nov 2002.
- [257] W. J. Weninger, S. H. Geyer, T. J. Mohun, D. Rasskin-Gutman, T. Matsui, I. Ribeiro, *et al.*, "High-resolution episcopic microscopy: a rapid technique for high detailed 3D analysis of gene activity in the context of tissue architecture and morphology.,," *Anat Embryol (Berl)*, vol. 211, pp. 213-21, Jun 2006.
- [258] H. Huang, C. MacGillivray, H.-S. Kwon, J. Lammerding, J. Robbins, R. T. Lee, *et al.*, "Three-dimensional cardiac architecture determined by two-photon microtomy," *Journal of Biomedical Optics*, vol. 14, Jul-Aug 2009.
- [259] C. Magnain, J. C. Augustinack, M. Reuter, C. Wachinger, M. P. Frosch, T. Ragan, *et al.*, "Blockface histology with optical coherence tomography: A comparison with Nissl staining," *NeuroImage*, vol. 84, pp. 524-33, 2014.
- [260] M. J. Ackerman, "The Visible Human Project," *The Journal of biocommunication*, vol. 18, pp. 14-14, 1991.
- [261] A. MacKenzie-Graham, E. F. Lee, I. D. Dinov, M. Bota, D. W. Shattuck, S. Ruffins, *et al.*, "A multimodal, multidimensional atlas of the C57BL/6J mouse brain," *Journal of Anatomy*, vol. 204, pp. 93-102, Feb 2004.
- [262] J. Rosenthal, V. Mangal, D. Walker, M. Bennett, T. J. Mohun, and C. W. Lo, "Rapid high resolution three dimensional reconstruction of embryos with episcopic fluorescence image capture," *Birth defects research. Part C, Embryo today : reviews*, vol. 72, pp. 213-23, 2004.
- [263] T. Ragan, J. D. Sylvan, K. H. Kim, H. Huang, K. Bahlmann, R. T. Lee, *et al.*, "High-resolution whole organ imaging using two-photon tissue cytometry," *Journal of Biomedical Optics*, vol. 12, Jan-Feb 2007.
- [264] G. J. Kazakia, J. J. Lee, M. Singh, R. F. Bigley, R. B. Martin, and T. M. Keaveny, "Automated high-resolution three-dimensional fluorescence imaging of large biological specimens," *Journal of Microscopy-Oxford*, vol. 225, pp. 109-117, Feb 2007.

- [265] D. Roy, G. J. Steyer, M. Gargesha, M. E. Stone, and D. L. Wilson, "3D Cryo-Imaging: A Very High-Resolution View of the Whole Mouse," *Anatomical Record-Advances in Integrative Anatomy and Evolutionary Biology*, vol. 292, pp. 342-351, Mar 2009.
- [266] D. A. Agard, Y. Hiraoka, P. Shaw, and J. W. Sedat, "Fluorescence microscopy in 3 dimensions," *Methods in Cell Biology*, vol. 30, pp. 353-377, 1989.
- [267] G. J. Steyer, D. Roy, O. Salvado, M. E. Stone, and D. L. Wilson, "Removal of Out-of-Plane Fluorescence for Single Cell Visualization and Quantification in Cryo-Imaging," *Annals of Biomedical Engineering*, vol. 37, pp. 1613-1628, Aug 2009.
- [268] K. P. Keil, V. Mehta, L. L. Abler, P. S. Joshi, C. T. Schmitz, and C. M. Vezina, "Visualization and quantification of mouse prostate development by *in situ* hybridization," *Differentiation*, vol. 84, pp. 232-239, Oct 2012.