A Clinical Raman Spectroscopy System for Real-Time Disease Diagnosis

by

Saumil J. Gandhi

S.B., Electrical Science and Engineering
Massachusetts Institute of Technology (2002)

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 9, 2003

Author____
Department of Electrical Engineering and Computer Science
May 9, 2003

Certified by____
Michael S. Feld
Professor of Physics
Director, G. R. Harrison Spectroscopy Laboratory
Thesis Supervisor

Accepted by____
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

A Clinical Raman Spectroscopy System for Real-Time Disease Diagnosis

by

Saumil J. Gandhi

Submitted to the
Department of Electrical Engineering and Computer Science

May 9, 2003

In Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

The advantages of using Raman spectroscopy for disease diagnosis have been investigated extensively in the past. Unlike conventional techniques that probe the structural and anatomical changes associated with disease, the Raman spectroscopy approach is based on detecting the molecular markers of disease progression. While *in vitro* studies have demonstrated the potential of Raman spectroscopy for diagnosing disease, several important obstacles remain for its use in a clinical setting. Problems in acquiring *in vivo* tissue spectra with acceptable signal to noise ratio in a short collection time and providing diagnostic information in real-time need to be addressed. The goal of this thesis is to present the design and implementation of a Raman spectroscopy system capable of providing real-time diagnostic information *in vivo*. Improvements made to ensure clinically acceptable excitation power and real-time calibration, data analysis, and disease diagnosis are discussed. The system was used to characterize atherosclerotic tissue in human femoral and carotid artery. The results of the *in vivo* experiments show that it is possible to acquire Raman spectra with excellent signal to noise ratio in 1 s. The ability to acquire tissue spectra and extract clinically relevant diagnostic parameters within 3 s allows physicians to assess the disease state in real-time.

Thesis Supervisor: Michael S. Feld
Title: Professor of Physics

Acknowledgements

The work presented in this thesis was performed at the George R. Harrison Spectroscopy Laboratory at the Massachusetts Institute of Technology. I would like to acknowledge the National Institutes of Health, the Lord Foundation, and Pfizer for funding this work.

I would like to thank my thesis advisor Professor Michael S. Feld and Dr. Ramachandra R. Dasari for giving me the opportunity to work on this project. It has been a challenging and rewarding experience that could not have been possible without their guidance and support. Next, I want to thank Jason Motz for being a great mentor, colleague, and friend. I owe a lot of my knowledge about Raman spectroscopy to Jason, who often went out of his way to answer my questions. He also provided invaluable feedback while I was writing my thesis. It has been my pleasure working with him for the past year. I want to thank my roommate Obrad Sceponovic for entertaining my scientific fancies during our nights out in Boston and pushing me to work harder.

Finally, I would like to thank my family for all their love, support, and encouragement. To my sisters Nepa and Niyati for being my two best friends. To my brother-in-law Sanjiv for pushing me to have a clearer view of my scientific interests. I hope to continue our late night chats. This work is a tribute to my parents who always guide me towards achieving my goals. I would not be where I am today without the opportunities they have given me.

To Mom and Dad

Contents

Chapter 1 Introduction **10**

1.1 Motivation 10

1.2 Theory of Raman Scattering 12

1.3 Biomedical Raman Spectroscopy 15

1.3.1 Previous *In Vitro* Work 15

1.3.2 Previous *In Vivo* Work 18

1.4 Thesis Work 19

Chapter 2 Real-Time Clinical Raman System **25**

2.1 Description of the *In Vivo* System 25

2.2 Safety 29

2.3 Real-Time Analysis 32

2.3.1 System Calibration 33

2.3.2 Data Acquisition and Analysis 36

2.3.3 Diagnosis 40

Chapter 3 *In Vivo* Experiments and Results **47**

3.1 Methods 48

3.2 Results 49

3.2.1 *In Vivo* Raman Spectra 49

3.2.2 Signal to Noise and Integration Time 54

3.3 Discussion 57

Chapter 4 Conclusion 60

4.1 Accomplishments 60

4.2 Future Directions 61

4.2.1 Side-Viewing Raman Probes 61

4.2.2 *In Vivo* Studies 62

Appendix A LabVIEW Implementation 64

Appendix B Matlab Subroutines 79

List of Figures

Figure 1.1: A photon with certain energy excites a molecule and scatters elastically (Rayleigh scattering) or inelastically (Raman scattering). 13

Figure 2.1: Clinical Raman system. 28

Figure 2.2: Real-time data acquisition process. 30

Figure 2.3: Teflon spectrum acquired with a single-ring Raman probe. 31

Figure 2.4: Dependency of laser power incident at the microscope objective on external modulation voltage. Solid line is the linear fit through measured data points. 32

Figure 2.5: Tylenol spectrum after background removal 34

Figure 2.6: White light spectrum used to correct for wavelength dependent system responses. 35

Figure 2.7: Aluminum spectrum for characterizing the probe background. 36

Figure 2.8: Data flow in the real-time clinical Raman system. 37

Figure 2.9: Raman spectrum of a non-atherosclerotic femoral artery (a) unprocessed, (b) after spectral response correction, (c) after background subtraction, and (d) after fluorescence removal. 40

Figure 2.10: Real-time display of tissue Raman spectrum, model fit, residual, fit contributions, and diagnosis of a normal tissue sample in the femoral artery. The collection time for the spectrum was 1 s with 120 mW excitation power. The total time for data acquisition and analysis was under 3 s. 43

Figure 3.1: Typical Raman spectra of femoral artery tissue from each of the three diagnostic classes: (a) non-atherosclerotic tissue, (b) non-calcified atherosclerotic plaque, and (c) calcified atherosclerotic plaque. The dotted line is the measured spectrum and solid line the model fit. The lower line in each graph is the residual (data minus fit). 50

Figure 3.2: Hematoxylin and eosin stained sectioning used for histologic confirmation of the diagnoses determined by analysis of the spectra shown in Figure 3.1 for (a) non-atherosclerotic tissue (4x), (b) non-calcified plaque (20x), and (c) calcified plaque (4x). 53

Figure 3.3: Raman spectra of non-calcified atherosclerotic femoral artery tissue at 1 s intervals. 54

Figure 3.4: The mean ± standard deviation of the fit contributions of five 1 s spectra from a tissue site in each of the three diagnostic classes 55

List of Tables

Table 2.1: Timing breakdown for data acquisition, analysis, and disease diagnosis in the real-time system. 44

Table 3.1: Relative fit contributions of the eight morphologic structures for the spectra shown in Figure 3.1. (a) Non-atherosclerotic tissue, (b) non-calcified plaque, and (c) calcified plaque. 51

Table 3.2: Correlation between mean of the fit contributions of the five 1 s spectra and the contributions for the entire 5 s collection time in all 34 cases. 56

Chapter 1

Introduction

1.1 Motivation

Cancer and heart disease are the two leading causes of death in the United States. About 7 million Americans are currently suffering from coronary heart disease, the most common form of heart disease. 1.2 million new cases of cancer are diagnosed each year. Cancer claimed 549,838 lives in 1999, while over 528,000 people died from coronary heart disease [1]. Early detection and treatment of disease is crucial for reducing the mortality rate. Conventional imaging, however, is inadequate for detecting the presence of early stages of disease. Structural and anatomical information probed by present techniques often fails to provide the very information needed by physicians to make an accurate diagnosis and prescribe proper treatment. Furthermore, time consuming and expensive histopathologic examination of a tissue biopsy is often needed to make a definitive diagnosis for most diseases. Such techniques are invasive and require a large number of biopsies from random locations for accurate diagnosis. In addition, histopathologic examination of arterial tissue is not possible since arteries cannot be biopsied.

Advances in biomedical spectroscopy have been driven by the need to provide minimally invasive, objective, and quantitative diagnostic information in a timely fashion. Applications of lasers and optical techniques such as reflectance, fluorescence, light scattering, and Raman spectroscopies are playing an increasingly important role in biomedical sciences. Fluorescence spectroscopy has been able to successfully diagnose atherosclerosis, as well as cancer in the colon, cervix, oral cavity, and lung. This technique, however, is limited due to the lack of sharp spectral features and relatively few fluorophores in tissue. Infrared spectroscopy has been used to study breast, lung, and colon cancers, but it is hindered by the effects of strong water absorption in that range.

Among the optical spectroscopy techniques, Raman scattering can provide the most detailed information about the chemical composition of the tissue under study. Unlike infrared spectroscopy, Raman spectroscopy can avoid the effects of water absorption by using excitation wavelengths in the near-infrared (NIR) region, where water absorption is minimized. NIR excitation also minimizes generation of the broad spectral features of fluorescence. Raman scattering, therefore, has been used extensively in biology and biochemistry to study the structure of biologically relevant molecules. Unlike conventional techniques that probe the structural and anatomical changes caused by disease, the Raman spectroscopy approach is based on detecting the molecular markers of disease progression. The relative peak intensities and spectral positions of various sharp bands yield fingerprints of relevant molecular components. Since the progression of disease is usually accompanied by chemical changes, the molecular composition of tissue can be used to provide important disease diagnostic information.

Much evidence indicates that Raman spectroscopy has the potential to provide real-time diagnostic information with histologic accuracy without removing tissue. The physician's ability to diagnose disease with detailed chemical analysis in real-time, combined with the potential to monitor disease progression over time is crucial for developing new treatment approaches and reducing patient mortality and morbidity.

1.2 Theory of Raman Scattering

Before discussing the use Raman spectroscopy for biomedical applications, we briefly recall the Raman effect. A detailed discussion of this topic can be found in McCreery [2]. The Raman effect is one of the basic interactions between light and matter. In the photon view, light incident on matter can be absorbed or scattered. The scattering can be either elastic or inelastic. Most of the time the light is scattered elastically and the emerging photons exit with the same energy $h\nu_0$, and frequency ν_0, as the incident photons, a process called Rayleigh scattering. Inelastically scattered Raman photons, however, are shifted up or down in frequency by the characteristic vibrational energy $h\nu_i$ of the molecule (Figure 1.1). During Stokes Raman scattering, the incident photon loses energy by exciting the molecule from its ground state to an excited vibrational state. As a result, the scattered photon appears at a lower frequency $\nu_s = \nu_0 - \nu_i$ relative to the incident photon. During anti-Stokes Raman scattering, the scattered light appears at a higher frequency $\nu_{as} = \nu_0 + \nu_i$ since it gains energy by interacting with the molecule which is initially in an excited vibrational state.

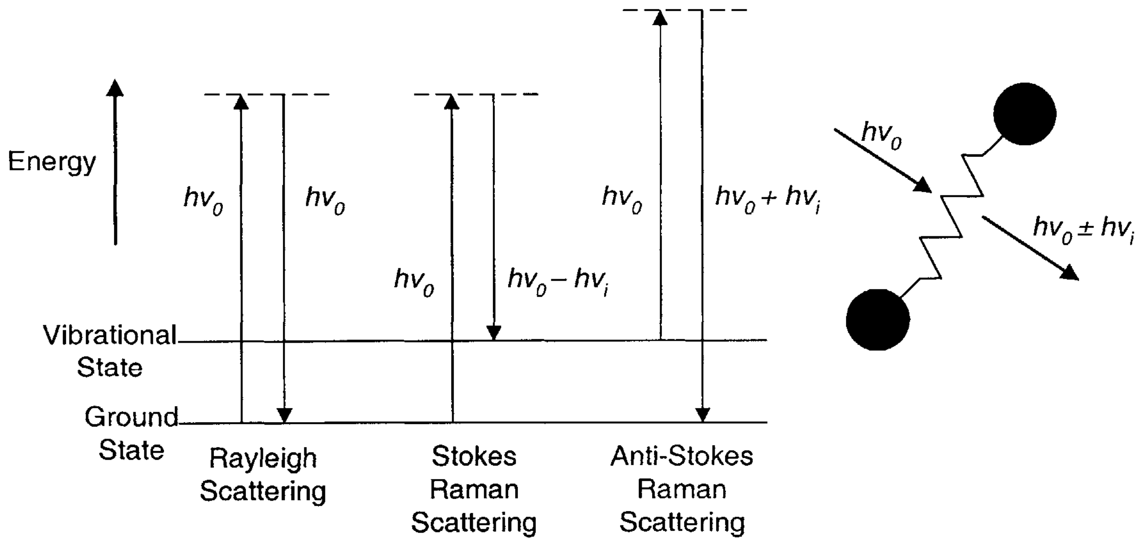


Figure 1.1: A photon with certain energy excites a molecule and scatters elastically (Rayleigh scattering) or inelastically (Raman scattering).

According to the classical view, light acts as an electromagnetic wave inducing a dipole moment in the molecule. The induced dipole, which reradiates scattered light with or without exchanging energy with vibrational states of the molecule, is determined as

$$P = \alpha E \tag{1.1}$$

where P is the strength of the induced dipole, α is the polarizability, and E is the incident electric field. The incident electric field of a monochromatic plane wave is

$$E = E_0 \cos 2\pi\nu_0 t \tag{1.2}$$

where ν_0 is the frequency of the incident laser light. A molecule with N atoms has $3N$-6 vibrational normal modes. The vibrational mode Q_i of the ith harmonic frequency ν_i is

$$Q_i = Q_i^0 \cos 2\pi\nu_i t. \tag{1.3}$$

To a first-order approximation, the polarizability of electrons in the molecule is modulated by the *i*th vibrational mode such that

$$\alpha = \alpha_0 + \left(\frac{d\alpha}{dQ_i}\right)Q_i. \tag{1.4}$$

Substituting Eq. (1.3) into Eq. (1.4),

$$\alpha = \alpha_0 + Q_i^0\left(\frac{d\alpha}{dQ_i}\right)\cos 2\pi\nu_i t. \tag{1.5}$$

The resulting strength of the induced dipole is obtained by multiplying Eq. (1.5) with Eq. (1.2) and using the trigonometric identity given by Eq. (1.6).

$$\cos A \cos B = \frac{1}{2}\left[\cos(A+B) + \cos(A-B)\right], \tag{1.6}$$

$$\therefore P = \alpha_0 E_0 \cos 2\pi\nu_0 t + E_0 Q_i^0\left(\frac{d\alpha}{dQ_i}\right)\frac{\cos 2\pi(\nu_0+\nu_i)t + \cos 2\pi(\nu_0-\nu_i)t}{2}. \tag{1.7}$$

As can be seen from Eq. (1.7) the induced dipole reradiates the incident light at three different frequencies. The first term, which is at the same frequency as the incident laser light, describes Rayleigh scattering. The second term at a higher frequency $\nu_0 + \nu_i$ represents the anti-Stokes Raman scattering, and the third term at $\nu_0 - \nu_i$ is Stokes Raman scattering.

The classical view provides an insight which is very useful for designing Raman spectroscopy systems with a high signal to noise ratio (SNR). The polarization and scattering intensities for both Rayleigh and Raman scattering are linearly dependent on laser intensity. More importantly, the value of $(d\alpha/dQ_i)$ is generally much smaller than

α_0. As a result, the Raman effect is several orders of magnitude weaker than Rayleigh scattering and requires the use of notch filters for rejecting the intense Rayleigh light.

It is important to note one limitation of the classical view. Eq. (1.7) predicts equal magnitudes for both Stokes and anti-Stokes Raman scattering. However, according to quantum mechanics, the probability of finding a molecule in a particular state is given by the Boltzmann distribution. Since the ground state is generally more populated than the vibrational states, the magnitude of Stokes Raman scattering is much larger than the magnitude of anti-Stokes Raman scattering. Therefore, Stokes Raman scattering is usually the observed Raman signal.

1.3 Biomedical Raman Spectroscopy

1.3.1 Previous *In Vitro* Work

The potential of Raman spectroscopy to diagnose disease by probing the biochemical makeup of tissue has been documented in several review articles [3-8]. Raman spectroscopy systems with excellent SNR have been proposed to study tissue *in vitro* [9-12]. Recent studies have investigated the ability to distinguish normal and diseased tissue in the brain, skin, cervix, gastrointestinal tract, breast, and arteries. Mizuno, *et al.* used Raman spectroscopy to study brain tissue and tumor [13, 14]. Raman spectroscopy has also been assessed for detecting Parkinson's disease [15] as well as brain edema [16]. Casepers, *et al.* have studied the molecular composition of different skin layers to understand the relation between skin disease and biochemical changes in the skin [17]. Distinctive molecular abnormalities in skin neoplasia have also been studied [18].

Several studies have investigated the use of Raman spectroscopy for detecting cancerous lesions in the cervix, gastrointestinal tract, and breast [19]. Mahadevan-Jansen, *et al.* reported 82% sensitivity and 92% specificity for differentiating precancerous tissue from all other tissues in the cervix [20]. Studies have been performed to demonstrate the potential of detecting premalignant gastrointestinal lesions [21, 22]. Raman spectroscopy has also been used extensively to characterize normal and cancerous breast tissue [23, 24]. Haka, *et al.* have reported 88% sensitivity and 93% specificity for identifying microcalcifications in benign and malignant breast lesions by probing the differences in their chemical composition [25, 26]. A chemical and morphological model of cancerous lesions in the breast has been developed to understand the state of disease progression [27]. The model relates Raman spectral features to diagnostic parameters used by pathologists.

Similar work has also been done to characterize arterial tissue using Raman spectroscopy. Several *in vitro* studies have been performed to characterize the biochemical makeup of atherosclerotic plaques [28-32]. Buschman, *et al.* were able to perform a chemical analysis of individual cellular and extracellular components of atherosclerotic lesions during different states of disease progression [33]. The study showed that Raman spectra of various morphologic structures could be modeled by a linear combination of basis spectra obtained from several biochemicals present in arterial tissue. Furthermore, the basis Raman spectra of eight different morphologic structures were used in a linear least squares minimization model to calculate the contribution of each morphologic structure to intact coronary artery tissue spectra [34]. The morphologic structures used to characterize coronary artery tissue spectra included calcified minerals, β-carotene, cholesterol crystals, foam cell/necrotic core, adventitial adipocytes, smooth muscle cells, collagen fibers, and elastic lamina.

A diagnostic algorithm was developed to separate the samples in a calibration data set into three categories: non-atherosclerotic tissue, calcified plaque, and non-calcified plaque [34]. Optimum separation between the three diagnostic classes was achieved by using logistic regression on fit contributions of various morphologic structures. The lines separating the three regions are

$$CM = -0.07 + 0.31\left(FCNC_{NCR} + CC_{NCR}\right) \tag{1.6}$$

$$CM = 0.17 - 0.48\left(FCNC_{NCR} + CC_{NCR}\right) \tag{1.7}$$

$$CM = -0.07 + 0.30\left(FCNC_{NCR} + CC_{NCR}\right) \tag{1.8}$$

where CM is the fit contribution of calcified minerals. $FCNC_{NCR}$, and CC_{NCR} are the fit contributions of foam cells/necrotic core, and cholesterol crystals, respectively, in a non-calcified region. Eq. (1.6) separates non-atherosclerotic tissue from non-calcified plaque. Eq. (1.7) separates non-atherosclerotic tissue from calcified plaque, while Eq. (1.8) separates non-calcified plaque from calcified plaque. The algorithm was tested on a validation data set and was able to correctly classify 64 (94%) of 68 coronary artery samples.

1.3.2 Previous *In Vivo* Work

There are three main obstacles for extending the *in vitro* Raman spectroscopy analysis to tissue *in vivo*. First, the Raman signal from tissue is extremely weak and might require prohibitive excitation powers or long collection times to obtain spectra with acceptable SNR. Second, the fluorescence signal from *in vivo* tissue is several orders of magnitude larger than the Raman signal and masks the sharp Raman features. Finally, large Raman and fluorescence background signals are generated in the optical fibers when the excitation light travels through the delivery fiber and the Rayleigh scattered light travels back through the collection fibers. The large backgrounds contaminate the tissue spectra and contribute significant shot noise that can overwhelm the signal of interest.

Raman spectrometer systems designed to minimize the effects of *in vivo* tissue fluorescence and maximize SNR have been proposed previously [35-37]. Extensive work has also been done in the past to develop high-throughput optical fiber Raman probes with proper filtering capabilities to suppress the effects of fiber background generated by the excitation light [38-41]. Some efforts have been made to characterize normal and diseased *in vivo* tissue with Raman spectroscopy [42,43]. However, inadequate filtering as well as low throughput and collection efficiencies have kept these systems from acquiring *in vivo* data with good SNR in clinically acceptable times. *In vivo* studies for detecting precancerous tissue in the cervix were only possible with collection times that exceeded several minutes [39, 44, 45]. Developments in endoscopic imaging have shown some promise for detecting premalignant gastrointestinal lesions without requiring a biopsy [46]. Preliminary work has also been done to determine the molecular composition of arterial wall *in vivo* [47].

The obstacles outlined above, however, continue to hinder the development of *in vivo* Raman spectroscopy. Much of the previous work was done with large excitation power and collection times on the order of many seconds or minutes to achieve acceptable SNR. Furthermore, the data calibration and analysis were performed off-line after all of the spectra were acquired. As a result, physicians were unable to assess the disease state in real-time.

1.4 Thesis Work

The development of an optical fiber Raman probe by Motz, *et al.* has significantly improved the SNR of previously used clinical Raman systems, thus allowing data acquisition with shorter collection times [48, 49]. The goal of this thesis is to present the design and implementation of a system capable of real-time data acquisition, analysis, and disease diagnosis.

Hardware and software improvements, made to ensure clinically acceptable excitation power and real-time calibration, data analysis, and disease diagnosis are discussed in Chapter 2. *In vivo* experiments performed on human femoral and carotid artery tissue are presented in Chapter 3. The goal of the *in vivo* study is to see whether the *in vitro* models and diagnostic algorithms developed by Buschman, *et al.* could be extended to a clinical environment [33, 34]. The experiments are also used to investigate whether the new system can be used to acquire *in vivo* Raman spectra with acceptable

SNR in short enough collection times to allow for real-time disease diagnosis. Finally, the conclusion and future directions are presented in Chapter 4.

1. Anderson R, *Deaths: Leading Causes for 1999.* National Vital Statistics Reports, 2001. **49**(11).
2. McCreery R, *Raman Spectroscopy for Chemical Analysis.* 2000, New York.
3. Hanlon EB, Manoharan R, Koo TW, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, and Feld MS, *Prospects for in vivo Raman spectroscopy.* Physics in Medicine and Biology, 2000. **45**(2): p. R1-R59.
4. Pappas D, Smith BW, and Winefordner JD, *Raman spectroscopy in bioanalysis.* Talanta, 2000. **51**(1): p. 131-144.
5. Kalasinsky VF, *Biomedical applications of infrared and Raman microscopy.* Applied Spectroscopy Reviews, 1996. **31**(3): p. 193-249.
6. Richards-Kortum R and Sevick-Muraca E, *Quantitative optical spectroscopy for tissue diagnosis.* Annual Review of Physical Chemistry, 1996. **47**: p. 555-606.
7. Manoharan R, Wang Y, and Feld MS, *Histochemical analysis of biological tissues using Raman spectroscopy.* Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy, 1996. **52**(2): p. 215-249.
8. Fabian H and Anzenbacher P, *New Developments in Raman-Spectroscopy of Biological-Systems.* Vibrational Spectroscopy, 1993. **4**(2): p. 125-148.
9. Brennan JF, Wang Y, Dasari RR, and Feld MS, *Near-infrared Raman spectrometer systems for human tissue studies.* Applied Spectroscopy, 1997. **51**(2): p. 201-208.
10. Baraga JJ, Feld MS, and Rava RP, *Rapid near-Infrared Raman-Spectroscopy of Human Tissue with a Spectrograph and CCD Detector.* Applied Spectroscopy, 1992. **46**(2): p. 187-190.
11. Jongsma FHM, Erckens RJ, Wicksted JP, Bauer NJC, Hendrikse F, March WF, and Motamedi M, *Confocal Raman spectroscopy system for noncontact scanning of ocular tissues: an in vitro study.* Optical Engineering, 1997. **36**(11): p. 3193-3199.
12. Puppels GJ, Colier W, Olminkhof JHF, Otto C, de Mul FFM, and Greve J, *Description and Performance of a Highly Sensitive Confocal Raman Microspectrometer.* Journal of Raman Spectroscopy, 1991. **22**: p. 217-225.
13. Mizuno A, Kitajima H, Kawauchi K, Muraishi S, and Ozaki Y, *Near-Infrared Fourier-Transform Raman-Spectroscopic Study of Human Brain-Tissues and Tumors.* Journal of Raman Spectroscopy, 1994. **25**(1): p. 25-29.
14. Mizuno A, Hayashi T, Tashibu K, Maraishi S, Kawauchi K, and Ozaki Y, *Near-Infrared Ft-Raman Spectra of the Rat-Brain Tissues.* Neuroscience Letters, 1992. **141**(1): p. 47-52.
15. Ong CW, Shen ZX, He Y, Lee T, and Tang SH, *Raman microspectroscopy of the brain tissues in the substantia nigra and MPTP-induced Parkinson's disease.* Journal of Raman Spectroscopy, 1999. **30**(2): p. 91-96.
16. Wolthuis R, van Aken M, Fountas K, Robinson JS, Bruining HA, and Puppels GJ, *Determination of water concentration in brain tissue by Raman spectroscopy.* Analytical Chemistry, 2001. **73**(16): p. 3915-3920.

17. Caspers PJ, Lucassen GW, Wolthuis R, Bruining HA, and Puppels GJ, *In Vitro and In Vivo Raman Spectroscopy of Human Skin.* Biospectroscopy, 1998. **4**: p. S31-S39.
18. Gniadecka M, Wulf HC, Nielsen OF, Christensen DH, and Hercogova J, *Distinctive molecular abnormalities in benign and malignant skin lesions: Studies by Raman spectroscopy.* Photochemistry and Photobiology, 1997. **66**(4): p. 418-423.
19. Mahadevan-Jansen A and Richards-Kortum R, *Raman Spectroscopy for the Detection of Cancers and Precancers.* Journal of Biomedical Optics, 1996. **1**(1): p. 31-70.
20. Mahadevan-Jansen A, Mitchell MF, Ramanujam N, Malpica A, Thomsen S, Utzinger U, and Richards-Kortum R, *Near-infrared Raman spectroscopy for in vitro detection of cervical precancers.* Photochemistry and Photobiology, 1998. **68**(1): p. 123-132.
21. Barr H, Dix T, and Stone N, *Optical spectroscopy for the early diagnosis of gastrointestinal malignancy.* Lasers in Medical Science, 1998. **13**(1): p. 3-13.
22. Bohorfoush AG, *Tissue spectroscopy for gastrointestinal diseases.* Endoscopy, 1996. **28**(4): p. 372-380.
23. Shafer-Peltier KE, Haka AS, Fitzmaurice M, Crowe J, Myles J, Dasari RR, and Feld MS, *Chemical basis for breast cancer diagnosis using Raman spectroscopy.* Lasers in Surgery and Medicine, 2002: p. 3.
24. Kline NJ and Treado PJ, *Raman Chemical Imaging of Breast Tissue.* Journal of Raman Spectroscopy, 1997. **28**: p. 119-124.
25. Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, and Feld MS, *Identifying microcalcifications in benign and malignant breast lesions by probing differences in their chemical composition using Raman spectroscopy.* Cancer Research, 2002. **62**(18): p. 5375-5380.
26. Haka AS, Shafer KE, Fitzmaurice M, Dasari RR, and Feld MS, *Distinguishing type II microcalcifications in benign and malignant breast lesions using Raman spectroscopy.* Modern Pathology, 2002. **15**(1): p. 137.
27. Shafer-Peltier KE, Haka AS, Fitzmaurice M, Crowe J, Myles J, Dasari RR, and Feld MS, *Raman microspectroscopic model of human breast tissue: implications for breast cancer diagnosis in vivo.* Journal of Raman Spectroscopy, 2002. **33**(7): p. 552-563.
28. Manoharan R, Baraga JJ, Feld MS, and Rava RP, *Quantitative Histochemical Analysis of Human Artery using Raman Spectroscopy.* Journal of Photochemistry and Photobiology B-Biology, 1992. **16**: p. 211-233.
29. Brennan JF, Römer TJ, Lees RS, Tercyak AM, Kramer JR, and Feld MS, *Determination of Human Coronary Artery Composition by Raman Spectroscopy.* Circ, 1997. **96**(1): p. 99-105.
30. Salenius JP, Brennan JF, Miller A, Wang Y, Aretz T, Sacks B, Dasari RR, and Feld MS, *Biochemical Composition of Human Peripheral Arteries Examined with Near-Infrared Raman Spectroscopy.* Journal of Vascular Surgery, 1998. **27**: p. 710-719.

31. Römer T, Brennan J, Bakker Schut T, Wolthuis R, can den Hoogen R, Emeis J, van der Laarse A, Bruschke A, and Puppels G, *Raman Spectroscopy for Quantifying Cholesterol in Intact Coronary Artery Wall.* Atherosclerosis, 1998. **141**: p. 117-124.
32. Romer TJ, Brennan JF, Fitzmaurice M, Feldstein ML, Deinum G, Myles JL, Kramer JR, Lees RS, and Feld MS, *Histopathology of human coronary atherosclerosis by quantifying its chemical composition with Raman spectroscopy.* Circulation, 1998. **97**(9): p. 878-885.
33. Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, and Feld MS, *Raman microspectroscopy of human coronary atherosclerosis: Biochemical assessment of cellular and extracellular morphologic structures in situ.* Cardiovascular Pathology, 2001. **10**(2): p. 69-82.
34. Buschman HP, Motz JT, Deinum G, Romer TJ, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, and Feld MS, *Diagnosis of human coronary atherosclerosis by morphology-based Raman spectroscopy.* Cardiovascular Pathology, 2001. **10**(2): p. 59-68.
35. Hanlon EB, Manoharan R, Koo T-W, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, and Feld MS, *Prospects for In Vivo Raman Spectroscopy.* Physics in Medicine and Biology, 2000. **45**(2): p. R1-R59.
36. Kaminaka S, Ito T, Yamazaki H, Kohda E, and Hamaguchi H, *Near-infrared multichannel Raman spectroscopy toward real-time in vivo cancer diagnosis.* Journal of Raman Spectroscopy, 2002. **33**(7): p. 498-502.
37. Shim MG and Wilson BC, *Development of an in vivo Raman spectroscopic system for diagnostic applications.* Journal of Raman Spectroscopy, 1997. **28**(2-3): p. 131-142.
38. Utzinger U and Richards-Kortum R, *Fiber Optic Probes for Biomedical Optical Spectroscopy.* Journal of Biomedical Optics, 2001.
39. Mahadevan-Jansen A, Mitchell WF, Ramanujam N, Utzinger U, and Richards-Kortum R, *Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue in vivo.* Photochemistry and Photobiology, 1998. **68**(3): p. 427-431.
40. Shim M, Wilson B, Marple E, and Wach M, *Study of Fiber-Optic Probes for in Vivo Medical Raman Spectroscopy.* Appl Spectrosc, 1999. **53**(6): p. 619-627.
41. de Lima CJ, Sathaiah S, Silveira L, Zangaro RA, and Pacheco MTT, *Development of catheters with low fiber background signals for Raman spectroscopic diagnosis applications.* Artificial Organs, 2000. **24**(3): p. 231-234.
42. Wolthuis R, Bakker Schut TC, Caspers PJ, Buschman HPJ, Romer TJ, Bruining HA, and Puppels GJ, *Raman Spectroscopic Methods for In Vitro and In Vivo Tissue Characterization*, in *Fluorescent and Luminescent Probes*, W.T. Mason, Editor. 1999, Academic Press: London. p. 433-455.
43. Bakker Schut TC, Withes MJH, Sterenborg HJCM, Speelman OC, Roodenburg JLN, Marple ET, Bruining HA, and Puppels GJ, *In Vivo Detection of Dysplastic Tissue by Raman Spectroscopy.* Analytical Chemistry, 2000. **72**(24): p. 6010-6018.

44. Utzinger U, Heintzelman DL, Mahadevan-Jansen A, Malpica A, Follen M, and Richards-Kortum R, *Near-infrared Raman spectroscopy for in vivo detection of cervical precancers.* Applied Spectroscopy, 2001. **55**(8): p. 955-959.
45. Scaiano JC, *Compact NIR-Raman probe for cervical tissue in vivo.* Photochemistry and Photobiology, 1998. **68**(3): p. IV-IV.
46. Dacosta RS, Wilson BC, and Marcon N, *New optical technologies for earlier endoscopic diagnosis of premalignant gastrointestinal lesions.* Journal of gastroenterology and hepatology, 2002.
47. Buschman HP, Marple ET, Wach ML, Bennett B, Schut TCB, Bruining HA, Bruschke AV, van der Laarse A, and Puppels GJ, *In vivo determination of the molecular composition of artery wall by intravascular Raman spectroscopy.* Analytical Chemistry, 2000. **72**(16): p. 3771-3775.
48. Motz JT, Hunter M, Galindo LH, Gardecki JA, Kramer JR, Dasari RR, and Feld MS, *Optical Fiber Raman Probe for Biomedical Spectroscopy.* In Preparation.
49. Motz JT, *Development of In Vivo Raman Spectroscopy of Atherosclerosis*, in *Health Sciences and Technology*. 2003, Massachusetts Institute of Technology: Cambridge, Massachusetts.

Chapter 2

Real-Time Clinical Raman System

In the previous chapter, we discussed *in vitro* studies that can prospectively classify atherosclerotic tissue with >94% accuracy. The development of a mobile Raman spectroscopy system based on optical fiber probes for light delivery and collection will allow us to perform similar studies *in vivo*. This chapter briefly describes the design and development of a compact Raman spectroscopy system suitable for clinical work. It also gives a detailed description of the software and hardware modifications made to provide real-time disease diagnosis in a clinical setting.

2.1 Description of the *In Vivo* System

The Raman system used for our *in vivo* work is similar to those described by Brennan, *et al.* [1] and Shim, *et al.* [2]. The system uses an 830 nm InGaAs diode laser (Process Instruments, Salt Lake City, UT). Tissue fluorescence is an obstacle for *in vivo* Raman measurements since it is generally several orders of magnitude larger than the Raman signal and masks the sharp Raman features. However, the use of near-infrared (NIR) excitation wavelengths, rather than the shorter visible wavelengths, reduces the fluorescence to a manageable level [3]. This is because fluorescence typically originates from the lowest excited electronic states of molecules. These states have energies corresponding to visible wavelengths and therefore cannot be excited by the lower energy NIR light. Although the fluorescence background can be further reduced by using excitation wavelengths longer than 830 nm, it prohibits the use of high-sensitivity charge coupled device (CCD) detectors. Excitation wavelengths shorter than 270 nm in the ultraviolet region also reduce the fluorescence background. However, these wavelengths have shallow penetration depths in tissue, allowing only superficial features to be probed. Furthermore, ultraviolet light at these wavelengths is also mutagenic [3].

As shown in Figure 2.1, the 830 nm laser beam is collimated by a pair of cylindrical collimating lenses CL_1 and CL_2. The collimated beam then passes through a bandpass filter (BP), which transmits greater than 90% at 830 nm and disperses other wavelengths away from the beam path. The filtered laser beam is incident on a mirror (M), which directs it through a 10x microscope objective (MO) for focusing onto the 200 μm core diameter excitation fiber. Transmission of the laser beam to the probe is controlled by a high-speed, 6 mm aperture, computer-controlled shutter (Vincent Associates, Rochester, NY) between the mirror and the microscope objective (cf. Section 2.2).

A bifurcated optical fiber catheter is used to transport the excitation and scattered light to and from the tissue sample at the distal end of the catheter. Fifteen collection fibers, each with a 200 μm core diameter, surround a single excitation fiber at the distal end of the probe. When the excitation laser light passes through the fused silica of the excitation fiber, it generates an intense background signal. The fiber background then undergoes Rayleigh scattering and enters the collection fibers along with the Raman light from tissue. The shot noise from the fiber background can often be larger than the tissue Raman signal itself. Therefore, a short pass filter is placed at the tip of the excitation fiber such that only the 830 nm excitation laser light reaches the tissue. In addition, a long pass filter is placed on the tips of the collection fibers in order to prevent the Rayleigh scattered excitation light from entering the collection fibers and generating additional fiber background.

At the proximal end, the fifteen collection fibers are aligned at the entrance of the spectrograph slit in a linear array. The numerical aperture (NA) of the collection fibers is *f*/# matched to the collection cone of the Holospec *f*/1.8 spectrograph (Kaiser Optical Systems, Ann Arbor, MI) (NA≈0.26) to conserve throughput. Conventional reflective spectrographs often introduce aberrations from the collimating and focusing optics. The Holospec spectrograph reduces the effects of this problem by using transmission optics as well as a holographic notch filter (NF) and grating (G) designed to correct for off-axis abberations. Spherical aberrations, however, still exist for large apertures. The notch filter is needed to remove the Rayleigh scattered 830 nm excitation light and prevent it from saturating the CCD detector. The grating used for the clinical system has a spectral range of 0-1850 cm^{-1} and provides a resolution of ~9 cm^{-1}.

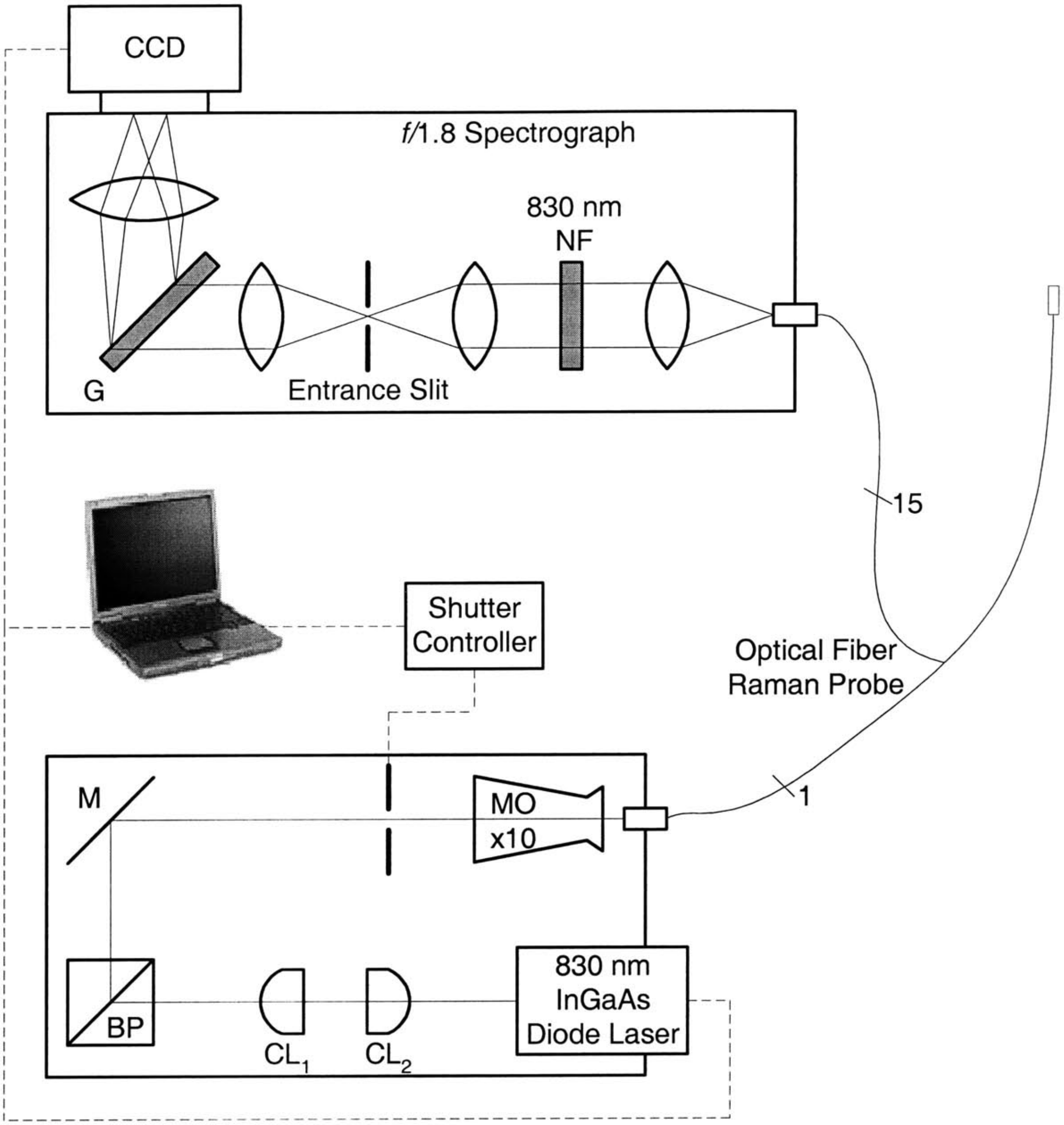


Figure 2.1: Clinical Raman system.

The dispersed light from the grating in the spectrograph is incident on a highly sensitive, back illuminated CCD detector (Princeton Instruments, Trenton, NJ). The CCD has a detector array of 1340 x 400 pixels. The quantum efficiency of the detector at 800 nm is ~75% and falls off to about 20% at 1000 nm. The full-well capacity of a single pixel is 250,000 electrons with a dark current of 0.01 electrons per pixel per second at a temperature of -100°C. The *in vivo* spectra are obtained by vertically binning the signal from the fifteen collection fibers. In order to minimize the noise resulting from dark current, pixels that do not receive light from the fifteen collection fibers are not binned.

2.2 Safety

The *in vivo* Raman system outlined in the previous section must adhere to several safety guidelines in the clinical environment. First, any part of the system that touches the patient directly or indirectly must be kept sterile. Next, the excitation laser power cannot exceed a predetermined threshold value. Finally, none of the stray laser light can leak out of the system. A few modifications must be made to the *in vivo* system in order to address these issues.

Accurate control of the excitation laser power is accomplished with LabVIEW, V.6.1 (National Instruments, Austin, TX). LabVIEW provides flexibility for interfacing with various devices such as the laser, laser shutter, and CCD detector.

Since we cannot risk desterilizing the probe tip by measuring the excitation power with a power meter, we use a sterilized block of Teflon to calibrate the excitation laser power. The diagram in the upper half of Figure 2.2 depicts the feedback loop for setting the proper excitation laser power in a sterile operating room. A Teflon spectrum obtained prior to the procedure is used to determine the expected signal (target intensity) for the desired excitation power. A sample Teflon spectrum is shown in Figure 2.3. The most intense features in the spectrum are due to the remaining quartz background from the probe. Peaks corresponding to Raman shifts of Teflon can be seen at 1060, 1130, 1295, and 1440 cm^{-1}.

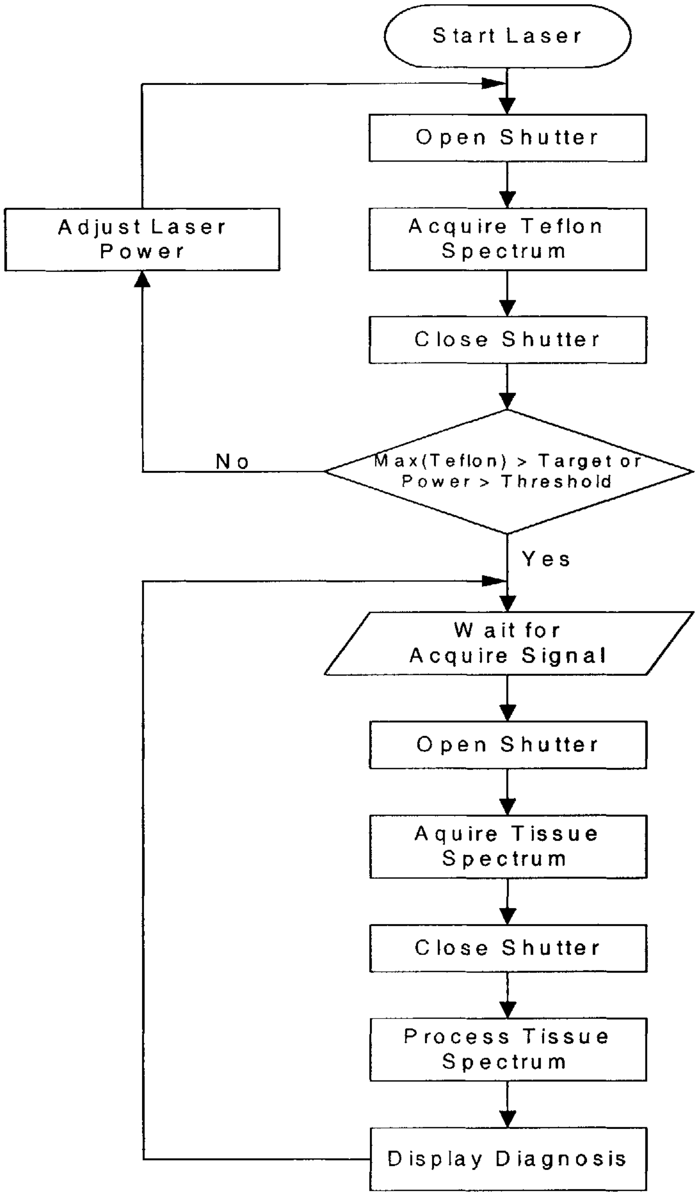


Figure 2.2: Real-time data acquisition process.

During the procedure, spectra of an identical sterilized Teflon block are taken and their peak intensities are compared to the target intensity. LabVIEW continues to adjust the laser power automatically until the target intensity is obtained, or until a predetermined threshold laser power is reached.

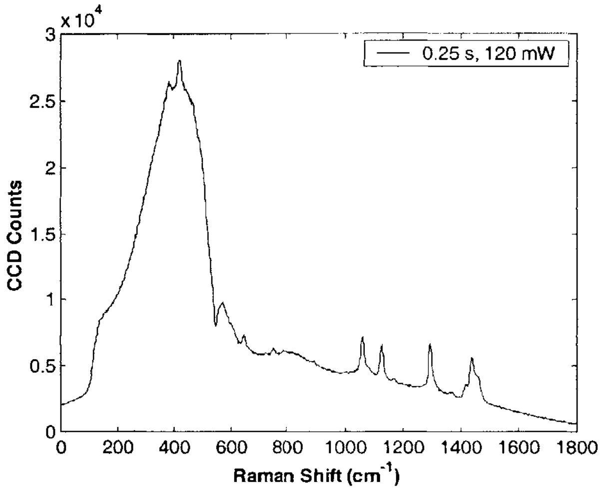


Figure 2.3: Teflon spectrum acquired with a single-ring Raman probe.

The laser power is modulated remotely through an analog waveform provided by a data acquisition card (National Instruments, Austin, TX). Figure 2.4 shows the dependency of laser power incident at the microscope objective (Figure 2.1) on the amplitude of the analog voltage waveform provided by LabVIEW. Since optical component losses differ from system to system, this relationship must be established empirically for each configuration. Although there is a non-linearity in the observed power at low voltages, the relationship is linear around our operating range of 100 mW. The linear relationship allows us to adjust the laser power with a feedback loop easily and reliably.

Finally, the utmost care must be taken to ensure that none of the stray light from the laser leaks out of the system. The optical components of the system, therefore, are encased in a tight black case. Furthermore, the laser beam is blocked with a high-speed, computer-controlled shutter. The shutter opens automatically just before data acquisition begins and closes immediately after the acquisition is complete (Figure 2.2). The shutter then remains closed until the system receives a signal through Labview for another acquisition.

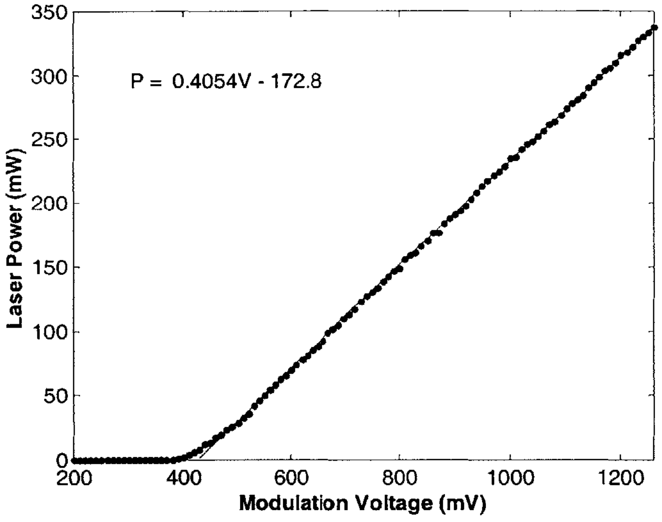


Figure 2.4: Dependency of laser power incident at the microscope objective on external modulation voltage. Solid line is the linear fit through measured data points.

2.3 Real-Time Analysis

Raman spectroscopy has wide-ranging diagnostic applications from cancer detection to Alzheimer's disease and atherosclerosis diagnosis and its potential to characterize *in vivo* tissue has been well established by Hanlon, *et al.* [3]. Recent advances in Raman probe design promise the collection of *in vivo* spectra in clinically realistic times (cf. Chapter 3). However, all of the analysis has previously been performed offline. In order to effectively use Raman spectroscopy as a minimally invasive tool for making real-time diagnosis in a clinical setting, we need to develop software that can acquire data and rapidly perform the necessary calibrations, signal processing, and data analysis. Such software will then allow us to extract the relevant diagnostic parameters in real-time.

LabVIEW is the primary platform for data acquisition and analysis in the real-time clinical system. In addition to providing flexibility for interfacing with various devices in the system, LabVIEW allows automation of calibration and data analysis routines by providing a direct interface to Matlab, V.6.5.0 (The Mathworks, Natick, MA). The modular design of the real-time system provides easily adaptable Matlab calibration and data analysis routines. Various model basis spectra for the particular disease we are studying can also be incorporated easily into the system. For the purpose of demonstrating the working real-time system, we will use atherosclerotic disease as an example. In the following sections, we first describe the system calibration and data analysis steps needed for the *in vivo* experiments described in Chapter 3. We then take a look at how these steps are automated to make a real-time diagnosis.

2.3.1 System Calibration

A spectrum of 4-acetamidophenol (Tylenol) is used for Raman shift calibration. As can be seen in Figure 2.5, the Tylenol spectrum has several sharp peaks in the 200-1700 cm^{-1} region, which can be identified easily. The pixel positions along the horizontal axis of the CCD are mapped to eighteen different peaks in the Tylenol spectrum with a fifth order polynomial fit for the wavenumber calibration.

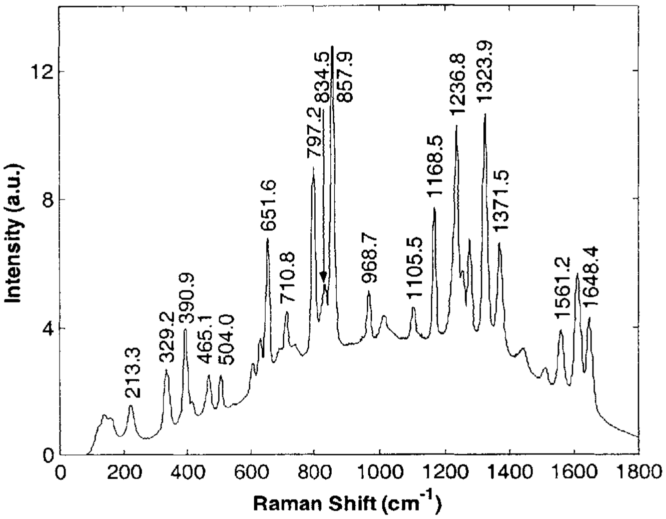


Figure 2.5: Tylenol spectrum after background removal.

Intensity calibration and accurate control of the laser power are necessary for ensuring clinically acceptable excitation power levels. Although in principle one of the known peaks of the Tylenol spectrum can be used to calibrate the intensity of the system, it is difficult to do so because of the sterilization constraints in a clinical setting. Therefore a block of Teflon, which can be easily sterilized, is used to calibrate the intensity and laser power (cf. Section 2.2).

The Raman light from tissue is severely distorted due to several different artifacts introduced by the system. First, there is a spectral distortion introduced by the CCD, since its quantum efficiency varies as a function of wavelength. The notch filter and the grating in the spectrograph are additional sources of spectral distortion. Finally, transmission through the filters in the probe tip is not constant across all wavelengths in the region of interest. In order to correct for these distortions, we divide the raw data from tissue by the normalized spectrum of a white light source which is diffusely scattered by a reflectance standard ($BaSO_4$) (Figure 2.6).

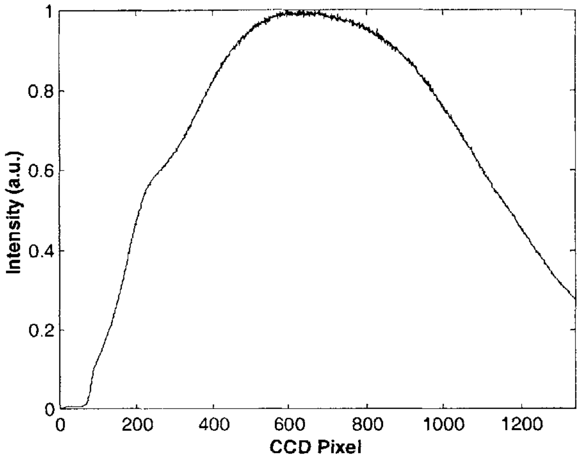


Figure 2.6: White light spectrum used to correct for wavelength dependent system responses.

The remaining fiber background that was not eliminated by the probe filters must also be removed from the raw tissue data. More than 95% the Rayleigh scattered excitation light is filtered out by the collection filters in the tip of the optical fiber Raman probe. The remaining excitation light, however, enters the collection fibers and introduces a large fiber (quartz) background to the Raman light from tissue. The quartz background from the excitation fiber also undergoes Rayleigh scattering and enters the collection fibers, adding to the fiber background generated by the excitation light. The fiber background is characterized by collecting the laser light reflected from a block of aluminum. This spectrum is then subtracted from the white light corrected tissue data to remove the unfiltered probe background. A sample aluminum spectrum with normalized intensity is shown in Figure 2.7.

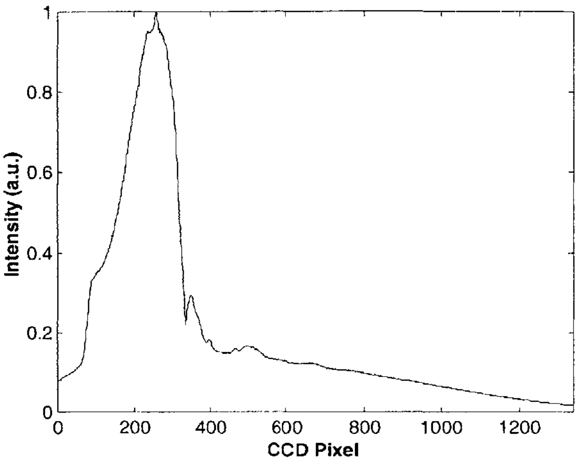


Figure 2.7: Aluminum spectrum for characterizing the probe background.

Finally, we need to correct for the effects of tissue fluorescence. Since the fluorescence features are generally much broader than the Raman features, we can isolate the tissue Raman signal by subtracting a fifth order polynomial from the raw data [4].

2.3.2 Data Acquisition and Analysis

Figure 2.8 shows the flow of data between different modules of the real-time clinical Raman system. As shown on the left-hand side of Figure 2.8, LabVIEW controls the laser, laser shutter, and CCD detector for data acquisition (See Appendix A). LabVIEW drivers to control the CCD are written by R^3-Software (Lawrenceville, NJ). The right-hand side of Figure 2.8 depicts LabVIEW interfacing with Matlab subroutines (See Appendix B) for real-time data analysis and atherosclerosis diagnosis.

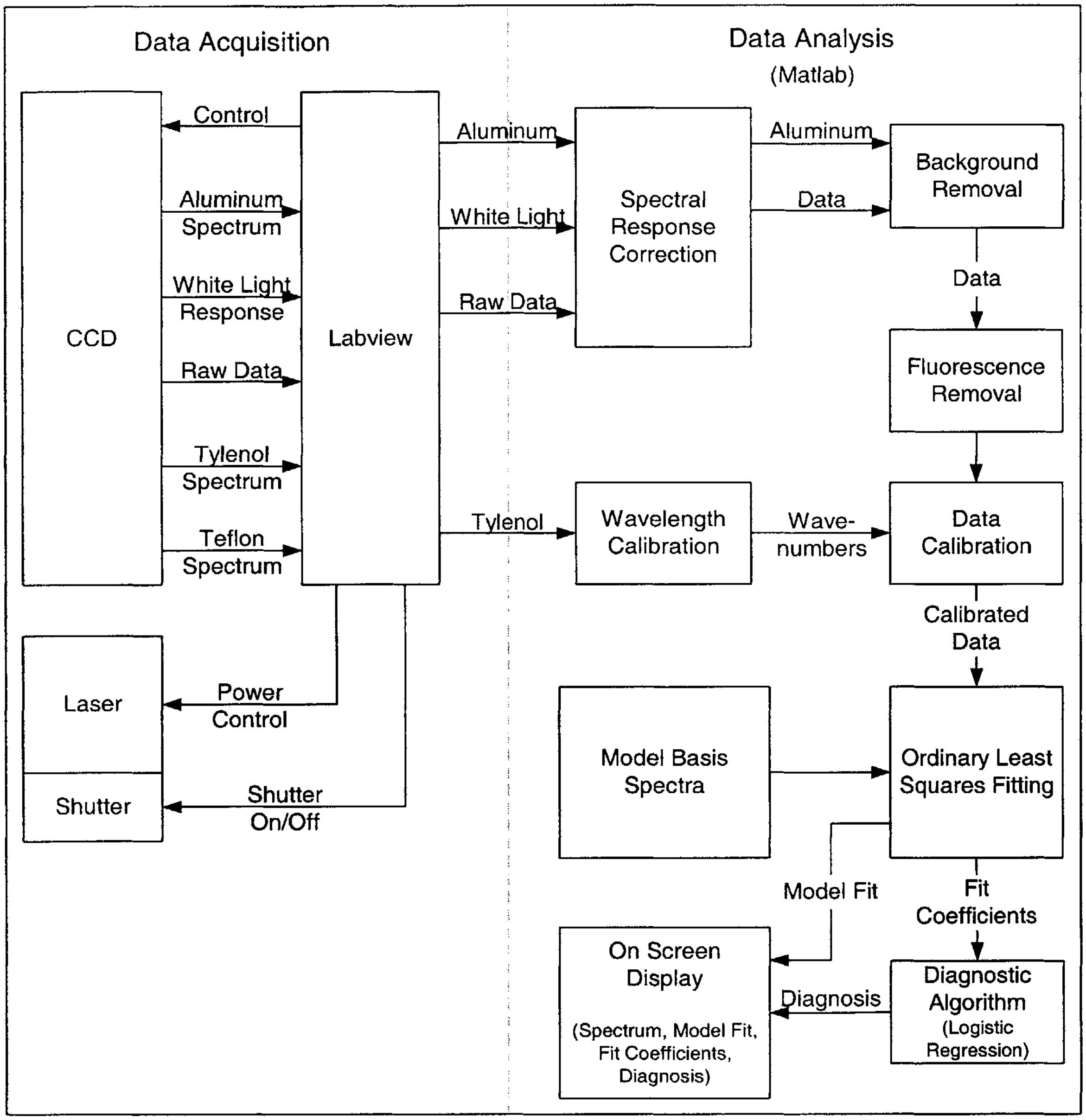
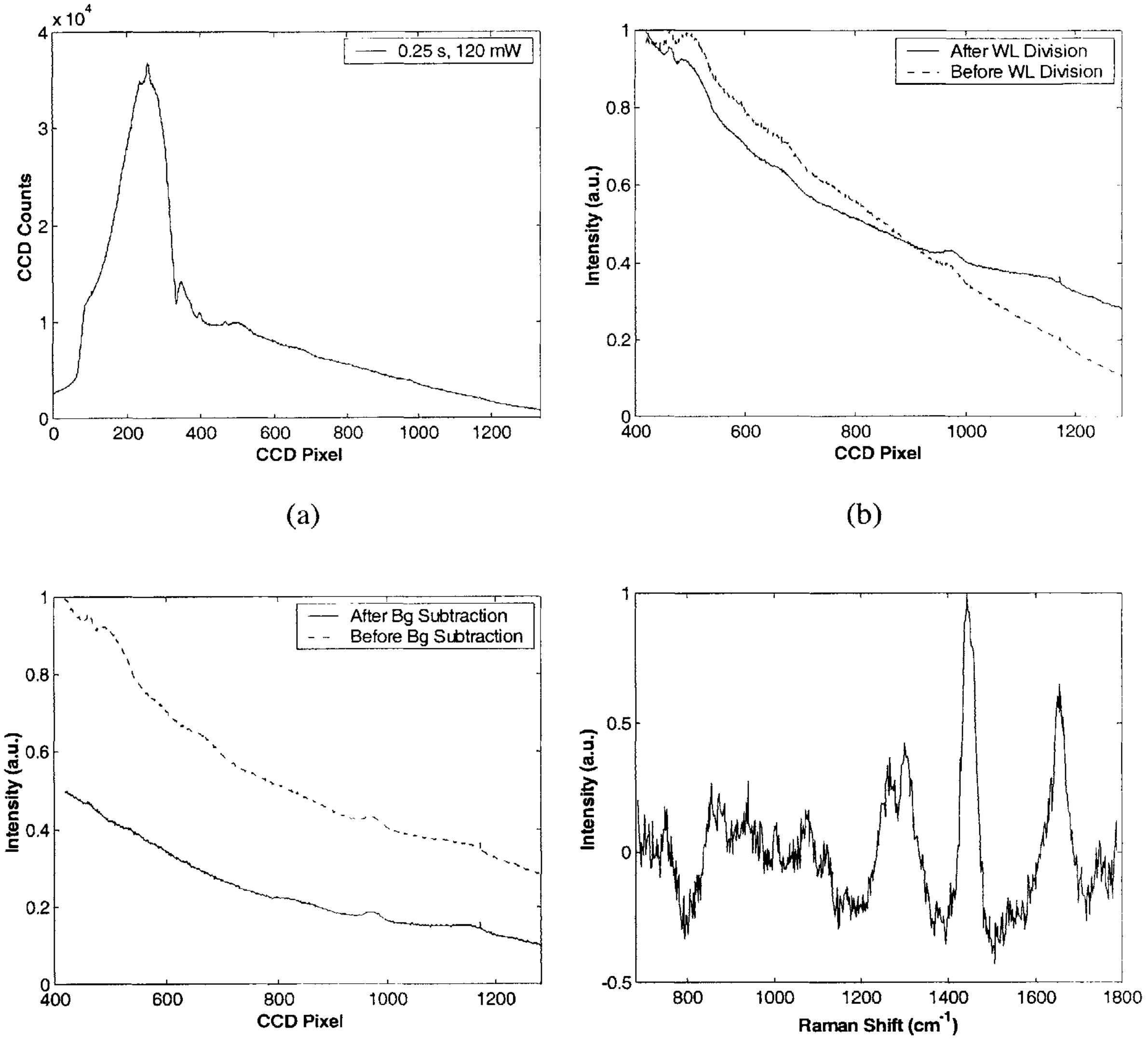


Figure 2.8: Data flow in the real-time clinical Raman system.

All spectra for system calibration are acquired prior to the clinical procedure. The proximal end of the excitation fiber is aligned with the microscope objective to maximize the laser power coupled into the probe. The linear array of collection fibers is aligned to focus the collected light onto the entrance of the spectrograph slit. After aligning the proximal ends of the optical fiber probe, LabVIEW is used to configure the acquisition time, region of interest, and binning parameters for the CCD. The Tylenol, white light, aluminum, and Teflon spectra are then acquired using LabVIEW and loaded into the system at some time prior to the *in vivo* procedure. Since the power of the white light source is arbitrary, the intensity of the white light spectrum is normalized before it is used for spectral response correction. The aluminum spectrum is then corrected for spectral response by white light division and normalized. Raman shift calibration is also performed at this time with the Tylenol spectrum. Finally, the model basis spectra are loaded into the system. The probe and the Teflon block are then sterilized for the *in vivo* procedure.

Another spectrum of the sterilized Teflon block is taken at the beginning of the procedure to set the laser power (cf. Section 2.2). The system is then ready to take spectra from tissue and analyze them in real-time. As we can see in Figure 2.9(a), the raw tissue spectrum from a normal femoral artery consists of small Raman features, which are barely discernable above the large fiber and fluorescence background.

We first perform a spectral response correction on the tissue data by dividing it with the white light spectrum. The tissue spectrum is truncated at this time to pixels corresponding to the region of interest (686-1788 cm^{-1}). Figure 2.9(b) shows the raw tissue data after its intensity is normalized and corrected for the spectral response. As we can see, the artifact introduced by low quantum efficiency of the CCD at longer wavenumbers is now partially corrected. As shown in Figure 2.8, the next step is to remove the fiber background from the tissue data by subtracting the aluminum spectrum. The relationship between the intensity of the aluminum spectrum and the tissue spectrum is dependent upon the tissue type, which is unknown *a priori*. Therefore, we subtract the same aluminum spectrum scaled with twenty different intensities to determine the optimal ratio for background removal. The spectrum that results in the lowest standard deviation of the residual between the data and the model fit is used for subsequent analysis. Figure 2.9(c) shows the tissue Raman spectrum after it is corrected for the effects of the fiber background. Finally, Figure 2.9(d) shows the Raman spectrum of tissue from a normal femoral artery after the effects of tissue fluorescence have been removed. The spectrum was collected in four accumulations of 0.25 s for a total collection time of 1 s with 120 mW excitation laser power. The intensity of the spectrum is normalized after fluorescence removal. Note that the pixels along the horizontal axis have been mapped to wavenumbers using the Raman shift calibration provided by the Tylenol spectrum. Removal of fluorescence reveals the fine detail contained within the Raman spectrum that cannot be seen in Figure 2.9(c).



(a) (b)

(c) (d)

Figure 2.9: Raman spectrum of a non-atherosclerotic femoral artery (a) unprocessed, (b) after spectral response correction, (c) after background subtraction, and (d) after fluorescence removal.

2.3.3 Diagnosis

The last step needed to extract clinically relevant parameters from the tissue Raman spectrum is carried out by fitting the data with an established Raman spectral model.

Basis spectra from the morphological model outlined by Buschman, *et al.* are used to fit the *in vivo* tissue Raman data [5]. The eight different basis spectra in the model characterize the following morphological structures: calcified minerals, β-carotene, cholesterol crystals, foam cells/necrotic core, adventitial adipocytes, smooth muscle cells, collagen fibers, and elastic lamina. In addition, we use basis spectra for epoxy, sapphire, hemoglobin, and water to account for components of the optical fiber probe tip as well as the *in vivo* environment.

Several linear regression approaches can be used to fit our measured spectrum with the model basis spectra, as outlined by Martens and Naes [6]. In the partial least squares regression method, the measured spectrum **y** is projected onto a few spectra

$$\hat{\mathbf{T}} = (\hat{\mathbf{t}}_1, \hat{\mathbf{t}}_2, \ldots, \hat{\mathbf{t}}_k)$$

instead of all of the basis spectra in the matrix

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i)$$

where $i>k$. By projecting the common features of the i model basis spectra in **X** onto the k spectra in $\hat{\mathbf{T}}$, the model in **X** is compressed into a more stable and easily interpretable model. However, since the morphological model basis spectra are a good characterization of the measured spectra, the ordinary least squares (OLS) method is utilized to fit the data. The problem is formulated as

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \tag{2.1}$$

where **y** is the measured spectrum, **X** is a matrix with one column for each of the i basis spectra, and **b** is a matrix of fit contributions we are trying to predict. The column vector

$\mathbf{e}$ accounts for the error due to measurement noise in the data as well as modeling errors. The goal then is to estimate $\hat{\mathbf{b}}$ such that the squared error

$$\|\mathbf{e}\|^2 = \|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2 \tag{2.2}$$

is minimized. This is accomplished by minimizing the scalar product

$$\mathbf{e}^{\mathrm{T}}\mathbf{e} = (\mathbf{y} - \mathbf{X}\mathbf{b})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\mathbf{b}). \tag{2.3}$$

Solving for the estimated $\hat{\mathbf{b}}$, we find

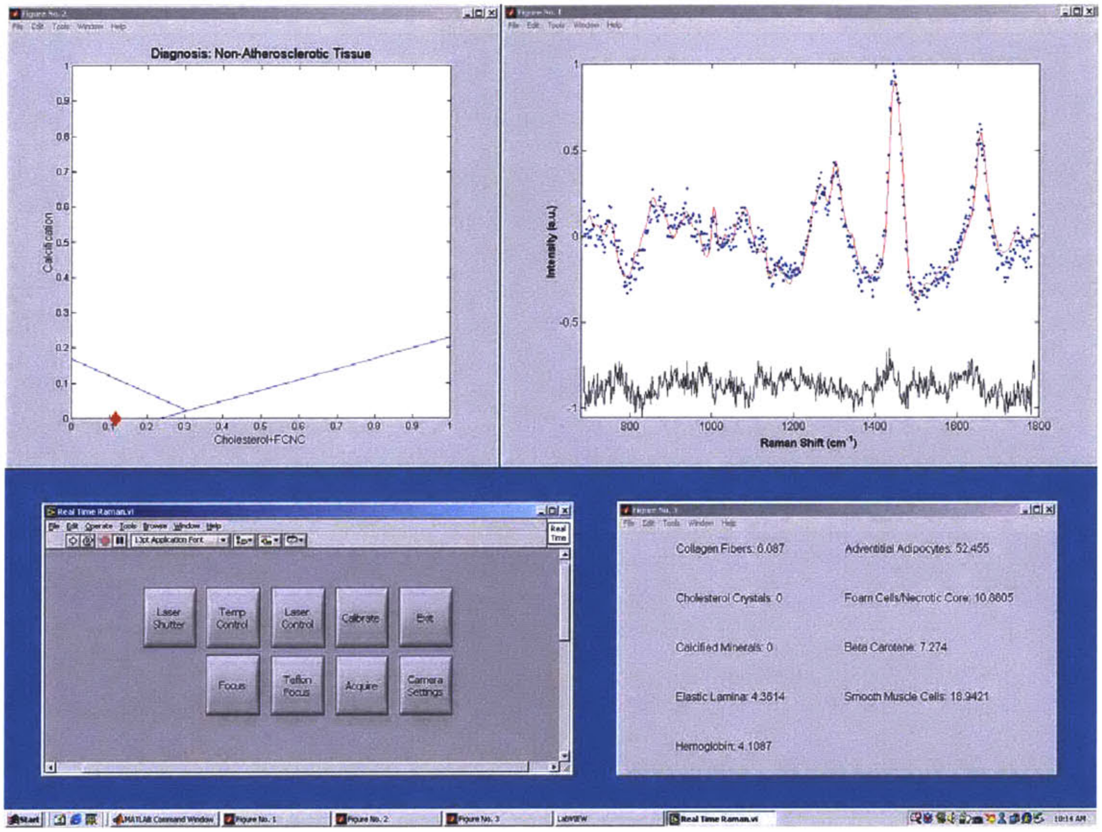
$$\hat{\mathbf{b}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}. \tag{2.4}$$

The model fit to the measured spectrum is then calculated as

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{b}} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}. \tag{2.5}$$

The resultant fractional fit coefficients, $\hat{\mathbf{b}}$, are used to make a diagnosis based on the *in vitro* diagnostic algorithm developed by Buschman, *et al.* [7].

Finally, the tissue Raman spectrum, model fit, residual between the data and model fit, fractional fit coefficients, and diagnosis are plotted on the computer screen (Figure 2.10). Depending on the speed of the computer, the entire process of data acquisition, analysis, and diagnosis takes 2-4 s. The timing breakdown for each step of the process is shown in Table 2.1. The reported times are for a 1.8 GHz Pentium 4 processor with 512 MB of RAM.



Collagen Fibers: 6.087

Adventitial Adipocytes: 52.455

Cholesterol Crystals: 0

Foam Cells/Necrotic Core: 10.8805

Calcified Minerals: 0

Beta Carotene: 7.274

Elastic Lamina: 4.3614

Smooth Muscle Cells: 18.9421

Hemoglobin: 4.1087

Figure 2.10: Real-time display of tissue Raman spectrum, model fit, residual, fit contributions, and diagnosis of a normal tissue sample in the femoral artery. The collection time for the spectrum was 1 s with 120 mW excitation power. The total time for data acquisition and analysis was under 3 s.

Process	Time (ms)
Open Laser Shutter	0.700
Collection Time	1000
Load acquired spectra	139
Spectral response correction	1
Background removal	10
Fluorescence removal	771
Ordinary least squares fitting	491
Diagnostic algorithm and on-screen display	411

Table 2.1: Timing breakdown for data acquisition, analysis, and disease diagnosis in the real-time system.

In conclusion, the necessary software and hardware modifications have been made to the previous *in vivo* system to allow for a real-time disease diagnosis in a clinical environment. The modular design of the real-time system grants us the flexibility to use models and diagnostic algorithms specific to a disease. In the future, the current diagnostic algorithms could be expanded to provide information about various stages and progression of disease as well as the ideal course of treatment.

Limitations to reducing the total time for data acquisition, analysis, and disease diagnosis are mainly imposed by two factors. First, the collection efficiency and throughput of the optical fiber Raman probe determine the time needed for collecting *in vivo* tissue spectra. Significant improvements in the design of the probe filters and collection optics may lead to higher throughput and collection efficiency, which would allow us to acquire *in vivo* tissue spectra in less than 1 s.

Reduction in time for data analysis is mainly limited by processor speed. As can be seen in Table 2.1, the most time consuming steps for data analysis is OLS fitting of the data. The relationship between the intensity of the aluminum spectrum and the tissue spectrum is unknown *a priori*. Therefore, the same aluminum spectrum scaled with twenty different intensities is needed to determine the optimal ratio for background removal. As a result, OLS fitting has to be carried out on twenty different spectra for each tissue site. Development of algorithms that rely on knowledge of *a priori* probabilities of probing a certain tissue type might allow us to reduce the time spent on analyzing the data.

1. Brennan JF, Wang Y, Dasari RR, and Feld MS, *Near-Infrared Raman Spectrometer Systems for Human Tissue Studies.* Appl Spectrosc, 1997. **51**(2): p. 201-208.
2. Shim MG and Wilson BC, *Development of an in vivo Raman spectroscopic system for diagnostic applications.* Journal of Raman Spectroscopy, 1997. **28**(2-3): p. 131-142.
3. Hanlon EB, Manoharan R, Koo T-W, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, and Feld MS, *Prospects for In Vivo Raman Spectroscopy.* Physics in Medicine and Biology, 2000. **45**(2): p. R1-R59.
4. Baraga JJ, Feld MS, and Rava RP, *Rapid Near-Infrared Raman Spectroscopy of Human Tissue with a Spectrograph and CCD Detector.* Appl Spectrosc, 1992. **46**(2): p. 187-190.
5. Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, and Feld MS, *Raman microspectroscopy of human coronary atherosclerosis: Biochemical assessment of cellular and extracellular morphologic structures in situ.* Cardiovascular Pathology, 2001. **10**(2): p. 69-82.
6. Martens H and Naes T, *Multivariate Calibration.* 1989, New York: John Wiley & Sons.
7. Buschman HP, Motz JT, Deinum G, Romer TJ, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, and Feld MS, *Diagnosis of human coronary atherosclerosis by morphology-based Raman spectroscopy.* Cardiovascular Pathology, 2001. **10**(2): p. 59-68.

Chapter 3

In Vivo Experiments and Results

We now present the results of a clinical study, which was performed to demonstrate the potential of Raman spectroscopy for *in vivo* atherosclerosis diagnosis in real-time. The study was approved by the MetroWest Medical Center Institutional Review Board, and the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects. The goal of the study was to see whether the *in vitro* models and diagnostic algorithms developed by Buschman, *et al.* could be extended to characterize *in vivo* atherosclerotic tissue in real-time [1, 2]. The results for nineteen of the twenty cases presented in this section were not available to clinicians in real-time when the experiments were performed. However, the real-time clinical system described in Chapter 2 is now capable of providing immediate diagnostic information to clinicians and was utilized in the final case.

3.1 Methods

Raman spectra were acquired from the peripheral vessels of patients undergoing carotid endarterectomy or femoral bypass surgery. A total of 74 different sites on 20 different patients were examined. 23 biopsies were obtained for 34 of those sites and evaluated by a blinded pathologist for confirmation.

The clinical system described in Section 2.1 was used to acquire the Raman spectra. The Teflon block used for excitation laser power calibration and the front-viewing optical fiber Raman probe were sterilized prior to the procedure (cf. Section 2.2). Lights in the operating room were turned off at the surgeon's command. Spectra were acquired from the proximal anastomosis sight during femoral bypass surgery and from the posterior arterial wall during carotid endarterectomy. The tip of the front-viewing optical fiber probe was held normal to the tissue surface as spectra were acquired at 0.25 s intervals for a total integration time of 5 s with ~100 mW excitation laser power. A biopsy was obtained and the excitation spot was marked with a stitch for histologic examination at a later time.

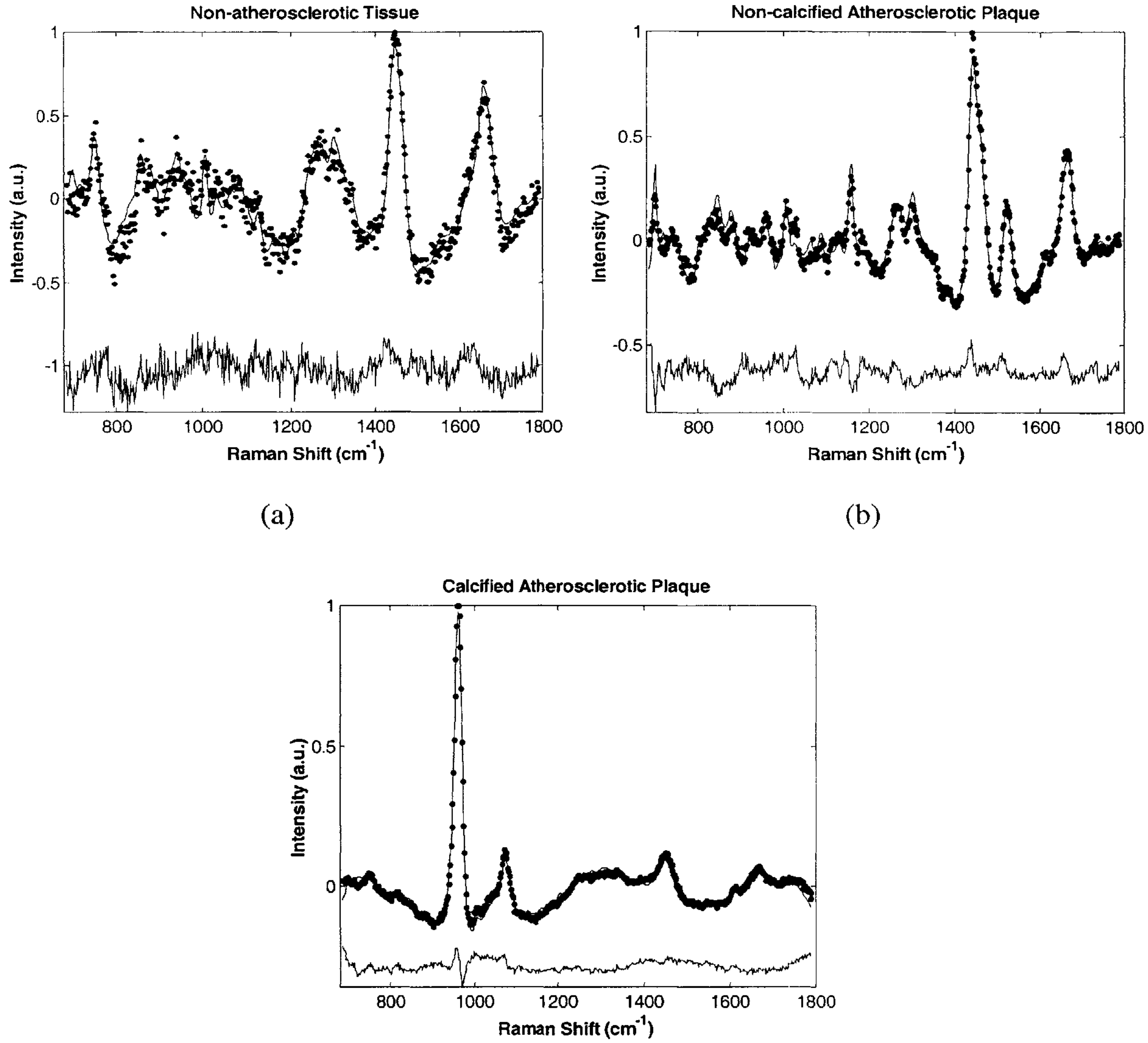
System calibration and data analysis were performed as described in Section 2.3. In short, the spectrum of a sterilized Teflon block was used for calibrating the intensity. Eighteen known peaks in the Tylenol spectrum were used to calibrate the Raman shifts. The spectral response correction was performed by dividing the tissue spectra with a spectrum of white light scattered from a reflectance standard. A spectrum of aluminum was used to characterize the probe background and was subtracted from the tissue spectra. Finally, a fifth order polynomial was subtracted from the measured spectra to remove the slowly varying fluorescence background.

3.2 Results

3.2.1 *In Vivo* Raman Spectra

Tissue Raman spectra were characterized through ordinary least squares (OLS) fitting with twelve basis spectra, which included a spectrum of epoxy, sapphire, water, and hemoglobin in addition to the eight morphologic structures used by Buschman *et al.* in their *in vitro* work [2]. The fractional fit contributions of certain morphologic structures were then used to classify each site as non-atherosclerotic tissue, calcified atherosclerotic plaque, or non-calcified atherosclerotic plaque.

Figure 3.1 shows typical Raman spectra (dotted line) of *in vivo* tissue samples from each of the three diagnostic categories. The OLS model fits (solid line) are also shown in Figure 3.1. Residuals (data minus fit) are shown on the same scale, but with an offset. The morphological model was able to account for the different features of the measured spectra for normal and diseased tissue. Excellent agreement between the calculated model fits and the measured spectra demonstrates that the *in vitro* morphologic basis spectra are able to characterize *in vivo* tissue.



(a)

(b)

(c)

Figure 3.1: Typical Raman spectra of femoral artery tissue from each of the three diagnostic classes: (a) non-atherosclerotic tissue, (b) non-calcified atherosclerotic plaque, and (c) calcified atherosclerotic plaque. The dotted line is the measured spectrum and solid line the model fit. The lower line in each graph is the residual (data minus fit).

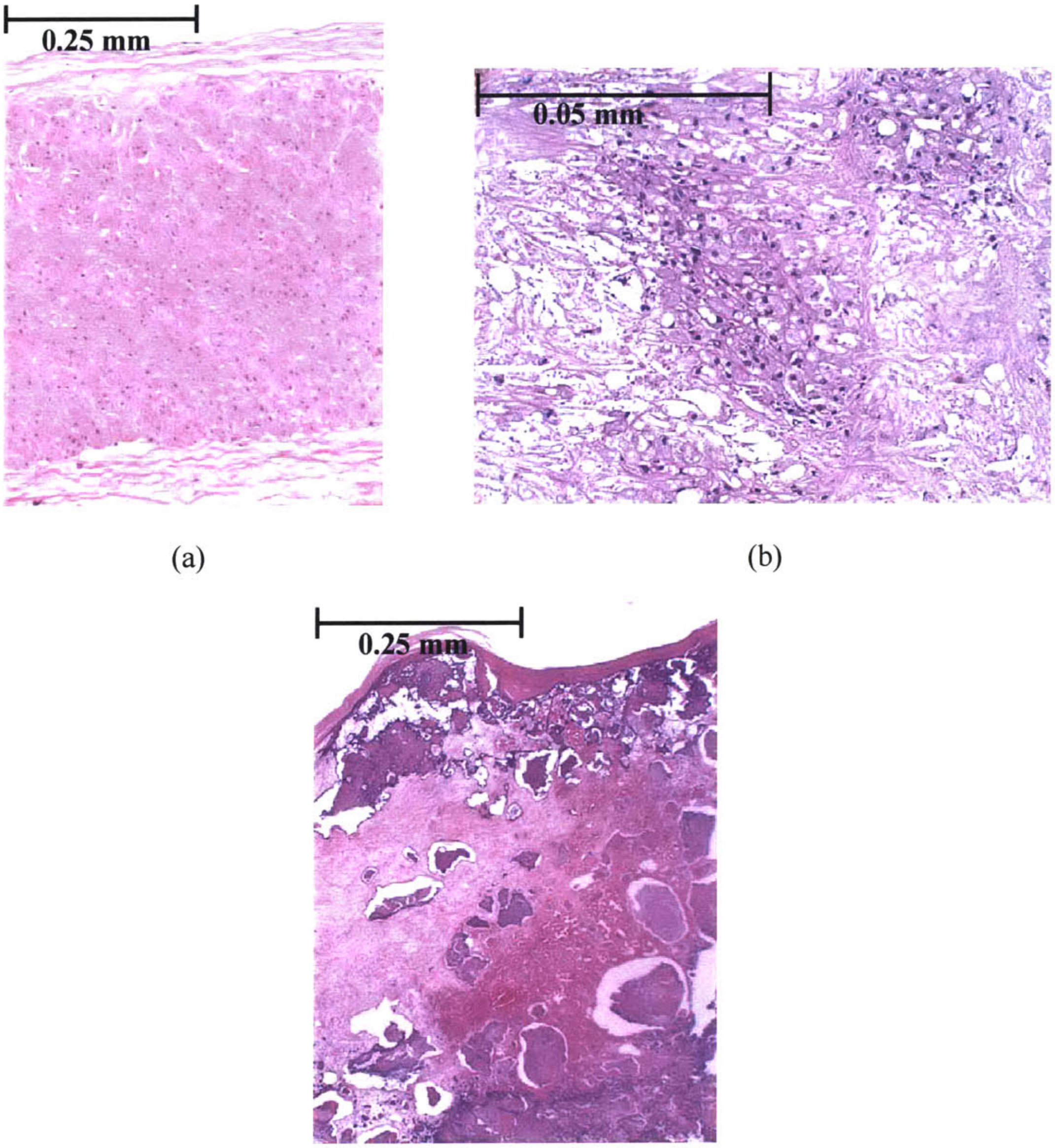
Model Component	**(a)**	**(b)**	**(c)**
Collagen Fibers (CF)	0.14	0	0.01
Cholesterol Crystals (CC)	0	0.28	0.35
Elastic Lamina (EL)	0	0.05	0
Adventitial Adipocytes (AA)	0.32	0.23	0.07
Foam Cells and Necrotic Core (FCNC)	0.19	0.38	0
Smooth Muscle Cells (SMC)	0.35	0.06	0.56
Calcified Minerals (CM)	0	0	0.94
β-Carotene (BC)	0.01	0.38	0

Table 3.1: Relative fit contributions of the eight morphologic structures for the spectra shown in Figure 3.1. (a) Non-atherosclerotic tissue, (b) non-calcified plaque, and (c) calcified plaque.

Table 3.1 shows the relative fit coefficients for each of the three spectra shown in Figure 3.1. The fit coefficients are consistent with the biochemical and morphological changes expected with disease progression. Since the intima is thin in non-atherosclerotic artery, we would expect to see a large spectroscopic contribution from smooth muscle cells as well as the adipose tissue in the adventitial layer. In addition, little or no contribution from cholesterol crystals and calcified minerals was seen. On the other hand, slightly lower contribution from adventitial adipocyte was seen in calcified and non-calcified atherosclerotic plaque due to a thickened intima. Furthermore, the contribution from foam cells/necrotic core and cholesterol crystals was much higher due to the presence of plaques. Finally, the Raman spectrum of calcified atheroscelrotic plaque was dominated by contribution from calcified minerals.

While the concentration of β-carotene in arterial tissue is low, the results from OLS fitting showed large differences in its fit contribution for different diagnostic classes. A large contribution of β-carotene was found in non-calcified plaques, since it is a lipophilic molecule. Consistent with Buschman, *et al.*, the fit contributions from the seven other morphologic structures were renormalized after removing the contribution from β-carotene [2]. This was done in order to prevent β-carotene from masking the contributions from other morphological structures. Similarly, a large contribution from calcified minerals was seen in calcified atherosclerotic plaques. Therefore, the fit coefficients of the six remaining morphologic structures were renormalized without the contribution from calcified minerals. As a result, the reported fit contributions are only relative and sum to a value greater than one.

The *in vitro* diagnostic algorithm outlined by Buschman, *et al.* was applied to the fit coefficients presented in Table 3.1 for each of the three different tissue sites [2]. Fit contributions of calcified minerals, foam cells/necrotic core, and cholesterol crystals were used by the diagnostic algorithm (cf. Section 1.3.1). Although the presence of β-carotene is an obvious identifier of non-calcified plaque, it was not used by the algorithm since it only provides correlative information about the actual disease progression in tissue. Blind evaluation of tissue biopsies by a pathologist confirmed the diagnoses determined by Raman spectroscopy for the three sample spectra shown before (Figure 3.2).



(a)

(b)

(c)

Figure 3.2: Hematoxylin and eosin stained sectioning used for histologic confirmation of the diagnoses determined by analysis of the spectra shown in Figure 3.1 for (a) non-atherosclerotic tissue (4x), (b) non-calcified plaque (20x), and (c) calcified plaque (4x).

Histologic examination of the 23 biopsy samples showed that 25 out of 34 sites were correctly classified into one of the three diagnostic classes by Raman spectroscopy.

3.2.2 Signal to Noise and Integration Time

The collection time for Raman spectra presented in the previous section was 5 s at ~100 mW. Figure 3.3 shows the Raman spectra of a non-calcified atherosclerotic tissue site at 1 s intervals within the 5 s integration time.

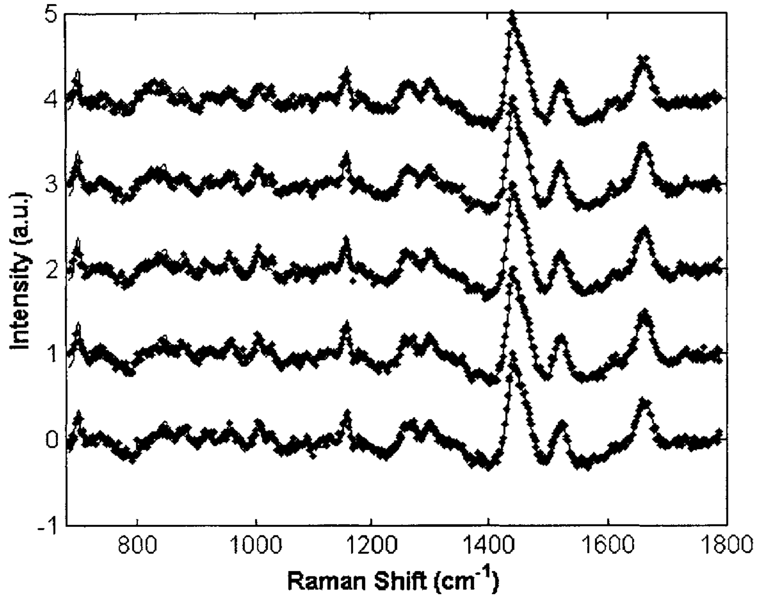
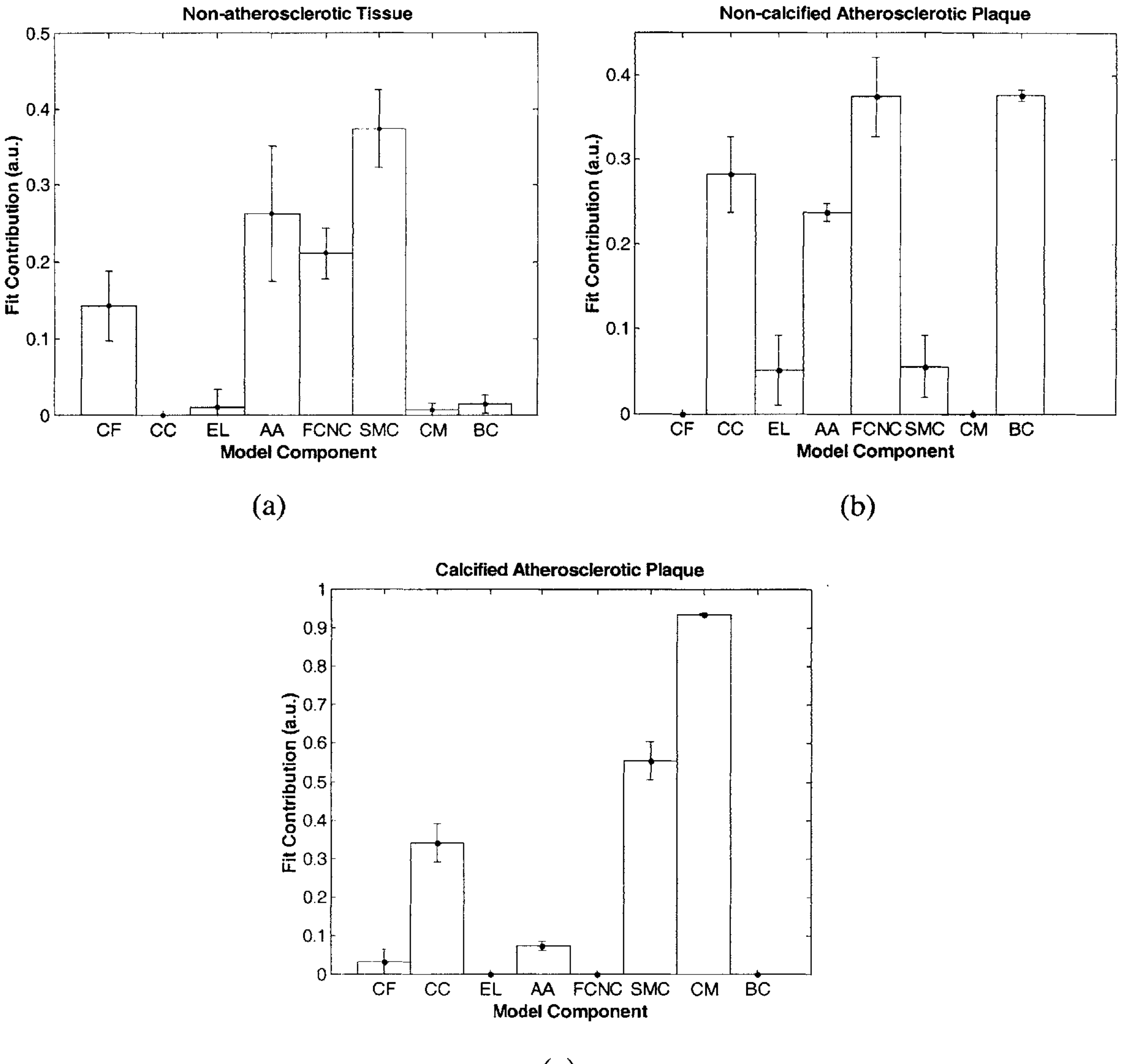


Figure 3.3: Raman spectra of non-calcified atherosclerotic femoral artery tissue at 1 s intervals.

As we can see, there was negligible variation in the spectra from each second to the next. We define the signal to noise ratio for the clinical system as

$$SNR = \frac{1}{\sigma(error)} \tag{3.1}$$

where *error* is the difference between the measured spectrum and the model fit and $\sigma(error)$ is its standard deviation. Although the SNR fell when the total integration time was reduced from 5 s to 1 s, we were able to extract the morphologic fit contributions without losing diagnostic accuracy. Figure 3.4 shows the mean ± standard deviation of the fit contributions from five 1 s spectra taken from each of the three different sites discussed in the previous section. Small error bars in Figure 3.4 indicate that the deviation in fit contributions from second to second was small. Furthermore, excellent correlation was found between the mean of the fit contributions of the five 1 s spectra and the contributions for the entire 5 s collection time in all 34 cases (Table 3.2).



(a)

(b)

(c)

Figure 3.4: The mean ± standard deviation of the fit contributions of five 1 s spectra from a tissue site in each of the three diagnostic classes.

Model Component	**Correlation (*R*)**
CF	0.9958
CC	0.9858
EL	0.7805
AA	0.9983
FCNC	0.9956
SMC	0.9850
CM	0.9996
BC	0.9994

Table 3.2: Correlation between mean of the fit contributions of the five 1 s spectra and the contributions for the entire 5 s collection time in all 34 cases.

The correlation coefficient *R* in Table 3.2 is defined as follows.

$$R(\mathbf{i},\mathbf{j}) = \frac{C(\mathbf{i},\mathbf{j})}{\sqrt{C(\mathbf{i},\mathbf{i})C(\mathbf{j},\mathbf{j})}} \tag{3.2}$$

where each of the 34 elements of vector **i** is the mean of the fit contributions of a particular morphologic structure for the five 1 s spectra. Vector **j** is the fit contribution of a particular morphologic structure for the 5 s spectra in all 34 cases. *C* is the covariance defined as

$$C(x,y) = \mathbf{E}[(x-\mathbf{E}[x])(y-\mathbf{E}[y])]. \tag{3.3}$$

3.3 Discussion

Results of these experiments show that the *in vitro* basis spectra of the morphological model and the diagnostic algorithm can be used to characterize atherosclerotic tissue *in vivo*. By using semi-quantitative biochemical and morphological information, we can diagnose atherosclerotic lesions in arteries without requiring a histologic examination of excised tissue. 25 (74%) of 34 cases were classified correctly into one of three different diagnostic classes, as opposed to 94% accuracy reported for the *in vitro* experiments [2]. The fit contributions varied significantly from sample to sample within the same diagnostic class. Two main reasons can explain the difference in accuracy between *in vivo* and *in vitro* experiments. First, 34 cases is not a large enough sample size to deduce a statistical estimate of the accuracy of the presented technique. Second, more importantly, in some cases there was poor registration between the spot from which the Raman spectra were taken and the spot examined by the pathologist. The rigid time constraints during surgery prevented the surgeon from properly marking the spot of data collection in some cases. Registration techniques need to be improved for future studies in order to definitively establish the accuracy of using Raman spectroscopy for *in vivo* diagnosis of atherosclerosis.

Nevertheless, the results indicate that real-time diagnosis of atherosclerosis is now possible. The analysis of integration time shows that the 1 s Raman spectra look almost identical to those collected over 5 s. Excellent signal-to-noise ratio can be achieved with 1 s integration time at an acceptable excitation laser power (~100 mW) for a clinical setting. There is an excellent correlation between the fit coefficients for tissue spectra collected over 5 s and over 1 s. The ability to acquire tissue spectra and extract fit contributions within 3 s allows for a real-time diagnosis of atherosclerosis.

The information provided by Raman spectra was used to classify tissue into one of three diagnostic classes. However, the detailed biochemical and morphological information can be used in the future to provide information about the stage of disease progression as well as the ideal course of treatment. By improving the current model to detect more subtle morphologic differences, a finer diagnostic categorization can be achieved. For example, by increasing our sensitivity for foam cells and necrotic core, atherosclerotic plaques (foam cells only) can be distinguished from atheromatous plaques (foam cells and necrotic core). Furthermore, the presence of foam cells, necrotic core, and cholesterol crystals are significant markers of plaque instability. By making the necessary modifications to the present model, we can use the semi-quantitative information provided by Raman spectroscopy to distinguish between stable and unstable plaques in real-time.

1. Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, and Feld MS, *Raman microspectroscopy of human coronary atherosclerosis: Biochemical assessment of cellular and extracellular morphologic structures in situ.* Cardiovascular Pathology, 2001. **10**(2): p. 69-82.
2. Buschman HP, Motz JT, Deinum G, Romer TJ, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, and Feld MS, *Diagnosis of human coronary atherosclerosis by morphology-based Raman spectroscopy.* Cardiovascular Pathology, 2001. **10**(2): p. 59-68.

Chapter 4

Conclusion

4.1 Accomplishments

A Raman spectroscopy system capable of providing real-time disease diagnosis in a clinical setting has been designed, developed, and tested. The implemented software package utilizes LabVIEW to interface with various components in the system. LabVIEW control of system components such as the laser and laser shutter allows data acquisition in accordance with clinical safety guidelines. In addition, Labview makes use of the computational power and versatility of Matlab for performing real-time data analysis while offering a customizable and user-friendly interface. The modular design of the system allows easy incorporation of models for any disease. Furthermore, it provides easily adaptable data analysis routines as well as diagnostic algorithms.

The system was used successfully to acquire *in vivo* tissue Raman spectra from human femoral and carotid artery. The results of the experiments showed that *in vitro* model basis spectra and diagnostic algorithm developed by Buschman, *et al.* can be used to provide chemical and morphological information about atherosclerotic tissue *in vivo* [1]. The development of a high-throughput optical fiber Raman probe with filters in the probe tip has made it possible to obtain spectra with excellent signal-to-noise-ratio in 1 s. Therefore, the real-time system can acquire data, perform data analysis, and present clinically relevant parameters within 2-4 s in a clinical setting. The development of a real-time system has also made it possible to scan large tissue areas for molecular markers of disease.

4.2 Future Directions

4.2.1 Side-Viewing Raman Probes

An optical fiber probe with a front-viewing collection and delivery geometry has already been designed and constructed for *in vivo* Raman spectroscopy and was used for the experiments presented in this thesis [2, 3]. However, the ability to obtain intravascular Raman spectra percutaneously is limited due to the inadequacy of the existing delivery and collection geometry. The current challenge is to develop an optical fiber probe that allows transluminal access to the coronary arteries while permitting the same efficiency in transmission of Raman photons and filtering of fiber background as the front-viewing probe.

There are several design criteria that need to be addressed for designing a side-viewing optical fiber probe. First of all, the probe diameter needs to be extremely small, ~1.5 mm in diameter. This means that the number of collection fibers might need to be reduced compared to the 15 used by the front-viewing design, resulting in lower collection efficiency. Preliminary simulations show that a probe with 9 collection fibers and lower, but sufficient, collection efficiency can be designed. In addition to having a small diameter, the probe needs to be flexible so that it can be inserted percutaneously into the vasculature through the femoral artery and guided to the coronary arteries. The collection and delivery optics also need to be designed such that a circumferential view of the arteries is provided. A lateral pull back with a circumferential side-viewing probe will allow us to generate a Raman image map of the entire arterial surface. Finally, balloon or saline flush technology needs to be combined with the Raman probe to minimize the effects of blood on the Raman signal.

The development of a side-viewing probe that meets the design criteria outlined above will allow the use of Raman spectroscopy as a minimally invasive tool for assessing coronary heart disease state in real-time.

4.2.2 *In Vivo* Studies

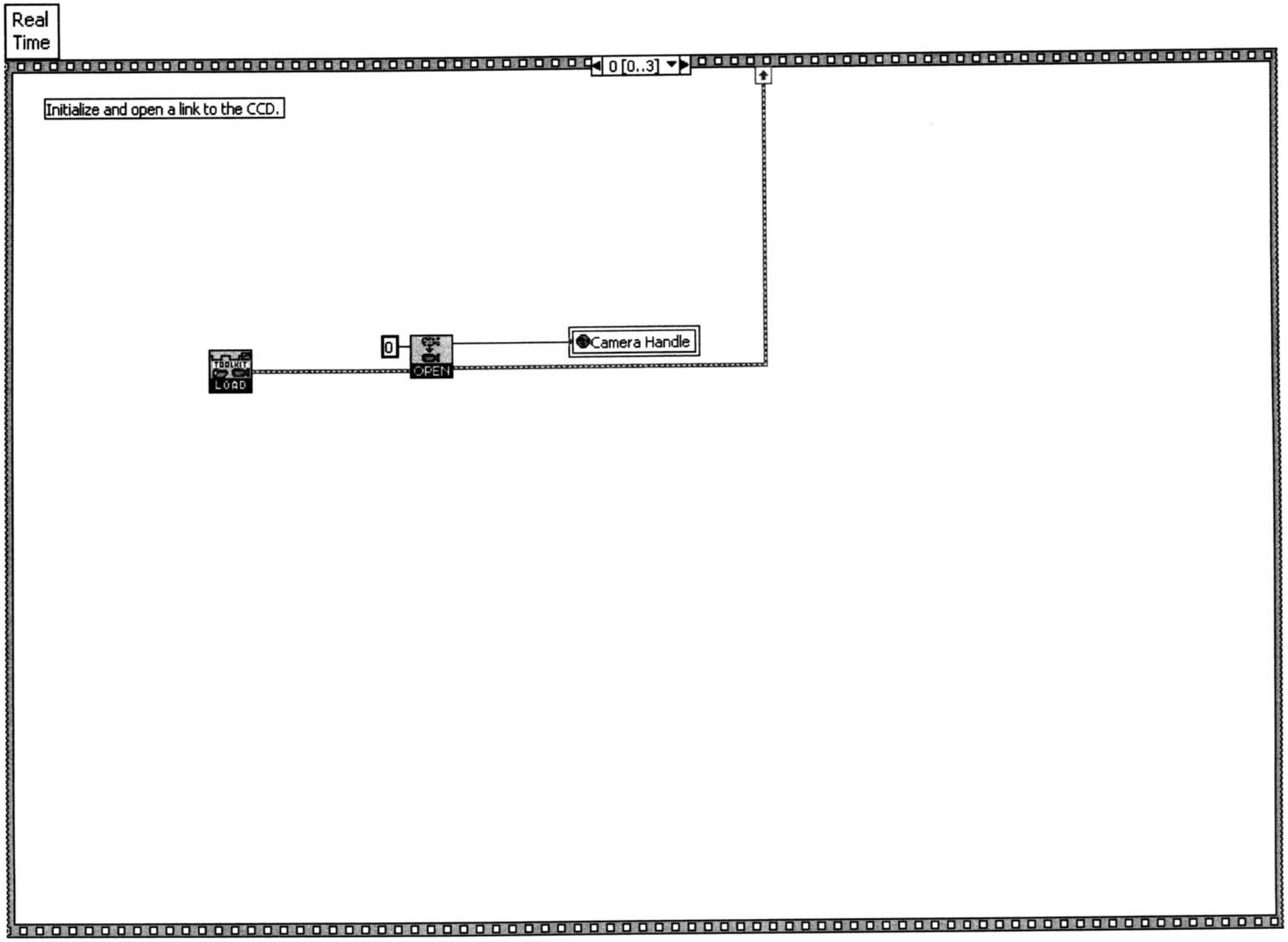
The work presented in thesis provides proof that Raman spectroscopy can be used for real-time diagnoses of disease in a clinical setting. However, extensive *in vivo* experiments need to be performed to evaluate its accuracy for diagnosing various diseases. Furthermore, diagnostic algorithms that allow us to distinguish stable from unstable plaques need to be developed in order to fully utilize the potential of Raman spectroscopy for disease diagnosis.

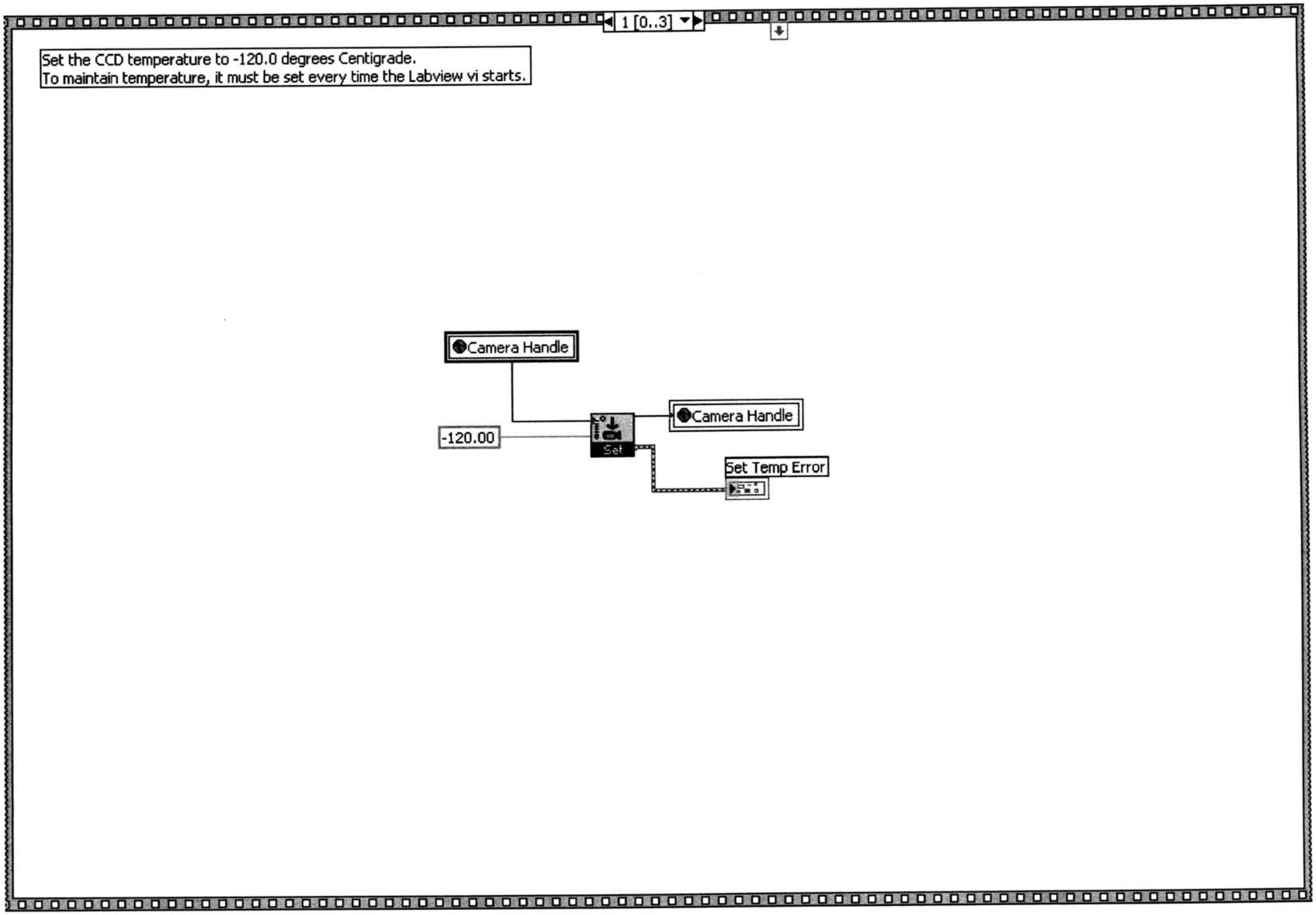
1. Buschman HP, Motz JT, Deinum G, Romer TJ, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, and Feld MS, *Diagnosis of human coronary atherosclerosis by morphology-based Raman spectroscopy.* Cardiovascular Pathology, 2001. **10**(2): p. 59-68.
2. Motz JT, Hunter M, Galindo LH, Gardecki JA, Kramer JR, Dasari RR, and Feld MS, *Optical Fiber Raman Probe for Biomedical Spectroscopy.* In Preparation.
3. Motz JT, *Development of In Vivo Raman Spectroscopy of Atherosclerosis*, in *Health Sciences and Technology.* 2003, Massachusetts Institute of Technology: Cambridge, Massachusetts.

Appendix A

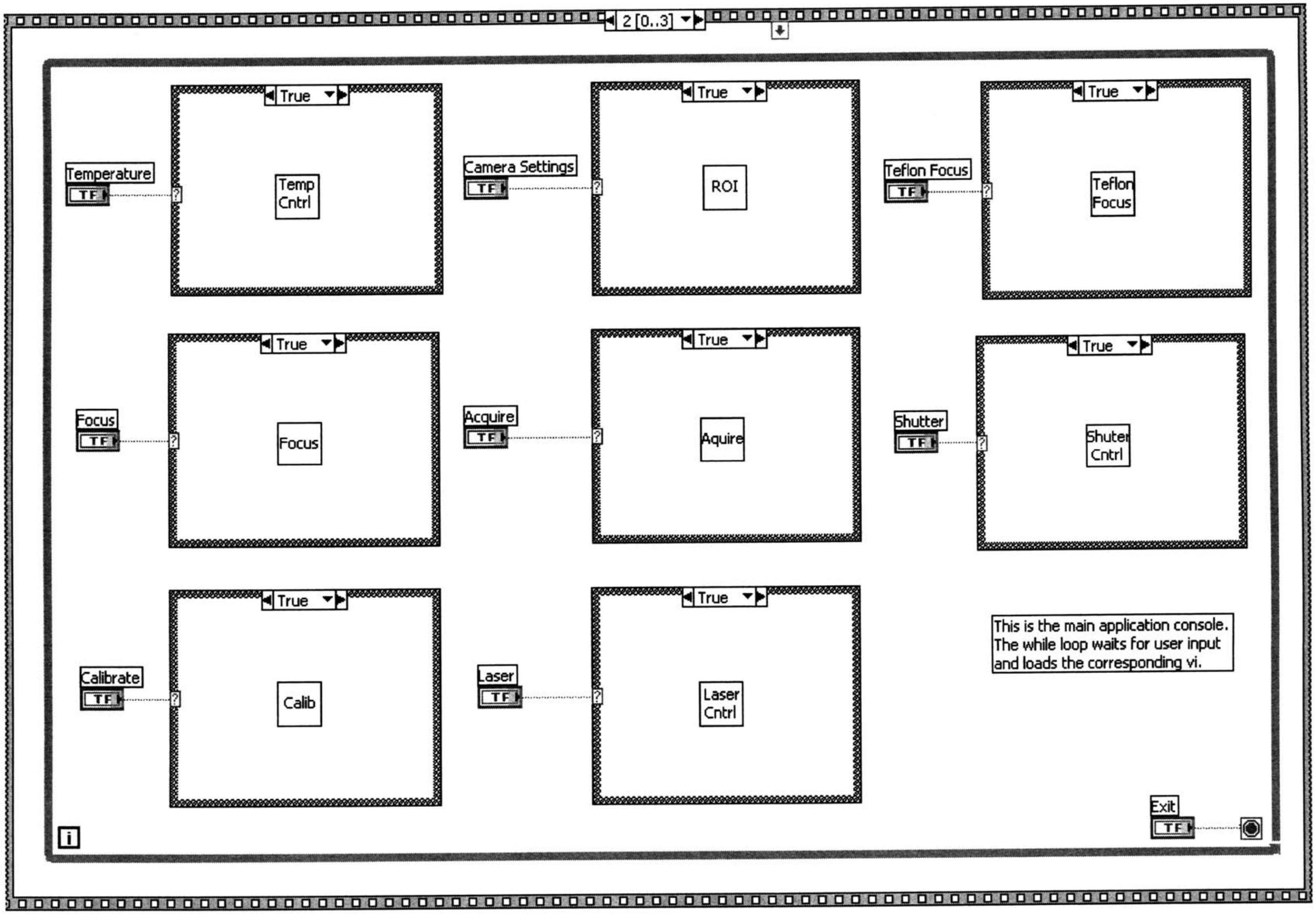
LabVIEW Implementation

Real Time Raman.vi

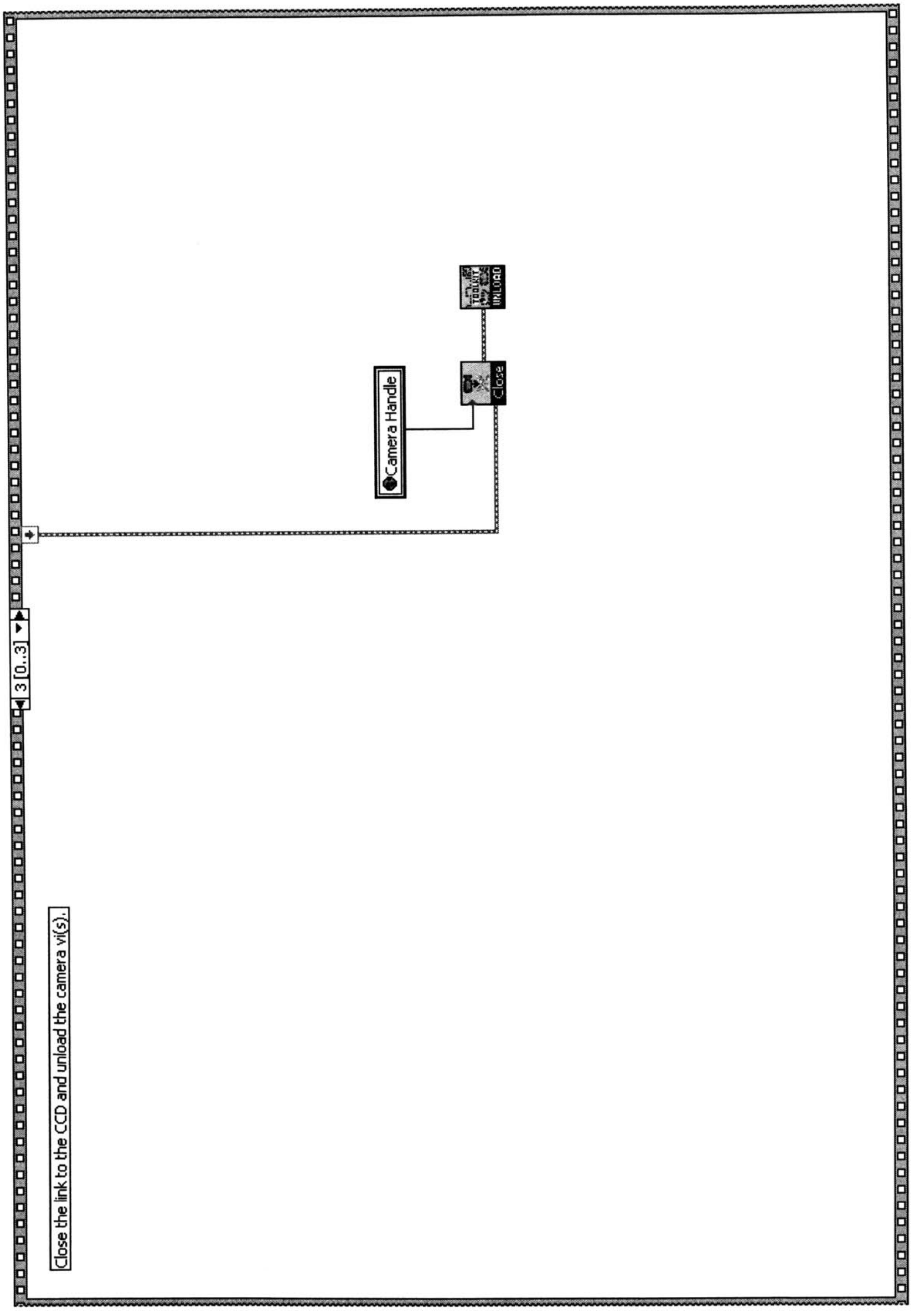




Set the CCD temperature to -120.0 degrees Centigrade.
To maintain temperature, it must be set every time the Labview vi starts.



This is the main application console. The while loop waits for user input and loads the corresponding vi.

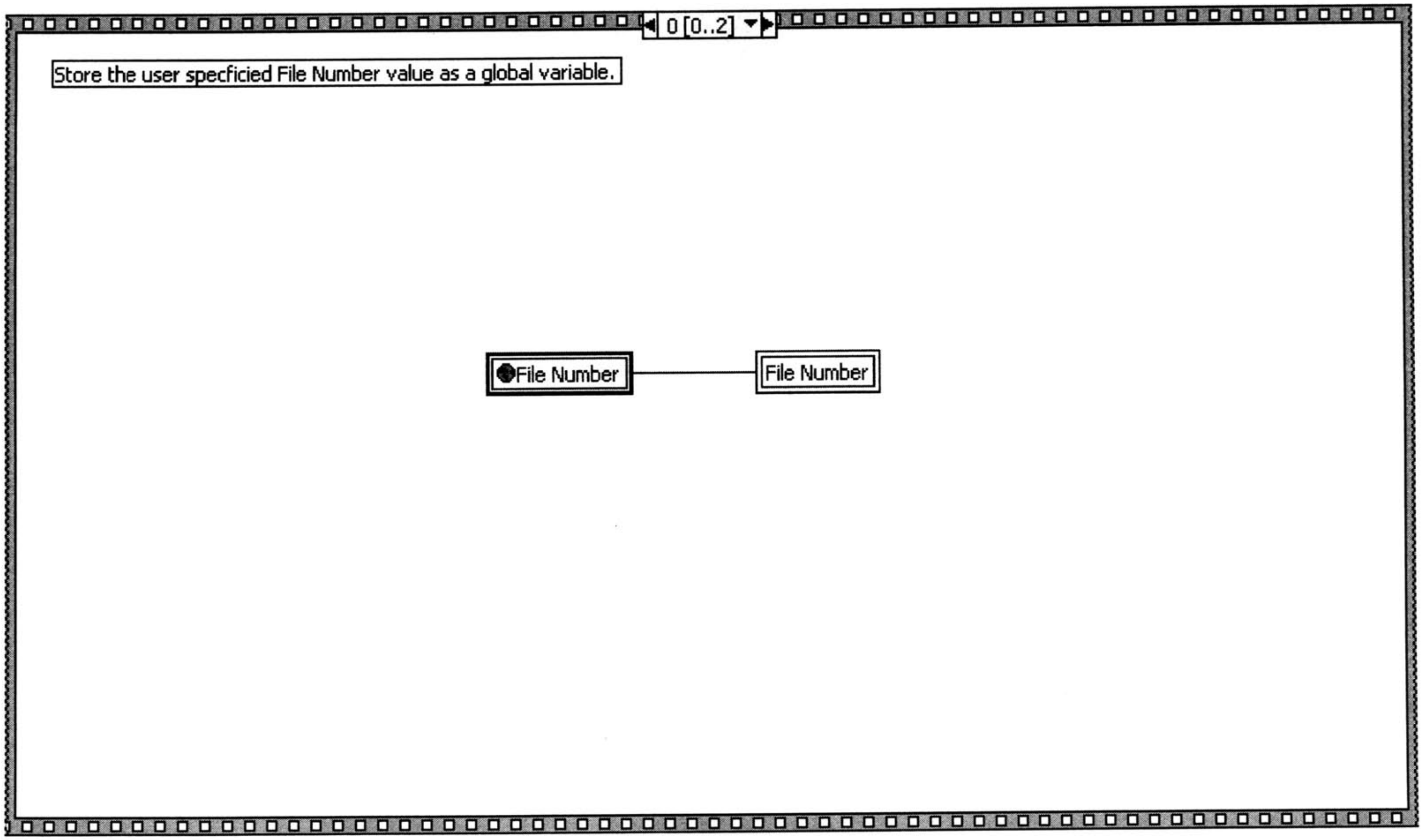


Close the link to the CCD and unload the camera vi(s).

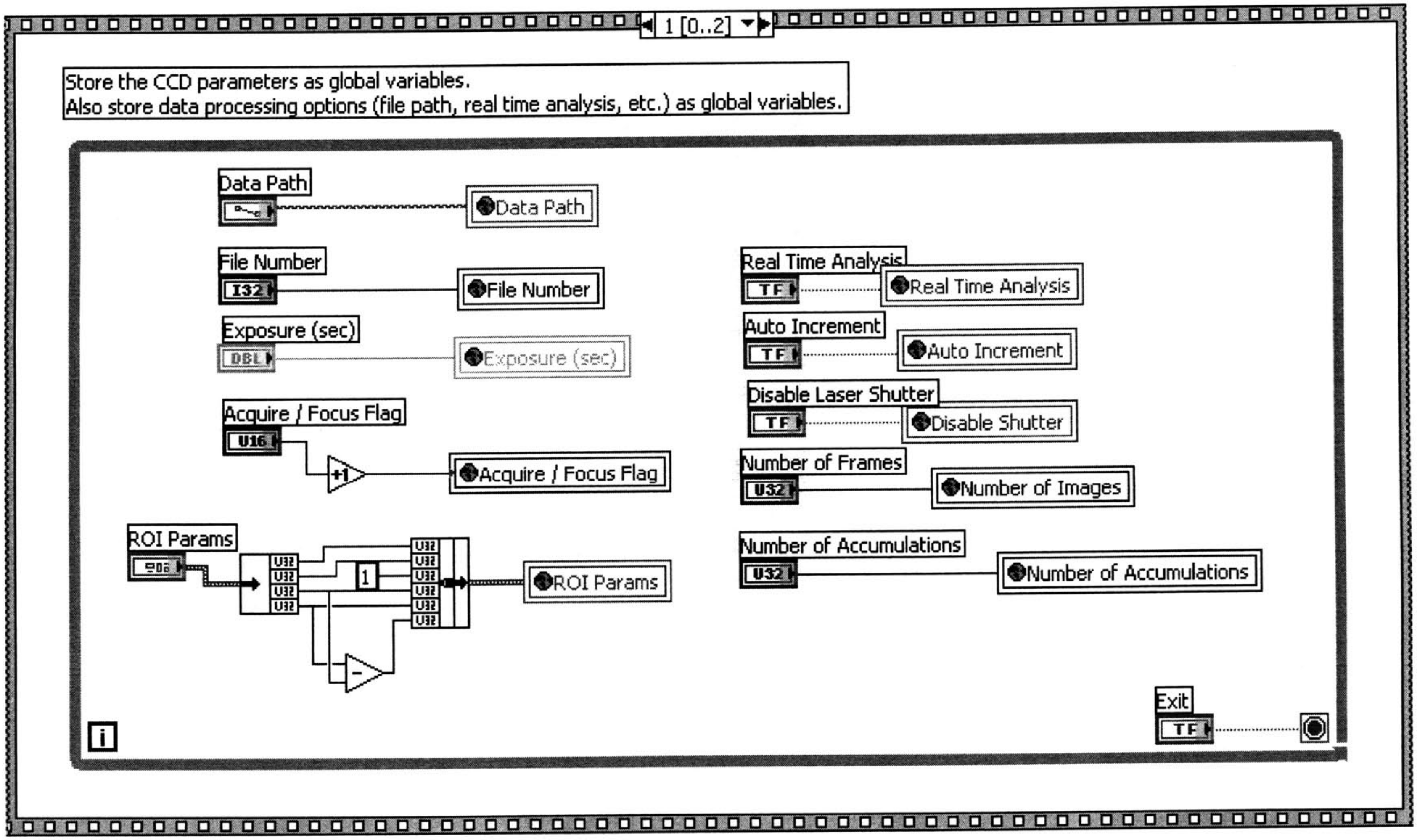
settings.vi

ROI

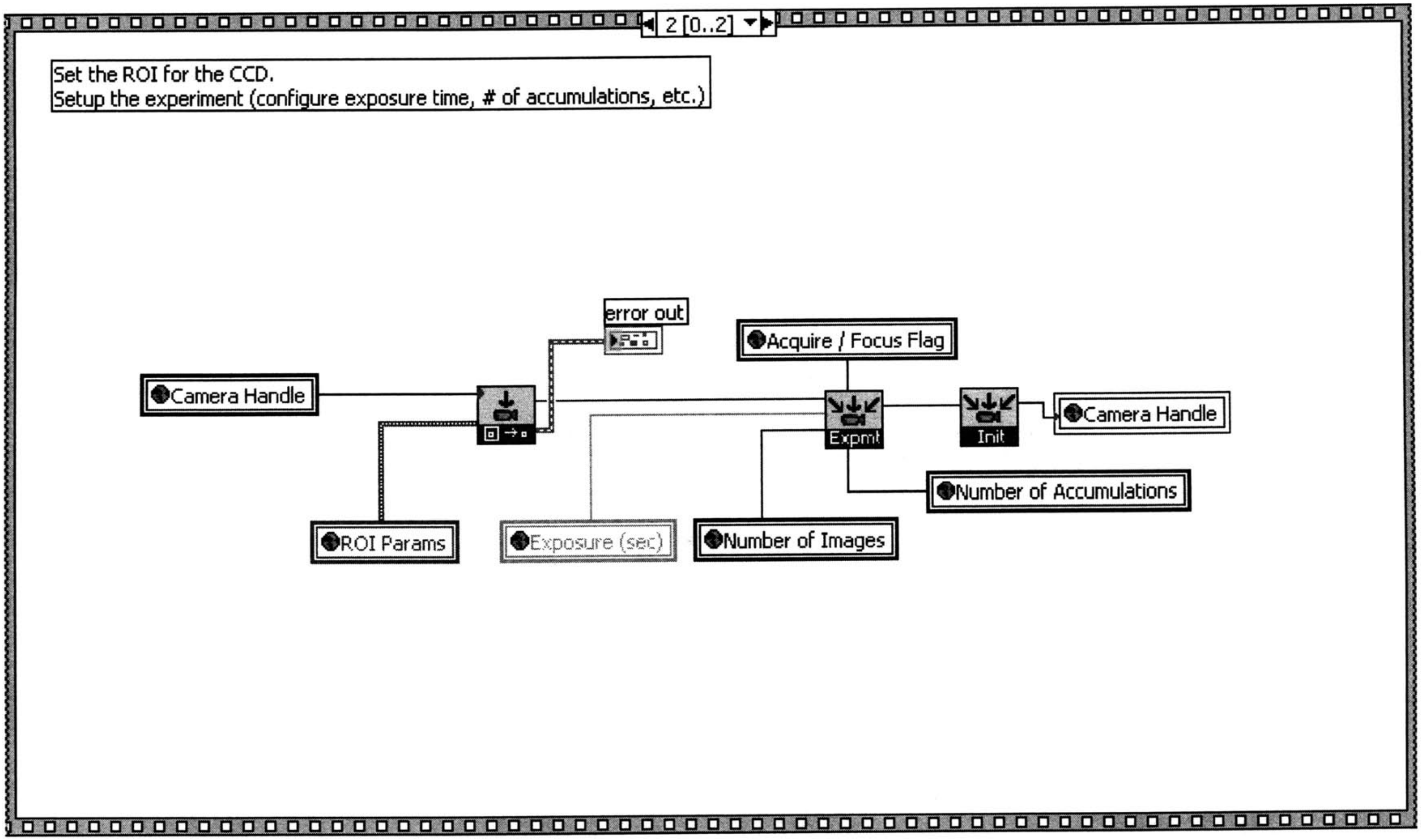
This vi is used to set the CCD parameters as well as real-time data analysis options.



Store the user specficied File Number value as a global variable.

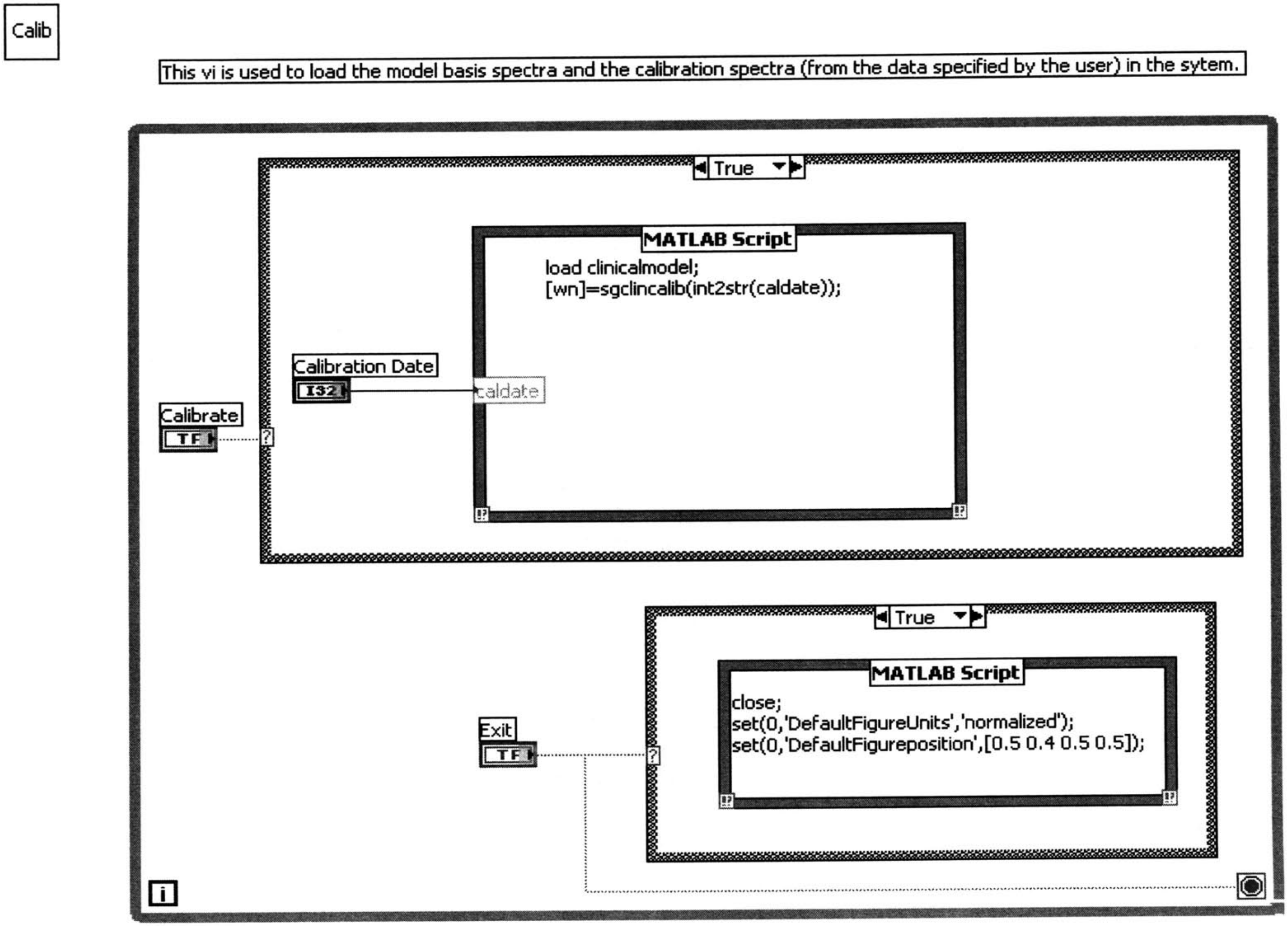


Store the CCD parameters as global variables.
Also store data processing options (file path, real time analysis, etc.) as global variables.



Set the ROI for the CCD.
Setup the experiment (configure exposure time, # of accumulations, etc.)

calibrate.vi

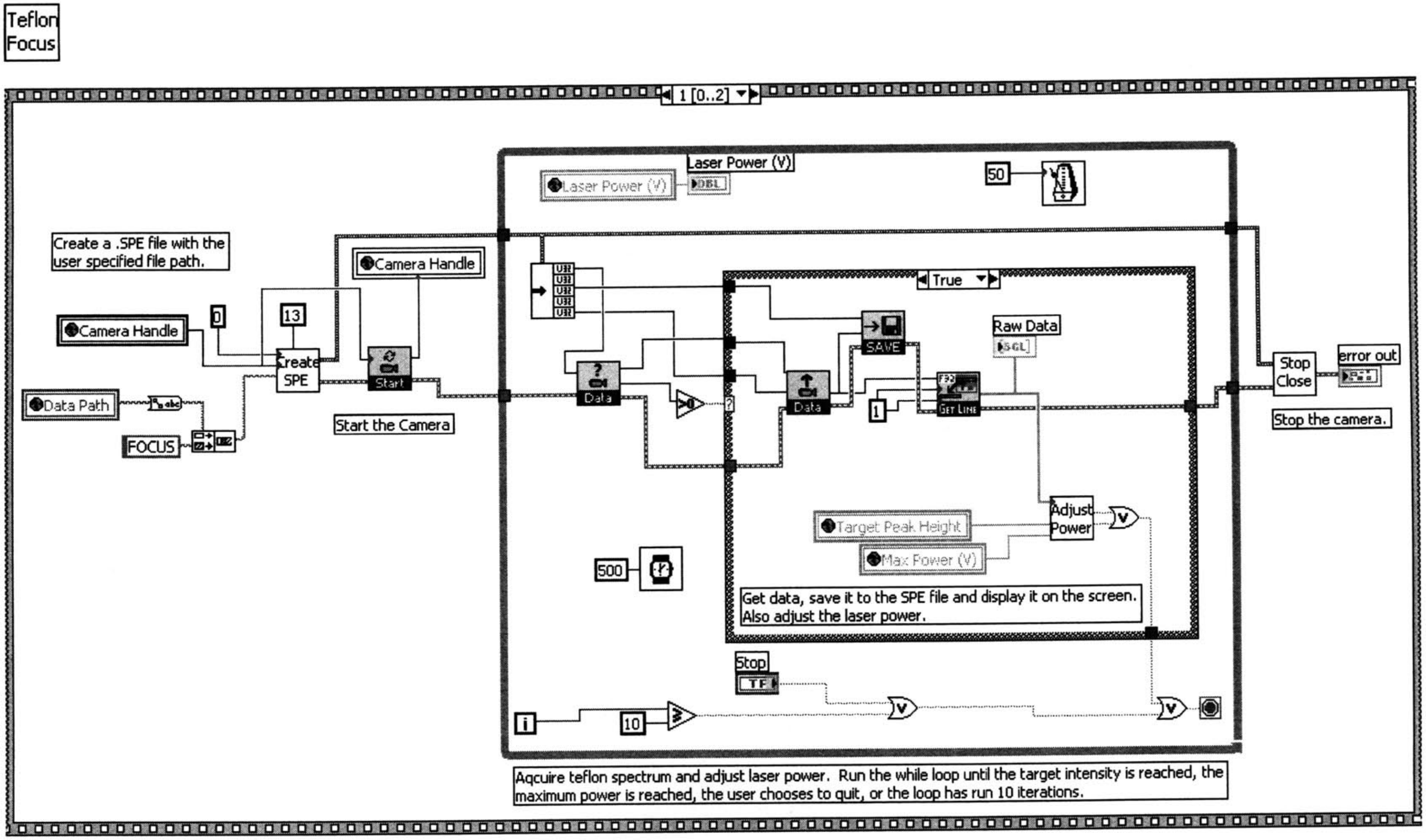


This vi is used to load the model basis spectra and the calibration spectra (from the data specified by the user) in the sytem.

```
load clinicalmodel;
[wn]=sgclincalib(int2str(caldate));
```

```
close;
set(0,'DefaultFigureUnits','normalized');
set(0,'DefaultFigureposition',[0.5 0.4 0.5 0.5]);
```

teflon focus.vi



Create a .SPE file with the user specified file path.

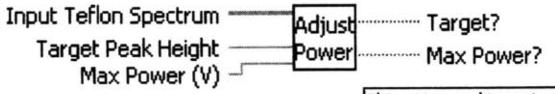
Start the Camera

Get data, save it to the SPE file and display it on the screen. Also adjust the laser power.

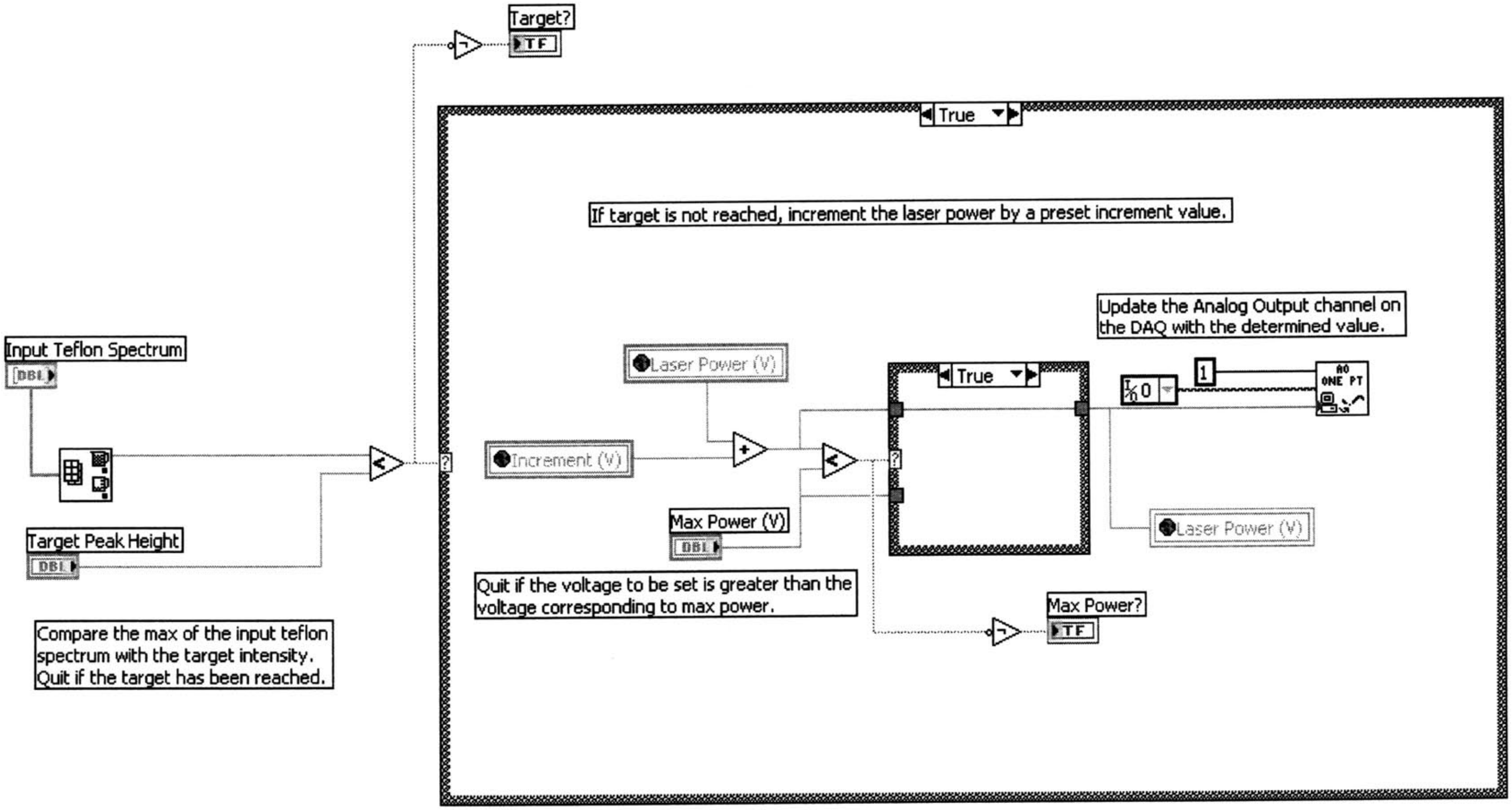
Stop the camera.

Aqcuire teflon spectrum and adjust laser power. Run the while loop until the target intensity is reached, the maximum power is reached, the user chooses to quit, or the loop has run 10 iterations.

adjust power (increment).vi



This vi is used to automatically change the voltage provided to set the laser power based on the value of the teflon intensity.


If target is not reached, increment the laser power by a preset increment value.

Update the Analog Output channel on the DAQ with the determined value.

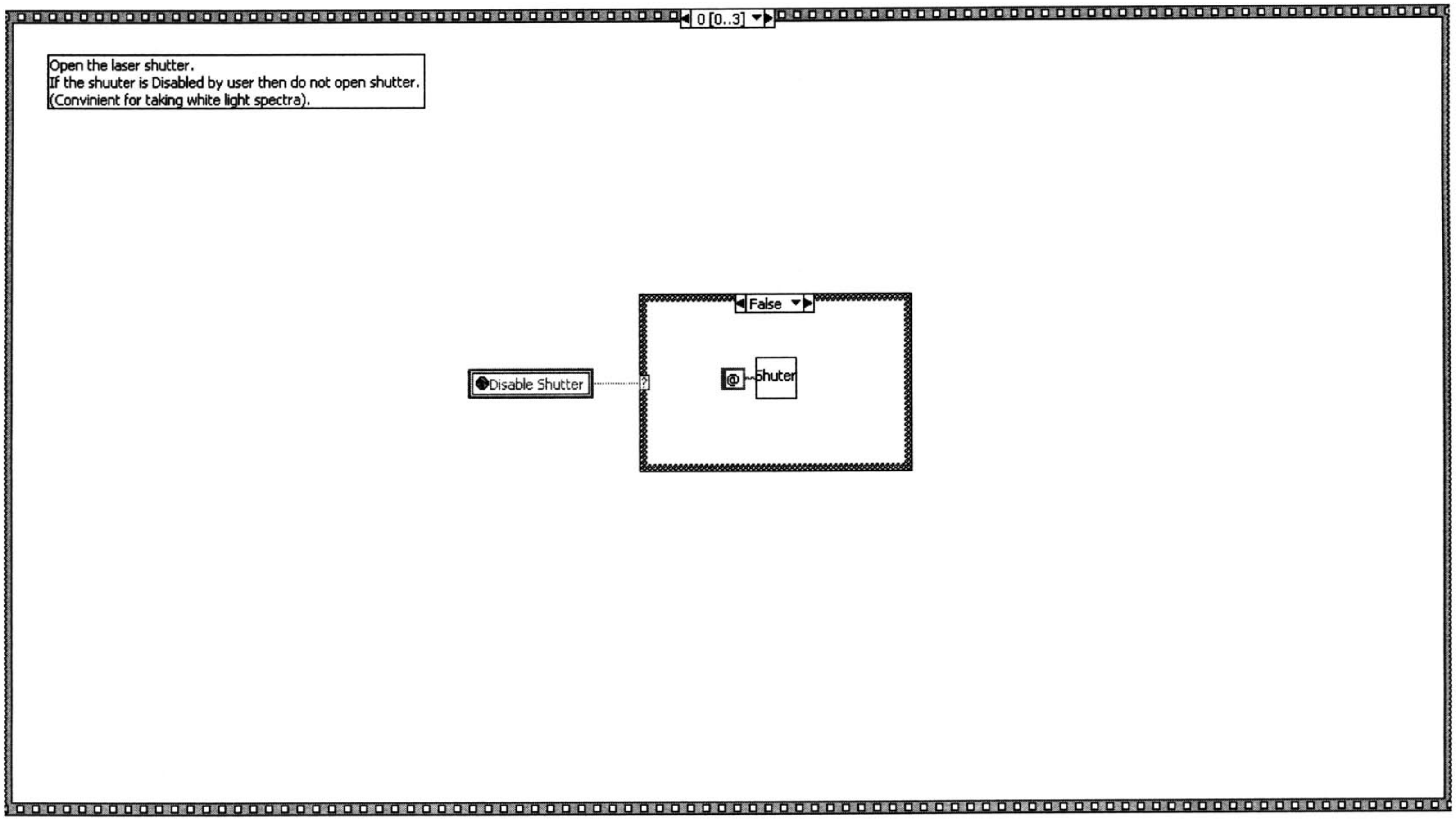
Quit if the voltage to be set is greater than the voltage corresponding to max power.

Compare the max of the input teflon spectrum with the target intensity. Quit if the target has been reached.

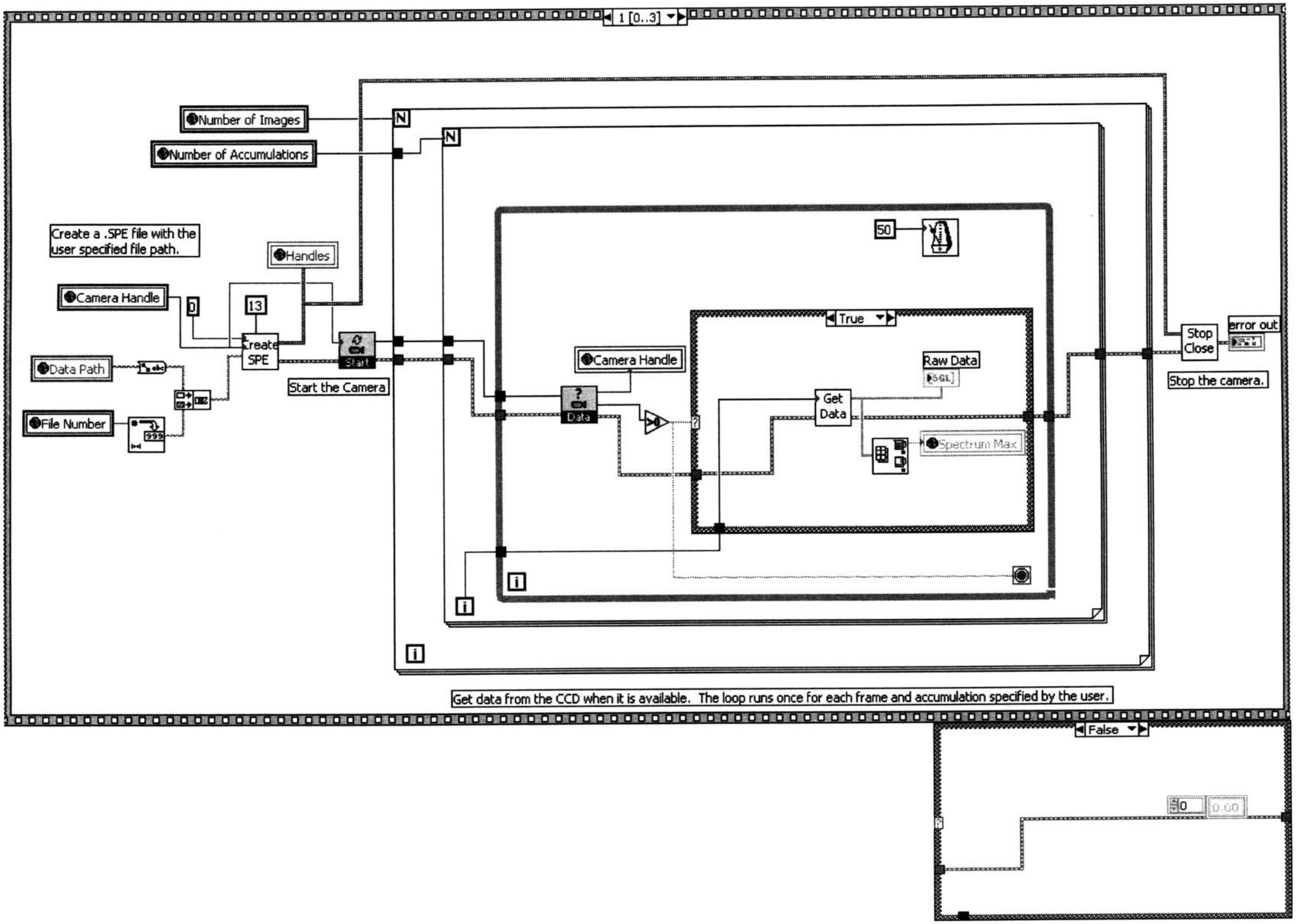
acquire.vi

Aquire

This vi is used to acquire the data and analyze it in real-time.



Open the laser shutter.
If the shuuter is Disabled by user then do not open shutter.
(Convinient for taking white light spectra).

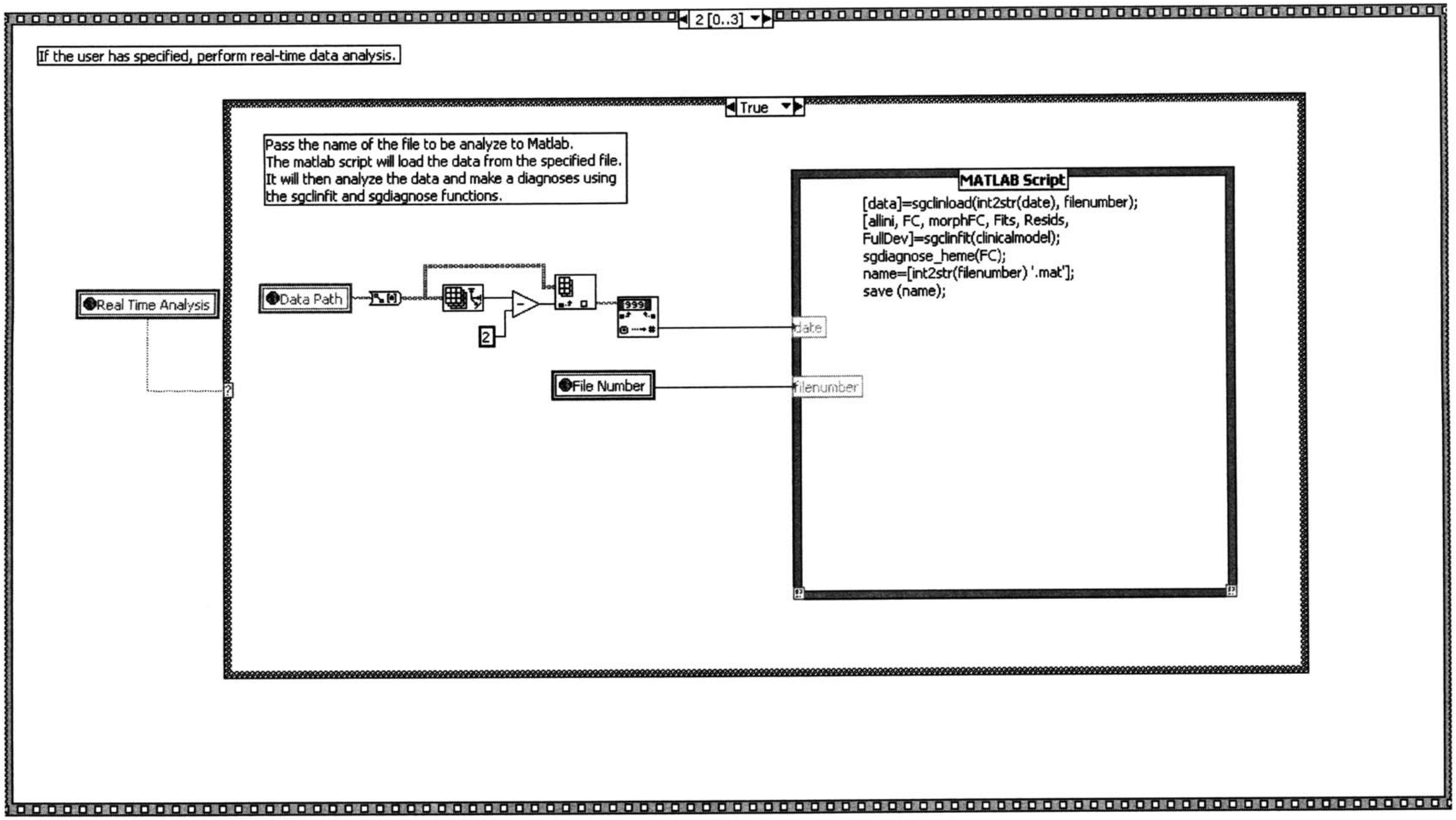


Create a .SPE file with the user specified file path.

Start the Camera

Get data from the CCD when it is available. The loop runs once for each frame and accumulation specified by the user.

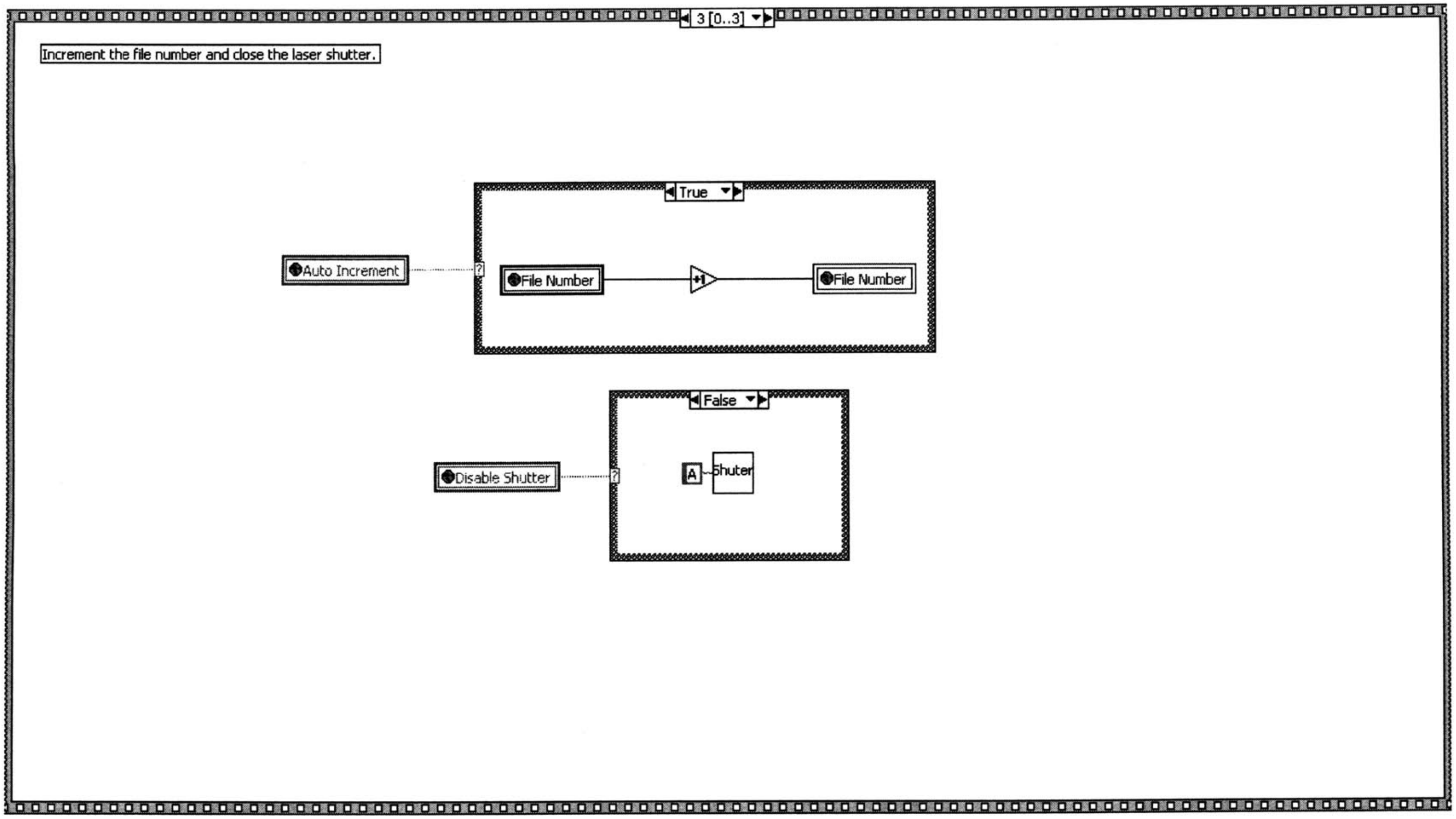
Stop the camera.



If the user has specified, perform real-time data analysis.

Pass the name of the file to be analyze to Matlab.
The matlab script will load the data from the specified file.
It will then analyze the data and make a diagnoses using
the sgclinfit and sgdiagnose functions.

```
[data]=sgclinload(int2str(date), filenumber);
[allini, FC, morphFC, Fits, Resids,
FullDev]=sgclinfit(clinicalmodel);
sgdiagnose_heme(FC);
name=[int2str(filenumber) '.mat'];
save (name);
```


3 [0..3]
Increment the file number and close the laser shutter.
Auto Increment
True
File Number
+1
File Number
Disable Shutter
False
A
Shuter

Appendix B

Matlab Subroutines

```
function [wn] = sgclincalib(Date);

%function [wn] = sgclincalib('Date');
%
%This program creates calibrations for clinical Raman data. The only output is the wavenumber file which
%is cut to the usable range. The only necessary input is the date of the procedure 'mmddyy', but the
%program relies on the pre-configured numbering scheme for clinical procedures. Global variables for
%various backgrounds necessary for proper analysis are also created. These include the white light
%spectrum for spectral correction in clinload.m, and a smoothed aluminum spectrum which is used in
%clinload.m for removal of the fiber background. Globals are also created for cutting of the data in
%clinload.m

clf

global wn
wn=[];
acet=[];
global alum
alum=[];
global white
white=[];
global minpix
minpix=[];
global maxpix
maxpix=[];

%Create wavenumber calibration
csfilename=sprintf('%s', 'mw_', Date, '_cs_');
acetname=sprintf('%s',csfilename,'2','.txt')
acet=load(acetname);
%acet=acet';
alumname=sprintf('%s',csfilename,'3','.txt')
alum=load(alumname);
%alum=alum';
wn=tylenol_ccf(acet/acet(1)-alum/alum(1));
plot(wn,acet/acet(1)-alum/alum(1));

%Cut wavenumber calibration to usable range
minpix=max(find(wn<680));
maxpix=min(find(wn>1800));
wn=wn(minpix:maxpix);
wn=wn';

%Load white light spectrum
whitename=sprintf('%s',csfilename,'4','.txt')
white=load(whitename);
%white=white';

%Cut white light and aluminmum spectra
alum=alum(minpix:maxpix);
white=white(minpix:maxpix);

%Normalize and correct data for spectral response
```

```
white=white/max(white);
alum=alum./white;
alum=alum/max(alum);
alum=smooth(smooth(smooth(smooth(alum))));
--------------------------------------------------------------------------------------------------------------
function [data] = sgclinload(Date, filenumber);

%function [data] = sgclinload('Date,' filenumber);
%
%This program loads data from in vivo Raman procedures. It should be preceeded by clincalib.m. It
%averages the data from all frames of a given acquisition. Then corrects the data for white light spectral
%response and cut the data into the usable range. It also normalizes the data by the maximum intensity.
%
%Load in vivo data
infilename=sprintf('%s','mw_', Date, '_in_', int2str(filenumber), '.txt');
global data
data=load(infilename);
%data=data';

global minpix
global maxpix
global white

%Average the data
if size(data,1)>1
  data=mean(data(1:size(data,1),minpix:maxpix));
  else data=data(1,minpix:maxpix);
end
%White light devision
data=data./white;

%Normalize data
data=data/max(data);



--------------------------------------------------------------------------------------------------------------



function [allini, FC, morphFC, Fits, Resids, FullDev] = sgclinfit(Model);

%function [allini, FC, morphFC, Fits, Resids, FullDev] = sgclinfit(Model);
%
%This program analyzes clinical Raman data. It should be preceeded by %clincalib.m and clinload.m.
%The only necessary input is the Model used for analysis. allini is a matrix of the clinical data. FC are the
%normalized fit %coefficients from the model for the spectrum in allini. morphFC are the normalized fit
%coefficients of the elements from the morphological model present in the Model used for analysis. Fits is
%the spectral fit to the data. Resids is the residual to the fit. FullDev is the standard deviation of the
%residuals created from the fits that %result from fitting the model to the data with various weightings of
%the aluminum background subtraction. FullDev should be checked to ensure that a minimum has been
%achieved, indicating proper subtraction of the aluminum background. Finally, the program plots the data,
%along with the fit and residual.
```

```
global wn
global alum
global data

FC=[];
morphFC=[];
Fits=[];
Resids=[];
FullDev=[];
meanFC=[];
stdFC=[];

P=[];
calibwn=[];
calibP=[];
devP=[];
normallFC=[];
allFits=[];
allResids=[];
x=[];
y=[];

%Subtract various weightings of the aluminum background
  for j=1:20,
    P(j,:)=data-alum/(2*0.1*j);
  end;

%Remove the fluorescence background
  P=chopfit(P,6,1);

%Calibrate the data matrix
  [calibP,calibwn]=calibrate2(P, wn, 686, 1788, 2);

%Renormalize the data
  for j=1:20,
     calibP(j,:)=calibP(j,:)/max(calibP(j,:));
  end

%Calulate the fit coefficients
  FC=nnfit(calibP, Model);

%Correct for oversubtraction of the sapphire spectrum,
  if FC(:,12)==0
     Model(12,:)=-Model(12,:);
     y=-1;
     %recalculate fit coefficients with inverted sapphire spectrum
       FC=nnfit(calibP, Model);
  else
     y=1;
  end

%Caluculate the fits
```

```
   allFits=FC*Model;

%Calculate residuals
allResids=calibP-allFits;

%Calculate the standard deviation of the residuals
   for j=1:20
      devP(j)=std(allResids(j,:));
    end

FullDev=devP;

%Find the smallest residual for each exposure
x=find(devP==min(devP));

%Normalize the all fit coefficients

   normallFC=FC(x,:);

clear FC

   for j=1:12,
      FC(j)=100*normallFC(j)/sum(normallFC);
   end
   FC(12)=y*FC(12);

%Normalize the morphological model fit coefficients

   for j=1:8,
      morphFC(j)=100*normallFC(j)/sum(normallFC(1:8));
   end

clear allini

   allini=calibP(x,:);

   Fits=allFits(x,:);

   Resids=allResids(x,:);

%Plot the data
set(0,'DefaultFigureUnits','normalized');
set(0,'DefaultFigureposition',[0.5 0.4 0.5 0.5]);
figure(1);
clf;

subplot(1,1,1);
plot(calibwn, allini, 'b.');
hold
plot(calibwn, Fits,'r');
plot(calibwn, Resids+2*min(allini),'k');
axis([680 1800 min(Resids)+2*min(allini) max(allini)]);
set(gca,'YTick',-1:.25:1);
```

```
set(gca,'YTick',-1:.25:1);
xlabel('Raman Shift (cm^{-1})', 'fontsize', 12);
ylabel('Intensity (a.u.)', 'fontsize', 12);
```

```
function f = sgdiagnose_heme(FC);

%Function applies the diagnostic algorithm to the fit coefficients and displays the diagnosis on the screen.

%function f = sgdiagnose(morphFC);
%
morphFC=[];
morphFC=[FC(1:8)];
morphFC=100*(morphFC/sum(morphFC));
nobetaFC=[];
nocmFC=[];

%Remover beta Carotene and re-normalize the fit coefficients.
nobetaFC=[morphFC(1:6),morphFC(8)];
nobetaFC=nobetaFC/sum(nobetaFC);

%Extract calcium mineralization.
cm=nobetaFC(3);

%Remove calcification and re-normalize the fit coefficients.
nocmFC=[nobetaFC(1:2),nobetaFC(4:7)];
nocmFC=nocmFC/sum(nocmFC);

%Extract cholesterol crystal + foam cell/nacrotic core.
ccfc=nocmFC(2)+nocmFC(5);


%Plot
set(0,'DefaultFigureUnits','normalized');
set(0,'DefaultFigureposition',[0.01 0.4 0.47 0.5]);
figure(2);
clf;

%Plot the normal or calcified vs. non-calcified plaque line.
plot([(0.07/0.30) 1],[0 0.23]);
hold;

%Plot the normal vs. calcified plaque line.
plot([(0.24/0.79) 0],[(0.17-0.48*(0.24/0.79)) 0.17]);

%Plot diagnosis.
plot(ccfc,cm,'rd','MarkerSize',12,'MarkerFaceColor','r');
```

```
if cm<(-0.07+0.30*ccfc)
  title('Diagnosis: Non-Calcified Plaque','fontsize', 14);
elseif cm<(0.17-0.48*ccfc)
  title('Diagnosis: Non-Atherosclerotic Tissue','fontsize', 14);
else
  title('Diagnosis: Calcified Plaque','fontsize', 14);
end

axis([0 1 0 1]);
xlabel('Cholesterol+FCNC', 'fontsize', 12);
ylabel('Calcification', 'fontsize', 12);

%Display fit coefficients
set(0,'DefaultFigureUnits','normalized');
set(0,'DefaultFigureposition',[0.5 0.03 0.49 0.15]);
figure(3);
clf;

axis([0 1 0 1]);
axis off;

text(0,1,['Collagen Fibers:' ' ' num2str(morphFC(1))],'fontsize',12);
text(0,0.75,['Cholesterol Crystals:' ' ' num2str(morphFC(2))],'fontsize',12);
text(0,0.5,['Calcified Minerals:' ' ' num2str(morphFC(3))],'fontsize',12);
text(0,0.25,['Elastic Lamina:' ' ' num2str(morphFC(4))],'fontsize',12);
text(0.5,1,['Adventitial Adipocytes:' ' ' num2str(morphFC(5))],'fontsize',12);
text(0.5,0.75,['Foam Cells/Necrotic Core:' ' ' num2str(morphFC(6))],'fontsize',12);
text(0.5,0.5,['Beta Carotene:' ' ' num2str(morphFC(7))],'fontsize',12);
text(0.5,0.25,['Smooth Muscle Cells:' ' ' num2str(morphFC(8))],'fontsize',12);
text(0,0,['Hemoglobin:' ' ' num2str(FC(9))],'fontsize',12);

set(0,'DefaultFigureUnits','normalized');
set(0,'DefaultFigureposition',[0.5 0.4 0.5 0.5]);
```
