Title
Techniques in two-photon microscopy for neuroscience

Permalink
https://escholarship.org/uc/item/74j959pq

Author
Driscoll, Jonathan Daniel

Publication Date
2011

Peer reviewed|Thesis/dissertation
Techniques in Two-Photon Microscopy for Neuroscience

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Physics

by

Jonathan Daniel Driscoll

Committee in charge:

Professor David Kleinfeld, Chair
Professor Gert Cauwenberghs, Co-Chair
Professor Alex Groisman
Professor Harvey J. Karten
Professor Tom O’Neil

2011
The dissertation of Jonathan Daniel Driscoll is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

University of California, San Diego

2011
# TABLE OF CONTENTS

Signature Page ................................................................. iii

Table of Contents .......................................................... iv

List of Figures ............................................................... vi

Acknowledgements ........................................................... vii

Vita and Publications ....................................................... viii

Abstract of the Dissertation ................................................. ix

1 Introduction ................................................................. 1
  1.1 Two-Photon Microscopy for Biology and Neuroscience ................. 1
  1.2 Blood Flow in the Brain ............................................. 3
  1.3 Organization of this Dissertation .................................. 6

2 Two-photon imaging of blood flow in cortex ................................ 8
  2.1 Abstract of Chapter ................................................. 8
  2.2 Introduction .......................................................... 8
  2.3 Protocol - Cranial Window Preparations in the Rat or Two-Photon Imaging of Blood Flow in the Cordex .............................. 9
    2.3.1 Materials ...................................................... 11
    2.3.2 Experimental Method ....................................... 13
  2.4 Example Application ................................................ 16
  2.5 Generation of Spatially Optimized Line Scans ......................... 20
  2.6 Summary ............................................................. 22
  2.7 Recipe ............................................................... 22
  2.8 Acknowledgements .................................................. 22

3 Photon counting, censor corrections, and lifetime imaging for improved detection in two-photon microscopy .............................. 24
  3.1 Abstract of Chapter ................................................. 24
  3.2 Introduction .......................................................... 24
  3.3 Methods ............................................................... 26
  3.4 Results ............................................................... 30
  3.5 Discussion ............................................................ 36
    3.5.1 Acknowledgments .............................................. 38
  3.6 Supplementary Material - Estimating the Emission Count .......... 40
  3.7 Supplementary Material - Photon Counting Circuitry ............... 46
    3.7.1 Fast Discriminator Circuitry ................................ 46
    3.7.2 CPLD Circuitry ................................................ 46
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Two-photon microscopy to measure blood flow and concurrent brain cell activity</td>
<td>54</td>
</tr>
<tr>
<td>4.1</td>
<td>Abstract of Chapter</td>
<td>54</td>
</tr>
<tr>
<td>4.2</td>
<td>Introduction</td>
<td>54</td>
</tr>
<tr>
<td>4.3</td>
<td>Choice of preparation</td>
<td>56</td>
</tr>
<tr>
<td>4.4</td>
<td>Cranial windows</td>
<td>56</td>
</tr>
<tr>
<td>4.5</td>
<td>Transcranial windows</td>
<td>57</td>
</tr>
<tr>
<td>4.6</td>
<td>Localization of active areas</td>
<td>57</td>
</tr>
<tr>
<td>4.7</td>
<td>Measurement of blood flow dynamics in single cortical vessels</td>
<td>58</td>
</tr>
<tr>
<td>4.8</td>
<td>Simultaneous imaging of blood flow and local cellular activity</td>
<td>61</td>
</tr>
<tr>
<td>4.9</td>
<td>Stimulated and basal hemodynamics in awake mice</td>
<td>64</td>
</tr>
<tr>
<td>4.10</td>
<td>Imaging of blood flow in deep cortical layers</td>
<td>66</td>
</tr>
<tr>
<td>4.11</td>
<td>Automatic Identification of Regions of Interest</td>
<td>66</td>
</tr>
<tr>
<td>4.12</td>
<td>summary</td>
<td>69</td>
</tr>
<tr>
<td>4.13</td>
<td>Acknowledgments</td>
<td>70</td>
</tr>
<tr>
<td>5</td>
<td>Conclusions and Future Directions</td>
<td>71</td>
</tr>
<tr>
<td>A</td>
<td>Ultrafast PMT Preamplifier and Discriminator</td>
<td>73</td>
</tr>
<tr>
<td>A.1</td>
<td>Component Overview</td>
<td>75</td>
</tr>
<tr>
<td>A.1.1</td>
<td>Fast DC coupled amplifier</td>
<td>75</td>
</tr>
<tr>
<td>A.1.2</td>
<td>Ultrafast Comparator</td>
<td>75</td>
</tr>
<tr>
<td>A.1.3</td>
<td>Level shifting</td>
<td>75</td>
</tr>
<tr>
<td>A.2</td>
<td>Circuit description</td>
<td>75</td>
</tr>
<tr>
<td>B</td>
<td>MpScope3 User and Developer’s Guide</td>
<td>77</td>
</tr>
<tr>
<td>B.1</td>
<td>Overview</td>
<td>77</td>
</tr>
<tr>
<td>B.2</td>
<td>Setting Up The Program</td>
<td>78</td>
</tr>
<tr>
<td>B.2.1</td>
<td>Software Drivers and Libraries</td>
<td>78</td>
</tr>
<tr>
<td>B.2.2</td>
<td>Hardware for MpScope</td>
<td>79</td>
</tr>
<tr>
<td>B.3</td>
<td>Using the Program</td>
<td>80</td>
</tr>
<tr>
<td>B.3.1</td>
<td>Initial startup</td>
<td>80</td>
</tr>
<tr>
<td>B.3.2</td>
<td>User Configuration</td>
<td>80</td>
</tr>
<tr>
<td>B.3.3</td>
<td>Using the Stages</td>
<td>80</td>
</tr>
<tr>
<td>B.3.4</td>
<td>Scripting with MATLAB</td>
<td>81</td>
</tr>
<tr>
<td>B.4</td>
<td>Developer’s Guide</td>
<td>82</td>
</tr>
<tr>
<td>B.4.1</td>
<td>Setting up and Using C#</td>
<td>83</td>
</tr>
<tr>
<td>B.4.2</td>
<td>Setting up and using NI-DAQmx</td>
<td>86</td>
</tr>
<tr>
<td>B.4.3</td>
<td>Automatically Generated Documentation (Doxygen)</td>
<td>87</td>
</tr>
<tr>
<td>B.4.4</td>
<td>Setting up and using HDF5 (file format)</td>
<td>88</td>
</tr>
<tr>
<td>B.4.5</td>
<td>HDF5</td>
<td>88</td>
</tr>
<tr>
<td>B.4.6</td>
<td>Setting up and Using the Galil Controller</td>
<td>90</td>
</tr>
<tr>
<td>B.4.7</td>
<td>Servo Power Control</td>
<td>92</td>
</tr>
<tr>
<td>B.4.8</td>
<td>Threading and Program Execution in C#</td>
<td>92</td>
</tr>
<tr>
<td>Bibliography</td>
<td>94</td>
<td></td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Setup for in vivo imaging of blood flow through a cranial window.</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Simultaneous measurement of diameter and velocity in multiple vessels using spatially optimized line scans.</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Photon counting system</td>
<td>28</td>
</tr>
<tr>
<td>3.2</td>
<td>Comparison of digital and analog images under conditions of low emission rates</td>
<td>31</td>
</tr>
<tr>
<td>3.3</td>
<td>Measured and estimated true photon counts and signal-to-noise ratios (SNR) for imaging under conditions of high emission rates</td>
<td>34</td>
</tr>
<tr>
<td>3.4</td>
<td>Fluorescence intensity and lifetime imaging of spectrally similar dyes in a mouse model for experimental stroke</td>
<td>37</td>
</tr>
<tr>
<td>3.5</td>
<td>Fast discriminator schematic</td>
<td>47</td>
</tr>
<tr>
<td>3.6</td>
<td>Counting circuitry for lifetime imaging.</td>
<td>48</td>
</tr>
<tr>
<td>3.7</td>
<td>Time delay and gating circuitry.</td>
<td>51</td>
</tr>
<tr>
<td>3.8</td>
<td>Measuring counts per laser repetition.</td>
<td>53</td>
</tr>
<tr>
<td>4.1</td>
<td>Intrinsic optical signal imaging for functional region targeting.</td>
<td>59</td>
</tr>
<tr>
<td>4.2</td>
<td>Example of automated cell segmentation and user defined fast scanning for functional imaging in rat parietal cortex.</td>
<td>63</td>
</tr>
<tr>
<td>4.3</td>
<td>Spontaneous and stimulus induced vascular dynamics in the cortex of awake mouse</td>
<td>65</td>
</tr>
<tr>
<td>4.4</td>
<td>Deep imaging of cortical angioarchitecture and blood flow.</td>
<td>67</td>
</tr>
<tr>
<td>A.1</td>
<td>Fast DC amplifier and Discriminator</td>
<td>74</td>
</tr>
<tr>
<td>B.1</td>
<td>MpScope 3 Main Components</td>
<td>78</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

The author would like to acknowledge all those who have contributed to this work. Thank you to the committee members, Professors Gert Cauwenberghs, Alex Groisman, Harvey J. Karten, and Tom O’Neil, and especially Professor David Kleinfeld, who oversaw the majority of the thesis work.

The author would also like to thank all the members of the Kleinfeld lab, who make the difficult task of multidisciplinary science not only possible, but intellectually satisfying as well. The innumerable discussions not only ranged across all the sciences, but across everything from music and politics to the shape world of the future as well. In particular, Andy Shih’s careful surgeries and thorough understanding of cerebral dynamics made much of the data presented here possible.

Chapter 2, in full, is a reprint of the material as it appears in Imaging in Neuroscience: A Laboratory Manual, F. Helmchen and A. Konnerth (eds), Chapter 84 “Two-Photon Imaging of Blood Flow in Cortex,” J.D. Driscoll, A.Y. Shih, P.J. Drew, I. Valmianski, G. Cauwenberghs, and D. Kleinfeld, Cold Spring Harbor Laboratory Press, 2011. The dissertation author was the primary author of this chapter.

Chapter 3, in full, is a reprint the material as it appears in The Journal of Neurophysiology, “Photon counting, censor corrections, and lifetime imaging for improved detection in two-photon microscopy,” J.D. Driscoll, A. Y. Shih, S. Iyengar, J.J. Field, G.A. White, J.A. Squire, G. Cauwenberghs, and D. Kleinfeld, Americal Physiological Society, 2011. The dissertation author was the primary author of this paper.

Chapter 4, in full, is a reprint of material as it is to appear in Optical Imaging of Cortical Dynamics, B. Weber and F. Helmchen (eds.), Chapter 3.3 “Two-photon microscopy to measure blood flow and concurrent brain cell activity,” A.Y. Shih, J.D. Driscoll, M.J. Pesavento, and D. Kleinfeld, Elsvier 2012, (in press). The dissertation author was a contributing author to this chapter.
VITA

2000
B. A. in Physics, Philosophy minor, magna cum laude, Washington University in St. Louis

2011
Ph. D. in Physics, University of California, San Diego

PUBLICATIONS

Publications (First Author):


Publications (Contributing Author):


ABSTRACT OF THE DISSERTATION

Techniques in Two-Photon Microscopy for Neuroscience

by

Jonathan Daniel Driscoll

Doctor of Philosophy in Physics

University of California, San Diego, 2011

Professor David Kleinfeld, Chair
Professor Gert Cauwenberghs, Co-Chair

Fluorescence imaging in biology has become a fundamental tool to understanding structure and function. In addition, the past two decades has seen the first implementation, and subsequent widespread use, of two-photon scanning laser microscopy as a tool to allow fluorescent data to be acquired in new modalities, such as imaging at depth, complete three-dimensional reconstructions, rapid acquisition of cellular activity, and accurate assessment of cerebral blood flow at the level of individual vessels.

This thesis examines the use of two-photon microscopy to study neurovascular coupling, the process by which computational mechanism in the brain actively reroute blood flow to regions of increased neural activity, where it is needed the most. Understanding the mechanisms through which this occurs is an active area of research, and represent a “sweet spot” in terms of complexity: the function of the system can be stated simply,
and many of the mechanisms involved are understood in isolation. However, putting everything together into a complete picture has remained elusive, and understanding how the brain accomplishes this seemingly simple task should provide insight into the general computational mechanisms of the brain.

Central to this research was the development of scanning hardware and software techniques that allow the rapid acquisition of the biologically relevant variables with high signal-to-noise ratios. This lead to the development of computer control software for creating user-defined scan lines, which can interface with machine learning programs for quickly and accurately identifying relevant regions in the field of view, and photon-counting hardware to make the most out of the limited amount of light that can be collected per unit time.

The photon-counting hardware and firmware, developed at UCSD, let to some unexpected and previously unanalyzed findings about what happens to the signal under high light conditions, and the resulting analysis and corrections are generally useful to two-photon microscopy. It is hoped that the data acquisition and analysis tools discussed here, and currently being used in the laboratory, will help advance the field of neuroscience, and provide some insight to long standing questions in brain function.
1 Introduction

1.1 Two-Photon Microscopy for Biology and Neuroscience

The phenomenon of fluorescence has been known for over 150 years. In 1852, Sir G.G. Stokes published a series of experiments in which he observed molecules that, when excited with light in the violet or ultraviolet wavelengths, emitted light in a redder portion of the spectrum [Stokes, 1852].

When light impinges upon a fluorescent molecule, photons with a characteristic energy can be absorbed, and the energy is stored by moving an electron to a higher energy state within the molecule. The molecule can stay in an excited state, typically on the order of nanoseconds, before reemitting the energy as a photon. However, since some of the energy of the excitation photon is lost to vibrational and other modes in the molecule, the reemitted photon has a shorter wavelength, a phenomenon known as Stokes shift.

The development of quantum mechanics in early part of the 20th century provided a theoretical basis for understanding the phenomenon, and also led to an interesting prediction. In her 1931 thesis dissertation, Maria Goeppert Mayer predicted that if two photons, each with half the energy need for an excitation, arrive at the molecule at nearly the same time, both could be absorbed and excite the molecule into a fluorescent state, just as in single photon fluorescence [Göppert-Mayer, 1931].

From the uncertainty principle, the uncertainty in an energy measurement made in a given time is given by

\[ \Delta E \Delta t \geq \frac{\hbar}{2} \] (1.1)

A photon has an energy of 1 eV at 1243 nm. For energies of about 2 eV, corresponding to a red photon, the characteristic time for two photon absorption would
be about 0.16 femtoseconds. That is, as a rough estimate, two photons must pass the absorption cross section of a fluorescent molecule within a fraction of a femtosecond in order to have the same effect as a single photon. Thus, the threshold for where two photon absorption can take place is at an extremely high fluence of light, far outside the reaches of what was available at the time the phenomenon was originally proposed.

Another 30 years passed before the two photon absorption phenomena was demonstrated experimentally. In 1962, Abella published his results of the two photon absorption, and subsequent fluorescent emission, from low pressure Cesium vapor. In this case, a continuous wave, 2 Watt laser was used [Abella, 1962].

Once again, another 30 years passed before the two photon phenomena was put to use as a microscope. In 1990, Denk and colleagues published the results from the first two-photon laser scanning microscope. By scanning the focal point of the laser excitation beam around the sample, the group was able to acquire fluorescent images, segmented into a plane [Denk et al., 1990].

The key technology was the invention of the pulse mode laser, which uses a crystal with non-linear index of refraction, such as Titanium-Sapphire, and dispersion correction prisms to condense the energy from a continuous pump laser into brief pulses of photons. A modern pulse laser, such as used in this dissertation work, may operate at 76.5 MHz, meaning the pulses are spaced 13 nanoseconds apart in time. Each pulse will only last typically a few hundred femtoseconds. This represents a compression of the laser energy by a five orders of magnitude.

The pulsed laser allows imaging with laser powers in the range milliwatt range — Denk used an excitation power of 3 mW — rather than the 2 Watts originally used by Abella. This lower laser power, three orders of magnitude in this case, is absolutely necessary for imaging biological preparations, as continuous laser powers on the order of Watt cause severe damage to biological tissue.

Focusing the light from a pulsed laser to a small point using a microscope objective leads to a small volume of high intensity light in which two photon excitation can occur. The size of this volume is ideally diffraction limited, and therefore a function of the wavelength of the excitation pulse and the numerical aperture (NA) of the objective. To get an idea of the resolution, the axial radius of the focal point is (see [Tsai and Kleinfeld, 2009]):
where $NA$ is the numerical aperture of the lens and $n$ is the index of refraction of the media. The resolution in the Z direction is given by the confocal distance,

$$z_{\text{focus,axial}} = \frac{\pi n}{\lambda_0} r_{\text{focus,lateral}}$$

For a typical setup of $NA = 0.8$, with an excitation wavelength of 800 nm passing through water ($n = 1.33$), this gives a focal volume of approximately $0.15 \mu m^3$.

The result is that two-photon excitation allows only a confined region of fluorescence to be excited. By scanning the beam in the X and Y direction, a two-dimensional image can be formed. More advanced scanning patterns such as 3-dimensional scans or scans across only regions of interest within a plane, will be discussed in later chapters.

## 1.2 Blood Flow in the Brain

The enormous metabolic needs of the brain requires an efficient means of routing the oxygen and glucose rich blood to entire organ, and a means of dynamically reallocating the nutrient rich blood to regions where it is needed the most, such as regions of high neural activity. The process by which the brain accomplishes this is known as “neurovascular coupling,” and is a system which is studied in order to elucidate how the brain accomplishes a relatively well defined, but still somewhat poorly understood task.

Oxygen rich blood from the heart is distributed across the surface of the brain by a series of arteries that branch successively into smaller vessels, called arterioles. The penetrating arterioles dive into the cortex itself, bring blood into the cerebral tissue. These penetrating arterioles continue to branch into smaller vessels, until they reach the size of a vessel capillary. It is in the capillaries, which are so thin the red blood cells can only pass through in single file, that the nutrients and gasses can be exchanged with the surrounding tissue.

The vascular architecture which returns the blood back to the heart is essentially a mirror of the arteriole structure. Capillaries carrying oxygen and nutrient depleted blood merge into larger venules. The ascending venules carry the blood back to the surface of the brain, where it is returned to the heart.
This branching of the vascular architecture in the brain allows for an extremely high surface area by for nutrient exchange. Moreover, even though the total blood volume is only approximately 2% of the brain by volume, neuronal soma have a mean distance of only 15\(\mu\text{m}\) from the nearest capillary, and a maximal distance of 25\(\mu\text{m}\) [Tsai et al., 2009].

The cerebral vascular architecture also contains means to control the rate of blood flow to various regions of the brain. Certain arterioles are surrounded by smooth-muscle cells, which can contract or relax in order to change the diameter of the vessel. Small changes in the diameter of cells can have a dramatic effect of the resistance to flow in that vessel. Laminar fluid flow in a pipe, which is a fairly good first order approximation, follows Poiseuille’s law

\[ Q = \frac{\pi D^4 \Delta P}{128L \eta} \tag{1.4} \]

Which relates the volume flux (flow) \(Q\) of a liquid with viscosity \(\eta\) through a pipe with length \(L\) and diameter \(d\). Note the quartic dependence on diameter: a contraction of a vessel by 25%, which is entirely within physiological range, will increase the resistance to the flow by a factor of greater than three.

Laminar flow models of blood flow are a useful approximation, and adequate in this case to demonstrate the strong dependence of flow on resistance. Systematic deviations from this model occur in actual blood flow through vessels because of effects such as the finite diameter of red blood cells, and the composition of the vessel walls. As a computational tool, it is useful to assume Poiseuille’s law is the correct description, and vary \(\eta\) as a parameter of the vessel radius as a correction factor. Note also that in capillaries, the red blood cells pass in single file, and are typically distorted, and the laminar flow approximation does not hold. See [Secomb et al., 1998, Pries and Secomb, 2005], for instance.

The smooth muscles on the arterioles are in turn responsive to a number environmental conditions, or vasoactive compounds, found in the blood and brain tissue. Gasses such as nitric oxide (NO) and peptides such as vasoactive intestinal peptide (VIP) cause vasodilation, while several neuropeptides, including somatostatin (SOM) and neuropeptide Y (NPY) can cause vasoconstriction or contraction [Kleinfeld et al., 2011].

One special cell type — the astrocyte, or astroglia — is commonly implicated in controlling cerebral blood flow. Astrocytes are the “other” type of cell in the brain, traditionally much less studied than neurons, and have had a traditionally underappreciated
role and in brain function.

Although astrocytes do not spike as neurons do, they show slower voltage and calcium changes, and can function as processing units in a similar way to neurons. For instance, it has been shown that astrocytes exhibit response tuning curves, much as neurons do. In the ferret visual cortex, astrocytes labeled with the calcium indicator Oregon Green Bapta (OGB1) responded to features in the visual stimuli such orientation, spatial frequency, and location in the visual field. The responses were similar to that of nearby neurons [Schummers et al., 2008].

Astrocytes are thought to be directly involved in neurovascular coupling, since their processes synapse directly onto the smooth muscles of the arterioles. Moreover, it has been directly shown that calcium transients in the end feet of astrocytes cause constriction in arterioles in vitro [Mulligan and MacVicar, 2004], leading to the idea that they play a direct role in neurovascular function, although it is not currently clear how this plays out in vivo.

Astrocytes also wrap their processes around neuronal synapses. This so called “tripartite synapse” consists of a presynaptic terminal, which signals to a postsynaptic cell across the synaptic cleft. The cleft is surrounded by an astrocytic process, which can absorb and respond to neurotransmitter spillover from the neurons. The astrocyte thus becomes a computational part of the circuit. It has been shown that the astrocytes are not merely “listening” passively, as astrocytic activity can influence the activity of nearby neurons as well. See [Araque et al., 1999] for a review.

The logical implication is that of all of this is that the system by which the brain actively reroutes blood to areas where it is needed, a process known as “neurovascular coupling,” can be thought of as a computational system within the brain, and can be studied as a complete system in order to gain insight to elucidate the way the brain accomplishes a seemingly simple, straightforward task.

Roughly, the inputs to the system are neural activity or local oxygen content, actuators are the smooth muscle cells around the arterioles, which in turn affect the flow of blood. How exactly this is accomplished is an active area of research. The simplistic picture, where neuronal activity results in neurotransmitter spillover, which is taken up by astrocytes which in turn activate smooth muscle cells in the arterioles, has not been reliably demonstrated in vivo. Further confounding the picture, neural activity can lead to both vasoconstriction and vasodilation [Kleinfeld et al., 2011].
1.3 Organization of this Dissertation

Chapter 2 begins with an overview of surgical techniques to create a cranial window in rat and mice in order to gain optical access to the cortex for the study of blood flow and neural activity. The chapter also introduces a technique developed by the dissertation author to create a user-defined scan paths across regions of interest in the field of view, a technique which has been used extensively with the laboratory for blood flow, neural activity, and stroke studies.

Chapter 3 discusses the use of photon counting in two-photon microscopy. Although photon counting is a well established technique, having a history going back nearly 40 years, some new and interesting phenomenon were discovered in the course of this dissertation research. Under high laser excitation light, photons begin to arrive too fast to be resolved by the detection circuitry: two photons arriving within a few nanoseconds only result in a single count, leading to a apparent fluorescence intensity decreases. In many imaging techniques, such as imaging of calcium indicator dyes, this non-linearity results in inaccurate data.

We develop a model of correcting for the missing counts, and examine what happens to the noise of the reconstructed data. Although this had been done previously in the case of constant fluorescent excitation illumination, the case of pulsed excitation had not been previously studied. Central to this study was the use of custom built, reprogrammable electronics, schematics of which are included at the end of the chapter.

Chapter 4 is a review of the use of two-photon microscopy in the study of blood flow and neural activity, and gives a summary of the current state of the field.

Part A of the appendix contains a proposed improved circuit for photon counting. This circuit allows for counting at extremely high frequencies, contains no AC coupling, and consists of only a few components per channel. Note that the circuit has not yet been built, but has been designed by the dissertation author, and given a “sanity check” by Allen White of the UCSD electronics shop. This circuitry would be useful, for instance, for anyone who wanted to commercialize flexible, reconfigurable photon counting devices for use in fluorescence scanning microscopy.

Part B of the appendix is a user and developer’s guide to the MpScope3 software. While working in this dissertation, it was determined that a rewrite of the current software which controls the two-photon microscope was required to address shortcomings of the previous software version, such as limited frame size, a hard limit in data file size, and the
inability to use newer data acquisition cards. In addition MpScope3 was written in order to allow for the addition of more advanced features at a later date, such as three-dimensional scanning, intelligent control and advanced automation through MATLAB, and extremely large (gigapixel) imaging, all of which are techniques which are “on the radar” for future use in the laboratory.
2 Two-photon imaging of blood flow in cortex

2.1 Abstract of Chapter

Cerebral blood flow plays a central role in maintaining homeostasis in the brain, and its dysfunction leads to pathological conditions such as stroke. Moreover, understanding the dynamics of blood flow is central to the interpretation of data from imaging modalities — such as intrinsic optical signaling and functional magnetic resonance imaging — that rely on changes in cerebral blood flow and oxygen level to infer changes in the underlying neural activity. Recent advances in imaging techniques have allowed detailed studies of blood flow in vivo at high spatial and temporal resolutions. We discuss techniques to accurately measure cerebral blood flow at the level of individual blood vessels using two-photon laser-scanning microscopy. By directing the scanning laser along a user-defined path, it is possible to measure red blood cell velocity and vessel diameter across multiple vessels simultaneously. The combination of these measurements permits accurate assessment of total flux with sufficient time resolution to measure fast modulations in flux, such as those caused by heartbeat, as well as slower signals caused by vasomotion and hemodynamic responses to stimulus.

2.2 Introduction

Neural activity is accompanied by both increases and decreases in local blood flow to active brain regions [Woolsey et al., 1996, Derdikman et al., 2003, Devor et al., 2008]. Detection of functional changes in cerebral blood flow are the basis of wide-field brain imaging modalities, including intrinsic optical signaling [Grinvald et al., 1986, Frostig et al., 1990] and functional magnetic resonance imaging [Ogawa et al., 1990]. An understanding
of the fundamental mechanisms of neurovascular coupling is critical for interpreting such data and is an active area of research.

Two-photon laser scanning microscopy is an ideal method for studying blood in individual vessels, from the level of the pial vessels down to subsurface capillaries. Two-photon microscopy has been used to acquire red blood cell (RBC) velocity [Kleinfeld et al., 1998, Belayev et al., 2002, Chaigneau et al., 2003, Hirase et al., 2004, Chuquet et al., 2007, Zhang et al., 2007] or lumen diameter [Devor et al., 2007, Devor et al., 2008, Shih et al., 2009] from a single vessel in a given imaging trial. However, the oxygen- and nutrient-carrying capacity of a vessel is proportional to its flux, which requires knowledge of both RBC velocity and vessel diameter. Since these two parameters can change independently, both must be measured simultaneously to accurately access total blood flow [Kontos, 1989, Shih et al., 2009].

Here, we review general techniques for animal preparation and measurement of blood flow with two-photon microscopy. We incorporate extensions to existing methods to accurately acquire flux data simultaneously across multiple vessels in a single trial. Central to these measurements is the ability to generate scan paths that smoothly connect user-defined lines of interest while maintaining high accuracy of the scan path.

2.3 Protocol - Cranial Window Preparations in the Rat or Two-Photon Imaging of Blood Flow in the Cordex

A craniotomy is made above the cortical area of interest; see [Mostany and Portera-Cailliau, 2008] for a demonstration of the original procedure [Kleinfeld et al., 1998]. In our example data, this corresponds to the somatosensory area of parietal cortex in rat. A custom made headframe is cemented to the skull to provide a means to secure the animal under the microscope (Fig. 2.1). As a means to protect the brain and prevent motion artifacts, the exposed region of cortex is covered with agarose prepared in an artificial cerebral spinal fluid and sealed with a microscope cover slip. Note that an alternative technique for mice makes use of a thinned and reinforced skull to avoid exposing the brain at the price of restricted optical access [Drew et al., 2010b]. All procedures must be performed in accordance with the relevant animal care guidelines.
**Figure 2.1**: Setup for in vivo imaging of blood flow through a cranial window. (A) To immobilize the head of the animal during imaging, we designed a metal frame that could be cemented to the skull, and then anchored to an optical setup. The frame is constructed of type 410 stainless steel with dimensions of 2.00 inches long, 0.61 inches wide, and 0.029 inches thick, and can be secured between two posts in a standard optical breadboard. An inset region, 0.015 inches deep, borders the frame window to hold a cover glass over the craniotomy. The window cover, 0.015 inches thick of type 301 stainless steel, is then secured to the frame with four screws, sandwiching a no. 1 cover glass in place. (B) A typical experimental setup. Note: The metal frame attached to the skull is immobilized between two anchoring posts inserted into an optical breadboard. (C) Diagram of cranial window preparation for rat. (D) A crosssectional view of the cranial window. (Inset) An inverted coronal view of surface vessels and deep microvessels that are targeted for occlusion.
2.3.1 Materials

CAUTION: See Appendix 6 for proper handling of materials marked with (!) See the end of the chapter for recipes for reagents marked with (R)

Reagents

- Dextrose, 5% (w/v) in saline
- Fluorescein-dextran (2 MDa, Sigma) or Texas Red-dextran (70 kDa, Invitrogen), 5% (w/v) in saline. Freeze the dextran-conjugated dye solution(s) in aliquots at −20°C for later use.
- Heparin, 20 U/mL in saline (!)
- Isoflurane anesthesia (for survival studies) (!)
- Lidocaine solution, 2% (v/v)
- Modified artificial cerebral spinal fluid (mACSF), free of carbonate and phosphate (R)
- Ophthalmic ointment
- Rose Bengal, 1% (w/v) in saline (filtered before use)
- Freeze the 1% Rose Bengal solution in aliquots at −20°C for later use.
- SurgiFoam (Edgepark)
- Urethane (!) or -cholalose anesthesia (for terminal studies)
- VetBond (3M)

Equipment

- Blood gas monitor (RapidLab 248; Bayer)
- Blood pressure monitor (BP1, World Precision Instruments for intra-arterial; XBP1000, Kent Scientific for tail cuff measurements)
- Catheter (Surflo; Terumo)
- Cover Glass, no. 1 thickness
- Dental acrylic (Grip Cement; Dentsply)
- Dental drill, air-powered (Silent Air)
- Drill burrs, 0.5- and 0.25-mm tip sizes (Henry Schein)
- Forceps, extra-sharp (Dumont no. 55; Fine Science Tools)
- Glass cutter
- Head-frame, custom-made (Fig. 2.1A)
- Heat pad, feedback-regulated (Harvard for rats; FHC for mice)
- Hemostats
- Isoflurane vaporizer (IsoTec)
- Kimwipes
- Optical breadboard with head holder (Fig. 2.1B) and devices for physiological support and monitoring
- PE50 tubing, for femoral artery/vein catheters (Intramedic)
- Periosteal elevator (Roboz)
- Pulse oximeter (Nonin for rats; Starr Life Sciences for mice)
- Scalpel blades
- Screwdriver, miniature
- Screws, self-tapping (#000-3/32; Small Parts Inc.)
- Stereotaxic frame (Kopf)
- Syringe needle, 26-gauge
- Two-photon laser-scanning microscope - Our TPLSM is a custom-design optimized for in vivo studies [Nguyen et al., 2006, Nguyen et al., 2009, Tsai and Kleinfeld, 2009].
2.3.2 Experimental Method

Anesthetizing the Rat

1. Anesthetize the rat. Common anesthetic choices include (1) urethane delivered intraperitoneally, that is, 1000 mg/kg body weight initial dose with 100 mg/kg supplements as required [Kleinfeld et al., 2000], or (2) initial isoflurane for surgery followed by a transition to intravenous delivery of $\alpha$-chloralose for imaging, that is, an initial bolus at 50 mg/kg body weight for induction and continuous delivery of 40 mg/kg for maintenance (Devor et al. 2007). Both urethane and $\alpha$-chloralose should be prepared fresh on the day of the experiment. Heating to 60°C and agitation is necessary to dissolve $\alpha$-chloralose. These anesthetics are not suitable for survival experiments.

2. Check for lack of toe pinch reflex to ensure an adequate level of anesthesia.

3. Secure the rat in a stereotaxic frame.

4. Apply opthalmic ointment to its eyes to keep them moist.

5. Inject 0.1 mL of 2% (v/v) lidocaine subcutaneously into the scalp as a local anesthetic before incision (in Step 11).

Monitoring Requirements

The following procedures need to be performed throughout the surgery and during imaging.

6. Monitor heart and breathing rates with a pulse oximeter. They should remain within a normal range: 300–400 beats/min and 60–120 events/min.

7. Maintain body temperature at 37°C using a feedback-regulated rectal probe and heat pad.

8. Inject 5% (w/v) dextrose in saline intraperitoneally at 3 mL/kg body weight every 2 h to maintain body fluids and energy requirements.

9. Install a femoral artery catheter to collect blood samples and a femoral vein catheter to administer drugs. Monitor blood gases every 2 h.
10. Monitor arterial blood pressure continuously from the femoral arterial line. Alternatively, the tail cuff method can be used to noninvasively measure blood pressure at intermittent time points.

**Cranial Window Surgery**

11. Shave and make a 4–5 cm incision down the midline of the scalp. Use a periosteal elevator to remove the thin periosteum from the surface of the skull.

12. Demarcate the location of the cranial window. We typically consider measurements over primary somatosensory cortex, which is the part of parietal cortex that nominally lies between –1 and –5 mm relative to the Bregma point and between 1 and 7 mm from the midline on the mediolateral axis for rats [Paxinos and Watson, 1986].

13. Attach a custom metal frame to the skull with dental acrylic. The frame holds the head of the animal rigidly to the optical apparatus (Fig. 2.1A,B).

   (a) Clean the contact regions on the bone of soft tissue.
   (b) Apply a thin layer of VetBond.
   (c) Introduce small self-tapping screws into the anterior and posterior aspects of the skull. (Note that one of the screws passes through an opening in the frame.) Mechanically link the screws to the frame with dental cement (Fig. 2.1C).
   (d) Secure the threads of the screws to the bone with a small dab of VetBond. Note that the temporalis muscle may need to be retracted in some cases, and the temporal ridge may need to be flattened with a dental drill. This is to ensure that the metal head frame is cemented tangential to the cortical surface when imaging the lateral aspects of the barrel cortex.

14. Perform a craniotomy above the brain region of interest using a high-speed drill.

   (a) Thin the skull throughout the entire window to approximately one-quarter of its original thickness using a 0.5-mm drill burr until the underlying pial vasculature becomes visible following application of mACSF. During drilling,
flush the window regularly with mACSF to reduce heat buildup and to remove blood and bone shavings.

(b) Carefully thin the edges of the window with a 0.25-mm burr until the bone begins to craze.

(c) Use forceps to gently separate the bone flap from the skull without protruding too deeply.

(d) Use two forceps to grasp adjacent corners of the loosened bone flap, and slowly peel it away from the underlying dura mater.

15. Reflect the dura to the edges of the window (Fig. 2.1D).

(a) Make a small incision using the cutting edge of a 26-gauge syringe needle. Bend the needle to an obtuse angle with hemostats to ensure that the cutting edge approaches the dura at a suitable angle.

(b) Use two sharp no. 55 forceps to gently lift the dura away from the cortical surface, starting at the incision site, and tearing in small increments. Whenever possible, tear around large dural vessels to avoid bleeding. Limit any dural bleeding with small pieces of Surgifoam soaked in mACSF, and Kimwipes twisted to a fine point with the fingers. Keep the cranial window moist with a piece of moist Surgifoam.

(c) Retract the dural flaps to the side of the window, and flush the cortical surface with mACSF. It is crucial to avoid any damage to pial vessels. Hemorrhaging will alter cerebral blood flow, accelerate edema, and severely degrade imaging quality.

16. Fill the interior of the chamber with 1.5% (w/v) low-melting-point agarose dissolved in mACSF (Fig. 2.1D) (Kleinfeld and Delaney 1996). Dissolve the agarose by heating it in a microwave. The temperature of the agarose must not exceed 37°C when it is applied to the brain.

17. Immediately seal the chamber using a cover glass as a window (Fig. 2.1D). Resealing the craniotomy is crucial to protect the cortex and suppress motion from cranial-pressure fluctuations due to heartbeat and breathing. One edge of the window can remain uncovered to allow insertion of electrodes or micropipettes.
18. Suture the skin together around the frame, and trail agarose around the cover glass to hold water for the dipping lens.

19. Stabilize the animal on an optical breadboard for imaging, using the frame as a head support. Our separate plate can be transported between surgical and imaging suites with the animal and all physiological monitoring devices assembled as one unit (Fig. 2.1B).

2.4 Example Application

A cranial window was generated over the rat somatosensory cortex (Fig. 2.2). Intrinsic optical imaging was used to determine the locations of the hindlimb and forelimb cortical representation [Drew and Feldman, 2009]. The cerebral vasculature of the hindlimb somatosensory region was mapped using a TPLSM image stack obtained with a low magnification 4x objective (Fig. 2.2A). A single penetrating arteriole and neighboring ascending venule were selected for measurement under a 40x objective (Fig. 2.2B).

A scan path was created to traverse along the length of the lumen center, and across the lumen width for each vessel. The accuracy of the scan path can be verified by comparing the value of the mirror position encoder with the control voltage. Scan accuracy in user-selected linear scan paths is $\sim 1\mu m$, whereas intervening segments that connect the user-defined regions and are not used in data analysis have a higher error, up to $\sim 1\mu m$ (Fig. 2.2C). The resulting line scan is a space-time image (Fig. 2.2D). Portions of the scan path along the vessel length appear as streaks within the line-scan image. These streaks represent nonfluorescent RBCs that move through a fluorescent background. The x-axis represents the distance traveled by the RBCs, and the y-axis is the time. The centerline velocity is then calculated from the slope of the RBC streaks (Fig. 2.2E, right panel). Previously, singular value decomposition was used to find the slope of these streaks and calculate the RBC velocity (Kleinfeld et al. 1998; Schaffer et al. 2006). A faster method that can calculate the velocity in near real time makes use of the Radon transform [Drew et al., 2010a]. In either case, the velocity is taken using windowed portions of the line scans, typically 25 msec of data for a Nyquist frequency of 20 Hzp.

Portions of the scan path across the vessel width capture the diameter of the vascular lumen, because the fluorescently labeled blood plasma provides high contrast
with the unlabeled region immediately outside the vessel. To increase the signal-to-noise ratio, multiple scans across the vessel diameter are averaged together before the profile is calculated. As with the velocity calculation, the diameter transform is taken on a windowed portion of the data. Typically, the same window size as for velocity is used so that both parameters can be calculated on the same timescale. Vessel diameter is calculated as full-width at half-max (FWHM) of the vessel profile (Fig. 2.2E, left). Since the intensity profile tends decrease in the center of the vessel, where a large volume of RBCs exclude the fluorescently labeled plasma, the two outermost half-max points are used. Linear interpolation is used to add subpixel accuracy to the diameter measurement.
**Figure 2.2:** Simultaneous measurement of diameter and velocity in multiple vessels using spatially optimized linescans. (A) Image of fluorescently stained vessels in somatosensory cortex of a Sprague Dawley rat. The forelimb and hindlimb representations across cortex were mapped using intrinsic optical imaging. (B) Image of a surface arteriole and venule, with scan pattern superimposed. Portions of the scan path along the length are used to calculate red blood cell (RBC) velocity, while portions across the diameter of the vessels are used to calculate diameter. Scans were acquired at a rate of 735 lines per second. (C) Scan path, colored to show the error between the desired scan path and the actual path the mirrors traversed. The error along linear portions of the image is $\sim 1\,\mu m$, and increases when the mirrors undergo rapid acceleration. The error between successive scans of the same path is $< 0.15\,\mu m$, several times lower than the point-spread function of a TPLSM. (D) Mirror speed as a function of time. Note that portions used to acquire diameter and velocity data are constant speed (top). The line scans generated from the path can be stacked sequentially as a function of time to produce a raw cascade image (bottom). (E) Vessel diameter is calculated as the full width at half-maximum of a time average of several scans across the width of a vessel (left). RBC velocity calculated from the angle of the RBC streaks. (F) Data traces of diameter, velocity, and flux for the arteriole and venule, processed to remove heart rate and smoothed with a running window. Both vessels show an increase in flux in response to forelimb stimulation. In the arteriole, this flux increase is due to simultaneous increase of lumen diameter and RBC velocity. In contrast, flux increase in the venule is due only to an increase in RBC velocity, as diameter is unchanged by stimulation.
We now consider combining the velocity and diameter measurements to calculate flux. In laminar flow, the velocity $v(r)$ of the fluid through the pipe with radius $R$ measured at a distance $r$ from the center has a quadratic profile, with a maximal velocity $v(r)$ in the center and tapering to 0 at the edges, that is,

$$v(r) = v(0) \left(1 - \frac{r^2}{R^2}\right)$$  \hspace{1cm} (2.1)$$

Measurements of RBC velocity as a function of distance from the vessel center show that this model is a close approximation [Schaffer et al., 2006], although systematic deviations result from the nonnegligible size of the red blood cells. This tends to flatten the profile, and prevents the velocity from reaching 0 at the edges of the vessel. Neglecting these effects, the integrated flux through the vessel is given by [Helmchen and Kleinfeld,
In our example, both vessel diameter and RBC velocity in the arteriole respond to somatotopic stimulation. The flux through the arteriole increases by 100% over baseline, compared with 35% and 40% for diameter and velocity measurements alone, respectively (Fig. 2.2F). Since the scan path runs at 733 Hz in this example, the diameter and RBC velocity traces are collected nearly simultaneously from both the penetrating arteriole and ascending venule.

2.5 Generation of Spatially Optimized Line Scans

The microscope laser is directed by a pair of fast x–y scan mirrors (Cambridge Technology, 6210H galvometer optical scanner with MicroMax 673xx dual-axis servo driver). The mirror controller employs a closed loop position feedback system to accurately map control voltage to mirror position. The position accurately tracks the control voltage, with $\sim 90\mu s$ delay.

To create an arbitrary scanning path, several full-frame images of the region containing the desired vessels are first acquired and averaged together to increase the signal-to-noise ratio. This image is loaded by custom software written in MATLAB (MathWorks), which allows users to interactively create lines of interest (i.e., scans across or along vessels) on the full frame image. The mirror voltages used to acquire the original full frame image are known, so that positions on the image can be mapped one-to-one to mirror voltages. Linear portions of the scan path, such as those used to track along a vessel to measure RBC speed and those that span the vessel to measure diameter, are given by

$$ P = P_0 + V_{\text{linear}}t $$

(2.3)

where $P$ and $V$ are two-dimensional vectors that contain $x$ and $y$ coordinates. Linear portions of the scan path, scanned at constant velocity, must be connected by a function that creates a physically realistic path for the mirrors to follow (Lillis et al. 2008). Although there are an arbitrary number of such functions, the simplest is a third-order polynomial given by

$$ F = \frac{1}{2} v(0) \pi R^2 $$

(2.2)
\[ P_{\text{spline}} = P_1 + V_1 t + C t^2 + D t^3 \] (2.4)

For computational convenience, the connecting spline is taken to start at \( t = 0 \), and end at \( t = \tau \). The initial these position \( P_1 \) and velocity \( V_1 \) of the spline are set to match the position and velocity of the endof the linear region preceding the spline. The additional spline parameters depend on the length (in time) of the spline, \( \tau \), and the final position and velocity of the spline, \( P_f \) and velocity \( V_f \), and are given by

\[ C = \frac{3P_f}{\tau^2} - \frac{3P_i}{\tau^2} - \frac{2V_i}{\tau} - \frac{2V_f}{\tau} \] (2.5)

and

\[ D = \frac{V_f}{3\tau^2} - \frac{V_i}{3\tau^2} - \frac{2C}{3\tau} \] (2.6)

The value of \( \tau \) used should be the smallest positive real value that does not subject the mirrors to an acceleration larger than a user-defined maximal value, \( m \), typically 100 V/msec\(^2\). Because the acceleration of the spline is a linear function of time, the extreme values of acceleration occur at the beginning and the end of the spline.

Candidate values for the shortest possible spline length are found by setting the mirror acceleration to \( m \) at the beginning at end of spline, and finding all positive real values for \( \tau \):

\[ 0 = \pm m\tau^2 + (2V_{i,x} + 4V_{f,x})\tau + (6P_{i,x} - 6P_{f,x}) \] (2.7)

\[ 0 = \pm m\tau^2 + (2V_{i,y} + 4V_{f,y})\tau + (6P_{i,y} - 6P_{f,y}) \] (2.8)

\[ 0 = \pm m\tau^2 + (4V_{i,x} + 2V_{f,x})\tau + (6P_{i,x} - 6P_{f,x}) \] (2.9)

\[ 0 = \pm m\tau^2 + (4V_{i,y} + 2V_{f,y})\tau + (6P_{i,y} - 6P_{f,y}) \] (2.10)

The actual time used is the smallest value that keeps the acceleration within limits at the beginning and end of the spline; that is, all of \( |2C_x| < m, |2C_y| < m, |2C_x + 6D_x\tau| < m, \) and \( |2C_y + 6D_y\tau| < m \) are true.
2.6 Summary

Two-photon laser scanning microscopy offers a means to obtain high-resolution images of RBC velocity and vessel diameter in vivo. These measurements can be combined to calculate the flux for a given vessel, which is a more accurate metric of the oxygen- and nutrient-carrying capability of blood.

Velocity and diameter measurements can change independently (Fig. 2.2F) and thus they must be measured nearly simultaneously to accurately access flux. This can be achieved with the spatially optimized line-scan algorithm described above. This technique can be extended to image other types of fluorescent signals, for instance, neural activity as indicated by calcium-sensitive dyes. In principle, this technique can be readily extended to scan in three dimensions as well. However, the speed of current mechanical z-axis scanners is currently much slower than what can be achieved when scanning in the x–y plane alone [Göbel et al., 2007, Göbel2007 and Helmchen, 2007].

2.7 Recipe

CAUTION: See Appendix 6 for proper handling of materials marked with (!). Recipes for reagents marked with (R) are included in this list. Modified Artificial Cerebral Spinal Fluid (mACSF), Free of Carbonate and Phosphate

1. NaCl 125 mM
2. Glucose 10 mM
3. HEPES 10 mM
4. CaCl2 (!) 3.1 mM
5. MgCl2 1.3 mM
6. Adjust the pH to 7.4 (from Kleinfeld and Delaney 1996).

2.8 Acknowledgements

We thank Quoc-Thang Nguyen for useful discussions on MPScope scan software. This work was funded by the NIH (grants EB003832, NS059832, and RR021907 to DK
and AG029681 to Gert Cauwenberghs), the NSF (grant DBI 0455027), and a postdoctoral fellowship from the American Heart Association to AYS.

3 Photon counting, censor corrections, and lifetime imaging for improved detection in two-photon microscopy

3.1 Abstract of Chapter

We present a high-speed photon counter for use with two-photon microscopy. Counting pulses of photocurrent, as opposed to analog integration, maximizes the signal-to-noise ratio so long as the uncertainty in the count does not exceed the gain-noise of the photodetector. Our system extends this improvement through an estimate of the count that corrects for the censored period after detection of an emission event. The same system can be rapidly reconfigured in software for fluorescence lifetime imaging, which we illustrate by distinguishing between two spectrally similar fluorophores in an in vivo model of microstroke.

3.2 Introduction

Two-photon laser scanning microscopy (TPLSM) [Denk and Svoboda, 1997, Denk et al., 1990, Denk et al., 1995, Helmchen and Denk, 2005, Yasuda et al., 2006] involves the periodic excitation of molecules by hundred-femtosecond laser pulses, followed by the emission of fluorescent light before the arrival of the next pulse. Despite the relatively large number of photons in each excitation pulse, roughly 100 million at the focus of the objective for nonsaturating intensities, the average number of fluorescence photons that
are detected is typically of order one or less [Neher and Hopt, 2001, Tsai and Kleinfeld, 2009]. Thus it is imperative to optimize all aspects of TPLSM, from ensuring that the excitation beam is free from temporal and spatial aberrations [Cui, 2011, Field et al., 2010a, Ji et al., 2010, Rueckel et al., 2006, Tsai et al., 2007] to the collection and detection of as many of the emitted photons as possible [Engelbrecht et al., 2008].

All measurements of light contain shot noise, which reflects the discrete nature of photons and limits the signal-to-noise ratio (SNR) for a pixel with an average count of $N$ photons to $\text{SNR} = \sqrt{N}$. Ideally, the photodetection process will not add to this noise. Photomultiplier tubes (PMTs) are universally used to convert the arrival of a photon into a pulse of current, which is then amplified, filtered or integrated, and digitized to generate a pattern of the intensity of fluorescent emission as a function of the position of the focus of the excitation beam. This pattern forms the optically sectioned planar image reported with TPLSM [Denk et al., 1990]. Photomultiplier tubes offer a unique combination of good quantum efficiency, high conversion gain, high speed, and linearity of the output current with incident intensity. However, the serial amplification of photocurrents along the dynode chain in a PMT results in a variation in the amplitude of the individual output pulses [Zworykin et al., 1936] that multiplies the variance of the shot effect [Shockley and Pierce, 1938]. The exact value of this additional noise depends on details of the PMT but is typically 20-70%.

Photon counting is commonly used as a means to circumvent variation in the intensity of the output pulses from the PMT. This technique is based on the assumption that current pulses that exceed a threshold amplitude are the result of photoelectrons generated by absorption of a single photon at the cathode of the PMT, as opposed to a thermally generated electrons. Photon counting devices employ a discriminator to convert analog pulses from the PMT into digital pulses, which are then summed. Since each photon will result in exactly one count, this eliminates a source of variability in the detection process. Photon counting is preferred so long as the rate of emitted photons is low, so that each emitted photon is counted before a second photon arrives. Although the chance of two photons arriving at the PMT at exactly the same instance is nil, dispersion of the electron current in the PMT, the capacitance of cables, and the response time of amplifiers leads to a photocurrent pulse with a width of one to a few nanoseconds. This creates a censored period after a photon arrives that briefly blocks the counting of subsequent photons. Uncertainty in the number of uncounted events constitutes a
source of variability. This uncertainty is negligible when the rate of fluorescent emission is low but will exceed the variability associated with the amplitude of the photocurrent pulses when the rate of emission is high. The best SNR is thus obtained by switching from photon counting to analog integration when the dominant noise source changes from analog variability to uncertainty in the count. However, algorithms to estimate the number of missed counts, as well as to identify the exact count at which the switch from counting to analog detection should occur, do not appear in past literature.

In this report we describe a reconfigurable photon counting system that achieves a resolution of 3.2 ns between successive pulses with a temporal accuracy of 1.5 ns per photon. It is used in conjunction with a maximum likelihood estimate of the number of censored pulses as a means to increase the dynamic range of counting. At the core of the system is a high-speed complex programmable logic device (CPLD) that operates together with commercially available components to transfer pulses to a master computer at the typical 0.5- to 2.5-MHz rates of pixel acquisition. Although photon counting has been previously utilized for TPLSM [Amir et al., 2007, Benninger et al., 2008, Buehler et al., 2005, Field et al., 2010b, Sheetz et al., 2008, Vučinić and Sejnowski, 2007], previous systems did not transfer the measured counts at these clock rates. Furthermore, the flexibility of our counter allows it to be rapidly reconfigured for fluorescent lifetime imaging (FLIM) [Berezin and Achilefu, 2010, Harvey et al., 2008, König et al., 1996, Yasuda et al., 2006] and potentially other modes of nonlinear imaging.

### 3.3 Methods

Optomechanics. All data were obtained with a custom-built two-photon microscope, as described previously [Tsai and Kleinfeld, 2009], and a tunable Ti:Al2O3 pulsed laser (Mira 900; Coherent) that operated at 75.5 MHz. We used x20, 1.0 numeric aperture (NA) (Zeiss) and x40, 0.8 NA (Olympus) dipping objectives. The microscope was controlled with MPScope 2.0 software (Nguyen et al. 2006, 2009), which synchronized data acquisition, both analog and digital, with bidirectional scanning of galvanometric mirrors.

Electronics. Signal acquisition begins with a PMT (H7422-40; Hamamatsu) that converts photons into brief current pulses with a conversion gain of $\sim 105$. For imaging by analog integration, the PMT output is switched (ZSWA-4-30DR; Mini-Circuits) to a current-to-voltage converter and a direct current-coupled amplifier, both with a 0.3 $\mu$s time constant, low-pass filtered with a 3-pole Bessel filter (model 3202; Kronhn-Hite) set
so that the -3 dB cutoff frequency of the filter, denoted \( f_c \), is equal to \( \sqrt{\frac{\ln(2) - \pi}{\ln(2) + \pi}} \), where \( f_{pixel} \) is the frequency of the pixel clock [Tan et al., 1999], and then digitized by an analog-to-digital board (PCI-6110; National Instruments), where it is sampled at the pixel clock rate (Fig. 3.1).

For imaging by photon counting, the output of the PMT is directed through a high-speed alternating current-coupled amplifier (ZPUL30P; Mini-Circuits), and the signal is converted to a transistor-to-transistor logic (TTL) pulse with a discriminator, either a commercial device (PRL-350TTL; Pulse Research Lab) or one of local design (Fig. 3.5)

The threshold is set above the thermal voltage noise and typically at 0.2 times the measure height of the current pulse. All lines are terminated at 50Ω. The TTL pulses are counted using a custom photon counting board, implemented with a CPLD (ispMach LC4256V; Lattice). The counts are read from the board at the pixel clock rate by a digital interface board (DIO-32HS, National Instruments) on a Windows-based computer. The firmware for the CPLD is written in ispLevel Classic (Lattice Semiconductor) and is flashed to the CPLD board through a standard parallel port interface. The CPLD implements a high-speed dual ripple counter (Fig. 3.6), in which the active counter alternates at each imaging pixel. This allows time for the counters to settle before being read by the computer interface board, which inputs 32 bits in parallel from the CPLD counters at the pixel clock rate.

For FLIM, a timing signal from the laser excitation pulse must be passed to the CPLD counting board. We split a small portion of the pulsed laser excitation light onto a high-speed photodiode (DET10A; Thor Labs). This signal is amplified (ZJL-4G; Mini-Circuits) and divided by a passive coupling device (ZFDC-10-5; Mini-Circuits). The larger portion of the signal is piped to a discriminator, where it is converted to a TTL logic level signal, while the weaker signal is directed to an oscilloscope to monitor the mode-locking of the excitation laser. Fluorescence lifetime data for each pixel are acquired by sequentially enabling a series of six separate counters on the CPLD device. These counters are gated so that the first counter is enabled for 0-1.5 ns after the initial laser excitation pulse, the second counter is enabled for 1.5-3 ns after the laser excitation pulse, and so forth. The end result is a six-bin histogram of arrival times for fluorescence photons from which the exponential decay time of the fluorophore, \( \tau \), can be calculated. The implementation of multiple counters for FLIM, acquisition of specific number of
Figure 3.1: Fig. 1. Photon counting system. Single photons arriving at the photomultiplier tube (PMT) are converted to discrete current pulses. For analog integration, the signal is routed to a custom-built direct current-coupled amplifier, filtered using a Bessel filter, and passed to the analog input of the data acquisition card. For photon counting, the signal from the PMT is sent to a wide-band amplifier and then into a discriminator to convert the pulses to a transistor-to-transistor logic (TTL)-level logic signal. These pulses are read by a custom-built board containing a complex programmable logic device (CPLD), which implements a dual counter scheme for each channel and sends the output to a digital interface card on the computer. In addition, a portion of the laser excitation beam is split off with a glass slide and sent to a high-speed photodiode. The output from the diode is amplified with a high-speed amplifier and split into 2 signals using a passive transformer-based splitter. One of the outputs from the coupler is sent to an oscilloscope, and the other is converted to TTL-level pulses with an additional discriminator channel, sent through an adjustable fine-scale analog delay line, and routed to the CPLD for use with fluorescence lifetime imaging microscopy (FLIM). I/V converter, current-to-voltage converter; BNC, Bayonet Neill-Concelman connector.
counts per pixel clock, and simultaneous data acquisition by two or more PMTs is detailed in 3.7 and Figs. 3.6–3.8.

The timing signal for the laser excitation pulse must be aligned to coincide with the start of the fluorescence excitation. We thus delay the signal from the excitation pulse to maximize fluorescence counts acquired in the first time bin. A coarse delay line, with a resolution of 1.5 ns, is implemented in the firmware of the CPLD (Fig. 3.7A). A fine-grained delay is achieved with a custom analog delay line that was constructed by switching among 12 lengths of impedance-matched (50 Ω) wire that incremented the delay from 0 to 1.65 ns in increments of 0.15 ns.

In vivo physiology. Our subjects were C57Bl6 mice and green fluorescent protein-expressing inhibitory interneuron (GIN) mice that ranged in mass from 20 to 30 g. All procedures were performed in anesthetized animals. Anesthesia was induced with 4% (vol/vol) isoflurane in 30% oxygen and 70% nitrous oxide and was maintained with 1-2% (vol/vol) isoflurane. The care and experimental manipulation of our mice were reviewed and approved by the Institutional Animal Care and Use Committee at the University of California at San Diego.

In the GIN animals, a craniotomy was performed to generate a cranial window with the dura left intact [Helmchen and Kleinfeld, 2008, Holtmaat et al., 2009]. Individual neurons were imaged as described [Drew et al., 2010b]. The C57Bl6 mice were used to form occlusions to single penetrating arterioles, as described previously [Blinder et al., 2010, Nishimura et al., 2007], with some modifications for mice. The pial vasculature was imaged through a thinned skull preparation reinforced with a glass coverslip [Drew et al., 2010b] check. The vasculature was labeled by injecting 50 µl of 2.5% (wt/vol) Texas red-dextran (70 kDa; Invitrogen) in phosphate-buffered saline (PBS) through the infraorbital vein. Targeted photothermogenesis was initiated by illuminating a target vessel with focused 530-nm laser light immediately following the injection of 50 µl of 1.25% (wt/vol) rose Bengal in PBS, also administered through the infraorbital vein. After confirmation of single vessel occlusion, the mice were allowed to survive for 2 days before reimaging. The vasculature was relabeled with another injection of Texas red-dextran. The nuclei of dying, membrane-disrupted cells were labeled by a single intraperitoneal injection of 1 mg/kg propidium iodide in PBS, as described previously [Cevik and Dalkara, 2003].

Immunohistology. Mice were injected with a fatal dose of pentobarbital and
transcardially perfused with PBS followed by 4% (wt/vol) paraformaldehyde in PBS. The brains were removed and cryoprotected in 30% (wt/vol) sucrose overnight. The forebrain was cryosectioned coronally at a thickness of 50 µm, and immunohistochemistry was performed on free-floating sections. Tissues were stained with primary antibodies for microtubule-associated protein-2 (MAP-2; no. AB5622, Chemicon) and secondary antirabbit Alexa 488 antibodies. Images were obtained using a Fluoview 1000 confocal microscope with an oil-immersion x63 objective.

3.4 Results

At low signal intensities, the average fluorescence that follows each laser pulse will contain, on average, significantly less than one photon. The signal must be integrated across multiple pulses, and the background electronic noise from the PMT, as well as from associated analog electronics, will reduce the ability to distinguish image features from the background. As an example, we used in vivo TPLSM [Svoboda et al., 1997] to image neurons in parietal cortex of mice that express green fluorescent protein in somatostatin-positive inhibitory cells (Bozza et al. 2004). The same field was imaged, using the same rate for the pixel clock, with our photon counting scheme (Fig. 3.2A) and with traditional analog integration (Fig. 3.2B). In general, the lower noise in the counting image reveals processes and details that are not seen in the analog images (n = 5 fields). A histogram of the counts and normalized intensity shows a clearly perceptible hump in the digital count (star in Fig. 3.2C), consistent with a non-Poisson distribution of the signal photons.

The scaling of excess noise with analog integration versus photon counting was directly tested by obtaining equivalent images of a sample, fluorescein-labeled cellulose that had a broad spatial distribution of emission intensities (Fig. 3.2D,E). The mean count per pixel is βr, where r is the number of laser pulses in the period of the pixel clock, i.e., r ≡ \( \frac{f_{\text{pulses}}}{f_{\text{pixel}}} \) ∼ 100, and β is the measured number of emissions per laser pulse (3.1). The variance for the photon counting data, denoted \( \sigma^2 \), scales identically with the mean count per pixel, i.e., \( \sigma^2 = \beta r \). In contrast, the analog signal showed excess variance compared with the Poisson limit (Fig. 3.2F). The data were well described by the phenomenological fit of \( \sigma^2 = (1 + \eta)\beta r \), with \( \eta = 0.57 \) for our conditions.

Correction to the photon count and variance. We analyzed the effect of near-simultaneous arrival of multiple photons, or photon pileup [Coates, 1968], that can
Figure 3.2: Comparison of digital and analog images under conditions of low emission rates. A: photon counting image of labeled neurons in a green fluorescent protein-expressing inhibitory interneuron (GIN) mouse at low signal level: single frame, pixel clock of 5 μs ($r = 383$ laser pulses per pixel), $\sim 125 \mu m$ below the pia, and $\sim 30 mW$ of incident laser power as measured through the objective at $\lambda = 900 nm$. B: image of the same field in A taken via analog integration using the same pixel clock. For comparison of the analog and digital images, the baseline amplitude has been subtracted from the analog image and the mean value of the intensity scaled to match the mean values from the photon counting data. Note that the fine processes are barely visible compared with the digital image in A. C: histogram for the images in A and B using a binwidth of 0.25 counts. Only discrete values, i.e., 0, 1, 2, . . . , are recorded with photon counting. The star indicates excess counts that correspond to signal photons. D: photon counting image of cellulose stained with rhodamine B at low signal level: 200-frame average, pixel clock of 3 μs ($r = 230$), and $\sim 6 mW$ of incident laser power at $\lambda = 800 nm$. E: image of the same field in D taken via analog integration using the same frame averaging and pixel clock. F: plot of the variance for the photon counting data, with the Poisson fit of $\sigma^2 = \beta r$, along with the variance for the analog data with a phenomenological fit of $\sigma^2 = \beta r$. 
occur when the lifetime of the fluorophore is comparable to the time resolution of the detection electronics. This is the typical case with genetically expressible indicators and most biologically relevant organic dyes, but not for quantum dots [Berezin and Achilefu, 2010, Shaner et al., 2005]. At modest to high imaging intensities, a single laser excitation pulse can result in two or more signal photons arriving at the PMT within the pulse-pair resolution of a single channel of the photon counter, denoted $\Delta$, where $\Delta = 3.2 \text{ns}$ for our system. Thus pairs or n-tuples of photons are counted as a single event, because all but the first photon are censored for a period $\Delta$. This phenomenon leads to the loss of information as to the exact number signal photons per laser excitation pulse.

How can we estimate the true count, $\alpha$, from the measured count, $\beta$? In principle, there is a monotonic relation between these quantities. However, as an approximation, one can estimate $\alpha$ from the measured distribution of counts with zero, one, two, etc., photons per laser pulse. We consider first the case of using only the number of samples with zero events, i.e., laser pulses that did not lead to a photocurrent, denoted $r_0$. A maximum likelihood estimate (MLE) yields an estimate, denoted $\hat{\alpha}$, in closed form (Supplementary Information on Count Estimate), i.e.,

$$\hat{\alpha} = -\ln \left( \frac{r_0}{r} \right)$$

(3.1)

and

$$\sigma^2 = \frac{r - r_0}{r \cdot r_0}$$

(3.2)

The value of $\hat{\alpha}$ for each pixel forms the image, and the ratio $\frac{\hat{\alpha}}{\sigma}$ forms the MLE of the SNR for that pixel.

An improved MLE for $\hat{\alpha}$ is formed by including $r_1$, the number of laser pulses that lead to exactly one measured count. This calculation incorporates the possibility that additional photons may have arrived in the censored period and yields an expression for $\hat{\alpha}$ that must be evaluated numerically (Supplementary Information of Count Estimate). We use the Newton-Raphson method to find the value of $alpha$ that maximizes the log-likelihood function $l(\alpha)$, where

$$l(\alpha) = r_0 \ln[p_0(\alpha) + r_1 \ln[p_1(\alpha)]] + (r - r_0 - r_1) \ln[1 - p_0(\alpha) - p_1(\alpha)]$$

(3.3)

with
\[ p_0(\alpha) = e^{-\alpha} \] (3.4)

and

\[ p_1 = \frac{e^{-\alpha} e^{-\frac{\Delta}{\tau}}}{1 - e^{-\frac{\Delta}{\tau}}} \] (3.5)

The variance is found by numerically computing the second derivative of \( l(\alpha) \) with respect to \( \alpha \), evaluated at \( \alpha = \hat{\alpha} \) (Eq. 3.3) so that

\[ \sigma^2 = -\left( \frac{\partial^2 l(\alpha)}{\partial \alpha^2} \bigg|_{\alpha=\hat{\alpha}} \right)^{-1} \] (3.6)

Note that the above expressions (Eqs. 3.1–3.6) can further include corrections for emissions that fail to occur with the time between laser pulses (Supplementary Information on Count Estimate). These emissions are not lost but can be counted as originating from a subsequent laser pulse.

We applied the above relations to images taken under relatively high laser intensities, for which the average count per laser pulse was well above one for bright regions of the image. We compared the image found with uncorrected counts (Fig. 3.3A), the image and SNR using the estimate of \( \hat{\alpha} \) at each pixel with just the number of pulses with no detected emission (Fig. 3.3B1 using Eq. 3.1 and Fig. 3.3B2) using Eq. 3.2), the image and SNR using the estimate of \( \hat{\alpha} \) at each pixel with both the number of pulses with zero and exactly one detected emission (Fig. 3.3C1 using Eqs. 3.3–3.5 and Fig. 3.3C2 using Eq. 3.6, and, last, the image obtained with analog detection (Fig. 3.1,D1,D2). These data illustrate three keys points. First, the dynamic range of the uncorrected image was compressed compared with the corrected images (cf. A with B1 and C1 in Fig. 3.3) and compared with the analog image (cf. A with D1 in Fig. 3.3). This occurs from the systematic loss of counts the high emission rates associated with bright pixels. Second, the SNR increases substantially when both \( r_1 \) and \( r_0 \) are used to estimate \( \hat{\alpha} \) for each pixel, compared with just \( r_0 \) (cf. B2 with C2 in Fig. 3.3). Last, the SNR when both \( r_1 \) and \( r_0 \) are used to estimate \( \hat{\alpha} \) for each pixel approaches that of analog recording at these high light levels, for which there is substantial censoring.

Crossover from photon counting to analog integration. At lower imaging intensities, photon counting offers an improved SNR compared with analog integration (Fig. 3.2F and Fig. 3.3E). At high intensities, however, the uncounted photons in the censored
Figure 3.3: Fig. 3. Measured and estimated true photon counts and signal-to-noise ratios (SNR) for imaging under conditions of high emission rates. The image is cellulose paper stained with rhodamine 6G, for which $\tau$, the exponential decay time of the fluorophore, is 4.1 ns. We used a pixel clock of $1.35\mu s$ ($r = 102$) and 5.6 mW of incident laser power, with $\lambda = 840\text{nm}$. We averaged 200 frames per image. A: plot of frame-averaged intensity, $\beta r$, for the raw counts. B: the image (B1) and SNR (B2) calculated from the number of samples with zero counts (Eqs. 3.1 and 3.2) on a pixel-by-pixel basis. C: the image (C1) and SNR (C2) calculated from the number of samples with both zero and one count (Eqs. 3.3–3.6) on a pixel-by-pixel basis. D: the image formed using analog integration with a low-pass filter frequency of 448 kHz (D1) and the SNR found by direct calculation of the frame-to-frame variability for each pixel (D2). The data were scaled so that the intensity of the averaged analog image matched that of the corrected digital image. E: plot of the variance as a function of the counts using data similar to that in B1, C1, and D1, but with an excitation power of 7.0 mW to increase the emission rate by 55%. The arrows mark the number of counts per pixel where analog integration will have a lower noise than photon counting together with the maximum likelihood estimate (MLE) for $\hat{\alpha}$. 
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
<th>Typical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{\text{pulses}}$</td>
<td>Frequency of laser pulses</td>
<td>75.5 MHz</td>
</tr>
<tr>
<td>$f_{\text{pixel}}$</td>
<td>Frequency of pixel clock</td>
<td>0.5-2.5 MHz</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Fluorescent lifetime of dye</td>
<td>$\sim 4\text{ns}$ for fluorescein</td>
</tr>
<tr>
<td>$\Delta$</td>
<td>Censored period or pulse-pair resolution</td>
<td>3.1 ns</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>True, mean number of photon counts per laser pulse</td>
<td>&lt; 3</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Measured, mean number of photon counts per laser pulse</td>
<td></td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>MLE number of photon counts per laser pulse</td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>Sample size or laser pulses per pixel clock, $f_{\text{pulses}}/f_{\text{pixel}}$</td>
<td></td>
</tr>
<tr>
<td>$r_0$</td>
<td>Number of samples with a count of zero per laser pulse</td>
<td></td>
</tr>
<tr>
<td>$r_1$</td>
<td>Number of samples with a count of exactly one per pulse</td>
<td></td>
</tr>
</tbody>
</table>

region result in additional variance that degrades the SNR of photon counting. There is a crossover point, above which analog integration yields a higher SNR for bright pixels and is preferable. In the example of (Fig. 3.3), obtained with relatively high incident intensity, the crossover point is $\hat{\alpha} = 0.8$ when only the number count for no emissions is used, but $\hat{\alpha} = 1.5$ when both the number of events with zero count and one count are used (Fig. 3.3E). There is a 0.45 chance of emitting two or more photons per laser pulse at the intensity that corresponds to $\hat{\alpha} = 1.5$, clearly demonstrating the utility of our approach to correct for censored counts.

In vivo detection of neuronal cell death using FLIM. The distribution of time between an excitation pulse and the returned fluorescence photon is a means to determine the lifetime of the fluorophores [de Grauw and Gerritsen, 2001, Suhling et al., 2005, Wang et al., 1990]. Thus FLIM can be exploited to separate multiple fluorophores with overlapping emission spectra but different fluorescence lifetimes within the same biological sample [Steinkamp et al., 1999]. We use FLIM to distinguish between two red emission dyes, Texas-red dextran and propidium iodide, in a model of microstrokes in which single penetrating arterioles are blocked by targeted photoexcitation [Blinder et al., 2010, Nishimura et al., 2007]. The single occlusion leads to a cessation of blood flow in a 500-µm-diameter neighborhood of the block and the eventual formation of a cyst. Both
dyes are essential to the assessment of neuronal damage. Texas red-dextran enables deeper vascular labeling than related molecules with green emission, such as fluorescein-dextran, and is used to locate the margin of the stroke. Propidium iodide is commonly used as a cell death indicator, because it preferentially stain cells with membranes disrupted by necrosis or apoptosis. Past studies have shown that propidium iodide can be used as a histological assay of cell damage when administered as an intravenous injection [Cevik and Dalkara, 2003]. We now extend this technique to image cell death in vivo.

Two-photon imaging based solely on intensity, or counts, yields individually labeled cells and vessels that are indistinguishable on the basis of brightness (Fig. 3.4A). Imaging based on lifetime clearly distinguishes between the two structures (Fig. 3.4B), consistent with the 4.1-ns lifetime of Texas red [Brismar et al., 1995] and the 6- to 12-ns lifetime of propidium iodide when bound to DNA [Hanley et al., 2002]. The merged counts and lifetime clearly highlights dying cells (red cells; arrowhead in Fig. 3.4C) and filled microvasculature (blue vessel; arrow in Fig. 3.4C). Post hoc histological analysis of this region shows that propidium iodide-labeled neurons are most prominent within the border of the microinfarction (Fig. 3.4D) and that labeled cells appear to be primarily neurons, based on immunohistochemistry with the pan-neuronal marker MAP-2 (Fig. 3.4E).

3.5 Discussion

Photon counting (Fig. 3.1) provides a simple way of increasing the SNR of images obtained with two-photon microscopy for low and medium imaging intensities (Fig. 3.2). At higher intensities, the increased SNR is offset by the arrival of multiple photons per laser pulse. The undercounting inherent in this process results in an additional noise source. Our analysis shows how to correct for undercounting in a manner that essentially doubles the count at which one needs to switch to analog integration (Eqs. 3.3–3.6). In fact, the correction for censored counts is sufficiently good to obviate
**Figure 3.4**: Fig. 4. Fluorescence intensity and lifetime imaging of spectrally similar dyes in a mouse model for experimental stroke. A: an intensity image collected with in vivo two-photon laser scanning microscopy (TPLSM) within infarcted tissue shows propidium iodide (PI)-labeled cells and Texas red-dextran (TR)-labeled vessels, which are indistinguishable by intensity alone. Images were acquired 140 μm below the pia, using 30 mW of incident laser power, at \( \lambda = 825 \text{ nm} \). Images were acquired at 512 X 512 resolution and then binned 3 X 3 for analysis and median filtered. We averaged 132 frames with \( r = 68 \) laser pulses per pixel. B: a fluorescence lifetime image demonstrating that labeled vessels and the cells can be distinguished on the basis of lifetime of the different fluorophores. C: a combined fluorescence lifetime and intensity image, with the color indicating lifetime and the saturation indicating number of photon counts, enables vessels (arrow) and cells (arrowhead) to be clearly distinguished. D: coronal section showing infarcted cortical region 2 days following occlusion of a single penetrating arteriole. Note that extravasation of the TR dye and PI labeling of necrotic cells is prominent within the infarct. An asterisk labels the approximate location of in vivo imaging. E: magnified view of the infarct in D showing strong microtubule-associated protein-2 (MAP-2) staining (green) of viable neurons within the stroke margin (open arrowheads). Within the stroke, intensity of MAP-2 labeling decreases and damaged neurons exhibit PI labeling of nuclei (arrowheads). Extravasation of TR (arrows) is more prominent within the infarcted region.
the need for analog detection, even at high emission rates (Fig. 3.3).

Theoretical analysis of the shot-to-shot variation in output pulses from the PMT predicts that the analog signal should have a multiplicative factor of \( \eta = (1 - g^{-d+1})/g - 1 \) excess noise power over the Poisson limit (Shockley and Pierce 1938), where \( g \) is the gain per stage in the PMT at a particular supply voltage and \( d \) is the number of stages. For \( g = 3.5 \) and \( d = 9 \), parameters for the operating point of our PMT in these studies, the theoretical value of \( \eta = 0.4 \) falls short of the observed value of \( \eta = 0.6 \). The reason for this excess is unclear, yet the implication is that photon counting is effective at higher intensities than suggested by theory.

Our system allows for uninterrupted counting at speeds up to 350 MHz with the use of relatively inexpensive commercially available parts. In addition, by creating the electronics around a reprogrammable CPLD core and by locking the counters to the onset of the excitation pulse, we are able to extend the photon counting device to acquire fluorescence lifetime images with 1.5-ns binwidths (Fig. 3.4). Locking of the counters to the laser pulse is also essential when alternate pulses are used to acquire different modalities, such as switching between focal depth or polarization [Carriles et al., 2009, Field et al., 2010b]. Other potentially useful configurations are possible with only a change in firmware, such as gating to isolate the extremely short-lived signals that are generated by second [Campagnola et al., 2001, Moreaux et al., 2000] and third [Débarre et al., 2006, 2007] harmonic process and coherent anti-Stokes Raman scattering [Ly et al., 2007].

Continuing advances in high-speed digital electronics and serial communication suggest that a next-generation photon counting system may make use of a field-programmable gate array with a USB 2.0 serial communication channel, rather than the current CPLD and parallel communication channel. Furthermore, advances in PMT design and high-speed signal conditioning should decrease the censored period below 3 ns.

### 3.5.1 Acknowledgments

We thank Earl M. Dolnick for involvement in the design of an analog integrator that predated this work, Quoc-Thang Nguyen for coding the first interface of MPScope with the photon counting device, Jing W. Wang for involvement in exploratory experiments, and Philbert S. Tsai for comments on the original manuscript, particularly with regard to the issue of censored counts.
The CPLD interface, as well as a variable time-delay and discriminator circuit, were developed by the University of California at San Diego (UCSD) Physics Electronic Shop. Confocal microscopy made use of the UCSD Cellular and Molecular Medicine Microscopy Shared Facility.

3.6 Supplementary Material - Estimating the Emission Count

For each experimental run, we assume that the number of photocurrent events, i.e., detected fluorescent emission events, $N$, is a Poisson random variable with mean $\alpha$. If $N = n \geq 1$, then let $W_1, \ldots, W_N$ be the times of the $N$ emissions; we assume that each $W_i$ is exponentially distributed and that the emission times are independent of the number of emissions. For all calculations below, our clock is in units of the fluorescent lifetime, $\tau$, of the emission times; thus, each $W_i$ has mean 1. Next, sort the emission times in increasing order to get the order statistics $W(1) < \cdots < W(n)$. In general, the joint distribution is a bit involved, but there is a convenient representation for order statistics from exponential random variables that facilitates many computations. In short, when $N = n$, start with independent and identically distributed (i.i.d.) exponential random variables $U = (U_1, \ldots, U_n)$, each with mean 1; then the joint distribution of $(W(1), W(2), \ldots, W(n))$ is the same as that of $V = (V_1, \ldots, V_n)$, where

$$V_k = \left[ \frac{U_1}{n} + \frac{U_2}{n-1} + \cdots + \frac{U_k}{n-k+1} \right], \text{ for } k = 1, \ldots, n, \quad (3.7)$$

[Cox and Hinkley, 1979]. Thus, calculations involving the order statistics can be recast as those involving $V$, which in turn are weighted linear combinations of the i.i.d. exponential random variables, $U$.

The distribution of an order statistic or the joint distribution of several order statistics is easy in principle, but the expressions are a bit involved. For example, suppose that $X_1, \ldots, X_n$ are i.i.d. random variables with continuous cumulative distribution function (cdf) $F$. Let $X_{(1)} < \cdots < X_{(n)}$ be the corresponding order statistics. Then, writing $C(n, j) = n!/[j!(n-j)!]$, we have

$$[X_{(k)} \leq x] = [\text{at least } k \text{ of the } X_i \text{ are at most } x],$$

so that

$$P(X_{(k)} \leq x) = \sum_{j=k}^{n} C(n, j) F(x)^j (1 - F(x))^{n-j},$$

and the probability density function (pdf) of $X_{(k)}$ is

$$f_{X_{(k)}}(t) = kC(n, k)f(x)F(x)^{k-1}(1 - F(x))^{n-k}. \quad (3.8)$$

For each experimental run, the observation period is $T$ and is normalized by the fluorescent lifetime, $\tau$. If an emission occurs after $T$, it is not observed in that run. If an
emission occurs before \( T \), there is a dead period of length \( \delta \), again in units normalized by \( \tau \), during which any emissions are censored. As a practical matter, we can observe at most \( T/\delta \) emissions during an experimental run. For the current hardware the observation period is 13 ms and the dead period is 3.2 ms, so we can observe at most four emissions.

In order to estimate \( \alpha \) from a series of such experiments, we need the probabilities of observing \( i \) emissions (or dead periods) for \( i = 1, \ldots, 4 \). Let \( D \) denote the number of emissions observed in an experimental run. We now evaluate

\[
p_i(\alpha) = P(D = i) = P(N = i)(1 - e^{-T}) + e^{-\alpha(1 - e^{-T})}.
\]

The computations of \( p_i(\alpha) \) for \( i = 2, 3, 4 \) are considerably more cumbersome, have progressively smaller contributions, and are thus omitted.

The proof of equation (3) is elementary. First, note that the generating function of a Poisson random variable is

\[
\sum_{n=0}^{\infty} P(N = n)\beta^n = e^{\alpha(\beta-1)}.
\]

Then note that the event \([D = 0]\) occurs if and only if either \([N = 0]\) or \([N \geq 1, W_1 > T]\). Conditioning on \( N = n \), we know from (1) that \( W_1 \) has an exponential distribution with pdf \( ne^{-nx} \); thus,

\[
p_0(\alpha) = P(N = 0) + \sum_{n=1}^{\infty} P(N = n)e^{-nT} = e^{-\alpha} + \sum_{n=1}^{\infty} e^{-\alpha} \frac{\alpha^n}{n!} e^{-nT} = \sum_{n=0}^{\infty} e^{-\alpha} \frac{\alpha^n}{n!} e^{-nT} = e^{-\alpha(1-e^{-T})}.
\]

The proof of equation (4) is more involved. We consider the cases \( N = 1, N = 2, \) and \( N \geq 3 \) separately. When \( N = 1 \), we observe one emission only if \( W_1 < T \), so

\[
P(D = 1, N = 1) = P(N = 1, W_1 < T) = P(N = 1)(1 - e^{-T}) = P(N = 1)[(1 - e^{-(T - \delta)}) + e^{-T}(e^\delta - 1)].
\]
Next, when $N = 2$, we observe only one emission if one of the following conditions holds: (a) $[T - \delta < W_1 < T]$, (b) $[W_1 \leq T - \delta, W_2 \leq W_1 + \delta]$, or (c) $[W_1 \leq T - \delta, W_2 > T]$. The probabilities of the three events are

\[
P(T - \delta < W_1 < T) = e^{-2(T - \delta)} - e^{-2T},
\]

\[
P(W_1 \leq T - \delta, W_2 \leq W_1 + \delta) = (1 - e^{-\delta})(1 - e^{-2(T - \delta)}), \quad (3.13)
\]

\[
P(W_1 \leq T - \delta, W_2 > T) = 2e^{-T}(1 - e^{-(T - \delta)}).
\]

These calculations come from the representation in equation (1). Thus, $P(D = 1, N = 2)$ is $P(N = 2)$ times these three terms.

When $N = n$ for $n \geq 3$, a similar listing of the configurations which allow us to observe one emission give the following probabilities. Once again, from the representation in equation (1), we have

\[
P(T - \delta < W_1 < T) = e^{-n(T - \delta)} - e^{-nT} = e^{-nT}(e^n - 1).
\]

Next,

\[
P(W_1 \leq T - \delta, W_n \leq W_1 + \delta) = P\left(\frac{U_1}{n} \leq T - \delta, \sum_{j=1}^{n} \frac{U_j}{n-j+1} \leq \frac{U_1}{n} + \delta\right)
\]

\[
= P[U_1 \leq n(T - \delta)]P\left(\frac{\sum_{j=2}^{n} U_j}{n-j+1} \leq \delta\right)
\]

\[
= (1 - e^{-n(T - \delta)})(1 - e^{-\delta})^{n-1},
\]

and

\[
P(W_1 \leq T - \delta, W_2 > T) = P\left[\frac{U_1}{n} \leq T - \delta, \frac{U_1}{n} + \frac{U_2}{n-1} > T\right]
\]

\[
= \int_{0}^{n(T - \delta)} e^{-(n-1)(T-x/n)} e^{-x} dx
\]

\[
= ne^{-(n-1)T}(1 - e^{-(T - \delta)}).
\]

Next, for $j = 2, \ldots, n - 1,$
\[ P(W^{(1)} \leq T - \delta, W^{(j)} \leq W^{(1)} + \delta, W^{(j+1)} > T) = P(A_1 \leq T - \delta, A_2 < \delta, \sum_{i=1}^{3} A_i > T), \]

where \( A_1 = U_1/n \) has pdf \( ne^{-nx} \),

\[ A_2 = \sum_{i=2}^{j} \frac{U_i}{n - i + 1} \]

has pdf

\[ g_{A_2}(y) = (j - 1)C(n - 1, j - 1)(1 - e^{-y})^{j-2}e^{-(n-j+1)y} \]

because it has the same distribution as that of the \((j - 1)\)st order statistic from a sample of size \((n - 1)\), and \( A_3 = U_{j+1}/(n - j) \). Thus, conditioning on \( A_1 \) and \( A_2 \), we get

\[ P[W^{(1)} \leq T - \delta, W^{(j)} \leq W^{(1)} + \delta, W^{(j+1)} > T] = \int_0^{T-\delta} \int_0^\delta e^{-(n-j)(T-x-y)}g_{A_2}(y)ne^{-nx} dy dx = C(n, j)e^{-(n-j)T}[1 - e^{-j(T-\delta)}][1 - e^{-\delta}]^{j-1}. \]

Combining these terms, we get

\[ p_1(\alpha) = \sum_{n=1}^{\infty} P(N = n)B_n = \sum_{n=0}^{\infty} P(N = n)B_n, \quad (3.14) \]

where

\[ B_n = e^{-nT}(e^{n\delta} - 1) + \sum_{j=1}^{n} C(n, j)(1 - e^{-\delta})^j e^{-(n-j)T}(1 - e^{-j(T-\delta)}) \]

\[ = e^{-nT}(e^{n\delta} - 1) + \sum_{j=0}^{n} C(n, j)(1 - e^{-\delta})^j e^{-(n-j)T}(1 - e^{-j(T-\delta)}) \]

\[ = e^{-nT}(e^{n\delta} - 1) + \frac{(1 - e^{-\delta} + e^{-T})^n - (e^{-(T-\delta)}(1 - e^{-\delta}) + e^{-T})^n}{(1 - e^{-\delta})}. \]

Note that \( B_0 = 0 \), as is the zeroeth term in the binomial expansion contained in \( B_n \). Finally, we use the generating function of a Poisson random variable to sum the series in equation (9) to get the expression in equation (4).

**Estimation of \( \alpha \)**

We now turn to the problem of estimating the Poisson parameter \( \alpha \). Suppose that we have \( r \) experimental runs, with \( r_i \) cases having \( i \) observed emissions. Below we
deal with two cases: use only $r_0$ or use both $(r_0, r_1)$. In both cases we seek the maximum likelihood estimate (MLE).

We first summarize the key facts about MLEs in general [Lehmann and Casella, 1998]. Suppose that $X^{(n)} = (X_1, \ldots, X_n)$ are independent and identically distributed observations with pdf $f(x|\theta)$, where $\theta$ is unknown. The likelihood is (up to a constant that does not depend on $\theta$)

$$ L(\theta|X^{(n)}) = \prod_{i=1}^{n} f(X_i|\theta). $$

More convenient is the log-likelihood $l(\theta|X^{(n)}) = \ln L(\theta|X^{(n)})$. The MLE is the value $\hat{\theta}_n = \hat{\theta}_n (X^{(n)})$ which maximizes $l(\theta|X^{(n)})$. When $l(\theta|X^{(n)})$ is sufficiently smooth, the MLE is a solution to $l'(\theta|X^{(n)}) = 0$ Under certain regularity conditions, $\hat{\theta}_n$ is unique and for large $n$, is approximately normal with mean $\theta$ and standard deviation $s = [-l''(\hat{\theta}_n|X^{(n)})]^{-1/2}$. Thus, an approximate 95% confidence interval for $\theta$ is $\hat{\theta}_n \pm 1.96 s$. With these general results, we now turn to the details of the two cases.

1. **Use only $r_0$.**

Suppose that there are $r$ experimental runs, with $r_0$ having no emissions and the remaining $(r - r_0)$ having at least one emission. The likelihood and log-likelihood functions for $\alpha$ are, respectively,

$$ L(\alpha|r, r_0) \propto p_0(\alpha)^{r_0}(1-p_0(\alpha))^{r-r_0} $$

and

$$ l(\alpha|r, r_0) = r_0 \ln(p_0(\alpha)) + (r - r_0) \ln(1 - p_0(\alpha)). $$

In the expression for $L$, we omit the combinatorial coefficient $C(r, r_0)$ because it does not depend on $\alpha$. In this case, maximum likelihood estimate (MLE) of $p_0(\alpha)$ is easy to compute: $\hat{p}_0(\alpha) = r_0/r$; thus, the MLE of $\alpha$ is

$$ \hat{\alpha}_{0r} = -\frac{\ln(\hat{p}_0(\alpha))}{(1 - e^{-T})} = -\frac{\ln(r_0/r)}{(1 - e^{-T})}. $$

Using the second derivative of the log-likelihood at the MLE, the approximate standard deviation is given by

$$ \hat{s}_{0r} = (1 - e^{-T})^{-1} \sqrt{\frac{(r-r_0)}{rr_0}}. \quad (3.15) $$
2. Use only \((r_0, r_1)\).

This time, suppose that there are \(r\) experimental runs, with \(r_0\) having no emissions, \(r_1\) having one emission, and the remaining \((r - r_0 - r_1)\) having at least two emissions.

Writing \(q(\alpha) = 1 - p_0(\alpha) - p_1(\alpha)\), the log-likelihood function is

\[
l(\alpha|r, r_0, r_1) = r_0 \ln(p_0(\alpha)) + r_1 \ln(p_1(\alpha)) + (r - r_0 - r_1) \ln(q(\alpha)).
\]

An explicit solution for the MLE \(\hat{\alpha}_{1r}\) is not available, but search using Newton-Raphson iterations with an analytic expression for the first derivatives will yield \(\hat{\alpha}_{1r}\). To do these calculations, we need the following expressions:

\[
l'(\alpha|r, r_0, r_1) = \frac{r_0 p'_0(\alpha)}{p_0(\alpha)} + \frac{r_1 p'_1(\alpha)}{p_1(\alpha)} + (r - r_0 - r_1) \frac{q'(\alpha)}{q(\alpha)},
\]

where

\[
\frac{p'_0(\alpha)}{p_0(\alpha)} = -(1 - e^{-T}) \quad \text{or} \quad p'_0(\alpha) = -(1 - e^{-T})p_0(\alpha)
\]

and

\[
(1 - e^{-\delta})p'_1(\alpha) = -(e^{-\delta} - e^{-T})e^{-\alpha[e^{-\delta} - e^{-T}]} + e^{-\delta}(1 - e^{-\delta})e^{-\alpha[1 - e^{-\delta}]}
\]

\[
\quad + (1 - e^{-\delta})(1 - e^{-T})e^{-\alpha[1 - e^{-T}]}
\]

allow for the calculation of \(l'(\alpha|r, r_0, r_1)\). Next, because \(p'_0/p_0\) does not depend on \(\alpha\), the second derivative is

\[
l''(\alpha|r, r_0, r_1) = \frac{r_1 p_1(\alpha)p''_1(\alpha) - [p'_1(\alpha)]^2}{p_1(\alpha)^2} + (r - r_0 - r_1) \frac{q(\alpha)q''(\alpha) - [q'(\alpha)]^2}{q(\alpha)^2},
\]

for which we need the expressions

\[
p''_0(\alpha) = (1 - e^{-T})^2 p_0(\alpha)
\]

and

\[
(1 - e^{-\delta})p''_1(\alpha) = (e^{-\delta} - e^{-T})^2 e^{-\alpha[e^{-\delta} - e^{-T}]} - e^{-\delta}(1 - e^{-\alpha[1 - e^{-\delta}]})
\]

\[
\quad - (1 - e^{-\delta})(1 - e^{-T})^2 e^{-\alpha[1 - e^{-T}]}.\]
3.7 Supplementary Material - Photon Counting Circuitry

3.7.1 Fast Discriminator Circuitry

We created a discriminator circuitry that allows for counting at speeds of 350 MHz and can be integrated into a single board which contains also contains the complex programmable logic device (CPLD) and auxiliary circuitry (Fig. S1). As an alternative strategy, fast discrimination can be accomplished with a commercial discriminator device (PRL-350TTL, Pulse Research Lab), albeit with the lower maximum effective counting rate of 280 MHz.

3.7.2 CPLD Circuitry

Critical to the implementation of our counting device is the use of an ultrafast CPLD, which allows digital counting and timing circuitry to be created in software and then downloaded into the device’s firmware. We use the Lattice IspMach LC4256V chip, which allows for 256 logic devices and runs at a maximum effective speed of greater than 400 MHz. The CPLD circuitry is created using ispLEVER Classic, and flashed to the chip using a standard parallel port interface.

Use of a CPLD results in a minimal digital hardware, since the counters, time gating, and delay line for synchronization with the laser pulses can all be implemented in the firmware of a single chip.

Fast counters.

There are many possible ways to implement digital counters. In the case of our photon counters, we designed for the following requirements. (i) The counters should be able to operate as quickly as possible, to allow for the possibility of multiple pulses within one laser repetition. (ii) The counters should not miscount in the case of two closely spaced pulses. (iii) The counters should avoid misreads by the computer (PC), since a pulse can occur even during a read, which could result in an erroneous value if the counter has not settled.

The solution is to use a dual ripple counter scheme (Fig. S2). Ripple counters are fast and stable, but have the disadvantage of a slower settling time, since only one flip-flop can change state at a time, and changes of more significant bits ripple through the system. Time multiplexing two separate counters allows for the fastest possible counting
**Figure 3.5**: Fast discriminator The discriminator threshold is set over the range of ±2.5 V by a front panel 10-turn potentiometer. The threshold voltage can be monitored at a front panel BNC. The discrimination is accomplished by an ultrafast comparator integrated circuit, (MAX9600, Maxim IC), which takes the input signal and the discrimination threshold voltage as inputs. The differential ECL output from the discriminator is converted to a ground referenced signal through a transformer and then amplified (GVA-84+, Mini-Circuits). The trimmer and inductor network provide power to the amplifier, and level-shift the output for TTL compatibility. Two amplifiers (HFA-1112IBZ, Intersil) provide buffered outputs for monitoring the input and output signals respectively.
Figure 3.6: Counting circuitry for lifetime imaging. Shown is an 8-bit counter for one channel. At any given time, only one of the counters is active, and the contents of the other inactive counter are connected to the output pins through a multiplexer. A rising edge of the SCANCLK from the data acquisition card means that a read has been completed; this clears the inactive counter, then switches the active and inactive counters after a short delay, giving time for the previously active counter to settle before being the next read cycle. The result is a counter which can operate up to 400 MHz, while avoiding the miscounting and read errors. For a typical photon counting setup, this is all the CPLD circuitry that is needed. In practice, three output channels can easily fit onto one CPLD chip. Four-bit counters (not shown) are identical, with the exception of four rather than eight flip-flops, and a 4-bit rather than 8-bit multiplexer at the output.

speed, without the possibility of counting or read errors. The dead time of this system is negligible, on the order of 1 ns.

Programmable digital delay line and multi-bin time gating.

In addition to functioning as a highspeed counter, the CPLD can be programmed to sort the incoming photon pulses into separate counters according to the arrival time relative to a synchronization signal (Fig. 3.7A) and thus execute fluorescent lifetime imaging microscopy (FLIM). The synchronization signal is derived from the laser pulse, and is sent through a series of gates that implement user-selectable delay lines. The delay line is set so that photon counts immediately following the excitation pulse fall into the first bin. In this configuration, a total of six time gated four-bit counters are used (Fig. 3.7B,C). The computer acquires this data as values from three eight-bit counters, and
sorts and analyses the data in post-processing.
**Figure 3.7:** Time delay and gating circuitry. (A) To capture lifetime decay information, a series of six gates are opened in sequence after a laser sync signal indicates the initial fluorescent excitation. In order to time the gates so that the initial fluorescence signal (after excitation) falls in the first gate, a digital delay line made from buffer gates is used. The length of the delay line is controlled by a 4-bit switch connected to the device (DLA), which allows the delay line to be tapped after a user-selected number of gates. The delayed laser-sync signal then uses a series of flip-flops to open six gates in sequence. The time between the opening of successive gates, as well as the active time of a given gate, is 1.5 ns. (B) When using the counting hardware for lifetime imaging, the counters can be run in one of three modes, selectable by an thumbwheel switch connected to the device. With the wheel set to 0, the photomultiplier tube (PMT) signal is connected to an 8 bit counter, in the standard configuration. With the wheel set to 1, channel 1 displays the result from the first time bin only. This is useful for setting the delay between the laser-sync signal and the PMT. With the wheel set to 1, the first 4 bits of channel 1 are used for time gated counter 1, and the remaining 4-bits are used for time gated counter 2. (C) In addition, four additional 4-bit counters are used to count pulses from the time bins 3, 4, 5, and 6.
Figure 3.7
(continued from previous page)

The overall design is similar to the standard photon counting setup, except that the FLIM design can be run in one of three modes, which are selectable using a thumbwheel on the counting device (Fig. S3B). Setting the thumbwheel to 0 implements the standard photon counting mode, using a single 8-bit counter. This is useful to obtain an overall image of the sample, or in cases where FLIM is not necessary or desired. Setting the thumbwheel to 1 implements a single 4-bit counter on channel one. In this mode, only counts from the first time bin (the bin which catches the initial excited photons) are displayed. This is useful for adjusting the delay lines on the laser-sync signal, and insuring that the counters do not rollover and exceed the maximum of 15 counts possible
for a 4-bit counter. And finally, setting the thumbwheel to 2 implements the full FLIM mode. In this mode, the six time gated channels, with 4-bits each, are multiplexed into the three imaging channels of the scope.

**Counting the number of pulses per laser repetition.**

The CPLD can be reconfigured to give the distribution of counts per laser repetition, in additional to acquiring the full number of counts. The key additional component is the use of flip-flops, configured so that the counters will inactivate after receiving either one or two pulses (Fig. S4). The flips-flops are reset by the laser sync signal. Because pulses arriving exactly during the reset can be missed, the laser-sync delay is adjusted to give the maximum number of pulses in the gated bins, thus ensuring the reset happens immediately before a new laser repetition, after the fluorescence has died out.

The number of times zero pulses occurred, r0, is calculated by subtracting the number of time one or more pulse occurred from the number of laser repetitions per pixel. Similarly, the number of times exactly one pulse occurred, r1, is calculated by subtracting the number of times two or more pulses occurred from the number of times one or more pulses occurred.
**Figure 3.8**: Measuring counts per laser repetition. In this configuration, three 8-bit counters are used. The upper counter records the number of times that one or more counts occurred during a laser repetition. The middle counter records the number of times two or more counts occurred during a laser repetition, and the third counter counts the full number of pulses. The PMT input to the upper two counters employ flip-flops, which are reset by the laser sync signal, to limit the number of counts per laser repetition.
4 Two-photon microscopy to measure blood flow and concurrent brain cell activity

4.1 Abstract of Chapter

The cerebral vascular system services the constant demand for energy during neuronal activity in the brain. Attempts to delineate the logic of neurovascular coupling have been greatly aided by the advent of two-photon laser scanning microscopy to concurrently image blood flow and the activity of individual neurons and astrocytes involved in the control of the flow. Here we review the procedures to generate optical access to the cortex for both rats and mice, determine the receptive fields of the exposed areas, and use two-photon microscopy to accurately measure blood flow in individual cortical vessels concurrent with local cellular activity. We illustrate the techniques with acute recordings from rats and chronic recordings from mice.

4.2 Introduction

Blood is a vital and limited resource in the brain. All aspects of neuronal and nonneuronal activity require a supply of oxygen and glucose — a need that constantly evolves with changes in brain activity [Fox and Raichle, 1986, Leybaert, 2005]. How is the distribution of blood controlled relative to these changing needs? Delimiting this phenomenon, commonly termed functional hyperemia or neurovascular coupling, remains an active area of research [Attwell et al., 2010]. However, recent studies have highlighted important conditions under which neural activity and blood flow become decoupled [Devor et al., 2008, Sirotin and Das, 2008, Jukovskaya et al., 2011], and thus
raise basic questions about neurovascular coupling [Kleinfeld et al., 2011]. One set of questions concerns the patterns of neuronal signals that lead to vasoactivity. A second set concerns the astrocytes that ensheath the vasculature and their role as intermediary cells that deliver signals from neurons to blood vessels.

The answer to the above questions depends on the ability to image blood flow and cells throughout the depth of cortex, 1.2 to 1.5 mm in rat and 1.0 to 1.2 mm in mouse. Imaging at this depth with subcellular resolution is facilitated by two-photon laser scanning microscopy (TPLSM) [Denk et al., 1990], an optical sectioning technique in which absorption of light to excite fluorescent molecules occurs only at the laser focus. Past studies have made use of TPLSM to examine vascular dynamics and blood flow in multiple brain regions, including somatosensory cortex [Kleinfeld et al., 1998, Zhang et al., 2005, Hutchinson et al., 2006, Schaffer et al., 2006, Wang et al., 2006, Brown et al., 2007, ?, Stefanovic et al., 2007, Winship et al., 2007, Zhang and Murphy, 2007, Devor et al., 2008, Shih et al., 2009, Sigler et al., 2009, Blinder et al., 2010, Fernández-Klett et al., 2010, McCaslin et al., 2010, Nishimura et al., 2010, Tian et al., 2010, Drew et al., 2011] and the olfactory bulb [Chaigneau et al., 2007, Petzold et al., 2008, Lecoq et al., 2009, Jukovskaya et al., 2011], down to depths of 600µm, which is sufficient to resolve vessels and neurons in layer 4. Recent advancements show that TPLSM can achieve imaging depths that allow single microvessels to be studied throughout the full depth of cortex [Kobat et al., 2009], and neuronal dynamics down to layer 5b [Mittmann et al., 2011], which is an important issue since vascular regulation in cortex appears to initiate in middle and deeper layers of cortex [Tian et al., 2010]. A further advantage is the concomitant use of exogenous and endogenous fluorescent based functional reporters to observe cellular activity, such as changes in intracellular $Ca^{2+}$ concentration [Wang et al., 2006, Chaigneau et al., 2007, Petzold et al., 2008], and the ratio of NADP to NAD+ [Mittmann et al., 2011, Kasischke et al., 2011] concurrent with blood flow and vessel diameter changes.

Here we discuss basic procedures of single and multi-vessel TPLSM imaging of blood flow dynamics, concurrent with cellular activity, in the somatosensory cortex of anesthetized and awake rodents. We further provide case studies. The equipment and algorithms used in these studies have been summarized elsewhere, including comprehensive reviews of basic methodology of vascular imaging [Shih et al., 2012], hardware [Tsai and Kleinfeld, 2009, Driscoll et al., 2011b] and software [Nguyen et al., 2006, Nguyen et al.,
2009] for TPLSM, and algorithms for data analysis [Drew et al., 2010a, Valmianski et al., 2010, Driscoll et al., 2011a, Kleinfeld and Mitra, 2011]. Additional work has addressed the use of TPLSM to image histological tissue with labeled vasculature [Tsai et al., 2003, Ragan et al., 2007, Tsai et al., 2009, Kleinfeld and Mitra, 2011, Tsai et al., 2011].

4.3 Choice of preparation

Both rats and mice have their place in cerebral blood flow imaging studies. The relatively large size of rats allows them to tolerate anesthesia better than mice and makes them the animal of choice for complex surgical procedures. Extended cranial windows can be fabricated to permit access to multiple regions of cortex, blood samples may be obtained at multiple time points in a procedure, and physiological parameters can be readily controlled. However, a current disadvantage of cranial windows, and thus rats, is that the imaging quality degrades within days. The use of mice have two advantages. First, they allow researchers to exploit the wide range of vascular related transgenic animals. Second, transcranial windows with a thinned skull may be fabricated. The use of transcranial windows obviates potential problems with inflammation and changes in cranial volume and are excellent for repeated imaging studies over many months. On the down side, blood sampling and physiological control is limited with mice compared with rats.

4.4 Cranial windows

The generation of a cranial window for optical access in rats and mice differ on a number of levels. In rats, the overlying bone must be completely removed. Further, the dura mater must be carefully resected to the edge of the imaging window for optical access [Levasseur et al., 1975, Morii et al., 1986, Kleinfeld and Delaney, 1996]. The window must then be resealed to restore intracranial pressure and to minimize motion artifacts caused by heart beat and breathing. Very large windows can be generated, i.e., 4 x 6 mm, to facilitate easy positioning of electrodes and cannulas. The clarity through such a cranial window is initially optimal as the materials overlying the pial surface cause minimal scattering. However, dural regrowth degrades the imaging quality such that repeated imaging is limited to about four days. As a result, chronically implanted windows for repeated imaging in rats are rarely reported and the use of pharmacological
agents to suppress inflammation may also affect the phenomenon under study, such as
the magnitude of injury in experimental stroke models [Tuor et al., 1993]. Longer lasting
windows may, in principle, be achieved by using inert substances such as Kwiskil silicone
for movement suppression [Dombeck et al., 2009].

Cranial windows in mice are surgically less demanding, as the dura is thin and
does not need to be removed for optical access. Detailed methods have been described
for cranial windows [Mostany and Portera-Cailliau, 2008, Holtmaat et al., 2009] and
transcranial, thinned skull windows [Yang et al., 2010].

4.5 Transcranial windows

An alternate method that is suitable for mice is to generate a stable transcranial
window, where the skull is thinned, polished and reinforced with a thin layer of glue and
cover slip [Drew et al., 2010b]. These windows, which may be as large as 2 x 2 mm, greatly
minimizes disruption of the intracranial milieu, reduces inflammation, and prevents bone
regrowth. While the imaging depth and clarity are somewhat reduced compared to
windows with complete bone removal, polished and reinforced cranial windows in mice
have been proven to give excellent clarity for TPLSM imaging months after the initial
surgery. This procedure has so far failed with rats because of the lower clarity of the
dura and skull.

4.6 Localization of active areas

Standard brain atlases provide approximate coordinates for different brain re-
gions as a means to locate the vasculature in relation to loci of neuronal activity. Yet
somatotopically refined maps are often of considerable value. The lissencephalic structure
of the rodent brain permits different regions in cortex to be further mapped to determine
receptive fields using a variety of classical tools. For sensory areas, these include surface
electrodes and intrinsic optical signal (IOS) imaging. The latter technique avoids any con-
tact with the brain and uses changes in the intensity of reflected light to report a change
in the ratio of oxy- to deoxyhemoglobin that occurs secondary to changes in neuronal
activity [Grinvald et al., 1986, Frostig et al., 1990]. Motor areas may also be mapped in
one of three ways: (i) measuring limb or vibrissa movement in response to intracortical
microstimulation with bipolar microelectrodes; (ii) measuring electromyogenic activity in
response to intracortical microstimulation; or (iii) through the use of focal illumination in conjunction with mice that express channelrhodopsin in projection neurons [Ayling et al., 2009][60].

We illustrate the mapping process with IOS imaging for the case of vibrissa primary sensory cortex (Fig. 4.1). A craniotomy was prepared, as described previously [Kleinfeld and Delaney, 1996], and individual vibrissae on the face of the animals were mechanically stimulated. Stimulation of a single vibrissa leads to a net decrease in reflectance within a localized region of vibrissa primary sensory cortex (Fig. 4.1A). This corresponds to a local decrease in blood oxygenation. The centroid of this signal and the overall shape of the active region remains roughly constant between blocks of trials (Fig. 4.1B), although the amplitude varies greatly between blocks. Nonetheless, IOS imaging serves as a useful tool for mapping the center of activation in cortex for the different vibrissa (Fig. 4.1C). The branching arrangement of pial arterioles and venules, which is unique to each animal, serve as fiducials to relocate the active region between intrinsic imaging and two-photon measurements.

4.7 Measurement of blood flow dynamics in single cortical vessels

We consider first the measurement of the flux of blood flow. This involves a user-defined scan profile for TPLSM to simultaneously measure both the speed of red blood cells (RBCs) and the diameter of vessels as a means to compute the volume flux of RBCs. The vasculature is labeled with an intravenous bolus of fluorescein- or rhodamine-conjugated dextran for green and red light emission, respectively. The cerebral vasculature within the window is first imaged at low resolution with a 5X objective (Fig. 2.2A); this image can be directly compared with that made with IOS imaging (Fig. 4.1C). High resolution imaging of surface pial vessels, penetrating vessels, and subsurface capillaries can then be performed in smaller regions with a high numerical aperture water dipping objective, such as a 20X to 40X objective with a numerical aperture ranging from 0.8 to 1.0 (Fig. 2.2B).

When the serum is labeled, RBCs exclude the dextran dye and appear as dark objects moving against a bright fluorescent background. We used custom software to direct the imaging laser beam in a user-defined path within the imaging plane (4.11) [Valmianski et al., 2010,Driscoll et al., 2011a]. Linear segments of constant scan speed traverse along
**Figure 4.1:** A female Long Evans adult (2 month) rat was anesthetized with isoflurane (2% in O2 for induction and < 1% sustained) and a 4 mm x 4 mm closed craniotomy performed over the vibrissa area of primary somatosensory cortex. The intrinsic optical signal (IOS) was obtained as the reflectance at 633nm as a function of time and trial, i.e., $R_i(x, y, t)$ where $i$ labels the trial, the sampling period is 50ms, and the pixel width is 8µm. The duty cycle for each trial was 18s, and we typically obtained 30 trials. We computed the smoothed, trial-averaged response as $R(x, y, t) \equiv \text{Gaussian}[\frac{1}{\text{median}}[R(x, y, t)] > \text{trial}]$ where the Gaussian filter had a width of $\sigma = 4$ pixels. The initial 4s of the average signal constitutes baseline, i.e., $R_{\text{base}}(x, y) \equiv R(x, y, t) > 0s \leq t \leq 4s$, the second 4s were used to deflect individual vibrissae ($\pm 12^\circ$ by 10Hz square wave filtered with a 6th order 100Hz Bessel low pass filter), and a period after the onset of stimulation was used to form the response, i.e., $R_{\text{response}}(x, y) \equiv R(x, y, t) > 4.5s \leq t \leq 9s$. We report the IOS as $\frac{\Delta R}{R} \equiv \frac{[R_{\text{response}}(x, y) - R_{\text{base}}(x, y)]}{R_{\text{base}}(x, y)}$. (A) An example of the intrinsic signal for deflection of the C2 vibrissa. The dark declivity indicates reduced reflectance of red light, suggesting an increase of deoxygenated hemoglobin and thus increased neural activity in that region. (B) Intrinsic signal for a profile line through the centroid of the activated region in panel A for three separate blocks of trials. The horizontal dashed lines are a 50% decrease of the intrinsic signal; the solid vertical lines indicate the centroids. (C) The thresholded IOS image, e.g., dashed line in panel B, for 14 different vibrissae across the cortical surface. The surface map was obtained by reflectance at 475 nm. Thresholds values were $B1 = -1$, $B2 = -2.5$, $B3 = 0$, $C1 = -2$, $C2 = -1$, $C3 = 0$, $C4 = -1.5$, $D1 = -2$, $D2 = -1$, $D3 = -1.5$, $D4 = -3$, $E1 = -0.5$, $E2 = -3$, $E3 = 0$; all $\times 10^{-4}$. In some areas the vessels were masked from the calculation, e.g., vibrissa D2.
the length of the center of the vessel and across the width of the vessel to measure RBC speed and lumen diameter, respectively. These linear scan segments are connected by polynomial splines, where connecting portions of the scan are accelerated to allow for rapid data collection across multiple vessels [Driscoll et al., 2011a] (Fig. 2.2C).

The resulting line scan data forms a space-time image, typically displayed with the individual scan lines stacked on each other (Fig. 2.2D). In principle, many vessels that lie in the same plane can be measured simultaneously. Portions of the scan path along the centerline of the vessel lumen reveal angled streaks within the cascade image. Moving RBCs in flowing vessels sampled at a sufficient rate will appear as diagonal streaks. The centerline velocity is proportional to the slope of the RBC streaks, measured from vertical (Fig. 2.2E). Determining this slope is most efficiently performed with a Radon transform of the data [Drew et al., 2010a].

A velocity time series is calculated by transforming successive time windowed portions of the line scans (right panel in 2.2F). The temporal spacing of successive windows must be close enough to resolve the highest velocity modulation frequency, the heart rate, which is 6 Hz for rats and 10 Hz for mice. In addition, the window size must be large enough to capture enough streak lines so that the Radon transform has sufficient data to calculate an accurate velocity value. We find that a window size of 40 ms with a spacing of 10 ms is a good compromise. In addition to heart rate, other physiological signals detected in the RBC velocity time series include breathing, at 1 Hz for rats and 2 Hz for mice, and vasomotion at 0.1 Hz for rat and 0.1 to 1 Hz for mice [Mayhew et al., 1996, Drew et al., 2010b, Kleinfeld and Mitra, 2011].

As with the velocity calculation, the diameter calculation is taken from a temporal windowed portion of the data (Fig. 2.2D). The same window size and spacing used for velocity is also used for diameter so that both parameters are concurrently calculated. Vessel diameter is defined as full width at half maximum of the vessel profile for each window (left panel in Fig. 2.2E); the intensity profile tends to increase near the edges due to the exclusion of RBCs from the glycocalyx. The two outermost half maximal points of these peaks are used to calculate the vessel boundary. Linear interpolation is used to add subpixel accuracy to the diameter measurement.

When the diameter of the vessel is much greater than that of the RBC, the flow is laminar and nearly parabolic [Rovainen et al., 1993, Schaffer et al., 2006]. The two vascular parameters, RBC velocity and lumen diameter, are combined to calculate the
volume flux, RBCs and plasma, for each vessel. The volume flux through the vessel is given by

\[ F = \frac{1}{2} v(0) \pi R^2 \]  

(4.1)

where \( v(0) \) is the time averaged RBC velocity at the center line of the vessel, and \( R \) is the radius of the lumen. This formula underestimates the flux as the nonzero spatial extent of the RBC flattens the parabola of Poiseuille flow.

As an example, we measure the change in flux at the level of single penetrating vessels in response to somatosensory stimulation. Both diameter and RBC velocity in the arteriole respond to stimulation (left column in Fig. 2F). The flux through the arteriole increases to a peak of 86% over baseline, compared with a much smaller peak increase of 29% and 24% for diameter and velocity measurements alone, respectively. The increase in RBC velocity is partially masked by a peak in the underlying vasomotor fluctuation, but remains a significant increase over an average one minute period of basal activity. In contrast to the arteriole, a neighboring venule exhibits no change in lumen diameter, but a 23% change in RBC velocity. As a result, the flux increases in the venule by 23% as well (right column in Fig. 2.2F). This increase in venular flux would not be detected with methods that measure diameter only.

### 4.8 Simultaneous imaging of blood flow and local cellular activity

An important goal is to identify the cell types and signaling pathways that regulate neurovascular coupling. The activity of various neuronal cell types and astrocytes can be monitored in superficial cortex by bulk loading the tissue with functional dyes, in this case Oregon Green Bapta-1 AM (OGB-1 AM), to detect changes in intracellular Ca\(^{2+}\) concentration [Stosiek et al., 2003, Garaschuk et al., 2006](Fig. 3A). The Ca\(^{2+}\) indicators may be injected by pipette into the active region through a vent in the cranial window under two-photon guidance, or prior to sealing the window during the initial craniotomy procedure. Astrocytes are selectively labeled with the astrocytic marker sulforhodamine 101 (SR101) in the same tissue to distinguish them from neurons in a second imaging channel [Nimmerjahn et al., 2004].
**Figure 4.2:** (A) A full field image of 19 neurons (N), 1 astrocyte (A), and 3 blood vessels, obtained at 4 frames/s, with a scan path superimposed on cells determined by our machine learning algorithm. All cells and vessels are scanned at 110 Hz. The green channel shows the fluorescence from OGB-1 and fluorescein while the red channel shows fluorescence from SR101. (B) Activity of neurons and an astrocyte, indicated in panel A, in response to a single weak electrical shock to the forelimb. (C) The Ca^{2+} response of the astrocyte (A1), the average neuronal response (N1 - N19), and the speed of red blood cells in one capillary (V1). All panels adapted from Valmianski et al. [42].
Figure 4.2
(continued from previous page)
We consider a particular field of view that contains 20 identified cells, i.e., 19 neurons and one astrocyte, along with three blood vessels. Even at this level it is tedious to localize each labeled cell and guide the laser scan path manually. We thus used a machine learning algorithm to locate each soma [Valmianski et al., 2010] in conjunction with full-field images from the OGB-1 emission channel. Cells that were colabeled with SR101 were automatically labeled as astrocytes; the coordinates of selected blood vessels were also marked. We then calculated the fastest scan path through each soma along with selected microvessel (Fig. 4.2A). The typical signal-to-RMS-noise ratio, which we define as the ratio of the peak of the response to the RMS noise during the baseline, for a change in intracellular Ca\(^{2+}\) induced by a single sensory stimulus is 10 (Fig. 4.2B). Further, the resulting line scan image contains stereotypical angled streaks for calculation of RBC speed, as well as intensity traces from selected neurons and astrocytes, similar to that seen in Figure 2.2C. The composite data permits changes in astrocytic Ca\(^{2+}\) levels, together with changes in the speed of RBCs in a nearby microvessel, to be compared with the composite neuronal activity (Fig. 4.2C).

4.9 Stimulated and basal hemodynamics in awake mice

We consider the quantification of blood flow both at and below the cortical surface in awake mice. Mice were habituated to head fixation and blood flow was measured in single vessels through a chronically implanted transcranial window (Fig. 4A). Robust arteriole dilations could be evoked by prolonged contralateral whisker stimulation (Figs. 4B and 4C, red). Small arterioles dilated proportionally more than larger arterioles and the rapid dilation that followed vibrissa simulation and vibrissae evoked dilation (Figs. 4.3A and 4.3B). In contrast, the delayed, second peak in the dilation does not show a significant dependence on vessel size (Fig. 4.3B), nor does the response to control stimuli (Fig. 4.3C). Pial venules, typically thought to be static in terms of diameter, show a delayed and weak dilation in the awake state that had no dependence on initial diameter (Fig. 4.3C). These data suggest that functional hyperemia changes detected by BOLD fMRI may be dominated by large changes in arteriole volume, in agreement with recent studies [Kim and Kim, 2011], rather than in venules [Buxton et al., 1998].

Individual traces of the arteriole diameter show a high degree of variation and spontaneous dilation (Fig. 4.3B). Similar to the case of stimulus induced changes, small arterioles dilated proportionally more than larger arterioles for both spontaneous
Figure 4.3: (A) Schematic of the experimental setup. The awake mouse is head fixed by means of a bolt and sits passively in an acrylic cylinder beneath the two-photon microscope. Air puffers for sensory stimulation are aimed at the vibrissa and as a control at the tail. (B) Individual dilation responses to 30 s vibrissae stimulation. (C) Plot of peak averaged dilation responses to 30 s vibrissae stimulation. Early arterial peaks, in the 0 to 10 s interval after stimulation, are denoted by red circles; regression slope $= 0.007 \mu m^{-1}$ ($r^2 = 0.15, p < 0.02$). Late arterial peaks, greater than 10s after onset, are denoted by red triangles; the linear regression (not shown) is not significant. Venules are denoted by blue dots; the linear regression is not significant. (D) Plot of peak value of the spontaneous dilations for arteries, in red, and veins, in blue. Grey area shows the $0.2 \mu m$ resolution limit of detectable changes. Lines show linear regressions; slope $= -0.004 \mu m^{-1}$ for arterioles is significant ($r^2 = 0.13, p < 0.001$), while that for veins (not shown) is not significantly different from zero. All panels adapted from Drew et al. [27].
events (Fig. 4.3D). The critical point raised by this data is that the magnitude of both prompt and delayed arterial responses induced by stimulation are similar in magnitude to spontaneous arterial dilations, with typical dilations of 30% and maximum dilations near 50% for both cases (Figs. 4.3B to 4.3D). Thus, from the point of control, stimulus induced changes in blood flow are small, i.e., on the order of the noise level. The corollary is that individual stimulus events cannot be distinguished from spontaneous dilations based on magnitude alone [Kleinfeld et al., 1998, Drew et al., 2008].

4.10 Imaging of blood flow in deep cortical layers

An important development in TPLSM is the ability to image deep cortical layers that are responsible for the output of neuronal processing [Helmchen and Denk, 2005]. Further, different cell types predominate in different layers, so that deep imaging permits neurovascular coupling to be studied in changing environments. Longer wavelengths of light penetrate deeper into tissue as a result of reduced scattering, absorption, and optical aberration by the tissue. In the example of Figure 5, a cranial window was prepared in a mouse and the dye Alexa 680-conjugated dextran was used to label the blood serum. The dye was excited with 1280 nm wavelength light from a Ti:sapphire pumped optical parametric oscillator and was found to enable imaging through the entire depth of cortex, i.e., 1 mm deep (Figs. 5A and 5B) [Kobat et al., 2009]. Red blood cell velocities in capillaries could be resolved as deep as 900µm into cortex (Figs. 4.4C and 4.4D).

4.11 Automatic Identification of Regions of Interest

In order to generate a standard two dimensional TPLSM image, the laser focal point is scanned systematically across the entire field of view in a raster pattern. Repeating the scan pattern generates a time series of frames, resulting in an easy to interpret movie of the entire field of view or a zoom of a smaller region. However, in many cases the relevant optical signal is spread through small portions of the field of view, such as when a sparse network of labeled neurons fire in response to stimuli. In such cases, it is advantageous to scan primarily across the active regions of interest, to maximize the number of signal photons acquired per unit of time. This not only maximizes the signal to noise ratio, but allows imaging on much faster timescales than the typical raster scan frame rates.
Figure 4.4: (A) Maximum intensity projections of mouse cortical vasculature in the coronal orientation. The image stack was collected over the entire depth of cortex through a cranial window with the dura intact. To reduce scattering and improve imaging depth, a long wavelength of excitation, i.e., 1280 nm, was used and the blood plasma was labeled by intravenous injection of Alexa 680 conjugated to dextran (Table 2). (B) Single planar image taken from panel A at the depth of the red line. (C) Magnified image taken from a vessel from the region in the red box in panel B. (D) Magnified image taken from a vessel from the region in the red box in panel B. All panels adapted from [Kobat et al., 2009].
Frequently, the field of view contains only a few obvious neurons, or regions of interest (ROIs), which can be selected by the user during the experiment. In cases where a large number of ROIs are present, such as a sparse network of labeled neurons which are responsive to stimuli, machine learning can be used to automatically detect the regions and create an optimized scan path which passes through all of them in a minimum amount of time. Typically, a computer algorithm is trained on experimental data which has been annotated by a human to mark relevant areas that the algorithm learns to recognize. During an experiment, the algorithm can then be used to quickly and automatically identify the ROIs. A typical workflow is [Valmianski et al., 2010]:

1. A region of cortex is labeled with a Ca\textsuperscript{2+} indicator. An initial, full frame movie is collected while the animal is stimulated so that a subset of the cells responds with increased fluorescence. Multiple training and validation datasets can be created by taking multiple movies over the same or similar regions.

2. After the experiment, the movies are marked by a human expert. Regions are labeled as a "cell" or "possibly a cell". Unmarked regions are considered to be not a cell. This training data set is used to create a pixel classifier, which is used to determine if a given pixel is or is not a cell, and a blob classifier, which uses the shapes of pixels to determine actual cell regions. The data is used to generate a set of mathematical rules, which can be implemented by the computer, to form a basis for the classifiers [Freund, 2009]. The classifiers operate on eight features in the dataset: the mean, variance, covariance, and correlation contained in the images, as well as the mean, variance, covariance, and correlation normalized by the standard deviation. After identifying pixels which are likely to belong to a cell, a morphological classifier using features such as size and eccentricity are used to delineate intact cells.

3. During subsequent experiments, a movie is taken of a region similar to that used in the original experiment. This movie, along with the classifiers created previously, is fed into the software, which automatically generates a set of ROIs along, with an initial laser scan path across them. The ROIs can then be modified by the user, typically by adding or removing ROIs.

4. The time it takes to scan across all the ROIs will depend on the order in which they are scanned. Optimizing the order is the classic traveling salesman problem, which
has been extensively studied in computer science and mathematics. In general, it is not possible to search through all possible paths, and some computational shortcuts are taken. We use the ANT optimization algorithm to search rapidly through a subset of all possible paths as a means to find the shortest path that travels through all regions [Di Caro and Dorigo, 1998].

4.12 summary

Two-photon microscopy provides a number of advantages that will aid the study of the mechanisms underlying neurovascular coupling and cerebrovascular disease in animal models, including i) the resolution needed to visualize single cortical vessels and their surrounding cells, ii) penetration depths of 250µm through a transcranial window and 500 µm with craniotomies at 800nm excitation, and imaging to 1000µm with longer excitation wavelengths, iii) reduced photodamage, iv) high speed user-defined line scans for near simultaneous measurement of RBC velocity, lumen diameter and local cellular activity, and v) the opportunity to image vascular dynamics deep in the cortex of awake mice.

A wide variety of functional fluorescent dyes can be exploited for studies of neurovascular control. Recent focus has been on Ca^{2+} sensors to study cellular activity. Exogenous sensors have the advantage of labeling both neurons and astrocytes, along with a meshwork of processes that intervene between cell bodies. However, the cellular source of a signal may be difficult to resolve when focusing on fine processes. Transgenic mice with different neuronal subtypes labeled with fluorescent proteins can aid the separation of responses from different cell populations [Sohya et al., 2007]. Genetically encoded Ca^{2+} sensors exclusively expressed in a specific cell type can be essential in separating signals from individual cells [Tian et al., 2009]. Finally, cyclic adenosine monophosphate sensors would also be of value since Ca^{2+} independent pathways may also be involved in neurovascular control.

The intermixing of cell types within a small tissue volume also hinders electrophysiological and pharmacological approaches to query the role of cell types in neurovascular coupling. New advances with light activated opsins [Lee et al., 2010] or engineered receptors with unnatural affinities for exogenous chemicals [Alexander et al., 2009] may help to unravel issues with cellular specificity [Kleinfeld et al., 2011]. However, caveats should also be considered, as activation of one cell type does not preclude activation
of non-targeted cells linked within the same circuitry. The manipulation of cell-specific vasoactive signaling cascades will be an important step in dissecting the chemical basis of neurovascular coupling, but will also be challenging as new tools to knockdown gene expression will need to be developed [Kleinfeld et al., 2011].

4.13 Acknowledgments

5 Conclusions and Future Directions

Techniques and tools to understand the neurovascular coupling system, and the brain in general, have progressed at an incredible pace in the past two decades. With the invention of the two-photon microscope 20 years ago, the subsequent widespread use of transgenic animals allowing for fluorescence imaging of specific proteins and systems, and an ever expanding library of reporter molecules such as calcium and voltage indicators, the ability to actually see what is happening in vivo in now entering a new era.

However, we still do not know much about how the brain actually accomplishes many tasks, even basic ones. For instance, although early experiments strongly implicated direct astrocytic control of vascular tone [Mulligan and MacVicar, 2004], experiments in this lab and others have so far failed to find a reliable connection between astrocytic activity and neighboring vessel diameter.

It may be that the brain has frequently been studied in conditions far from that found in normal, healthy animals, and clear answers to brain function have been confounded by looking at partially damaged or non-functioning systems. In the extreme case, in vitro experiments in slice can demonstrate features of neurovascular coupling, such as vessel diameter change in response to astrocytic calcium transients, but exhibit timescales an order of magnitude slower or more than that found in vivo [Mulligan and MacVicar, 2004]. Even in anesthetized animals, the use of anesthesia itself can have a profound impact on the neurovascular coupling system — certainly, an anesthetized animal behaves differently than an awake one! In fact, recent research has compared the anesthetized state with that of a “reversible coma,” certainly not the desired state for understanding normal brain function.

Future work in lab is refocusing on the use of awake, behaving animals to study the brain in its natural state. In addition, surgical techniques which allows for optical
access to the brain through a thinned skull circumvent the trauma associated with full craniotomies [Drew et al., 2010b]. In the future the use of ablation lasers with feedback from second harmonic generation and atomic spectra for surgery may allow an additional level of less invasive imaging, as many of the effects of heating and vibration associated with traditional surgical techniques can be avoided [Jeong et al., 2011].

In addition, current research in the lab is focused on the use of scanning systems with close to gigapixel resolution over roughly a square centimeter, with the goal of looking neurovascular coupling on a much larger scale, for instance, bilaterally across the two hemispheres of the brain. In conjunction with this, scanning techniques have been developed which allow for three dimensional scanning [Göbel et al., 2007], though technical challenges remain. The use of advanced reported dyes and cells, such phosphorescence oxygen sensors or CNiFER [Nguyen et al., 2010] reporter cells, open the possibility of actually seeing the state variables and computational mechanisms of the system, rather than just the presumptive inputs (neural activity) and outputs (vessel diameter).

If the neuroscience technology and techniques over the next two decades progress at anything near the pace of the previous two, we may very well be on the cusp of a major breakthrough in understanding how the brain work; regardless of how things turn out, the path forward will certainly be challenging and rewarding
A Ultrafast PMT Preamplifier and Discriminator

The current laboratory electronics setup for photon counting involves the use of several separate electronics enclosures, connected together through 50Ω BNC wiring. Although this is an extremely practical setup, since most of these components are commercially available, there are a number of disadvantages.

1. The high-speed amplifier, ZPUL-30P, is AC coupled. Although this amplifier a wide-bandwidth, pulse amplifier, the lower corner frequency on the amp is $2.5kHz$, giving a decay constant $\tau$ of roughly $64\mu s$. For most low-to-mid intensity imaging, the AC coupling is a non-issue, since the PMT current pulses are spaced far apart, and thus do not change the baseline voltage offset of the output, which remains near zero volts. However, at high light intensities, the current from the PMT can cause the baseline voltage to drift, potentially resulting in missed pulses. In imaging, this shows up as apparent banding around high intensity regions. A DC coupled amplifier would avoid these issues.

2. Since the CPLD takes a CMOS level input, the discriminator must drive a fairly high voltage signal ($>4V$) across the BNC connectors, in order for the digital signal to still be above threshold at the receiving end. This is because the 50 Ohm output impedance and the 50 Ohm termination form a resistor divider network, cutting the voltage in half. This requires fairly complex circuitry. A better solution would be a system where the discriminator is integrated with the CPLD chip, which would reduce the necessary voltage swing by half, allowing much simpler, standard comparator and level translation chips to be used.

3. Although not a part of the original project, if the counting board were to be commercialized or used by other labs, there is a cost effectiveness to creating a
Figure A.1: The proposed circuit includes a two stage PMT pre-amp with a bandwidths of in excess of 1 GHz, fast discriminator, and a logic conversion stage for interfacing with the CPLD. See text for a detailed circuit description.

small, integrated system, using a minimal part count. Integrating the system into one box not only would make the setup much more practical and user friendly, but also allows for higher performance with less cost.

As a note for future applications, this section describes a proposed circuit to integrate a high-speed (> 1 GHz) preamplifier, with a high speed comparator and level translation circuit, which can be integrated into a board containing the CLPD counter and digital interface circuit. The discriminator and level translation circuitry operate with rise times > 750 ns, faster than the typical rise times of the PMT pulses, and slightly faster than the ∼ 400 MHz maximum counting speed of the CPLD.
A.1 Component Overview

A.1.1 Fast DC coupled amplifier

The ADA4960-1 chip by Analog Devices is a new chip, which provides 18 dB of gain in a single stage when used differentially, and 9 dB of gain as a single-ended amplifier. By cascading two gain stages, a 27 dB amplifier can be made, giving a gain of 22. This is approximately the same gain as the ZPUL-30P amplifier by mini-circuits, and provides enough gain to run the comparator. Each stage has a $-3dB$ bandwidth of 1.2 GHz, and a rise time of a few hundred ps, faster than the rise time of the PMT pulses.

Furthermore, the ADA4960-1 chip can be DC coupled from end to end, eliminating any AC coupling artifacts, requires very few supporting components, and is relatively inexpensive.

A.1.2 Ultrafast Comparator

The ADCMP572 by Analog Devices comparator chip is one of a number of possible chips to convert the analog PMT pulse to a digital signal. As with all ultrafast comparators, the output is differential, and the output signal must pass through a level translator before being passed to the CPLD.

A.1.3 Level shifting

Most ultrafast comparitors have differential outputs, typically ECL, PECL, LVDS, or CML. In this case, the ADCMP572 chip has CML outputs, which must be translated into CMOS / TTL levels in order to be compatible with the CPLD inputs.

The MC100EPT21 level translation chip can operate in excess of 500 MHz, and still maintain a swing from $<0.5V$ to $>2.2V$. Note that the datasheet lists a typical maximum output frequency of 350 MHz, since this is the maximum rate for a full swing of the output; however, the CPLD will still trigger on a reduced swing output. In addition, this chip requires minimal supporting components.

A.2 Circuit description

The ADA4960-1 chip is typically run using a $Vcc = +5V$ signal and ground, with the inputs and outputs running on a DC offset of $\frac{Vcc}{2}$. In order to DC couple the inputs and outputs, the chip can be run from $\pm2.5V$, meaning the offset in this configuration is
0V. The input voltage range in this configuration is $\sim 250mV$, and the output range is $3.5V$ PP.

The first gain stage gives a differential voltage gain of 18 dB, with is then fed into a second ADA4960-1 chip. Since the second chip must supply a single-ended output, it adds an additional 9 dB of gain to the system.

The comparator is run from a +3.3V source, and can work with input levels from -0.2 V to 1.2V. Since PMT pulses result from and electron cascade, they are negative going pulses by default. By using the negative input from the second stage, the comparator can trigger on a positive going pulse, which simplifies design.

Finally, a translator must be used to convert the CML output from the comparator into a single ended. All commercially available ultrafast comparators have some form of differential output, which is not compatible with the single ended CMOS / TTL level required by the CPLD chip. The MC100ETP21 chip accomplishes the translation with minimal fuss. Note, the specs on the chip state $> 275MHz$, operation speed but this is extremely conservative. For the voltage swings necessary to operate the CPLD, the chip can operate at least to $500MHz$, due to the rise and fall times of approximately 750ps.

One final note this circuit has not, as of yet, been built and tested. Other modifications may be necessary. It is presented here as a reasonable future possibility for further development.
B  MpScope3 User and Developer’s Guide

B.1  Overview

MpScope is a software suite designed to control a two-photon scanning microscope. The software coordinates hardware, such as the analog and digital DAQ cards and motion stages, handles streams large datasets to file, and includes both a graphical interface and the ability to script through MATLAB.

MpScope was originally written by Quoc Nguyen in the Kleinfeld lab at UCSD. The software is distributed free of charge, and is currently used in over two dozen labs worldwide. Version 2.0 added the ability to use photon counting hardware. However, this version was written in Pascal using older data acquisition card drivers, and has a number of other features in need of improvement. For instance, MpScope 2.0 has limitations with the file format (files are limited to 2 GB of data and a 4096 frames, and can only be opened on 32-bit Windows machines), scan size (512 x 512 image resolution is hard-coded throughout the software), and data acquisition hardware.

Thus, MpScope was rewritten to add in additional features, and create a stable, scalable code base for additional, more complex features, such as 3D scanning.

The 3.0 version is a complete rewrite of the code using C# and the modern .NET platform. The newer version was written by Jonathan Driscoll in the Kleinfeld lab. In order to easily handle large file sizes in a transparent, cross-platform way, the new software writes data in HDF5 format, which is free and well supported file format from the supercomputer world. In addition, the newer NI-DAQmx drivers allow for a broader range of hardware options, and the ability add additional cards for additional data acquisition.

This document describes MpScope 3.0, first from a practical, user viewpoint, and
Figure B.1: The C# components, shown in beige, center around the Engine, which is the core of the program. The C# forms, which provide user interaction, are shown in light beige. Additional components, shown in white, and typically accessed through the .NET interface.

then more in depth, for those who may want to modify the code at a later date.

B.2 Setting Up The Program

B.2.1 Software Drivers and Libraries

MpScope needs several drivers and software libraries to in order to work properly. This software is located in the “MpScope3_Install” directory of the MpScope 3.0 distribution. Below are the steps required to install the software

- **Install HDF5 file support.** The HDF5 library must be installed on the system for file access. To install, simply copy the files in "HDF5 library," on the Installation disk, to the “Windows\System32” directory. Note that many programs, such as MATLAB, already can read and write HDF5 files.

- **Install NI-DAQmx.** If the system does not have NI-DAQmx installed, install this now. The latest version as of this writing is 9.4, but in general the latest available version (available from the National Instruments website) should be installed. If the system is going to be used for developing code, be sure make sure the to install the development libraries as well. When installing, select custom installation, and
add the .NET 3.5 language support. (Note that MpScope 3.0 uses .NET version 3.5, since not all libraries are available for the later 4.0 version of the runtime.)

- **Install the Galil Controller Software.** Run the Galil .NET software installer, which allows MpScope to talk to the stages.

- **Install the HDF5 viewer (optional).** Although not strictly necessary, it is convenient to be able to open the data files to look at the contents and structure of the file. The HDFView program is a simple, java based program which allows viewing of the MpScope data files. The latest version can be downloaded from the web.

- **Install Visual Studio Express 2010 for C# (optional).** For code development, install Microsoft Visual Studio Express 2010.

### B.2.2 Hardware for MpScope

- **Main Scanning Card.** The PCI-6110 data card is standard, and is required to be installed as the main card. Note that other cards may work without modification of the code, but this has not yet been tested.

- **Auxillary Analog Input Card.** The PCI-6052 data card can be added for additional analog data acquisition. Again, other cards may work without modification of the code, but this has not yet been tested.

- **Photon Counting Card.** The PCI-6534 can be used for acquiring digital photon counting data. Again, other cards may work without modification of the code, but this has not yet been tested.

- **RTSI.** In order for the cards to synchronize correctly, they must be connected by a RTSI bus. This is a physical cable running between the cards, which connects the scan clock and data trigger, as well as other signals. After the cable is installed, it must be registered in NI-DAQmx. (Note, this step is new to MpScope 3.0, since it uses NI-DAQmx.) To set up the cable, open MAX (Measurement and Automation Explorer), right click on “Devices and Interfaces,” and select “create new”.

- **Galil Stages.** MpScope 3.0 is designed to work with stage controllers made by Galil Motion Control. Support for other stages may be added in the future.
B.3 Using the Program

B.3.1 Initial startup

On the first program startup, the program will load the default file configuration. Before actually scanning, the program must find the location of the data cards. Open the Configuration with the Configuration button, and select the “Cards” tab. Select “Search for Cards.” This should automatically select the correct cards.

The main card, which is used for analog output (to the scanners) and input (for analog data reads) is the PCI-6110.

The Auxillary Analog card is currently the PCI-6052, which is used for additional analog inputs (and will be used for additional outputs) The digital input card, is currently the which runs the photon counters, is correctly the PCI-6534.

Other cards may be added in the future, and may actually work “as is”. The above listed cards will be detected and selected automatically, if available. After selecting the cards and clicking “okay,” remember to save the default configuration.

B.3.2 User Configuration.

The user configuration contains most of the state variables and parameters of the program, such as frame size, pixel rate, which channels are active, etc. These parameters are all set from the configuration form. To set the configuration, press the configuration button, and set the correct values on the form. Press the okay button to use the new values, or cancel to not use the new values. Note, the configuration is not saved to disk at this point. The user configuration can be saved to disk by selecting “File → Save ...”, or when prompted on exit.

B.3.3 Using the Stages.

The stage motion works by clicking on any of the available viewers. To move 1/10th the distance, hold the shift key while clicking. In order to scroll in Z, make sure “scroll in Z” is selected for the mousewheel action, and use the mousewheel. Again, holding the shift key moves the distance by 1/10th the full distance.

Note that the mousewheel can also be used to rotate and zoom the image.
B.3.4 Scripting with MATLAB

Previously, scripting in MpScope was handled through a Visual Basic scripting interface. In MpScope 3.0, scripting is handled through MATLAB. This allows users to write more powerful scripts (for instance, integrating data processing features and other user MATLAB scripts), without having to lean on an additional programming language. Note, in order to use this feature, you must have MATLAB 2009a or later, which is the first version to contain the .NET interface.

Since MpScope 3.0 is written in C#, all the public static functions are automatically available to the outside world, including MATLAB scripts. In order to use them, the assembly (which is just the MpScope executable) must be loaded into MATLAB. Following is a simple script to demonstrate control through MATLAB:

```matlab
% query to make sure .NET works on this system
disp(['Today''s day, from .NET: ' num2str(System.DateTime.Now.Day)])

% add the assembly, to allow access to functions
% note, actual location of the .exe file will vary
MpScopeAssembly = NET.addAssembly('C:\MpScope.exe');

% functions are called using the normal dot notation
% for instance, try a simple function from the library
MpScope.MpLib.SquareNumber(5)  % returns 25

% start and MpScope session as a separate thread
MpScope.Program.RunAsThread();

% make sure this has time to load
pause(5)

% scanning can be initialize by running StartScan
% (this is equivalent to pressing the StartScan button)
MpScope.Engine.StartScan();

% wait for scanning to complete by polling
% note, this time can also be used to accomplish other
% calculations, such as data analysis of previous files
while( MpScope.Engine.IsScanning() )
    pause(.2)
end

% the scan can also be stopped manually:
```
An HTML file of all the available commands is available.

**B.4 Developer’s Guide**

MpScope was originally written by Quoc Nguyen for the Kleifeld laboratory in UCSD, and is used in dozens of laboratories around the world. MpScope is fast, stable, freely available to researchers.

In order to address some limitations of the original software, expand the capabilities, and create a easy to understand and use platform which can be expanded for future needs (for instance, 3d scanning, automation, etc.), the code has been completely rewritten as as MpScope3 (MpScope, version 3.0).

Many of these changes are internal, and transparent to the user. Below is an overview of the main new features. These will be discussed in more detail later in the document.

- **User Interface Improvements.** the User Interface is greatly simplified. Scanning cards are detected automatically, and the mouse can be used for navigation by clicking (for stage movement) and scrolling on the image (for zoom, rotation, and Z movement).

- **Scan Waveform Improvements.** Arbitrary high scanning resolutions are supported (beyond the previous 512 x 512 scan limitations). The scan path is now calculated using a user-defined mirror acceleration, which remains constant at different zoom levels; this allows consistent imaging at different resolutions and zoom levels, and also allows for cases where the mirror acceleration must be lower than standard (for instance, when driving large mirrors) or higher (for instance, with AOD systems) than standard.

- **Written in C#/.NET** MpScope was originally written in Delphi Pascal. The newer code uses C#, a programming language similar C++/Java, which is an extremely powerful way to code Windows Software. C# uses the .NET framework, the native and preferred was to interface with many hardware devices as well as Windows system calls. Moreover, the C# IDE, Visual Studio, has a number of
features which make programming easier (such as Intellisense and refactoring), and is available as a free download from Microsoft.

- **Small, clean code base.** The code base of MpScope3 is extremely compact, and well documented. Many of the features implemented in the original code have been implemented in a simpler, more compact way; for instance, user configuration data is written as an .xml file format, requiring only a few lines of code to save and retrieve the data, which is stored in a transparent and platform independent file format.

- **NI-DAQmx.** In order to use the latest data acquisition cards from National Instruments, the software now uses NI-DAQmx, rather than the older NI-DAQ. This allows for the faster, newer cards to be used; it also greatly simplifies the coding required to use the cards, since many of the complicated data acquisition features (such as double-buffering for fast continuous data acquisition) are now handled mostly by the NI-DAQmx drivers.

- **New File Format** The previous version of MpScope used a custom file format, the .mpd file. In order for platform independence and a simpler file format interface, MpScope now uses the HDF5 file format. This format has a large user and support base, and has been well established for rapid access to large, structured datasets for decades. It is completely platform independent, meaning that it can be used on 64-bit systems, Mac, etc., and is supported natively in MATLAB.

- **Scripting Through MATLAB** - One of the powerful features of MpScope is the ability to script the commands through Visual Basic. Since MpScope3 is built on .NET, publically accessible functions can be called through MATLAB, which is the preferred scripting language. It is possible that other languages with operate through .NET, such as IronPython, will also work, though this has not been tested.

B.4.1 Setting up and Using C#

C# was chosen as the programming language for MpScope3, since it offers a fast and powerful way of programming Windows applications. C# is probably the most widely used language to program Microsoft Windows applications, and uses familiar c-style syntax. C# uses managed code, which makes the code stable and easy to debug,
and allow for direct memory access (“unsafe” code) for performance critical sections, where necessary.

Other languages could have been used, but each has its own drawbacks. Java is platform independent, but correspondingly, accessing the hardware and low-level system calls (required for speed) must be done through wrappers. Java also cannot do direct memory manipulations operations. Similarly, MATLAB is extremely high-level, and has a correspondingly limited user interface options and programming capabilities, such as data structures and threads. C++ has no particular advantages over C# (which was designed to be an improved C++), has an awkward interface into .NET, and lacks native features such as dynamic allocation of multidimensional arrays, a native thread library, etc.

The newest version of the C# IDE (Integrated Development Environment) is Visual Studio 2010, and uses .NET 4.0 (backwards compatible with previous versions of .NET). The Express version, which still contains basically all of the functionality of the Professional version, can be downloaded for free from Microsoft. You must register each copy within 30-days, but the registration is also free.

After installing the program, it is recommended to change the following settings to work with the MpScope3 code. Note, many of these changes (such as the selection of .NET 3.5) have already been made in the MpScope3 project.

- To enable the complete menu options, go to Tools → Settings → Expert Settings.
- Similarly, Select Tools → Options, and select “Show all Settings.” To view line numbers in the code editor, go to Text Editor → All Languages, and select Display Line Numbers.
- MpScope3 uses “K&R” type formatting rules, in which the open brace character “{” appears at the end of the line, rather than on a separate line (the default setting). To keep the indentation consistent, go to Tools → Options ... the go to Text Editor → C# → New Lines, and uncheck all the “New line options for braces.”
- Many useful library compiled for .NET (such as the Galil and HDF5 libraries) have not been updated to use the latest version, .NET 4.0, which is the default for Visual Studio 2010. This can be changed to by selecting Project → MpScope Properties, and changing the Target Framework to .NET Framework 3.5.
For an overview of the C# language, many resources are available, such as “C# 4.0 in a Nutshell” by Albahari & Albahari. The sample chapter on Threading is available for free on the Internet. In addition, there are numerous programming notes and tutorials produced by Microsoft. Some useful ones are:

- Threading Tutorial
- Threading (C# Programming Guide)
- Events Tutorial, Events
- (C# Programming Guide), etc.

Following are some tips and tricks for getting started Visual Studio and C#.

- To create a new project Go to File → New Project ... and select “Windows Form Application.” Be sure to name the project appropriately, and then select “OK.”

- Note that renaming files happens intelligently. So, for instance, the file “Form1” can be renamed to something more descriptive, like “FormMain,” and VS will automatically rename references to this object (and the code will still compile). This is known as refactoring, and can be used to intelligently rename other variable or functions as well - right click on a name, and select Refactor → Rename ...

- The Project can be run by pressing “F5,” (or pressing the “play” button), and can be stopped by pressing “shift+F5”. This runs in debug mode which is useful for development. The code will run without the debugger, by pressing control+F5.

- To auto-format a selection, go to Edit → Advanced → Format Selection (or Format Document)

- To make a Release version (as opposed to a “Debug” version), Go to “Build → Configuration Manager,” and select Release.

- External .NET libraries are added to the project through “references” to other .NET assemblies. These are added though “References” in the Solution Explorer. For example, MpScope adds references and the NI-DAQmx, XML, HDF5, and the Galil controller. Once a reference is added, a “using” statement in the class allows the IDE and compiler to use the reference, without typing in the full path to the function.
B.4.2 Setting up and using NI-DAQmx

In order to use NI-DAQmx, the latest drivers must also be downloaded and installed.

For development, download the NI-DAQmx 9.3. Note that this version (9.3 or later) is required to run with .NET 4.0. However, since MpScope uses the .NET 3.5 framework, this should also be installed. When installing NI-DAQmx, select custom installation, and add the .NET 3.5 language support, as well as the Examples. Note that the examples are one of the central ways of learning to use NI-DAQmx, since the documentation is somewhat diffuse. The examples will install to “C:\Users\Public\Documents\National Instruments\NI-DAQ\Examples”.

To use NI-DAQmx in C#, You must included the line “using National Instruments.DAQmx;” and add the references to NI-DAQ. To add the reference, in the Solutions Explorer, right click the project, go to “add a reference,” and select “National Instruments DAQmx” and “National Instruments Common.” Do not add “Common Native,” this will result in a compile error.

Double Buffering, and Asynchronous Callbacks

The most algorithmically difficult part of using the cards is the continuous acquisition and processing (descanning the frames, writing to file) of the data. This is accomplished by NI-DAQmx using a behind-the-scenes circular buffering scheme, documented in Waveform and Single-Point I/O in NI-DAQ 7 (and later versions).

NI-DAQmx uses delegates to control callbacks from the card. To understanding how these work check out the ms example. When a delegate is constructed, it takes the name of the function to do the processing in the constructor.

Synchronizing across the RTSI bus

In order for the various DAQ cards (the main card, analog input card, and digital counting card) to synchronize their start, and use the same clock, the RTSI cable must be physically installed and configured in max. See the NI RTSI overview for more information. To set up the cable, open MAX, right click on “Devices and Interfaces,” and select “create new”
More NI-DAQmx Documentation

The documentation for NI-DAQmx is somewhat diffuse and can be difficult to find.

There are a number of Examples, which can be installed when the drivers are installed.

Learn 10 Functions in NI-DAQmx and Handle 80 Percent of Your Data Acquisition Applications– a good initial overview of the NI-DAQmx system

Using NI-DAQmx in Text Based Programming Environments - a good general overview of programming in C# (and Visual Basic)

Programming NI-DAQ in Visual C# .NET

Text Based NI-DAQmx Data Acquisition Examples - includes additional examples, not installed with NI-DAQ

Timing and Synchronization Features of NI-DAQmx

Much of the main documentation is installed in a help file, which is available from the Start Menu. “C:\ProgramData\Microsoft\Windows\Start Menu\Programs\National Instruments\NI-DAQ\Text-Based Code Support” contains a hyperlinked help file “NI-DAQmx \version\ .NET Framework Help”, containing most of the API Information, etc. Note that National Instruments steers users towards using the “Measurement Studio,” which is a separate package (available for a fee). MpScope3 does not use this, sticking to the core classes instead.

B.4.3 Automatically Generated Documentation (Doxygen)

Standard XML tags to add internal documentation to the code are becoming more standard. For instance, typing “///” into Visual Studio IDE will automatically generate a stub for a “summary” comment,

/// <summary> comments look like this </summary>

A number of parsers are able to look through all the source file, and assemble these comments, along with the function stubs, into usable documentaion (such as html, microsoft help files, or RTF document files). One of the most popular and easy to use tools is doxygen. Doxygen for windows has a helpful user GUI, which allows for an easy way to customize what the help files look like. The MpScope3 release contains a settings file for Doxygen, which is an easy way to create documentation of all the publicly available function calls, as well as any comments associated with them.
B.4.4 Setting up and using HDF5 (file format)

The file format in MpScope 2.0 and earlier was based on the Microsoft Streams file access implementation. Although this approach has a number of advantages (fast, and fully supported on Windows machines), there are a also number of disadvantages. The file write and read (MPFILE.OCX active X control, and the Mp2Mat.m programs) must be separately maintained. This adds to code complexity, and potential errors. Similarly, the file format does not currently work with Windows 64-bit (since it would require a 64 bit active X control), or other operating systems, such as Mac or Linux. The solution is to use one of the established formats for reading large data sets. There are several possibilities, such as CDF, netCDF, FITS, and HDF5. All of these are commonly available, and supported by MATLAB.

The HDF5 file format was chosen since it is widely supported (available for many platforms, has been ported to .NET, and is natively supported in MATLAB), well maintained, easy to use, and fast. HDF5 is was created by the National Center for Supercomputing Applications (NCSA) at the University of Illinois. The data can be written in whatever format is convenient (multiple 2d frames, 1d data, using doubles, bytes, etc.), making it easy to mix imaging, electrophysiology, configuration data together in the same file.

Note, Another close contender is CDF / netCDF. In the netCDF vs. HD5 debate, a number of posts point to HDF5 as being superior, due to the hierarchical nature of storage (grouping), and the built-in compression support. (Newer versions of netCDF implement these features, by using a HDF5 format.) See Comment on HDF5 vs netCDF and Another comparison.

HDF5 files can have extensions hdf, h4, hdf4, h5, hdf5, he4, he5. MpScope3 uses “.h5” Mathematica will import .h5, but not cdf. To simplify, all MpScope3 files use the “.h5” extension.

B.4.5 HDF5

Tools to use HDF5 with C# are available here. Note that the test programs were written with Visual Studio 2005 - later version of VS will import the projects with no problem, though the project must be set to target .NET 3.5.

If you run the test program (CSharpExample1) as-is, it will crash with the note: “The specified module could not be found. (Exception from HRESULT: 0x8007007E)”
This is due to the supporting files (the HDF5 libraries) not actually being installed in a location where they can be found by the program!

The binaries can be downloaded from here. For a program to compile, a reference to “HDF5DotNet” must be included. In addition, the following files (libraries) must be added to C:\Windows\System32\ hdf5dll.dll, hdf5dll.lib, hdf5_hldll.dll, hdf5_hldll.lib, zlib 1.dll, zlib 1.lib, szip.dll, szip.lib ... from the \hdf5DotNet\HDF5DotNet\Release folder into the Windows\System32 folder. Not having these will result in a run-time error.

For 64-bit windows, you must copy the HDF5DotNet.dll, hdf5dll.dll, szlibdll.dll, zlib1.dll, msvcr71.dll, msvcr90.dll into the SysWOW64 directory, instead of the system32 directory. This has not yet been tested.

There is a free HDF5 viewer, hdfview, which allows browsing and editing HDF files. HDFView is JAVA based ... note that occasionally on launch, it will generate a “LaunchAnywhere Error Windows Error 6 occurred while loading the Java VM” error. This is not an error in the underlying HDF file.

An HDF5 group is analogous to a file system directory

In HDF, there are three classes of named objects: group, dataset, and named datatype. An HDF5 group is analogous to a file system directory.

The datatype are analogous to “double,” “byte,” etc.

The datasets contain the actual data. These can be a single value (a scalar), an 1d array, or a matrix of 2 or higher dimensions (such as imaging frames).

The HDF5 dataspace describes the layout of the elements of a multidimensional array. Conceptually, the array is a hyper-rectangle with one to 32 dimensions. A typical dataspace would be a frame, for instance. The HDF5 datatype object describes the layout of a single data element

Attributes - Any HDF5 named data object (group, dataset, or named datatype) may have zero or more user defined attributes

The basic steps for writing an HDF file are on p. 21 of the User’s Guide, section 2.1.

- Create the file
- Create a dataset
  - Create and initialize a dataspace for the dataset
  - Define a datatype for the dataset
– Create and initialize the dataset

Note that HDF5 uses C storage conventions, assuming that the last listed dimension is the fastest-changing dimension. This is the same as C#.

The File Format Data in an HDF5 file is arranged in Groups (essentially, folders), and these groups may contain attributes. The Groups may contain other groups, or contain data (like arrays, etc.).

Since data is acquired in chunks, corresponding to frames of the image, data is also stored as a series of these frames. This is true both for imagining data, which is written frame-by-frame, as well as the analog data, which is written in “frames” as well.

Each frame in the Analog data sequence spans the same time as the imaging frame. That is, if Imaging frames are taken 200 ms apart, then the analog data frames would be 200 ms apart as well. (Typically, however, they will be downsampled from the full scanning speed, and thus contain fewer samples per frame. When reading the file, the analog frames can be strung back together for a continuous time sequence, or left as separate frames.

In general, the .h5 files generated by MpScope3 have the following format: /Config - the config group does not actually contain data, but rather has the scanning attributes attached to it.

/Image - the main imaging data, up to four channels Ch1 Ch2 .... /AnalogMain - Analog data, taken from the main scanning board. Note, data on the National Instruments board is acquired as doubles, and then converted to shorts (signed 16 bit numbers), for compact storage. To convert back to a voltage value, the samples can be multiplied by the “ChannelPrecision” attribute. Ch1 Ch2 ...

/AnalogAux - Analog data, taken from the auxiliary analog input board. The format is similar to that for the data taken by the main board.

B.4.6 Setting up and Using the Galil Controller

To use the Galil controllers, GalilTools must be installed - this installs all of the necessary .dll’s as well. Software can be downloaded from Galil. For the .NET framework, download DMC .NET API Framework 2 (VS2005), which supports all the Visual Studios. Instructions are in the current app note.

Note that these target Visual Studio 2005 - however, they will work fine with Visual Studio 2010, as long as the project is set to use .NET 3.5 (rather than 4.0).
Creating a .NET Application without Measurement Studio


To get to the same help topics from within Visual Studio, go to Help Contents. Select Measurement Studio from the Filtered By drop-down list and follow the previous instructions.

Photon Counting The photon counting in MpScope 2.0 currently uses the NI PCI-6534. The newer version of this board, recommended by NI, is the NI PCIe-6536 (25 MHz) or the NI PCIe-6537 (50 MHz).

The appropriate cable for the PCI-6534 is SH68-68-D1 Cable, which is available in .4, 1.0, 2.0 and 5.0 meter lengths. For the NI PCIe-6536, use the Note that this is unshielded - they also make shielded cables (SHC68-H1X38), but these are much more expensive. (Seems the limit is 2M on these?)

**Turn Calculations - Y scan mirror**

For the Y turn, it’s easiest to calculate for the turn starting at \( t = 0 \), and lasting for a time \( T \). The mirror position at the start of the turn is 0, and \( Y \) (the frame height) at the end. Note that in this case, the mirror velocity will be \( -V \).

\[
P(t = 0) = 0 \text{ (starts at point 0)} \quad P'(t = 0) = V \text{ (y scan velocity at the start)}
\]

\[
P(t = T) = Y \text{ (full Y height at the end of the turn)} \quad P'(t = T) = V \text{ (y scan velocity at the end)}
\]

This can be solved with a spline:

\[
P(t) = a + bt + ct^2 + dt^3 \quad (B.1)
\]

With the first two constraints (initial position and velocity), this becomes:

\[
P(t) = Vt + ct^2 + dt^3 \quad (B.2)
\]

Adding the second two constraints (final position and velocity) gives
\[ Y = VT + cT^2 + dT^3 \quad \text{(B.3)} \]

(final position) and

\[ V = V + 2cT + 3dT^2 \quad \text{(B.4)} \]

(final velocity)

These can be solved to give:

\[ c = -\frac{3dT}{2} \quad \text{(B.5)} \]

(from final velocity equation)

\[ d = \frac{2(VT - Y)}{T^3} \quad \text{(B.6)} \]

(substituting in \( c \) to final position equation)

\[ P(t) = Vt + ct^2 + dt^3 \quad \text{(B.7)} \]

**B.4.7 Servo Power Control**

The power can be controlled by a servo, typically the HS-785. This is controlled by the counter1 output on the main board. Note, from the Hitec site: “All Hitec servos require a 3-4V peak to peak square wave pulse. Pulse duration is from 0.9ms to 2.1ms with 1.5ms as center. The pulse refreshes at 50Hz (20ms).”

**B.4.8 Theading and Program Execution in C#**

MpScope is a type of program knows as a “Windows Forms Application.” Like most programs, at initialization there is only one execution thread. This is true, even if the program spawns off multiple child windows (like Viewers, or the Configuration form for instance).

In a forms application, after initializing, application will just wait for user input. This could be in the form of a button press, or an action such as a mouse over. Whenever such an action occurs, an event is placed on a message queue to execute the function (such as a button press handler), and the thread begins processing that function. If another event occurs (such as a different button press, or a resize, etc.) while that function is
executing, the function is not preempted. Rather, after the function exits, the message queue will execute the next event in sequence.

The advantage of this execution scheme is that nothing in the application has to be thread safe. A button press handle in a child window can access the same data as the main window, without the possibility of a “torn read” or other errors that can occur in a multithreaded application. Keeping application mostly single-threaded makes development and debugging much, much easier.

The disadvantage is that sometime a function which is takes a long time to execute can cause the user interface to become unresponsive. For instance, if a button handler contains the call “System.Threading.Thread.Sleep(5000)”, then the UI will become unresponsive for 5 seconds, and will not draw updates not handle any additional button presses. For this reason, it is occasionally useful to spawn off different threads, but frequently this can be handled with a Windows Timer.

The Timer class (System.Windows.Forms.Timer) allows a function to schedule a function to be run at some later time. This can be useful, say, when a stage update must occur at some later time (i.e., after the stage has moved). Scheduling a timer essentially adds an event to the message queue, and is similar to just calling the function at a later time. Note that timers will not result in a preempting of the current thread, since they do not run in additional threads.

As a subtle point of confusion, there is an additional type of timers other than the Forms Timer, such as “System.Threading.Timer”. These are multithreaded, and will not be discussed here.

Scripting in MATLAB - note that the situation is similar with MpScope is being run from MATLAB. Only one function in MpScope can run at a time, so the whole program is intrinsically thread-safe. However, be careful when running from BOTH the GUI and MATLAB! In this case, since both MATLAB and the GUI are running separate threads, it is possible to call functions in an thread-unsafe matter, for instance by clicking a button and then issuing the same command in MATLAB.
Bibliography


