

NORTHEASTERN UNIVERSITY

UNDERGRADUATE THESIS

**An Improved Method for Real-Time
Ratiometric Quantitative Fluorescent
Microscopy in the Pharynx of *C. elegans***

Author:
Sean Johnsen

Supervisor:
Dr. Javier Apfeld

*A thesis submitted in fulfillment of the requirements
for the degree of Bachelor of Science*

in the

Department of Biology

April 19, 2019

“...the grandest discoveries of science have been but the rewards of accurate measurement and patient long-continued labour in the minute sifting of numerical results.”

Lord Kelvin

NORTHEASTERN UNIVERSITY

Abstract

Faculty Name
Department of Biology

Bachelor of Science

**An Improved Method for Real-Time Ratiometric Quantitative Fluorescent
Microscopy in the Pharynx of *C. elegans***

by Sean Johnsen

The abstract goes here...

Acknowledgements

The acknowledgments and the people to thank go here...

Contents

Abstract	ii
Acknowledgements	iii
1 Introduction	1
1.1 Fluorescent biomarkers enable real-time quantification of cytosolic protein oxidation	1
1.2 Limitations of current pipeline	2
1.2.1 Inter-frame movement results in measurement error	2
1.2.2 Segmentation and centerline estimation sometimes require manual input	3
1.3 Aims	4
2 Methods	5
2.1 Edge information helps to differentiate the pharynx from the intestine	5
2.2 Flexible measurement boundaries account for anterior-posterior movement	6
2.3 Midlines	7
2.3.1 A transmitted light image helps to anchor the midlines	7
2.3.2 A spline-based centerline algorithm improves quality around the tip	8
2.3.3 Dorsal-ventral tip movement is addressed with frame-specific midlines	8
2.4 Channel Registration picks up the slack	8
3 Results	9
3.1 Definition of Error	9
3.2 Collection of test data	9
3.3 A partially synthetic data set increases statistical power	10
3.4 Reductions in Manual Intervention	11
3.4.1 Incorporation of edge information increases segmentation accuracy	11
3.4.2 Incorporation of transmitted-light information increases centerline stability around the posterior bulb	11
3.5 Channel-Specific Masks and Midlines Reduce Error	11
4 Discussion	12
4.1 Future Directions	12
4.1.1 Automatic movement detection	12

List of Figures

1.1	Ratios of images to redox state	2
1.2	Pharyngeal contractions lead to boundary issues	2
1.3	A ratio image reveals inter-frame movement	3
1.4	The goal of segmentation	3
1.5	The problem with static thresholding	4
2.1	Overview of the improved segmentation algorithm	5
2.2	Overview of the improved segmentation algorithm	6
2.3	Flexible measurement boundaries	7
2.4	Previous centerline estimation algorithm	7
3.1	Test imaging schematic	9
3.2	Error by strategy in the natural data set	10
3.3	Average intensity over time in technical replicates	10
3.4	Error by strategy in the synthetic data set	11

Chapter 1

Introduction

Oxidation and reduction (redox) reactions play a vital role in biology. These reactions are characterized by a flow of electrons between chemical species. The species gaining electrons is said to have been oxidized, while the species losing electrons is said to have been reduced. Redox reactions are well-known to be involved in cellular respiration and metabolism more generally. More recently, redox control of the proteome has been the subject of great interest, mediating such events as differentiation and apoptosis. The oxidation of cysteine thiols to disulfide bonds provides an interface between redox chemistry and proteomic signal transduction.

Many chemical species in the cell may exist in either an oxidized or reduced form. These pairs, termed *redox couples* facilitate electron transfer in a variety of cellular processes. The NADH/NAD⁺ couple, for example, transfers high energy electrons wrought from the oxidation of sugars in the citric acid cycle to O₂ via oxidative phosphorylation.

The healthy cell actively maintains a steady-state disequilibrium of each of these redox couples. The "redox hypothesis" attributes aging and age-related disease to the misregulation of this network. Quantifying redox chemistry in live cells will allow for a deeper understanding of the regulatory mechanisms that mediate these vital processes.

1.1 Fluorescent biomarkers enable real-time quantification of cytosolic protein oxidation

In a previous paper, our lab demonstrated the use of the genetically-encoded redox-sensitive fluorescent protein roGFP1_R12 to quantify the cytosolic redox state in the pharyngeal muscle of *C. elegans* in real-time. These redox biosensors are leading to quantitative and mathematical models of the genetics and dynamics of intracellular and organismal-level redox signaling.

A key feature of the roGFP family of sensors is their ratiometric nature, which corrects for differential expression of the sensor, as well as cell thickness (Hanson 2004). Because the oxidized and reduced form of the protein have differing emission spectra, the emission intensity at differing excitation wavelengths may be used as estimates of concentration in the Nernst equation to estimate the sensor's reduction potential. Thanks to the transparent nature of *C. elegans*, fluorescence micrographs reveal spatial reduction potentials at subcellular resolution (Figure 1.1).

Due to the cellular architecture of the pharynx, almost all of the variation in redox state follows the posterior-anterior axis. Thus, we can model the pharynx 1-dimensionally as the redox state along its posterior-anterior axis without losing

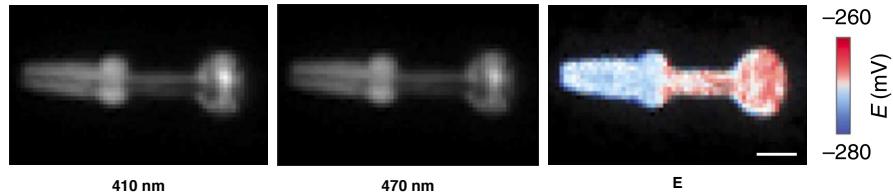


FIGURE 1.1

significant information. Computationally, this is achieved by (1) estimating the centerline of the pharynx then (2) measuring the intensity of the images under this estimated centerline.

As will be discussed, a fundamental limitation of this analysis arises when the animal moves in the time between capturing the first and second frame in the pair of images required for each animal.

1.2 Limitations of current pipeline

1.2.1 Inter-frame movement results in measurement error

The pharynx of *C. elegans* acts as the animal's feeding muscle. It contracts along its anterior-posterior axis, functioning as a pump to bring in food. This contraction poses a problem for ratiometric analysis. Ordinarily, dividing intensities pixel-by-pixel is appropriate because pixel-coordinates in both images capture the same point in physical space. If the animal moves, however, the same point in space is represented by different pixel coordinates in the first and second image. To understand why, consider Figure 1.2.

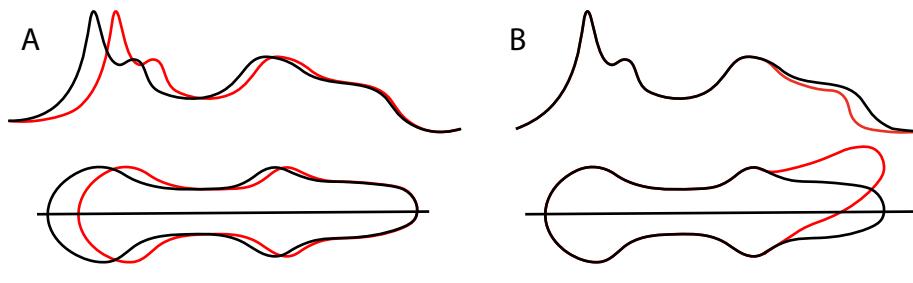


FIGURE 1.2: A cartoon shows the two most common modes of inter-frame movement. **A** Posterior bulb contractions lead to a shrink or stretch of the intensity profile. **B** Dorsal-ventral movement of the anterior tip leads to premature truncation of the intensity profile as the midline does not follow the curve.

On the bottom we see the outline of the pharynx in each frame. The pharynx has contracted in one frame (red) and is elongated in the other (black). When the intensities along the posterior-anterior axis are plotted above, it is clear that the contraction has lead to unwanted compression of the intensity profile.

Another common mode of interframe movement is represented similarly in Figure 1.2. This is movement of the tip of the pharynx dorsal-ventrally. Dorsal-ventral tip movements result in a loss of information about the tip in one frame, as depicted in red.

During imaging, animals are paralyzed with the acetyl choline agonist levamisole. However, complete paralysis is not possible due to concentration limitations. Inter-frame movement is currently addressed by visually screening the ratio images for a highly textured appearance, as shown in Figure 1.3. These animals are excluded from analysis. However, visual screening is a time-intensive process, requires training, and is subject to experimenter error and bias.

talk about limitations

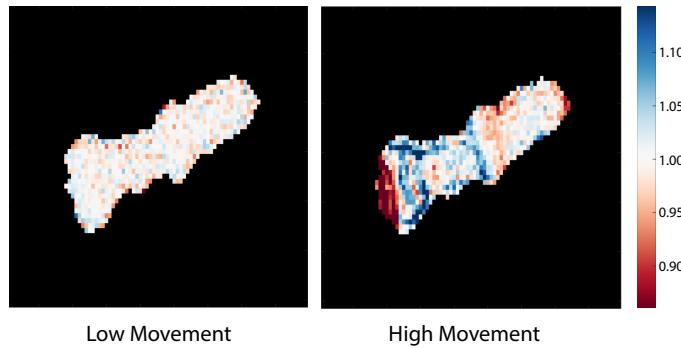


FIGURE 1.3: Ratio images reveal inter-frame movement. Both images are the result of dividing one image taken at 410nm with a second at the same wavelength. The animal on the left did not move between frames while the animal on the right did.

1.2.2 Segmentation and centerline estimation sometimes require manual input

As noted, the visual screen for inter-frame movement is a manual step. Two other processes in the current pipeline also require manual supervision. The first is segmentation. Segmentation is the process by which pixels corresponding to objects in an image are given salient labels. In this case, the task is to separate the pharynx from everything else (Figure 1.4).

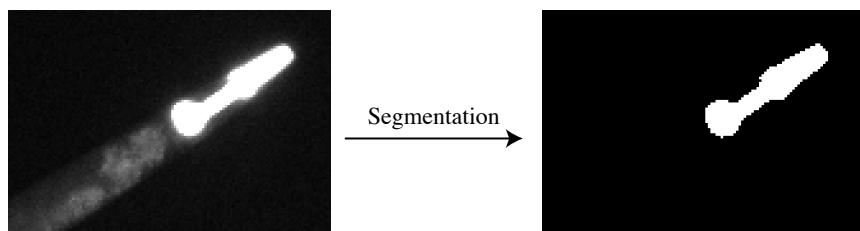


FIGURE 1.4: The goal of segmentation. On the left, the original image showing fluorescence of roGFP in the pharynx and autofluorescence in the gut. On the right, a binary image consisting of value 1 in the pixels with a pharynx and value 0 elsewhere.

Because the transgenic animals express roGFP with pharynx-specific promoters, most fluorescence occurs only in the region of interest. However, the intestine of *C. elegans* exhibits autofluorescence in response to the wavelengths of light required for roGFP excitation. This poses a problem for the static thresholding algorithm currently used to segment the pharynx. This algorithm assigns any value greater than a threshold the value 1 and those below the threshold 0. Static thresholds work well when the distribution of brightness is different for each class of object in the

image, but this is not the case when the intestine autofluoresces, resulting in ill-formed segmentation masks (Figure 1.5).

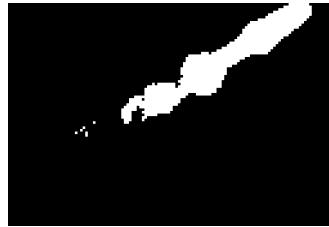


FIGURE 1.5: The problem with static thresholding. Autofluorescent intestinal tissue results in a segmentation mask that must be manually corrected.

As described in 1.1, the 1-dimensional model of the pharynx requires the accurate estimation of the tissue's centerline. While the current centerline estimation algorithm (described in 2.3) works well in the interior regions of the pharynx, it is unstable around the terminal regions, both the anterior and posterior. These portions of the centerline must be visually inspected and corrected as well.

figure of midline that needs correction

1.3 Aims

This thesis aims to address the limitations of the current analysis pipeline described in 1.2.1 and 1.2.2. To achieve this, a new pipeline was written in MATLAB to process and analyze these images end-to-end with little to no necessary manual input. The improved pipeline decreases the time required to analyze this data, standardizes the analysis, reduces human error, and mitigates movement-induced error.

Chapter 2

Methods

2.1 Edge information helps to differentiate the pharynx from the intestine

As described in 1.2.2, intestinal autofluorescence causes the static thresholding algorithm to perform insufficiently in separating the pharynx from the rest of the animal and the background. Histogram-based threshold detection such as Otsu's method also perform poorly when the histograms are not bimodal, as is the case with bright intestinal autofluorescence.

To address these issues, a new segmentation algorithm was developed. This algorithm exploits *edges*, sharp changes in brightness which are information-rich regions in images. The Sobel operator creates an image emphasizing edges by discretely differentiating the image in the vertical and horizontal directions (Figure 2.1). By thresholding the resultant image, using morphological operations to remove noise, and selecting the largest region, we isolate the pharynx from gut autofluorescence.

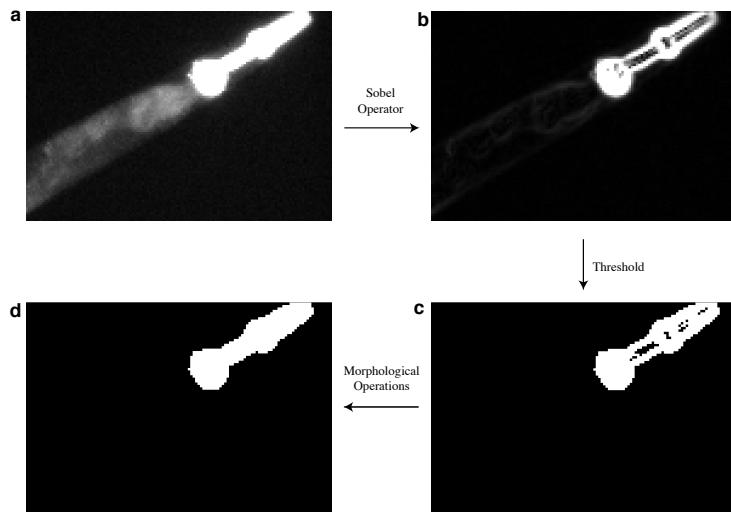


FIGURE 2.1: An overview of the improved segmentation algorithm. **a** The original image, displaying fluorescence of roGFP in the pharynx and the autofluorescence of the gut. **b** The resultant image after application of the sobel operator. **c** The resultant binary image after thresholding the edge-emphasized image. **d** The final segmented binary image, after performing morphological cleaning operations.

For reasons described in 2.3.1, an image taken with transmitted light must also

be segmented. Even though the distribution of intensities in transmitted light images are very different from fluorescence images, this edge-based segmentation algorithm still performs robustly (Figure 2.2). This highlights another benefit of the improved method: image brightness independence. If one particular genotype fluoresces dimly, an edge-based segmentation method still performs well, while the static thresholding method requires the manual change of the threshold parameter.

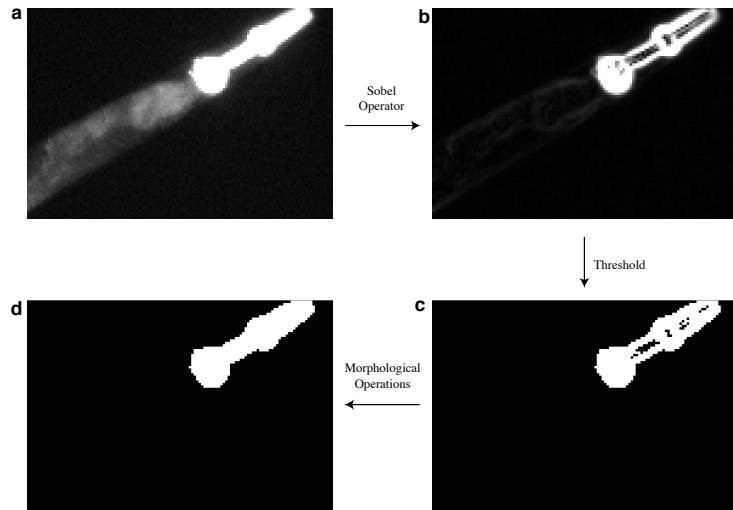


FIGURE 2.2: This should be an image of transmitted light before/after segmentation.

2.2 Flexible measurement boundaries account for anterior-posterior movement

As described in 1.2, interframe movement is a large source of error. To understand why, consider the how the previous pipeline takes measurements. First, a single mask using the image taken at 410nm is computed. This mask is then applied to both the 410nm and 470nm images. Pixels corresponding to 1 in the mask retain their value while pixels corresponding to 0 in the mask are assigned the value 0. Measurements are taken of the masked images. Thus, if the animal pumps, measurement might start in the gut or halfway through the posterior bulb, depending on if the animal contracted during the first or second frame. If the animal contracts during the first frame and extends in the second, the mask will be too "short" and measurement in the second frame will start in the middle of the posterior bulb. If the animal is extended in the first frame and contracts in the second, the mask will be too "long" and measurement in the second frame will start in the gut. This error is accounted for via visual inspection of ratio images so as to exclude those animals who have moved from analysis.

The new approach (summarized graphically in Figure 2.3) tries to account for this movement without the need for manual exclusion. Here, segmentation masks are not applied to the images before measurement as before. Instead, they are used only in the calculation of the midlines (described in 2.3). Intensity profiles of the unmasked images are taken under the midlines. The intensity profiles are trimmed to the point where they cross a threshold then resized to the same length via linear

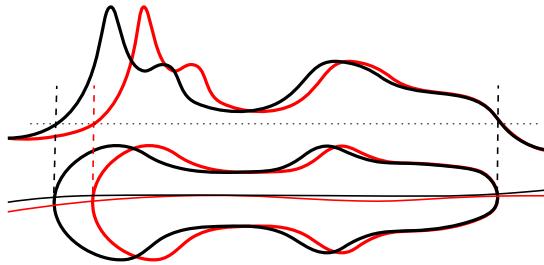


FIGURE 2.3: A cartoon depicting the flexible measurement boundaries in the revised measurement pipeline

interpolation. Allowing measurement boundaries to move flexibly is a simple yet effective change in reducing error introduced by inter-frame pharyngeal contractions.

2.3 Midlines

The previous centerline estimation algorithm takes as input an image of the segmented pharynx aligned horizontally along its anterior posterior axis and works as follows (depicted in figure 2.4).

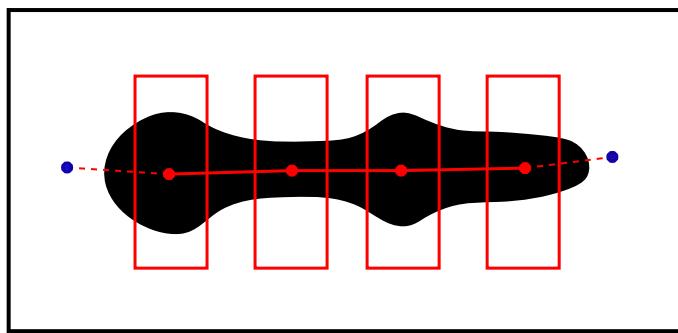


FIGURE 2.4: Cartoon representation of the previous centerline estimation algorithm.

Boxes are drawn at specific static coordinates. The centroid of each shape is calculated (red points). Lines (solid) connect these points. The y-coordinates of the terminating points (blue) are determined via the point-slope method $y = mx + b$ where m is the inverse of the slope of the neighboring line, x is a fixed constant and b is the y-coordinate of the neighboring point.

2.3.1 A transmitted light image helps to anchor the midlines

A major problem faced by the current centerline estimation algorithm is instability around the posterior bulb. Small positional changes in the neighboring points cause the terminal point to move dramatically. These instabilities must be manually corrected.

To address this concern, the new centerline estimation algorithm incorporates positional information from a transmitted-light image. The entire body of the worm is visible in transmitted light. We segment the transmitted light image with the algorithm described in 2.1, resulting in a binary mask. This mask is combined with

the pharyngeal mask by cutting off the transmitted-light mask when the pharyngeal mask starts. It is this combined mask that is fed into the centerline estimation algorithm. The points from the transmitted-light mask serve to constrain and anchor the previously unstable posterior-bulb region of the estimated centerline.

2.3.2 A spline-based centerline algorithm improves quality around the tip

The centerline estimation around the tip of the pharynx is subject to instabilities for similar reasons as discussed in 2.3.1. In this case, the transmitted-light frame does not help as the body of the worm does not extend past the pharynx. To address this tip instability, a new centerline estimation algorithm was created.

The method takes as input a binary image created via the methods described in 2.3.1. Pixels with the value 1 are treated as points in 2D space. A smoothing B-spline is fit to these points, resulting in a continuous function. Spline parameters were chosen by visual inspection.

2.3.3 Dorsal-ventral tip movement is addressed with frame-specific mid-lines

We saw in 2.2 a strategy to combat the major mode of inter-frame movement, contractions of the posterior bulb. However, there is another less common mode: dorsal-ventral movement of the tip. The spline algorithm described above allows the centerline to follow the curve of the tip without human intervention. To address inter-frame movement, the new pipeline draws a midline independently in each frame. These frame-specific midlines are used to take profile measurements.

2.4 Channel Registration picks up the slack

The frame-specific midline approach discussed in 2.3.3 introduces problems of its own. Specifically, the length of the measurement vector may be stretched nonlinearly due to differences in the arc length of each midline. That is, some sections of the intensity profile measured under these lines might be stretched while others are compressed. To approach these nonlinear stretches and compressions, we utilized a functionalized version of the dynamic time warping algorithm.

explanation of fda registration

Chapter 3

Results

3.1 Definition of Error

To assess the degree to which changes in the analysis pipeline affect the quality of measurement, a definition of error must first be decided upon. Error (ϵ) is defined to be a function over pharynx space (p) which is the difference between the intensity measured in one frame and another. The difference is normalized by the average intensity of the two frames at 410 nm.

$$\epsilon(p) = \frac{I_{410_1}(p) - I_{410_2}(p)}{\frac{I_{410_1}(p) + I_{410_2}(p)}{2}}$$

To calculate redox potentials, we take the first frame at 410 nm and the second at 470 nm. When we quantify error, however, we must use a single wavelength as the changes in emission spectra would confound the errors.

3.2 Collection of test data

A large number of technical replicates were collected. A single animal was plated on 5.5 mM levamisole and imaged 57 times in pairs of two images both at 410 nm. A paired-sample t-test indicates no significant difference ($p=0.41$) between the first and second intensity profile.

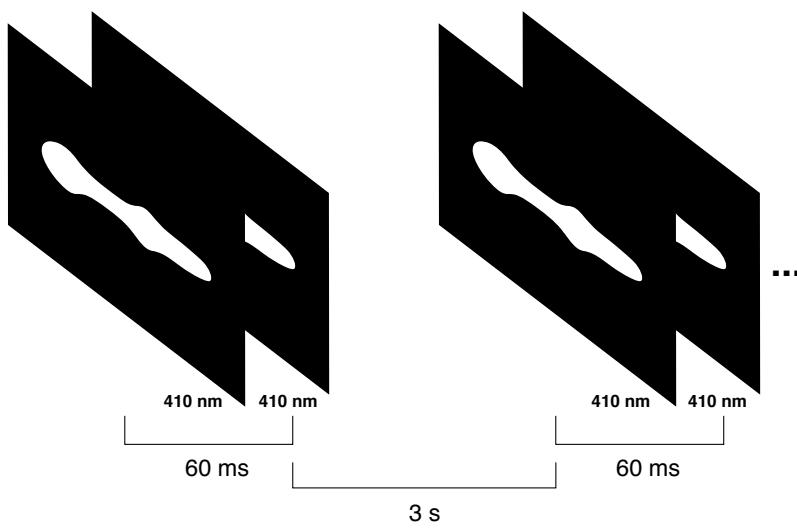


FIGURE 3.1

A set of ratiometric images (collection described in 3.2) was hand-classified according to the degree of movement (0 - 3) in four regions: the posterior bulb, anterior bulb, sides of the tip, and tip. Figure 3.2 shows the mean I_{410}/I_{410} measured under the midline and colored by degree of movement in the posterior bulb. A ratio of 1 indicates no error.

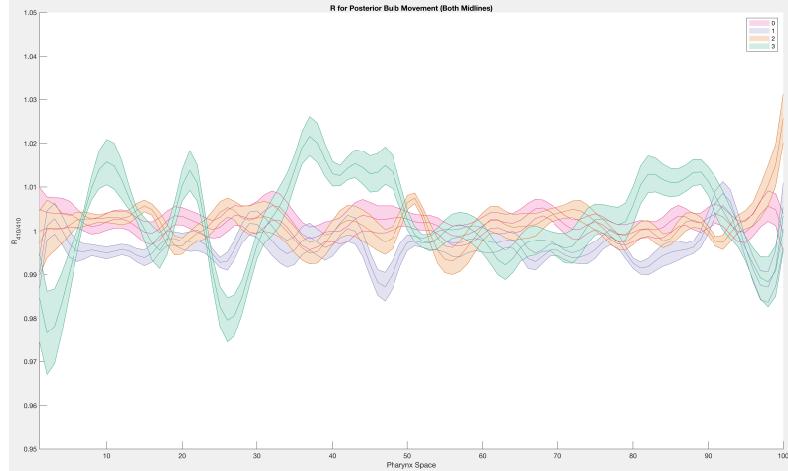


FIGURE 3.2: Relative error

3.3 A partially synthetic data set increases statistical power

Instead of comparing pairs of measurements taken back-to-back, we can compare all possible pairs. Given n images, we generate $\binom{n}{2} = \frac{n!}{2!(n-2)!}$ pairs. However, repeated excitation seems to attenuate the emission amplitude of the sensor (Figure 3.3). To account for this effect (known as photobleaching) we generate pairs starting at the 15th imaged pair and exclude outliers (Figure 3.3).

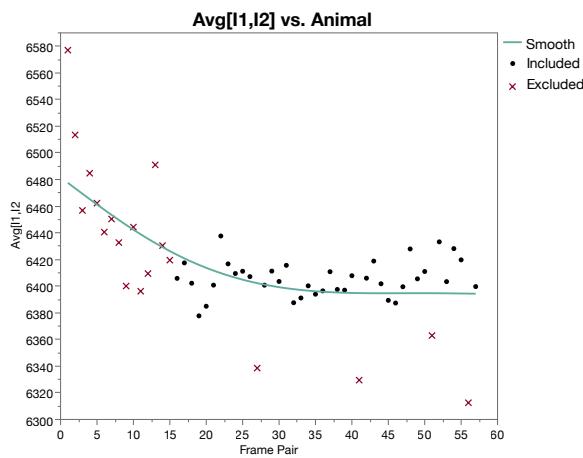


FIGURE 3.3: The average intensity over time in technical replicates.
Red xs mark image pairs excluded from the recombination.

Since the average time between frames in this synthetic data set is 10 seconds, there is a large degree of inter-frame movement. Because the true test data has only do this calculation

12 pairs which display inter-frame movement, this partially synthetic data set allows us to be more confident about the changes in error seen using different analysis strategies.

real number

3.4 Reductions in Manual Intervention

3.4.1 Incorporation of edge information increases segmentation accuracy

To compare the accuracy of segmentation methods, we first create a "gold-standard" set of images which were segmented by hand to ensure accuracy. The segmentation methods were applied to the same set of images and the resultant masks were compared via the total difference measure described in Yang et al. (A Supervised Approach to the Evaluation of Image Segmentation Methods).

show figures comparing % requiring manual intervention for segmentation and for midlines

3.4.2 Incorporation of transmitted-light information increases centerline stability around the posterior bulb

To compare the stability of the centerline around the posterior bulb, we

3.5 Channel-Specific Masks and Midlines Reduce Error

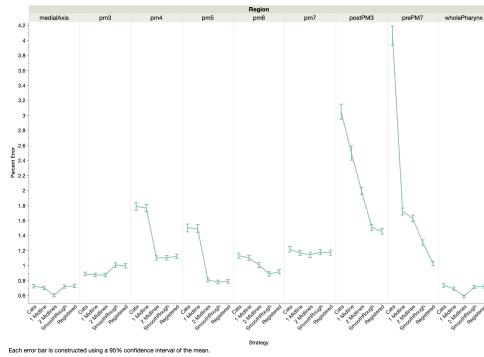


FIGURE 3.4

Chapter 4

Discussion

4.1 Future Directions

4.1.1 Automatic movement detection

This thesis has focused on the mitigation of the error introduced by inter-frame movement. As noted, not all error can be removed by these methods. As such, it would be useful for an automated classification system to detect inter-frame movement. The fundamental difficulty with this task is that in an experimental setting the image pairs consist of one image at 410 nm and another at 470 nm. Because the excitation spectra of roGFP is different at each wavelength, it is difficult to determine if the differences in each image's intensity profile is due to movement or true biochemical activity.