RECONSTRUCTING IMAGES FROM *IN VIVO* LASER-SCANNING MICROSCOPE DATA

by

Mihaela Obreja

B.S. in Mathematics, University of Bucharest, 2000

M.A. in Statistics, University of Pittsburgh, 2007

Submitted to the Graduate Faculty of

the Department of Statistics

School of Arts & Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF STATISTICS

SCHOOL OF ARTS & SCIENCES

This dissertation was presented

by

Mihaela Obreja

It was defended on

November 19, 2010

and approved by

William F. Eddy, Ph. D., Professor, Department of Statistics,

Carnegie Mellon University

Satish Iyengar, Ph. D., Professor, Department of Statistics

Leon J. Gleser, Ph. D., Professor, Department of Statistics

Henry W. Block, Ph. D., Professor, Department of Statistics

Justin C. Crowley, Ph. D., Professor, Department of Biological Sciences,

Carnegie Mellon University

Dissertation Advisors: William F. Eddy, Ph. D., Professor, Department of Statistics,

Carnegie Mellon University,

Satish Iyengar, Ph. D., Professor, Department of Statistics

RECONSTRUCTING IMAGES FROM *IN VIVO* LASER-SCANNING MICROSCOPE DATA

Mihaela Obreja, PhD

University of Pittsburgh, 2010

Two-photon laser-scanning microscopy can be used for *in vivo* neuro-imaging of small animals. Due to the very high resolution of the images, any brain motion can cause significant artifacts; often the tissue may get displaced by 10 or more pixels from its rest position. To scan an image of 512 lines it takes about $1s$. During this time, at least 3 heart beats and 1 respiration happen moving the brain. Therefore some tissue locations are scanned several times while others are missed. Consequently, although the images may appear reasonable, they can lead to incorrect conclusions with respect to brain structure or function. As lines are scanned almost instantaneously ($\approx 1ms$), our problem is reduced to relocating each line in a three-dimensional stack of images to its "correct" location. In order to model the movement process and quantify the effect of the physiological signal, we collected hybrid image data: fixing y and z, the microscope was set to scan in the x direction for several thousands of times. Classifying these lines using Normalized Cross-Correlation kernel function, we were able to track the trajectory that the line follows due to brain motion. Based on it, we can estimate the number of replicates that we may need to reconstruct a reliable image. Also, we can study how it relates with the physiological values. To address the motion effects, we describe a Semi-Hidden Markov Model to estimate the sequence of hidden states most likely to have generated the observations. The model considers that at the scanning time the brain is either in "near-to-rest" (S_1) state, or in "far-from-rest" (S_2) state. Our algorithm assigns probabilities for each state based on concomitant physiological measurements. Using Viterbi's approach we estimate the most likely path of states and we select the lines

observed in S_1. Because there is no gold standard we suggest comparing our result with a stack of images collected after the animal is sacrificed. Conditioned on inherent experimental and technological limitations, the results of this work offer a description of the brain movement caused by physiology and a solution for reconstructing reliable images from *in vivo* microscopy.

Acknowledgements

Each journey has its ups and downs, but while taking it we can transform ourselves or not. Since I started my adventure at University of Pittsburgh, my journey through this thesis project has opened my mind and it has made me grow as a person and as a researcher.

For showing me the "light" I have to deeply thank my Professors. My advisor, Dr. Bill Eddy guided me to adjust my "senses" so that I could distinguish the significant steps forward. While initiating this journey, Dr. Justin Crowley helped me understand its meaning and he was very receptive to any new exploratory ideas. Dr. Satish Iyengar, my co-advisor, believed in the success of my efforts and always provided good advice. Dr. Leon Gleser showed me ways to enjoy the journey and windows which could bring more light on my search. Dr. Henry Block offered good comments and understanding. While attending their classes, Dr. Taeyoung Park, Dr. Allan Sampson, Dr. Tom Savits, Dr. David Stoffer and Dr. Wesley Thompson showed me paths that could be turned into shortcuts to bring me closer to the target.

Also, I thank David Whitney and Santosh Chandrasekaran for crossing their paths with mine and providing me with tools that helped me advance. I am thankful to Mary Gerber, Kim Thomas and Margie Smykla for their kind words and prompt help.

I am grateful to my family, especially my brother, Iulian Obreja who convinced me to begin this adventure while being an example of successful researcher himself. My parents, Maria and Ion were always fascinated by my journey and they have always been a refreshing source of motivation.

This journey has been made more pleasant by the presence of my friends. Among others who listened to me along the way, I want to thank Lucia Talpes, Diana Luca, Gary Klein, Missy Sovak, and Dongli Zhou for sharing their experience and thoughts. Thanks to all those who, near-by or far-away, supported and encouraged my perceptions or offered helpful criticism.

I can not imagine this long trip without the presence of my friend Emil Talpes who stood by me at all the heights and depths of my path. Words are not enough to express my gratitude for his unwavering support.

to Maria, my mother

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Beauty is in the eyes of the beholder... (Plato).

In recent years we have witnessed the continuous emergence of techniques for measuring high throughput biological data. The abundance of details revealed by these methods has clarified some questions, but it has also opened new areas of research. The attention of the research community has gradually shifted from reductionism towards a more complex understanding of the various biological systems. Even though it has started with intuition and experimental observation, modern biological research requires the development of computational and statistical tools. Such tools are needed in order to model increasingly complex systems and analyze large quantities of data.

Developing such models relies on interdisciplinary collaboration. While the overarching goal of research is to gain insight into biological systems, each discipline tends to focus on its specific domain. For example, doctors and biologists would generally target applications and experimental hypotheses, while statisticians emphasize data acquisition, developing of computational modeling, analysis, and validation tools.

In this thesis I present several techniques which allow the accurate reconstruction of brain tissue images collected through *in vivo* microscopy. While the microscopy process can produce very detailed images, the movement in living tissue introduces significant uncertainty. In all living organisms, the biological components work together to promote development and sustainability, and therefore, they play a pivotal role in all the processes that occur in these organisms. Imaging the living brain produces noisy pictures, aspect which should be carefully considered when drawing conclusions about the structure or functionality of the tissue.

From an application point of view, our goal is to develop tools that can generate joint

models with:

1) Quantifiable capabilities, i.e. we need to allow users to evaluate not only whether, for example, the images are corrupted by motion artifacts, but also by how much;

2) Predictive capabilities, in predicting for example the optimal number of replicates necessary for the reconstruction of a "good" image;

3) Subject -specific capabilities, i.e. the model needs to be adaptable to differences between species used for experiments;

4) Computational efficiency, since working with large dimensional data is always constrained by memory size computational complexity.

In this thesis I lay out our research from a statistical angle. However, the main directions of research in all domains involved will ultimately converge. For example, experimental learning and applications generate hypotheses; hypotheses influence the type of data acquired, and thus the development of data acquisition techniques, but also how much we model and at what level of detail.

In the following sections I present a short introduction to the field of neuroscience. Next I describe briefly various imaging techniques available, as well as some of the problems that researchers face when analyzing images produced by these instruments. This information will give a general sense about the scientific context of our project and will support our motivations and research goals.

1.1   BACKGROUND

To study the structure of brain cells, scientists have had to overcome several obstacles. The first one was the size of the features involved, which is beyond the limit of what can be seen by naked eye. Therefore, the development of compound microscope in the late seventeen century contributed significantly to the progress of cellular neuroscience.

Technical advances in microscopy during the early 1800s gave scientists their first opportunity to examine animal tissue at high magnifications. In 1839, German zoologist Theodore Schwann proposed what is now known as cell theory, speculating that all living tissues are

composed of microscopic units called cells ([2]).

In 1873, Camillo Golgi discovered that that a small percentage of neurons become entirely darkly colored when the brain tissue is soaked in a silver chromate solution (now called the Golgi stain).). This process reveals the complete neuronal cell body, showing that neurons have at least two distinguishable parts: a central region that contains the cell nucleus and numerous thin tubes that radiate away from the central region. The swollen region containing the cell nucleus is the soma or the cell body. The thin tubes that radiate away from the soma are called neuritis and they are of two types: axons and dendrites ([2]) (see fig.1.1).
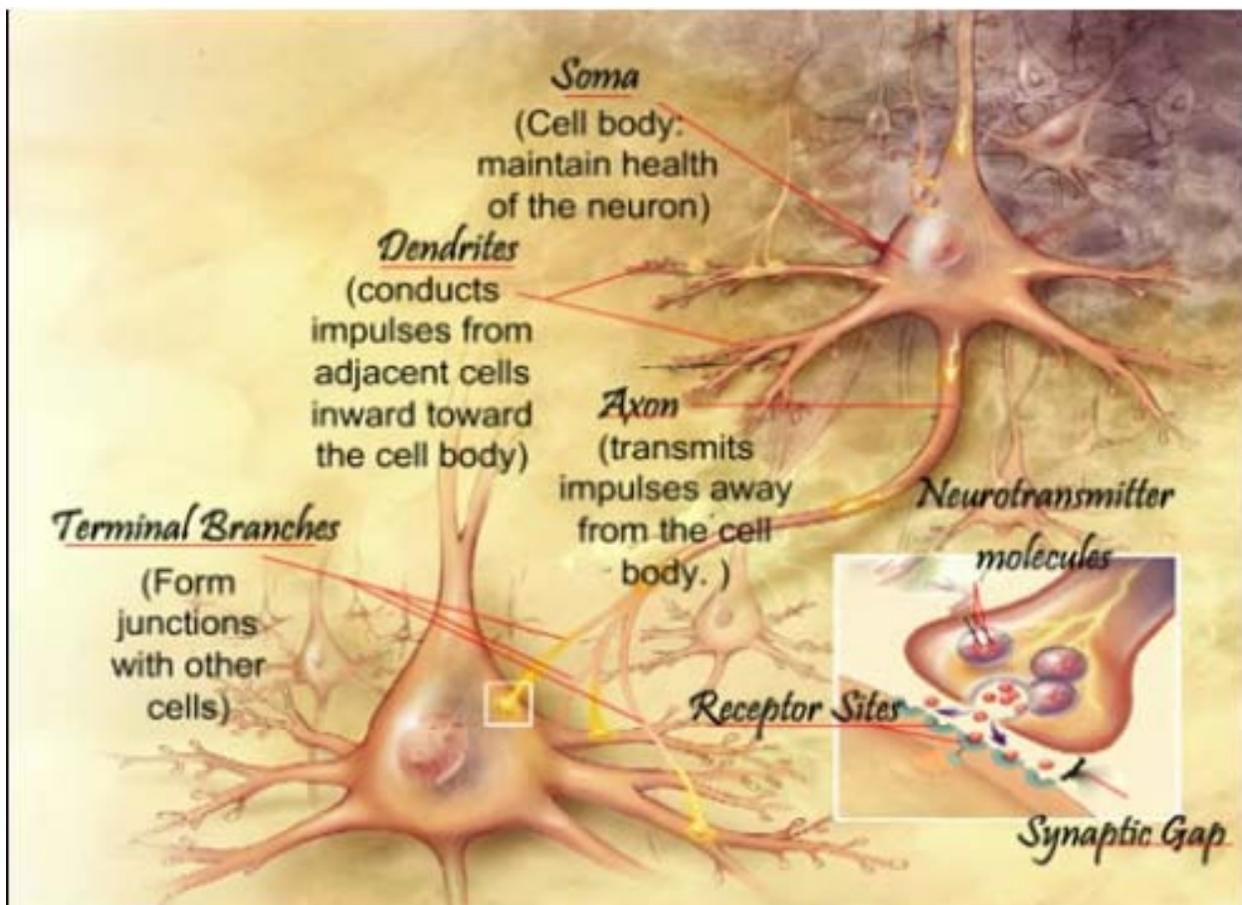
Figure 1.1: A schematic view of two adjacent neurons from http://www.elearningsource.info

The cell body usually gives rise to a single axon which is of uniform diameter throughout its length. Axons can travel over great distances in the body (about 1m), so histologists have recognized that they must act like "wires" carrying the output of the neurons. Dendrites on

the other hand, rarely extend to more than 2 mm length. Many dendrites extend from the cell body and generally taper to a fine point. Since dendrites come in contact with many axons, early histologists have speculated that they must act as the antennae of the neuron which receive incoming signals ([2]).

In 1888, Santiago Ramon Cajal learned about Golgi's method. He extended Golgi's credo that "The gain in brain is mainly in the stain" by labeling many neuronal regions. Over the next 25 years, he used the Golgi stain to work out the circuitry of many regions of the brain. He argued forcefully that the neurites of different neurons are not continuous with one another and must communicate by contact, not continuity. This idea that the nervous system is made up of discrete individual cells was known as the neuron doctrine ([14]).

Due to technical and experimental limitations, obtaining scientific evidence was only possible after the development of the electron microscope in 1950s. The increased power of the electron microscope, described briefly below, made is possible to observe that the neurites of different neurons are not continuous with one another. Thus, the starting point in the explanation of the brain must be the study of an individual neuron.

Neurons have a soma diameter of about $5-50\mu$m, while neurites are considerably smaller at $0.5-4\mu$m. As 100μm is near the limit of resolution for the unaided eye, the light microscope was a necessary development before neuronal structure could be studied. . However, this type of microscopy has inherent limitations imposed by the properties of optical lenses and the wavelength of visible light. With the standard light microscope, the finest possible resolution is about 0.1μm. The space between neurons measures only about 0.02μm(20nm), making it difficult to check whether the neurites were continuous from one cell to the next. This issue could not be resolved until 1950's when the electron microscope was developed and applied to biological specimens ([2]).

The electron microscope uses an electron beam instead of light in order to illuminate the specimen and form magnified images. The limit of resolution for this microscope is about $0.1nm$, bringing out interesting insights into the fine structures the neurons.

Today, state of the art microscopes use laser beams to illuminate the tissue and computers to create digital images. Unlike the traditional methods of light and electron microscopy, which require tissue fixation, these new techniques give neuroscientists their first chance to

peer into brain tissue that is still alive. More details regarding this stage in the neuroscience evolution are presented in the next section.

Bear, et al. (2007) do an overview on how brain research is conducted nowadays and why its continuation is important to society. "History has shown that understanding how the brain works is a big challenge. To reduce the complexity of the problem, neuroscientists break it into smaller pieces for systematic experimental analysis. This is called the reductionist approach. The size of the unit of study defines what is often called the level of analysis. In ascending order these levels are molecular, cellular, systems, behavioral and cognitive. The brain has been called the most complex piece of matter in the universe. Brain matter consists of a fantastic variety of molecules, many of which are unique to the nervous system. These different molecules play many different roles that are crucial for brain function: messengers that allow neurons to communicate with one another, sentries that control what materials can enter or leave neurons, conductors that orchestrate neuron growth, archivists of past experiences. The study of the brain at this most elementary level is called **molecular** neuroscience. The next level of analysis is **cellular** neuroscience which focuses on studying how all these molecules work together to give the neuron its specific properties. Among the questions asked at this level are: How many different types of neurons are there, and how do they differ in function? How do neurons influence other neurons? How do neurons become "wired" together during fetal development? How do neurons perform computations? From the functionality angle, constellations of neurons form complex **circuits** that perform a common function: vision, for example. Thus, we can speak of the "visual system" with its own distinct circuitry within the brain. At this level of analysis neuroscientists study how different circuits analyze sensory information, form perceptions with the external world, make decisions, and execute movements. How do neural systems work together to produce integrated behaviors? For example, are there different forms of memory accounted for by different systems? Where in the brain do "mind-altering" drugs act, and what is the normal contribution of these systems to the regulation of mood and behavior? What neural systems account for gender specific behavior? Where in the brain do dreams come from? These questions are studied in **behavioral** neuroscience. Perhaps the greatest challenge of neuroscience is understanding the neural mechanism responsible for the higher levels of human

mental activity, such as self awareness, mental imagery, and language. Research at this level, called **cognitive** neuroscience, studies how the activity in the brain creates the mind."([2])

The miracle of life is not completely understood and the history of modern neuroscience is still being written. Unexplained phenomena, many of them unnoticed until very recently, capture researchers' attention all over the world. We hope that our efforts will help bring some explanations.

1.2   *IN VIVO* IMAGING

Imaging internal body tissue requires non-invasive methods as much as possible, in order to reduce the external perturbations. Such perturbations can impact the understanding of the relationships among all components within a cell and how they respond to different stimuli. For a long time, image processing methods have been applied for improving the quality of resultant images. Although there is a broad spectrum of applications for these techniques, due to the nature of input data I will focus mainly on neuro-imaging methods.

Obtaining detailed and reliable information about brain organization is an important technical challenge in neuroscience. Understanding the neural structure and its functions require accurate methods to visualize live brain tissue in a non-destructive way. Therefore, a variety of methods and instruments have been developed to provide insights into the neural activity.

In the 1970's, A. M. Cormack and G. N. Hounsfield introduced Computerized Axial Tomography (CAT). Ten years later, the development of radioligands allowed Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) of the brain. A big step forward in non-invasive methodology was made by P. Mansfield and P. Lauterbur with the Magnetic Resonance Imaging, opening the way for a veritable explosion of technical refinements and MR applications. These imaging techniques allow access to deeper structures, but they offer significantly lower spatial resolution and longer acquisition time compared to light microscopy ([28]). Based on a concept first described by M. G. Mayer ([16]), two-photon microscopy is a newly developed imaging technique which enables

6

the study of three dimensional biological specimens with submicrometer resolution ([31]). Extended to multi-photon microscopy, this method was pioneered and developed into a practical tool by W. Denk at Cornell University who revealed its superiority to confocal microscopy. Multi-photon microscopy allows deeper penetration, efficient light detection and reduced phototoxicity ([4]).

Two Photon Laser Scanning Microscopy (TPLSM) *in vivo* has become a turning point for the progress of science, as it can be employed to generate three-dimensional microscopic imaging in thick tissue. Among other properties, depth discrimination has permitted the tracing of learning dependant dendritic spine morphology ([19], [20]). Hence, the experimental design becomes more interesting for researchers. Not only can the spatial distribution of the microscopic features be measured, but also their developmental progress can be monitored over time. Manipulation of brain activity, sensory experience, learning, alteration in gene expression, and pharmacological manipulations are only some of the experiments that take advantage of TPLSM. Also, the subject under the microscope can be used as its own control, improving significantly the conclusions drawn by reducing sources of variation.

All high-resolution imaging techniques require mechanical stability of the subject or an exposure time that is fast relative to the rate of movement. Since these techniques often rely on collecting several replicated or spatially related images, it is extremely important to control the causes that determine differences among replicates. Analyzing the recorded images reveals the presence of motion artifacts ([10]) which can be data acquisition related or subject related ([21]). Process-related distortions may reduce the image quality because of signal drops and pixel noise, while physiological causes (e.g., respiration or heart beat) may impact the quality as well. Sirovich and Kaplan ([30]) report seemingly disastrous signal-to-noise ratios of 0.001 in intrinsic signal optical imaging data, with much of the noise coming from physiological processes unrelated to the stimulus of interest.

In our research, we look for ways to improve the reliability of TPLSM using statistical methods for reducing movement artifacts. While we focus on images of the neocortex / visual cortex, our approach is applicable to a wide array of imaging problems.

1.3 MOTIVATIONS AND GOALS

Eppur si muove (Galileo Galilei)

Past research has determined that during development, the optic nerves originat from the eyes, cross at the optic chiasm and project to the Lateral Geniculate Nucleuous (LGN). The output of the LGN connects with the primary visual cortex (see fig.1.2). However, the way neurons change and communicate over time in the visual cortex is still under research.
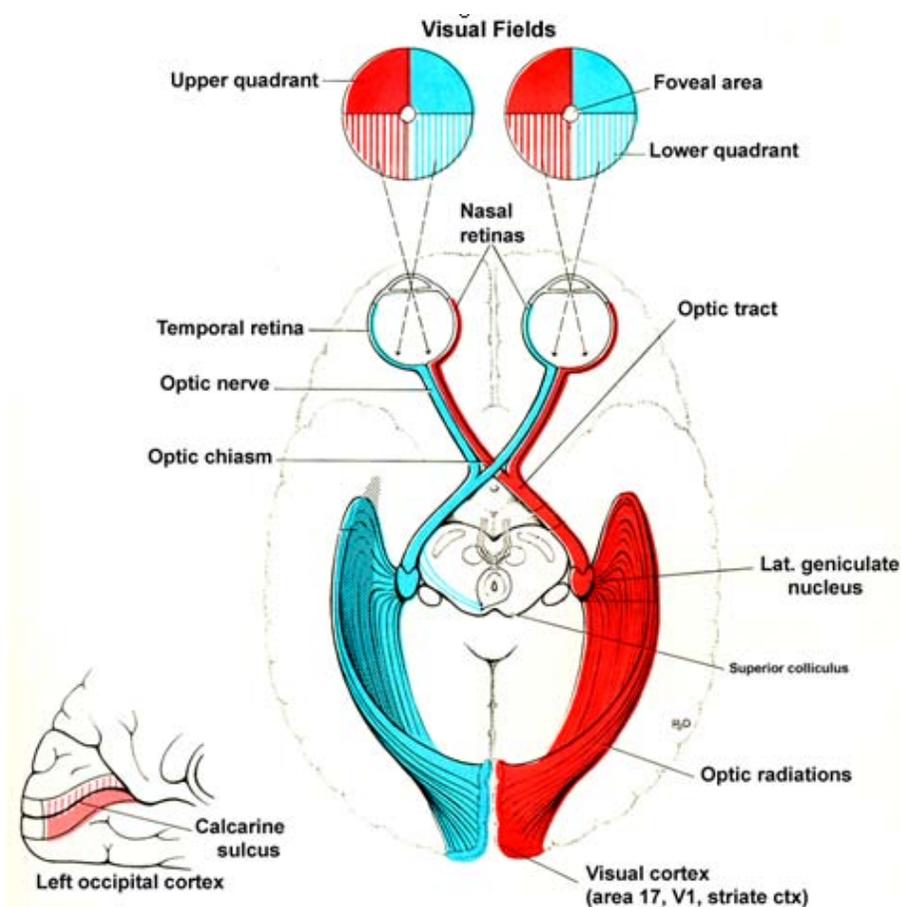
Figure 1.2: A schematic representation of the visual system (from [33]).

To learn more about the growth of the neurons in the visual cortex, researchers have tried to observe their structural changes over time. This requires a nondestructive way to image on the *in vivo* tissue in the brain of a neonatal animal (e.g., ferret or mouse), at different times during its development. The ferrets are of special interest as they are born blind, with

8

the neurons of the visual cortex forming synaptic connections during the first months of life. TPLSM offers a major advantage due to its dramatically reduced effects of light scattering, high resolution, deep tissue penetration and minimal tissue damage.

The microscopic areas in the brain that are scanned are of the order of several hundreds of micrometers or less. Each dendrite represents a possible series of connections with other neurons, therefore monitoring the dynamics of dendritic structure gives valuable information about the development of the nerve and neuroactivity.

The pattern generated from laser-excited fluorescence can be analyzed in order to construct a three-dimensional, micro anatomical, structural image. It is common practice to take multiple images within the same experimental condition and average them to increase the signal-to-noise ratio. However, brain movement in the *in vivo* experiment can make the simple averaging of this data impossible.

In this set-up, limiting the mechanical movement of the subject is critical. A small movement from one pixel into the neighboring one can cause a significant change in the value of the observed signal and it can blur the final image data ([9]). Thus, when correlated against stimulus presentation, the movement artifacts can be misinterpreted and the final experimental conclusions can be considerably affected.

One major source of noise consists of body dynamics, primarily those processes which cause a movement rate higher than the scanning rate. This introduces distortions in the image acquisition - the positions of the pixels get shifted, while some real data is never scanned. A visual example is presented in Fig1.3 where the same scenario in the brain has been scanned three consecutive times. By creating difference images between any two replicates we can easily notice that the main feature in the image has moved from one scan to the other.

Another way of representing these feature displacements is by constructing an RGB system, which is based on the way the color is perceived by the human eye. To represent any color, we specify the intensity of red light (R), green light (G) and blue light (B), the so called primary colors. In the RGB color model, each pixel of an image is a mixture of these
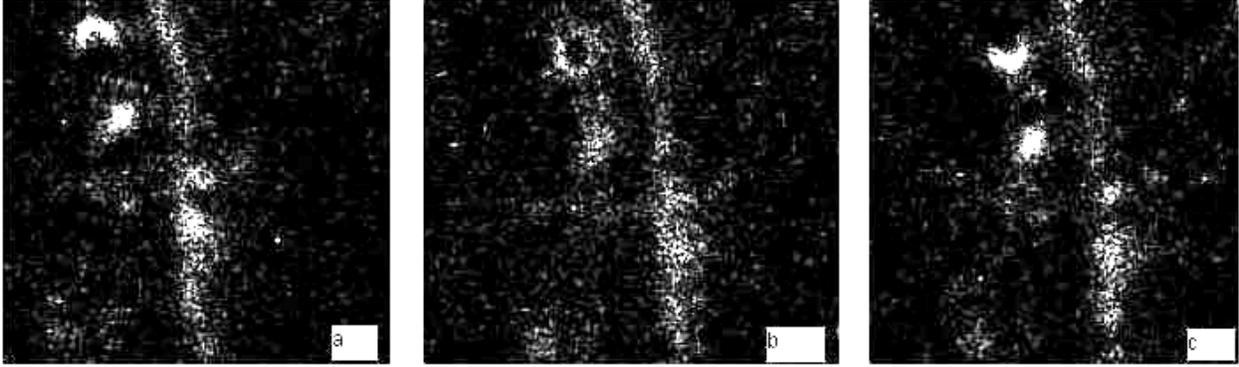
Figure 1.3: The difference images between any two replicates of scans in a group of three, for the same z-slice of rodent brain (the difference between first and second (a), second and third (b), and third and first (c)). The higher intensity values point out the large difference in pixel intensity for the same location, i.e. the brain position changed during these recordings. Note: the features have been magnified for better visualization.

primary colors. It is based on a Cartesian coordinate system in which the color space can be shown in a cube, as in Fig1.4.
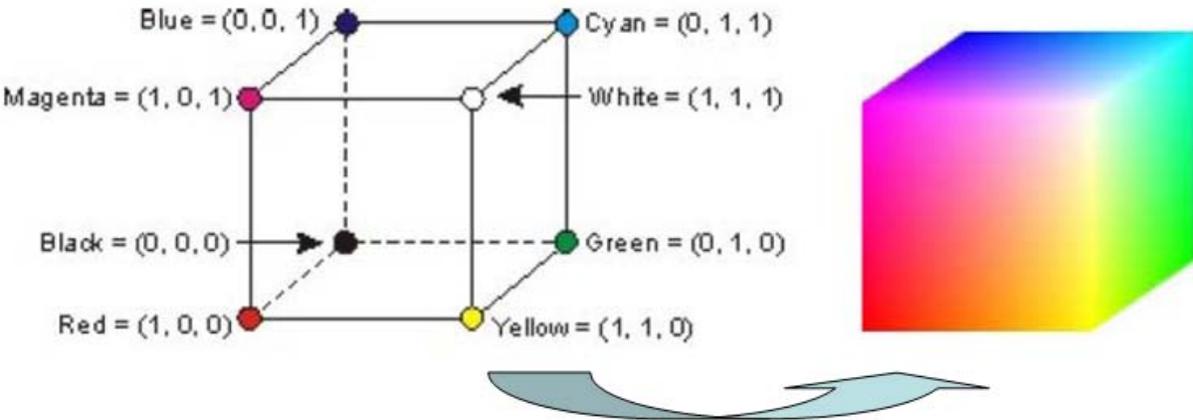
Figure 1.4: Color Cube system: over-imposing three color coded images creates a new image with pixel intensities varying accordingly to the color cube. Its origin is black (no red, no green, no blue), the opposite corner (on the diagonal) is white (same amount of red, green and blue), while all the other points are combinations of certain amounts of red, green and blue ([17])

Thus, we can associate any color with a three dimensional quantity given by its RGB components. This specifies a point inside a cube, the sides of which correspond to the range of possible intensities for each of the primary colors.

One use for the color in image display is to assign different pixel values in a grayscale image. In our case, the intensities of the pixels from the first replicate had the values translated to the red range, those from the second replicate were translated to the green range, while those from the third one were translated to the blue range. According to Glasbey, et al. (1995) ([15]), this is called pseudo color model, since it does not reflect the actual color variation in the scene imaged. For us it represents the variation of the pixel intensity from one image to the other. Sometimes this representation has advantages over a standard grey scale display, since the human eye is better able to see color differences and compare colors in different parts of display. Many image features and comparisons can be more readily made in color than in grey levels, since the problem of sensitivity to the surrounding area is much reduced ([15]).

An equal response in each of the three components produces a perception of grey or white. Consequently, if there was no movement between the three replicates, when combining the pseudo-intensities for each pixel position we should get an image on a grey scale range. The colored version of fig1.5 shows variations of red, green or blue indicating movement artifacts.
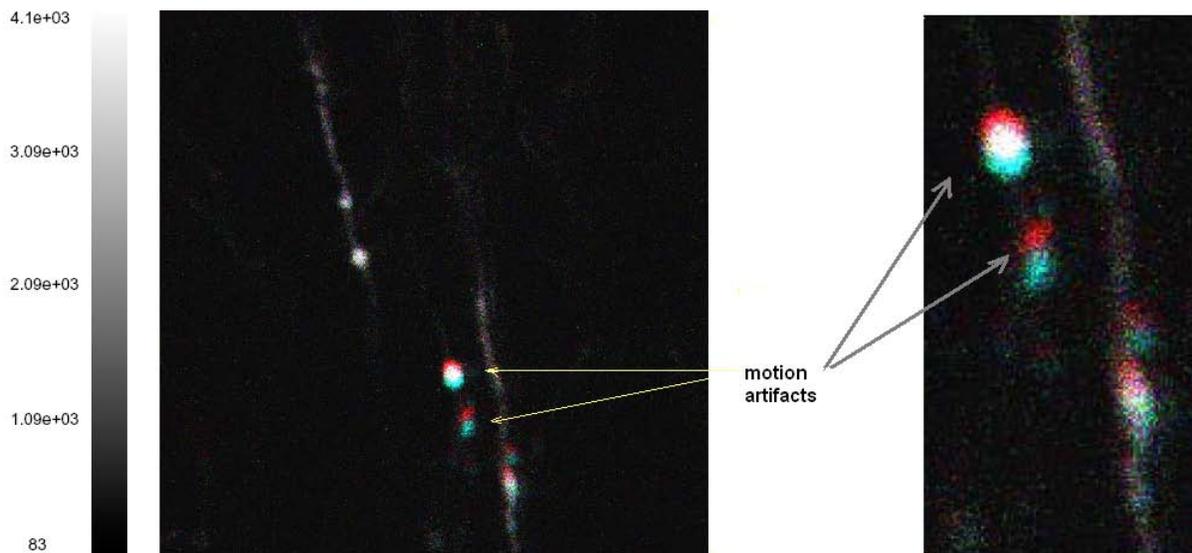
Figure 1.5: The entire RGB version of fig1.3: the images collected at 3 consecutive times at the same depth (z) have been color -coded in red (1st), green (2nd) and blue (3rd), and then over-imposed. If the pixel position was not changed then the combined image would have only variations of grey. The presence of red, green and blue colors indicates the displacement of the scanned feature from one scan to the other. Note: The right side magnifies the main feature which seems to have been moved from one time to the other.

Identifying the problem in the image space does not offer a way to solving it. This research project is concerned with finding solutions for minimizing the motion artifacts and reconstructing reliable images of the brain. The task may seem simple, as all we need is to estimate the true positions of the pixels and then put them in the right place. However, the imaging domain is rich in information and the movement is three dimensional. With so many sources of variability, it becomes difficult for computer analysis.

Estimating position parameters in rigid body transformations involves solving equations with about 7 unknowns: 3 for space translations and 4 for rotations ([34]). The input data we use comes from images with 512x512 pixels which brings the amount of computations to a very expensive level. At the same time, by the nature of these experiments, the moving process is continuously affecting each scan, at different degrees, so there is no gold standard image, free of artifacts that can be used as target in the usual registration methods. Consequently this registration problem might not be considered as a rigid body set-up, but rather as a dynamic system. Entities of interest which are measurable directly become, in fact, latent variables in our model. These variables - such as feature location - need to be inferred from directly measurable data.

It is the goal of this thesis to provide relatively inexpensive ways to describe the movement in the brain by estimating the amount of shift that pixels may be translated under the physiological pressure.

By learning as much as possible about the biological and physical phenomena, we were able to identify several constraints that helped us base our approaches to solve this problem. One example might be the empirical knowledge that the physiological process is almost periodic, forcing the brain to move back and forth inside the skull.

Simple mathematical evaluations of the rates of physiological and scanning processes are presented in more details in the Preliminary Analysis chapter, section 4.1. They prove that the scanning process is fast enough so that scanning one line can be considered instantaneous. Therefore we reduced the number of parameters by estimating positions of lines rather than pixels. However, the scanning process is not fast enough while acquiring an entire image, as there are at least 3 heart beats and 1 respiration happening during this time.

Constructing synthetic models help us build probabilistic arguments about various as-

pects of the problem. The scanning process (presented in more detail in a subsequent section) can be simplified to drawing scanned lines from scenes within the brain. If we define "success" as the event of scanning pixels within lines at any of their possible locations in the "real" scanned scene, counting the actual "successes" implies the use of a binomial distribution. As the probability of getting no "success" is positive, we need to be aware that there are lines of pixels that are never scanned.

Consequently, the estimation problem gets more complicated by the fact that we are given less input and asked to come with more output. That is, the amount of constraint on each quantity to be computed is less. Another interesting challenge would be to determine what we should use to fill in these gaps.

Experimental observations suggest that we need to collect multiple images of the same slice in a time lapse manner to increase the incidence of "true"' signal in the data. In our case the same specimen is imaged under different stages of the neuronal development. Hence, in order to keep the animal in a relatively normal biological state, the experimenters need to optimize the number of scans, minimizing therefore the tissue damage caused by photo bleaching.

The objective of TPLSM is to obtain information about the three dimensional shape of a scene of interest, stacking the output as a list of 2D images. When movies are created from these stacks (e.g., through animation in the z dimension), we can observe, as expected, the movement of the brain occurring mostly due to heart beat and respiration. The general goal of our research would be to adjust the movement in all three directions and thus to reconstruct a movie where the physiological artifacts are no longer significant. However, the vectorial decomposition of the signal on each spatial coordinate allows us to focus our research locally, in a lower dimensional plane. Throughout this thesis I describe each method we approached locally for fixed slices in the brain, in a two dimensional set-up, and then extend them including the inter-slice dimension.

The local practical problem that we plan to solve is: given r views of the same scene in the brain, taken at relatively close but different times, we need to determine a single image that is adjusted for the movement effects. This problem has two essential characteristics: it is a parameter estimation of the line locations and a reasonable reconstructed image with

the main attributes -smoothness and continuity- preserved.

The estimation problem is complex: we need to understand the relationship between the physiological signal and the output image. Then, for each pixel, we need to find the most reasonable location that can be predicted from modeling this association. Next, the digitization of the image may require some adjustments as the estimated positions may be located at a fractional number of pixels. The two main questions of interest would be when and how is the movement happening. As the "how" part depends to considerable extent on the "when" part, we will address it latter in this thesis by describing the computational algorithms used. It seems that the key would be to determine those lines scanned when the blood pressure in the brain is smallest, and thus the movement is minimal. We call this the "near-to-rest" state of the brain. It is biologically impossible to have a completely resting state while the animal under experiment is alive. While cardiac or respiratory events happen, the pressure in the brain changes forcing the tissue to enter a "far-from-rest" state.

In this context, we could look at the observed images as divided into blocks of lines that come from different states. Considering local dependencies within images, both block sizes and classification rules can vary according to context. One way to address these connections would be using a hidden Markov model (HMM) as a general framework for context dependent classifiers. Since we are primarily interested in the "near-to-rest" state, we could estimate the block of lines that have been scanned in that state and consider them as reference for the final image.

1.4   PREVIOUS WORK

The process-related noise that corrupts TPLSM pictures is generally addressed through technological advancements. Motion artifacts however, have not been studied in much detail. Very few research teams implement post-hoc adjustments which take motion noise into consideration, generally assuming that the brain under anesthesia is free of movement.

Relative motion between the microscope and brain was apparent during some experiments described by Dombeck and his collaborators ([6]). While scanning the brain in awake

adolescent and adult mice, they were interested in several forms of neural dynamics such as persistent activity. Before collecting TPLSM images, the team applied head restraints to the subject, employing in the same time an air-supported spherical treadmill to allow walking and running. They adjusted the residual lateral shifts to reduce the image distortions for post-processing the output. For quantifying in-focal-plane brain motion, the team corrected the frames back to a reference image selected from the resting state. They used an offline line-by-line algorithm based on an Hidden Markov Models ([27]), where the probability of a given sequence of observations (offsets) is a sum of two components, summed over all the time points in the sequence:

1. the fit of the line scanned at a given time point compared against the reference image at the given offset;

2. the probability of the observation transitioning from the offset given at the previous time point to the current offset.

The correction algorithm they developed remapped the recorded line by placing it at the correspondent x and y shifted location in a reconstructed image.

Since we study scanned images in resting, anesthetized subjects, the brain motion that we want to adjust is mostly caused by heart beat and respiration. The approach of Dombeck, et al. ignores these causes, even though they still affected the image taken when the mouse was not running.

Greenberg and Kerr ([18]) developed an automated correction of fast motion artifacts based on the Lucas-Kanade framework, operating directly on the motion-distorted two photon imaging data of awake animals. Like Dombeck, et al., they did not consider external signals such as heart beat or respiration either. However, when estimating the displacement they used a template image obtained during subsequent anesthesia which they later generated as an average of multiple scans of the same slice. Estimating the movement trajectories as piecewise linear functions, they used a gradient descent approach which iteratively estimates the displacement parameters.

Mizrahi, et al. ([24]) have scanned at high speed each region of interest for $4-10$ times (pixel dwell time $0.64 - 0.8\mu$sec/pixel). Then an offline correction algorithm was applied to

adjust for the movement artifacts:

1. The first stack was considered as a reference for a three dimensional alignment of the rest of the stacks. These have been spatially shifted so that each component frame would match with its correspondent reference. Maximal match was defined as the maximal value of a two-dimensional cross correlation between two images.

2. Within each frame, the correction for the movement caused by the heartbeat was done as follows. Each frame was divided into segments of $10 - 20$ consecutive lines which were compared with the appropriate segments in the first stack by a two-dimensional cross correlation function. Then they were shifted to the location of the peak of the cross correlation function.

3. Corresponding frames in consecutive stacks were averaged to get a single "average stack".

4. Averaged images obtained at the previous step were filtered using a two-dimensional low-pass filter (cutoff, 5 cycles/μm).

Although the above mentioned studies performed their motion correction in different experimental set-ups, our algorithms confirmed some of their findings, such as existence of lines in a slice which are never scanned 4.3.

1.5   DISSERTATION OUTLINE

The following chapters describe several approaches that we used in order to capture the essence of this project. Given a computer image, the intent is to learn from it or to convert it to a more accurate version.

The experimental instrumentation and some of the underlying physical principles are presented in **Chapter 2**. Having a better understanding of the phenomenology clarifies the meaning of pixel values in a microscopy image. It also drives the decision process when analyzing the true signal versus noise. Although recorded at different rate and through a different device, the physiological data plays an important role in making this decision.

Chapter 3 reviews the pre-processing methods that are generally utilized for decreasing the measurement related noise and it introduces a new filter for image thresholding. As most of these methods represent transformations of the original image, the gain or loss associated with their use justifies the superiority of our novel algorithm. Simplifying the original image to an optimum degree, we use for analysis a highly correlated image with very few possible values for pixel intensities. Consequently, most of the background noise is reduced while the features gain a more consistent shape. Subsequently, the attention is focused on the position of these features.

Chapter 4 explores a series of preliminary analyses which provide leading guidelines for the rest of the thesis. The registration methods proposed in this chapter, together with the basic characteristics of the scanning process re-emphasize the main artifacts that corrupt a TPLSM image: some lines are scanned multiple times, while others are never scanned. In the absence of a "gold" standard image to which the observed images can be registered, we need a more sensitive way to capture the motion process.

Chapter 5 describes a modified experimental set-up which fixes the y and z coordinates while scanning the brain tissue in the x direction for several thousands times. We start by analyzing this data using simple methods, and in **Chapter 6** we move to more complex models which estimate the trajectory of the brain movement.

Chapter 7 offers a probabilistic perspective on images, justified by the results described in the previous chapters. Using hidden Markov models we reconstruct images based on the physiological values and several replicates for each image.

I conclude in **Chapter 8** by discussing the methods and results in the context of our initial research problem. This chapter emphasizes the novel contributions of our work while pointing out potential directions for future research.

2.0 DESCRIPTION OF THE INSTRUMENTS AND THE DATA

Study lends a kind of enchantment to all our surroundings (Honore de Balzac).

The correction of neuro-imaging data by accounting for physiological processes in live animals has not been thoroughly studied in the past. In general, such correction needs to be applied when the captured data includes both the experimental induced activity (like growths of fluorescent neurons or homodynamic response) as well as brain movements that are not related to the experiment, such as heart beat and respiration. Some studies (e.g., fMRI-based approaches) solve this problem by averaging over several replicas to extract a signal which is small relative to the noise in the acquired data. These techniques introduce a new source of noise, making the picture more blurred. However, it has not yet been decided what the golden standard image should be in cases where the physiology is still affecting the scanning process.

Somewhat more general techniques for image reconstruction are those based on an image registration process. Their main purpose is to overlay two or more views of the same scene, taken by different sensors and/or at different times so that the pixels of the same coordinates in the images correspond to the same real data.

As we will clarify in more detail in the following sections, our scanning process is not fast enough so the subject's physiology moves the brain while the images are taken. Thus, the information gained from different replicate images is usually of a complementary nature: the data missing from one image may be captured in a different one. Given that useful data is scattered across the entire set of replicates we need a method to properly integrate it in a single image.

In the next section I briefly describe the imaging technique and the data we gather for analysis.

18

2.1   TWO-PHOTON LASER SCANNING MICROSCOPY

TPLSM is a unique method that allows a specimen to be imaged in three dimensions, moving the focal plane to a depth of up to one millimeter without damaging the tissue.

It is based on the interaction of two distinct phenomena: fluorescent dyes and light excitation. Fluorescence occurs when an electron in the outer shell of an atom absorbs most of the energy of a photon and gets excited into a higher energy state. This new high energy state is unstable. A short time later the electron returns to its stable state and emits a photon of lower energy. In general, there is an inverse proportional relationship between the wavelength and the energy of the photon. Hence, the wavelength of the emitted photon is longer than the wavelength of the incident photon. The difference in energy is absorbed by the tissue as heat.

Under sufficiently intense laser illumination, the simultaneous absorption of two infrared photons offers enough energy to generate the previously described energetic excitation. This event can occur if the sum of the energies of the two photons is greater than the energy gap between the molecule's ground and excited states. This process is nonlinear as the action of one infrared photon is not as effective as half of the action of two. Since it depends on the simultaneous absorption of these photons by a fluorescent molecule, the probability of this event is a quadratic function of the excitation radiance.
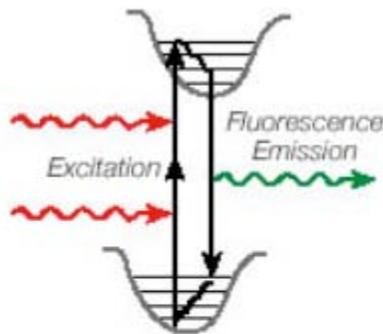
Figure 2.1: The underlying process necessary for TPLSM: two photons with wave length of 860-890nm combine their energy to excite the fluorophore. The resulting photon with wave length around 500nm is captured by the microscope and translated into pixel value ([31].

As applied in the laser-scanning microscope, two- photon fluorescence excitation illuminates a volume of tissue at the focal point with high intensity light at twice the excitation wavelength. The high intensity illumination is attained by focusing a beam from a high-energy pulsed laser delivering bursts of 100 femtosecond to 1-2 picosecond pulses at high frequencies (100 MHz). By using a laser with twice the required wavelength a fluorescence event can only occur if two photons arrive at the same electron shell simultaneously [32].

This microscope is able to focus at various depths in the visual cortex, by moving the objective. In our experiments, during the scanning process a laser beam creates fluorescent emissions in the visual cortex which is then detected by a photomultiplier tube (PMT). This phenomenon results in a voltage change recorded as pixel intensity value. The resulting value, originating from an illuminated volume element, represents one pixel in the recorded image. As the laser scans the z plane of interest, a whole image is obtained pixel-by-pixel and line-by-line.

The beam is focused across the sample in the horizontal plane by using one or more oscillating mirrors (see Fig2.2). At each position, the number of excited electrons is directly proportional to the number of fluorescent molecules hit by the converging light. As a result, the intensity of the fluorescent light (the brightness of the pixel in the resultant image) is proportional to the density of the fluorescent molecules at the corresponding position in the tissue.

Images are built up by scanning the laser illumination across the specimen, changing the focal plane, and repeating the scanning as with the conventional confocal laser-scanning microscope.

As So, et al. point out, depth discrimination is one of the most important properties of two photon microscopes equipped with high numerical aperture objectives. For a spatially uniform specimen, most of fluorescence signals (80%) are generated confined to a region 1 μm thick about the focal point. This precision results from the quadratic dependence of the fluorescence probability on the spatial distribution of the excitation radiance ([?]). Appreciable two-photon fluorescence occurs only at the microscope focal volume, where the photon density is high; negligible fluorescence is excited outside of this volume [32]. A critical
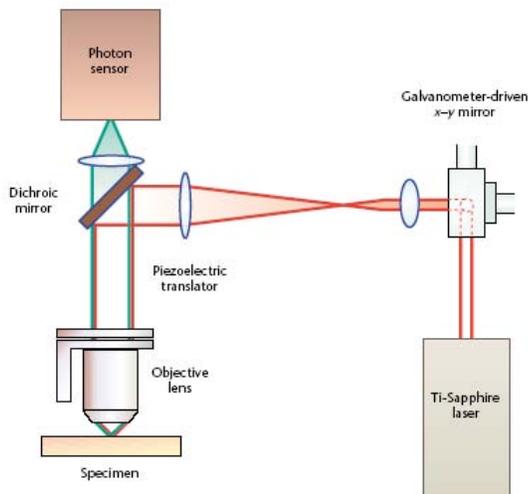
Figure 2.2: Two Photon Laser Scanning Microscope optical system diagram: typically consists of a high-peak-power pulsed laser, a high-throughput scanning microscope and high-sensitivity detection circuitry (from [32]). The excitation light is reflected by a dichroic mirror to the microscope objective and is focused in the specimen. Two-photon induced fluorescence is generated at the diffraction- limited volume. Images are constructed using a galvanometer-driven x-y scanner and a piezo-objective z driver (PMT). The emission signals are collected by the same objective and transmitted through the dichroic mirror.

component is its light source: a high-radiance light source of order of $10^{10} - 10^{12}$ W/cm^2 is required for efficient excitation. High repetition rate ($100MHz$), ultra fast (femtosecond or picosecond pulse widths) lasers, such as Titanium-sapphire are the most widely used light sources. The higher peak power and the lower duty cycle of these lasers minimize average power deposition in the specimen while maximizing two-photon excitation efficiency.

In practice, TPLSM depends on the fluorophore which exists in the scanned tissue. In our case the method for labeling neurons with fluorescent dyes is either by altering their genetic structure or by injecting specific viruses. Combined, these methods have enabled repeated imaging of neurons in intact brain.

The data we use come from two types of animals with different neuron-labeling techniques. One technique requires transgenic mice with foreign genes inserted into their genome in order to interact with the laser scanner ([12]). Alternatively, we use normal ferrets which get a lentivirus injected in their brain to express Green Fluorescent Protein (GFP) ([5]). This causes the neurons to fluoresce when subjected to extra energy from outside photons.

By varying the focal depth we are able scan an entire 3D structure into a stack of 2D images. Each image in the stack corresponds to a slice taken at a constant depth within the

visual cortex tissue.

TPLSM reduces overall bleaching and photo-damage, as the fluorescence depends on the square of the incident light intensity. In turn, the light intensity decreases approximately as the square of the distance from the focus. Because of the nonlinear relationship ($\approx$fourth power), the only dye molecules which get excited are the ones very near the focus of the beams. The tissue above and below the plane of focus is merely subjected to infrared light that is relatively harmless. When working in thick tissue, such as brain slices or developing embryos, TPLM offers major advantages due to the dramatically reduced effects of light scattering. Additionally, since both photons must arrive at the same place at the same time, fewer fluorescent events are taking place away from the focal plane, sharply reducing the dissipation of heat. Photo damage to the cells is very much reduced outside the focal plane, which makes the method particularly useful for imaging of living tissues.

The experiments were performed on a Zeiss LSM 510 Metal $NL0$. A Coherent Mira Titanium Sapphire laser pumped by a $10W$ Verdi diode laser was used to excite GFP at $865 - 890nm$. A closer view is shown in fig2.3, while a more detailed one is presented in D1.

Figure 2.3: Two Photon Laser Scanning Microscope used in Dr. Justin Crowley's lab

2.1.1 Image Data

The natural way to display images is to use pixel values which specify the brightness of every point in the image. The pixel values represent a measured physical property of the object

being studied. If this property is the amount of emitted, reflected or transmitted light, then the display will look like a monochrome photograph. A physical measurement of the amount of light is its intensity. This is proportional to the energy in the light, and is the square of the amplitude of the light waves.

The human eye is not simply an instrument which registers the amount of light it receives. It is specifically trying to detect shapes in the scene, as it is very sensitive to the local contrasts at immediate surroundings of an object. This is best illustrated by optical illusions, one of each is shown in fig2.4. In this picture the horizontal rows seem to change their height when actually they do not. The main lesson to be derived is that the eye is not to be trusted for objective assessment of the absolute intensity of different parts of image display.
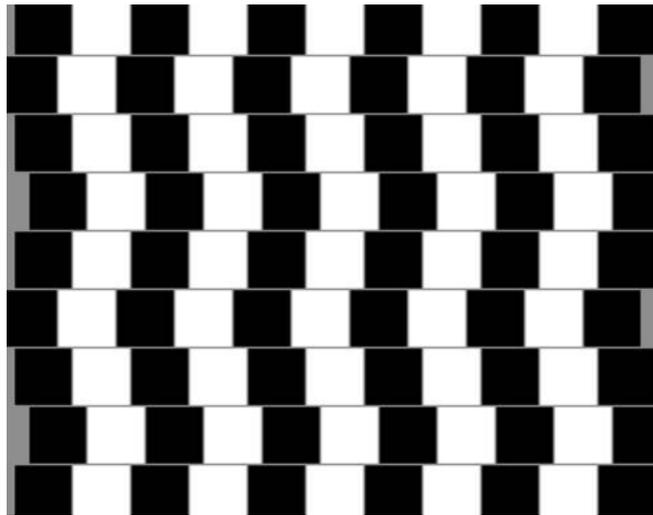
Figure 2.4: Optical Illusion: Because the black and white squares are shifted from one line to the other, the horizontal rows seem to have variable width.

Based on the physical processes described above, image data consists of pixels and their associated coordinates: intensity on a grey scale (an integer value between 0 and 4095), spatial locations (in the y plane or "row" $(1-512)$, in the x plane or "within row" $(1-512)$, and in the z plane or "slice" in the depth of visual cortex plane). The image is scanned several times for each "slice", adding a new variable for time. The dimensions of the pixels depend on the zoom used. They are usually 0.31 μm, but in some experiments they may reach 0.15 μm. This high resolution increases the precision of the scan and the amount of detail present in the picture.
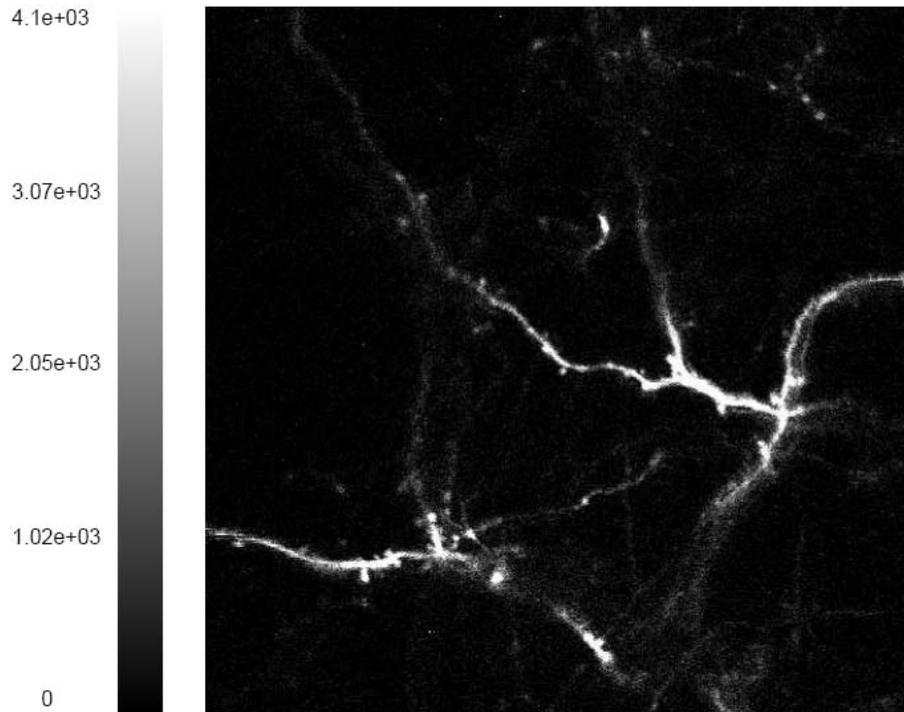
Figure 2.5: A slice in the visual cortex of a mouse's brain. The scanning is displayed as 262144 pixels having intensities between 0 and 4095.

The image coordinate system is subject to change, varying with the position of the rodent under the microscope. In the general set-up, the animal is facing the pedestal of the microscope, immobilized on a small table at an angle of 45 degrees. The image coordinates with respect to the position of the animal are approximately as follows: x direction is from the left of the animal towards its right, y is from anterior to posterior and z is from dorsal to ventral. A schematic visualization is presented in fig.2.6.

In the images we analyze, the features of interest are the dendrites and axons. They are fine appendages of the neurons which appear during nerve development. The high resolution capabilities of TPLSM allow fine details to be captured. By studying how these features evolve while the animal subject is growing we can get valuable insight into the development of the brain tissue.
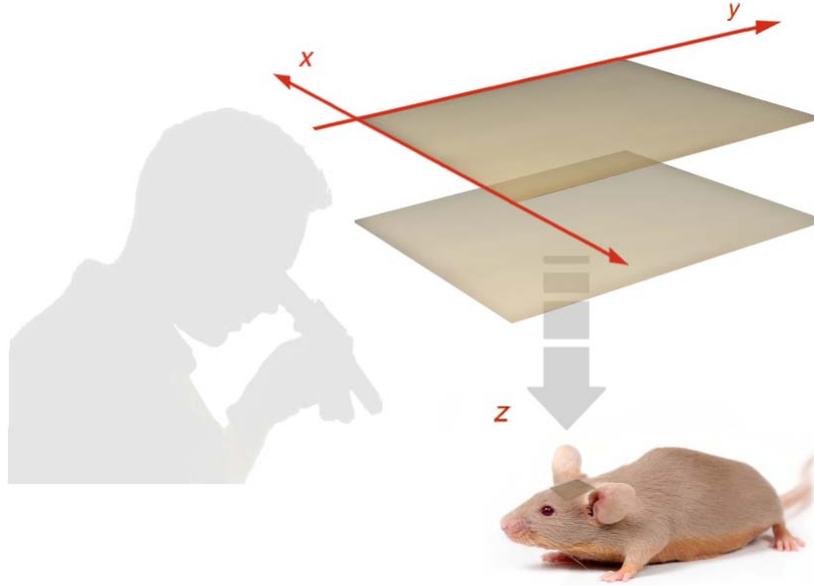
Figure 2.6: Typically, the animal is mounted under the microscope at 45° with respect to the microscope. As the laser beam advances from left to right in this 45° rotated direction (x), each line is created sequentially from the anterior towards the posterior of the animal.

2.2 BRAIN MOREY: BRAIN MOTION REMOVAL SYSTEM

As illustrated by fig1.3, the brain is moving while the scans are taken (more details on this movement will be described in the next sections). Due to this fact, we need to learn as much as possible about the physiology and its contribution to the displacement of the brain. Daniel E. Smith, formerly student in Dr. Crowley's Lab, Department of Biological Sciences, Carnegie Mellon University, developed a LabView program called Brain Morey which helps us in this direction. Besides collecting the physiology values that activate during the experiment, the system can be programmed in modes that help us compensate for respiratory artifacts. Also, the recording can be triggered so that image data can be acquired at equally spaced time points during a pulse cycle.

2.2.1 Physiological data

Before, during, and after the scanning process, the physiological and experimental data is acquired through the Brain Morey program. The data is generally sampled in eight channels, at a rate of $5kHz$. The eight channels are EKG (Electrocardiogram), POX (Pulse-

Oxygenation), RSP (Respiration), END(Pulse at end of experiment), LSM(Wait pulse), STP(Scan stop pulse), SYN(Line sync). The EKG channel records the electrical signal from the heart; the pulse-oximeter channel is a proxy measure for the blood flow, while respiration is measured as the expansion of the chest. The other channels (END, LSM, STP, SYN) connect with the scanning microscope monitoring the experimental steps. Fig2.7 displays the values of these channels for 3 seconds during an experiment on a mouse.

2.3   FIASCO

Considering the nature and the size of our data, almost all analysis and image processing have been done with a special software called FIASCO[1], developed at the Department of Statistics at Carnegie Mellon University. This system was developed as a flexible, adaptive package for data processing and statistical analysis. Thus, the user has a lot of freedom in determining the order and the processing steps. Also, various data collection types are available to be used even if this might be computationally intensive ([8]) The scanned images along with corresponding physiologies are stored in specific file formats which can be decoded by FIASCO. Written in the C programming language, FIASCO offers primitives for a series of mathematical operations and statistical methods on very large data sets.

2.4   SUMMARY

This research combines the above mentioned instruments with a new *tool box* that we create at each level of the analysis. The goal is to set up a collection of programs accessible to anybody who is interested in correcting images for movement caused by physiological events. The difficulty of the proposed algorithms varies from simple to more complex, according to the problem they are designed to solve.

---

[1]FIASCO stands for Functional Image Analysis Software Computational Olio ([22])
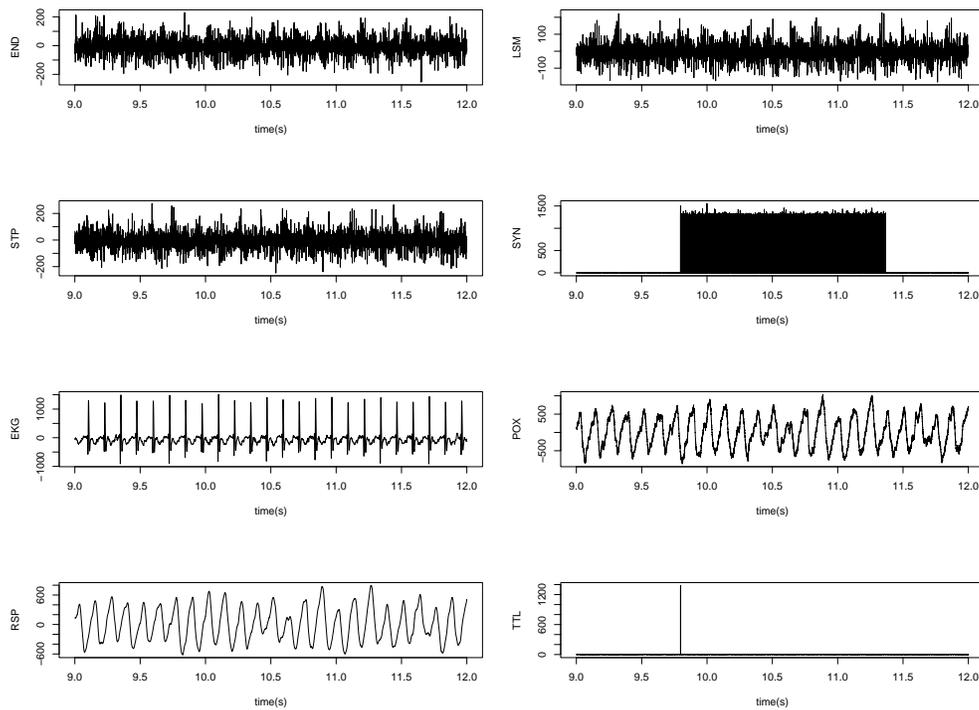
Figure 2.7: Physiological and system- related data presented in eight channels starting with second 9 during a mouse experiment. END represents the pulse at End of experiment from microscope; LSM wait pulses from microscope; STP scan Stop pulses from microscope; SYN line Syncs from the microscope; EKG electrocardiogram; POX Pulse Oxygenation ratio; RSP strain gauge around the diaphragm-currently unused shows the linkage effect of channel 6; TTL spikes when the microscope starts scanning a slice.

3.0   FILTERING IMAGE DATA

The obvious is that which is never seen until someone expresses it simply. (Kahlil Gibran)

Most images are affected at some extent by noise, i.e. unexplained variation in data. Image analysis is often simplified if this noise is filtered out. Thus, filters have been created to operate on the image and accentuate the features of interest. Filters change a pixel's value taking into account the values of neighboring pixels. Depending on the operations used, filters may introduce some artifacts in the data.

In the following sections we are going to present the general methods used for filtering noisy images, corrupted by random noise. We intend to clarify the features scanned via TPLSM by altering the initial image as little as possible. Therefore we develop a new method to filter out some of the variations in pixel intensity.

3.1   GENERAL IMAGE FILTERS

As Glasbey stated in [15], linear filters are weighted linear combinations of the pixel values in the image. If the weights are positive then the filtering effect is to smooth/blur the image, while negative weights emphasize edges. Although the moving average filter is simple and fast, it has two major drawbacks: it is not isotropic (i.e. circularly or spherically symmetric), and smoothes further along diagonals than along rows and columns; weights have an abrupt cut-off rather than decaying gradually to zero, which leaves discontinuities in the smoothed image (see fig.3.1). These drawbacks are overcome by Gaussian filters, which are separable and, at least to a lattice approximation, circularly symmetric.

As with all smoothing operations, there is a fundamental trade-off between variance and

bias: a filter which operates in a larger neighborhood will be more effective at reducing noise but will also blur edges.

If an image has been contaminated by noise and blurring of known forms, then filters can be constructed which optimally (in some sense) restore the original image.

Formally, we could state that the observed image f is a blurred version of an underlying 'true' image g, with added noise. We envisage that g is a clear image, unaffected by noise and with distinct boundaries between objects. We can use information about the nature of the degradation to design a filter which will smooth f and enhance the edges, such that it gets as close as possible to restoring g. Provided that we can consider g to be the realization of a random process and subject to various technical conditions, the Wiener filter is the best linear predictor (in the least square sense) of g from f ([15]). However, some non-linear filters may out-perform it. In filtering to reduce noise levels, linear smoothing filters inevitable blur edges, because both edges and noise are high-frequency components of images. Nonlinear filters, median ones for example, are able to simultaneously reduce noise and preserve edges. However, they can be computationally expensive to use. Also, they can generate spurious features and distort existing features in the image ([15]).

In our case, the laser beam captures fluorescent pixels from neighboring slices in the brain, contaminating the pixels in the actual scanned slice. We may include the physics of the scanning process later, and develop suitable algorithms for deconvolution, to reverse the blurring process. The shot noise, however, can be identified and its impact diminished. The moving median filter is similar to the moving average, except that it produces the median rather than the mean of the pixel values in a square window centered at the current pixel coordinates.

In practice we would select the best window length such that the background noise is considerably reduced, while the feature structure is minimally affected. As seen in Fig.3.1 and Fig.3.2, the median filter is preferred versus the average one, as the intensity of the pixels does not change much than the original version and the features are clearer. However, some of the methods we used for the analysis are quite sensitive to different sources of variation in the data, so, when working with images where pixel values can be an integer number between
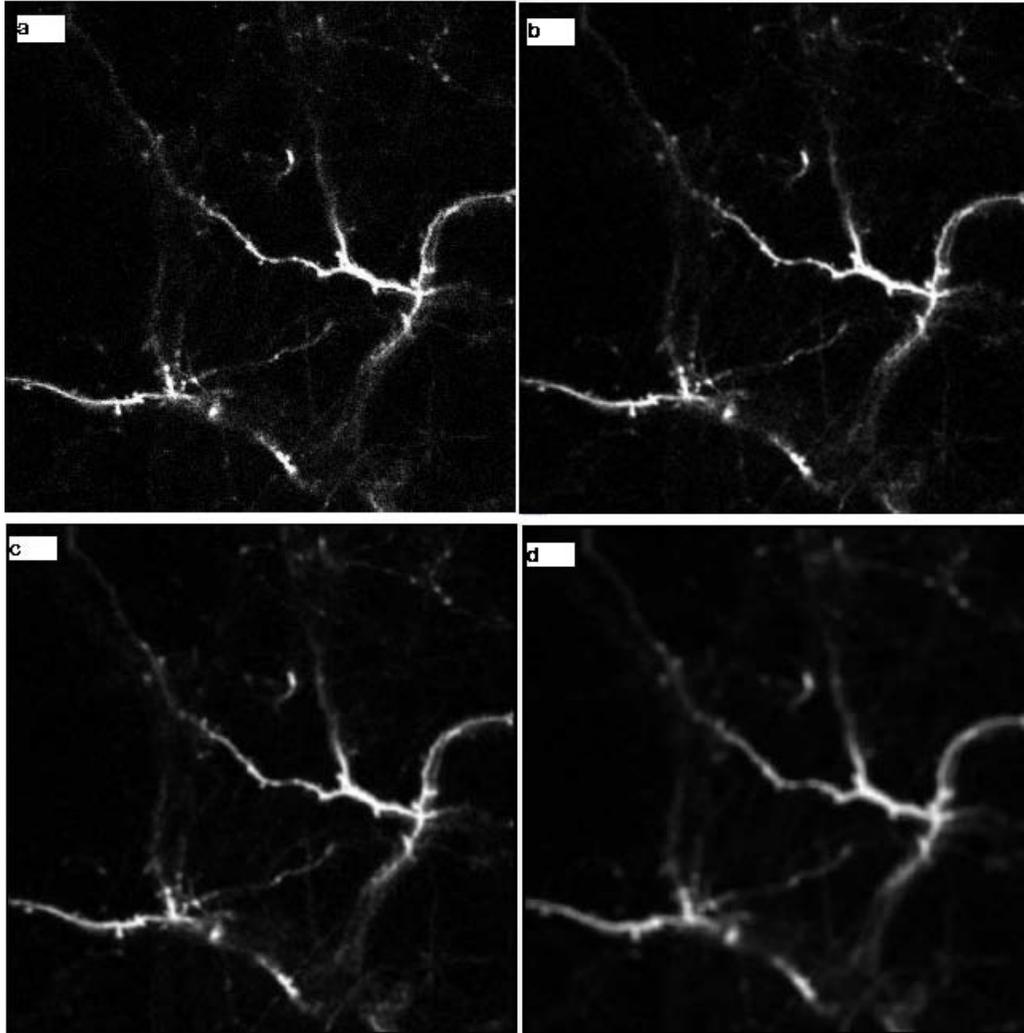
Figure 3.1: 2D Moving average -filtered versions of an image: (a) original image, (b) moving average with window-length= 3, (c) moving average with window-length= 5, (d) moving average with window-length= 9. As we increase the window-length, the background noise is reduced, but the image becomes more blurred.
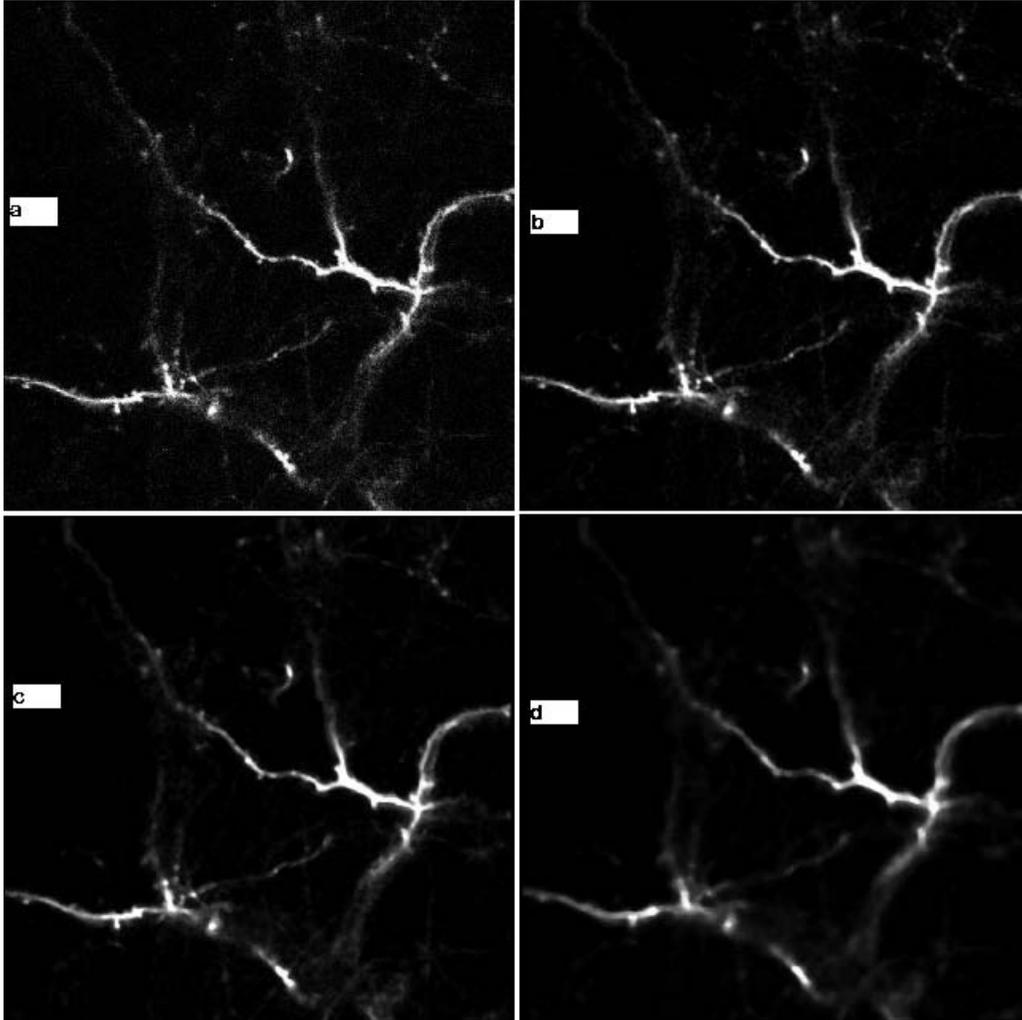
Figure 3.2: 2D Moving median -filtered versions of an image: (a) original image, (b) moving median with window-length= 3, (c) moving median with window-length= 5, (d) moving median with window-length= 9. As we increase the window-length the background noise is reduced, but the edges of the features in the image are visibly affected.

0 and 4095, we have a lot of variability introduced by pixel intensities. Consequently, we would like to adjust for this variation as much as possible.

3.2 CORRELATION-BASED FILTER

In our case, an improvement to the previous methods would consider the distribution of the intensities of the pixels in the image. Therefore we developed an algorithm which creates a new image with a smaller number of intensity values. The new intensities are assigned based on thresholds which are selected to maximize the correlation function between the new image and the initial image. More details and examples are presented in the following subsections.

3.2.1 Bi-colored Image

Krishnan Padmanabhan, former student in Dr. Crowley's lab, at Mellon Institute, has noticed that for any image with M possible pixel intensities we can construct an equivalent bicolor image, where each pixel can be 0 or 1 ([26]). The criteria to decide the value each pixel gets in the created image is based on a threshold selected such that the correlation between the new image and the original one is maximized. Essentially, the image created in this way will assign a value of 0 for the background and a value of 1 for the features. There is still some loss of information as the continuity principle may no longer hold, and some gaps in the features can be created. However, the variability due to pixel intensity is significantly decreased.

Considering an initial image $\underline{x}$, let us denote with M the total number of grey levels and with N, the total number of pixels in image $\underline{x}$ (in our case M is 4046 while $N = 512 * 512 = 262144$).

For any t taking values between $min(\underline{x}) + 1$ and $max(\underline{x}) - 1$, the bicolor image will have a pixel with value of 0 if the correspondent pixel has an intensity smaller than t in the original image ($\underline{x}$) and 1 otherwise.

correlation between original Image and binary image constructed with the threshold value
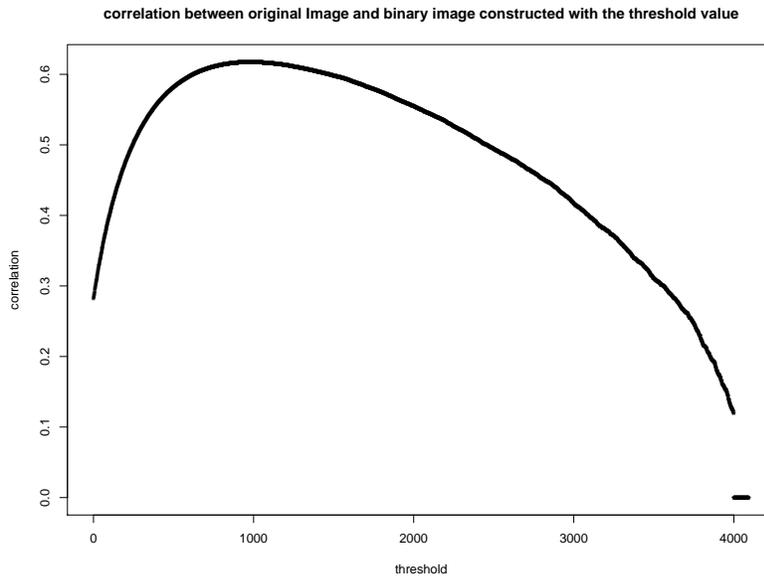
Figure 3.3: Correlation between original image and the transformed one (using only one threshold) as a function of the threshold value.

As an example, if $\underline{x}$ is the slice imaged in Fig2.5, then for each $1 \leq t \leq 4094$, we can construct a bicolor image $(\underline{y}(t))$ whose correlation coefficient with $\underline{x}$ will be a function of t.

Out of all these bicolor images, we want to select the one that preserves best the information from the original image. As plotted in Fig.3.3, the correlation function admits a maxima of 0.617 at $t = 978$.

The thresholded image (fig3.4) constructed with the maximal t value seems to preserve to a good extent the information captured in the original image.

The distribution of the pixel intensities is considerably simplified, but there is still some loss of information as some original pixels may have come from a feature but they were estimated as background in the newly created image. To minimize this loss we can consider a larger range for the intensities in the new image. Consequently, the correlation between the original image (with M grey levels) and the new one (with $K << M$ grey levels) will increase.
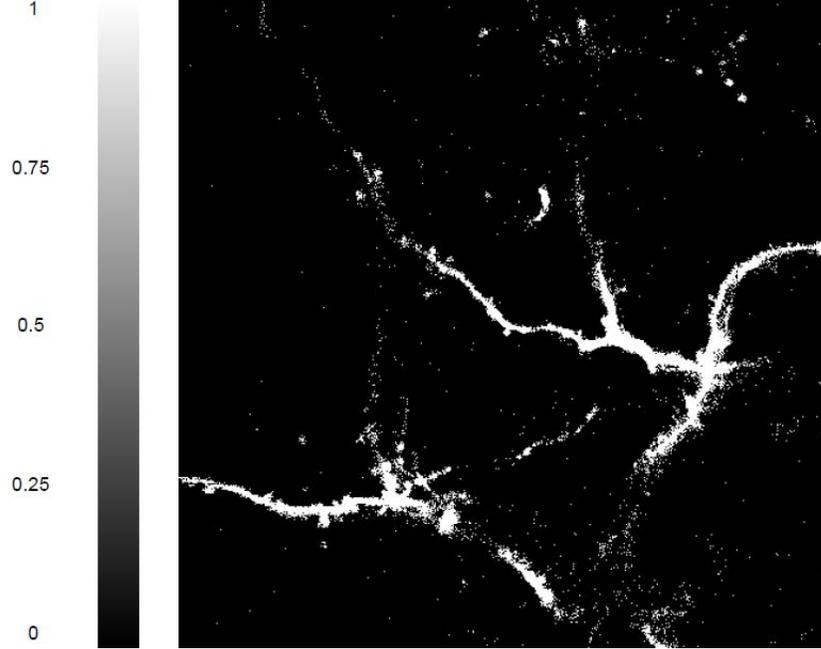
Figure 3.4: The bicolor version of fig 2.5 constructed with the threshold that maximizes the correlation function from fig 3.3. Each pixel can have an intensity value of 0 (background) or 1 (feature).

We extended Krishnan's idea for at least two thresholds ($K \geq 2$). The best number of thresholds can be estimated such that it optimizes a decision function. The details with some examples will be presented further in this section.

3.2.2 K-Colored Image

For the two thresholds case, we will define a new image with only three grey levels.

Let $0 < k_1 < k_2 < M$ be two thresholds that will be used to construct the new image $\underline{y}(k_1, k_2)$. We want to find those thresholds such that the correlation between $\underline{x}$ and $\underline{y}(k_1, k_2)$ is maximized. This estimation problem is well defined as the correlation function is bounded $[-1; 1]$.

For each $1 \leq i \leq N$, $\underline{y}$ is given by

$$y(i) = \begin{cases} 0 & \text{if } x(i) \leq k_1 \\ 1 & \text{if } k_1 < x(i) \leq k_2 \\ 2 & \text{otherwise} \end{cases}$$

The summaries of these images that are necessary to calculate the correlation function are:

- $\overline{x} = \frac{\sum_{j=0}^{M} j n_j}{N}$

- $Var(\underline{x}) = \frac{\sum_{j=0}^{M} n_j (j - \overline{x})^2}{N}$

- $\overline{y} = \frac{S_1 + 2S_2}{N}$

- $Var(\underline{y}) = \frac{S_0 (0 - \overline{y})^2 + S_1 (1 - \overline{y})^2 + S_2 (2 - \overline{y})^2}{N}$

where n_j is the number of pixels in $\underline{x}$ with intensity equal to j and S_k is the number of pixels in $\underline{y}$ that have intensity grey level k, with $k = 0, 1, 2$.

Let denote

- N_k as the number of pixels of $\underline{x}$ whose grey level is larger than k,

- $\overline{x_k}$ as the mean of the intensities of those pixels in $\underline{x}$ with grey level larger than k.

Then, the correlation function $corr(\underline{x}, \underline{y})$ is:

$$\frac{N_{k_1}(\overline{x_{k_1}} - \overline{x}) + N_{k_2}(\overline{x_{k_2}} - \overline{x})}{\sqrt{(\sum_{j=0}^{M} n_j (j - \overline{x})^2)((N - N_{k_1})(0 - \overline{y})^2 + (N_{k_1} - N_{k_2})(1 - \overline{y})^2 + N_{k_2}(2 - \overline{y})^2)}} \qquad (3.1)$$

Generalizing for K thresholds, for each $1 \leq i \leq N$, the new image $\underline{\mathbf{y}}_{k_1,\ldots,k_K}$ is given by:

$$y(i) = \begin{cases} 0 & \text{if } x(i) \leq k_1 \\ 1 & \text{if } k_1 < x(i) \leq k_2 \\ \vdots & \\ K - 1 & \text{if } k_{K-1} < x(i) \leq k_K \\ K & \text{otherwise} \end{cases} \qquad (3.2)$$

with $min(\mathbf{x}) < k_1 < \ldots < k_K < max(\mathbf{x})$.

Then the correlation function to be maximized becomes:

$$corr(k_1, \ldots, k_K \,|\, \underline{x}) = \frac{\sum_{i=1}^{K} N_{k_i}(\overline{x_{k_i}} - \overline{x})}{\sqrt{(\sum_{j=0}^{M} n_j (j - \overline{x})^2)(\sum_{i=0}^{K} (N_{k_i} - N_{k_{i+1}})(i - \overline{y})^2)}} \qquad (3.3)$$

with some adjustments:

- $\overline{y} = \frac{\sum_{i=1}^{K} N_{k_i}}{N}$

- $N_0 = N$ and $N_{K+1} = 0$

As the variance of $\underline{x}$ is constant with respect to $k_1 \ldots k_K$ and $\overline{y} = \frac{\sum_{j=1}^{K} N_{k_i}}{N}$, the correlation function in 3.3 is proportional to:

$$\rho(k_1, ..., k_K \,|\underline{x}) \;\; = \;\; \frac{\sum_{i=1}^{K} P(X > k_i)(\overline{x_{k_i}} - \overline{x})}{\sqrt{\left(\sum_{i=0}^{K} \left(P(X > k_i) - P(X > k_{i+1})\right)(i - \sum_{j=1}^{K} P(X > k_j))^2\right)}}, (3.4)$$

where X is the random variable represented by the value of a randomly selected pixel in original image $\underline{x}$.

Its cumulative distribution function is given by $P(X \le k_i) = 1 - \frac{N_{k_i}}{N}$. We present more details on calculations in Appendix B.

We would like to estimate the K thresholds as the solution that maximizes the above correlation function 3.4. A *grid search* which calculates all the values of the correlation function for any possible value of the thresholds requires a large number of steps $(O(M^K))$. Since the complexity of this exhaustive algorithm is exponential, calculating all possible values for function 3.4 is practically impossible. Although for $K = 1, \ 2, \ and 3$ we were able to get the entire data set of the values of function 3.4, for $K \ge 4$, this attempt was no longer possible.

From a simplistic point of view, our problem is to determine a K-dimensional $(k_1, ..., k_K)$ solution that maximizes a function, subject only to the constraint that the thresholds maintain a strictly increasing order.

$$min(\underline{x}) < k_1 < k_2 < ... < k_K < max(\underline{x}) \tag{3.5}$$

One possible way to obtain an optimized solution without iterating through the entire solution space is to use search algorithms based on various heuristics. Out of this category, we used a genetic algorithm based on Emanuel Falkenauer's idea. In 1997 he proposed an evolution concept for optimally solving some complex problems, such as clustering or partitioning, by making the characteristics of the groups of items equivalent to genes [11]

Genetic algorithms are a family of computational models inspired by evolution. These algorithms encode a potential solution to a specific problem on a simple chromosome-like data structure and apply recombination operators to these solutions in order to obtain new candidates. The recombination operators are generally designed such that they preserve critical properties of the potential solutions.

An implementation of a genetic algorithm begins with a population of (typically random) chromosomes. Each of these chromosomes represents a possible solution to the given problem. At every step a new population is created based on the current one by applying 2 operators, one for recombination (crossover) and one for mutation. The algorithm evaluates each chromosome (solution) and allocates reproductive opportunities such that those chromosomes which represent a better solution to the target problem are given more chances to "reproduce" than those which are poorer solutions. The "goodness" of a solution is typically defined with respect to the current population. At the end of each step, the "best" chromosomes survive in the population and then participate in the algorithm at the next step. Iteratively, these points migrate in the search space towards better solutions and the algorithm generally ends after a certain number of steps or when the best candidate stops improving. Most users of genetic algorithms typically are concerned with problems that are nonlinear. This often implies that it is not possible to treat each parameter as an independent variable which can be solved in isolation from the other variables.

The genetic algorithm considers also that as long as the number of "good solutions" to a problem is sparse with respect to the size of the search space. Therefore random search or search by enumeration of a large search space is not practical for a problem solving. On the other hand, any search other than the random search introduces some bias in terms of "how" and "where" it looks for better solutions.

In our case, we represented the chromosome as a potential solution $(k_1, ..., k_K)$ and we used the correlation function 3.3 as a measure for the "goodness" of each chromosome. The recombination function (crossover operator) simply selects two chromosomes and splits them at a random position. The front-end of one solution is combined with the back-end of the other and the result is kept in the population if it is a viable candidate. The mutation operator selects a single chromosome and changes one of its components k such that the
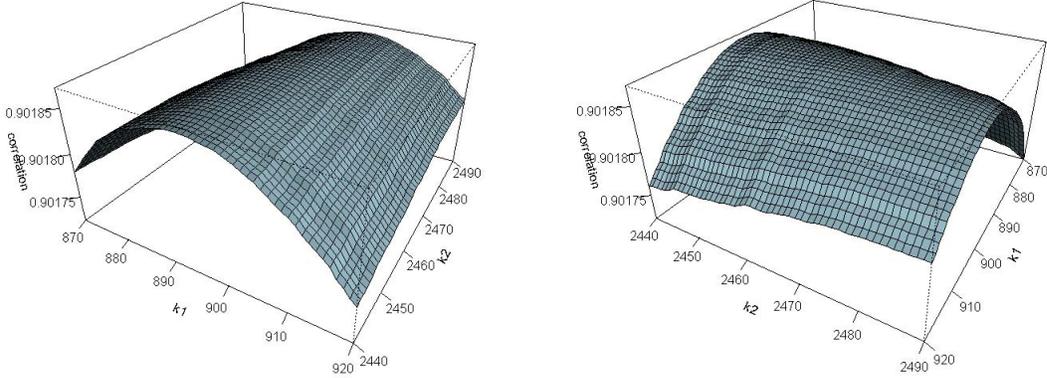
Figure 3.5: The correlation function between the original image (Fig.2.5) and any tri-colored image constructed with 2 thresholds that are displayed on x (k_1) and on y (k_2). Note: There are presented two views from different angles at a magnified resolution around the maximization region. Due to high dimensionality, 2D versions of the image with all the values for the correlation function are presented in B1.

result remains a viable candidate. In general, genetic algorithms define viability as the ability of each solution to satisfy the initial constraints. In our case, a solution is viable if all its K components maintain a strictly increasing order.

At each step, we started with $N - sol$ possible solutions (chromosomes) and applied the crossover operator $C - op$ times and the mutation operator $M - op$ times. For all viable candidates (up to $N - sol \ + C - op \ + M - op$) we computed the correlation function 3.4 and the top N solutions were maintained as an initial population for the next step. The iterative process was continued until the best solution stopped improving for $S (= 100)$ steps in a row.

One big caveat of this algorithm is that it does not guarantee a global optimizer. As our function has not a closed analytical form in $(k_1, ..., k_K)$, it is difficult to get the Hessian matrix. However, we think that the solution found by our algorithm is satisfactory if the correlation function is large enough (i.e. $1 - corr(\widehat{k_1}, ...\widehat{k_K}) \leq \delta(\underline{x})$, where $\delta(\underline{x})$ is a small number that depends on the original image) and the estimates are not on the boundary of the solution space.

For $K = 1$ and $K = 2$, calculating all the possible values of 3.3 and plotting them against the possible values for each threshold, the correlation function looks concave, hence it admits a global maxima. For a visual check, see fig3.3 and fig3.5 (B1).

38

Using the genetic algorithm idea, we were able to find very fast a solution for any K. The solutions for $K = 1$ and $K = 2$ were matched with those from the exhaustive search. Fig3.6 presents the resulting images, created using the thresholds estimated in this way. Table3.1 tabulates the results for $1 \leq K \leq 10$.

As expected, increasing the number of thresholds will improve the new image: the correlation gets larger and the information loss decreases. When plotted against the number of thresholds used, the maximum correlation for 3.3 has a fast increase after which it slows down. This fact will help us determine the best number of thresholds that will create a new image with less pixel variation. For the image 2.5, the plot below (fig 3.7) suggests that we can use a small number (less than 20) as a proximal value for K: let's use $K = 10$, for example.

The "optimal chosen" number of thresholds are used to get a filtered image (Fig 3.8) with only 11 grey levels for its pixels. As one can observe, the main features from the initial image appear very clearly, while some parts of other features, shaded originally, are emphasized now.

3.2.3 Theoretical Characteristics

Although we are interested in reducing the number of pixel intensities preserving the main characteristics of the image, further theoretical steps could be taken. We would like to investigate whether the function 3.4 could be simplified.

The probabilistic form for the component variables changes the domain of definition from $\mathbf{N}$) to $(0, 1)$. The cumulative distribution function of a random variable is distributed $Uniform(0, 1)$, therefore 3.4 is equivalent to:

$$\propto \frac{\sum_{i=1}^{K} (1 - u_i)(\overline{x_{k_i}} - \overline{x})}{\sqrt{\sum_{i=0}^{K} (u_{i+1} - u_i)(i - \sum_{j=1}^{K} u_j)^2}}, \tag{3.6}$$

where u_j is varying as a U(0,1) variable, such that $u_{j+1} \geq u_j$ for any $0 \leq j \leq K$.
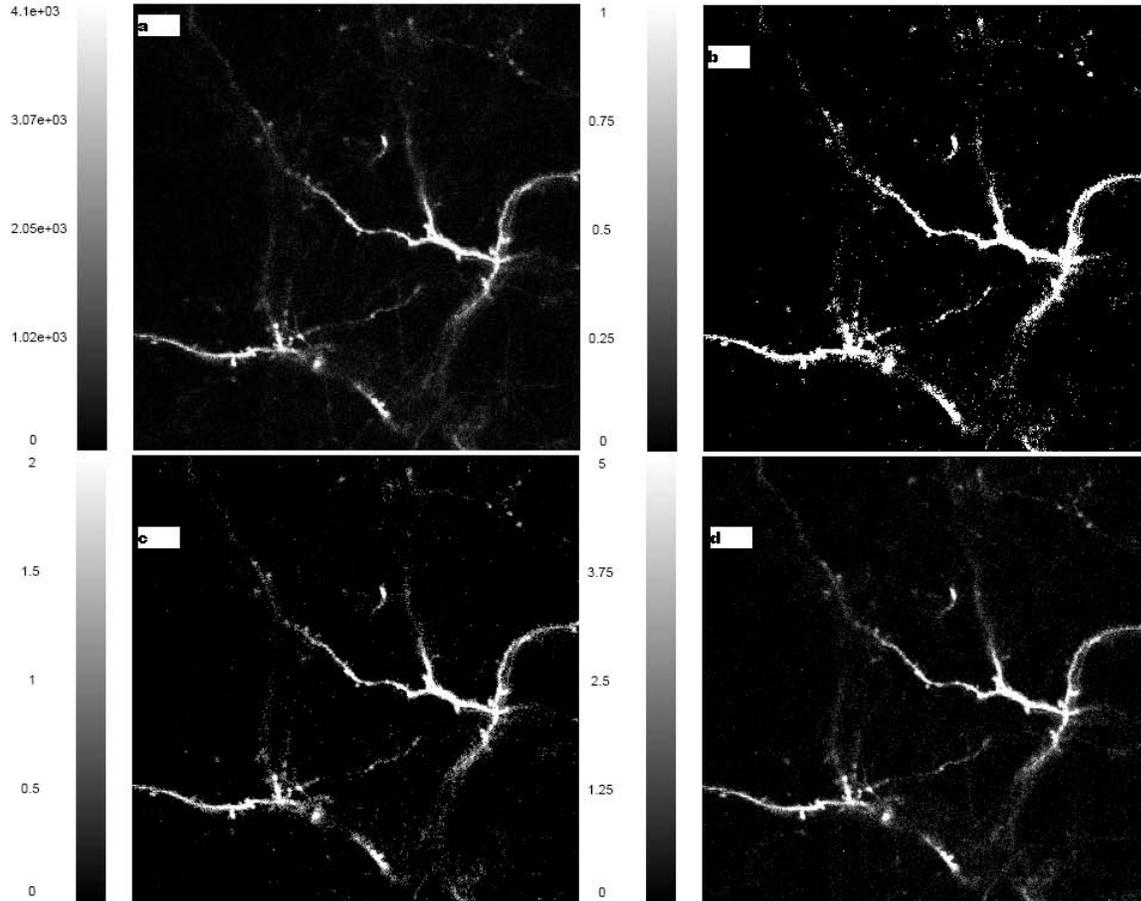
Figure 3.6: Correlation based -filtered versions of an image: (a) original image, (b) bicolor image (K=1, t=978, max-cor=0.617), (c) tricolor image (K=2, $t_1 = 896$, $t_2 = 2462$, max-cor=0.90187), (d) hexacolor image (K=5, $t_1 = 418$, $t_2 = 1129$, $t_3 = 1840$, $t_4 = 2550$, $t_5 = 3261$). As we increase the number of thresholds, the loss of feature information is reduced and the correlation increases. Note: the grey scale is modified, too.

K	Max-corr	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}
1	0.617	978									
2	0.90187	896	2462								
3	0.93283	643	1746	2850							
4	0.95087	508	1374	2239	3107						
5	0.96257	418	1129	1840	2550	3261					
6	0.97073	354	959	1563	2169	2771	3375				
7	0.97656	309	839	1371	1902	2430	2961	3495			
8	0.98092	270	739	1209	1677	2146	2614	3084	3556		
9	0.98426	242	664	1085	1509	1931	2353	2777	3198	3622	
10	0.9868	218	604	988	1371	1755	2140	2525	2910	3292	3676

Table 3.1: The estimates that optimize the correlation function 3.3 for different number of thresholds. These values are used to construct a new image (as defined by 3.2) whose correlation with original image is high (first column).
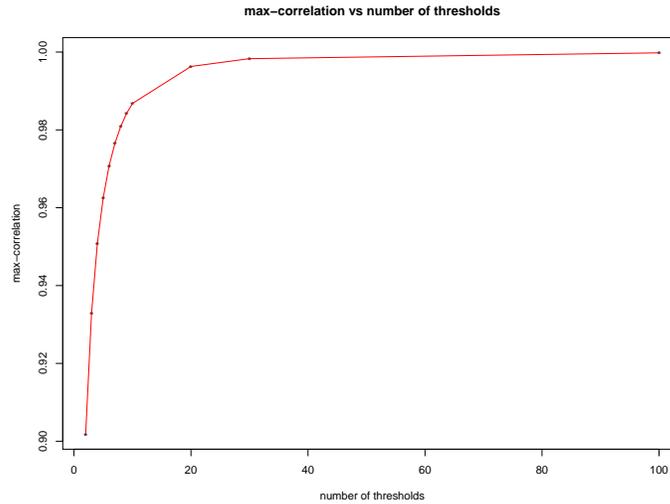
Figure 3.7: The trajectory of the maximum correlation between the original image and the K color image, with K varying on the x axis. Note: around K=10 the increase in the max-correlation slows down significantly.
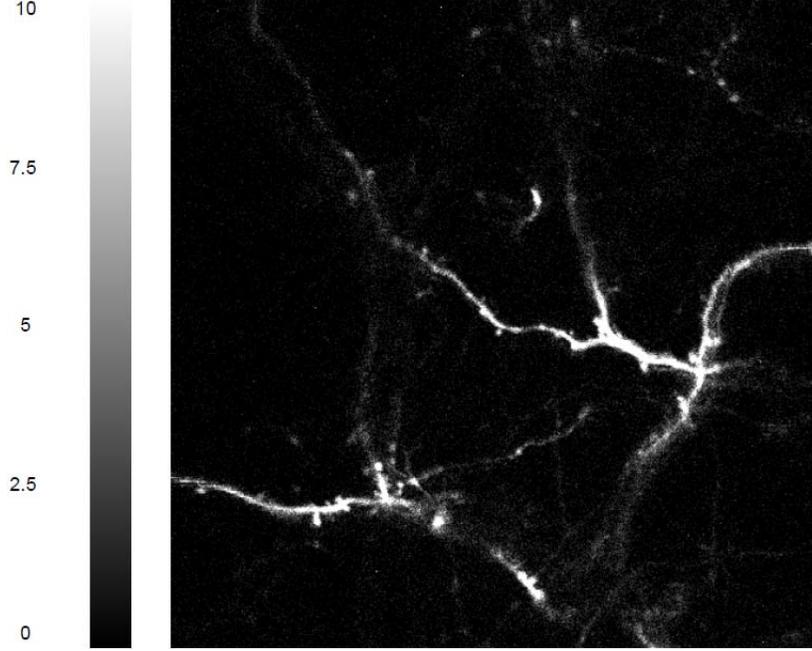
Figure 3.8: The correlation -filtered image constructed based on K=10 thresholds (see Table 3.1). The resemblance with initial image (Fig2.5) is striking, while some other features, originally diluted in the background, appear clearer.

Taking advantage of the telescopic sums from the previous expression, we can write the argument of the square root function at the denominator as:

$$3(\sum_{i=1}^{K} u_i)^2 - \sum_{i=1}^{K}(2K + 2i - 1)u_i + K^2. \tag{3.7}$$

We present in more details how we got to this form in Appendix B.

This function, 3.7, is quadratic in each u_i, thus continuous. As the definition domain is a convex set (simplex on $(0;1)^K$), the support of 3.7 is bounded away from 0. Its Hessian is $6 * \underline{1}x\underline{1}^T$ which is a positive definite matrix (see appendix Ai). Therefore this function admits a global minimum.

The square root of a quadratic function is a hyperbola function. As the domain in our case is defined for positive numbers only, the resulting function is half of a hyperbola.

Incorporating this information into the correlation function 3.6 and considering that the numerator is a monotonically decreasing function in each u_i we need further investigation on the characteristics of this function.

42

Image filter	Parameters	Number of grey levels	Correlation with initial Image
2D-Mean	window=3	3700	0.9244
	window=5	3734	0.9046
	window=9	3431	0.8553
2D-Median	window=3	3700	0.922
	window=5	2882	0.90846
	window=9	3431	0.8658
Correlation based	3 thresholds	4	0.933
	5 thresholds	6	0.9626
	10 thresholds	11	0.9868

Table 3.2: Comparisons of different image filters.

3.3 CONCLUSIONS

A visual examination yields that our proposed filter is producing more satisfactory images when compared with mean and median filters. Table 3.2 summarizes some statistics calculated based on theses resulted images.

Even though the thresholds estimated by our method give very good results, we investigated the correlation function in more detail. We produced a simplified version of the K dimensional function and we plan to study further more its characteristics.

4.0   PRELIMINARY ANALYSIS

A little knowledge that acts is worth infinitely more than much knowledge that is idle.
(Kahlil Gibran)

4.1   TIMING BETWEEN SCANS AND PHYSIOLOGY

While collecting images from the brain of neonatal mice, the circulatory and respiratory processes happen at a rate of 5Hz and about 1.5Hz, respectively.

The pixels have very small sizes (0.3x0.3x1 $micron^3$) allowing high resolution imaging. Depending on the level of detail, the time to record a pixel is varying in a narrow range of microseconds. The data we used was collected at 0.92 to 2.56μs. Therefore, one line of 512 pixels is scanned almost instantaneous (around 0.5-1 ms) with respect to the physiology rates. An image of 512 lines, however, takes approximately 1 second, while recording an entire stack may take several minutes. Due to various physiological fluctuations during the scanning of a slice (heartbeat, breathing, etc), the features of interest can shift from one image to another. Since measuring the size and position of the dendrites and axons is in fact the goal of these experiments, random variations introduce an undesirable error. These sizable deformations caused by brain movement keep the researchers from successfully aligning the images and comparing them.

Since it is not possible to eliminate the factors that affect the image recording, we need to be aware of the artifacts even if they are not always obvious to the naked eye. Based on the observed flaws, we tried several methods to account for this displacement both globally and locally.

4.1.1 Matching physiological data with image data

As pixels are scanned sequentially over a period of up to several seconds, normal physiological events impact pixel location. Changes in internal blood pressure during the cardiac and respiratory cycles are the principal causes for the brain movement in the anesthetized rodent. For the small animals that we use in our study, a normal heart rate is about 5 beats per second while the respiration rate is close to two events per second (between one and two respirations). These physiological characteristics are collected using the Brain Morey software, which can trigger the imaging in synchronicity with the cardiac cycle to ensure that the variation remains consistent. In order to have control over the respiration events, the experimental rodent is anesthetized and fitted with an endotracheal tube for artificial respiration.

The pulse-oximeter provides us with a measured value of blood activity that we can use later in the analysis. This device summarizes the physiological channels into a data series which oscillates at the same rate as the EKG.

As Shumway, et al. stated, the regularity of these physiological measurements can be best expressed in terms of periodic variations in the underlying phenomenon through Fourier frequencies (typified by sines and cosines). From the regression point of view, we may imagine a system responding to various driving frequencies by producing linear combinations of sine and cosine functions. On these grounds, the time domain approach may be thought of as a regression of the present on the past. The frequency domain offers an approach that may be considered as regressing of the present on periodic sines and cosines ([29]).

The most frequently used tool to disentangle complex processes to simpler forms is the Fourier transform. For a general function $u(t)$ with $t \in \mathbf{Z}$ that satisfies the absolute summability condition:

$$\sum_{t=-\infty}^{\infty} |u(t)| < \infty \tag{4.1}$$

the Fourier transform would be a pair of the form:

$$u_{(}t) = \int_{-1/2}^{1/2} U(w) exp(2\pi iwt) dw \tag{4.2}$$

and

$$\mathbf{F}(u) = U(w) = \sum_{t=-\infty}^{\infty} u(t)exp(-2\pi iwt) \tag{4.3}$$

The representation in 4.2 is the spectral density of u_t, while 4.3 corresponds to its inverse or Fourier transform.

For a discrete number of times, when $0 \leq t \leq (N-1)$, $u(t)$ and $U(w)$ can be seen as n-dimensional vectors. In this context, using the regression point of view we can formulate the following relationship:

$$\underline{u} = W\underline{U}, \tag{4.4}$$

where N*W = $\begin{pmatrix} 1 & 1 & .. & .. & 1 \\ 1 & e^{i2\pi/N} & .. & .. & e^{i2\pi(N-1)/N} \\ .. & .. & & .. & .. \\ 1 & e^{i2\pi(N-1)/N} & .. & .. & e^{i2\pi(N-1)(N-1)/N} \end{pmatrix}$.

Therefore, estimating $\underline{U}$ implies actually regressing $\underline{u}$ on W, which, based on Euler formula is a collection of sines and cosines. Also, W is full rank with $W^{-1} = (e^{-i2\pi kj/N})$, with $0 \leq k \leq N-1$, and $0 \leq j \leq N-1$. The properties of W make the Fourier transform to have, among others, the following characteristics:

- Linearity

$$\mathbf{F}(af(t) \pm bg(t)) = a\mathbf{F}(\omega) \pm b\mathbf{F}(\omega) \tag{4.5}$$

- Time scaling

$$\mathbf{F}(f(at)) = \frac{1}{|a|}\mathbf{F}(\frac{\omega}{a}) \tag{4.6}$$

- Time shifting

$$\mathbf{F}(f(t-t_0)) = e^{-j\omega t_0}\mathbf{F}(\omega) \tag{4.7}$$

- Frequency shifting

$$\mathbf{F}(f(t)e^{j\omega_0 t}) = \mathbf{F}(\omega - \omega_0) \tag{4.8}$$

As it can be observed in 2.7, our physiology data (POX) is fairly periodic but embedded in noise, so we need to isolate the underlying signal. In order to do that we can take advantage of the properties of this transform to construct linear filters. Therefore, we modify the spectral characteristics of the time series in a predictable way.

A set of specified coefficients a_t, with $t \in \mathbf{Z}$ and $\sum_{-\infty}^{\infty} |a_t|$, is used to transform an input series, x_t into:

$$y_t = \sum_{r=-\infty}^{\infty} a_r x_{t-r} \tag{4.9}$$

Under these conditions, y_t in 4.9 exists as a limit in mean square.

In our case, a correct choice of the coefficients will help smooth over the physiology series (POX) by removing the high frequency noise. Setting $a_r = 0$ for $r \geq tr_x$, with tr_x being a threshold value depending on the input, we keep the range of frequencies that should correspond to the underlying phenomena (low pass filter).

An example of using this algorithm is presented in fig.4.1. The top panel shows the input POX series during one experimental second, while the bottom panel shows the smoothed output of the signal driven by the heart rate.

One of the advantages of using this POX series is the fact that it was designed to capture both the heart signal and the respiration. Thus, we need to do some spectral analysis to identify the dominant frequencies which describe the oxygenation of the blood at an extremity.

Through the shape of its power spectrum, the POX series describes the respiration and heartbeat rates (see fig.4.1). At the same time, being recorded at the leg or tail, the pulse-oximeter may be considered a better factor in explaining the effect of the blood pressure in the brain than the EKG.

In order to work with the image and physiological data sets together, we need to merge them by matching the physiology value to the time when each pixel was recorded. Synch pulses indicate the times of acquiring image data. After some initial exploratory analysis, we determined that the system records roughly three physiological samples for every line scanned in a picture (see fig.4.2).

47
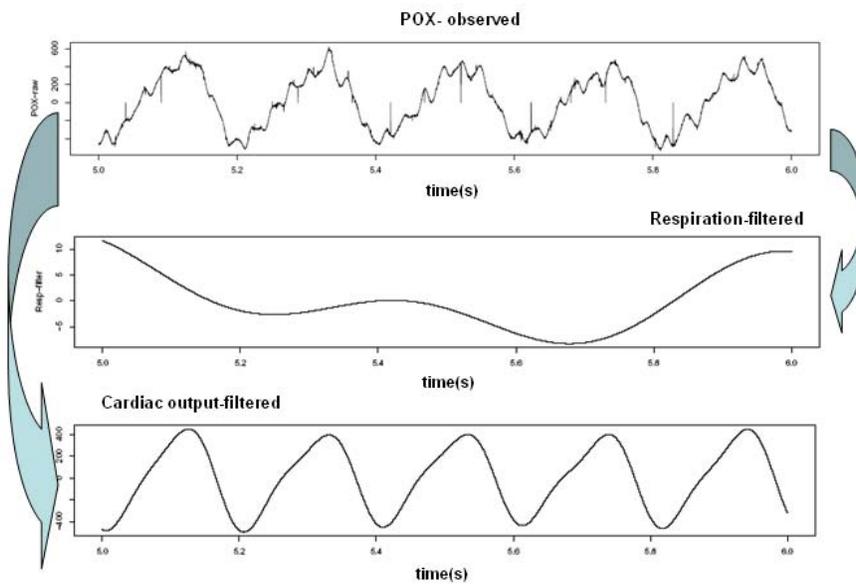
Figure 4.1: The respiration (around 2Hz) and the heart beat (5Hz) were recovered from the POX time series (1st panel) via the spectrum power and a notch filter around the suggested rate.
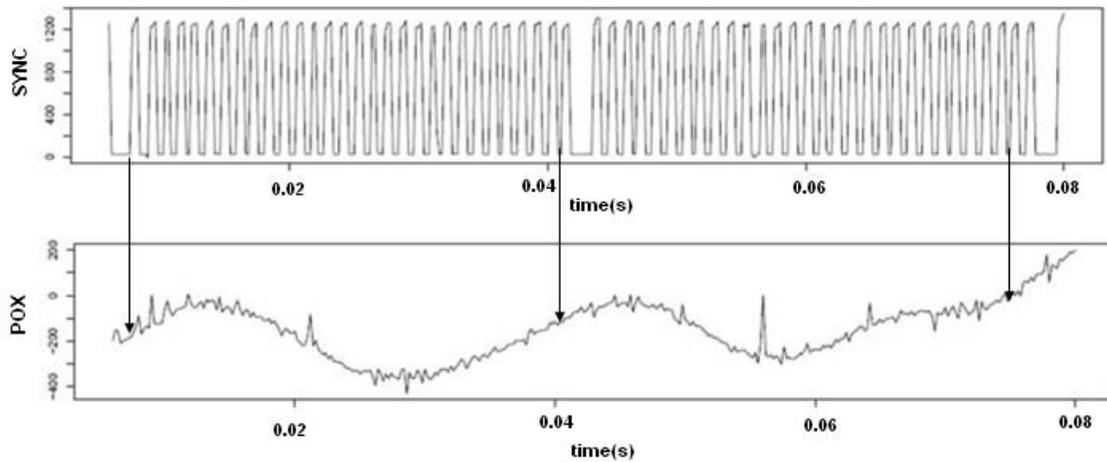
Figure 4.2: SYNC data collected at the same time with EKG and POX. After identifying the moment when the SYNC series spikes high values (the same as the moment when a line started to be scanned), the time to scan a line takes 3 to 6 samples of POX series.

The microscope records each pixel in the line sequentially, scanning a picture of 512 lines in approximately one second. In this interval, the movement in the brain causes some pixels to have an observed position shifted from the true one. Our observations suggest that each individual pixel can move up to ten pixels in any of the three dimensions. This causes some lines to be scanned multiple times, while some others are never scanned.

4.1.2   Estimating the true position of a pixel

To construct an image that is free of movement, we start with a set of "replicates", images taken at the same period in the animal development. They represent the same brain area, and, in the absence of movement noise they should show the exact same features.

In some of the methods we used, we combine the information collected from these replicates with the physiological values obtained at each collection time. The final goal is to be able to describe the distribution of the pixel displacement based on the pulse-oximeter (physiological) data, and consequently, to be able to recreate a stable image. In general ([10]), this will involve two steps:

1. determination (i.e., estimation) of the amount of motion and
2. correction of the data for that amount of motion

4.1.3   A synthetic example

In this section we illustrate theoretically the process of scanning a line.

Let us suppose that each line can be scanned only in three different positions p_i, where $1 \leq i \leq 3$, with probability of scanning the line in position p_i being $\frac{1}{3}$. Under the assumption that scanning is independent of movement, the number of positions X where each line could be scanned is a random variable $Bi(3, \frac{1}{3})$. Then $P(X = 0) = \frac{8}{27}$.

When we extend the conditions above to an arbitrary number of positions n, X (the variable counting in how many of these n possibilities the line was actually scanned) would be $Bi(n, \frac{1}{n})$. Consequently, $P(X = 0) = (\frac{n-1}{n})^n$.

It seems that increasing the number of positions will increase the probability of missing the line. Taking the limit in the above equation, we get: $lim_{n->\infty}(1 - \frac{1}{n})^n = \frac{1}{e}$ which is

always a positive quantity.

In reality, as we saw in previous sections, the process is more complex, but this simple example suggests that the scanning process can skip lines or can scan lines several times with a positive probability.

4.1.4 Filling the gaps caused by missing lines

Due to tissue movement, it is quite possible for the microscope beam to advance different distances between pixels. It takes approximately 500 microseconds to scan a single line (0.92 μs per pixel for 512 pixels per line plus the fly back time). This time only corresponds to about .25% of a heartbeat period and .1% of a respiratory cycle. Given the relatively short intervals, we can assume that a constant process is a good approximation for the subject's physiology for one scan line. For this reason, it makes sense to average the three physiological sample values recorded for each line. In our subsequent analysis we use one physiological value for each observed line in the picture. Depending on the experimental design, the time to scan an image with 512 lines is approximately 1 second. During this interval there are around 5 heart beats and 1.5 respiratory events.

On the horizontal direction (x axis), the scanning speed is fairly quick compared to the normal physiological processes, so we will assume that there are not significant errors. However, the time taken by the beam between lines on the vertical direction (y axis) is significantly larger and this can result in lines being scanned twice or lines missing altogether.

Registration techniques can provide some solutions for filling the gaps caused by such missing lines. Based on good estimations and suitable maps that minimize the bias, these procedures will predict the positions and intensities for the missing pixels. One challenge will be again to find the best model that will map this information.

4.2 REGRESSION METHODS

We define an image as a 2D array $I(x, y)$ of pixel intensities located at x and y distance with respect to an origin which is fixed at the upper left corner. We intend to identify parameters included in a general mapping $I_j(x, y) = g(I(f(x, y))$ that describes the phenomena of interest, where I_j, $1 \leq j \leq J$, are replicates of the "true" image of the brain I at a certain position on z axis. Even if in time there is a noticeable photo-damaging effect, two consecutive "repetitions" are acquired fast enough to assume that g is the identity function (or a white noise). As f function depends mainly on the physiology, we can consider it as a continuous function of time.

In the absence of the targeted image (I with previous notations) and without any additional information beyond what is contained in the data itself, there is no "correct" registration method that will relocate the pixels in the "exact" positions in the brain. However, we need some homogeneous transformations preserving the observed information. In the case of interpolation, a combination of a rotation and a two-dimensional translation will be required in order to estimate 3 parameters. If there is also a scaling problem the number of parameters will increase, affecting the amount of computations. Also, as the previous example has shown, the probability of missing lines is positive directing our search into this conditional space of scanned pictures.

In a preferred scenario, we would like to estimate the effect of the physiology upon the movement of pixel position on x, y, and z dimensions.

Considering the scanning rate fast enough to avoid displacement on the x direction, we used regression-like methods to get an estimate for the line (the y dimension) in the observed image. The underlying assumption here is that our high scanning speed eliminates errors on the x axis except for a rigid shift, so the initial images contain mainly movement errors on the y and z directions. Therefore, our first regression model 4.10 uses only the movement between the lines of the images taken at same z position.

As we develop our approach from a 2D perspective (intending to extend then to 3D), let us fix z for the rest of this section as the depth in the brain where the 3 replicates have been taken. Let $I_j(., y(t))$ be the observed intensity of a pixel located at time t on the line y(t)

in the jth replicated image. At the same time, the physiology $P_j(t)$ impacts on y-dimension, determining the true position of the line to be affected with $\beta * P_j(t)$.

We propose to model this aspect of the movement as:

$$I_j(., y(t)) = I_j(., y(t) + \beta * P_j(t)) + \epsilon(., t), \qquad (4.10)$$

where $I_j(., y(t) + \beta*P_j(t))$ is the intensity of the pixel located at the true y position.

We estimate β such that combining the information from different replicates would produce the same intensity for the pixel located on the true y position. In other words, the estimate of the physiological effect minimizes the mean of the squared deviances of pixels intensities measured at the same time in two consecutive "versions" of the zth slice:

$$SSQ = \frac{1}{512^2} * \sum_{x=1}^{512} \sum_{t=1}^{512} \left(I_1(x, y(t) + \beta * P_1(t)) - I_2(x, y(t) + \beta * P_2(t)) \right)^2 \qquad (4.11)$$

Next, $I_j(x, y(t) + \beta * P_j(t))$ is approximated by a Taylor expansion around $I_j(x, y(t))$:

$$I_j(x, y(t) + \beta * P_j(t)) = I_j(x, y(t)) + I_j^1(x, y(t)) * \beta * P_j(t), \qquad (4.12)$$

with $I_j^1(., y(t)) = \frac{I_j(., y(t)+k) - I_j(., y(t)-k)}{2}$, (see fig.4.3) where $k = 1 \ldots 4$ and $t = 1 \ldots 512$.

The custom made software, FIASCO, enabled us to get an estimate of β by minimization algorithms. We have used the Nelmin algorithm extensively, which is based on the Nelder-Mead simplex algorithm ([25]). Using this estimate for β, we can obtain an estimate of the true position of the lines and reconstruct a steady image. Even after applying this algorithm, some of the lines will still overlap due to rounding error and oscillations in physiology data. This leaves gaps in the resulting image, which we need to fill.

Repeating the algorithm for j = 1 and 3, and respectively j =2 and 3, we get two other estimates of the steady image that we can combine with the first one and fill in the gaps
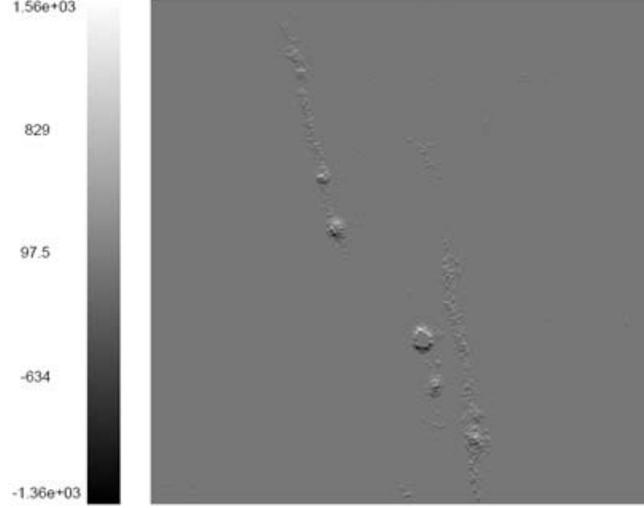
Figure 4.3: The derivative of spatial mapping in y coordinate as defined in 4.12, for k=1. We notice that the neighboring variations within pixels intensities are preserved.

wherever possible. However, some artifacts remain in the estimated steady image as it can be seen in Fig.4.4 from results section 4.3.

We extended this regression idea by elaborating on the algorithm so that β is estimated through an iterative algorithm. We start with an initial estimate for the true image as the arithmetic mean of the observed images. Then following the OLS approach we iterate over β to minimize the mean square error:

$$\widetilde{\beta}^{(s+1)} = argmin \sum_{j=1}^{3} \sum_{x=1}^{512} \sum_{t=1}^{512} (I_j(x, y(t)) - I^{(s)}(x, y(t)) - \beta * P_j(t) * I^{(s)1}(x, y(t)))^2 \quad (4.13)$$

Here $\widetilde{\beta}^{(s+1)}$ is the estimate of the physiological effect when the estimation of the true image at step s is $I^{(s)}(., y(.))$.

The algorithm converges when the difference between two successive estimates becomes very small. At convergence $I^{(s)}(., y(.))$ will be the desired estimate.
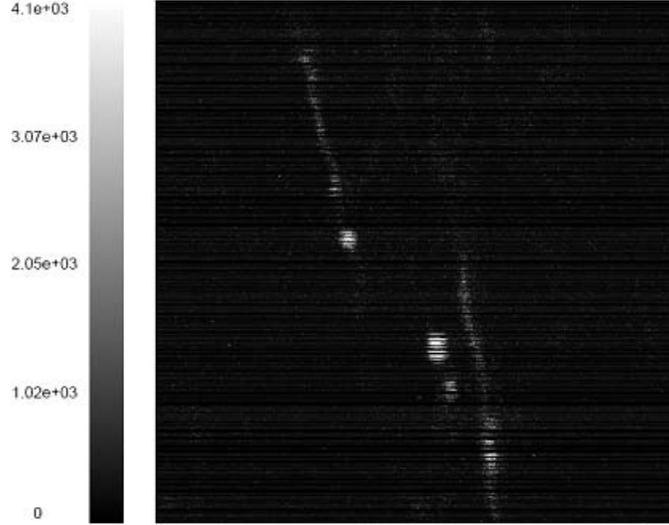
Figure 4.4: An estimate of the image adjusted for POX effect on y axis. It is reconstructed after estimating β with model4.10 using the replicates scanned at 1st and 2nd time.

4.3   RESULTED IMAGES

Based on the first regression model 4.10 we estimated the y positions adjusted for the effect of the concomitant POX value using the replicates taken at 1st and 2nd time. Moving the lines at these positions we obtained an estimated image that presents y positions with no tisue input. Hence the lines that are skipped during the scanning process are noticeable in this resulted image (4.4).

We repeated the algorithm of getting an image estimate by adjusting the POX effect on the replicates scanned at 2nd and 3rd time. The same process of "skipping"' lines was present. We calculated the rate at which the "gaps" appear, and, as expected they are quite close from one estimate to the other (see 4.5). The lines outputted in this way have been combined into a "better" image, in order to fill in tissue information captured at that depth.

The iterative model produced a very small estimate for 4.13 at first step, reducing the chances to go further. Also the choice of initial "guess" for the free-of-movement image and for filling the gaps in the next estimates might bias the results, by considering lines which are never scanned actually.

182 missing lines(.35)

190 missing lines(.37)
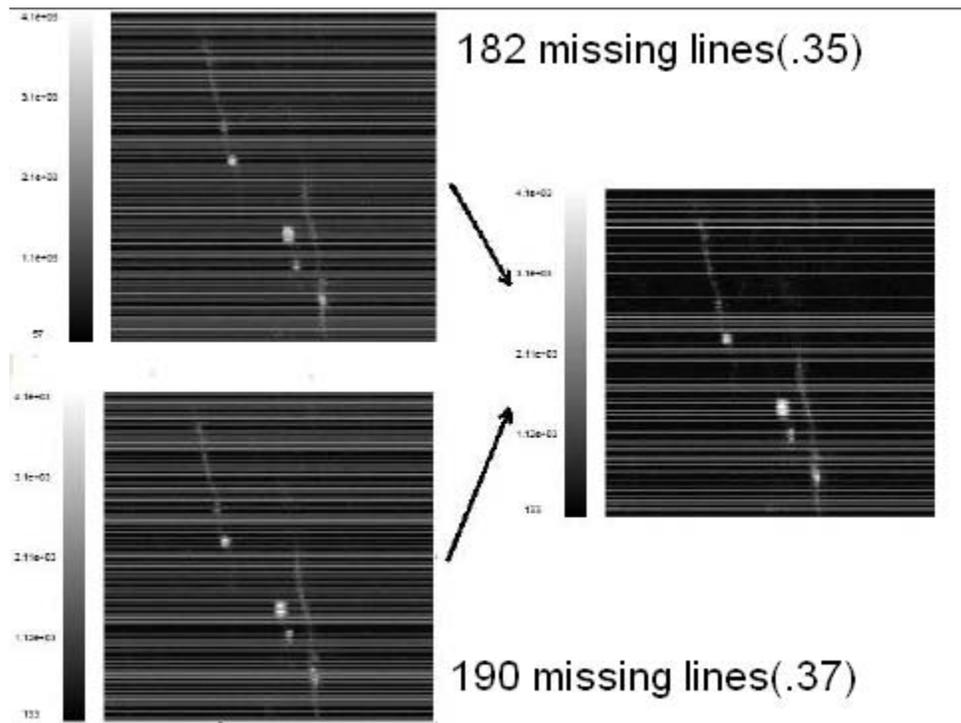
Figure 4.5: The two estimates of the image (adjusted for POX effect on y axis applying model 4.10 twice on pairs of replicates scanned at 1st, 2nd, and 3rd time) have been combined to get a "better" estimate.

4.4   CONCLUSIONS

In the previous section we approached the observed images from a classical statistic point of view which assumes complete independence over time. Using a statistically simple, but computationally intensive linear regression model, we showed that a pixel-wise treatment of the data gives reasonable results. It also points out that the true signal is contained in several scans of the same scene in the brain.

As we point out in section 4.1, the microscope does not record pixels fast enough so that the lines of an entire image to be considered independent over time. The empirical observations described in the next chapter exhibit the repetitive behavior of the image data. Therefore, we need to extend our research approach to describe and quantify this phenomenon. The β estimate that we produce with previous model (4.10) is computed under the general assumption that all lines in the image are affected linearly by the physiology. However, we need a better understanding of the underlying process which determines the movement of the brain.

5.0 REPEATED SCANNED LINE- EXPERIMENTS

If coming events are said to cast their shadows before, past events cannot fall to leave their impress behind them. (Helena P. Blavatsky)

Exploratory analyses exposed in the previous chapter 4 emphasize the need to have a better description for the movement in the brain. In order to accomplish this goal, we have performed a series of new experiments.

We have sampled several times through a middle line from a predetermined slice (Fig.2.5), recording the corresponding physiologies. At a fixed (y, z) position in the brain, a single line of pixels has been scanned in the x direction repeatedly. For a better comprehension of the features that we captured, the upper part of Fig.5.1 shows a region of twenty one lines around this location. The lower part of Fig.5.1 shows the image created from the lines acquired when scanning through the mentioned location for the first 500 times.

If the scanning process was not affected by the movement of the cerebral tissue, the features scanned in the first line would line up perfectly with all the other lines. However, the observed lines vary in a nearly periodic way. We show later in this section that this variation is strongly correlated with the physiological events. The change from one time to the other is more or less visible based on the shape and dimension of the feature captured in the extensively sampled line. In this example, the variation in pixel positions is visible especially for the first feature, as it originates from an oblique axon body.

As we explained in more detail in the previous section, scanning a line is quite fast with respect to the physiology rate. Thus, we can assume that we have a single measure for the physiology during the scanning of a line.

We are interested to study the relationship between animal's physiology quantified by the POX series and the imaging process. A useful graphical representation of this relationship
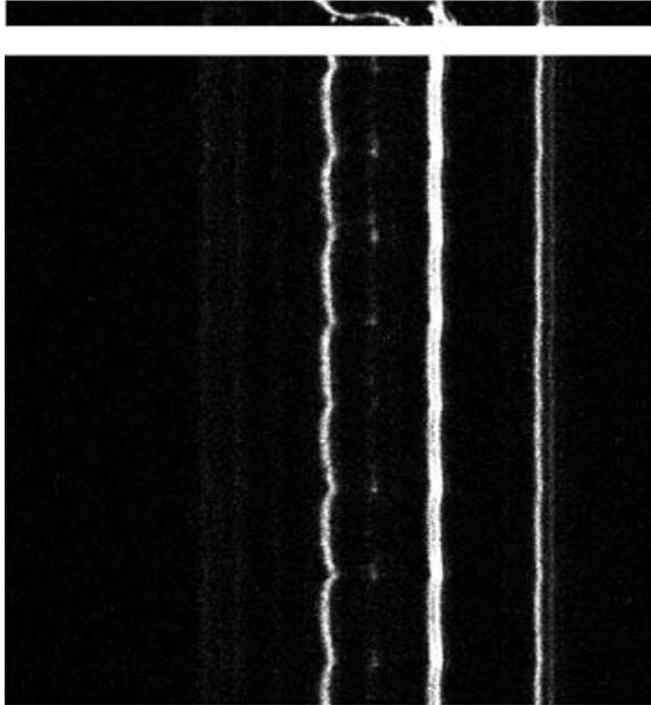
Figure 5.1: The middle line of upper panel was selected and sampled extensively for almost 37 seconds. First 500 lines are represented in the lower image. We can notice that, in fact it was not scanned the same line, as the feature position within the line varies in a nearly periodic way.

would be a scatter plot. Having one POX value per line and each scanned line representing a "replicate" of the same outcome of interest, we summarized each line as described above, with its average of the pixel intensities. When plotted versus the correspondent physiology as in Fig.5.2, the mean intensity value for each line is nearly cycling accordingly with the heart rate.

If there was no movement, except for some discoloration from the bleaching effect of the laser beam on the tissue, then the mean intensity value of each scanned line would display a clearly linear negative trend in time. However the brain is moving locally on "ellipses" -like trajectories due to different velocities for each direction. We can conclude that the physiology represented by the POX data is directly affecting the image quality.

In this thesis I present several approaches we considered to account for the effect of physiology on the image data. Adapting the design of the experiment to run repeated line -scans may give us a lot of useful insights for *in vivo* TPLSM.

In the following sections I describe directions we considered in order to model the way

Change in Line-Mean intesity versus POX
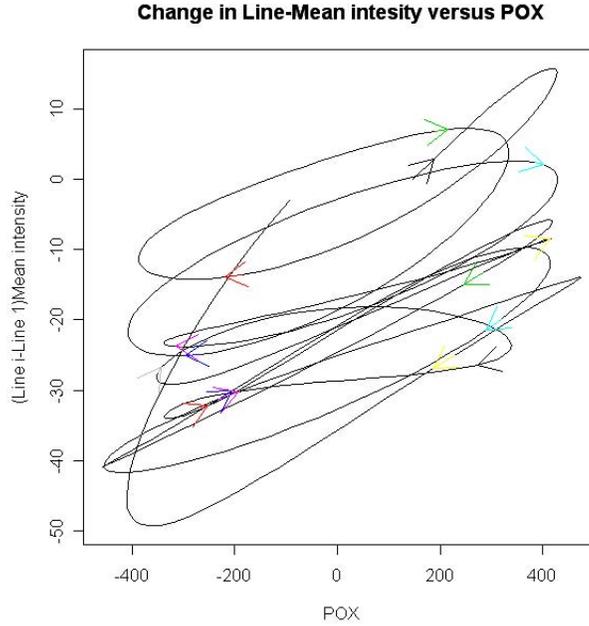
Figure 5.2: Mean intensity change for the first 512 lines with respect to first line, versus their correspondent POX values. Note: the arrows indicate the line-scanning succession, pointing out the motion of the tissue due to the physiology (POX).

these lines vary. Each step made us aware of the sources of variation and also of the limitations due to the position where they have been captured at the scanning time.

5.1   MISSING DATA IMPUTATIONS

Before we start any analysis on this data, we need to run some additional steps that will prepare the input variables. Matching the physiological measurements with the image data is an important step, which sometimes requires extra work. For this experiment, for example, the SYNC channel which monitors when the scanner started to produce image data, misses to record indicators for times when pixels are acquired. Usually the SYNC series spikes when the microscope begins to scan a line. Once its value is larger than 1000, it increases for about 3 sample points after which it goes back to being just noise. In fig.5.3 we note that there are random gaps which need to be treated separately.

For this experiment there appear to be scanned 10000 lines, but only 9770 have a cor-
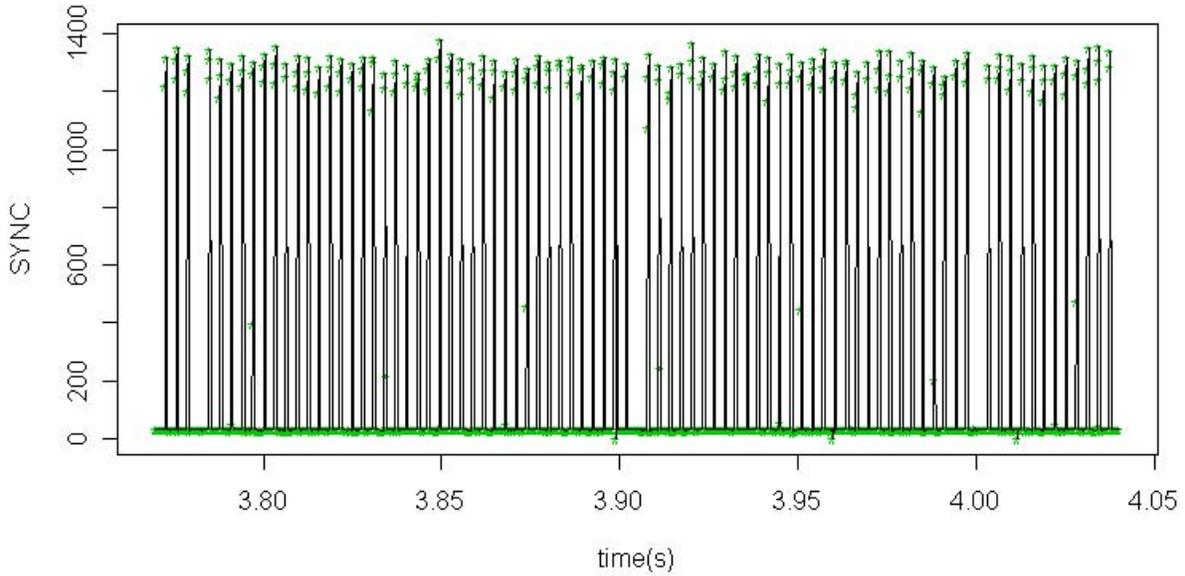
Figure 5.3: Drop outs of SYNC series.

respondent SYNC spike. Also, as the pixel time in this case was 2.56μs, to scan a line (of 512 pixels) it takes about 1.3ms. This time corresponds to about 6 sample points in these series. Similarly, we estimate the duration to scan an entire line (of 512 pixels) and fly back to be on average 2.46ms, with variance of 0.074ms. Converted in sample points this period is equivalent to 11.37 sample points, and 0.34 sample points for the variance. Considering that the scanning process happens in a fairly regular fashion, we are able to estimate the physiological value for a line. Therefore, in this case, we estimated 10000 POX values corresponding to the physiology during the scanning process.

5.2 ANOVA MODELS

In order to get a better view of the data for each observed line (fig.5.1) we selected an area of interest where the displacement is clearer. A median filter was applied to this portion to reduce some of the shot noise in the background (see fig.5.4). The trajectory of the feature captured in this region is highly correlated with the physiology ($r = 0.81$), supporting our

60

previous findings. In the absence of movement, the intensities of the correspondent pixels in each new recorded line should be very close with those already scanned. Accordingly to this simple rationale, we would like to model the distribution of the movement and estimate the "steady" locations.
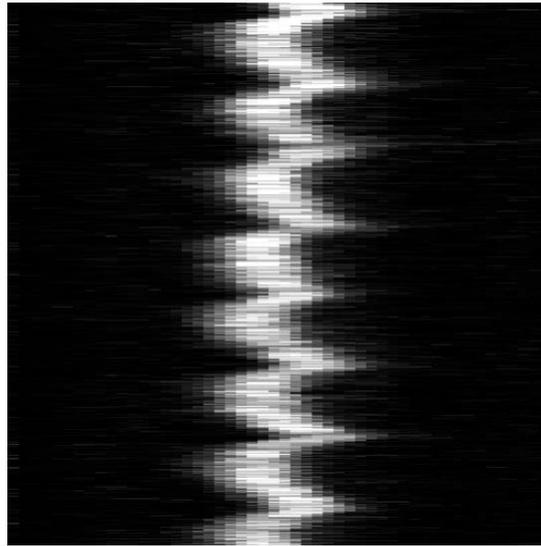
Figure 5.4: A fragment of the line-experiment: fifty pixels around the first feature captured were applied a 3 length median filter on x dimension in order to clarify the image.

Let us denote with $\lambda(x)$ the "true" line from the (y, z) position in the brain, and with $L_t(x)$ the observed lines scanned during this experiment. Considering that a line is scanned almost instantaneously when compared against the physiology rate, we assume that the movement of the pixels within their line is negligible. Thus, for the simplicity of the notations, we drop the dependence on x.

A simple model for the scanned data would be:

$$L_t(i, j) = \lambda_j + \theta_i + \gamma_j + e_t(i, j) \tag{5.1}$$

where e_t is some error associated with the scanning process, i corresponds to the time factor $(1 \leq i \leq 500)$, and j to the position within the line $(1 \leq i \leq 512)$.

Assuming that the errors are white noise, we could fit this model to a $2Way$ ANOVA model relating the observations with time and position factors. The fitted values of this model are presented in fig.5.5.
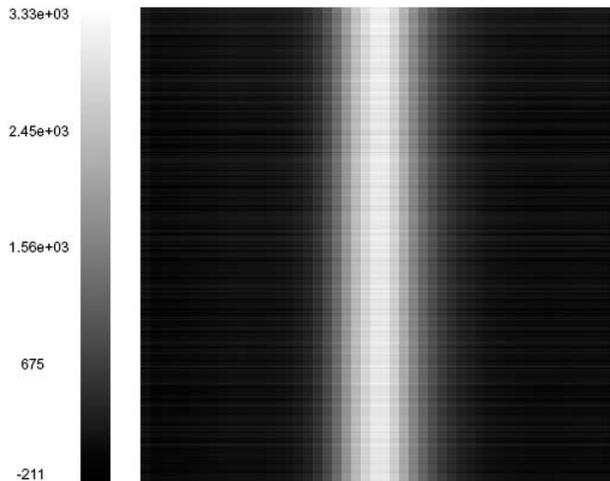
Figure 5.5: Fitted lines with 2Way ANOVA model- partial region around main feature.

As presented in fig.5.6, the residuals still vary in a regular fashion, pointing out that there are some other aspects of this process that have not been captured by this model.

We included the interaction between these two factors, but the model did not seem to improve; the residuals exhibited approximately the same variation in a regular fashion.

Also, we "fixed" a randomly selected line and shifted all the others to it using the Time shifting property as described in 4.7. Even though we hoped to account for some of the variation in this way, the model fit did not seem to get significantly improved. The residuals displayed high values and variation in a regular fashion (as in fig. 5.6).

5.3   EXTENDED EXPERIMENTAL SET-UP

Performing repeated scanned line experiments seem to have a high potential in capturing the motion of the brain caused by heart beat and respiration. This dynamic process happens in all the spatial directions, affecting clearly the images.

A series of experiments have been done on neonatal mice or ferrets scanning a line several hundreds of times for a slice, and then repeating the process for a different slice. This new experiment allows the acquisition of several hundreds of "replicates" at each slice in a stack of several tens of slices.
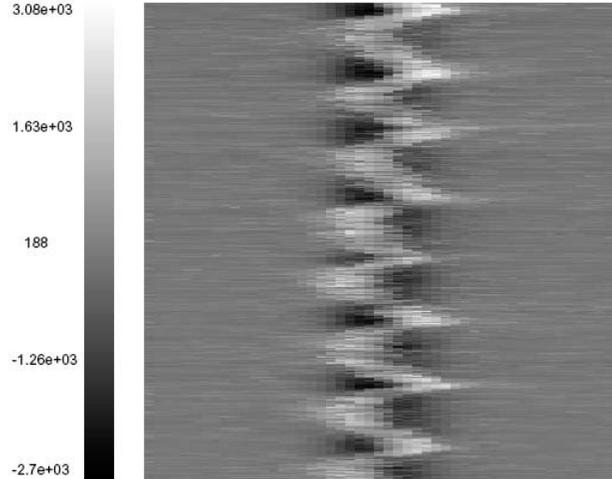
Figure 5.6: Residuals from fitting the line data with 2Way ANOVA model- partial region around the main feature.

Furthermore, orientation of the scanning has been changed using different angles while repeating the previous experiment. In other words, a line has been acquired for several hundreds of times at different z locations in the scanned volume of the cerebral tissue. Then the direction of scanning has been rotated with an angle θ selected upfront, and the scanning process has been repeated.

5.3.1 Example: Stack of repeated scanned lines (ferret)

Scanning a line repeatedly in the x direction gives us indications about the movement in x and y directions. In order to expand our tracking algorithms to z direction, we extend the experimental set-up such that it incorporates this dimension.

Hence, several experimental steps have been projected in advance.

1. Pre-specify the number of slices (z) that a volume of the cerebral tissue will be scanned in.

2. Scan a regular stack of 2D images for the volume of interest.

3. Fix the y coordinate within each z and specify the number of times the microscope should scan in x direction.

4. Repeat step 2.
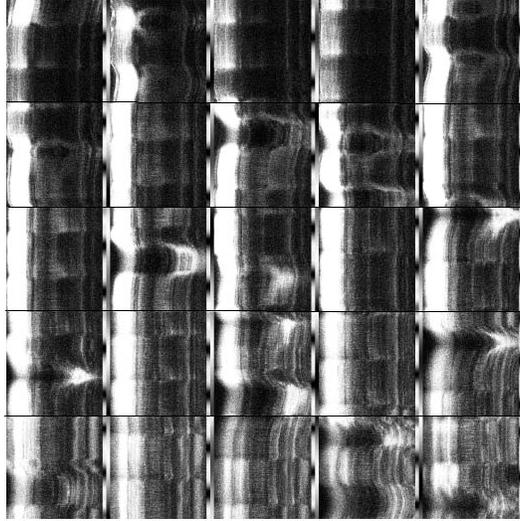
Figure 5.7: Repeated scanned line experiment: extended set-up on neonatal ferret. The process of scanning the middle line of a slice for 600 times has been applied to several slices. Note: on the right side of each slice we added indicators for POX and Respiration estimates concomitant with the lines.

Line data collected in this way is represented in fig.5.7 at it comes from the brain of a ferret. Although the stack has 161 slices, a set of 25 middle ones is shown. The actual scenes captured at these depths in the brain have been imaged before and after the line experiments. The set of 25 corresponding to those z's presented in fig.5.7 for the repeated scanned lines, is figured in 5.8. The POX values recorded at the same time are attached at the right side of each image. Briefly, we notice that the physiological rate for a ferret is lower than for a mouse (about 3 heartbeats per second). Looking at the images taken before and after the line data, the features moved significantly from one time to the other. See also D2 which is the RGB version constructed with the images taken 1st time repeated twice.

5.3.2 Example: Stack of repeated scanned lines (mouse)

Similar experiments, as described in the previous section 5.3.1 have been performed on mice, too. Fig.5.9 presents a group of lines collected from 5 adjacent slices.

Even though the between lines variation is apparently not so obvious, we notice the large artifacts due to respiration. When we selected lines from the stack scanned before and after the line experiment, we observed that the regions that the lines were highly likely to belong
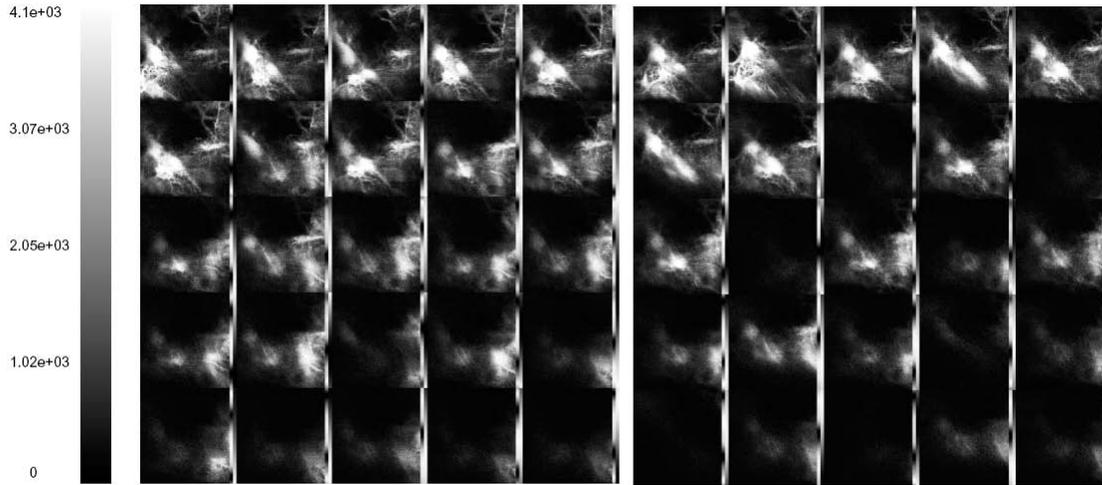
Figure 5.8: 25 consecutive slices from a stack (z=1,...,161) of images scanned in the brain of a ferret. The images on the left side present the slices taken before the line experiment, while those on the right side present the slices taken after. Visually we notice that the brain moved significantly in z direction as the features had not been scanned on the same positions at one time versus the other. Note: on the right side of each slice we added indicators for POX and Respiration estimates concomitant with the images.
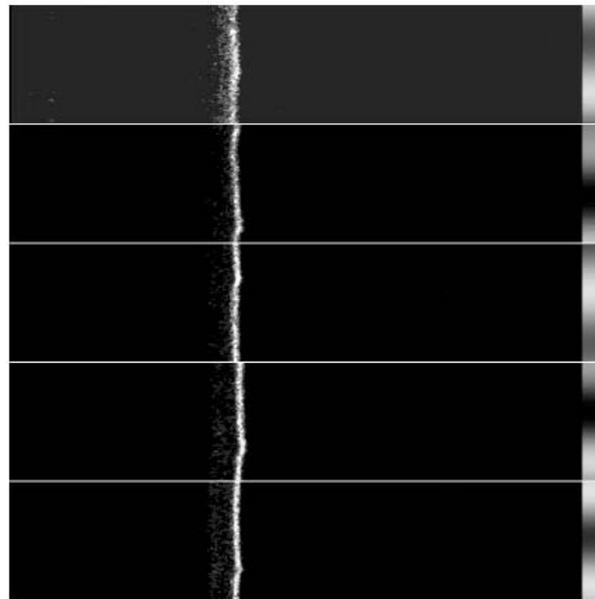
Figure 5.9: Repeated scanned line experiment: extended set-up on mice. The process of scanning the middle line of a slice for 600 times (300 shown) has been applied to 101 slices (only 5 shown). Note: on the right side of each slice we added indicators for POX estimates concomitant with the lines.

to, did not exhibit significant variation in y direction (see 6.20 below).

5.3.3   Example: Rotated scans

In order to explore the heterogeneity of the moving tissue under the microscope, the repeated scans of the line have been taken at different angular direction within the (x, y) plane of a pre-specified z slice.

In this case, the experimental steps that have been projected in advance are as following:

- Pre-specify the coordinates where the imaged slice (z) will be scanned.
- Depending on the number of directions, determine the angle with respect to the standard reference system that will give the rotated directions on which the scanner will collect pixels sequentially.
- Repeat the previous two steps at different z locations.

An excerpt of this kind of data collected at one z position is presented in fig.5.10. The actual acquisition directions were rotated with an angle $\theta = 22.5°$, but for simplicity we included only four sets at 45° angle separation.

5.4   CONCLUSIONS

To our knowledge, this kind of experiments has not been performed so far by any other lab. We believe that the information that is captured through these experiments is very important for learning about the movement process. However, the results depend to a large extent on the shape of the feature that has been included in the volume of interest. Therefore, some of these experiments may observe lines which do not vary for the entire period during the scanning through a slice. Under regular experimental conditions (i.e. the rodent's physiology is happening at the usual rates), we believe that this apparent lack of variation is caused by the particular shape of target feature. In such cases, the region where the repeated scan of a middle line has been done is similar over a relatively large number of lines.
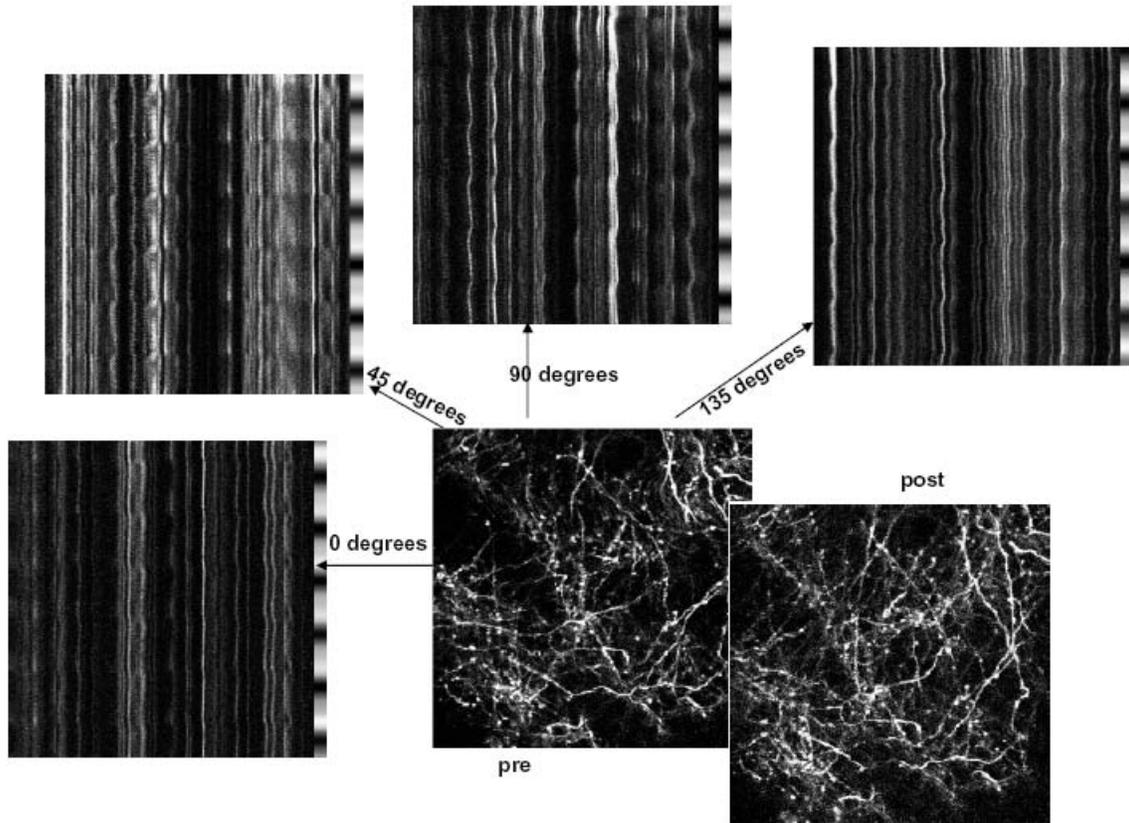
Figure 5.10: Repeated scanned line experiment: extended set-up. Fixing the y at the middle region of a slice, the scanner collected lines (in x direction) for several hundred times at different angles. With respect to the images taken before (pre) and after (post) the lines have been collected for 1500 times, the ones taken at $0°$ correspond to the horizontal direction; the ones taken at $90°$ correspond to the vertical direction.

When we noticed variation in the observed "replicates", a simple model as ANOVA has not been able to describe completely the factors that determine this variation. The following chapter presents other methods that help us track the movement and relate L with the POX values.

6.0   TARGET TRACKING AND CLASSIFICATION PROBLEM

The search for truth is more precious than its possession. (Albert Einstein)

From a practical point of view, we want to track the brain movement, identifying the true position of the features of interest during the scanning process. According to Liu ([23]), target tracking is an example of dynamic modeling that uses a linear Gaussian state-space model to describe the movement of the target object.

He describes a 2D tracking model as being characterized by:

- State Equation:

$$
\begin{pmatrix} v_{t,1} \\ v_{t,2} \end{pmatrix} = \begin{pmatrix} v_{t-1,1} \\ v_{t-1,2} \end{pmatrix} + \begin{pmatrix} \epsilon_{t,1} \\ \epsilon_{t,2} \end{pmatrix},
$$

$$
\begin{pmatrix} s_{t,1} \\ s_{t,2} \end{pmatrix} = \begin{pmatrix} s_{t-1,1} \\ s_{t-1,2} \end{pmatrix} + \begin{pmatrix} v_{t-1,1} \\ v_{t-1,2} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \epsilon_{t,1} \\ \epsilon_{t,2} \end{pmatrix},
$$

where $s_t = (s_{t,1},\ s_{t,2})^T$ is the position of the object at time t and $v_t = (v_{t,1},\ v_{t,2})^T$ is its current velocity vector. The error $\epsilon_t = (\epsilon_{t,1},\ \epsilon_{t,2})^T$ is distributed as $N(\mathbf{0}, \sigma_v^2 \mathbf{I})$.

This model describes the speed (vector) of the object as Gaussian random walk, where the position of the object follows the change of its speed. We assume, however, that a noisy version of the object's true position $(y_{t,1},\ y_{t,2})^T$ is observable.

- Observation equation:

$$
\begin{pmatrix} y_{t,1} \\ y_{t,2} \end{pmatrix} = \begin{pmatrix} s_{t,1} \\ s_{t,2} \end{pmatrix} + \begin{pmatrix} e_{t,1} \\ e_{t,2} \end{pmatrix},
$$

where the observation noise $e_t = (e_{t,1},\ e_{t,2})^T$ follows $N(\mathbf{0}, \sigma_y^2 \mathbf{I})$.

By writing $x_t = (s_{t,1}, \ s_{t,2}, \ v_{t,1}, \ v_{t,2})^T$ and $y_t = (y_{t,1}, \ y_{t,2})^T$ we can rewrite the foregoing system more briefly as

$$
\begin{aligned}
x_t &= Gx_{t-1} + \epsilon_t, \ \epsilon_t \text{ from } N(\mathbf{0}, \sigma_v^2 A); \\
y_t &= Hx_{t-1} + e_t, \ e_t \text{ from } N(\mathbf{0}, \sigma_y^2 \mathbf{I}),
\end{aligned}
$$

where

$$
G = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ A = \begin{pmatrix} \frac{1}{4} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & 1 \end{pmatrix}, \ H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}^T
$$

Mathematically, the tracking task is accomplished by an $on-line$ estimation of the object's position, $(s_{t,1}, \ s_{t,2})$, based on all information available before time t. If the y_t are always observable at time t, this estimation task can be achieved rather efficiently via the Kalman filter, because of the linear Gaussian structures being employed.

However, if we consider a cluttered environment, then at time t we observe a set of points, $z_t = z_{t,1}, \ldots, z_{t,k_t}$ in a 2D detection region of area Δ. In this region, the number of false signals follow a spatial Poisson process with rate λ. The set z_t includes the true measurement $y_t = (y_{t,1}, \ y_{t,2})^T$ with probability p_d. Other z's are treated as uniform within the detection range. This model for tracking in clutter is no longer a linear Gaussian system. According to ([23]), there has not been a universally effective algorithm for dealing with non-linear and non-Gaussian systems. Depending on the features of individual problems, some methods have been extended to apply the Kalman filter to such cases. Most of these methods are based on local linear approximations of the non-linear system.

6.1   MIXTURE KALMAN FILTER

Avitzour ([1]) modeled the tracking problem as a state-space model with the state variable $x_t = (x_{t,1}, x_{t,2})$, where $x_{t,1}$ is the location of the target on a straight line and $x_{t,2}$ is the target

velocity. Then, as described previously, x_t evolve in the following way:

$$x_{t,1} = x_{t-1,1} + x_{t-1,2} + \frac{1}{2}w_t,$$

$$x_{t,2} = x_{t-1,2} + w_t,$$

where the noise terms w_t are i.i.d., $N(0, w_0^2)$.

If for every t we could uniquely identify the object, our observation would have been the object's location $x_{t,1}$ plus a small random noise. In other words, we would observe

$$z_t = x_{t,1} + u_t,$$

where u_t are i.i.d. Gaussian noise with distribution $N(0, u_o^2)$.

When the detection is not precise, in a window with area Δ we observe m_t objects whose positions will construct the observation vector y_t. Among these m_t recorded locations, at most one corresponds to the true target we are interested in tracking (i.e. is equal to z_t). The occurrence of the confusing objects is assumed to follow a Poisson process with rate λ. We further assume that there is only a probability $p_d < 1$ for the observation window to actually include the target's location (z_t) in y_t. Therefore, the distribution of m_t is $Bernoulli(p_d) + Poisson(\lambda\Delta)$, and the false signals are uniformly distributed in the detection region.

By introducing an indicator variable I_t,

$$I_t = \begin{cases} 0 & \text{if the target object is not in the detection aria} \\ k & \text{if the } k^{th} \text{ object corresponds to the target} \end{cases}$$

we can formulate this problem as a state-space model.

- When $I_t = 0$,

$$\begin{aligned} p(y_t \,|\, x_t, I_t = 0) &= \Delta^{-m_t} \frac{(\lambda\Delta)^{m_t}}{m_t!} e^{-\lambda\Delta} \\ &= \frac{\lambda^{m_t}}{m_t!} e^{-\lambda\Delta} \end{aligned}$$

71

- When $I_t = k$,

$$p(y_t \mid x_t, I_t = k) \quad = \quad \frac{\lambda^{m_t-1}}{(m_t-1)!} e^{-\lambda\Delta} \frac{1}{u_0\sqrt{2\pi}} exp\left[-\frac{(y_{t,k}-x_t)^2}{2u_0^2}\right]$$

Since *a priori* $P(I_t = 0) = 1 - p_d$ and $P(I_t = k) = \frac{p_d}{m_t}$, we have:

$$f_t(y_t, I_t \mid x_t) \propto \begin{cases} (1-p_d)\lambda, & \text{if } I_t = 0; \\ p_d(2\pi u_0^2)^{-1/2} exp-\frac{(y_{t,k}-x_t)^2}{2u_0^2}, & \text{otherwise.} \end{cases}$$

As I_t is not observable, we need to sum out the I_t to obtain the observation distribution $f_t(y_t \mid x_t)$.

Using $q(x_t \mid x_{t-1})$ to denote the state evolution relationship, we can obtain a sequence of auxiliary distributions, $\pi(x_1, \ldots, x_t)$ iteratively:

$$\pi_t(x_t, x_{t-1}) \propto f_t(y_t \mid x_t) q(x_t \mid x_{t-1}) \pi_{t-1}(x_{t-1})$$

and, because of the Markovian structure among the x_t,

$$\pi_t(x_t) \propto \int f_t(y_t \mid x_t) q(x_t \mid x_{t-1}) \pi_{t-1}(x_{t-1}) dx_{t-1}.$$

With this formulation, the current position of the target can be estimated as $E_{\pi_t}(x_t)$.

Next, let us consider an important feature of this tracking model. If we know the values of the target trajectory Λ_t, the tracking system becomes linear and Gaussian. Therefore, the computation of the Bayes estimator, $E(x_t \mid y_1, \ldots, y_t)$ can be achieved exactly by a standard Kalman filter.

Let $\Lambda_t = (I_1, \ldots, I_t)$ be the trajectory up to time t. Conditional on Λ_t we can integrate out x_t exactly. This feature enables the user to design a Sequential Importance Sampling (SIS) system only on the reduced space of Λ_t ([3]). Furthermore, this model becomes a mixture of Kalman filters.
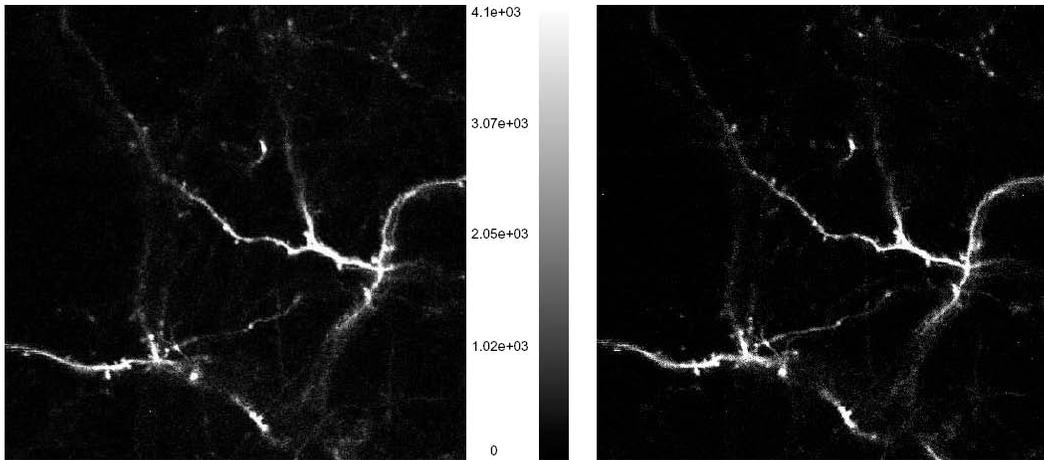
Figure 6.1: The images taken before (t=0) and after ($t = T + 1$) the line experiment. The scenes captured by these images look similar, despite the relatively long time lapse.

6.2 LINE DETECTION

It is natural that models based on predicting the present as a regression on the past, such as those provided by the state-space models, would be attractive to statisticians who are trained to view nature in terms of linear models. In fact, the difference equations used for representing such models are simply the discrete equivalents of linear differential equations that may, in some instances, provide the ideal physical model for a certain phenomenon. An alternate version of the way nature behaves is based on a decomposition of an empirical series into its regular components.([29])

As the Kalman filters make several distributional assumptions for the observations, we intend to take a simpler approach. We want to determine the regularity of the positions where the lines have been scanned during the experiments described in the previous chapter.

An overview of the data used for analysis in section 5.2 summarizes the following: we have repeated scanned lines, collected at times $t = 1, \ldots, T, T = 10000$, along with an image ($I^0$) recorded before (when $t = 0$) and another one (I^{T+1}) assessed after (when t=$T+1$) (see 6.1). Even though taken at about 37 seconds apart, the scenes captured by these images look similar.

However, there are small shifts as can be seen in fig 6.2.

Let us denote with $L_1^0, \ldots, L_{512}^0$ the lines that belong to I^0 and with $L_1^{T+1}, \ldots, L_{512}^{T+1}$ the
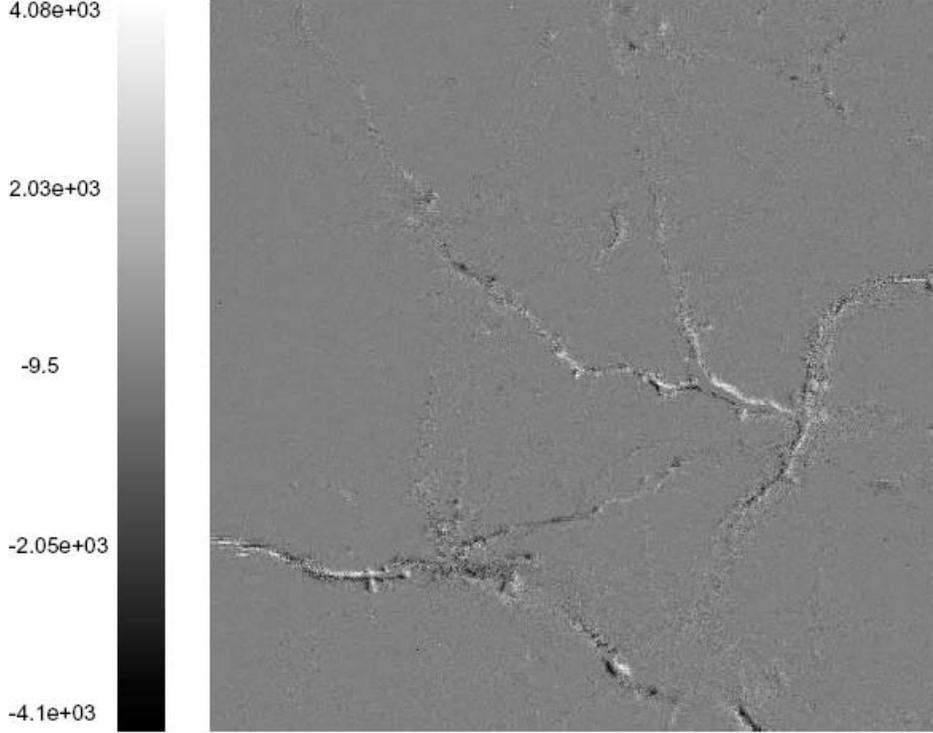
Figure 6.2: The difference image between the pre-image (t=0) and post-image ($t = T + 1$). The variation of the pixel intensities exposes the relative small shifts between the two images.

lines that belong to I^{T+1}.

Although the positions of the lines in I^0 and I^{T+1} may not be the true ones from the brain, we would like to learn how the lines scanned at times $1 \leq t \leq T$ moved with respect to these ones. In other words, conditioning on the fact that line L_j^0 has been scanned at $t = 0$ on $y = j$, we want to estimate the time $(1 \leq t \leq T)$ when the scanner captured the "same" line.

The experiment was designed so that fixing z and y in the region of interest, the microscope acquires continuously 512 pixels over the x direction then returns for several hundreds of times. For fig 2.5, y was fixed at the middle of the image, therefore at $y_0 = 256$. Empirical observation of the movement process exposed the fact that a line may move up to 10 pixels. Consequently, we selected *apriori* a band $\delta = 21$ of lines from I_0 around the middle one, and we used them to define several classes of interest: $C_1, \ldots, C_\delta$. Finally we will identify which classes have actual components that have been observed during $t = 1, \ldots, T$.

We developed an algorithm that searches among the lines $L^1, \ldots, L^T$ and classifies them

74

within the considered classes. As a matching measure we started with the Euclidean norm:

$$L^t \text{ belongs to class } C_{\hat{d}} \text{ iff } \hat{d} = argmin_{1 \leq d \leq \delta} \left\| L^t - L^0_{y_0 - d} \right\|_2 \tag{6.3}$$

However, this measure does not consider the possible drifts in the brain which may cause some lines (at some $1 \leq t \leq T$) to be shifted with respect to $t = 0$ or $t = T + 1$. To compensate for this event, we used the lagged correlation between two lines (we will call it furthermore, cross-correlation function).

The cross-correlation function between two vectors $\underline{x}$ and $\underline{y}$ for lag d is defined as:

$$r_{x,y}(d) = \frac{\sum_i (x(i) - \overline{x})(y(i - d) - \overline{y})}{\sqrt{\sum_i (x(i) - \overline{x})^2} \sqrt{\sum_i (y(i - d) - \overline{y})^2}} \tag{6.4}$$

We also prefer this function because of its properties:

- $-1 \leq r_{x,y}(d) \leq 1$;

- $r_{x,y}(d) = r_{a_x x + b_x, a_y y + b_y}(d)$, thus invariant to an affine transformation.

For long vectors (for us $1 \leq i \leq 512$) calculating this function is quite expensive in computer time and memory. Taking advantage of the useful applications of the Fourier transform, introduced briefly in the previous section 4.1.1, we can reduce this complexity. The algorithm that we developed is based on the generalization of the following Autocorrelation theorem.

Before formulating this result, we will start with a brief discussion of the relevant background.

The definition of the convolution between two functions is:

$$h(x) = f * g \equiv \int_{-\infty}^{\infty} f(t)g(x - t)dt \tag{6.5}$$

which actually multiplies f by the time reversed and shifted function g. It allows one function to smear or broaden another (see fig.6.3). Therefore it can be useful when we are interested to see how two functions are related to each other.
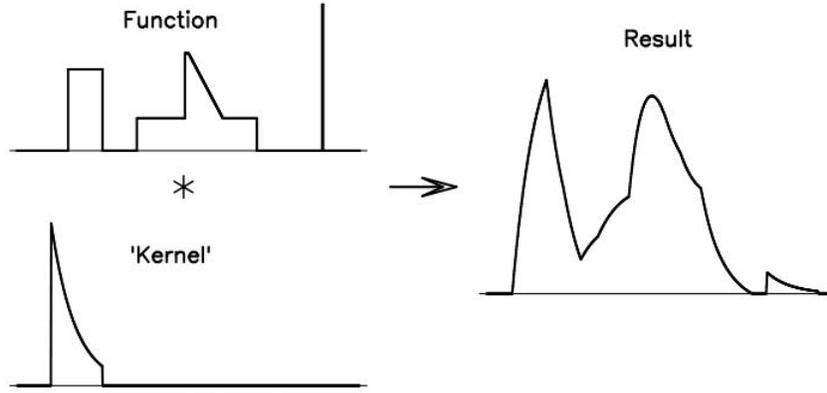
Figure 6.3: Convolution between a delta-like function and a kernel function.

The Convolution Theorem turns a convolution into the inverse Fourier transform of the product of two Fourier transforms:

$$f(t) * g(t) = \mathbf{F}^{-1}(F(w)G(w)) \tag{6.6}$$

where $\mathbf{F}$ is the Fourier transform, and F(w)= $\mathbf{F}$(f(t)) and G(w)=$\mathbf{F}$(g(t)). A proof of this result is presented in Appendix C.1.

In this context we can state now the Autocorrelation theorem, also known as Wiener-Khinchin theorem:

$$f(t) * f(t) = \mathbf{F}^{-1}(F(w)F^*(w)) = |F|^2 \tag{6.7}$$

with $F^*(w)$ being the complex conjugate of $F(w)$. Appendix C.2 sketches the proof for this result.

The generalization of 6.7 is the Cross-correlation theorem which states that the Fourier transform of the cross-correlation of two functions is equal to the product of the individual Fourier transforms, where one of them has been complex conjugated. As we are interested the relationship between two lines (seen as functions of time), we compute the correlation between L_i^a and L_t with

$$\underline{\rho}(L_i^a, L^t) = \mathbf{F}^{-1}(F(L_i^a)F(L^t)^*) \tag{6.8}$$

76

Using this similarity measure to quantify the coherence between two lines, we can classify

$$L^t \text{ belongs to class } C_{\hat{d}} \text{ iff } \hat{d} = argmax_{1 \leq d \leq \delta}[max_{lag}\underline{\rho}(L^t, L^0_{y0-d})] \tag{6.9}$$

6.2.1 Class Estimation

Based on the classification rules presented in the previous section, we ran several sets of analyses.

First we considered that t goes from 1 to 512. Therefore, we computed a 512x21 similarity matrix (S_k^{Euclid}) among the lines $L^1, \ldots, L^{512}$ and the 21 lines from the images taken at time $k = 0$ and $k = T + 1$, $L^0_{245}, \ldots, L^0_{265}$, as:

$$S_{t=k}^{Euclid}(i, j) = \left\| L^i - L^k_{244+j} \right\|_2, \tag{6.10}$$

where $k = 0$ or $T + 1$, $1 \leq i \leq 512$, and $1 \leq j \leq 21$.

When performing these measurements, the value of the pixel intensity for each line varies from 0 to 4095. Also, as we are interested in finding the lines that match best, i.e. the feature location is preserved despite the variation in pixel intensity. Therefore, we filtered the before and after images as well as the lines taken at these first 512 times with the filter described in section 3.2.2. After trying several K values (to determine how many thresholds we need to estimate), the one that seems to optimize all these images was selected to be 5. We applied the same steps for the filtered data and we calculated the similarity matrices in this case.

The images created with the values of these matrices are presented in fig 6.4.

We note that out of 21 possible classes, displayed as columns in the similarity matrices, only a few have actually lines captured at these times. Using 6.3 from previous section, we identify the classes as shown in fig.6.5.

We note that the classes identified with the filtered data do not differ too much from the ones which use unfiltered data. After a few random checks, we are confident that the filtered data gives us a better precision.

Next, we applied the cross-correlation algorithm on the lines collected for $1 \leq t \leq 2560$.
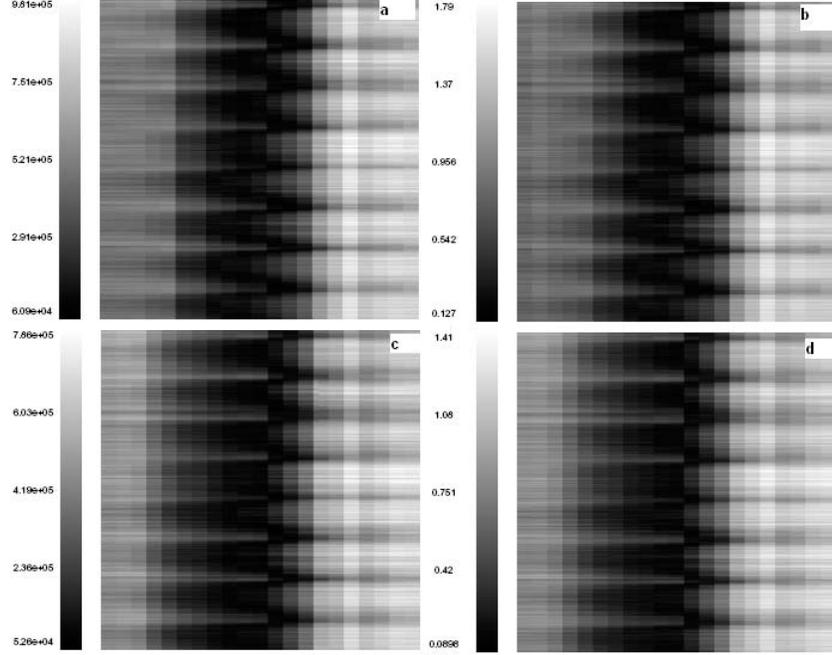
Figure 6.4: The similarity matrices computed with Euclidean distance: (a)lines observed for the first 512 times versus 21 lines from I^0;(b) filtered version from (a); (c)lines observed for the first 512 times versus 21 lines from I^{T+1};(d) filtered version from (c).

The definition 6.4 that we generalized in 6.8 is sensitive to variation in pixel intensities. For data coming from the repeated line experiment, the within line variance ranges from 327065 to 631555. Calculated on filtered data (as described above), the within-line variance ranges from 0.5 to 1.1.

The following plot 6.6 shows that after filtering the lines, the amplitude of the variances decreased significantly while the between-line variation was preserved.

Consequently, we used the subset of 2650 filtered lines as input for our classification algorithm 6.9. We calculated the similarity matrices as:

$$S_k^{cross-cor}(i,.) = \mathbf{F}^{-1}(F(L_{iMod\delta}^k)F(L^{[i/\delta]})^*), \tag{6.11}$$

where k is 0 or $T+1$, $1 \leq i \leq$ 2560x21, $[i/21]$ is the integer part of the division, and $iMod21$ is the remainder of this operation. Each line in these matrices is 512 positions long corresponding to the possible lags.
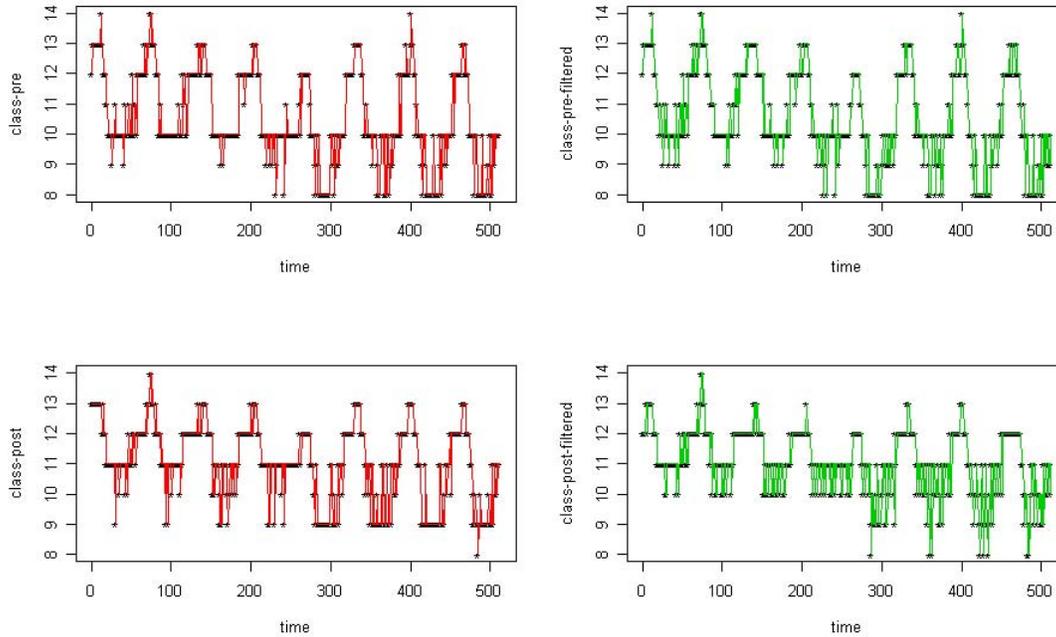
Figure 6.5: The classes estimated using the Euclidean norm. Note: the plots on the left represent the classes estimated with initial data; the plots on the right represent the classes estimated with filtered data.

Fig.6.7 represents an excerpt of the similarity matrix when k=0 and $1 \leq i \leq 500 \mathrm{x} 21$ $(1 \leq t \leq 500)$.

For class detection we are interested in maximizing the cross-correlation series over the lag variable. Then we apply the classification rule in order to choose the class where the scanned line fits best. A more detailed look into the selection process is included in Fig.6.8.

The estimated classes are presented in fig6.9:

6.2.2 Movement Process

The estimated classes give us indications about the amplitude of the brain movement captured by the scanner during these times. If there was no movement, the lines scanned at times $1 \leq t \leq 2560$ should fit in the class C_{11}^k corresponding to the line in the middle of the image I^k, with k=0 or $T+1$. However, the results presented in the previous subsection show that the brain has moved during the experiment.

79
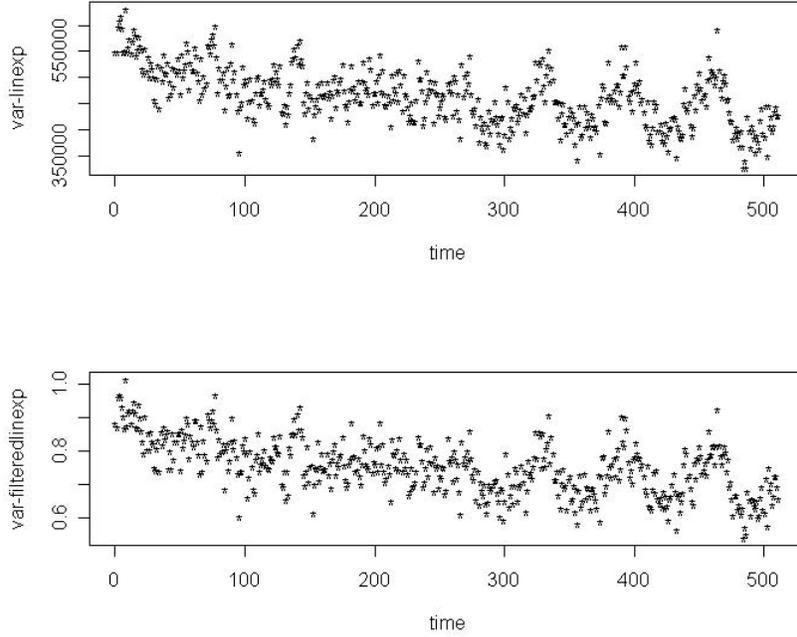
Figure 6.6: The within line variances for observed data (up) and for the filtered data (bottom).

We define the shift in the y axis as:

$$\nu^k(t) = E[C^k(t)|L^t \text{ moves}] - E[C^k(t)|L^t \text{ does not move}] \tag{6.12}$$

where k is 0 or T+1 and $1 \le t \le T$ and $C^k(t)$ is the class where the line L^t fits in best.

We estimate these shifts using the classes estimated with the cross-correlation algorithm. A graphical representation for the smoothed version of the amount of movement in y direction is shown in fig.6.10

Using the definition of the cross-correlation series between two series u and v, the lag where the maximum is attained corresponds to the shift in x direction necessary for $L^1, \ldots, L^T$ to match the selected classes. If the brain was not moving, then the cross-correlation series would be maximized at lag=0.

Using 6.8, we define the shift in the x axis as:

$$\eta^k(t) = E[argmax_{lag}(\underline{\rho}(L^k_{\hat{d}}, L^t))|L^t \text{ moves}] - E[argmax_{lag}(\underline{\rho}(L^k_{\hat{d}}, L^t))|L^t \text{ does not move}] \tag{6.13}$$

80
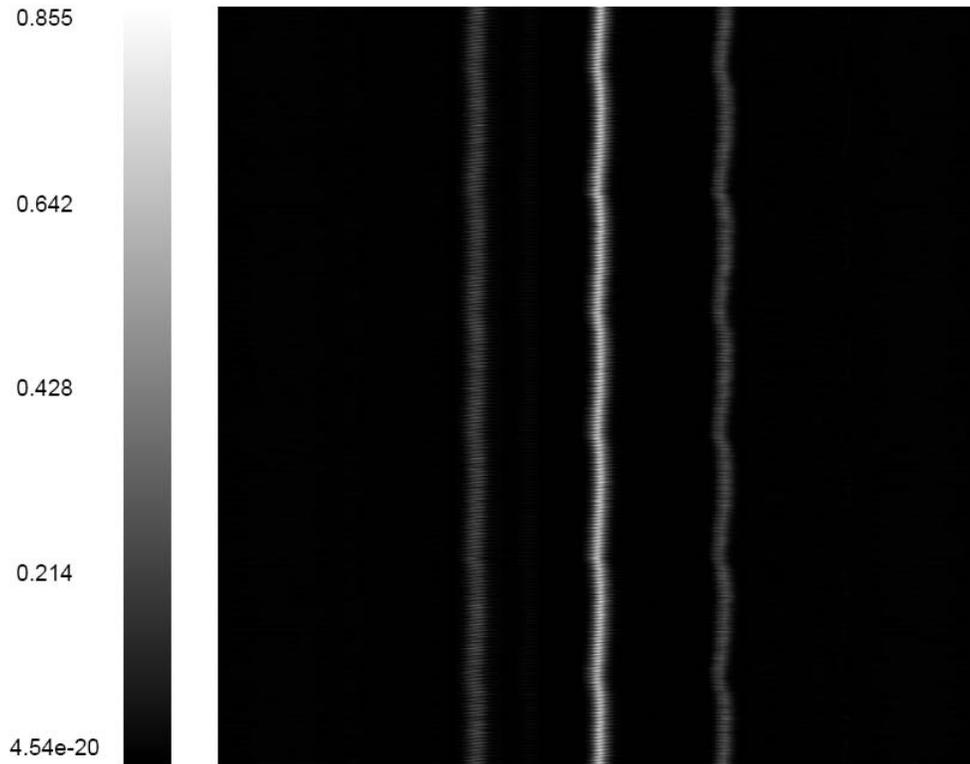
Figure 6.7: The cross-correlation series for the group of 21 lines selected from I^0 with lines acquired at $1 \leq t \leq 500$ (see 5.1)
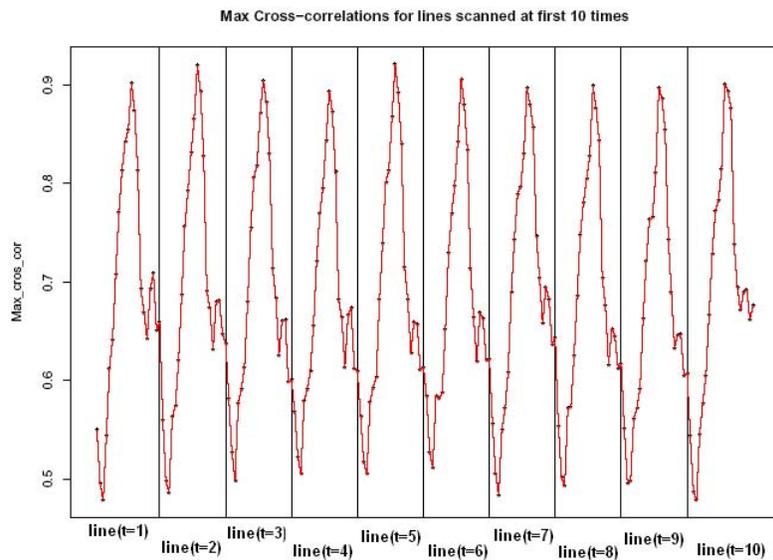
Figure 6.8: The maxima of the cross-correlation series corresponding to the groups of 21 lines selected at $t = 0$ with 10 lines collected at $t = 1 \ldots 10$ during the line experiment (see also fig 5.1)
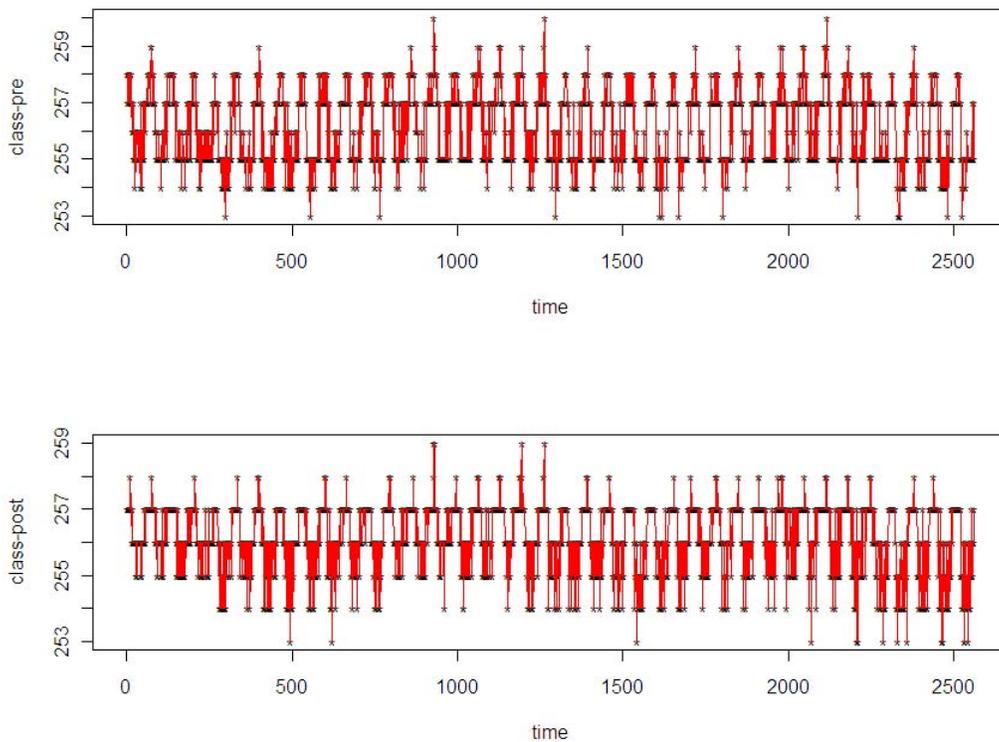
Figure 6.9: The classes estimated using the cross-correlation measure. Note: the upper plot is using the classes defined from I^0, while the lower plot is using the classes defined from I^{T+1}
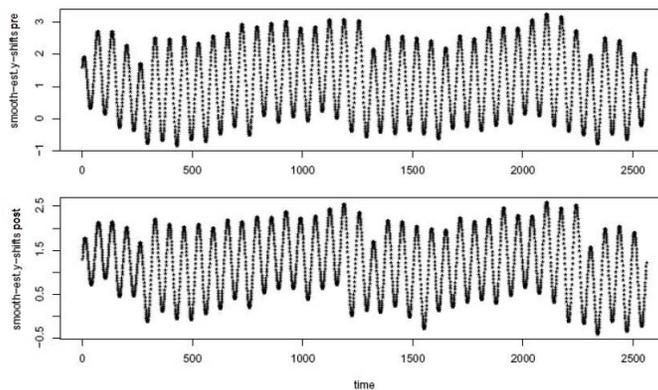
Figure 6.10: Estimated shift in y direction -smoothed version. Note: the upper pannel correspond to t=0, while the bottom panel correspond to $t = T + 1$.
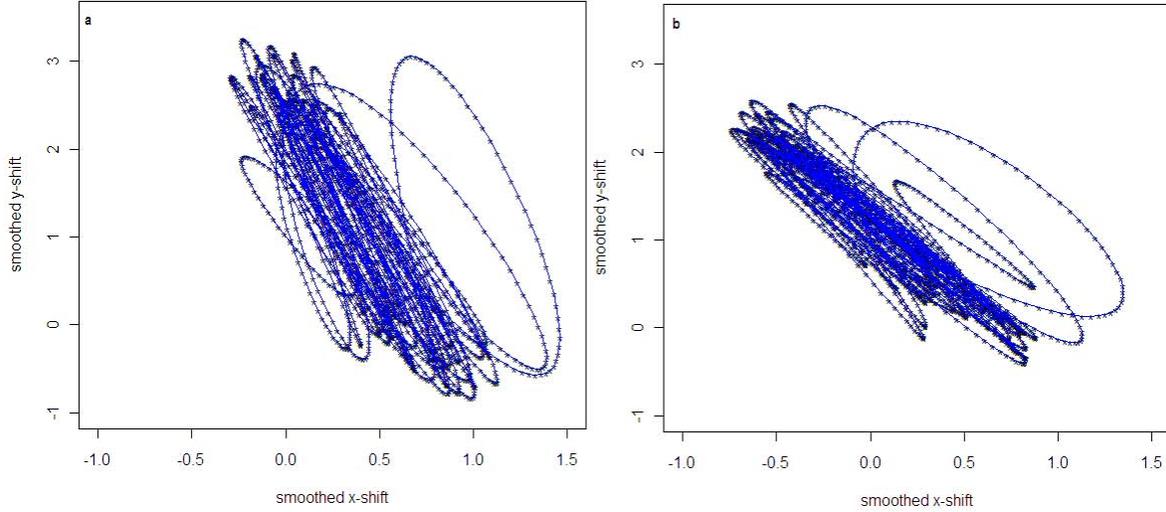
Figure 6.11: The x-y movement for the lines scanned first 2560 times in the repeated-scanned line experiments, with respect to I^0(a) and with respect to I^{T+1}(b)

where k is 0 or T+1 and $1 \leq t \leq T$ and corresponds to the class where the line L^t fits best.

As the second term in 6.13 is 0, the shift in x direction will be estimated directly from the similarity matrices 6.11. We can describe the movement in (x,y) plane by estimating the shifts in x and y directions. Fig. 6.11 represents the estimated trajectory that the brain is following during the scanning of $1 \leq t \leq 2560$. We notice that the movement is fairly periodic with larger amplitude in the y direction. Also, the few estimates that seem to have an "unusual" behavior are very likely to be related to the respiration.

A more detailed examination of the relationship between the shift of line at time t and the estimated POX(t) value recorded simultaneously, shows a statistically significant association. As the observed values for the POX series vary roughly from -700 to 700, the change in the shift value for each additional POX unit would be very small. Therefore, we rescaled this measure with the one-to-one transform:

$$n(p(t)) = \frac{p(t) - min_{s \leq 2560}p(s)}{2}, \tag{6.14}$$

which preserves its spectral characteristics. Calling the image set of POX through this isomorphism as normalized POX, we used it in several regression models to estimate its

Reference Image	Outcome	normalized POX Effect	Std. Error	p-value
I^0 Image	y-Shift	2.7795	0.0915	< 0.0001
I^0 Image	x-Shift	-0.5558	0.0472	< 0.0001
I^{T+1} Image	y-Shift	1.737	0.0711	< 0.0001
I^{T+1} Image	x-Shift	-0.8261	0.05	< 0.0001

Table 6.1: Estimates of the effect of the normalized POX values on the shifts in x and y directions.

impact on the shifts on y and x direction.

Formulating the models as:

$$out^l(t) = \beta_0^l + \beta_1^l n(p(t)) + \epsilon^l(t), \tag{6.15}$$

where $1 \leq t \leq 2560$, $out^l(t)$ is the shift estimate at time t, and ϵ^l is white noise (from $N(0, \sigma^l)$).

Table 6.1 presents the impacts of normalized POX signal on the y and x shifts $(\hat{\beta}_1^l)$, estimated through the linear model set-up 6.15. These values tell us how many pixels the line moved on average while scanning it for 2560 times.

Let us define a POX cycle as the period that the device needs to go from one minimal value to the next one. During the acquisition of an image with 512 lines, about 5 heartbeats happen corresponding to around 4-5 POX cycles. As we are interested in identifying moments when the brain is in a minimal motion state, or "near-to-rest" , monitoring the movement trajectory per POX cycle may give us useful indicators about the transition to a "far-from-rest"' state. Fig.6.12 represents the brain movement for the first 9 POX cycles, conditioned on the pre-experimental image (I^0). Fig.6.13 represents the brain trajectory for the same 9 POX cycles, as before, but conditioned on the post-experimental image (I^{T+1}).
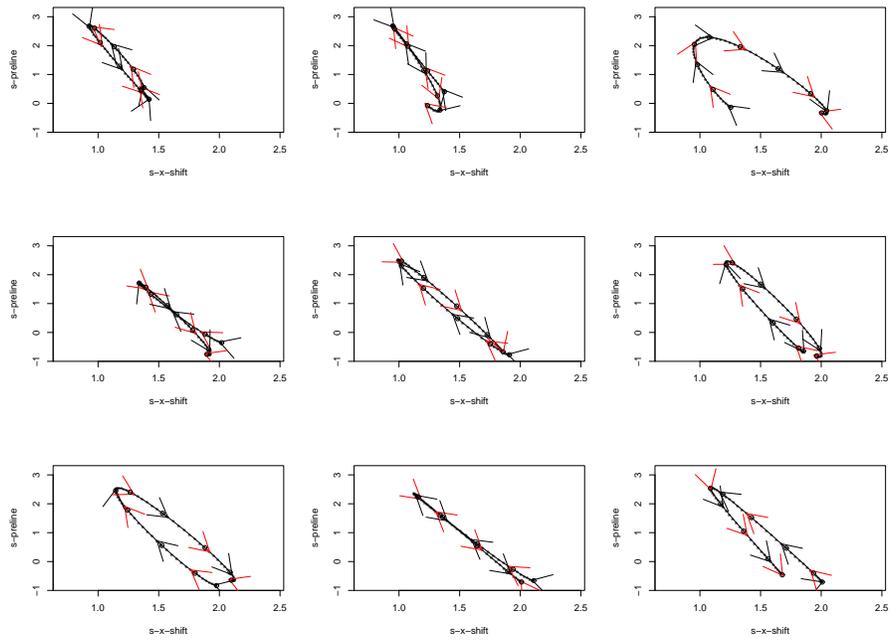
Figure 6.12: The x-y movement in the repeatedly-scanned line data, for a group of 9 heart beats (when compared with respect to I^0).
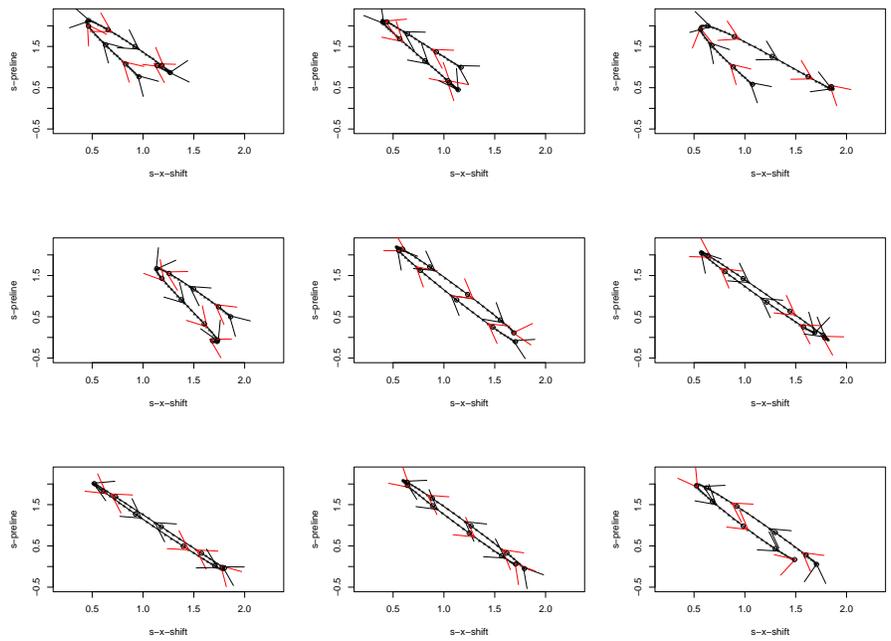
Figure 6.13: The x-y movement in the repeated-scanned line data for a group of 9 heart beats (when compared with respect to I^{T+1}).
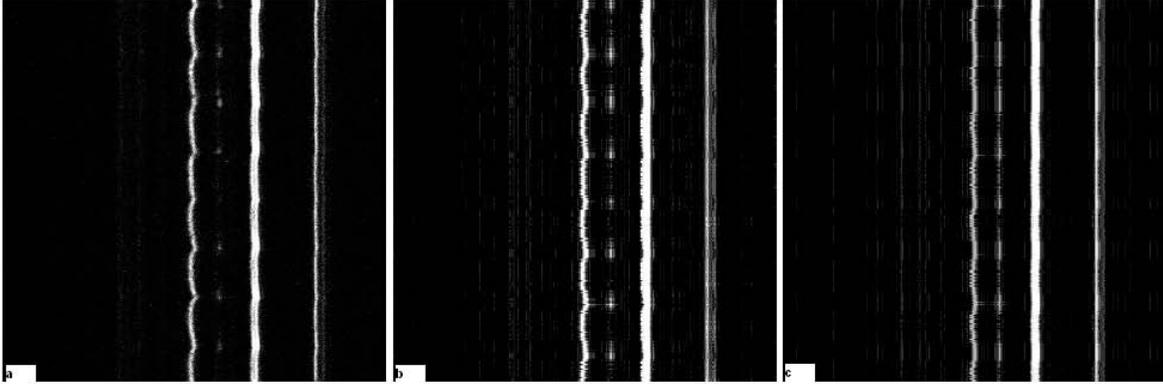
Figure 6.14: The fitted lines for 5.1(a) estimated with the classification model applied on I^0 (b), and I^{T+1} (c) (first 500 times).

6.3 MODEL FIT

The classification models defined in the previous section 6.2 describe the brain motion in (x,y) plane. Fitting these models to the observed lines at $1 \leq t \leq 2560$ helps us measure the goodness of fit. The fitted values estimated with the second classification model are presented in fig 6.14 (first 500 lines) or D3/D4(all).

Visually, the model displays the almost periodic variation for the positions of the features between the lines as the observed ones.

When representing the residuals, fig.6.15 (first 500 times) or D5 (all), it seems that there is still unexplained variation in the data. This fact is expected as we did not include the z component in the spatial modeling of the line movement. A closer look over the values used for classifying with the cross-correlation measure, shows that these values vary also in a relatively broader range (see fig.6.16 and 6.17). Further analyses might use a Support Vector Machine classifier to determine which lines might come from a different slice than I^0 or I^{T+1}. However, we would not be able to know how further away might be the neighboring slice. To have an idea about that we need to collect repeated lines from several adjacent slices, i.e. fixed y, but variable z.

The means squared error for these models was 0.2419 and 0.296 respectively. To check whether there is any event that might have triggered large residuals at some specific times, we divided the 2560 times into 5 groups of 512 lines. We computed the MSE for each subgroup,
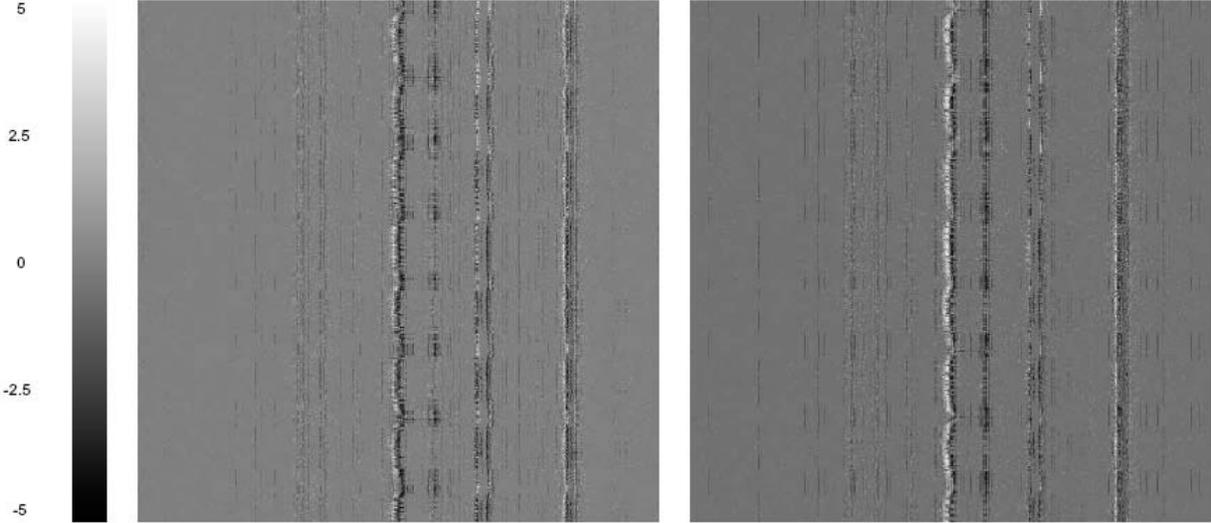
Figure 6.15: The residuals estimated with the classification model applied on I^0 (left), and I^{T+1} (right) (for the first 500 times).

yielding values between 0.239 and 0.249, 0.274 and 0.317 respectively. The variance of these estimates is 0.000022 and 0.00028, respectively. These statistics show that one model offers a somehow better fit than the other, but the differences are very small.

Overall, we can conclude that for the period of time considered for this analysis the motion process happened in a relatively regular cyclic fashion.

A closer look on these scatter-plots (fig. 6.12 and 6.13) points out the fact that for smaller normalized POX values the movement is relatively small when compared with large normalized POX values. Also, regressing the y shifts estimates versus the POX values yield significant positive associations (see table 6.1). Therefore we can conclude that the POX value could be an indicator of the state in which the brain is at that time. We will continue to study this relationship in more detail in the next chapter and use it to reconstruct images.

6.4   APPLICATIONS ON EXTENDED LINE DATA

We mentioned in last chapter, 5.3 about several extensions we considered in the repeated line scan experiments. For the data coming from z-stacks acquired from this kind of experiments,
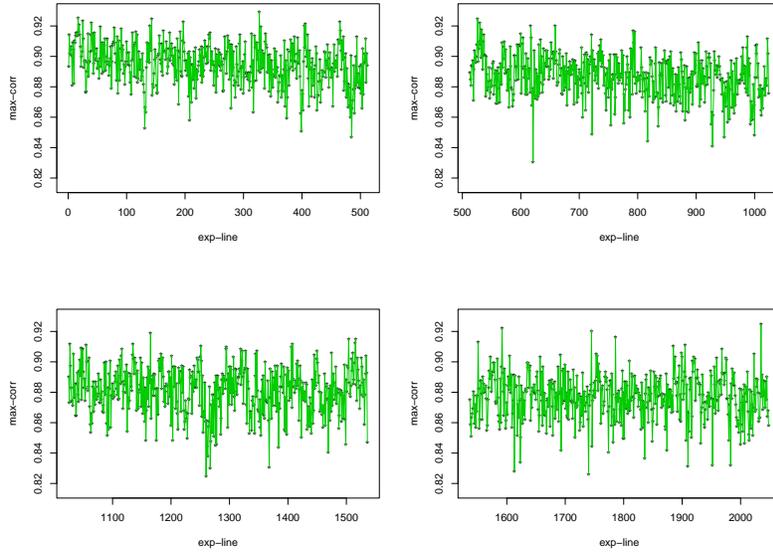
Figure 6.16: Maximum values from cross-correlating lines at times $1 \leq t \leq 2048$ with lines from I^0.
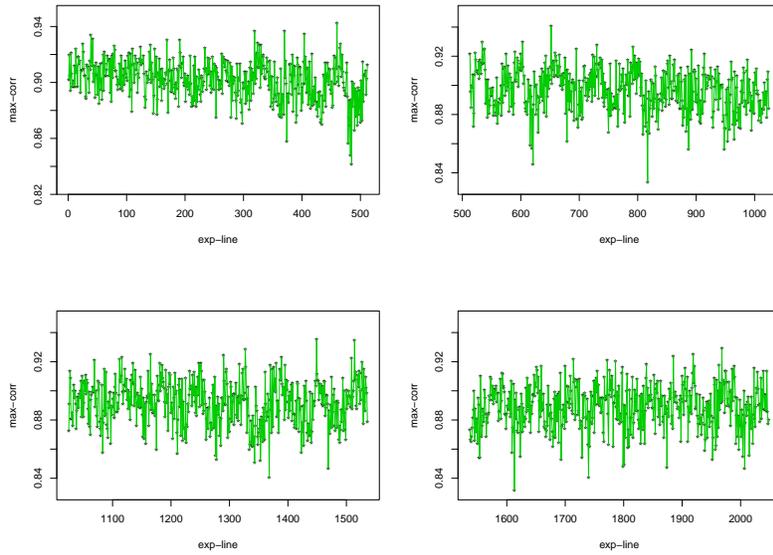
Figure 6.17: Maximum values from cross-correlating lines at times $1 \leq t \leq 2048$ with lines from I^{T+1}.

we tried to expand the analysis steps presented in this chapter to include the z dimension as well.

6.4.1 Stack of repeated scanned lines (ferret)

In order to visualize the between variation among the lines of one slice, we computed a distance matrix given by the Euclidean norm (see 6.18). We notice the periodic similarities among some of the lines represented with darker shades.

Next, using the Euclidean norm, we fixed a line (t=60) from D6 (z=104) and we did a search in the set of the lines scanned at the same z and at 10 adjacent slices (5 above and 5 bellow). The similarity matrix is presented in 6.19, where the vertical axis corresponds to the depth (z) and the horizontal one to time (t).

6.4.2 Stack of repeated scanned lines (mouse)

Similar steps as those described in previous sections have been taken with the z-stack line-data collected from a mouse. We selected a sample of lines which have clearer image/variation data. For each slice, let us denote the times when these lines where scanned as $t = 1, 2, ..., T$. Then the stack of pre- (x, y)- images was taken at $t = 0$.

The cross-correlation values used to determine how similar are the lines recorded at times are presented in fig6.21.

The pre-slices that matched the lines in the considered sample (scanned at $t = 1, 2, ..., T$) are represented in fig6.22. We notice that the lines scanned during the line- experiment were not varying too much from one slice to the other. Consequently, most of them matched the lines from 2nd and 6th slices of the pre z- stack.

6.5 CONCLUSIONS

Although with inherent experimental limitations, the analysis presented in this chapter gives us numerical statistics which estimate the amplitude of the movement. First, we conditioned

89
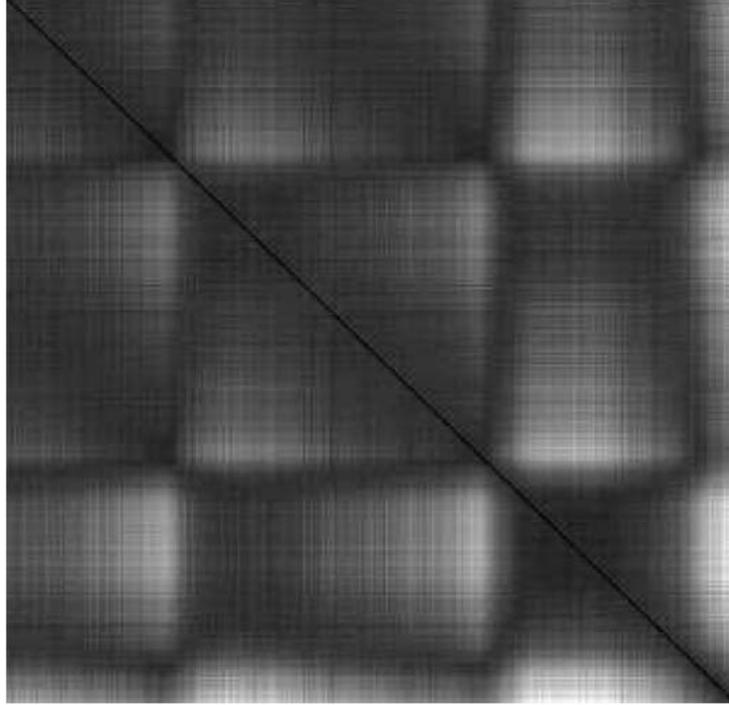
Figure 6.18: A distance matrix calculated for the lines of the slice presented in D6. t^h line in the matrix represents the Euclidean distances of the line scanned at time t with all the other lines(299) recorded from the same slice z. Note: darker shades corresponding to smaller values point out periodic similarities among lines at different times.
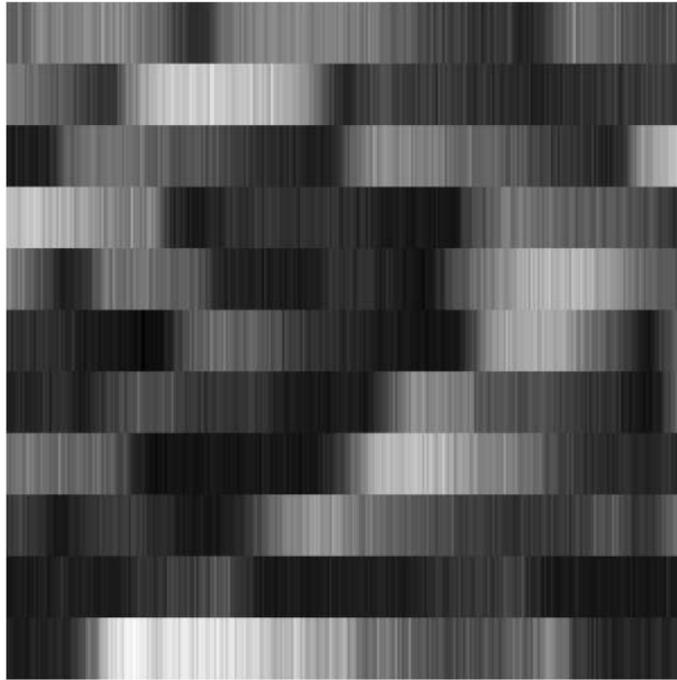
Figure 6.19: The similarity matrix corresponding to Euclidean norm of a "fixed" line (t=60) at z=104 with the set of the lines taken at the same z and at 10 adjacent slices ($89 \leq z \leq 109$). The vertical axis corresponds to the depth (z) and the horizontal one to time (t). Darker shades correspond to smaller distances.
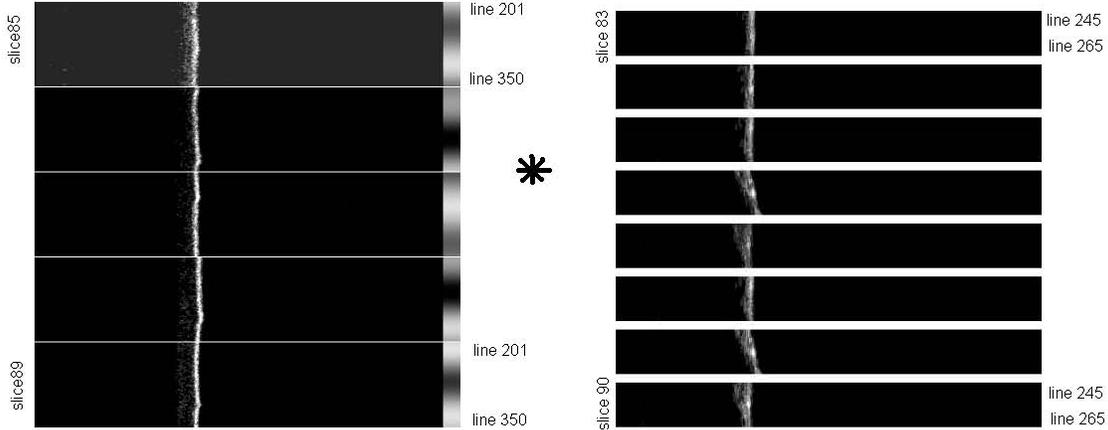
Figure 6.20: The segments of middle 21 lines of 8 slices adjacent of the z-stack taken before the line-data was acquired (right side) are cross-correlated with the groups of lines scanned for 300 times in 5 adjacent slices. Note: the depths of the slices taken before include the depths of the slices taken during the line-experiment.

our analyses on the region in the brain where the repeated line experiment was executed at two different times (before and after), relatively further apart from each other (a few tens of seconds). As the physiological data was not recorded for those images, we do not know the POX values for their middle lines. However, we expected that under normal physiological conditions, the lines scanned before, (at $t = 0$) should be rescanned at certain times during the repeated line scan experiment. We plan to apply these classification models to the rest of the lines, up to $T = 10000$. We expect to recover the results stated at this section.

Because of the brain movement, we expect that the features which were scanned in I^0 to appear moved in I^{T+1}. To have an idea about the amplitude of the motion, we cross-correlated the δ lines selected from the region of interest and defined in the previous analyses as classes (C_l^k). The shifts in x and y directions were on average small (1-3 pixels), but not zero. This might explain also the relatively different trajectories estimated through the two models.

As we are constrained by the lack of clear information about features at neighboring depths in the cerebral tissue, we need to be aware that some lines scanned during the repeated line experiment may come from different regions in the brain than those captured in I^0 or I^{T+1}. The z spatial component has not been integrated in this algorithm. When fitting the classification model to I^0 or I^{T+1}, the estimated lines look very close to the
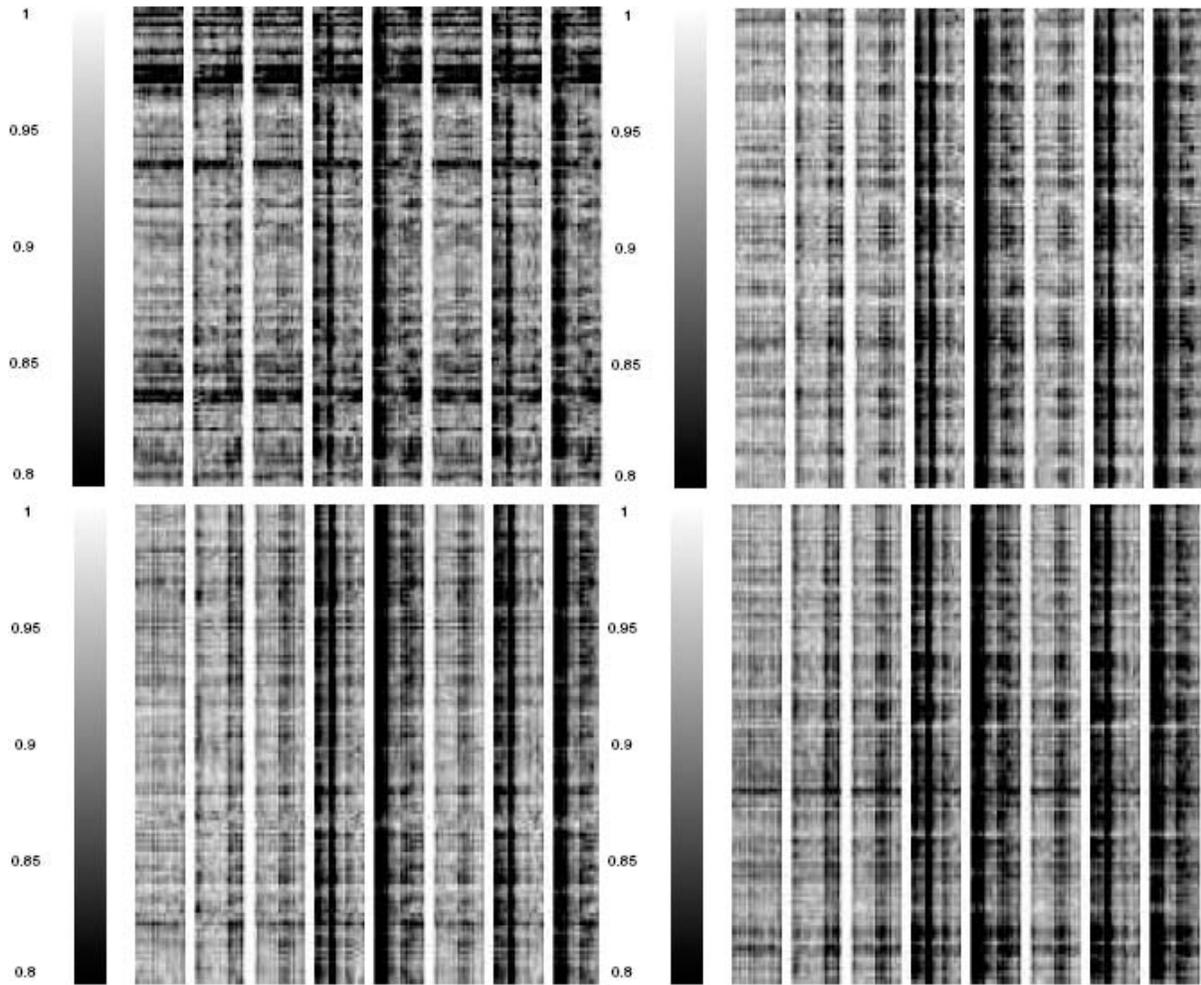
Figure 6.21: The cross-correlation values of the groups of lines from the extended line-experiment and from the pre-stack slices, as described in fig6.20.
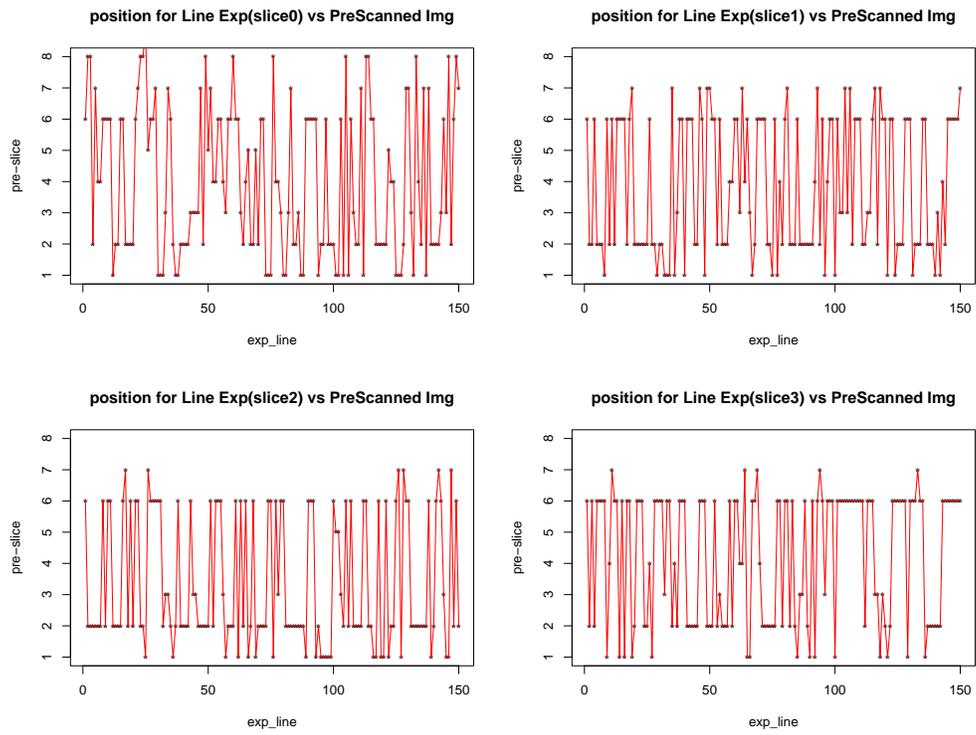
Figure 6.22: The estimated slices (on y axis) from the pre z stack whose middle lines maximized the cross-correlation values of the groups of lines (on x axis) described in fig 6.20.

Shift direction	correlation (t=0 vs t=T+1)	test statistic (t_{2558})	p-value
x-Shift	0.56	34.19	< 0.0001
y-Shift	0.75	56.83	< 0.0001
x-Shift (smoothed)	0.95	153.9	< 0.0001
y-Shift (smoothed)	0.97	201.8	< 0.0001

Table 6.2: Tests of the association between the estimated shifts with the two cross-correlation models (conditioned on I^0 versus I^{T+1}).

observed ones, but there is still unexplained variation. Also, we compared the trajectories identified when classifying with respect to I^0 with the ones estimated with respect to I^{T+1}. The test statistic used was:

$$t = r\sqrt{\frac{n-2}{1-r^2}}, \tag{6.16}$$

which is a random variable from a t distribution with n-2 degrees of freedom.

As summarized in table 6.2, cross-correlating the two series yielded that the y-shifts are highly correlated. Similar finding is true for x-shifts. The two cross-correlation series were maximized at the same $lag = 1$, suggesting that the description of the process is quite consistent from one model to the other.

7.0 IMAGE RECONSTRUCTION METHODS

Yesterday is but today's memory and tomorrow is today's dream. (Kahlil Gibran)

7.1 HIDDEN MARKOV MODELS

Other methods that we have considered for estimating the "good" image rely on combining the replicates of the image taken at the same z position using the physiological values. The Pulse-oximeter device displays the percent oxygenation of arterial hemoglobin. Thus, it gives an indication about the cardiac and respiratory cycle, and it can be used as a cue about potential tissue movement. As explained at the beginning of the Preliminary Analysis section (4), after some spectral analysis of the POX wave we can recover the heart beat and the respiration signals. The EKG is recorded simultaneously with POX data, so when the two series are overlaid (fig.7.1) we notice that they are very well related. The POX is minimal when the heart is relaxing, i.e. in the diastolic phase of its cycle.

Consequently, we could intuitively think that among lines scanned at the same (y,z) position for multiple times (not necessary consecutively), the one whose correspondent POX value is smaller, has a higher probability to be recorded at the true (y,z) coordinates in the brain tissue. In order to apply this method, the three copies of the same z slice where recorded in the triggered mode such that their physiologies start at equally spaced moments during the cardiac cycle.

One way to describe the underlying process of scanning lines in a systematic fashion from a dynamic environment is by using hidden Markov model framework. We can look at the brain movement as following a continuous path determined mainly by the cardiac

EKG and POX during 1 second of scanning
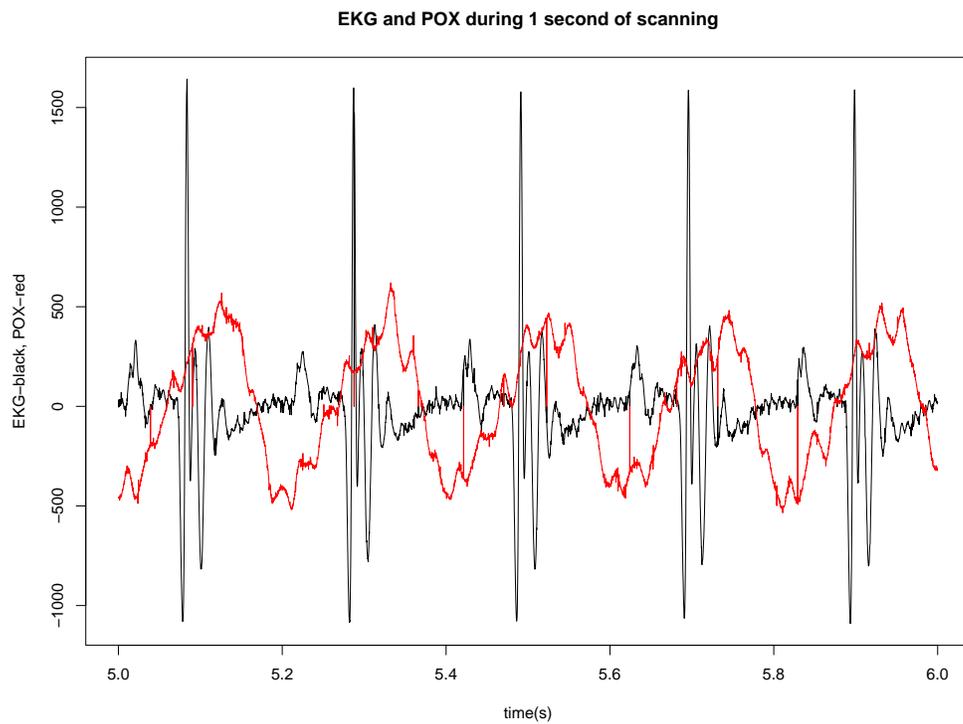
Figure 7.1: Physiology measured while scanning the brain. EKG is colored in black, while POX is colored in red. Note: the minimal POX values correspond to diastole phase in the cardiac cycle.
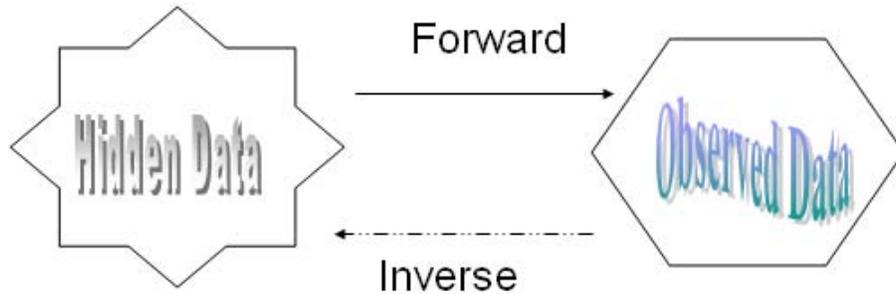
Figure 7.2: A scheme of the underlying process described by Hidden Markov Models: we observe data from a dynamic system whose states are hidden to direct knowledge. Estimate HMM to fit well the observations (forward) then infere about the latent states (inverse).

and respiratory cycles, and we would like to combine the steadiest states with the most dynamic ones into one scoring system. Decisions made based on this system would be probabilistic "labels" for the position of the line. Hidden Markov models (HMMs) are the formal foundation for making probabilistic models of linear sequence "labeling" problems. ([27], [7]). They provide a conceptual toolkit for building complex models by simply drawing an intuitive picture.

In our case the HMM invokes two states, one for each of the labels that we might assign to one scanned line: "near-to-rest" (S_1) and "far-from-rest" (S_2). Each state has its own emission probability determined by the physiological value. For each state, transition probabilities of stationing or moving to a new state describe the linear order in which we expect the states to occur.

It is useful to imagine an HMM generating a sequence of states. When we visit a state, we emit a residue from the state's emission probability distribution. Then, we choose which state to visit next based on the state's transition probability distribution. The model generates two strings of information. One is the underlying state path (the labels), as we transition from state to state. The other is the observed sequence, each observation being emitted from one state. The state path is a Markov chain, meaning that the state we go to next depends only on the state we are currently in. Since only the observations sequence was given, the underlying state path is hidden so it becomes a hidden Markov chain.

An HMM is a full probabilistic model: the model parameters and the overall sequence "scores" are all probabilities. Therefore, we can use the Bayesian probability theory to

manipulate these numbers in standard, powerful ways, including optimizing parameters and interpreting the significance scores. Models of dynamic behaviors are often best expressed in terms of a sequence of events or phenomena that occur over time. This approach defines the set of valid states and describes the system dynamics in terms of stochastic transitions among them.

We denote $\pi(i)$ as the probability of being in the relevant state S_i, and $b_j(k)$ as the probability assumed in the model for the process which distorts S_j to produce O_k. The probability $P(\underline{O}, \underline{\Sigma} | HMM, \underline{\theta})$ for an HMM with parameters $\underline{\theta}$ to generate a state path $\underline{\Sigma}$ and an observed sequence $\underline{O}$ is the product of all emission probabilities $(b_j(k))$ and transition probabilities $(a_{i,j})$ that are used. The goal of our method is to estimate the hidden sequence of states using estimates of $\underline{\theta}$ ($b_j(k)$'s and $a_{i,j}$'s).

Bayes theorem states that:

$$P(x\,|y) \propto P(y\,|x)P(x)$$

where the constant of proportionality does not depend on x.

Thus, the right-hand side of 7.1 is the key to defining such quantities as the maximum a posteriori (MAP) estimate of x, given y.

7.1.1 Standard HMM limitations

The standard HMM framework restricts the distribution of the time spent in one state to be geometric. Thus, if t_d is the duration of the system in state i, $1 \leq i \leq M$, its distribution is:

$$p_i(t_d = d) = a_{ii}^{d-1}(1 - a_{ii}),$$

where a_{ii} is the "cycling"' transition probability of state i and d is the number of time-steps.

In reality, based on prior knowledge we may want a different kind of distribution. Normal physiology for *in vivo* experiments reveals that the transition from the "near-to-rest"' state to the "far-from-rest"' one is more likely to occur about half way through the cardiac cycle.

The problem of modifying the standard Markov model to allow for arbitrary state-durations can be address by using semi-Markov models ([13]). A semi-Markov model has the following generative description:

- when entering state i, the process remains in this state t_d time-steps;
- the distribution of t_d is given by $p_i(t_d)$;
- at time $t_d + 1$ the process enters another state according to a transition matrix A;
- the process repeats.

The state duration distributions can be modeled using parametric distributions or non-parametrically by mixtures, kernel densities, etc.

Also in the standard HMM framework, the observed y_t 's are often modeled as Gaussians or mixtures of Gaussians. For Gaussian, this implies a piecewise constant process with one mean μ_i per state i with additive Gaussian noise. Mixtures allow switching between multiple means per state, but still imply a constant regression process as a function of time. However, there are many examples of real-world time-series where this piecewise - constant model is inappropriate.

Rabiner [27] shows that the concept of hidden states, where the observation is a probabilistic function of the state, can be used effectively. An observable Markov model outputs a set of states at each instant time, where each state corresponds to a physical (observable) event. In our case, we are given a sequence of observations from which we want to infer the hidden state path. Potentially, there could be multiple state paths that generate the same sequence. Our problem becomes finding the one with the highest probability. For most problems there are so many possible state sequences that it may be very computational costly to consider all of them. The efficient Viterbi algorithm is guaranteed to find the most probably state path given a sequence and an HMM, with less computational effort.

In our case, we start with a HMM frame work for 2D images. In order to increase the signal-to-noise ratio, the image acquired at a pre-specified z depth inside the brain is collected for three consecutive times. Therefore, at each y position the system produces noisy versions of the "true" line located at those (y,z) coordinates.

We hypothesize that these lines are observed when the brain is oscillating in one of the two hidden states: 'near-to-rest' (S_1) and "far-from-rest" (S_2). Also, based on the results of the analyses presented in the previous chapter, we estimate the output probabilities for each state based on the physiological POX measure. Therefore, at every time t when a line is recorded, POX(t) indicates the probability of being scanned in S_1 or S_2.
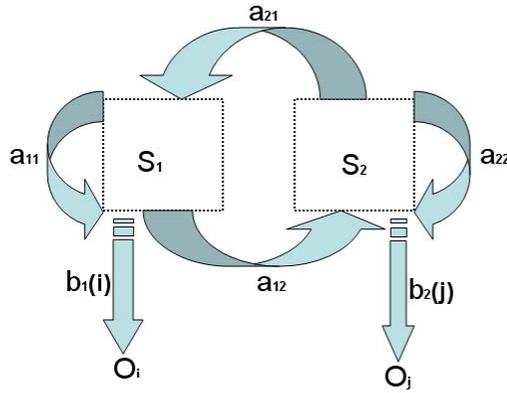
Figure 7.3: Hidden Markov Model idea: Two hidden states S_1 and S_2 that produce observations O_i and O_j with probability $b_1(i)$ and $b_2(j)$, respectively. The transition from one state to the other happens with probability a_{ij}, while the probability of remaining in any state is a_{ii}.

Formally, there are 3x512 lines scanned during several seconds when the brain was moving for several times. Coming from the same cerebral tissue and from neighboring spatial locations, these lines are not independent in this observed sample. However, they are collected by the microscope under the triggered mode to ensure that the variation due to physiological motion remains consistent across each replicate. This allows us to assume that the lines scanned at the same y position in the three images are sufficiently separated in time that they might be considered relatively independent.

7.1.2   Probabilities Estimation

A schematic representation of the scanning process described above is shown in .

For the semi- HMM probabilistic set-up we propose, we need to specify the following:

- S = the set of states = S_1, S_2
- O = the observations = $O_1, O_2, \ldots, O_M$
- $\pi(i)$ = probability of being in state S_i at time 0 (i.e., in initial states)
- A = transition probabilities = a_{ij}, where a_{ij} = P(entering state S_j at time $t+1$ | being in state S_i at time t)

- B = output probabilities = $b_j(k)$, where $b_j(k)$ = P(producing O_k at time t | being in state S_j at time t)

Based on the results stated in previous section 6.2.2, we can use the POX(t) value as an indicator of the state where the line is scanned in at time t. We would like to include this information in further analyses done within the HMM framework. First we integrate it into an algorithm for estimating the necessary probabilistic set-up. Next, as presented in the following section, we use it for predicting the number of replicates necessary for this model.

Let us define the state in which the line might have been scanned in as:

$$
L_t \text{ is in } \begin{cases} S_1 & \text{if } L_t \in C_{10} \text{ or } C_{11} \text{ or } C_{12} \\ S_2 & \text{otherwise} \end{cases} \tag{7.1}
$$

Then we relate the times when the line was scanned in S_1 with the physiological POX value. Generally, during an experiment, the POX ranges on average from -600 to 600. This range is preserved after estimating the POX value corresponding to the line scanned at time t. As described in section 5.1, when estimating the POX value for each line, the higher frequency noise from the observed POX measurements have been filtered out at some degree by this moving average algorithm. In fig.7.4 it can be noticed how the POX values changed in the matching process and after filtering more the higher frequency noise with the Fourier transform.

The POX measure is a continuous scalar, but displayed in a discrete way to match the line image data. We have categorized it even further dividing the overall range in N_p equal bins. Next, based on the classes estimated in the previous chapter (6.2.1) and the definition 7.1, we estimated the output probabilities b_i, which, conditioned on each state should be a probability mass function. Fig. 7.5 shows the estimated probabilities $b_i(k)$ of scanning a line O_k in state S_i. A better estimation is obtained when we use the smoothed POX version (see 7.6). In this case we used $N_p = 52$ (or $N_p = 40$ for the smoothed version). However, increasing or decreasing N_p will make b_i look smoother or more sparse, but it will not change the overall shape.
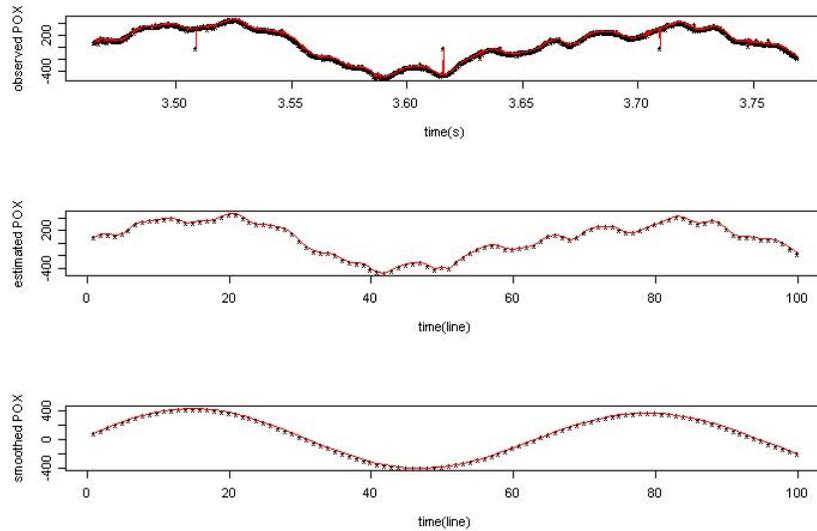
Figure 7.4: POX data during the repeated scanned line experiment: The first panel represents the observed POX for about 300ms. This time corresponds to scanning 100 lines, whose estimated POX values are represented in the second panel. The last one shows the smoothed version of the estimated values.
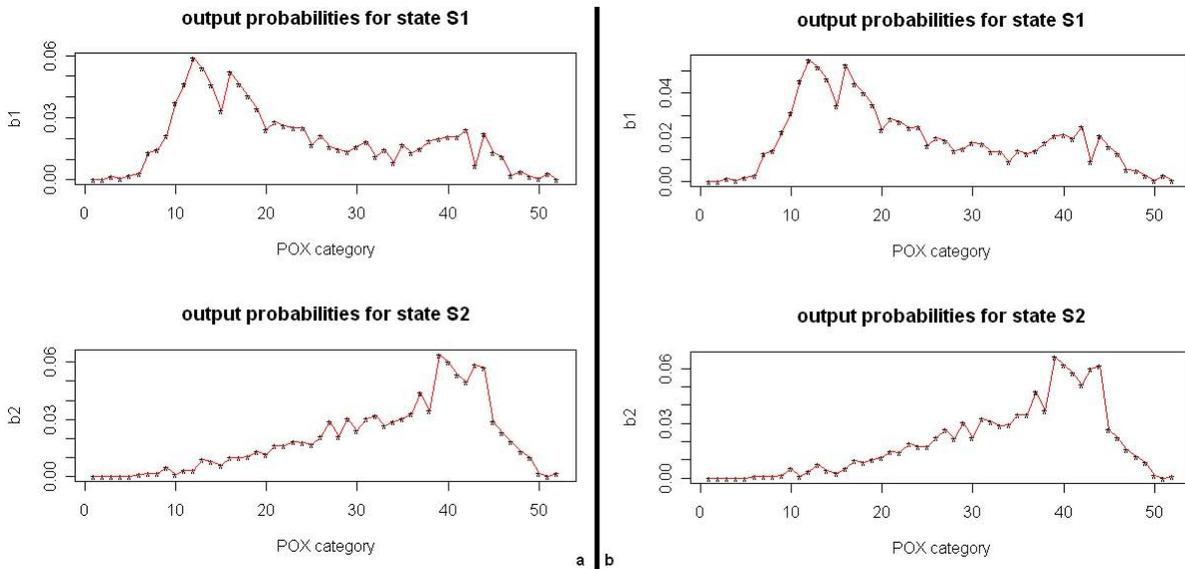
Figure 7.5: Estimated output probability mass function using the POX values corresponding to lines collected in the repeated scanned line experiment for $1 \leq t \leq 2560$. Not: for the left panels (a) we used the classes estimated with respect to I^0 (pre), for the right ones (b) we used the classes estimated with respect to I^{T+1} (post)
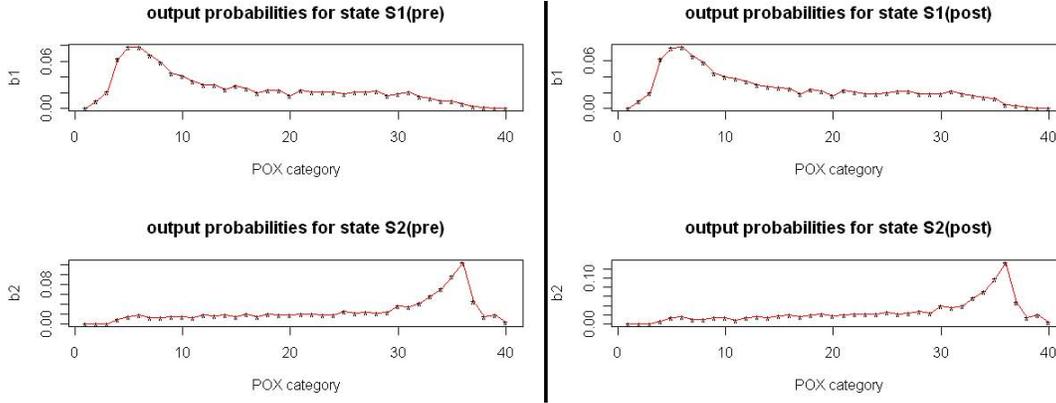
Figure 7.6: Estimated output probability mass function using the smoothed POX values corresponding to lines collected in the repeated scanned line experiment for $1 \leq t \leq 2560$.

Consequently, we were able to get the estimates of the posterior probabilities d_i, where:

$$d_i(k) = P(S_i|O_1, O_2, ..., O_k) \tag{7.2}$$

,

with the constraint:

$$\sum_{i \leq 2} d_i(k) = 1.$$

For a graphical representation, they are shown in fig. 7.7.

As explained in more details in next section, the rate of lines scanned in S_1 within a POX cycle (as defined in 6.2.2) is about 50%. Depending on the heart rate, about half of the cardiac cycle is spent in state S_1, and about half in state S_2. Consequently, we consider the transition matrix:

$$(a_{ij}) = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

In a general set-up, with more than 2 states we could define it as:

$$\frac{\text{number of times going from state l to state k}}{\text{number of times going from state l to any other state}}$$

103
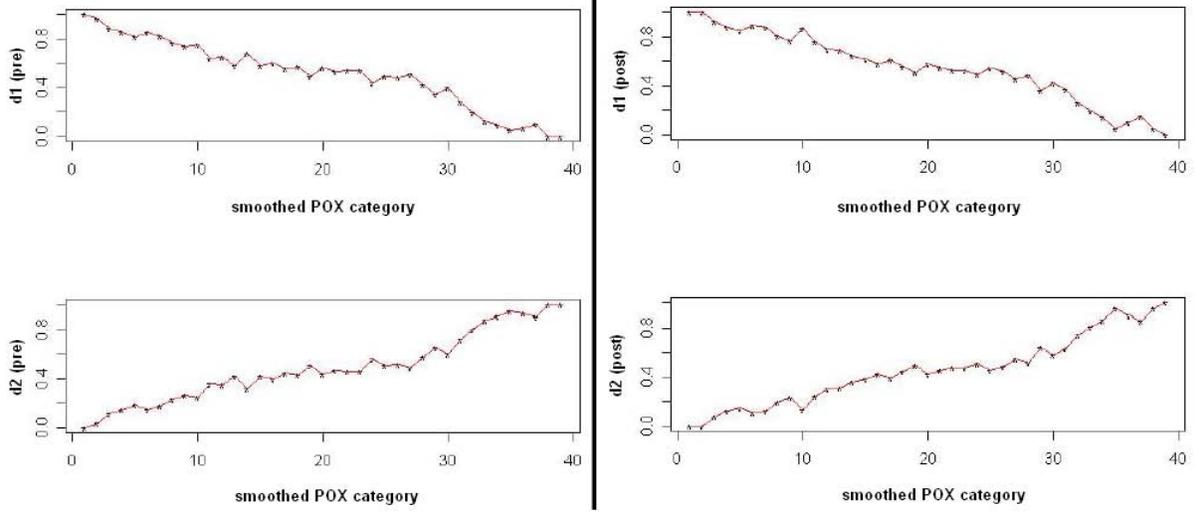
Figure 7.7: Estimated posterior probabilities using the smoothed POX values corresponding to lines collected in the repeated scanned line experiment for $1 \leq t \leq 2560$.

7.1.3 Viterbi Algorithm

In order to predict the state sequence $\underline{S} = S_1, ..., S_T$ based on the observed data we consider the following likelihood function:

$$\delta_t(i) = Pr[S_1 S_2 \ldots S_t = i, O_1 O_2 \ldots O_t \,|\underline{\theta}] \tag{7.3}$$

Defined in this way, $\delta_t(i)$ computes the probability along a single state path at time t which accounts for the first t observations and ends in state S_i.

Viterbi algorithm finds the optimal state sequence which maximizes $\delta_t(i)$, equivalent to maximizing:

$$Pr[S_1 S_2 \ldots S_t = I | O_1 O_2 \ldots O_t, \underline{\theta}]. \tag{7.4}$$

This optimization criterion is called the rule of Maximum A Posteriori (MAP). The amount of computation in the Viterbi algorithm is at the order of $M^2\text{T}$, while the memory required is at the order of MT.

Generally the algorithm involves an exhaustive search of all the M^T possible state sequences, evaluating the likelihood that the observed lines come from those states. The steps are:

1. Initialization:
$$
\begin{aligned}
\delta_1(i) &= \pi(i)b_i(1) \\
\phi_1(i) &= 0, \text{for } 1 \le i \le 2
\end{aligned}
$$

2. Induction:
$$
\begin{aligned}
\delta_t(i) &= (\underbrace{max[\delta_{t-1}(j)a_{ji}]}_{1 \le j \le 2})b_i(k) \\
\phi_t(i) &= \underbrace{argmax[\delta_{t-1}(j)a_{ji}]}_{1 \le j \le 2}, \text{for } 2 \le t \le T
\end{aligned}
$$

3. Backtracking the path:
$$
\begin{aligned}
\Sigma_T^* &= \underbrace{argmax[\delta_T(j)]}_{1 \le j \le 2} \\
&= \phi_{t+1}(\Sigma_{t+1}^*), \text{for } t = T - 1, \ldots, 1.
\end{aligned}
$$

where $\Sigma_1^* \ldots \Sigma_T^*$ is the estimated state sequence.

In our case, Viterbi algorithm maximizes an objective function $G(\underline{S})$ which can be written as a sum of "gain" functions depending on one state and its preceding one:

$$
G(\underline{S}) = g_1(S_1) + g_2(S_2, S_1) + + g_T(S_T, S_{T-1}) \tag{7.5}
$$

Suppose in the optimal state sequence Σ^* the state at the lth position is Σ^*=k. Maximizing $G(S_1, , S_T)$ is equivalent to maximizing the following two functions separately:

$$
\begin{aligned}
G_{l,k}(S_1, , S_{l-1}) &= g_1(S_1) + g_2(S_2, S_1) + + g_l(k, S_{l-1}) \\
G_{l,k}^*(S_{l+1}, , S_T) &= g_{l+1}(S_{1+1,k}) + + g_T(S_T, S_{T-1}),
\end{aligned}
$$

which involve only the states before and after l .

Using the estimated probabilities from the previous section 7.1.2, we compute the g functions based on POX values.

To do that we used the smoothed values of the POX collected simultaneously with images whose normalized version is plotted in fig.7.8.

We modified slightly the Viterbi algorithm reducing it to the search of the state path $\hat{\Sigma}_1^*, \ldots, \hat{\Sigma}_{512}^*$ that selects only S_1. The lines estimated with high probability of being scanned in this state path have been assembled in a new image, as described in section 7.2.

7.1.4 Prediction of the Number of Replicates

With the rodent being alive under the microscope, the observed images include lines scanned when the brain was in S_1 or in S_2. Given the classification models from previous chapter **??** describing the brain movement caused by physiological events, we can predict the number of lines that are scanned in S_1 during a POX cycle.

From the repeated scanned line experiment used for analysis in chapter **??** during a POX cycle there are scanned about 67 lines, out of which approximately 40 have higher probability to be collected in S_1. Table 7.1 summarizes these counts and rates for 7 POX cycles that would normally happen while scanning 512 lines.

Consequently, we need at least two replicates to be able to reconstruct images with lines that have high probabilities to be scanned in S_1. Increasing the number of replicates will provide lines with higher probabilities than those coming from only 2 replicates.

7.2 RECONSTRUCTED IMAGES

Applying Viterbi algorithm to this data provided us with a sequence of lines most likely to be scanned in S_1 state. Assembling them allows us to recreate an image with no gaps. One outcome of the HMM reconstruction method is presented in fig.7.9.

A good quality image should present the features smooth and continuous. Our resultant image has still some discontinuities that need to be addressed. We estimated the length and
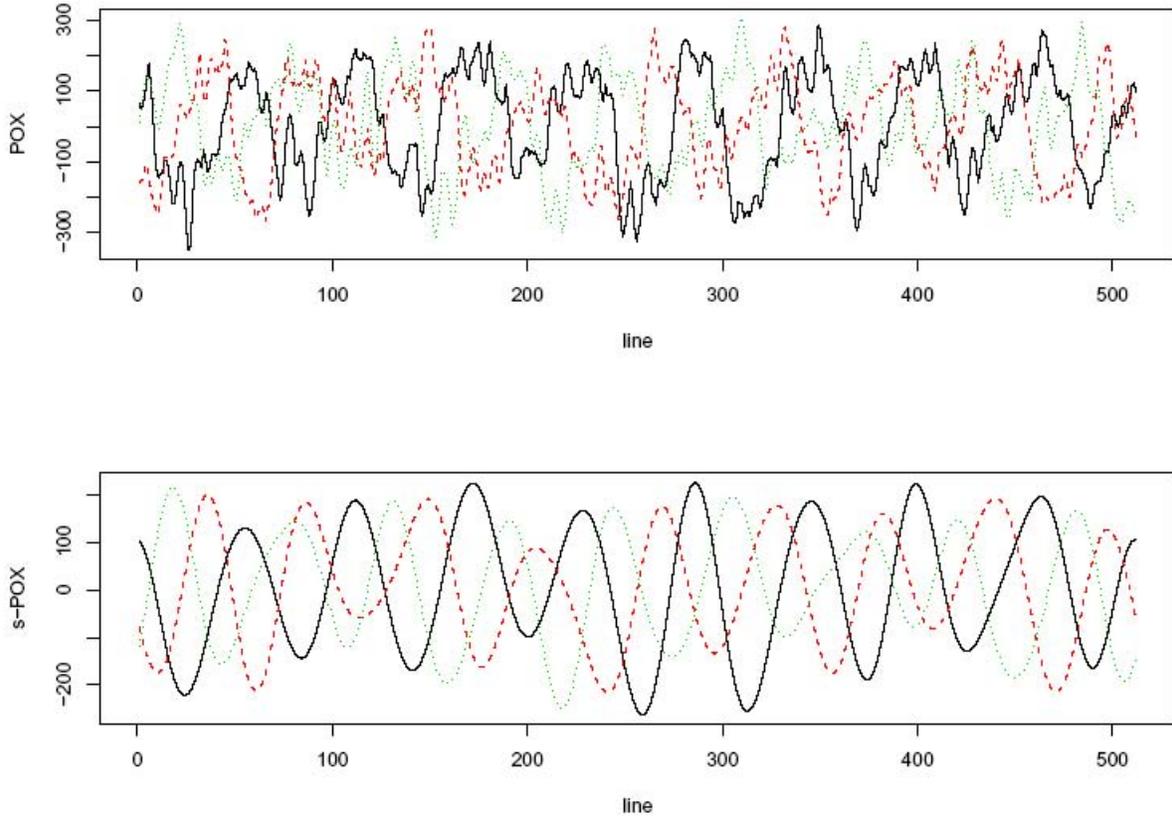
Figure 7.8: Observed POX (upper) vs Smoothed POX (lower) data. Each color represents the physiology associated with one of the three replicates collected for this experiment: black (full line) for 1st time, red (−) for 2nd time, green (..) for 3rd time.

event	lines per event (counts)	lines in S_1 (counts)	lines in S_1 (%)
POX cycle 1	71	36	0.51
POX cycle 2	65	32	0.49
POX cycle 3	65	41	0.63
POX cycle 4	65	48	0.74
POX cycle 5	63	38	0.60
POX cycle 6	69	43	0.62
POX cycle 7	71	44	0.62
average for all 7 cycles	67	40.2857	0.602 (0.007)

Table 7.1: Estimated rates of scanning lines in S_1 during a POX cycle.
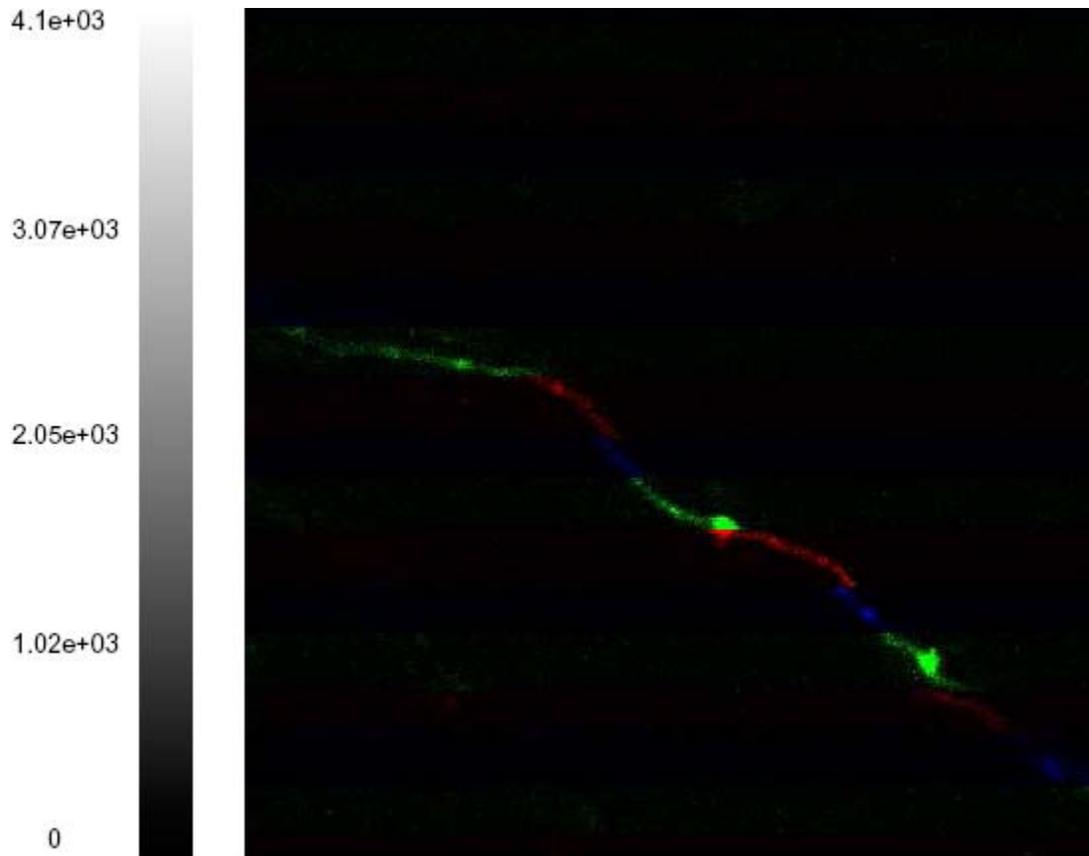
Figure 7.9: An estimate of the "true"' image based on HMM: it includes the line that corresponds to the estimated "near to rest" (S_1) state. The lines coming from 1st replicate are color-coded in red, the ones coming from 2nd replicate are color-coded in green, and those coming from 3rd replicate are color-coded in blue. Note: there are small drifts at the junction regions.
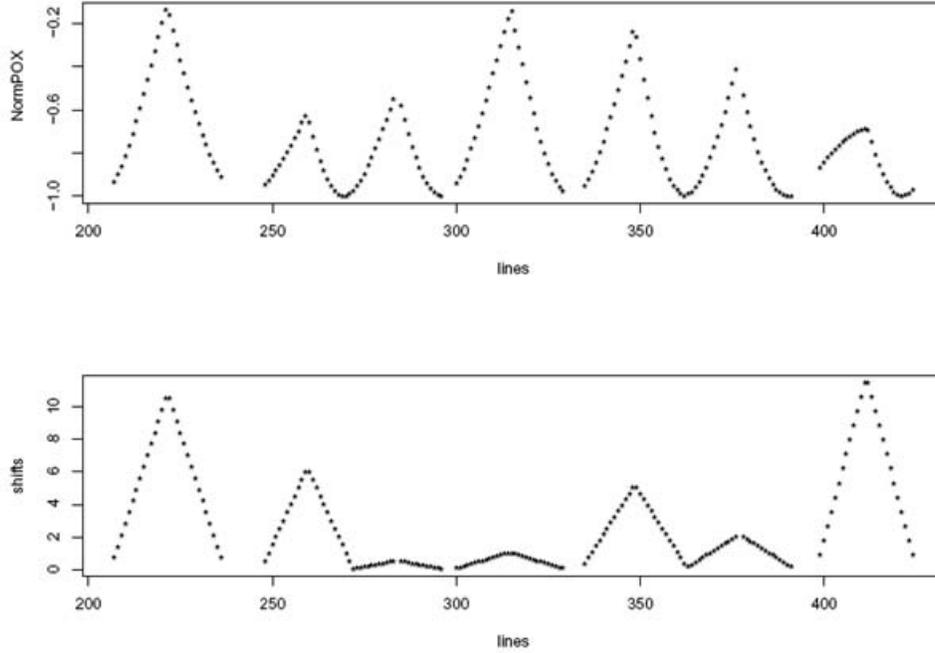
Figure 7.10: Estimated drifts for the reconstructed image 7.9 (bottom). Concomitant normalized POX values are presented above.

direction of the shifts that the lines need to be translated with, such that the correlation of the shifted lines to be maximized.

$$\hat{\delta}x(\underline{u}, \underline{v}) = argmax_{lag}(\rho(\underline{u}, \underline{v})), \tag{7.6}$$

where $\underline{u}$ and $\underline{v}$ are two adjacent lines at a junction point. Fig. 7.10 presents the estimated drifts for the resultant image 7.9 and emphasizes the association between these shifts and the POX values at those times.

The Shifting property of Fourier transform 4.7 aligned the lines estimated to be scanned in S_1 producing an improved reconstructed image (**??**), free of physiological movement artifacts.

7.2.1 Fatal experiment

At this point, we needed a validation procedure to compare the model fit. In the absence of a true reference image, we compared the resulting images with the observed ones, using a least square approach and a correlation method. This approach has not yielded any satisfactory

109
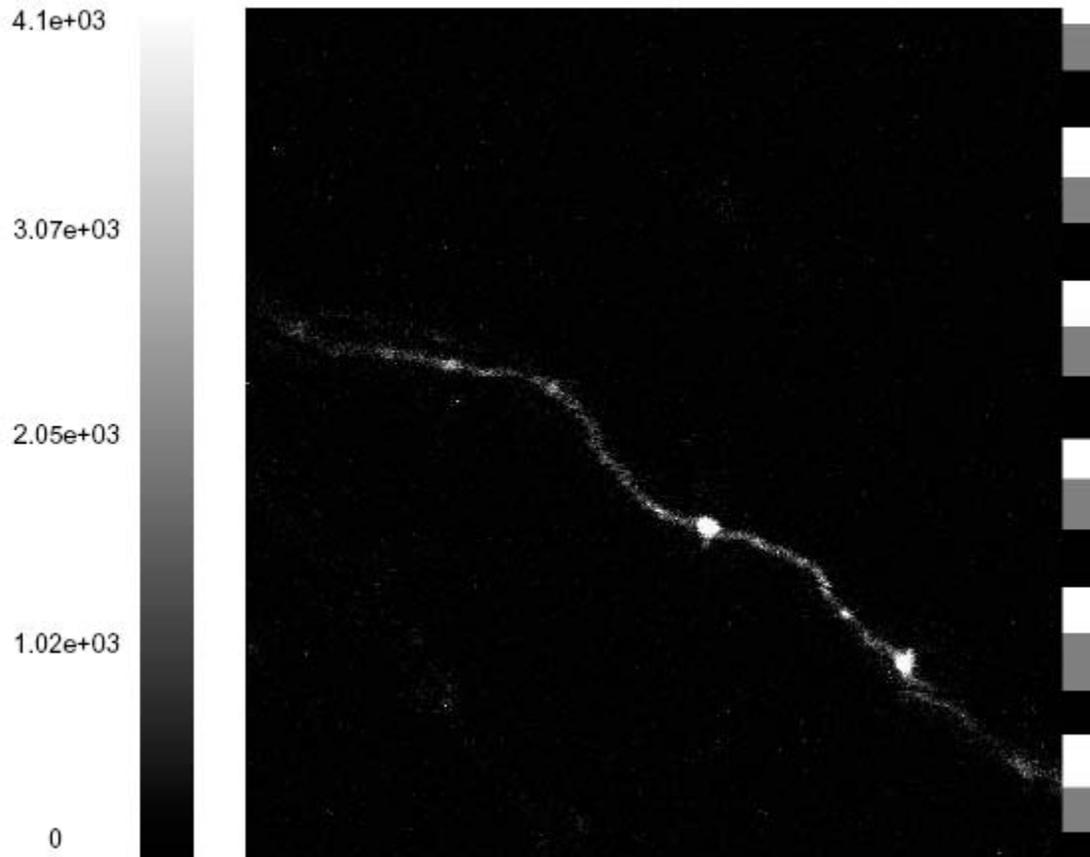
Figure 7.11: An improved estimate of the "true" image based on HMM (7.9): the lines corresponding to the estimated "near to rest" (S_1) state are aligned to preserved the continuity of the feature. The panel in the right indicates which replicate captured the line in this state with black for the 1st and white for the last.
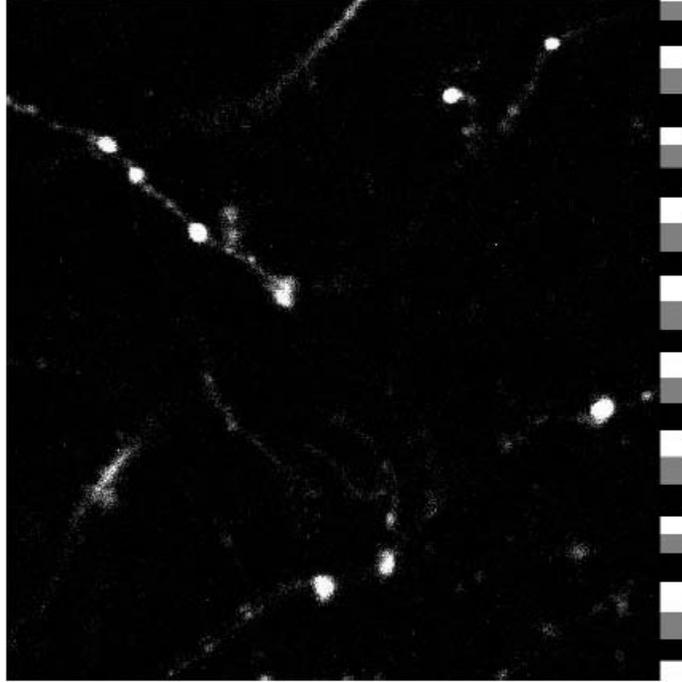
Figure 7.12: The estimated image through the semi-HMM applied on data collected in triggered mode. Note: the panel in the right side indicates the time when the lines have been scanned in S_1

results.

However, a new set-up experiment (fatal for the specimen) would provide us with an image free of physiological movement artifacts. We hope that we could use it to validate our reconstruction algorithm.

The experimental steps have been designed as following:

1. Perform z-stacks of triggered images while the animal is alive (3 replicates).

2. Inject the specimen with high dose of lethal substance and image its tissue.

The semi-HMM framework has been applied on images collected during this kind of experiments. A resultant estimate is shown in fig 7.12.

The estimated image (7.12) has been compared with a sample of images scanned while the brain was not affected by the blood pressure any more. Due to natural changes occurring in the spatial distribution of the fluorescent features, we considered several slices scanned at a depth that included the z position when the animal was alive. The 2D cross-correlation function of the estimated image with the sample of slices from the dead brain was maximized

111
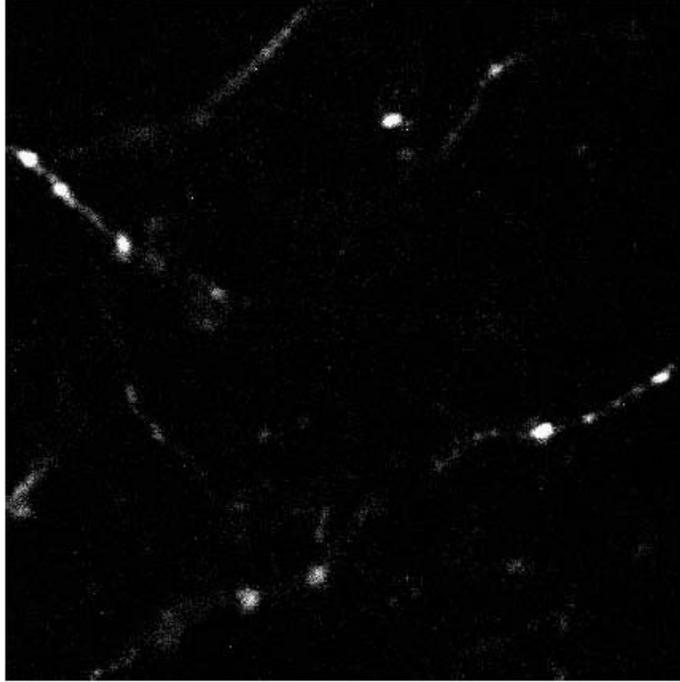
Figure 7.13: The slice taken when the animal was dead that maximizes the 2D cross-correlation function with the HMM-estimated image of the live brain.

at the image shown in fig 7.13.

We notice that the features moved considerably from one image to the other, making their comparison quite difficult. Future work on this direction might prove whether our suggestion was helpful in validation the reconstruction procedure.

7.3   CONCLUSIONS

Using the results on the movement of the brain concluded in the previous chapter (**??**), we proposed a data driven set-up using a semi-hidden Markov model framework. Given that the observed lines come from three replicates of the same scene in the brain at different times, we estimated the ones with high probability to be scanned in S_1 state. Put altogether, these lines create a rough estimate of the scene in the brain with the image artifacts due to physiological movement being minimized. After identifying the discontinuities in the resultant estimate, an improved image has been produced.

8.0 DISCUSSION

It is a frequent misconception that scientific research results in simple clear answers to questions. The truth is that almost every answer results in a whole battery of new questions. But the research serves to increase our understanding so that we know how to frame the new questions and try to tackle them. (William Greenough)

This dissertation presents a data-driven framework for reconstructing images from Two Photon Laser Scanning Microscopy. We show that the data-driven approach allows us to generate subject -specific, quantifiable, predictive models in spite of current measuring limitations. We describe a classification model for tracking the movement of the brain and then another probabilistic model for reconstructing images. Each model has specific strengths for various aspects of the modeling and together they offer relatively simple and sufficiently accurate estimations of biologically relevant measures.

We instantiate the framework by developing a collection of computational algorithms, analysis and visualization tools. We adjust the acquisition technique to collect hybrid image data, with lines scanned repeatedly for several hundreds of times. This approach enables us to observe the variations in feature positions among the scanned lines. Next we incorporate some of these results and experimental information into a Hidden Markov Model with two states: "near-to-rest" and "far-from-rest". As we are interested in the image regions where the movement was minimal, the lines observed in this state are estimated using the Viterbi algorithm.

8.1 IMAGE FILTER -CONTRIBUTIONS

When working with high resolution imaging techniques, details matter at a high extent. Image analysis methods are necessary to describe and extract biological phenomena of interest. In our case, we analyze large sets of high bit depth neuronal images where fine neurites must be identified and tracked over time. Therefore we pay increased attention to each pixel value and we develop a novel thresholding algorithm to increase the signal to noise ratio. We consider several values within the pixel intensities range that will act as thresholds for a new image with fewer levels on the grey scale. Consequently, we define a "many- to- one" transformation that applies to each pixel of the input image.

Choosing the correct threshold values is critical, as they will determine which pixels are labeled as signal and which are noise. We estimate them by maximizing the correlation between the original image and the transformed image.

The advantages of using this novel algorithm may be emphasized from several points of view:

- Visually, the new image has less background noise, while some of the features have clearer contours than in the original picture.

- The spatial variation of the pixels is preserved, due to the properties of the correlation function. The filtered image is highly correlated with the initial one, even though the pixel values are significantly reduced.

- From the analysis perspective, the number of parameters to estimate is significantly decreased. Also, the within-feature variation of pixel intensities gets significantly low, increasing the precision of other analysis methods. For example, classifying lines based on between features variation has been improved not only in the matching results, but also in computer time and memory.

- Although the function to maximize has not a closed analytical form, the simple underlying idea makes it accessible to a vast research population. The use of the genetic algorithms in a C environment has allowed us to estimate a multi-dimensional solution of this optimization problem. However, after taking several theoretical steps into a deeper

analysis of this function, we became aware of new characteristics that could be revealed by future work.

8.2   BRAIN MOTION -CONTRIBUTIONS

Tracking the trajectory of a target is usually done using time-space models. In our case, we take a classification approach conditioned on some inherent constraints. The fact that we are not able to capture an image free of movement artifacts is an ongoing limitation for which we need to adjust our analysis. Therefore, for the repeated scanned line experiment we track the position of a line by looking it up within images taken before and after this experiment (defined as classes). We start with an Euclidean distance classification rule which we develop into a cross-correlation framework.

Specific contributions of this work include:

- Conditioned on the "before" and "after" images, we estimate the shifts in (x,y) plane and we represent the trajectory of a line under the microscope. This method has also lead to the creation of a brain motion database of unprecedented detail. Although it was designed for multiple species, angles and directions, most of the results presented in this thesis come from a single direction data acquisition. Future work would enrich this database with more details and complexity.

- The classification model is quantifiable, estimating the amplitude of the movement during the experiment. It also has predictive capabilities in computing the expected position under "regular" physiological conditions.

- The number of lines scanned in each class can be used to compute how much time the brain spends in that position. This statistic could be used to determine how many replicates are necessary for the image reconstruction model.

- The trajectory obtained with respect to the "after" image is consistent with the one estimated with the "before" image. This gives a measure of confidence in the accuracy of our results.

- Fitting these classification models ("before" and "after") to the repeated scanned lines points out the goodness of fit both visually but also computationally via the mean squared error.

- A step further, the extensive use of a correlation based thresholding filter increased the accuracy by decreasing the variation in pixel intensities. This novel algorithm requires as sole input the usual image with M grey levels and outputs a highly correlated image with K grey levels, $K << M$.

- The class matching was performed without the need of physiological measurements, but statistical tests confirm the significant association that exists between these processes. A careful analysis of this relationship gives indices regarding the "near-to-rest" state of the brain.

8.3   IMAGE RECONSTRUCTION -CONTRIBUTIONS

The image reconstruction system we developed takes as input biological images (with several replicates scanned in a triggered mode) and the concomitant physiological measurements, estimating new images constructed from the regions where the brain movement was minimal.

The results of this work bring innovations to the field because:

- It includes the effect of subject's physiology in estimating the periods when the brain motion was minimal.

- It offers a probabilistic solution to main aspects related to movement artifacts:

    1. some lines are scanned several times, therefore we need to decide which line was in the "near-to-rest " state;

    2. some lines are never scanned, therefore we need to find lines that can fill these gaps, considering the inherent correlations with the neighboring pixels.

- It estimates reliable *in vivo* images overcoming the current imaging technology limitations. The advantages that TPLSM offers to science could be used with higher confidence to draw conclusions about complex structures.

We used a simplistic model that assumes the system oscillating only between two states. The estimated classes represented in fig6.10 do not expand to more than 8 y-positions. Therefore the scanning process can not enter in more than 3 states, as, from the classification model, we considered at least 2 classes per state. The transition probabilities were considered fixed, hence the results are biased by the assumptions we made. However we could extend the model by placing some random error around these probabilities.

Ultimately, this work creates a modeling approach that has the potential to more simply and efficiently explain the brain motion caused by heart beat and respiration, improving the quality of *in vivo* images.

8.4   FUTURE WORK

There are limitations to our work, some caused by the current limited data availability, and some due to our computational approach. From the repeated scanned line data, we modeled the brain motion caused by subject's physiology on (x,y) plane. However the movement is three directional, therefore future experiments would consider this component when designing the set-up. The models that we develop in this research could be then extended to a 3D space. Also, a challenging project would be to extend these analyses to fractional pixel manipulations.

There are further limitations in our computational modeling of the lines observed in the "near-to-rest"' state. We estimate the transition/output probabilities based on process related factors, but the estimation could always be improved. Also, as the time to scan a pixel can be at least 0.64 μs, scanning an image of 512x512 pixels will last at least 400 ms which will correspond to at least 2 heart-beats. Even though this speed is too high to capture the necessary amount of detail for the intended area, the scanner is still not fast enough to take an image free of movement artifacts. Consequently, we do not have a golden standard image to validate the estimates of chapter 7. Further improvements in the technological field would be able to verify our models and might impact the improvement of TPLSM.

The collection of tools presented in this dissertation solves significant challenges in the

process of reconstructing images from *in vivo* laser scanning microscopy. Being able to identify "real"' features will allow advances in tracking the progress of brain development and in modeling the neural dynamic over the longitudinal course. It will open also new perspectives in the manipulation of brain activity, learning about sensory experience, alteration in gene expression and pharmacological manipulations.

APPENDIX A

BIOLOGICAL TERMS

Hystology: the microscopic study of the structure of tissues.

in vivo: inside the living body.

neuron: the information-processing cell of the nervous system, also called nerve cell. Most neurons use action potentials to send signals over a distance. All neurons communicate with one another using synaptic transmission.

optic chiasm: called after the Greek letter χ, this is the structure in which the right and left optic nerves converge and partially cross to form the optic tracts.

photoreceptor: a specialized cell in the retina that transduces light energy into changes in membrane potential.

retina: a thin layer of cells at the back of the eye that transduces light energy into neural activity.

synapse: the region of contact where a neuron transfers information to another cell.

APPENDIX B

CORRELATION- BASED THRESHOLDING IMAGES

B.1 K THRESHOLDS CORRELATION FUNCTION

By definition, the correlation function between two vectors of length N $\underline{x}$ and $\underline{y}$ is

$$corr(\underline{x}, \underline{y}) = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{N\sqrt{var(\underline{x})var(\underline{y})}} = \frac{\sum_{i=1}^{N}(x_i - \overline{x})y_i}{N\sqrt{var(\underline{x})var(\underline{y})}} \tag{B.1}$$

as $\overline{y}\sum_{i=1}^{N}(x_i - \overline{x})$ is zero.

When $\underline{y}$ is defined as

$$y(i) = \begin{cases} 0 & \text{if } x(i) \leq k_1 \\ 1 & \text{if } k_1 < x(i) \leq k_2 \\ 2 & \text{otherwise} \end{cases}$$

then B.1 becomes:

$$\frac{\sum_{j=0}^{k_1} n_j(j - \overline{x}) * 0 + \sum_{j=k_1+1}^{k_2} n_j(j - \overline{x}) * 1 + \sum_{j=k_2+1}^{M} n_j(j - \overline{x}) * 2}{\sqrt{N * var(\underline{x})(\sum_{j=1}^{k_1} n_j(0 - \overline{y})^2 + \sum_{j=k_1+1}^{k_2} n_j(1 - \overline{y})^2 + \sum_{j=k_2+1}^{M} n_j(2 - \overline{y})^2)}}$$

which simplifies to:

$$\frac{N_{k_1}(\overline{x_{k_1}} - \overline{x}) + N_{k_2}(\overline{x_{k_2}} - \overline{x})}{\sqrt{(\sum_{j=0}^{M} n_j(j - \overline{x})^2)((N - N_{k_1})(0 - \overline{y})^2 + (N_{k_1} - N_{k_2})(1 - \overline{y})^2 + N_{k_2}(2 - \overline{y})^2)}}$$

which is 3.1.

The same algorithm is applied when generalizing $\underline{y}$ to K thresholds.

Then, dividing the numerator and denominator in 3.1 by N we get the 3.4.

B.2   DENOMINATOR OF THE CORRELATION FUNCTION

The denominator in 3.6 has under the square root a sum that can be decomposed based on the expansion of the square term.

$$\sum_{i=0}^{K}(u_{i+1}-u_i)(i-\sum_{j=1}^{K}u_j)^2 = \sum_{i=0}^{K}[(u_{i+1}-u_i)i^2 - 2i(u_{i+1}-u_i)S_K + (u_{i+1}-u_i)S_K^2] \quad \text{(B.2)}$$

where $S_K = \sum_{j=1}^{K}u_j$ does not depend on i.

Therefore, the sum in B.2 is decomposed into 3 telescopic sums, as following:

$$[\sum_{i=0}^{K}(u_{i+1}-u_i)]S_K^2 = [u_1 - u_0 + u_2 - u_1 + ... + u_{K+1} - u_K]S_K^2 \quad \text{(B.3)}$$

$$= S_K^2. \quad \text{(B.4)}$$

$$-2[\sum_{i=0}^{K}i(u_{i+1}-u_i)]S_K^2 = -2[u_2 - u_1 + 2u_3 - 2u_2 + ... + Ku_{K+1} - Ku_K]S_K^2 \quad \text{(B.5)}$$

$$= -2[K - \sum_{j=1}^{K}u_j]S_K \quad \text{(B.6)}$$

$$= 2S_K^2 - 2KS_K. \quad \text{(B.7)}$$

$$\sum_{i=0}^{K}(u_{i+1}-u_i)i^2 = u_2 - u_1 + 4u_3 - 4u_2 + ... + K^2u_{K+1} - K^2u_K \quad \text{(B.8)}$$

$$= K^2 - \sum_{j=1}^{K}(2j-1)u_j. \quad \text{(B.9)}$$

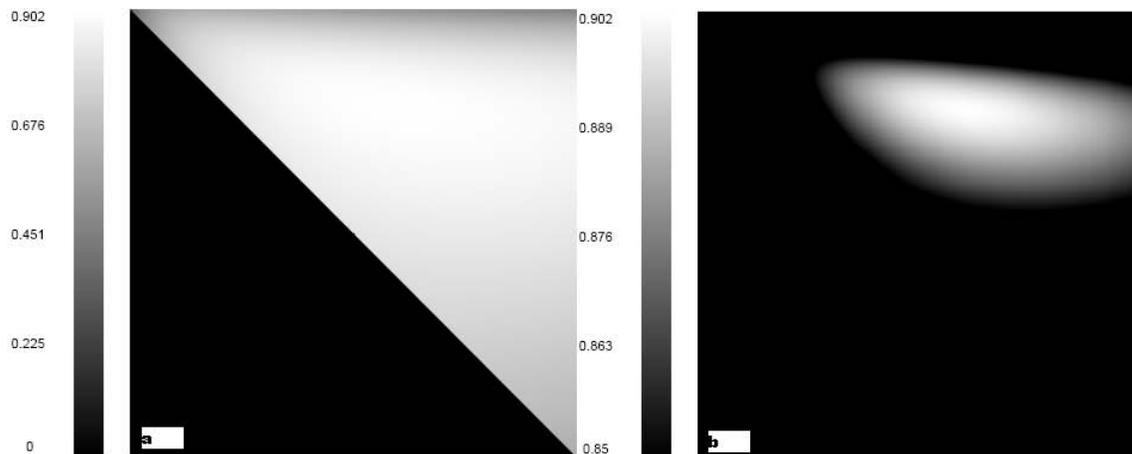as $u_{K+1} = P(X < M) = 1$ and $u_0 = P(X < 0) = 0$.

Figure B1: The correlation function between the original image (Fig.2.5) and any tri-colored image constructed with 2 thresholds that are displayed on y (k_1) and on x (k_2). (a) is showing all the values, (b) the grey scale is changed to enhance the concavity of the function. Note: the actual values are on the upper right side of the image, while in the lower left there are only zeros.

Summing these components (B.4 + B.7 + B.9) we get

$$K^2 - \sum_{j=1}^{K} (2K + 2j - 1)u_j + 3(\sum_{j=1}^{K} u_j)^2$$

APPENDIX C

RESULTS ON FOURIER TRANSFORM

C.1 CONVOLUTION THEOREM

$$\mathbf{F}(f(t) * g(t)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)g(t-x)dx\,exp(-iwt)dt \tag{C.1}$$

$$= \int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} g(t-x)exp(-iwt)dtdx \tag{C.2}$$

$$= \int_{-\infty}^{\infty} f(x)G(w)exp(-iwx)dx \tag{C.3}$$

$$= \int_{-\infty}^{\infty} f(x)exp(-iwx)dxG(w) = F(w)G(w) \tag{C.4}$$

C.2 AUTOCORRELATION THEOREM

$$\mathbf{F}(f(x)f^*(x-t)dx) = \int_{-\infty}^{\infty} exp(-iwt) \int_{-\infty}^{\infty} f(x)f^*(t-x)dxdt \tag{C.5}$$

$$= \int_{-\infty}^{\infty} f(x)[\int_{-\infty}^{\infty} exp(iwt)f(x-t)dt]^*dx \tag{C.6}$$

$$= \int_{-\infty}^{\infty} f(x)[\int_{-\infty}^{\infty} exp(-iws)f(x+s)ds]^*dx \tag{C.7}$$

$$= \int_{-\infty}^{\infty} f(x)[F(w)exp(iwx)]^*dx \tag{C.8}$$

$$= \int_{-\infty}^{\infty} f(x)exp(-iwx)dxF^*(w) = F(w)F^*(w) = |F(w)|^2 \tag{C.9}$$
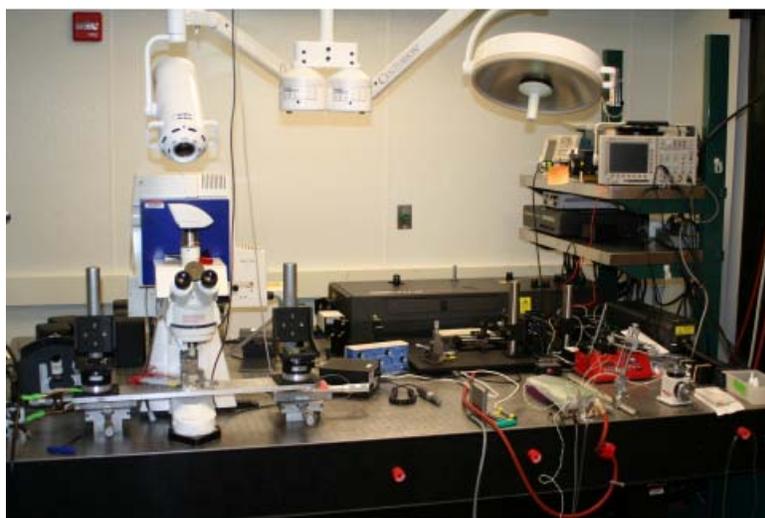
APPENDIX D

ADDITIONAL REPRESENTATIONS

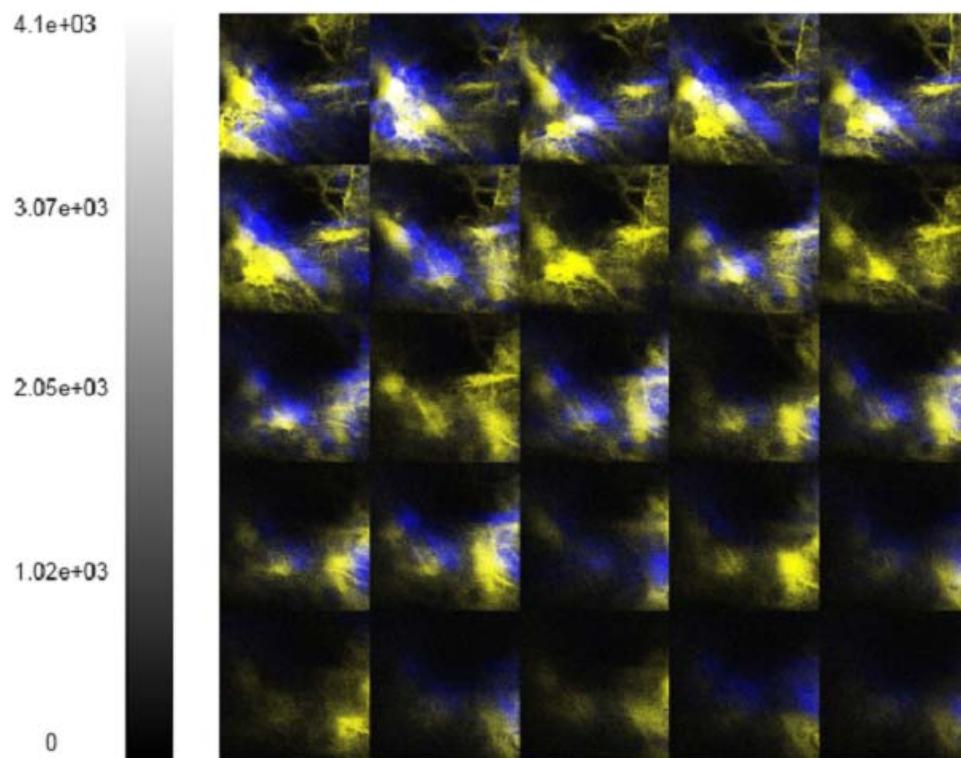Figure D1: An extended view on the TPLSM from Dr. Crowley's lab.

Figure D2: The RGB version of the z-stack collected from a ferret before and after the repeated scanned line experiment. The color variation shows that the brain moved at a large amplitude.
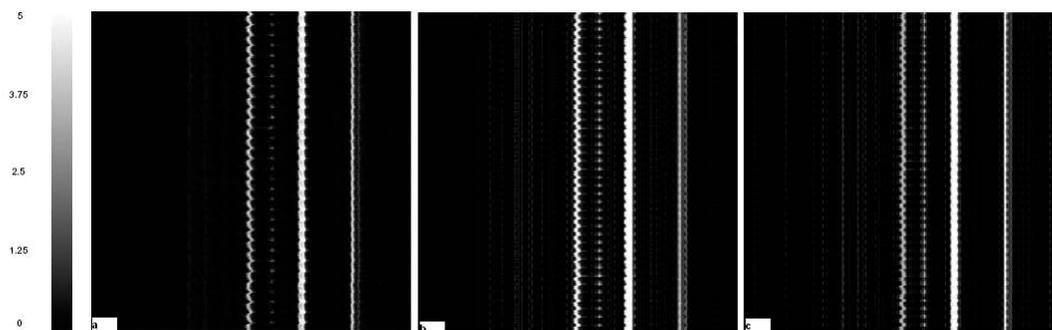

Figure D3: The fitted values for the lines scanned at the first 2560 times, estimated with the classification model applied on I^0 (b), and I^{T+1} (c). The observed lines are shown in (a).
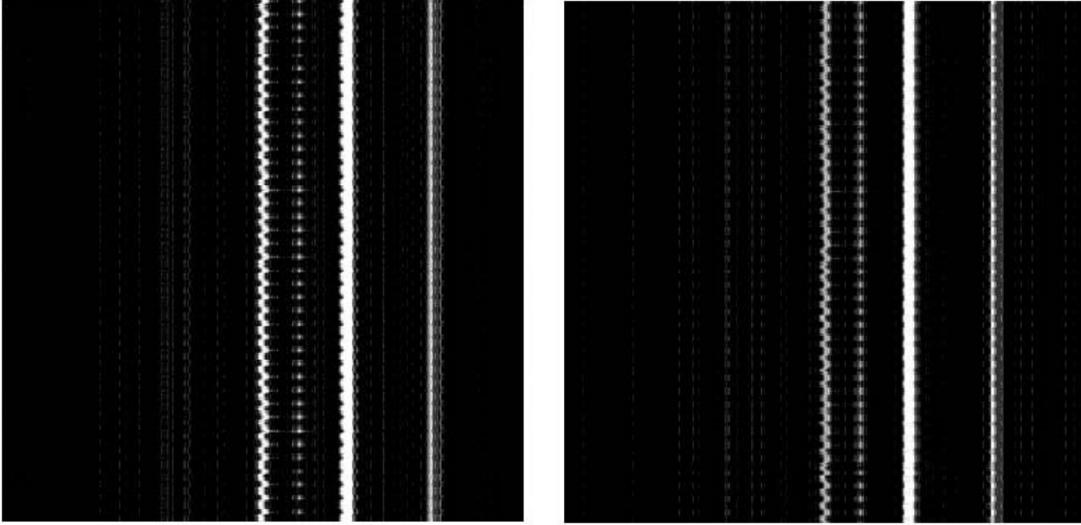
Figure D4: The fitted lines for 5.1 estimated with the classification model applied on I^0 (left), and I^{T+1} (right).
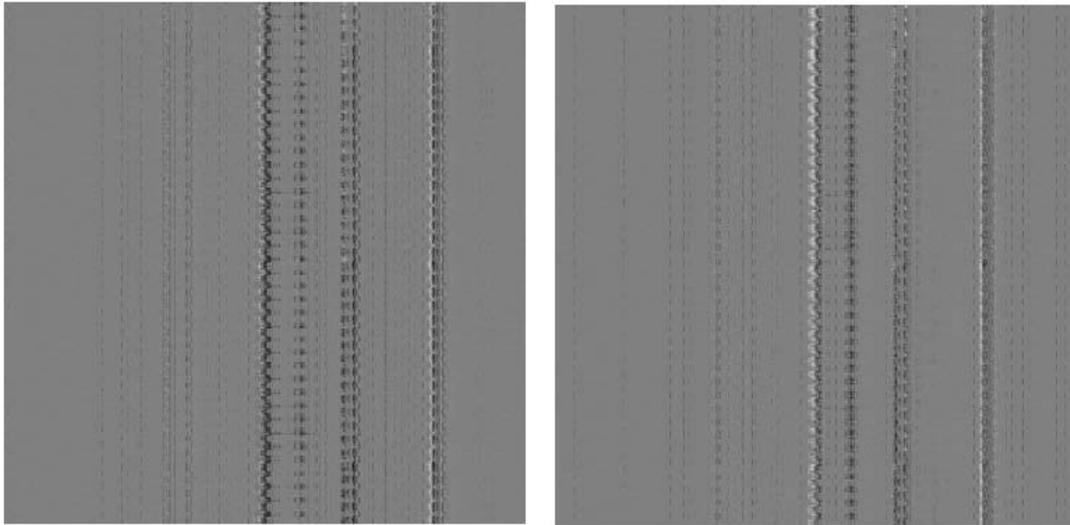

Figure D5: The residuals estimated with the classification model applied on I^0 (left), and I^{T+1} (right).
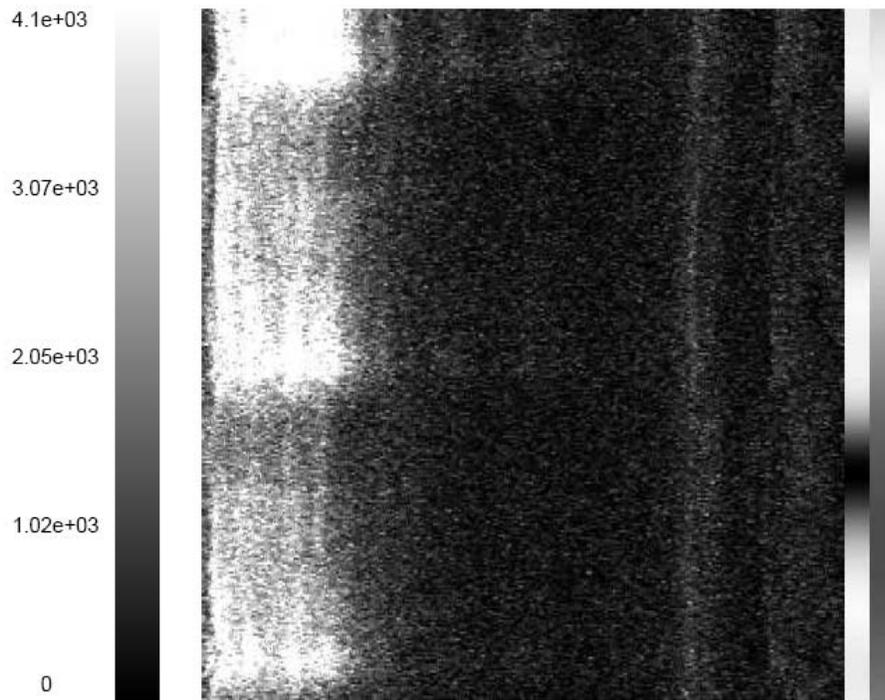
Figure D6: The lines scanned for 600 times during a repeated scanned line experiment on a neonatal ferret. Note: on the right side are added indicators for estimated POX and respiration values during this time.

BIBLIOGRAPHY

[1] D. Avitzour. Stochastic simulation bayesian approach to multi-target tracking. *IEE Proceedings-Radar Sonar and Navigation*, 142:41–44, 1995.

[2] M.F. Bear, B.W. Connors, and M.A. Paradiso. *Neuroscience: exploring the brain.* Lippincott Williams and Wilkins, 2007.

[3] R. Chen and J. S. Liu. Mixture kalman filters. *Journal of the Royal Statistical Society*, B(62), 2000.

[4] W. Denk, J. Strickler, and W. Webb. Two-photon laser scanning fluorescence microscopy. *Science*, 4951(248):73–6, 1990.

[5] T. Dittegen, A. Nimmerjahn, S. Komai, P. Licznerski, J. Waters, T.W. Margrie, F. Helmchen, W. Denk, M. Brecht, and P. Osten. Lentivirus-based genetic manipulations of cortical neurons and their optical and electrophysiological monitoring in vivo. *PNAS*, 101(52):18206–18211, 2004.

[6] A.D. Dombeck, A.N. Khabbaz, F. Collman, T.L. Adelman, and D.W. Tank. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron*, 56:43–57, 2007.

[7] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison. Probabilistic models of proteins and nucleic acids. In I.N.Bankman, editor, *Biological sequence analysis*. Cambridge University Press, 1998.

[8] W.F. Eddy, M. Fitzgerald, C. Genovese, Mockus A., and D.C. Noll. Functional imaging analysis software- computational olio. *COMPSTAT Proceedings in Computational Statistics, Physica-Verlag, Heidelberg*, pages 39–49, 1996.

[9] W.F. Eddy, M. Fitzgerald, and D.C. Noll. Improved image registration using fourier interpolation. *Magnetic Resonance in Medicine*, (36):923–931, 1996.

[10] W.F. Eddy and T. Young. Optimizing the resampling of registered images. In I.N.Bankman, editor, *Handbook of Medical Image Processing*, pages 603–612. Academic Press, 2000.

[11] E. Falkenauer. *Genetic Algorithms and Grouping problems.* John Wiley and Sons Ltd, 1997.

[12] G. Feng, R.H. Mellor, M. Bernstein, C. Keller-Peck, Q. T. Nguyen, M. Wallace, J.M. Nerbonne, J.W. Lichtman, and J. R. Sanes. Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of gfp. *Neuron*, (28):41–51, 2000.

[13] J.D. Fergusson. Variable duration models for speech. *Proc. Symposium o the Application of Hidden Markov Models to Text and Speech*, pages 143–179, 1980.

[14] S. Finger. *Origins of neuroscience: a history of explorations into brain function.* Oxford University Press US, 2001.

[15] C.A. Glasbey and G.W. Horgan. *Image analysis for the biological sciences.* Wiley, 1995.

[16] M. Goeppert-Mayer. Uber elementarakte mit zwei quantensprungen. *Annalen der Physik*, 401(3):273–95, 1931.

[17] R.C. Gonzalez and R.E. Woods. *Digital image processing.* Pearson Education, 2002.

[18] D.S. Greenberg and J.N. Kerr. Automated correction of fast motion artifacts for two-photon imaging of awake animals. *J. Neuroscience*, (176):1–15, 2009.

[19] A.J. Holmaat, J.T. Trachtenberg, L. Wilbrecht, G.M. Shepherd, X. Zhang, G.W. Knott, and K. Svoboda. Transient and persistent dendritic spines in the neocortex in vivo. *Neuron*, (45):279–291, 2005.

[20] A.J. Holmaat, L. Wilbrecht, G.W. Knott, E. Welker, and K. Svoboda. Experience-dependant and cell-type-specific spine growth in the neocortex. *Nature*, (441):979–983, 2006.

[21] A.I. Holodny. *Functional Neuroimaging: a clinical approach.* Informa Healthcare, 2008.

[22] N.A. Lazar, W.F. Eddy, C.R. Genovese, and J.S. Welling. Statistical issues in fmri for brain imaging. *International Statistical Review*, (69):105–127, 2001.

[23] J. Liu. *Monte Carlo strategies in scientific computing.* Springer, 2001.

[24] A. Mizrahi, J.C. Crowley, E. Shtoyerman, and L.C. Katz. High-resolution in vivo imaging of hippocampal dendrites and spines. *J. Neuroscience*, 13(24):3147–3151, 2004.

[25] J.A. Nelder and R. Mead. A simplex method for function minimizations. *Computer Journal*, (7):308–313, 1965.

[26] K. Padmanabhan, W. F. Eddy, and J.C. Crowley. A novel algorithm for optimal image thresholding of biological data. *J.Neurosci Methods*, ([Epub ahead of print]), 2010.

[27] L.R. Rabiner. A tutorial on hidden markov-models and selected applications in speech recognition. *Proc.IEEE*, (77):257–286, 1989.

[28] M. Rudin and R. Weissleder. Molecular imaging in drug discovery and development. *Nat.Rev.Drug.Discov.*, 2(2):123–131, 2003.

[29] R. Shumway and D. Stoffer. *Time Series analysis and its applications with R examples.* Springer, 2006.

[30] L. Sirovich and E. Kaplan. Methods for in vivo optical imaging of the central nervous system. In Frosting R.D., editor, *In Vivo Optical Imaging of Brain Function*, pages 43–76. CRC Press LLC, 2001.

[31] P.T. So, C.Y. Dong, B.R. Masters, and Berland K.M. Two-photon excitation fluorescence microscopy. *Annu.Rev.Biomed.Eng.*, 2:399–429, 2000.

[32] P.T.C. So. Two-photon fluorescence light microscopy. In www.els.net, editor, *Enciclopedia of life sciences.* Macmillan Publishers Ltd, Nature Publishing Group, 2002.

[33] R.C. Truex and M.B. Carpenter. *Human neuroanatomy.* Williams and Wilkins, 1969.

[34] J.S. Welling, W.F. Eddy, and T.K. Young. Rotation of 3d volumes by fourier-interpolated shears. *Graphical Models*, 68:356–370, 2006.