A Computational Framework for the Structural Change Analysis of 3D Volumes of Microscopic Specimens

Madhusudhanan Balasubramanian
Louisiana State University and Agricultural and Mechanical College, madhu@csc.lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Part of the Computer Sciences Commons

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_dissertations/3066

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
A COMPUTATIONAL FRAMEWORK FOR THE STRUCTURAL CHANGE ANALYSIS OF 3D VOLUMES OF MICROSCOPIC SPECIMENS

A Dissertation
Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
The Department of Computer Science

by
Madhusudhanan Balasubramanian
B.E. in Electrical and Electronics Engineering, Bharathiyar University, 1998
M.S. in Systems Science, Louisiana State University, 2003
May 2006
Acknowledgments

I would like to thank Dr. S. Sitharama Iyengar who has served as my Major Professor, Dissertation Advisor, Master’s thesis committee member and a mentor during my tenure at LSU, Baton Rouge. From the initial stages of taking graduate courses to my dissertation preparation, he constantly nurtured my research interests, provided me the freedom to explore, critiqued my research ideas, provided references to collaborate with various researchers in the field of computational biology / medicine and gave moral support and encouragement in times of need.

I thank Dr. Peter Wolenski who has served as a Minor Professor in Mathematics in my Doctoral Dissertation work. The graduate math-tuneup workshop in the summer of 2002 provided me an opportunity to meet and interact with Dr. Wolenski. I gained much of my interests in Hilbert Spaces during the Summer of 2003, when I worked with him on identifying an optimal basis discovery approach for a glaucoma study and I learned a lot by constantly interacting with him from thereon.

With my interests orienting towards computational biomedicine, Dr. Roger Beuerman provided me opportunities to work on several interesting problems. Those were crucial to strengthen my interests in the field, help me define my dissertation topic and as well choose a career path in the interdisciplinary field of computational biology / medicine. I appreciate and thank him for his efforts in explaining the clinical aspects of the research, for his careful evaluation of the analysis, techniques and the results. I thank Dr. Doris Carver for her guidance and for serving in my dissertation committee. My sincere thanks to Dr. Bijaya Karki for partially funding my research, for his guidance and for serving in my dissertation committee.

My interactions and discussions with Dr. Hilary Thompson, Associate Professor, LIONS Research Laboratories, New Orleans, LA, have always been educating, inspiring and goal defining. I have learnt and further developed my interests in Bio-Statistics while working with him on several projects. I thank him for his continued guidance and support. I would like to thank Dr. Sundar, Professor, Department of Mathematics for his help in teaching me “Hilbert spaces” on a special reading class. Also, I would to thank Dr. David Koppelman, Associate Professor, Department of Electrical and Computer Engineering, LSU for serving as a Dean’s Representative in my dissertation committee and for his guidance.
The classes I took with Dr. John Tyler, Professor, Department of Computer Science, LSU, on “Computational Methods”, “Introduction to Wavelets” and “High Performance Computing” and the lectures on “Scientific Computing” with Dr. Paul Seylor, Visiting Professor of Mathematics, University of Illinois at Urbana-Champaign, assisted me to lay a strong foundation on cluster computing and high performance computational methods. I would like thank them and the other faculty members in the Department of Computer Science, LSU. My sincere thanks to Juan Reynaud for being available for discussions, for assisting my research activities at the LIONS Eye Research Laboratories and peer reviewing several of my research manuscripts. Special thanks to Balaji Vasudevan for his time and efforts to review my research manuscripts.

I would like to thank Dr. Jost Goettert and Dr. Yohannes Desta with the LSU Center for Advanced Microstructures and Devices (CAMD) for their consultation and supplying microstructures for calibrating the white-light confocal microscope. The research support from the Biological Computation and Visualization Center (BCVC) and the Louisiana Biomedical Research Network (LBRN) have been invaluable for conducting this research during my doctoral studies. My sincere thanks to Dr. Harold Silverman for providing the BCVC research support. I extend my thanks to all the committee members with the LBRN for awarding a financial support for my research proposals during two summers and an academic year.

Thanks to Steven Coorpender, LIONS Eye Research Laboratories, New Orleans, LA, John Quebedeaux, Biological Science, LSU, Dr. Honggao Liu and other members of High Performance Computing team, LSU for arranging computer hardware for this research.

I have no words to express my gratitude and appreciation for my parents, sisters and brother. I do not know how I could reciprocate the sacrifices my parents have made all along. They are the source of my inspiration and consoling. This dissertation research and all of my other endeavors would not have been possible without their support, encouragement and guidance.

LSU provided me another opportunity to get acquainted with wonderful people. I thank my wonderful friends Alagiri, Arun, Balaji, Larkin, Lucas, Prak, Prakash, Rajesh, Ramesh, Ryan, Robby, Suressh and Vijai for providing me an excellent environment to work through my dissertation and for sharing great moments. I thank the almighty for the motherly love, care and affection.
Table of Contents

Acknowledgments .......................................................... ii
List of Tables .................................................................. vi
List of Figures ................................................................ vii
Abstract ........................................................................ x

Introduction .................................................................... 1

Chapter 1. Confocal-4D: An Architecture for Real-time Tracking
and Visualization of White-light Confocal Microscope Optical
Serial Sections ................................................................. 4
  1.1 Introduction ............................................................... 4
  1.2 Hardware Arrangement ............................................... 8
  1.3 Tracking Optical Serial Sections Using Template Matching 10
    1.3.1 Optical Section Dimension Reduction Using Wavelet Multi-
         resolution Analysis .................................................. 13
  1.4 Software Architecture for Real-time Tracking of Optical Sections 15
    1.4.1 Building Blocks of a Multithreaded Realtime Optical Section
         Tracking System ...................................................... 16
    1.4.1 Server Architecture .............................................. 20
    1.4.2 Master-Client Architecture .................................... 21
    1.4.3 Servant-Client Architecture .................................... 21
  1.5 Results .................................................................... 23
  1.6 Conclusion ............................................................... 26

Chapter 2. Adaptive Noise Filtering and Restoration of White-
light Confocal Microscope Images .................................. 28
  2.1 Introduction ............................................................... 28
  2.2 Karhunen-Loeve Expansion ......................................... 31
    2.2.1 Direct Method to Compute the Optimal Basis ................. 32
    2.2.2 Method of Snapshot .............................................. 33
  2.3 Adaptive Noise Filtering and Restoration of WLCM Images .. 33
  2.4 Performance Evaluation of the Restored Images Using a Contrast Measure ................................................. 34
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lamina cribrosa optical section images at 12mmHg tracked using the template matching algorithm presented in section 1.3 after increasing the chamber pressure to 24mmHg and 60mmHg</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>Performance evaluation of DIC coarse-fine optical flow estimation with (CNP-DIC) and without (DIC) Compute-and-Propagate framework</td>
<td>57</td>
</tr>
</tbody>
</table>
# List of Figures

1. Anatomy of eye ............................................. 1
2. Schematic representation of a white-light confocal microscope ... 4
3. Schematic representation of aberrant light rays from adjacent imaging points passing through the detector pinholes ................. 5
4. Principle of operation of a Nipkow scanning disk .................... 7
5. White-light confocal microscope at LSU Eye Center, New Orleans, LA with DC-Mike Actuator and DC-Motor Controller/Driver added for real-time z-axis scanning and volume acquisition .................. 9
6. WLCM z-Stacks .............................................. 10
7. Template matching algorithm to identify the location of best match 11
8. Overall system architecture for real-time tracking of optical sections 16
9. Multithreaded server design for confocal control and image acquisition 17
10. State transition diagram of the dual-buffer operation ............... 19
11. Servant-client architecture for tracking optical sections ............ 20
12. Multithreaded architecture of the master-client for receiving the sensor confocal volumes and coordinating the tracking in a cluster facility 22
13. Correlation surface using normalized correlation measure and FFT based correlation measure ........................................ 25
14. Volume of a random cotton fiber constructed from optical sections using ray tracing ........................................... 27
15. Images of random cotton fibers restored using Lucy-Richardson restoration algorithm; contrast measure increases exponentially due to ringing artifacts .............................................. 37
16 Experimental PSF of the WLCM determined by imaging 5-micron diameter spheres .................................................. 39

17 Ensemble of random cotton fiber images, their KL decomposition and the inverse restored images using median and adaptive KL prefiltering ................................................................. 40

18 Random cotton fibers (a, d, g) restored using inverse filter deconvolution after prefiltering noise using median (b, e, h) and adaptive KL filters (c, f, i) .......................................................... 41

19 Image of a random cotton fiber (a) and lamina cribrosa optical section at 12mmHg IOP level (d) restored using inverse filter deconvolution after prefiltering the image noise using median and adaptive KL filters .................................................. 42

20 Contrast measure performance comparison of median and adaptive KL prefilters with the inverse restoration method .......................................................... 43

21 Computing the ringing metric of an image restored using LR iterative restoration algorithm .................................................. 44

22 An image sequence showing the changes in a scene in a short interval δt .................................................................................. 45

23 A scene depicting the movement of an object undergoing transformation .......................................................... 47

24 Digital Image Correlation based optical flow estimation .......................................................... 50

25 Discrepancy between the ground truth optical flow $U_{GT}$ and the estimated optical flow vector $U_{Est}$ .................................................. 51

26 CNP framework providing an initial value for the iterative differential or coarse-fine correspondence method .................................................. 53

27 Schematic representation of the propagation of the regional flow vectors (sparse) to successive levels in the Compute-And-Propagate framework for optical flow computation .................................................. 55

28 Cotton image $I_{t_0}$ shifted 10 pixels diagonally to generate $I_{t_0+\delta t}$. The purpose is to demonstrate the validity of the proposed error metric for evaluating optical flow results .................................................. 56

29 Optical flow estimation of a scene containing a clock and a mug .......................................................... 59
| 30 | Optical flow estimation of a scene exhibiting movement of cotton fibers observed using a white-light confocal microscope | 61 |
| 31 | Optical flow estimation of a scene exhibiting a rotating tabletop globe | 62 |
| 32 | Optical flow estimation of a scene exhibiting a rotating tabletop globe with some lateral movement | 63 |
| 33 | Optical flow estimation of the images of lamina cribrosa of cow retina at intraocular pressure levels of 12mmHg and 60mmHg | 64 |
| 34 | Optical flow estimation of the images of lamina cribrosa of cow retina at intraocular pressure levels of 12mmHg and 60mmHg restored using KL filter | 65 |
| 35 | Optical flow estimation of the images of a rotating sphere | 66 |
| 36 | Euclidean and IMED error metrics of the DIC and CNP-DIC methods on the experimental image sequences | 67 |
| 37 | Convergence rate (Iteration counts) of the DIC and CNP-DIC methods on the experimental image sequences | 68 |
Abstract

Glaucoma, commonly observed with an elevation in the intraocular pressure level (IOP), is one of the leading causes of blindness. The lamina cribrosa is a mesh-like structure that provides axonal support for the optic nerves leaving the eye. The changes in the laminar structure under IOP elevations may result in the deaths of retinal ganglion cells, leading to vision degradation and loss. We have developed a comprehensive computational framework that can assist the study of structural changes in microscopic structures such as lamina cribrosa. The optical sectioning property of a confocal microscope facilitates imaging thick microscopic specimen at various depths without physical sectioning. The confocal microscope images are referred to as optical sections. The computational framework developed includes: 1) a multi-threaded system architecture for tracking a volume-of-interest within a microscopic specimen in a parallel computation environment using a reliable-multicast for collective-communication operations 2) a Karhunen-Loève (KL) expansion based adaptive noise prefilter for the restoration of the optical sections using an inverse restoration method 3) a morphological operator based ringing metric to quantify the ringing artifacts introduced during iterative restoration of optical sections 4) a $l_2$ norm based error metric to evaluate the performance of optical flow algorithms without a priori knowledge of the true motion field and 5) a Compute-and-Propagate (CNP) framework for iterative optical flow algorithms. The realtime tracking architecture can convert a 2D-confocal microscope into a 4D-confocal microscope with tracking. The adaptive KL filter is suitable for realtime restoration of optical sections. The CNP framework significantly improves the speed and convergence of the iterative optical flow algorithms. Also, the CNP framework can reduce the errors in the motion field estimates due to the aperture problem. The performance of the proposed framework is demonstrated on real-life image sequences and on z-Stack datasets of random cotton fibers and lamina cribrosa of a cow retina with an experimentally induced glaucoma. The proposed framework can be used for routine laboratory and clinical investigation of microstructures such as cells and tissues, for the evaluation of complex structures such as cornea and has potential use as a surgical guidance tool.
**Introduction**

The eye is a complex organ that generates the visual information from the light entering the eye. Figure 1 shows the schematic representation of the cross-section of the eye. Maintaining its spherical shape is crucial for the proper optical functioning of the eye and is accomplished by the aqueous humor flow. The aqueous humor is produced by the ciliary body into the posterior chamber and enters into the anterior chamber through the lens and iris. The aqueous humor distributes nutrients and immune responses in case of inflammation or infections to the lens, cornea and trabecular meshwork. Also, aqueous humor flow removes metabolic wastes by draining into the schlemm’s canal through the trabecular meshwork. The angle between the iris and cornea is referred to as the angle. A proper clearance at the angle is needed for maintaining a normal pressure, referred as intraocular pressure (IOP), in the eye measured in millimeters of mercury (mmHg). The light entering the eye through the cornea is focussed on to the retina. Retina is a photoreceptive layer consisting of millions of axons originating from the neuron cell bodies within the ganglion nerve fiber layer. The axonal fibers group together to form the optic nerve and leave the eye at the blind spot located in the posterior end of the eye.
The optic disc region is usually referred as optic nerve head or papilla and exhibit a natural ‘cup’ shape due to the arrangement of the optic nerves leaving the eye. Lamina cribrosa is a mesh-like structure present in the optic nerve head region. The lamina acts as a pressure barrier between the intraocular space and the retrobulbar space and also provides axonal support for the optic nerves leaving the eye [17].

Glaucoma is an ocular disease usually observed with an elevation in the IOP levels. When left undiagnosed and untreated, it may cause progressive optic nerve damage and loss of vision [6]. An imbalance in the aqueous humor inflow into the posterior chamber and outflow from the posterior chamber through the trabecular meshwork causes an elevation in the IOP level. Based on the source of the blockage at the trabecular meshwork, glaucoma is classified as primary or secondary. Primary glaucoma is caused by a raised iris blocking the trabecular meshwork (open angle) or due to a blockage at the angle (closed angle). Secondary glaucoma is due to an elevation in the IOP level from an injury, a tumor or cataract [7]. An acute or sustained elevation in the IOP level causes optic nerve damage. Increased cupping in the ONH region is generally observed prior to visual field loss. Under IOP elevations, the lamina cribrosa undergoes surface structure variations [1, 5, 4, 2] and backward bowing [3]. It is speculated that the laminar deformation at elevated IOP levels causes insults and trigger a series of biochemical events leading to the death of the retinal ganglion cells [17]. The associated mechanical pressure may decrease the blood flow to the optic nerve head region. Therefore it is very important to study and understand the laminar deformation, in vivo, to understand its role in glaucoma. A confocal microscope can be used to image thick microscopic specimens such as lamina cribrosa at various depths without physically sectioning the specimen [8].

The changes in the structure of lamina cribrosa, under experimental glaucoma condition, can be studied by identifying the changes at various z-axis depths of the specimen using the confocal optical section images. The deformation to the structure may involve an axial and lateral movement of structures including contraction, expansion, regional rotation and skewing of structures. Analyzing dynamic process such as glaucomatous condition in the eye, requires continuous evaluation of the structures involved. Thus a volume-of-interest within the structure have to be scanned continuously to observe and quantify the structural changes. We used a custom built white-light confocal microscope available at LIONS Eye Research Laboratory, LSU Eye Center, New Orleans, LA for conducting the research presented in this dissertation. The images acquired using a white-light confocal microscope suffer from blur due to non-ideal lens, imaging conditions and light contributions from out-of-focus layers. Therefore, the images need to be restored prior to analyzing the optical sections.

In this dissertation, we have developed a comprehensive computational framework to assist the study of structural changes in microscopic specimen such as
lamina cribrosa. The computational framework developed includes:

1. *A multithreaded architecture for realtime tracking* of microscopic structures,

2. *An adaptive noise filtering algorithm* using Karhunen-Loève expansion to filter random artifacts in the optical section images,

3. *A ringing metric* to quantify the ringing artifacts introduced during iterative restoration of images

4. *A $l_2$ norm based error metric* for performance evaluation of optical flow algorithms and

5. *A Compute-and-Propagate framework* for improving the speed and convergence of the iterative optical flow algorithms. The optical flow algorithm will be used for identifying the in-plane structural changes between the corresponding optical sections.

The main body of this dissertation is organized as three main chapters. Each chapter contains a separate introduction, main body of research, results and conclusion sections to maintain clarity in the presentation. In chapter 1, we present the hardware and software architectures for realtime tracking and visualization of volumes of interest in microscopic specimen using the white-light confocal microscope. We have developed an adaptive noise filter using Karhunen-Loève expansion for removing random noise artifacts in the confocal images. Chapter 2 discusses the adaptive noise filter for prefiltering the optical section images prior to restoring the images using inverse filter restoration method. A ringing metric for quantifying the ringing artifacts in the images restored using iterative restoration algorithm is included in chapter 2. In chapter 3, we present the Compute-and-Propagate framework for use with iterative optical flow estimation algorithms. The conclusion chapter summarizes the contributions of this dissertation. We used lamina cribrosa images of a cow retina and images of random cotton fibers to evaluate the performance of the realtime tracking algorithm and the optical section restoration algorithms. In addition to the lamina and cotton optical section images, we used pictures of real-life scenes to evaluate the performance of the Compute-and-Propagate framework for iterative optical flow algorithms. A review of linear algebra relevant to this dissertation is included in the appendix chapter.
Chapter 1. Confocal-4D: An Architecture for Real-time Tracking and Visualization of White-light Confocal Microscope Optical Serial Sections

1.1 Introduction

A confocal microscope allows imaging thick microscopic structures at various depths without physically sectioning the specimen. The images obtained using a confocal microscope are referred to as optical sections. The concept of confocal microscopy was invented by Marvin Minsky in 1955 [9]. Figure 2 shows a schematic representation of a confocal microscope. The light from a light source passes through

---

Figure 2: Schematic representation of a white-light confocal microscope
a light-source pinhole and gets reflected by a dichroic mirror. An objective lens focuses the thin ray of light and illuminates a small volume in the specimen. Assuming that the objective lens has a focal length $f$ and an imaging plane is at a distance $i$ from the objective to collect the reflected light from the specimen, then from the lens equation,

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f} \tag{1}$$

a layer at a depth $o$ will be on focus. The layer is referred as an *on-focus* layer. However, the light from the layers neighboring to the on-focus layer, referred as *out-of-focus* layers, will also be collected by the objective lens. If a finitely small pinhole is placed at the focal length $f$ of the objective lens, the light coming from the out-of-focus layers can be effectively blocked at the pinhole as shown in the figure 2. As indicated in the original work of Marvin Minsky [10], the challenge will be in choosing an optimal pinhole size as well as allowing sufficient light to pass through.

Figure 3: Schematic representation of aberrant light rays from adjacent imaging points passing through the detector pinholes

For imaging an entire on-focus layer in the specimen, the whole area needs to be imaged point-by-point. An initial opinion about the area scanning for image formation might be to place an array of pinholes and image the entire layer so as to parallelize and avoid the point-by-point scanning method of image formation. However, it should be noted that it is the point-wise imaging that allows
elimination of the out-of-focus light contribution \cite{10}. For example, consider placing additional pinholes (Detector pinholes 1 and 3) adjacent to the pinhole in the original configuration (Detector pinhole 2) to detect imaging points $P_1$, $P_2$ and $P_3$ as shown in figure 3. The light rays from the imaging points $P_1$ and $P_3$ are not shown for clarity. It can be observed that using an array of pinholes for imaging multiple imaging points in parallel will seriously compromise the confocal principle and will result in a blurred image in the image plane due to aberrant rays from the out-of-focus planes passing the pinholes. We can avoid this problem by keeping the pinholes sufficiently apart and scanning the imaging area. Confocal microscopes use a Nipkow scanning disk for real-time imaging. Figure 4 shows the schematic representation of a Nipkow scanning disk. To compensate for the reduced light throughput due to smaller pinhole size, modern confocal microscopes use brighter light sources such as Laser and Mercury (HBO) and Xenon (XBO) arc-discharge lamps. Confocal Scanning Laser Microscopes (CSLM) use a rotating mirror arrangement to scan the imaging area using laser beams, but the scan speed is relatively slower compared to the Nipkow disk based point-scanning confocal microscopes. A patent pending Disk Scanning Unit (DSU) from Olympus \cite{11} uses a random pattern of slits that form virtual pinholes when the disk spins at 3000 rpm. It is claimed that the DSU provides a high light throughput while achieving a high scanning rate available in point-scanning systems.

We use a custom-built white-light confocal microscope (Advanced Scanning Limited, New Orleans, LA) available in the LIONS Research Laboratories, LSU Eye Center, New Orleans, LA, illuminated by a xenon light source \cite{13}. White light sources such as xenon lamps provide a non-coherent illumination and reduces speckle noise compared to coherent laser light sources \cite{12}. The structure of a thick microscopic specimen such as thick cells, tissues and cotton fibers can be reconstructed by building a z-Stack of optical sections acquired using a confocal microscope at a regular z-axis depth intervals. Confocal microscopes have been used for studying the surface cells of the human cornea \cite{13, 16}, examining the three-dimensional structure of the lamina cribrosa of the retina in relation to the effects of elevated intra-ocular pressure levels \cite{17, 18}, for the analysis of bacterial infections in the cornea due to contact lens wear \cite{15, 19}, for morphologic characterization of the cells in cornea \cite{13}, for studying corneal wound healing after surgery \cite{14} and for studying cellular structures in Gums, Skin and bones \cite{20}.

For observing cellular events in tissues, analyzing microscopic structural changes under various diagnostic conditions and for the investigation of pathological states of tissues, we need the ability to observe site specific changes \textit{in vivo}. Assuming that the specimen under observation has a definite shape and structure, the changes in its structure can be studied from their optical sections. In present laboratory and clinical practices, a trained technician or a physician manually locate the pathological site by referring to the optical sections obtained in its initial state. We have designed and developed a hardware arrangement and a novel
software architecture that can convert a confocal microscope into a 4D-confocal microscope with the ability to track a reference volume-of-interest in a specimen in approximate real-time. In section 1.2 we discuss the hardware arrangement for
z-axis control of the confocal microscope and an image acquisition hardware. In section 1.3, we investigate the choice of an optimal matching measure for tracking optical section images of a microscopic specimen with image dimension-reduction using wavelets for tracking the optical sections in real-time. We discuss a software architecture for real-time optical section image acquisition, tracking a reference volume-of-interest using optical sections in a cluster computing environment and visualization of tracked z-Stack data-sets in section 1.4. In section 1.5 we discuss the results of the presented methods on z-Stack datasets of lamina cribrosa of retina in an experimental glaucomatous condition and on random cotton fibers. We conclude this chapter in section 1.6.

1.2 Hardware Arrangement

By adding an additional z-axis control, a 2D point-scanning confocal microscope or a laser scanning microscope can be converted into an area-scanning confocal microscope generating a z-Stack of volume datasets at the end of each scan. Figure 5 shows the white-light confocal microscope (WLCM) available at the LSU Eye Center with additional hardware added for real-time volume acquisition. It should be noted that the hardware additions similar to the proposed additions are commonly used for volume acquisition [21]. An ultra-high-precision DC-Mike actuator M-227.25 (Physik Instrumente, Irvine, CA) is coupled with the confocal objective lens control as shown in the figure 5. The actuator controls the object distance \( o \) in the lens equation \( 1 \) by generating a linear motion up to 25mm for imaging a specimen at various depths. A DC-Motor controller C-862 sends control commands to the actuator. The DC-Motor controller is connected to a computer through an RS232 cable.

The M-227 actuator allows a velocity up to 1000 \( \mu m/s \). The WLCM can acquire 30 images/sec. Thus it takes approximately 33 milliseconds to complete a single scan. Depending on the z-axis resolution of the microscope and the required sample distance between adjacent optical sections, the velocity of the actuator can be adjusted to operate in a scan-mode. Multiple images can be acquired at a given depth to compensate for additional processing at each z-axis depth. Depending on the required z-axis interval between the optical sections, the scan-mode velocity can be adjusted. With the DC-Motor controller and actuator control, a selective volume-of-interest (VOI) in a specimen can be imaged at a finer z-axis interval. Thus we can achieve a more accurate 4D representation of the specimen. We use a monochrome frame grabber DT3155 (Data Translation, Marlboro, MA) PCI card for image acquisition. The frame grabber allows a frame acquisition rate of 30 frames/second.

In our implementation, the users of the WLCM can specify a z-axis depth range \( (D_{Begin} \rightarrow D_{End}) \), a depth interval \( D_{interval} \) and define a VOI for observation and
Figure 5: White-light confocal microscope at LSU Eye Center, New Orleans, LA with DC-Mike Actuator and DC-Motor Controller/Driver added for real-time z-axis scanning and volume acquisition.

analysis. Once the scan depth $D_{End}$ is reached, the DC-Motor controller adjusts the velocity of the actuator to a reset-mode to quickly bring the focus back to the top of the VOI at depth $D_{Begin}$. Once the focus is brought back to the top of the VOI, the actuator velocity is set to operate in scan-mode. Typically in scan-mode, the actuator travel speed is slower to allow image acquisition at a finer interval. Velocity of the actuator in the reset-mode is close to its maximum velocity to quickly bring back the focus to the top of the VOI. An alternative scanning method will be to scan the VOI both in a top-down and bottom-up fashion to eliminate the time lost in bringing the focus back to the top of VOI at the end of each scan. With top-down and bottom-up scanning, once a VOI is defined, the WLCM will operate in a scan-mode alone, generating a stream of z-Stack volumes for visualization and further analysis.
1.3 Tracking Optical Serial Sections Using Template Matching

With the hardware additions described in section 1.2, the WLCM can generate a stream of z-Stack datasets. To track a VOI \( Vol_{Ref}(t_0, N) \) acquired in a reference state of the specimen at time instant \( t_0 \), we need to identify the optical sections in the current state of the specimen, represented by the z-Stack \( Vol_{Sensor}(t_1, M) \) at time instant \( t_1 \), matching with the optical sections in \( Vol_{Ref}(t_0, N) \). Here, \( N \) is the number of optical sections in the reference z-Stack \( Vol_{Ref} \) and \( M \) is the number of optical sections in the newly acquired z-Stack \( Vol_{Sensor} \) at time \( t_1 \). We keep \( Vol_{Sensor} \) size \( M \) larger than the \( Vol_{Ref} \) size to account for any volume expansions.

A recently acquired z-Stack volume will be referred as a sensor z-Stack or volume, a terminology commonly used in tracking systems. Figure 6 shows a graphical representation of the z-Stacks.

![Figure 6: WLCM z-Stacks](image)

The volume tracking problem can be redefined as tracking the individual optical sections in the reference z-Stack. For each optical section image \( I_{Ref}(i) \) in the reference volume, we need to find the corresponding optical section \( I_{Sensor}(j) \) in the sensor volume. After acquiring the reference volume at time instant \( t_0 \) the features at a z-axis depth \( i \) may have undergone one of the following changes:

1. The portion of the structure at depth \( i \) moved to a new depth \( j \)
2. All or part of the structure initially observed at depth \( i \) moved laterally by \( (\delta x, \delta y) \) and axially to a new depth \( j \)

If the only possible changes to the specimen are the axial movement of the optical sections, then the sensor optical sections are misaligned in the z-axis but spatially...
registered with the reference optical sections. Then an optical section $I_{\text{Ref}}(i)$ in
the reference z-Stack $Vol_{\text{Ref}}(t_0, N)$ can be tracked by determining its similarity
(or dissimilarity) with all the optical sections $\{I_{\text{Sensor}}(j)\}_{j=1}^M$ in the sensor volume
$Vol_{\text{Sensor}}(t_1, M)$. The optical section in the sensor volume resulting in a highest
similarity (or lowest dissimilarity) will be chosen as the tracked optical section.
This can be repeated for all the optical sections in the reference volume to build a
z-Stack of tracked optical sections. When there are multiple best matches (with ap-
proximately same similarity or dissimilarity measures), an optical section $I_{\text{Sensor}}(j)$
(among the best matching optical sections) closest to $I_{\text{Ref}}(i)$, depth wise, will be
chosen as the best match.

![Template matching algorithm](image)

**Figure 7:** Template matching algorithm to identify the location of best match

When there is a lateral movement of structures, a direct image similarity mea-
surement may not provide an optimal match. If the volume acquisition rate is kept
sufficiently high, then any lateral structural movement can be kept to a minimum
between subsequent volume acquisitions. Therefore, a majority of the structures
in the center of the reference optical section $I_{\text{Ref}}(i)$ can still be observed in the
corresponding optical section in the sensor volume $Vol_{\text{Sensor}}(t_1, M)$ with a lateral
shift. Using a template region that contains a large portion of the optical section
image (say 80% of the size of the optical section image) from the center, we can
search for an optimal matching location in each of the optical section images in
the sensor volume $Vol_{\text{Sensor}}(t_1, M)$ using a template matching algorithm [67]. The
template matching algorithm constructs several moving windows as shown in figure
7 to identify a region in a sensor optical section $I_{\text{Sensor}}(j)$ that best matches with
the template from the reference optical section. A measure of similarity (or dis-
similarity) at the best matching location will be used as the measure of similarity
(or dissimilarity) between the optical section $I_{\text{Ref}}(i)$ and optical section $I_{\text{Sensor}}(j)$.
Similarly, the measure of match between a given optical section $I_{\text{Ref}}(i)$ in the
reference volume with the rest of the optical sections $\{I_{\text{Sensor}}(j)\}_{j=1}^M$ in the sensor
volume can be computed. The optical section $I_{\text{Sensor}}(j)$ with the highest measure
of match will be identified as the tracked optical section in the sensor volume.
As mentioned earlier, when there are multiple best matches, the matching optical
section closest to the reference optical section (in the z-axis) will be chosen as the
corresponding optical section.

Similarity or a dissimilarity measure used in the template matching algorithm can be chosen from Area Based Matching (ABM) methods and Feature Based Matching (FBM) methods [68]. In the ABM methods, the correspondence between a pair of images is estimated from their pixel intensity values. In FBM methods, the image features are extracted for estimating the correspondence between images. The reliability of the FBM methods depend on the \textit{a priori} knowledge about the structure and the reliability of the feature extraction algorithm. ABM methods do not require any \textit{a priori} knowledge of the structure. Hence we choose the ABM method of matching for volume tracking. The ABM methods use a measure of match or mismatch to estimate the degree of similarity between two images. The following image similarity and dissimilarity measures are available.

1. Normalized inner-product measure (Similarity measure)

\[
NProd(I_{Ref}(i), I_{Sensor}(j)) = \sum_{k=1}^{M} \sum_{l=1}^{N} \frac{I_{Ref}(i, k, l) \cdot I_{Sensor}(j, k, l)}{\|I_{Ref}(i)\| \cdot \|I_{Sensor}(j)\|}
\]  

2. Normalized correlation measure (Similarity measure)

\[
\frac{\sum_{k=1}^{M} \sum_{l=1}^{N} (I_{Ref}(i, k, l) - I_{Ref}(i)) (I_{Sensor}(j, k, l) - I_{Sensor}(j))}{\sqrt{\sum_{k=1}^{M} \sum_{l=1}^{N} (I_{Ref}(i, k, l) - I_{Ref}(i))^2 \sum_{k=1}^{M} \sum_{l=1}^{N} (I_{Sensor}(j, k, l) - I_{Sensor}(j))^2}}
\]  

3. Measure of absolute difference (Dissimilarity measure)

\[
Norm_{L1}(I_{Ref}(i), I_{Sensor}(j)) = \sum_{k=1}^{M} \sum_{l=1}^{N} |I_{Ref}(i, k, l) - I_{Sensor}(j, k, l)|
\]  

4. Euclidean distance (Dissimilarity measure)

\[
Norm_{L2}(I_{Ref}(i), I_{Sensor}(j)) = \sqrt{\sum_{k=1}^{M} \sum_{l=1}^{N} |I_{Ref}(i, k, l) - I_{Sensor}(j, k, l)|^2}
\]  

According to the convolution theorem [67], convolution between two functions \(f(x, y)\) and \(h(x, y)\) of size \(M \times N\) defined as

\[
f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n)h(x - m, y - n)
\]  
in the spatial domain is equivalent to the product of their fourier transforms \(F(u, v)\) and \(H(u, v)\):

\[
f(x, y) * h(x, y) \Leftrightarrow F(u, v)H(u, v)
\]
Template matching using the normalized correlation measure in equation (3) is nothing but the convolution of the template with an optical section image in the sensor volume. After normalizing the images and padding zeros to the template image (the zero-padded template image size equal to the size of the sensor optical section image), the template based correlation can be implemented in the frequency domain as in equation (7). Algorithm 1 summarizes the computational steps involved in tracking the optical sections in a reference volume.

### Algorithm 1

**Algorithm for tracking optical sections in** $Vol_{Ref}(t_0, N)$

1: **procedure** TrackOpticalSections($Vol_{Ref}(t_0, N), Vol_{Sensor}(t_1, M)$)
2: Let $\{I_{Ref}(i)\}_{i=1}^{N}$ be the optical sections in $Vol_{Ref}(t_0, N)$
3: Let $\{I_{Sensor}(j)\}_{j=1}^{M}$ be the optical sections in $Vol_{Sensor}(t_1, M)$
4: Initialize $axialDisplacement \leftarrow 0$
5: Initialize $Vol_{Tracked}(t_1, N)$ volume for storing the tracked optical sections
6: for $i = 1$ to $N$ do
7: Initialize $bestCorrMeasure$ to the lowest similarity or highest dissimilarity measure possible
8: Initialize $bestMatchDepth \leftarrow i$
9: for $j = i + axialDisplacement$ to $M$ do
10: $tempCorrMeasure \leftarrow$ Similarity / dissimilarity measure between optical sections $I_{Ref}(i)$ and $I_{Sensor}(j)$ computed using template matching
11: if $tempCorrMeasure$ better than $bestCorrMeasure$ then
12: Store $bestCorrMeasure \leftarrow tempCorrMeasure$
13: Store $bestMatchingDepth \leftarrow j$
14: end if
15: if $bestCorrMeasure =$ Maximum correspondence measure then
16: Exit the inner for-loop
17: end if
18: end for
19: Update $Vol_{Tracked}(t_1, i)$ with $Vol_{Sensor}(t_1, bestMatchingDepth)$
20: Update $axialDisplacement$
21: end for
22: **end procedure**

### 1.3.1 Optical Section Dimension Reduction Using Wavelet Multi-resolution Analysis

The number of moving windows in the spatial domain can be significantly reduced by decreasing the dimension of the optical sections for tracking purpose. Among other transforms, multi-resolution analysis using wavelets is one of the most popular approaches [67]. Wavelet series expansion uses a set of scaling functions $\Phi = \{\varphi_i\}$ and wavelet functions $\Psi = \{\psi_j\}$ that form an orthogonal complement pair for the space $L^2(\mathbb{R})$ of measurable, square integrable functions containing the
where, $\oplus$ represents the orthogonal complement.

The choice and construction of the scaling and wavelets functions are discussed here [23]. Using the wavelet decomposition, the signal $f(x)$ can be expressed as the sum of the scaling functions and wavelet functions. The scaling function coefficients in the expansion of a signal are referred to as scaling or approximation coefficients and the wavelet functions coefficients are referred to as wavelet or detail coefficients. Assuming that the scaling and wavelets are orthonormal basis functions, the 1D function $f(x)$ can be expanded as

$$f(x) = \sum_i \langle f(x), \varphi_i \rangle \varphi_i + \sum_i \langle f(x), \psi_i \rangle \psi_i$$

Here, $\langle f(x), \varphi_i \rangle$ gives the scaling or approximation coefficients and $\langle f(x), \psi_i \rangle$ gives the wavelet or details coefficients. The scaling functions are constructed by integer translation and scaling of a real, square-integrable function $\varphi(x)$ as follows.

$$\varphi_{j,k}(x) = 2^{j/2} \varphi(2^j x - k)$$

where, $k$ translates $\varphi$ to cover the entire length of the signal $f(x)$ and $j$ scales $\varphi$ to analyze $f(x)$ at various resolutions. The corresponding wavelets functions can be constructed as

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$$

The factor $2^{j/2}$ normalizes the scaling ($\varphi$) and wavelet ($\psi$) functions. The input parameter $2^j x$ to the functions $\varphi$ and $\psi$ controls the width of the functions $\varphi$ and $\psi$. It can be observed that $j$ determines the number of scaling and wavelet functions required to expand the signal $f(x)$.

The discrete wavelet expansion of a 1D discrete signal $f = (f_1, f_2, f_3, \ldots, f_N)$ of length $N$, at level $L$ is given by

$$f = \sum_{i=1}^{N/2^L} \langle V^i_L, f \rangle \cdot V^i_L + \sum_{\text{level}=1}^{L} \sum_{i=1}^{N/2^\text{level}} \langle W^i_{\text{level}}, f \rangle \cdot W^i_{\text{level}}$$

Here, $V^i_L$ represents the $i^{th}$ scaling function and $W^i_L$ represents the $i^{th}$ wavelet function at level $L$. The Haar wavelet transform uses a support of two integer positions. Therefore, the wavelet coefficients may miss the high frequency changes in the signal at the locations of transition from even to odd integer indices. The Daubechies wavelet uses an overlapping windows like many other wavelet transforms. The
scaling and wavelet coefficients of the Daubechies wavelets are as follows.

\[
\text{Scaling coefficients } V_1 = \begin{pmatrix}
  h_1 & h_2 & h_3 & h_4 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\
  0 & 0 & h_1 & h_2 & h_3 & h_4 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
  0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & h_1 & h_2 & h_3 & h_4 \\
  h_3 & h_4 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

\[
\text{Wavelet coefficients } W_1 = \begin{pmatrix}
  g_1 & g_2 & g_3 & g_4 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\
  0 & 0 & g_1 & g_2 & g_3 & g_4 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
  0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & g_1 & g_2 & g_3 & g_4 \\
  g_3 & g_4 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

where, \( h_1 = \frac{1+\sqrt{3}}{4\sqrt{2}} \), \( h_2 = \frac{3+\sqrt{3}}{4\sqrt{2}} \), \( h_3 = \frac{3-\sqrt{3}}{4\sqrt{2}} \), \( h_4 = \frac{1-\sqrt{3}}{4\sqrt{2}} \), \( g_1 = h_4 \), \( g_2 = -h_3 \), \( g_3 = h_2 \) and \( g_4 = -h_1 \).

The scaling and wavelet coefficients for wavelet analysis of 2D functions such as images can be obtained using the following 2D scaling and wavelet functions.

- **2D scaling function** \( \varphi(x, y) = \varphi(x)\varphi(y) \)
- **2D horizontal wavelet function** \( \psi_H(x, y) = \psi(x)\varphi(y) \)
- **2D vertical wavelet function** \( \psi_V(x, y) = \varphi(x)\psi(y) \)
- **2D diagonal wavelet function** \( \psi_D(x, y) = \psi(x)\psi(y) \)

We will use Daubechies wavelet transform to reduce the dimension of the optical section images. Note that the approximation coefficients are the scaled down version of the original signal. After \( L \)-level wavelet decomposition as in equation (11), the approximate coefficients \( \langle V_i^j, f \rangle \) are of size \( X/2^L \times Y/2^L \), where \( X \times Y \) being the size of the original image. The dimension of the original image is reduced by a factor of \( 2^L \times 2^L \) by using only the approximation coefficients at the level \( L \). Care should be taken not to compromise the tracking accuracy of the template matching algorithm while reducing the dimension of the optical sections.

### 1.4 Software Architecture for Real-time Tracking of Optical Sections

A DC-Mike actuator M-227.25 attached to the confocal objective can be controlled using a set of vendor supplied control routines from a computer connected to the DC-Motor controller C-862. The frame grabber PCI card DT3155 is added to the computer controlling the confocal objective. In the processor controlling the confocal objective and acquiring confocal volumes, other processing should be kept to minimum to achieve real-time control and volume acquisition. The
acquired volumes may have hundreds of optical sections and therefore tracking individual optical sections requires tremendous computing power. Each of the optical sections in a reference volume can be tracked independently. The task of tracking optical sections can be distributed to various processors in a cluster-computing environment. It should be noted that a prior knowledge about the overall z-axis displacement may help the tracking algorithm to converge faster as described in section 1.3.

1.4.1 Building Blocks of a Multithreaded Realtime Optical Section Tracking System

Figure 8 shows the overall system architecture for real-time tracking in a cluster-computing environment. Preemptive multi-tasking support in the Operating System (OS) allows efficient, multiple threads of execution commonly referred to as threads [25]. The OS uses various scheduling techniques to allocate processor time for each of the active threads. Besides a thread’s regular waiting time planned by the scheduling algorithm, the thread is put to sleep or its processor time is given to another waiting thread when there is a blocking operation. For example, during a data read operation from a disk or through a network connection, the thread will
be inactive and therefore its processor time can be reallocated to other threads running in the system. When multiple threads need access to the same datasets, an effective data sharing mechanism must be adopted to ensure data integrity.

Figure 9: Multithreaded server design for confocal control and image acquisition

In a preemptive multitasking OS such as Windows (95 and above) and Unix, data sharing objects such as critical sections, mutex, semaphores and event objects are available. These objects allow controlled read and write access to the datasets. Available data communication components include sockets, mailslots and named and unnamed pipes. Sockets are software abstraction to allow bi-directional com-
munication between processes. Socket communication can be broadly classified as
stream and datagram modes of communication. The stream mode of socket com-
munication allows reliable data transfer between processes. The datagram mode
of communication does not guarantee a reliable communication. Pipes use sockets
in stream mode for exchanging large datasets between processes. Mailslots are
light-weight communication components that are used for exchanging short mes-
sages between threads. In our multi-threaded system design, we use the following
components.

1. **Sockets** operating in stream mode for reliable data transfer in cluster envi-
   ronment,

2. **Pipes** for streaming the z-Stack datasets to a cluster-facility for tracking,

3. **Mailslots** for message exchange between threads and

4. **Critical sections** for data sharing among multiple threads.

In the overall system architecture shown in figure 8, the processor controlling the
confocal and building the z-Stack will be referred to as *server*. The processor that
coordinates the tracking in the cluster-computing environment will be referred to
as *master client* and the processors in the cluster which share the task of tracking
optical sections will be referred to as *servant clients*.

The reference z-Stack volume and the sensor z-Stack volumes that the mas-
ter client in the cluster receives, should be broadcasted to all the workstations
in the cluster. A z-Stack volume of dimension 640 × 480 × 100 is approximately
30MB and an efficient broadcast mechanism must be used for realtime tracking.
Message Passing Interface (MPI) is the standard way of communication between
the workstations in a cluster [26]. The current MPI standards for the collective
communication operations such as broadcasting (MPI_BCAST) use point-to-point
unicast mode of communication [27]. Significant communication performance
improvement is possible by avoiding the point-to-point communication for collective
communications in the cluster. The broadcast and multicast mode of communica-
tions allow broadcasting a packet to multiple workstations in a networked envi-
nronment. In broadcast mode, a data packet is received by all the computers. A
multicast group is created prior to using a multicast mode of communication. Only
the workstations interested in receiving the group messages, will join the multicast
group. It does not require a membership to send data to a multicast group. At the
end of communication, a workstation can leave the multicast group. Multicast uses
class D IP addresses in the range 224.0.0.0 through 239.255.255.255. A multicast
packet sent to a multicast group with Time-To-Live (TTL) as 1 restricts the data
to be broadcasted only to the local network. When multicast routers are available,
a multicast data packet with TTL greater than 1, will be delivered across networks
[30].
Multicast is built over a connectionless mode of communication called Universal Datagram Protocol (UDP). UDP does not guarantee the delivery of a data packet and is a major drawback of the multicast for group communications in cluster computing. Application level protocols are common in UDP communication mode to ensure reliability. Several methods have been proposed for a reliable multicast communication using application level protocols for use in MPI environment [27, 28, 29]. The reliable multicast protocols maintain a list of all the hosts (ACK or NAK protocols) or a group representative (Ring-based protocol) or a list of representative workstations (Tree-based protocol) to acknowledge the receipt of a data packet or to report a missing packet. In the traditional ACK based protocol, the receivers or the workstations send an ‘ACK’ packet to the sender for every multicast packet received. At the sender, a multicast packet will be removed from the buffer only after receiving ACKs from all the receivers in the group. This requires huge buffer capacity. In this work, we use a variant of the ACK based protocol to ensure multicast reliability. The sender (Master client) keeps a list of all the available receivers (Servant clients) in the group. After sending a multicast packet to the group, the sender requests an acknowledgement (ACK) from all the workstations. The workstations that did not receive the multicast packet, request a retransmission and the sender resends the packet only to the requesting

Figure 10: State transition diagram of the dual-buffer operation

<table>
<thead>
<tr>
<th>State</th>
<th>Primary Buffer Status</th>
<th>Secondary Buffer Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Empty</td>
<td>Empty</td>
</tr>
<tr>
<td>1</td>
<td>Filled by the 'scanVolume' thread</td>
<td>Empty</td>
</tr>
<tr>
<td>2</td>
<td>Filled by the 'scanVolume' thread</td>
<td>Currently Full; streamed by the 'realtimeStream' thread</td>
</tr>
<tr>
<td>3</td>
<td>Currently Full; streamed by the 'realtimeStream' thread</td>
<td>Filled by the 'scanVolume' thread</td>
</tr>
<tr>
<td>4</td>
<td>Empty</td>
<td>Full or being filled by the 'scanVolume' thread</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State Input</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>End of a z-Stack reached</td>
</tr>
<tr>
<td>b</td>
<td>An optical section acquired by the 'scanVolume' thread</td>
</tr>
<tr>
<td>c</td>
<td>An optical section streamed by the 'realtimeStream' thread</td>
</tr>
<tr>
<td>d</td>
<td>z-Stack streaming complete</td>
</tr>
<tr>
<td>e</td>
<td>System reset / Stop tracking</td>
</tr>
</tbody>
</table>
workstation through a reliable (stream) mode of communication.

1.4.1 Server Architecture

Figure 9 shows a multithreaded system design of the server for confocal z-axis control, image acquisition and streaming of the acquired z-Stack datasets to a cluster facility for tracking. The ‘scanVolume’ thread controls the velocity of the actuator control for the WLCM, resets the z-axis position of the confocal objective to the beginning of the volume-of-interest (defined by the user) for subsequent scanning and acquires the optical section images. The acquired optical section images are stored in a reference-buffer or in a dual-buffer. The reference-buffer is used to store a reference volume for tracking. The real-time tracking of a reference volume is theoretically an indefinite process and the ‘scanVolume’ thread will generate a stream of sensor volumes. A recently generated sensor volume is stored in one of the buffers in the dual-buffer. At the end of the current sensor volume, the ‘scanVolume’ thread switches to an available buffer in the dual-buffer for building a new sensor volume. The ‘realtimeStream’ thread is a normal priority thread that sends the reference volume and continuously streams the sensor volumes, as they are acquired, to a cluster facility (master client) for tracking. The dual-buffer facilitates parallel operations of volume acquisition and streaming. Critical sections are used to ensure the integrity of the dual-buffer data. Figure 10 shows the finite-state model of the dual-buffer operations. The state transitions indicate...
the status of the primary and secondary buffers in the dual-buffer during tracking. The state transitions due to the ‘scanVolume’ and ‘realtimeStream’ threads are shown separately for clarity. The ‘realtimeStream’ thread uses named-pipe to stream the sensor volumes to the cluster facility. Mailslots are used to synchronize the parallel threads of execution.

1.4.2 Master-Client Architecture

Figure 12 shows a multithreaded architecture of the master-client that receives the sensor volumes, coordinates the tracking algorithm in the cluster facility and renders the tracked 3D volumes for visualization. The master-client uses a similar set of threads of execution as in the server architecture for building the reference and sensor volumes and broadcasting them to the servant-clients in the cluster facility for tracking. The ‘scanVolume’ thread receives the streamed reference and sensor volumes from the server connected to the confocal microscope. The ‘realtimeStream’ thread broadcasts the confocal volume datasets to the servant-clients through a multicast group.

We use a reliable ACK-based multicast broadcast with stream mode of unicast connections for receiving ACK (or resend request) from the servant-clients in the cluster as explained in section 1.4.1. A standard stream (TCP) mode of unicast connection blocks the thread of execution on data read/write and further there will be 10’s and may be 100’s of processors in the cluster. It will be wasteful of the system resources to span individual threads or groups of threads to block at the data read/write steps. We use an I/O communication port for asynchronous data exchange in the unicast communication mode for reliable multicast. Also the master-client uses the unicast stream communication channel to assign a set of reference optical sections to each of the servant-clients for tracking.

The servant-clients post the tracking results into the ‘msTrkResult’ mailslot available in the master client. The ‘realtimeTracking’ thread builds a z-Stack of tracked optical sections and sends the tracked volume to a real-time volume visualization facility. VolumePro 1000 (TeraRecon, Inc., Concord, MA) provides high quality, real-time rendering capability for PCs. VolumePro uses ray-casting algorithm and a nearest neighbor or a trilinear interpolation algorithm for volume rendering. The highly parallel architectural design of the VolumePro enables real-time visualization of 4D datasets.

1.4.3 Servant-Client Architecture

Figure 11 shows the architecture of the servant-client. Each servant-client joins the multicast group to receive the reference and sensor volumes for tracking. The servant-clients use a stream mode of communication to receive the assigned optical sections for tracking, sending multicast data acknowledgements, sending data
Figure 12: Multithreaded architecture of the master-client for receiving the sensor confocal volumes and coordinating the tracking in a cluster facility.

resend requests and receiving private retransmissions of the reference / sensor vol-
umes. It uses a reference-buffer to store the reference volume and a tracking-buffer to store the sensor volumes. The total tracking time per volume depends on the number of servant-clients available for processing. The servant-client architecture does not require high-end processors. An approximate real-time tracking can be achieved using a few servant-clients. Therefore each servant-client receives subsequent sensor volumes without much delay and does not necessarily require a dual-buffer for parallel volume reception and tracking. To the contrary, a dual-buffer may deter the performance of a servant-client. The servant-client, upon completing the tracking of the optical sections, post the results to the ‘msTrkResult’ mailslot available in the master-client.

1.5 Results

A DC-Mike actuator control (as described in section 1.2) was added to the WLCM at the LIONS Research Laboratories, New Orleans, LA as shown in figure 5 for z-axis control. The DC-Motor controller was calibrated from the actuator design parameters (2048 encoder counts per micron of linear motion). The tracking algorithm with Daubechies wavelet based dimension reduction (section 1.3) and the software architectures for confocal control, confocal volume streaming and cluster architectures (section 1.4) were implemented in Visual C++ 6.0. Visualization Toolkit (VTK) [34] opensource libraries were used for visualizing the optical sections. For volume visualization, the VolumePro 1000 libraries (VLI) with VTK were used. A small cluster of computers with the following specifications was setup at the LIONS Research Laboratories.

1. Server-Confocal control: Pentium III Dual Xeon Processor, 1GB RAM, 10/100/1000 Gigabit Ethernet PCI adapter card,
2. Master-client (cluster): Pentium III Dual Xeon Processor, 1GB RAM, VolumePro 1000 with 2GB onboard video memory, 10/100/1000 Gigabit Ethernet PCI adapter card and

The workstations in the cluster facility were connected through a gigabit ethernet switch.
Table 1: Lamina cribrosa optical section images at 12mmHg tracked using the template matching algorithm presented in section 1.3 after increasing the chamber pressure to 24mmHg and 60mmHg

<table>
<thead>
<tr>
<th>mmHg</th>
<th>Method</th>
<th>Z-Axis Depth in μm</th>
<th>% Match</th>
<th>Time (Sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Visual⁵</td>
<td>0 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340 360</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>NProd²</td>
<td>40 40 60 100 100 100 120 140 120 120 120 120 120 120 120 120 120 120 120 120 120 120</td>
<td>15.79</td>
<td>18.30</td>
</tr>
<tr>
<td></td>
<td>NProd²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corr³</td>
<td>0 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340 360</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FFT⁴</td>
<td>0 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340 360</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L₁ᵇ</td>
<td>100</td>
<td>17.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L₂ᵇ</td>
<td>360</td>
<td>19.00</td>
</tr>
<tr>
<td>60</td>
<td>Visual⁵</td>
<td>110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NProd²</td>
<td>110 110 110 110 110 110 110 110 110 110 110 110 110 110 110 110 110 110 110 110 110 110</td>
<td>5.26</td>
<td>19.95</td>
</tr>
<tr>
<td></td>
<td>Corr³</td>
<td>110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FFT⁴</td>
<td>110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L₁ᵇ</td>
<td>78.95</td>
<td>43.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L₂ᵇ</td>
<td>84.75</td>
<td>16.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ Visual Matching  
² Normalized Innerproduct  
³ Normalized Correlation Measure  
⁴ FFT based Correlation Measure  
⁵ L₁ Norm  
⁶ L₂ Norm
The posterior segment of fresh cow eyes, cut equatorially, is positioned in a custom-built pressure chamber. The pressure in the chamber is monitored using a 12-bit A/D board connected to a pressure transducer. The confocal objective is focussed on to a reference laminar layer (0 microns). Initially the chamber pressure is set to 12 mmHg. Images of the lamina cribrosa were acquired at a regular 20 microns interval for a total depth of 360 microns. Laminar images were acquired after adjusting the chamber pressure to a steady state levels of 24 mmHg and 60 mmHg [17, 18].

In order to validate the tracking performance of the similarity and dissimilarity measures in the ABM method, a trained technician visually matched every image at 12 mmHg with the images at 24 mmHg and 60 mmHg. Table 1 shows the tracking performance of the area based measures in the template matching framework. The FFT based correlation and normalized correlation measures result in a high overall matching rate of 89.5% and were invariant to any illumination changes. The optical section images at 12 mmHg matched with the images at 24 mmHg.
at the same depth, indicating minimal or no laminar deformation. The optical section image acquired at 0 micron (reference layer) at 12 mmHg matched with the image acquired at 110 microns at 60 mmHg, indicating a 110 microns z-axis displacement. The subsequent images at 20 microns interval from 0 micron at 12 mmHg and 110 microns at 60 mmHg matched respectively until a depth of 280 microns at 12 mmHg was reached.

The FFT based template matching (using correlation measure) resulted in a slightly different correlation surface than the normalized correlation measure as shown in figure 13 however, the location of best match remained the same in both the methods. The tracking performance of the FFT and normalized correlation measures remained the same using the 3rd level Daubechies scaling coefficients. Level-3 Daubechies wavelet expansion was found to be optimal to reduce the dimension of the optical section images by 1/64 (i.e. $\frac{1}{12}$ × $\frac{1}{12}$) without compromising the tracking accuracy. Experiments were conducted on random cotton fibers and provided similar tracking results.

The multicast implementation for collective communication in the cluster for broadcasting the z-Stack datasets proved advantageous. On several experimental trials, only 10 to 20 optical sections were retransmitted while broadcasting $\sim$10,000 optical sections to the workstations in the multicast group in the cluster. The cluster used in the study used only 4 workstations and the number of retransmissions is expected to increase with the increase in the number of workstations in the cluster. Figure 14 shows the volume of a random cotton fiber rendered using the VolumePro1000 board.

1.6 Conclusion

We have designed and developed a hardware and software architecture that can convert a Confocal Microscope (2D) into a 4D-Confocal Microscope with the ability to track a reference volume-of-interest in a microscopic specimen. The ABM measures of FFT and normalized correlation measures in the template matching framework were found to be suitable for tracking the optical section images. The Level-3 Daubechies Wavelet expansion provides a significant reduction in the dimension of the optical section images and an increase in the tracking speed without compromising the tracking accuracy. Use of multicast communication mode for broadcasting confocal volume datasets to the workstations in the cluster facility improves the communication performance in a cluster setup. The software architectures allow computationally intensive tracking and similar processing such as finite element analysis to be delegated to a cluster-computing facility. The real-time tracking and visualization architecture developed has potential use as a surgical guidance tool. The 4D-Confocal with tracking facility can be used for: 1) routine laboratory and clinical observation of cellular structures of tissues such as
Figure 14: Volume of a random cotton fiber constructed from optical sections using ray tracing

cornea or skin, *in vivo*, including the precise location of sites of pathology within the tissue volume 2) *in vitro* confocal study of the cellular organization of tissues with the use of various markers for the visualization of the localization of cell organelles or proteins and 3) evaluation of complex materials such as determining the composition of paper with different fiber types. The architectures can be extended to other imaging modalities such as PET, CT, etc., that build 3D volume datasets from optical section images.
Chapter 2. Adaptive Noise Filtering and Restoration of White-light Confocal Microscope Images

2.1 Introduction

The size of the pinholes plays an important role in determining the lateral (x-y) and axial (z-axis) resolution of the optical section images acquired using a confocal microscope. The light from the out-of-focus planes cannot be completely blocked due to the limitation on the size of the pinholes. The aberrant light rays from the adjacent points on the focal plane and from the out-of-focus planes affect the contrast of the image observed in the image plane. A larger pinhole size permits the light from the points in the focal plane adjacent to the point of interest to reach the image plane. Also, a really small pinhole size will significantly reduce the light throughput of the scanning system thereby requiring a very powerful light source to achieve a good quality image.

A white-light confocal microscope (WLCM) allows imaging tissues in vivo. The use of a white-light source reduces the speckle noise in the observed images mainly due to the interaction of various wavelengths present in the white light. The characteristics of the WLCM, especially for accounting the light from adjacent and out-of-focus planes, can be described using an impulse response of the imaging device. Assuming that the impulse response $h$ at an imaging point of the WLCM is spatially invariant, an image acquired at the image plane can be represented as

$$ g = h * f $$ \hspace{1cm} (12)

where, $g$ is the image observed at the image plane, $f$ is the true object image at the object plane, $*$ is the convolution operator and $h$ is the impulse response of the WLCM. If the pinhole sizes are optimal, then the impulse response will resemble a dirac delta function. However due to non-ideal lens and pinhole size, each image point (pixel) will receive photic contributions from adjacent pixel positions and from out-of-focus planes. Therefore, the observed image in equation (12) is usually blurred. A non-ideal impulse response of a confocal microscope is usually referred to as a smearing or point spread function [35]. With random additive noise $n$ introduced by the imaging system, the imaging model in equation (12) becomes

$$ g = f * h + n $$ \hspace{1cm} (13)
Image restoration algorithms estimate the original image \( f \) from the observed image \( g \) with the knowledge of the impulse response \( h \) of the imaging instrument \[67, 37, 36\]. It should be noted that the image restoration algorithms, as opposed to image enhancement algorithms, do not aim to improve the visual score of a human visual system, but rather attempt to compensate or remove the error introduced due to non-ideal impulse response \[67\]. The image restoration algorithms can be broadly classified as:

1. Direct methods,
2. Iterative methods and

Since image restoration is reverse of the blurring or convolution process, the image restoration algorithms are commonly referred to as image deconvolution or image deblurring algorithms \[67\].

The direct and iterative methods derive an estimate \( \hat{f} \) of the original image \( f \) using the knowledge of the impulse response \( h \) of the system and the image acquired at the image plane \( g \) \[38\]. Blind deconvolution algorithms attempts to derive an estimate of the impulse response \( \hat{h} \) and the original image \( \hat{f} \) \[39\]. The direct methods compute an estimate of the original image using direct matrix computations in the frequency domain \[40\]. These include inverse restoration method, least squares method and Wiener filters \[37, 67\]. The direct methods are faster but the solution computed is sensitive to the noise present in the acquired image \( g \) and are usually ill-conditioned. Iterative regularization techniques incorporate the knowledge about the imaging system noise and / or the true solution \( f \) as constraints in deriving a physically acceptable solution; an example of a physically unacceptable solution include negative pixel intensities in the restored image. A regularization technique minimizes a stabilizing functional \( \| c \hat{f} \| \) subject to a bounding error \( \| g - H \hat{f} \| \leq \epsilon \), where \( c \) is a regularization operator. The challenge with regularization method lies in choosing an appropriate stabilizing functional / regularization operator \( c \) to yield a better estimate of the original image. The projection onto convex sets (POCS) method defines a closed convex set of constraints \( \{ C_i \}_{i=1}^m \) to be obeyed while restoring the images. If \( P_i \) denote the projection onto the \( i^{th} \) constraint then the iteration

\[
\hat{f}_k = (P_1, P_2, \ldots, P_m)^k \hat{f}_0
\]

will find a solution or a point \( \hat{f}_k \) satisfying all the constraints. However the performance of the constrained optimization techniques depend on the constraints used for the restoration. An object oriented based MatLab toolbox is available for iterative restoration of images \[41\].
Assuming that the noise in the observed image \( g \) is insignificant and can be ignored, the image model in equation (12) in the frequency domain becomes

\[ G = HF \quad (14) \]

where, \( G, F \) and \( H \) represents the frequency domain representations of the observed image \( g \), original image \( f \) and the impulse response \( h \) of the imaging system respectively. \( H \) is most commonly referred to as the Optical Transfer Function (OTF) of the microscope. Using a simple inverse restoration operation, the original image can be restored as

\[ \hat{F} = H^{-1} G \quad (15) \]

It can be observed that the confidence on the restored image relies heavily on the OTF. If the OTF matrix \( H \) is singular i.e. at least one of the eigenvalues of \( H \) is zero, then the inverse filter solution does not exist and the problem is ill-posed. Also, a near singular \( H \) makes the restoration ill-conditioned. Even a small perturbation to the image model in equation (14) with an additive noise will make the inverse deconvolution unstable.

The performance of the image restoration algorithms, more specifically that of the inverse restoration algorithm, can be improved by pre-filtering the image noise \[42\]. If multiple WLCM images of a specimen at the same z-axis depth are acquired at the same instant, then the images typically will represent the same scene. The only variations among the images in the ensemble per z-axis depth at a given instant will be due to the random noise introduced by the system. In this chapter, we present a novel adaptive noise filtering algorithm using Karhunen-Loève (KL) expansion for use in confocal microscopes. We use the KL transform to decorrelate an ensemble of WLCM images acquired per z-axis position to filter the random variations in the pixel intensities due to the scanning system and additive noise. Deconvolution is a deblurring process and the restored images typically exhibit sharp edges with an overall increase in the high frequency contents of the image. Therefore we use a gradient based contrast metric to measure the quality of the restored images. We will compare the performance of the proposed adaptive KL filter with that of a median filter.

Iterative deconvolution algorithms derive an estimate of \( f \) from \( g \) in iterative steps. A usual stopping criteria for an iterative deconvolution is an error functional of the form \( \min \| f_{k+1} - \hat{f}_k \| \) where, \( \hat{f}_k \) is the estimate of the original image at the \( k^{th} \) iteration. We observed that the images restored using iterative deconvolution algorithms such as Lucy-Richardson restoration algorithm, exhibit prominent oscillations at higher number of iterations around the edges and at pixel locations with sharp intensity transitions. These spurious oscillations around the edges are known as ringing artifacts. Thus an error functional based on a difference measure cannot accurately represent the amount of ringing introduced in the images during restoration. These ringing artifacts can be easily identified by a human.
The objective quality metrics such as peak-signal-to-noise ratio (PSNR) and mean squared error (MSE) metrics do not correlate with the quality assessed by a human visual system [51]. The assigned image quality metric should confirm with the human visual system to quantify the amount of ringing introduced during restoration. Hence assigning an objective image quality metric is a very difficult problem [52]. Image quality assessment can be broadly classified as no-reference methods and full-reference methods. The full-reference methods compute a quality metric of a given image using a reference or undistorted-original image. No-reference methods do not require a reference image in assigning a quality metric. Recently several image quality metrics have been proposed to measure the blur and blocking artifacts introduced by image/video compression algorithms [53, 54, 55]. A frequency based ringing metric will have difficulties in differentiating the ringing artifacts from the image features. Spatial analysis techniques as in [54] require extensive row-by-row processing and are not suitable for real-time image quality measurement. We have developed a novel ringing metric using simple binary morphological operations to measure the amount of ringing artifacts present in an image.

Section 2.2 contains a brief overview of the KL expansion. In section 2.3 we present the adaptive noise filtering and inverse restoration algorithm using KL expansion. In section 2.4 we describe a sobel operator based contrast measure for evaluating the performance of the WLCM image restoration. We present our ringing metric for use with iterative deconvolution algorithm in section 2.5. In section 2.6 we present and discuss the results of 1) WLCM image restoration using adaptive KL noise prefiltering and inverse restoration algorithm and compare the results with median prefiltering algorithm and 2) the performance of the ringing metric in estimating the ringing artifacts introduced in the WLCM images restored using Lucy-Richardson deconvolution algorithm. We conclude this chapter in section 2.7.

### 2.2 Karhunen-Loève Expansion

Karhunen-Loève (KL) expansion uses an optimal set of orthogonal basis vectors that span an entire ensemble of signal [45]. It is also referred as Proper Orthogonal Decomposition (POD), Hotelling Transform and Principal Component Analysis (PCA) [44]. KL expansion decorrelates a given ensemble of signal by discovering an orthogonal basis set that is optimal for the signal under consideration as opposed to analyzing the signal using an off-the-shelf wavelet basis or Fourier transform. It has been used for characterizing an ensemble of human faces with few optimal
image basis called eigen-pictures thereby reducing the number of coefficients in the 
KL expansion for each image [46].

Using KL expansion, a continuous second-order process \( x \in \mathbb{R}^{M \times N} \) with a 
covariance \( R(x, x') \) can be expanded as

\[
x = \sum \alpha_n \phi_n(x)
\]

with \( E(\alpha_i, \alpha_j) = \sqrt{\lambda_i} \delta_{ij} \) and \( \langle \phi_i, \phi_j \rangle = \delta_{ij} \). Here \( E \) is the expectation, \( \{ \lambda_i \} \) are the 
eigenvalues and \( \{ \phi_i \} \) are the eigenvectors of the covariance matrix \( R \) of \( \{ x_i \} \). In 
discrete case,

\[
R \Phi = \Lambda \Phi \quad (17)
\]

The existence of \( \Lambda = \{ \lambda_i \} \) and \( \Phi = \{ \phi_i \} \) in equation (17) is guaranteed by the 
Mercer’s theorem [47] analogous to the spectral decomposition of symmetric matrices [48]. For a given ensemble \( \{ x_i \}_{i=1}^L \), KL expansion guarantees the best k-term 
orthogonal expansion among all the orthogonal transforms [49]. The eigenvectors 
of the covariance matrix \( R \) forms an optimal orthogonal basis for the ensemble 
\( \{ x_i \}_{i=1}^L \). The orthogonal basis vector \( \phi_i \) corresponding to the largest eigenvalue, 
is in the principal direction of the ensemble \( \{ x_i \}_{i=1}^L \). When the elements in the 
ensemble \( \{ x_i \}_{i=1}^L \) are correlated, for example the ensemble formed using multiple 
snapshots of images of a given scene, the principal basis vector alone can effectively 
represent the entire ensemble with the least expansion error when compared to the 
other available orthogonal expansions. This property is desirable in pattern recog-
nition applications. The energy retained during a k-term expansion is given in 
terms of the eigenvalues of \( R \) as follows.

\[
\sum_{i=1}^k \frac{\lambda_i}{\sum_{j=1}^{MN}} \times 100\% \quad (18)
\]

The eigenvectors and eigenvalues of \( R \) can be computed using a direct computa-
tion or using a reduced computation technique called the method of snapshots [50].

Let, \( \{ x_i \}_{i=1}^L \) be the expectation-centered ensemble set, where \( x_i \in \mathbb{R}^{M \times N} \) and 
\( E(x) = 0 \). Let, each of the column formatted ensemble element \( x_i \) form the \( i^{th} \) 
column of the ensemble matrix \( X \). For example, given an ensemble \( \{ x_i \}_{i=1}^L \) of gray 
scale images of size 256 \( \times \) 256, each of the column formatted images \( x_i \in \mathbb{I}^{65536 \times 1} \) 
will form the \( i^{th} \) column of the ensemble matrix \( X \in \mathbb{I}^{65536 \times L} \). Here, \( \mathbb{I} \) represents 
the space of gray scale image vectors and \( L \) is the number of elements in the 
ensemble. Now, the ensemble covariance matrix \( R = XX^T \), where \( R \in \mathbb{R}^{MN \times MN} \).

2.2.1 Direct Method to Compute the Optimal Basis

The eigenvectors \( \{ \phi_i \}_{i=1}^{MN} \) of the ensemble covariance \( R \) in equation (17) that form 
the orthogonal basis for the ensemble can be derived using a singular value decom-

position (SVD) of the ensemble matrix $X$ as follows.

$$X = U \Lambda V$$

Covariance $R = XX^T$

$$= U \Lambda V V^T \Lambda^T U^T$$

$$= U \Lambda^2 U^T$$

The covariance $XX^T$ is a symmetric matrix and therefore an eigendecomposition as in equation (20) exists. The left singular matrix $U$ forms the basis of the column space of the ensemble matrix $X$ as in equation (19) and therefore the ensemble elements $\{x_i\}_{i=1}^L$ which form the column space of $X$ have an expansion in $U$. Solving the eigenvalue problem in equation (20) requires solving an $MN \times MN$ system. In the above example of gray scale image ensemble, this would require solving a $65536 \times 65536$ system. Although computational resources are available for solving such a massive eigenvalue problem, it is unnecessary for the problem under consideration.

2.2.2 Method of Snapshot

The dimension of the optimal orthogonal basis needed to describe the ensemble is $L$, where $L$ is the ensemble size. The ensemble matrix $X$ is singular and does not require a full dimension to describe the ensemble elements. Thus a reduction in the basis computation can be achieved using a reduced SVD approach as follows.

$$X^T X = V^T \Lambda U^T U \Lambda V$$

$$= V^T \Lambda^2 V$$

The right singular vectors $V$ and the eigenvalues $\Lambda^2$ can be computed from equation (21). This requires solving a $L \times L$ system. Further, the left singular vectors $U$ can be computed from equation (19) using $X$, $V$ and $\Lambda$. The left singular vectors $U$ span the column space of the ensemble matrix $X$ and thus form the basis of the ensemble $\{x_i\}_{i=1}^L$. The reduced SVD approach for determining an optimal ensemble basis [45] provides a significant reduction in computation, when $L \ll MN$.

2.3 Adaptive Noise Filtering and Restoration of WLCM Images

The WLCM optical section images are digitized using a frame grabber (DT 3155) PCI card that allows a maximum frame acquisition rate of 30 frames per second. As mentioned in chapter 1, the frame acquisition rate will be limited by the scanning rate of the confocal microscope. To adaptively filter noises in the acquired images, we propose to use 3 frames of images acquired per z-axis position. Each of the images is converted to a column formatted image $X_i \in \mathbb{F}^{MN \times 1}$. An ensemble
matrix $X \in \mathbb{I}^{MN \times 3}$ is constructed using the column-formatted images. The vector space $\mathbb{I}^{MN \times 1}$ represent the space of grayscale WLCM images. From the KL expansion using the method of snapshot, the dimension of the ensemble space can be reduced to three. An optimal basis for this ensemble matrix $X$ can be obtained by computing the eigenvectors of the covariance matrix $XX^T \in \mathbb{R}^{3 \times 3}$ and the left singular matrix $U \in \mathbb{R}^{MN \times 3}$ of $X$ as described in section 2.2.2. Now the column formatted ensemble images can be expressed using the optimal basis as follows.

$$X_i = \sum_{j=1}^{3} \langle X_i, U_j \rangle U_j$$

(22)

Since the ensemble matrix is formed using the images of the same scene, we found that only the first principal left singular vector of $X$, denoted $U_{max}$, is sufficient to retain more than 90% of the energy. Note that the bases $\{U_1, U_2, U_3\}$ are arranged in the order of their contribution in the ensemble space. The first principal component $U_1$ points in the direction of the maximal ensemble variance and therefore typically represents the features present in the image ensemble. The second and third principal components are available to depict the differences between the three images in the ensemble and contain less than 10% of the total ensemble energy. Since the ensemble was formed using the images of the specimen at the same z-axis depth, the coefficients of the second and third principal components can be dropped to eliminate the variations observed between the images in the ensemble. These coefficients typically represent the noise and random pixel variations due to the scan lines introduced by the Nipkow scanning disk. Thus any image from the ensemble reconstructed using only the first principal component will be a noise filtered image of the specimen at the given z-axis depth given by

$$X_{Filtered} = \langle X_i, U_1 \rangle U_1$$

(23)

where, $X_i$ in equation (23) can be any image from the ensemble. Now the filtered image can be restored using the inverse restoration algorithm. Algorithm 2 in page 35 summarizes the automatic noise filtering and the restoration steps.

## 2.4 Performance Evaluation of the Restored Images Using a Contrast Measure

If the impulse response of the WLCM were ideal, $h(x, y) = \delta(x, y)$. However a non-ideal lens and pinhole dimension results in a point spread function (PSF) that acts as a low pass filter resulting in a blurred image. Image restoration process restores an image using a prior knowledge about the PSF of the imaging device. The blur in the restored images is expected to decrease and cause an eventual increase in the high frequency details in the image. Autofocusing in widefield microscopy requires a similar focus measure to identify an optimal focusing point with less blur [57, 58, 59, 60, 61]. The common choices to measure the blur in the images are:
Algorithm 2 Algorithm for adaptive KL prefiltering and inverse restoration

1: **procedure** KL Prefilter-Inverse Restoration
2: Acquire image ensemble \( \{x_i\}_{i=1}^3 \) at a given z-axis position
3: Form column-formatted image matrices \( \{X_i\}_{i=1}^3 \) and ensemble matrix \( X \)
4: Compute the covariance matrix \( R = XX^T \)
5: Compute the optimal ensemble basis \( \{U_i\}_{i=1}^3 \) using a reduced SVD as in section 2.2.2.
6: Filter the noise and random components in the image ensemble by dropping the second and third principal components. Determine the noise filtered image as in equation (23).
7: Using the optical transfer function \( H \) and the noise prefiltered image \( X_{\text{Filtered}} \), restore image as \( X_{\text{Restored}} = H^{-1}X_{\text{Filtered}} \) where, \( X \) represents the FFT of \( X \).
8: Energy retained during the adaptive KL prefiltering = \( \lambda_{\text{max}} / \sum_{i=1}^3 \lambda_i \times 100\% \)
9: **end procedure**

1. Image pixel intensity variance
2. \( l_1 \) norm of the image gradient \( \nabla I \) (first derivative)
3. \( l_2 \) norm of the image gradient \( \nabla I \) (first derivative)
4. \( l_1 \) norm of the image Laplacian \( \nabla^2 I \) (second derivative)
5. \( l_2 \) norm of the image Laplacian \( \nabla^2 I \) (second derivative)

We use an image contrast measure of \( l_2 \) norm of the image gradient to estimate the significance of the adaptive KL filter presented in section 2.3 especially for improving the performance of inverse restoration of WLCM images. Image gradient is computed as

\[
\nabla I = \begin{pmatrix} G_x \\ G_y \end{pmatrix} \\
= \begin{pmatrix} \partial I/\partial x \\ \partial I/\partial y \end{pmatrix}
\]

Now the magnitude of the image gradient at all the pixel locations can be computed as

\[
\nabla I = \text{magnitude}(\nabla I) = \sqrt{G_x^2 + G_y^2}
\]

(24)
where the $G_x$ and $G_y$ are computed by convolving the image $I$ with the sobel operator kernels

$$S_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad S_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

respectively. Using the image gradient in equation (24), contrast of an image $I \in \mathbb{R}^{M \times N}$ can be computed as follows.

Contrast measure = $\sum_{i=1}^{M} \sum_{j=1}^{N} (\nabla I(i,j) - \nabla I)^2$ (25)

where, $\nabla I$ is the mean image gradient computed as

$$\nabla I = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \nabla I(i,j)$$

### 2.5 A Ringing Metric for Use with Iterative Deconvolution Algorithms

We attempted to quantify the amount of deblurring achieved in each iterative step using the contrast measure defined in equation (25) in an iterative algorithm such as Lucy-Richardson (LR) algorithm [67] defined as

$$f_{k+1}(x,y) = \hat{f}_k(x,y) \left( \frac{h(-x,-y) * g(x,y)}{h(x,y) * \hat{f}_k(x,y)} \right)$$ (26)

We observed that the contrast measure indicated an increasing trend during LR iterative restoration steps. We expected the contrast measure to converge to a maximum contrast value. But the contrast measure was exponentially increasing at each iterative step. Upon careful evaluation of the restored images at various iterative steps, we found that the LR restoration introduced oscillations around edges and at pixel locations with sharp intensity transition at higher number of iterations. The oscillations artifacts are commonly referred to as ringing. Figure [15] shows the contrast plot of an image of a random cotton fibers at various iterative steps of LR restoration algorithm.

The ringing artifacts arise primarily due to the inverse of the PSF involved during restoration. The PSF $h$ in the imaging model $g = h * f$ integrates pixel
Figure 15: Images of random cotton fibers restored using Lucy-Richardson restoration algorithm; contrast measure increases exponentially due to ringing artifacts.

intensities from adjacent locations to cause blurring in the acquired image $g$. The restoration process computes pixel gradients using the inverse of the optical transfer function. Therefore, the locations in the image with sharp intensity transition exhibit pixel intensity overshoot (extreme positive values) or undershoot (extreme negative values) [37]. The ringing artifacts appear as oscillations around the edges or locations with sharp intensity transitions. This led us to develop a ringing metric to quantify the amount of ringing introduced during iterative steps to assist identifying a terminal point of an iterative restoration algorithm.

The proposed ringing metric uses the edge profiles of the observed image $g$ and
the edge profiles of the images restored \( \hat{f}_k \) to isolate the ringing artifacts introduced during restoration. The area surrounding the edge profiles of the original image are the regions to be observed for ringing artifacts. The edge profile \( E_{\text{ref}}(g) \) of the original image \( g \) can be extracted using canny edge detector \[67\]. From the edge profile \( E_{\text{ref}}(g) \), the ringing region around the edges can be defined by dilating \( E_{\text{ref}}(g) \) with a \((r \times r)\) structuring element \((SE)\), where \( r \) is the approximate width in number of pixels from an edge to cover the ringing artifacts. The dilation operation is defined as follows.

\[
E_{\text{RingRegion}}(g) = E_{\text{ref}}(g) \oplus SE = (E_{\text{ref}}^c(g) \ominus \bar{SE})^c
\]  

(27)

where, \( \oplus \) is the binary dilation and \( \ominus \) is the binary erosion defined as

\[
E_{\text{ref}}^c(g) \ominus SE = \{x : SE_x \subset E_{\text{ref}}\}
\]

\(E_{\text{ref}}^c(g)\) is the binary complement of \(E_{\text{ref}}(g)\), \(SE\) is the structuring element, \( \bar{SE} \) is the rotation of the structural element by \(180^\circ\) and \(SE_x\) is the translation of \(SE\) by \(x\) \[62\]. At the end of each iterative step, the edge profile \(E(\hat{f}_k)\) of the restored image \( \hat{f}_k \) can be determined. Now the edges and any ringing artifacts around the original edges \(E_{\text{ref}}(g)\) can be selected from \(E(\hat{f}_k)\) by a simple pixelwise logical AND operation between \(E_{\text{RingRegion}}(g)\) and \(E(\hat{f}_k)\). Since restoration preserves the edges present in the original image \(g\), the additional edges observed around the reference edge profile in the restored image typically represent the ringing artifacts. The ringing metric is defined as

\[
\text{Ringing metric} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} (E_{\text{RingRegion}}(g) \text{ AND } E(\hat{f}_k)) - \sum_{i=1}^{M} \sum_{j=1}^{N} E_{\text{ref}}(g)}{\sum_{i=1}^{M} \sum_{j=1}^{N} E_{\text{ref}}(g)}
\]  

(28)

2.6 Results

The PSF of the WLCM was experimentally determined by imaging 5-micron diameter micro-spheres under the usual imaging conditions. Several frames of micro-sphere images at the same focus were averaged to reduce the noise sensitivity of the PSF. A single micro-sphere is isolated and cropped from the image and was used to determine the PSF of the WLCM. The background of the cropped micro-sphere image is kept to a minimum to avoid oscillations in the places of sharp intensity transitions in the restored images. Figure 16a shows the image of a micro-sphere and figure 16b shows the PSF of the WLCM. Figures 17a, 17b, 17c show the image ensemble of a random cotton fiber at the same z-axis depth. The KL basis computed using reduced SVD method is shown in figures 17d, 17e and 17f. It is clear from the optimal ensemble basis that the first principal vector retains most of the energy in the ensemble (95.5%). Figure 17g shows the inverse restored image after
median prefiltering and figure 17h shows the inverse restored image after adaptive KL prefiltering. Restoration results on similar cotton fibers and images of lamina cribrosa of cow retina at 12mmHg of IOP level are shown in figures 18 and 19. A contrast measure plot in figure 20 shows the contrast measures of the restored images after median and adaptive KL prefiltering. The adaptive KL prefiltering outperforms the median prefiltering in all the experimental cases. Also, a visual evaluation of the image restored after adaptive KL prefiltering shows a significant improvement over the median prefiltered images.

Figure 21b shows an image of random cotton fibers superimposed with its edge profile. Figure 21d shows the edge profile $E(\hat{f}_6)$ of a restored image at iteration # 6 of LR deconvolution algorithm superimposed with the reference edge profile $E_{\text{ref}}(g)$ and a binary edge mask $E_{\text{RingRegion}}(g)$ for selecting the region around $E_{\text{ref}}(g)$. Ringing artifacts can be observed around the edges in the restored image $\hat{f}_6$ in figure 21c. Visual inspection of the images restored at higher iterative steps of LR deconvolution algorithm confirms with the increasing trend of the ringing metric plot shown in figure 21e.

2.7 Conclusion

We have presented an adaptive noise filtering technique by discovering an optimal basis for 3-image-frames ensemble acquired per z-axis position. The noise and random components are filtered by dropping the second and third principal components of the basis. A reduced SVD or snapshot method of determining the basis makes this technique suitable for real-time restoration of WLCM images. The presented automatic noise filtering algorithm significantly reduces the random image artifacts and noise in the WLCM images. The automatic noise filtering algorithm proposed here does not require a prior knowledge about the type of noise present in the system and hence adaptive to the images being restored. The proposed adaptive noise filtering algorithm can be used in prefiltering stage of image deconvolution algorithms to improve the convergence of the iterative restoration.
Figure 17: Ensemble of random cotton fiber images, their KL decomposition and the inverse restored images using median and adaptive KL prefiltering algorithms and to improve the quality of the restored images using a simple inverse restoration method.
Figure 18: Random cotton fibers (a, d, g) restored using inverse filter deconvolution after prefiltering noise using median (b, e, h) and adaptive KL filters (c, f, i).

A near-singular nature of the point spread function introduces ringing artifacts during image restoration, especially at higher iterations of iterative deconvolution algorithms. Hence a robust feedback metric is essential to identify an optimal terminal condition of the iterative deconvolution algorithms. The ringing metric proposed here is less sensitive to noise amplification during restoration and is computationally efficient making it suitable for real-time image quality evaluation.
Figure 19: Image of a random cotton fiber (a) and lamina cribrosa optical section at 12mmHg IOP level (d) restored using inverse filter deconvolution after prefiltering the image noise using median and adaptive KL filters.

Visual inspection of the images restored using LR deconvolution algorithm confirms with the increase in the ringing artifacts during restoration as indicated by the ringing metric.
Figure 20: Contrast measure performance comparison of median and adaptive KL prefilters with the inverse restoration method.
Figure 21: Computing the ringing metric of an image restored using LR iterative restoration algorithm

3.1 Introduction

Image synthesis is one of the important steps for automatic visual analysis of a scene. Determining the motion field of a scene or an object in a scene is an important problem in the field of Computer Vision. The marvelous synchrony of the eye and the brain of insects, primates and human beings for visual interpretation of objects, movements and their transformations have been an inspiration for the pursuit of computer vision systems [63]. The characteristics of a moving object including its 3D structure can be partially described and reconstructed from an estimate of the motion field [64]. Figure 22 shows a sequence of images of a scene containing a clock. The hour and minutes hand in the clock moved a bit between the images. These movements between the two instances of the scene can be described by a motion field. Besides the method of physically measuring the motion fields, the non-contacting method of motion field estimation using digital images is attractive and has gained significant attention.

Figure 22: An image sequence showing the changes in a scene in a short interval $\delta t$

Digital images of a scene or an object under observation when imaged at a reference state and a transformed state form the basis of the non-contacting method
of motion field estimation. The pattern of pixel brightness changes in a scene observed at an interval is usually referred to as optical flow. Under the same illumination conditions, the image pixel brightness changes observed between any two images in an image sequence of a scene are either due to object transformation, scene changes or camera movements. Assuming that the object transformations \((\delta x, \delta y)\) does not change the brightness of the object \(E\),

\[
\frac{dE}{dt} = 0
\]

i.e. \(E(x,y,t) - E(x+\delta x,y+\delta y,t+\delta t) = 0\) (29)

Now the Taylor series expansion of \(E(x+\delta x,y+\delta y,t+\delta t)\),

\[
E(x+\delta x,y+\delta y,t+\delta t) = E(x,y,t) + \delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} + \cdots
\]

(30)

Omitting the higher order terms in equation (30) and substituting in equation (29),

\[
\delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} = 0
\]

Dividing by \(\delta t\) and applying \(\lim_{t \to 0}\),

\[
\lim_{t \to 0} \frac{\delta x \frac{\partial E}{\partial x}}{\delta t} + \frac{\delta y \frac{\partial E}{\partial y}}{\delta t} + \frac{\partial E}{\partial t} = 0
\]

\[
\frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0
\]

\[
E_x u + E_y v + E_t = 0
\]

\[
(\nabla E)^T \tilde{U} + E_t = 0
\]

(31)

Equation (31) is referred to as Image Brightness Constancy Equation (IBCE) and was introduced by Horn & Schunck in 1981 [65]. Here, \(\tilde{U} = (u, v)\) form the flow velocity or the pixel velocity at \((x,y)\), \(E_x\) and \(E_y\) are the spatial brightness gradients in the \(x\) and \(y\) directions respectively and \(E_t\) is the temporal intensity gradient.

The IBCE in equation (31) uses the spatial brightness gradient \(\nabla E\) in determining the optical flow. Therefore only the components of the optical flow in the direction of the brightness gradient \(\nabla E\) can be estimated from the IBCE. The algorithms that use the IBCE for estimating the optical flow typically use a narrow support of region around each pixel position \((x,y)\) in determining the image gradients. Therefore they have trouble estimating the optical flow components in the direction normal to the image gradients. Also, the size of the scene observation window has effects on the estimated optical flow components. For example, consider a rectangular object moving in a scene as shown in figure [23]. The observation window \((W)\) plays a prominent role in determining the true displacement vector \(\tilde{U}\). If the window \(W\) does not cover the prominent features of the object under
transformation as in figure 23, only the component \( u \) of \( \vec{U} \) in the direction of the brightness gradient, i.e. \( E_x \), can be determined from the IBCE in equation (31). This is the classical aperture problem in motion field estimation [64].

In section 3.2, we provide a review of the existing optical flow techniques and a detailed background on the digital image correlation based coarse-fine search method of optical flow estimation. We present a distance based error metric for the evaluation of optical flow techniques without the knowledge of a ground truth of the optical flow field \( \vec{U}_{GT} \) in section 3.3. In section 3.4 we describe a Compute-and-Propagate (CNP) framework for computing a sparse flow field and to supply an initial estimate of flow velocities for the iterative optical flow algorithms. Further we demonstrate the integration of the CNP framework with the Digital Image Correlation (DIC) Coarse-Fine Optical Flow estimation technique [71]. We present the results of the CNP framework with DIC Coarse-Fine optical flow estimation technique on: 1) A rotating sphere 2) Cotton fiber images acquired using a white-light confocal microscope 3) A rotating globe 4) A rotating globe with small displacement 5) A clock sequence with movement in hour and minute hands and 6) Images of lamina cribrosa of retina acquired before and after an experimental glaucomatous condition in section 3.5. We conclude this chapter in section 3.6.

### 3.2 Background

Barron, et. al [78, 79, 80] have conducted an extensive survey and performance comparison of the most popular optical flow algorithms. The optical flow techniques can be broadly classified as:

1. Differential technique,
2. Phase based technique,
3. Energy based technique and
4. Correspondence based technique.

The differential techniques use the IBCE in equation (31) with additional constraints in computing the optical flow [65, 81, 82, 83]. A least-squares method of solving the IBCE is presented in [64]. The pioneering Horn & Schunck algorithm [65] uses a smoothness constraints in the scene in solving the IBCE. The differential techniques use the spatial brightness gradients $\nabla E$ and the temporal brightness changes in estimating the optical flow. Due to the narrow support involved in calculating the spatial and temporal brightness gradients, they are suitable for estimating smaller displacements, usually in the sub-pixel range involving small transformations. Therefore, the images of the scene need to be acquired at a very high frame rate.

The Energy based methods estimate the optical flow from a set of velocity-tuned filters tuned to different spatial orientations and temporal frequencies [84]. The phase based methods derive an estimate of the optical flow from a family of spatiotemporal velocity-tuned linear filter [85]. Both the energy-based and phase-based techniques require a large temporal support and require a minimum of 21 image frames in the sequence for an optimal performance [79].

The correspondence based techniques use a measure-of-match (similarity) or a measure-of-mismatch (dissimilarity) to establish correspondence between regions in the image sequences. The correspondence based optical flow techniques can be further subclassified as:

a. Area based matching,

b. Feature based matching and

c. Relational matching.

The Area Based Matching (ABM) methods use gray scale windows, known as templates, from an image $I_{t_0} \in \mathbb{I}^{M \times M}$ of a scene at an initial condition, to search and locate an appropriate matching location in the image $I_{t_0+\delta t} \in \mathbb{I}^{M \times M}$ of the same scene after a time interval $\delta t$ [68], where $\mathbb{I}^{M \times M}$ represents the space of gray scale images of size $M \times M$. The choice of the template window size is crucial for including essential structures of the object or structures from the scene to obtain an unique displacement vector $\vec{U}$ resulting from the object or scene transformation. The Feature Based Matching methods use features such as lines, edges and points extracted from $I_{t_0}$ and $I_{t_0+\delta t}$ for establishing a regional correspondence. Kalman filter is one of the most popular FBM method [66]. The Relational matching methods use a relationship between the features in $I_{t_0}$ to identify the changes in $I_{t_0+\delta t}$. In this work, we use an ABM method for demonstrating the performance of the
The commonly used correspondence measures in the ABM methods are [68, 69]:

1. Measure of absolute difference (Dissimilarity measure)

\[
\text{Norm}_{L1}(I_{t0}, I_{t0+\delta t}) = \sum_{i=1}^{M} \sum_{j=1}^{M} |I_{t0}(i, j) - I_{t0+\delta t}(i, j)|
\]  

(32)

2. Euclidean distance (Dissimilarity measure)

\[
\text{Norm}_{L2}(I_{t0}, I_{t0+\delta t}) = \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{M} |I_{t0}(i, j) - I_{t0+\delta t}(i, j)|^2}
\]  

(33)

3. Normalized correlation measure (Similarity measure)

\[
\frac{\sum_{i=1}^{M} \sum_{j=1}^{M} (I_{t0}(i, j) - \bar{I}_{t0})(I_{t0+\delta t}(i, j) - \bar{I}_{t0+\delta t})}{\sqrt{\sum_{i=1}^{M} \sum_{j=1}^{M} (I_{t0}(i, j) - \bar{I}_{t0})^2 \sum_{i=1}^{M} \sum_{j=1}^{M} (I_{t0+\delta t}(i, j) - \bar{I}_{t0+\delta t})^2}}
\]  

(34)

The digital image correlation (DIC) based optical flow estimation techniques [71, 72, 73, 74, 75, 76, 77, 86] use a measure of similarity as in equation (34). For an optimal match, algorithms using a dissimilarity measure need to minimize the measure-of-mismatch and the algorithms using a similarity measure need to maximize the measure-of-match. The normalized correlation measure in equation (34) is robust to changes in the lighting condition.

To determine the unique displacement vectors, the DIC based optical flow algorithms assume that the transformation or deformation in an infinitesimal region is homogeneous. An unique displacement vector is identified by comparing a small region around each pixel \((x, y)\) in \(I_{t0}\) with a set of regions centered at \((x \pm k\delta x, y \pm k\delta y)\) in the image \(I_{t0+\delta t}\). The incremental values \((u = n\delta x, v = m\delta y)\) yielding a maximum similarity measure will be the optical flow in the corresponding region. As it can be seen, the uniqueness of the solution depends on various factors including 1) the choice of the window size and 2) the search area. After an initial estimate of the transformation, DIC based optical flow algorithm proceed to a fine search step where any interpixel displacements are determined using a fine searching or an iterative algorithm such as Newton-Raphson method [74]. Figure 24 depicts the DIC search process. An infinitesimal region centered at pixel \((x, y)\) in \(I_{t0}\) is searched in the image \(I_{t0+\delta t}\) over a small region of size \(2k\delta x \times 2k\delta y\) around the pixel \((x, y)\). During the coarse-fine search steps, the search boundaries are progressively adjusted based on the convergence of the search [77].
The differential, energy-based and phase-based optical flow methods require an image sequence with higher temporal support, typically in the range of 10 to 21 frames of the scene, for flow field computation [79]. The optical flow algorithms have been successfully used for: 1) non-contacting method of strain measurement in experimental mechanics [88, 99, 100], 2) deformation analysis of microscopic structures [97, 98], 3) measuring plant growth [89, 90] and 4) several biomedical applications such as measuring strain distributions in bovine articular cartilage [92], quantifying the motion of orbital tissues for diagnosing orbital disorders in the eye [93], measuring strain in bovine retina [94], studying the biomechanics of knees [95], quantifying the heart wall motion [95, 96] and cardiac motion estimation [91]. The intention of this work is not to evaluate the choice of the class of algorithm for a problem in hand and reevaluate their performances. However, most of the practical uses of optical flow algorithms require analyzing a before and after snapshot of the scene or object under investigation. DIC based flow field estimation techniques are robust to contrast and brightness changes in the scene and require only two frames of the scene. We demonstrate the performance improvements of the proposed Compute-and-Propagate framework with the DIC based Coarse-Fine search method.

### 3.3 A Novel $l_2$ Norm Based Evaluation of Optical Flow Estimates

The confidence of the flow field velocity estimates need to be determined to assess the performance of an optical flow algorithm. The most popular performance measure uses 1) an angular error $\theta_{\text{Error}}$ and 2) a difference in magnitude [78]. The angular error $\theta_{\text{Error}}$ between the vectors $\vec{U}_{GT}$ and $\vec{U}_{\text{Est}}$ shown in figure 25 can be estimated as

$$\theta_{\text{Error}} = \arccos(\vec{U}_{GT} \cdot \vec{U}_{\text{Est}})$$

(35)
Figure 25: Discrepancy between the ground truth optical flow $\vec{U}_{GT}$ and the estimated optical flow vector $\vec{U}_{Est}$

The difference in their magnitude can be computed as

$$\|\vec{U}_{GT} - \vec{U}_{Est}\|$$  \hspace{1cm} (36)

The ground truth of the object or scene displacement need to be known a priori for computing these performance metrics. Usually, the true deformation or motion field of a natural scene or object undergoing a deformation is unknown. Therefore the angular error and magnitude difference measures are not suitable for practical applications. We propose a new $l_2$ norm based optical flow evaluation method that does not require a prior knowledge of the true optical flow field.

Let,
- $I_{t_0}$ be the image of the scene prior to deformation,
- $I_{t_0+\delta t}$ be the image of the scene acquired after an interval $\delta t$,
- $[\vec{U}]$ be the true flow field matrix,
- $[X]$ be the pixel coordinate matrix of the images $I_{t_0}$ and $I_{t_0+\delta t}$ and
- $[\vec{U}_{Est}]$ be the estimated optical flow field matrix.

Given the pixel intensity coordinates $[X]$ of the initial image $I_{t_0}$, a ground truth optical flow field $[\vec{U}_{GT}]$ indicates that the pixel intensities of $I_{t_0}$ at $[X]$ have moved $[\vec{U}_{GT}]$. Thus the final image $I_{t_0+\delta t}$ can be represented using the initial image $I_{t_0}$, pixel intensity coordinate matrix $[X]$ and the ground truth optical flow field matrix $U_{GT}$ as follows.

$$I_{t_0+\delta t} = T_1([X], [\vec{U}_{GT}]) \cdot I_{t_0}$$  \hspace{1cm} (37)

Alternatively,

$$I_{t_0} = T_2([X], [\vec{U}_{GT}]) \cdot I_{t_0+\delta t}$$  \hspace{1cm} (38)

Now, an estimate of the image $I_{t_0}$, referred as $\hat{I}_{t_0}$, can be computed from the final image $I_{t_0+\delta t}$ using the optical flow matrix $[\vec{U}_{Est}]$ and the initial pixel coordinate matrix $[X]$ from equation (38). The deviation of the estimated optical flow field from the ground truth can be identified from the dissimilarity between $\hat{I}_{t_0}$ and $I_{t_0}$ as follows.

$$\text{Flow field estimation error} = \|\hat{I}_{t_0} - I_{t_0}\|$$  \hspace{1cm} (39)
If the estimated optical flow is equal to the ground truth, then the images \( \hat{I}_{t_0} \) and \( I_{t_0} \) will be same. Therefore, the measure of dissimilarity will be \( \sim 0 \). A larger flow field estimation error in equation (39) indicates a larger deviation of the estimated optical flow field from the ground truth. \( T_1 \) and \( T_2 \) in equations (37) and (38) are the interpolation functionals [102]. The most common choices are polynomial interpolation functions [72, 103], bicubic, spline and bicubic-spline interpolation algorithms [101].

The error metric defined in equation (39) measures the Euclidean distance between \( I_{t_0} \) and \( \hat{I}_{t_0} \) and estimates the degree of dissimilarity between two images by determining their pixel-to-pixel correspondence. Thus it is sensitive to even small pixel movements and object transformations. Recently a modified Euclidean distance measure called Image Euclidean Distance (IMED) is proposed for evaluating the dissimilarity between images [70]. The main reason for the performance degradation in using Euclidean distance for dissimilarity measurement comes primarily from the pixel-to-pixel correspondence matching. Basically, this is due to the orthogonal coordinate system employed for measuring an image similarity distance.

The IMED overcomes this drawback by giving varying weight to the adjacent pixels by considering a non-orthogonal basis. To successfully account for the displaced pixels, the IMED uses a non-orthogonal basis that assigns a varying weight to the adjacent pixels while comparing a pixel value between images. If there is any pixel displacement, it is expected to be near the original pixel position. Therefore, a gaussian kernel based pixel weighing scheme would be a natural choice for image similarity measurement. The problem of determining an optimal non-orthogonal basis is avoided by using the metric coefficient matrix \( G \) induced by the non-orthogonal basis for IMED computation. With a symmetric positive definite matrix \( G \), the IMED between \( I_{t_0} \) and \( \hat{I}_{t_0} \) can be computed as a \( G \)-innerproduct as follows.

\[
IMED = \langle I_{t_0}, \hat{I}_{t_0} \rangle_G = I_{t_0}G\hat{I}_{t_0}
\]

If the standard deviation \( \sigma \) chosen is far less than the image dimension, then the formation of metric coefficient matrix \( G \) and hence the IMED computation can be significantly reduced as follows.

\[
IMED = \langle I_{t_0}, \tilde{I}_{t_0} \rangle \quad (40)
\]

where \( \tilde{I}_{t_0} \) is computed by filtering \( \hat{I}_{t_0} \) using a gaussian filter of standard deviation \( \sigma \). The error metrics induced from the Euclidean distance in equation (39) and the IMED in equation (40) will be used to evaluate the performance of the CNP framework.

52
3.4 Compute-and-Propagate Framework with the Digital Image Correlation Based Coarse-fine Optical Flow Algorithm

A scene undergoing transformation typically involves:

1. scene translation due to camera or scene movement,

2. object transformation (translation, rotation, dilation, contraction and skewing) and

3. scene transformation.

Assuming that the initial and final images are of the same scene, the optical flow at a given pixel location \((x, y)\) often is related to the optical flow observed in its neighboring pixel locations \((x + \delta x, y + \delta y)\). For example, all the pixels in an object undergoing translation in a scene, will have nearly the same optical flow velocity. With no changes in a scene, a camera movement will result in an uniform optical flow field in the direction of the camera movement. Thus, in real-life situations change analysis typically have homogenous motion field in portions of a scene (involving objects in the scene) or parts of an objects. Differential optical flow techniques use a narrow spatial support and correspondence techniques such as coarse-fine methods search a small area around the original pixel location to determine the optical flow in that region. It can be observed that the computation and the search is repeated for almost all the pixel locations.

The proposed Compute-and-Propagate framework generates a sparse optical flow field based on the overall regional transformations in the scene. The generated sparse flow field can be used as an initial value for the iterative differential
Algorithm 3 Algorithm for computing the sparse flow field (FF) using Compute-and-Propagate (CNP) framework

1: procedure COMPUTEANDPROPAGATE($I_{t_0}$, $I_{t_0+\delta t}$, cnpLevels)
2:     prevFF(1, 1) ← DetermineOverallDisplacement
3:     for currCNPLevel ← 1 to cnpLevels do
4:         Initialize prevLevelFFRowColSz ← $2^{currCNPLevel-2}$
5:         Initialize currLevelFFRowColSz ← $2^{currCNPLevel-1}$
6:         Initialize currCol ← 1
7:         for prevCol ← 1 to prevLevelFFRowColSz do
8:             for prevRow ← 1 to prevLevelFFRowColSz do
9:                 currFF(currRow, currCol) ← prevFF(prevRow, prevCol)
10:                currRow ← currRow + 1
11:             currFF(currRow, currCol + 1) ← prevFF(prevRow, prevCol)
12:                currRow ← currRow + 1
13:         end for
14:     currFF ← computeCNPRegionalFF(currFF, currCNPLevel, $I_{t_0}$, $I_{t_0+\delta t}$)
15: end procedure

16: procedure computeCNPRegionalFF(currFF, currCNPLevel, $I_{t_0}$, $I_{t_0+\delta t}$)
17:     Initialize imgRowSz ← rowSize($I_{t_0}$ or $I_{t_0+\delta t}$)
18:     Initialize imgColSz ← colSize($I_{t_0}$ or $I_{t_0+\delta t}$)
19:     Initialize ffRowColSz ← $2^{currCNPLevel-1}$
20:     for ffCol ← 1 to ffRowColSz do
21:         for ffRow ← 1 to ffRowColSz do
22:             cnpRowMin ← $\lfloor imgRowSz \ast (ffRow - 1)/ffRowColSz \rfloor + 1$
23:             cnpRowMax ← $\lfloor imgRowSz \ast ffRow/ffRowColSz \rfloor$
24:             cnpColMin ← $\lfloor imgColSz \ast (ffCol - 1)/ffRowColSz \rfloor + 1$
25:             cnpColMax ← $\lfloor imgColSz \ast ffCol/ffRowColSz \rfloor$
26:             Choose a template from the initial image $I_{t_0}$ from the region $(cnpRowMin : cnpRowMax, cnpColMin : cnpColMax)$
27:             Initial regional optical flow computed for this region using CNP is available in currFF(ffRow, ffCol)
28:             With the knowledge of the overall flow field in the region, identify the best matching region in the final image $I_{t_0+\delta t}$ using a normalized correlation measure as in (34)
29:             Update the flow field in the region in currFF(ffRow, ffCol)
30:         end for
31:     end for
32: end procedure
and coarse-fine search based optical flow estimation techniques as shown in figure 26. The initial estimate of the optical flow generated by the CNP framework can help 1) the optical flow estimation algorithm to converge faster and 2) to converge to a true solution. Figure 27 shows a schematic representation of the sparse flow field computation using CNP framework. The CNP framework estimates an overall scene movement $[U]_1$ at level-1 and sub-divides the whole region into four sub-regions. All the four sub-regions use the sparse flow field estimate $[U]_1$ from level-1 to identify regional transformations in the four sub-regions. Each of the four sub-regions are further sub-divided into four sub-regions. At each subsequent level $i+1$ in the CNP framework, the sparse flow field from the regions in the previous level $[U]_i$ are propagated to the appropriate regions. The region subdivision, regional optical flow field computation and propagation continues up to a user defined depth / level $N$. Thus the CNP framework starts by estimating the overall scene movements and at each new level derives the regional movements. For regional optical flow field estimation, a normalized correlation coefficient measure (equation 33) based template matching algorithm can be used. Fast fourier based correlation based template matching can be faster depending on the region size 67. Algorithm 3 in page 54 layouts the computational steps involved in the CNP framework.

The proposed CNP framework can be efficiently implemented in a parallel cluster. After a region is subdivided at a given level in the CNP framework, the translations in each of the subdivided regions can be independently computed in a parallel node. In this work, we have implemented the proposed CNP framework in MatlabMPI 104, a parallel Matlab facility, in a IBM © Server pSeries TM650 105 in a cluster computing environment. The iterative schemes and computationally intensive parts of the algorithms are written in ANSI-C to create appropriate dynamic linked libraries (.dll) using mex-c compiler for calling the routines from the MATLAB environment.
3.5 Results

Figure 28: Cotton image $I_{t_0}$ shifted 10 pixels diagonally to generate $I_{t_0 + \delta t}$. The purpose is to demonstrate the validity of the proposed error metric for evaluating optical flow results.

The proposed error metric is suitable for evaluating an estimate of motion field in real-life image sequences with unknown ground truth of the resulting motion field. To demonstrate the validity of the proposed error metric, we use an image of random cotton fibers acquired using a white-light confocal microscope (WLCM) to generate an initial and a final image with a known motion field. The cotton fiber image itself is used as the initial image $I_{t_0}$. The image $I_{t_0}$ is translated 10 pixels diagonally with an optical flow of $(u \approx 7.07, v \approx 7.07)$ at each pixel resulting in the final image $I_{t_0 + \delta t}$. When the estimated optical flow $(\hat{u}, \hat{v})$ (using either DIC with or without CNP framework as in section 3.4) is equal to the true motion field, the image reconstructed $\hat{I}_{t_0}$ using $(\hat{u}, \hat{v})$ will be identical to the initial image $I_{t_0}$, omitting...
the borders for any occluding regions. Thus, the value of the proposed Euclidean error metric in equation (39) and IMED in equation (40) will be \( \sim 0 \). Figure 28 shows the image of random cotton fibers at an initial stage and a final stage with 10 pixels diagonal movement. A quiver plot of the optical flow field estimated using DIC with CNP framework in figure 28c indicates a diagonal movement of the cotton fibers. A difference image \( I^t_0 - \hat{I}^t_0 \) shown in figure 28d and the error magnitudes confirms that the image \( \hat{I}^t_0 \) reconstructed using the estimated optical flow \((\hat{u}, \hat{v})\), which is same as the true motion field, is identical to the initial image \( I^t_0 \) resulting in an Euclidean distance and IMED of 0. This test ensures the validity of the proposed error metrics. It should be noted that the choice of \( \sigma \) in computing the IMED in equation (40) determines the error tolerance in the optical flow estimates. With \( \sigma = 1 \), IMED solution will converge to the standard Euclidean distance error metric in equation (39). In this work, we use \( \sigma = 3 \), allowing a pixel deformation up to 3 pixels in the reconstructed image \( \hat{I}^t_0 \). In other words, an optical flow estimate error greater than 3 pixels, in all directions, will be penalized.

Table 2: Performance evaluation of DIC coarse-fine optical flow estimation with (CNP-DIC) and without(DIC) Compute-and-Propagate framework

<table>
<thead>
<tr>
<th>Scene Description</th>
<th>Iteration Counts</th>
<th>Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DIC</td>
<td>CNP-DIC</td>
</tr>
<tr>
<td>Clock</td>
<td>9178021</td>
<td>8364060</td>
</tr>
<tr>
<td>Random cotton fibers</td>
<td>3560111</td>
<td>6349.72</td>
</tr>
<tr>
<td>Rotating globe</td>
<td>8541639</td>
<td>1693.80</td>
</tr>
<tr>
<td>Rotating globe with lateral movement</td>
<td>7754842</td>
<td>5228.87</td>
</tr>
<tr>
<td>Lamina Cribrosa at IOP levels of 12mmHg and 60mmHg</td>
<td>2949741</td>
<td>3894.27</td>
</tr>
<tr>
<td>Lamian Cribrosa after restoration</td>
<td>3447427</td>
<td>7180.17</td>
</tr>
<tr>
<td>Rotating sphere</td>
<td>6788882</td>
<td>828.58</td>
</tr>
</tbody>
</table>

We use the following images from a variety of scenes to evaluate the performance improvement using CNP framework with DIC optical flow algorithm. All the images except the sphere image sequence are real-life image sequences.

1. A lateral clock movement from an initial position, with small changes in the positions of the minute and hour hands (figure 29).
2. Random cotton fibers moved diagonally from an initial position; images are acquired using a WLCM (figure 30)

3. A rotating tabletop globe (figure 31)

4. A rotating tabletop globe with some lateral movement (figure 32)

5. Lamina cribrosa of retina at intraocular pressure levels (IOP) of 12mmHg $I_{0}$ and 60mmHg $I_{0+\delta}$ (figure 33)

6. Images of lamina cribrosa at 12mmHg and 60mmHg IOP levels restored using adaptive Karhunen-Loève filter (figure 34)

7. A rotating sphere (figure 35)

The rotating sphere images were obtained from [87]. The results of the optical flow field estimation using DIC method with and without CNP framework is shown in the figures 29, 30, 31, 32, 33, 34, 35. The results include 1) a quiver plot of the estimated optical field from the respective methods, 2) a difference image and 3) their error magnitudes (Euclidean distance and IMED). It can be observed from the quiver plots and the error magnitudes that the optical flow estimation using CNP framework gives an improved performance in terms of deriving a more accurate motion field. Table 2 contains the error magnitudes of Euclidean distance and IMED for DIC optical flow method with and without CNP framework and the number of iterations they needed to converge. The IMED method shows a reduction in the error metric with the CNP framework and an improvement over the Euclidean error metric. It can be seen that the CNP framework significantly reduces the error measure and converges with a fewer number of iterations when there are multiple large transformations.

Figure 36 shows the comparative error metrics of Euclidean and IMED on the performance of the DIC and CNP-DIC methods on the experimental image sequences. Figure 37 shows the comparative convergence rate of the DIC and CNP-DIC methods on the experimental image sequences in terms of the number of iterations they needed to converge. It can be observed that the DIC method with CNP framework consistently outperforms and shows an improvement over the DIC method when there are multiple large motions.

3.6 Conclusion

We have presented novel error metrics for evaluating the performance of optical flow algorithms and a Compute-and-Propagate framework for use with iterative optical flow estimation algorithms. The methods are demonstrated with the iterative DIC optical flow algorithm. One of the drawbacks of the traditional optical flow performance measures is the requirement of a prior knowledge about the ground truth motion field. Even a few small differences in the estimated optical
Figure 29: Optical flow estimation of a scene containing a clock and a mug
flow field may indicate and result in a significant reduction in the performance of the optical flow method. The proposed error metrics allow the evaluation of optical flow methods on image sequences with no prior knowledge about the true motion field. The IMED metric takes into account the possibility of small errors in the estimated optical flow field and allows a tolerance in the flow field estimates, without penalizing the optical flow method. An optical flow method converging closer to the true motion field will receive a relatively lesser IMED measure. The CNP framework demonstrates significant performance improvement in terms of converging to the true motion field and convergence rate in a variety of scenes. The CNP framework accounts for homogeneous pixel movements or transformations. Thus regions without any transformations are readily identified at a certain level in the CNP framework depending on the size of the region with homogeneous motion field. The regions undergoing transformation typically involve a group of pixels or a region which can identified using the CNP framework. The initial estimate from the CNP framework will allow the iterative optical flow algorithms to converge faster and to converge to the true motion field in the scene. As the CNP framework considers overall large transformations at various levels and regions, it can considerably reduce the errors in the optical flow estimates due to the aperture problem.
Figure 30: Optical flow estimation of a scene exhibiting movement of cotton fibers observed using a white-light confocal microscope.
Figure 31: Optical flow estimation of a scene exhibiting a rotating tabletop globe
Figure 32: Optical flow estimation of a scene exhibiting a rotating tabletop globe with some lateral movement

(a) Initial Image $I_{t_0}$

(b) Final Image $I_{t_0+\delta t}$

(c) Optical flow $\vec{U}_{Est}$ computed using CNP-DIC method

(d) Difference Image $I_{t_0} - \hat{I}_{t_0}$ computed using CNP-DIC method; $\|I_{t_0} - \hat{I}_{t_0}\| = 2386.04; IMED(I_{t_0}, \hat{I}_{t_0}) = 821.77$

(e) Optical flow $\vec{U}_{Est}$ computed using DIC method

(f) Difference Image $I_{t_0} - \hat{I}_{t_0}$ computed using DIC method; $\|I_{t_0} - \hat{I}_{t_0}\| = 5228.87; IMED(I_{t_0}, \hat{I}_{t_0}) = 3504.74$
Figure 33: Optical flow estimation of the images of lamina cribrosa of cow retina at intraocular pressure levels of 12mmHg and 60mmHg.
Figure 34: Optical flow estimation of the images of lamina cribrosa of cow retina at intraocular pressure levels of 12mmHg and 60mmHg restored using KL filter.
(a) Initial Image $I_{t0}$

(b) Final Image $I_{t0+\Delta t}$

(c) Optical flow $\tilde{\mathbf{U}}_{Est}$ computed using CNP-DIC method

(d) Difference Image $I_{t0} - \tilde{I}_{t0}$ computed using CNP-DIC method; $\|I_{t0} - \tilde{I}_{t0}\| = 828.58; IMED(I_{t0}, \tilde{I}_{t0}) = 332.10$

(e) Optical flow $\tilde{\mathbf{U}}_{Est}$ computed using DIC method

(f) Difference Image $I_{t0} - \tilde{I}_{t0}$ computed using DIC method; $\|I_{t0} - \tilde{I}_{t0}\| = 828.58; IMED(I_{t0}, \tilde{I}_{t0}) = 332.10$

Figure 35: Optical flow estimation of the images of a rotating sphere
Figure 36: Euclidean and IMED error metrics of the DIC and CNP-DIC methods on the experimental image sequences.
Figure 37: Convergence rate (Iteration counts) of the DIC and CNP-DIC methods on the experimental image sequences
Conclusion

We have developed a comprehensive computational framework for studying structural changes in microscopic specimens using a confocal microscope. The realtime tracking and visualization architecture presented in chapter 1 can be used for regular laboratory and clinical observation of microscopic structures. The realtime tracking feature in a confocal microscope can be used for patient care in a longitudinal health-care setup. For example, a reference state of the patient’s site of pathology, such as cornea and retina, can be imaged and tracked upon patient’s subsequent visits more precisely using the tracking architecture. The behavior of such micro-structures can be observed in realtime under various experimental conditions using the visualization facility.

The KL filter presented in chapter 2 filters random noise artifacts. In point-scanning confocal microscopes, the KL filter can significantly reduce the scan-lines in the optical section images. Also use of KL filter prior to restoration allows the use of a simple inverse deconvolution method to restore the images. A stack of optical section images can be restored in realtime in a cluster environment using KL prefiltering and inverse deconvolution method. An exponential increase in the contrast measure of the images restored using iterative restoration algorithm such as Lucy-Richardson algorithm indicated amplification of ringing artifacts at higher iterations. The ringing metric presented in chapter 2 can be used with iterative restoration algorithms to quantify the amount of ringing introduced in the restored images and identify an appropriate terminal condition for the iterative algorithms.

The Compute-and-Propagate framework presented in chapter 3 demonstrates significant improvement in the speed (in terms of number of iterations) and convergence (in terms of an error measure) of Digital Image Correlation based iterative optical flow estimation method. This framework can be used in differential and correspondence based iterative optical flow algorithms that use an initial flow estimates for motion field estimation. Also we have developed a $l_2$ norm based error metric that can be used for evaluating the performance of optical flow algorithms without a prior knowledge about the true motion field.

The framework presented here can be extended for use with other imaging modalities such as PET, CT, MRI, etc., that acquires optical section images at various depths to build a 3D representation of the structure. Inverse restoration with the KL noise prefiltering and the ringing metric can be directly applied for use
with other imaging systems. The $l_2$ norm based error metric for evaluating optical flow algorithms and the CNP framework is applicable for all imaging modalities. The framework developed is comprehensive and will facilitate studying the structural changes of microscopic structures such as lamina cribrosa and can be used for understanding the properties of materials, for understanding the underlying mechanism involved in the structural changes, for regular clinical use and longitudinal patient care.
Bibliography


[34] Visualization Toolkit, http://www.vtk.org


Appendix: Review of Linear Algebra

Definition 1. In any vector space, a vector \( x \) is said to be a linear combination of vectors \( \{\phi_k\}_{k=1}^n \) if \( x = \sum_{k=1}^n \alpha_k \phi_k \) where \( \{\alpha_k\}_{k=1}^n \) are the scalar coefficients of \( \{\phi_k\}_{k=1}^n \).

Definition 2 (Linear subspace). Let \( V \) be a vector space. Let \( N \) be a set of vectors \( \{x_i\}_{i=1}^n \) such that \( x_i \in V \). The set \( N \) is said to be a linear subspace of \( V \) if:

1. \( N \) contains a zero-vector
2. \( \forall x, y \in N, x + y \in N \) and
3. \( \forall x \in N \) and a scalar \( \alpha, \alpha x \in N \)

Definition 3 (Linear independence). Let \( N \) be a set of vectors in vector space \( V \). Any vector \( x \in V \) is said to be linearly independent of \( N \) if \( x \) cannot be expressed as linear combination of vectors in \( N \). The set of vectors \( N \) is said to be linearly independent if each vector in \( N \) is linearly independent of every other vectors in the set \( N \).

Definition 4 (Basis and dimension of a vector space). A set of vectors \( N \) is said to be a basis for a vector space \( V \) if the set \( N \) is linearly independent and every vector in \( V \) can be expressed as a linear combination of vectors in \( N \). The dimension of the vector space \( V \) is the number of vectors in its basis \( N \).

Definition 5 (Distance function in a metric space). A metric space \( X \) is equipped with a distance function \( d(x, y) \) for every combination of points \( x, y \in X \) with the following properties:

1. \( d(x, y) > 0 \), when \( x \neq y \)
2. \( d(x, y) = 0 \), when \( x = y \)
3. \( d(x, y) = d(y, x) \) and
4. \( d(x, z) \leq d(x, y) + d(y, z) \)

Definition 6 (Cauchy sequence and complete metric space). A sequence of vectors \( \{x_i\} \) is said to be convergent if there exists a limit vector \( x \) such that \( d(x_n, x) \to 0 \) as \( m, n \to \infty \). A metric space \( X \) is complete if and only if every Cauchy sequence of points in \( X \) converges.
Definition 7 (Vector norm). A normed linear space $X$ is a metric space with the distance metric $d(x, y)$ defined by a norm $\|x - y\|$. The norm assigns a scalar length to each vector with the following properties:

1. $\|x\| = 0 \iff x = 0$
2. $\|ax\| = |a| \cdot \|x\|$, $\forall x \in X$ and $\forall a \in \mathbb{R}$ and
3. $\|x + y\| \leq \|x\| + \|y\|$, $\forall x \in X$

Given a vector $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$, a $p$-norm is defined as

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

Special cases of a $p$-norm include:

1. Euclidean norm: $\|x\|_2$
2. $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$
3. $\|x\|_{\infty} = \max_i |x_i|$

Definition 8 (Banach Space). A complete normed linear space is a Banach space.

Definition 9 (Innerproduct). An innerproduct space $X$ is a linear space equipped with a real-valued function called an innerproduct on pairs of elements of $X$ denoted $\langle x, y \rangle$, where $x$ and $y$ are elements of $X$ with the following properties.

1. $\langle x, x \rangle \geq 0$, $\forall x \in X$, with $\langle x, x \rangle = 0 \iff x = 0$
2. $\langle x, y \rangle = \langle y, x \rangle$, $\forall x \in X$
3. $\langle ax, y \rangle = a \langle x, y \rangle$, $\forall x, y \in X$, $\forall a \in \mathbb{R}$

A norm induced by an innerproduct $\|x\| = \sqrt{\langle x, x \rangle}$, $\forall x \in X$

Definition 10 (Hilbert space). A complete innerproduct space $\mathbb{H}$ is called a Hilbert space.

Definition 11 (Vector orthogonality). A vector $x \in \mathbb{H}$ is orthogonal to a vector $y \in \mathbb{H}$, if $\langle x, y \rangle = 0$. A set of vectors $\{x_i\}_{i=1}^{n}$ is said to be orthogonal set if and only if $\langle x_i, x_i \rangle = 0$ whenever $i \neq j$

Definition 12 (Vector orthonormality). A set $S$ of vectors $\{x_i\}_{i=1}^{n}$ is said to be orthonormal, if: 1) $S$ is an orthogonal set and 2) $\|x\| = 1$, $\forall x \in S$

Definition 13 (Orthonormal basis). A set $S$ of vectors $\{x_i\}_{i=1}^{n}$ is an orthonormal basis for the vector space $\mathbb{V}$, then any $x \in \mathbb{V}$ can be expressed as a linear combination of the basis as $x = \sum_{i=1}^{n} \langle x, x_i \rangle x_i$. 

81
**Definition 14** (Rank of a matrix). Rank of a matrix $A \in \mathbb{R}^{M \times N}$ is the maximum number of linearly independent rows or columns of a matrix. Thus the rank of a matrix $A$, $\text{Rank}(A) \leq \min(M, N)$. A matrix is said to be rank deficient if its rank is less than $\min(M, N)$.

**Definition 15** (Row and column space of a matrix). The set of all linear combinations of the column vectors of a matrix $A$ forms the column space of $A$. Similarly, the set of all linear combinations of the row vectors of $A$ forms the row space of $A$.

**Definition 16** (Orthonormal matrix). The columns of an orthonormal matrix are orthonormal to each other. Let $Q \in \mathbb{R}^{M \times N}$ be an orthonormal matrix. Then $QQ^T = I$, where $I$ is the identify matrix. For any matrix $A \in \mathbb{R}^{M \times N}$,

$$
\|QA\|_2 = \|\langle QA, QA \rangle\|
= \|\langle QQ^TA, A \rangle\|
= \|\langle A, A \rangle\|
= \|A\|_2
$$

(41)

Thus an orthonormal projection of a set of vectors preserves their Euclidean distance.

**Theorem 1** (Spectral theorem). A matrix $A \in \mathbb{R}^{N \times N}$ is symmetric if and only if there is a real orthogonal matrix $X$ such that $X^TAX = \Lambda$. Also it can written as

$$
A = X\Lambda X^T
= \sum_{i=1}^{N} \lambda_i x_i x_i^T
$$

(42)

**Theorem 2** (Singular value decomposition). Let $A \in \mathbb{R}^{M \times N}$. There is a decomposition $A = U\Lambda V^T$ such that the diagonal elements of the diagonal matrix $\Lambda$ are non-negative and arranged in a non-increasing order. The matrices $U$ and $V$ are orthonormal matrices and are called left and right singular matrices respectively.
Vita

Madhusudhanan Balasubramanian was born in 1977 in Chidambaram, India. He received his Diploma in Electrical and Electronics Engineering from Muthiah Polytechnic, Annamalainagar, India, in 1995. He earned his Bachelor of Engineering in Electrical and Electronics degree from PSG College of Technology, Coimbarore, India, in 1998. He worked as a software engineer from August 1998 to December 1999 for IMR Global Inc. (currently with CGI Inc.), Bangalore, India, and as a senior systems engineer from December 1999 to June 2001 for Wipro Technologies, Bangalore, India. He came to Louisiana State University, Baton Rouge, in August 2001 to pursue a doctoral degree in computer science. He also has received a Master of Science in Systems Science degree from LSU, Baton Rouge in December, 2003. His master’s degree thesis was “Computer Assisted Eye Fungal Infection Diagnosis”. He is currently a doctoral candidate in the Department of Computer Science and will be graduating in May 2006.