Acquisition and Analysis of Cryogenic Transmission Electron Microscope Biological Images

Lewis Collier

URI CSC592

Spring 2005

Lewis-Collier@Cox.net

Overview

- Acknowledgments
- Biological Impetus
- Previous CTEM LDL Research
- Introduction to TEM
- Existing CTEM Image Processing
- Project Importance
- Project Plan
  - Image Acquisition Support Software
  - Image Processing Analysis Software
  - Acquire and Analyze Images
- Project Results
  - Documentation
  - Software
  - Analysis
  - Problems
- Future Work
- Demonstration of Prototype Software

Acknowledgments

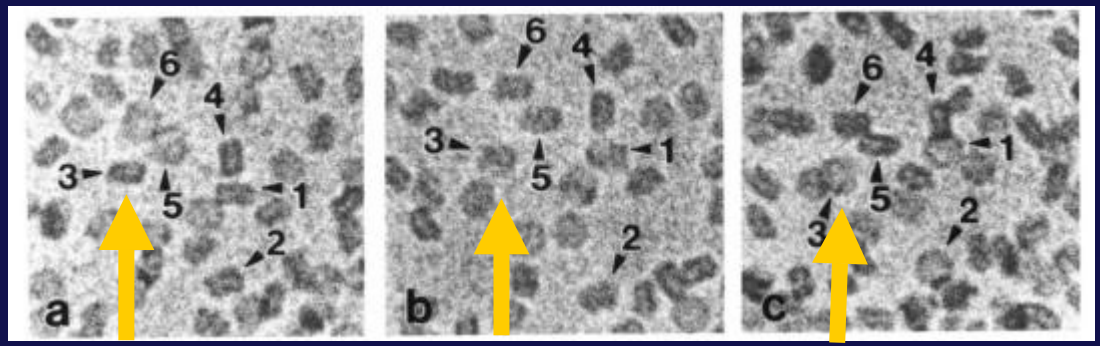
- Dr. Martin – for suggesting the project and helping understand the biology
- Dr. Peckham – for encouraging me to take this course and helping me find Dr. Herve' when I needed to find him
- Dr. Herve' – for providing good ideas for image processing and for the image processing library base classes
- Paul Johnson – for describing the JEOL TEM system
- Matt Kayala – for generating test images
- PFAST team – for their many helpful comments during the semester reviews

Biological Impetus

- Biologic target of interest is LDL (low density lipoprotein)
  - Lipid (fat) & protein package (“suitcase for cholesterol”)
  - HDL : smaller cells with more protein than fat (lipid), carry cholesterol back to liver
  - LDL : larger cells with more fat than protein, carry cholesterol to cells that need it
- Shape (and more importantly SIZE) of LDL is important to understand bio-mechanisms
  - Small number of bigger LDL particles correlates to good health
  - Big number of smaller LDL particles correlates to unhealthy

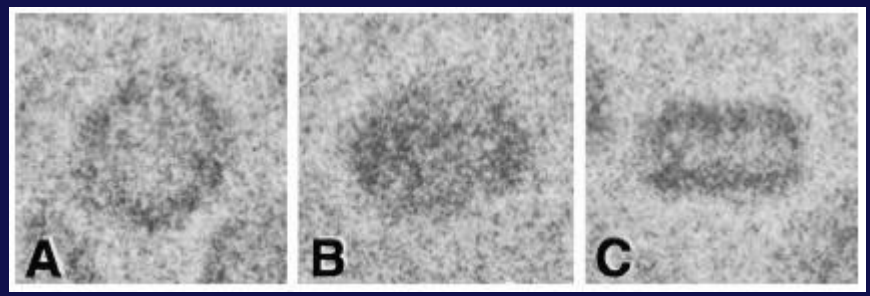
Previous CTEM LDL Research

- Antwerpen, et. al (Univ. AZ Tucson) 1994
  - CTEM images of human LDL onto photo paper
  - Analysis of Cryo-TEM process and possible errors
  - Simple projections at -45°, 0°, and +45°
  - Estimated shape as discs 21.4nm diameter by 12.1nm high
  - Did not estimate the thickness of the outer band



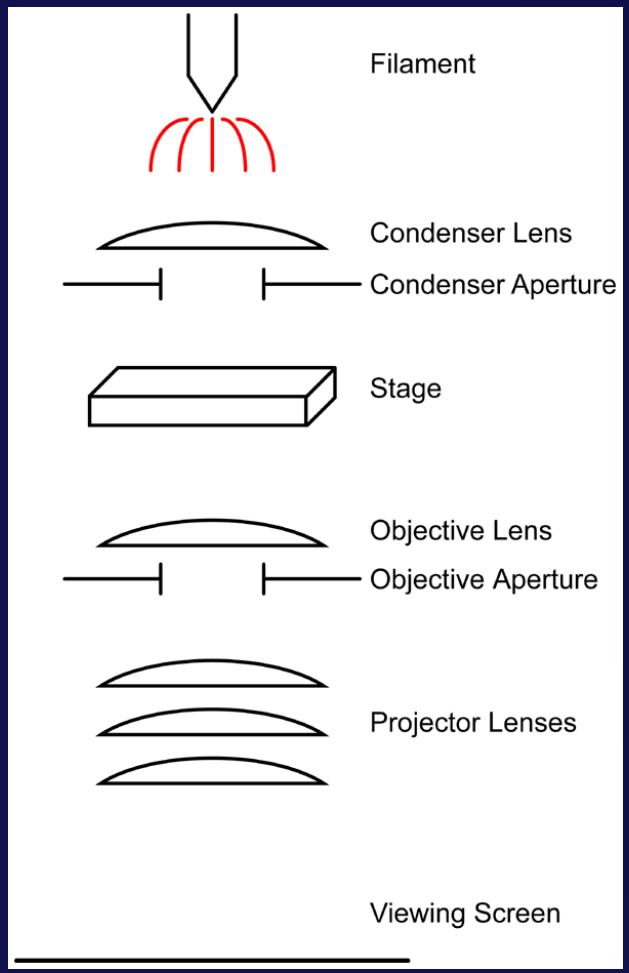
Previous CTEM LDL Research

- Antwerpen, et. al (Univ. AZ Tucson) 1997 and 1999
  - CTEM images of human LDL onto photo paper
  - Broke LDL into classes (VLDL, IDL, and LDL)
  - Simple projections at -45°, 0°, and +45°
  - Estimated shape of classes of discs
  - Used Adobe Photoshop measurement tools to take metrics



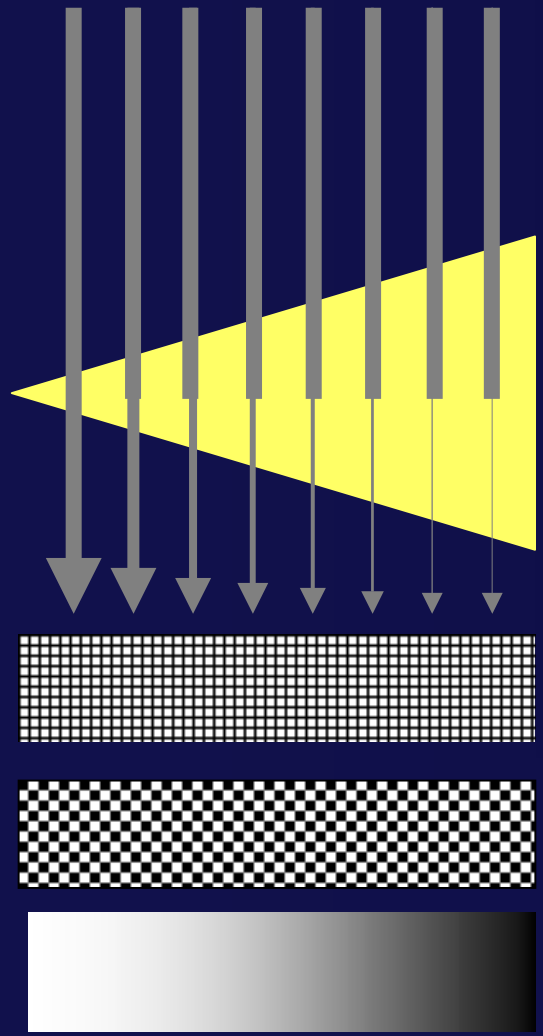
-45 degrees 0 degrees 45 degrees

Introduction to TEM



- Electrons accelerated through sample on stage
  - Electron energy
  - Electron density
  - Many tradeoffs between energy and image quality.
- Image Setup
  - Condenser Aperture
  - Objective Aperture
  - Magnification
  - Focusing
  - Magnification Calibration
  - Resolution Test
- Sensor Setup
  - Contrast Transfer Function

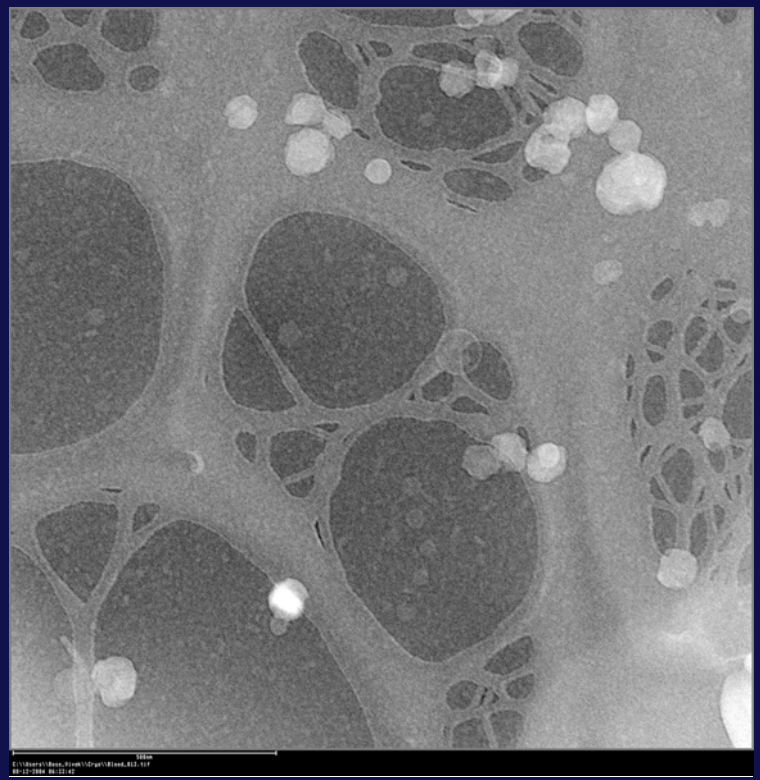
Introduction to TEM



- E Beams
- Sample
- Scintillator
- CCD
- Gray Levels

- Electrons pass through sample
  - More dense regions let less electrons pass through
- Electrons hit Scintillator
  - Electrons converted to photons
  - Photons passed to sensor
- CCD Sensor
  - 2Kx2K pixels
  - 12 bits per pixel
  - Dark areas are more dense
    - Less electrons pass through
  - Light areas are less dense
    - More electrons pass through

Introduction to TEM



- Invert gray scale for further processing
- Look for objects of the desired shape and sizes
  - Correlation processing
  - FFT Correlations

Existing CTEM Image Processing

- Dr. Tim Baker at USCD (formerly at Perdue)
  - http://cryoem.ucsd.edu/
  - Approximately 20 special purpose programs for virus modeling
  - Many images converted into 3-D model
- HREM
  - “AutoTuning Software for your TEM”
  - PCI hardware for processing speedup
  - Only works with Gatan Cameras (URI has TVIPS)
- Xmipp : An Image Processing Package for EM
  - Used for 3-D reconstruction from large data sets

…Just to list a few…

Project Importance

- LDL size is tied to many health issues but size measurements vary wildly
  - LDL size is known to vary in diabetics (25nm to 27nm)
  - Heart Disease is tied to small LDL (12 and 17nm vs. standard 25 and 27nm) *AND* small HDL as well
  - Increases of only 0.98nm in diameter of LDL provided a decrease in the progression of atherosclerosis in diabetics

- Recent (2004) HPGC measurements lend credibility to the discoid shape model but find varying heights.
  - Volume/mass ratios matched discs but not spheres
  - Volume more determined by height than diameter
  - Diameter and height seemed to be unrelated

- NMR can provide ratios of sizes but not shape details

A method of determining the actual size and shape of lipoproteins is needed to allow for further modeling and analysis

Project Plan

- Provide computer vision tools to allow for LDL geometric modeling from a small number of images to solve the following problems:
  - Long exposures can melt sample
  - TEM cannot provide slices for contour 3-D reconstruction
  - Rotation of sample for 3-D reconstruction is problematic
    - Requires addition of inert markers for focusing
    - Requires multiple focus cycles which can melt sample
  - Multiple samples for statistical analysis require more processing than can be performed using manual tools

Project Plan

- Image Acquisition Support Software
  - TEM Settings for optimal processing
  - ROI analysis for selection of “target rich” areas
  - In-situ analysis to help in image selection for small angle rotations
- Image Processing Analysis Software
  - Model discs with geometric parameters
  - Process images to locate candidate lipoproteins
- Image Acquisition and Analysis
  - Try acquisition support software with the TEM
  - Qualify analysis software with models
  - Analyze C-TEM images to try to locate and measure LDLs

Support Software Plan

- TEM Settings for optimal processing
  - Bright and Dark State Corrections
    - Remove outlier pixels
    - Provide background images for downstream processing
  - Focus settings
    - Provide assistance in adjusting the TEM settings in order to get the best possible image for object detection
    - Select ROI for magnification and rotation angles
  - Contrast Settings
    - Assist in setting the exposure time and electron beam voltages in order to maximize contrast

Analysis Software Plan

- Image Processing Analysis Software
  - Model discs with geometric parameters
    - Diameter, height, wall thickness, center opacity
    - Model CCD and TEM noise
  - Process images to locate candidate lipoproteins
    - Template matching (correlation “manually” and with FFTs)
    - Investigate limits over geometric models (how well do we do)
    - Investigate other methods
- Measure geometric parameters
  - Correlation against known models
  - FFT-based correlation (multiply spectrums and do IFFT)

Analysis Plan

- Use the acquisition support software with the TEM
- Analyze generated discs to understand limits of pattern matching
- Process C-TEM images to try to locate LDLs
  - Previous work had done -45^{o}, 0^{o}, and +45^{o} with refocus
  - Our goal was to allow for small +/- 5^{o} rotations and use computer vision to process shapes and sizes rather than relying on human intervention with Photoshop

Project Results

- Documentation
- Software
- Analysis
- Setbacks

Project Documentation

- System Performance Specification (SPS)
  - Defines what the system must do without telling *HOW* to do it.
- Software Requirements Specification (SRS)
  - Describes the requirements that the software must meet.
  - Can include “how”.
  - Traceable back to the SPS.
- Final Project Report
- These Slides☺

Project Software

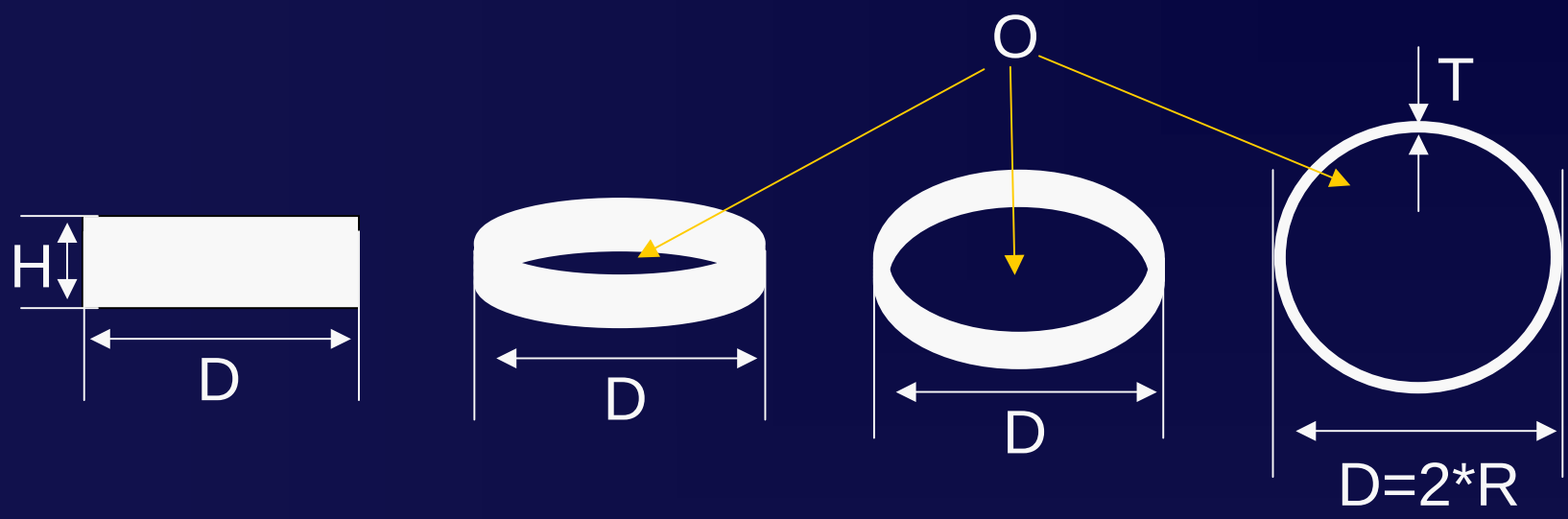
- 3rd Party SW
  - Dr. Herve's CSC583 base classes (framework w/ no code)
  - Code Warrior v8.3, GLUT v3.7.6, QuickTime v6.5.2
  - 2-D FFT from *Takuya OOURA, Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-01 Japan*
  - LZW and TIF2TGA converter code was accessed as well
- C++ code written by me this semester and used in this project
  - ~40 general purpose image processing routines (~3kSLOC)
  - ~20 specific routines (~2kSLOC)

Project Analysis

University of Rhode Island

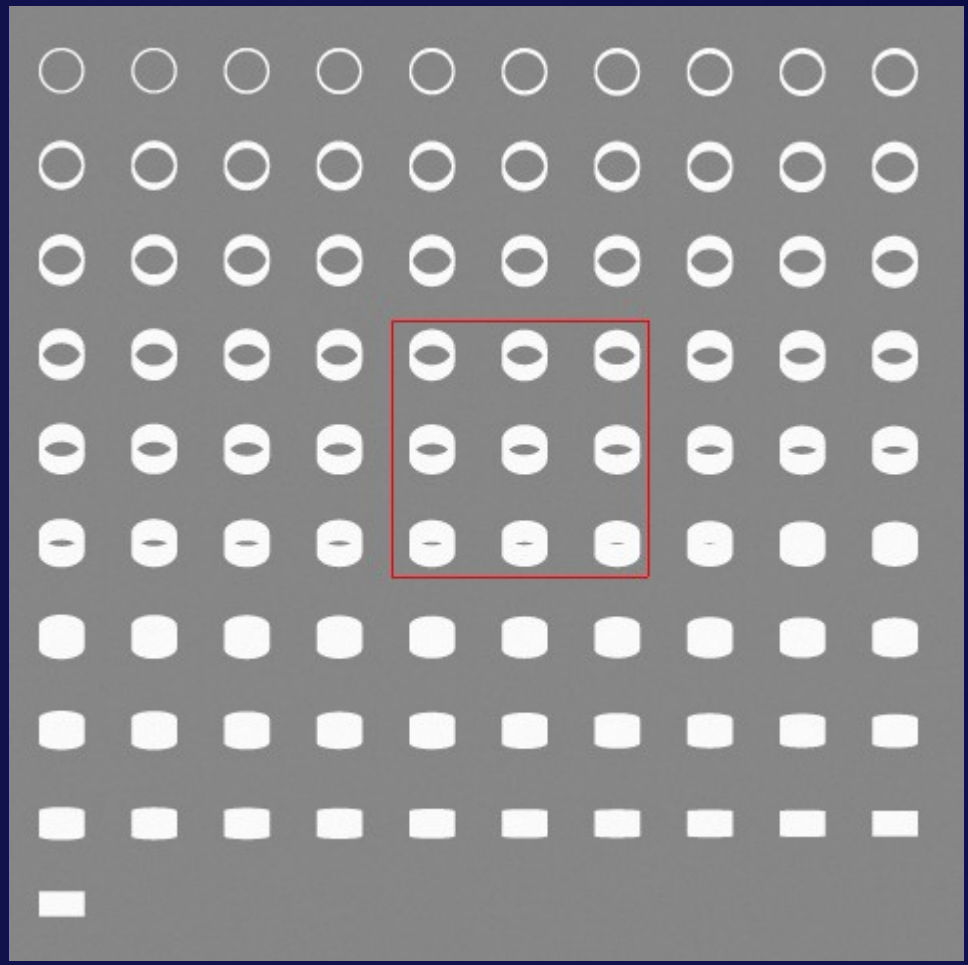
- Build simple model of discs
  - Defined by Radius, Height, and Wall Thickness
  - Maya model includes opacity of inner region
- Compare “Manual” Correlation results
  - Single image vs. 91 tilt angles (~ 14 mins on P4-1.6Ghz)
  - Use image combined 4:1 for faster processing
  - R=41:45; H=18:22; T=3:7 (125 tests ~ 10 hrs)
- Compare FFT-Correlation results
  - Use full 2Kx2K image
  - R=37:49; H=15:25; T=1:9 (1287 * 150 secs ~ 55 hrs)

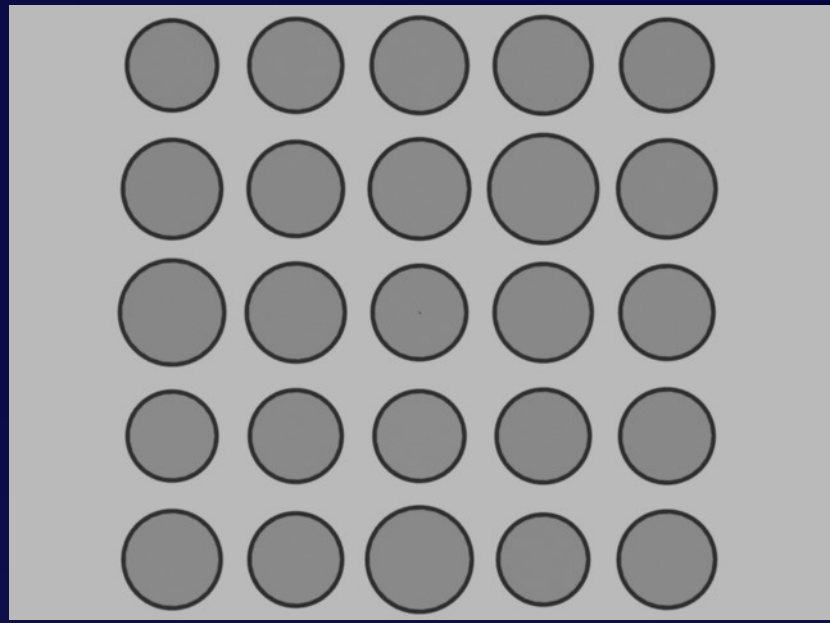
Disc Model



- H = Height in pixels (1pixel ≈ 0.24nm)
- D = diameter in pixels (define as R)
- T = wall thickness in pixels
- O = opacity (% of full scale fill factor)
- A = horizontal tilt angle in degrees (90 = rectangle, 0 = circle)

Disc Models

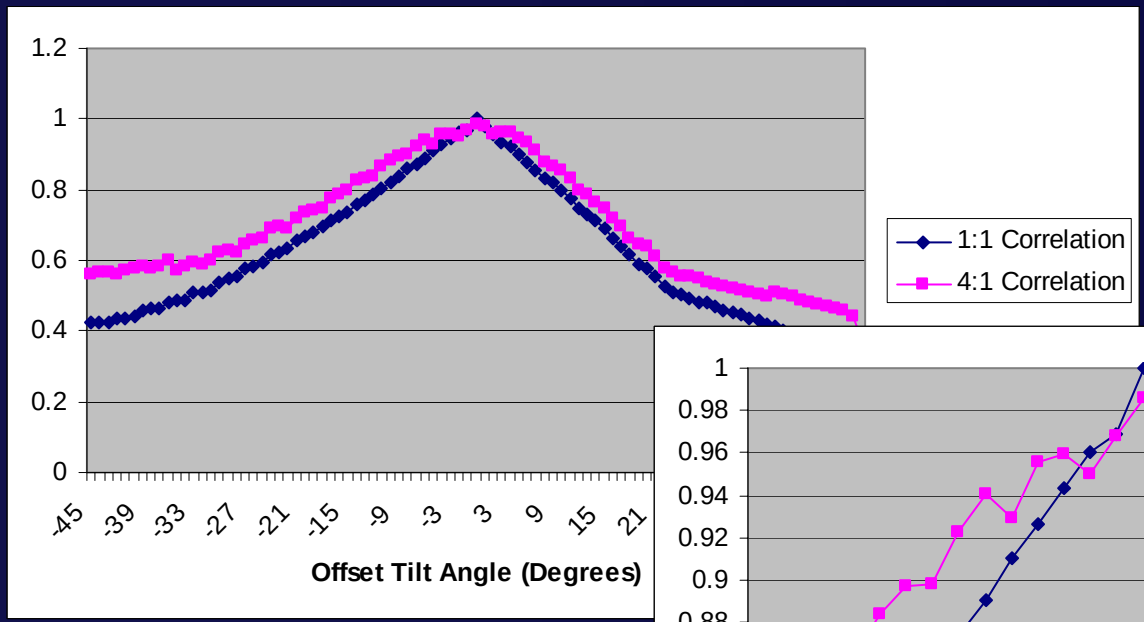


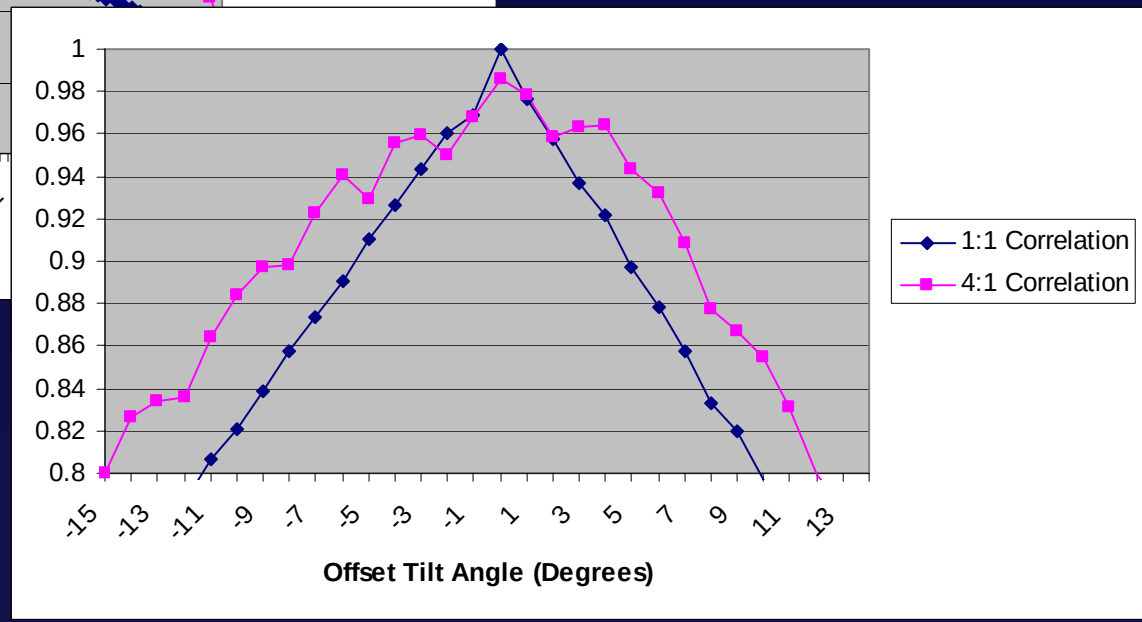


Discs generated with Maya 3-D modeling and Rendering software.

Auto-generated discs in my software.

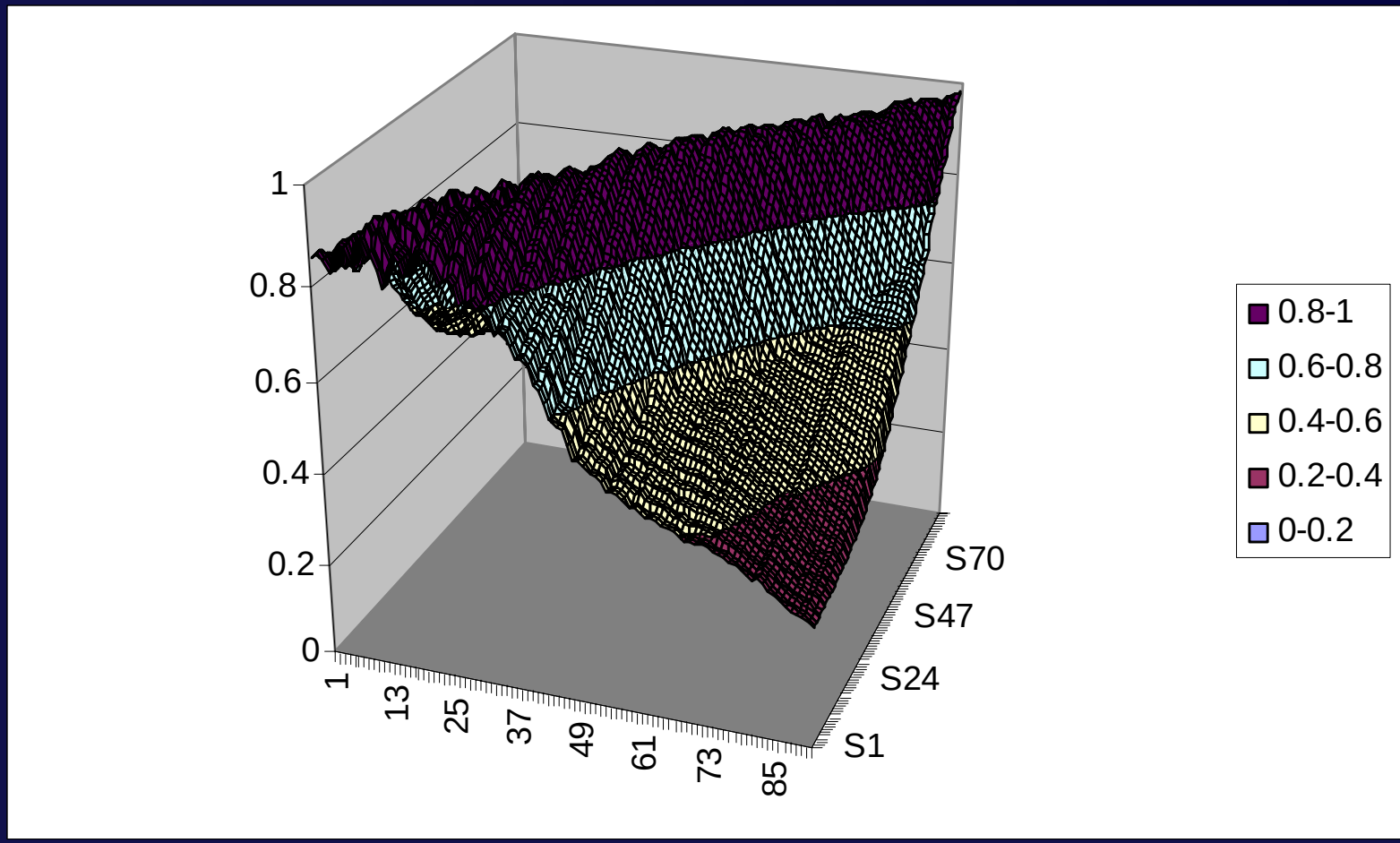
Correlation Results





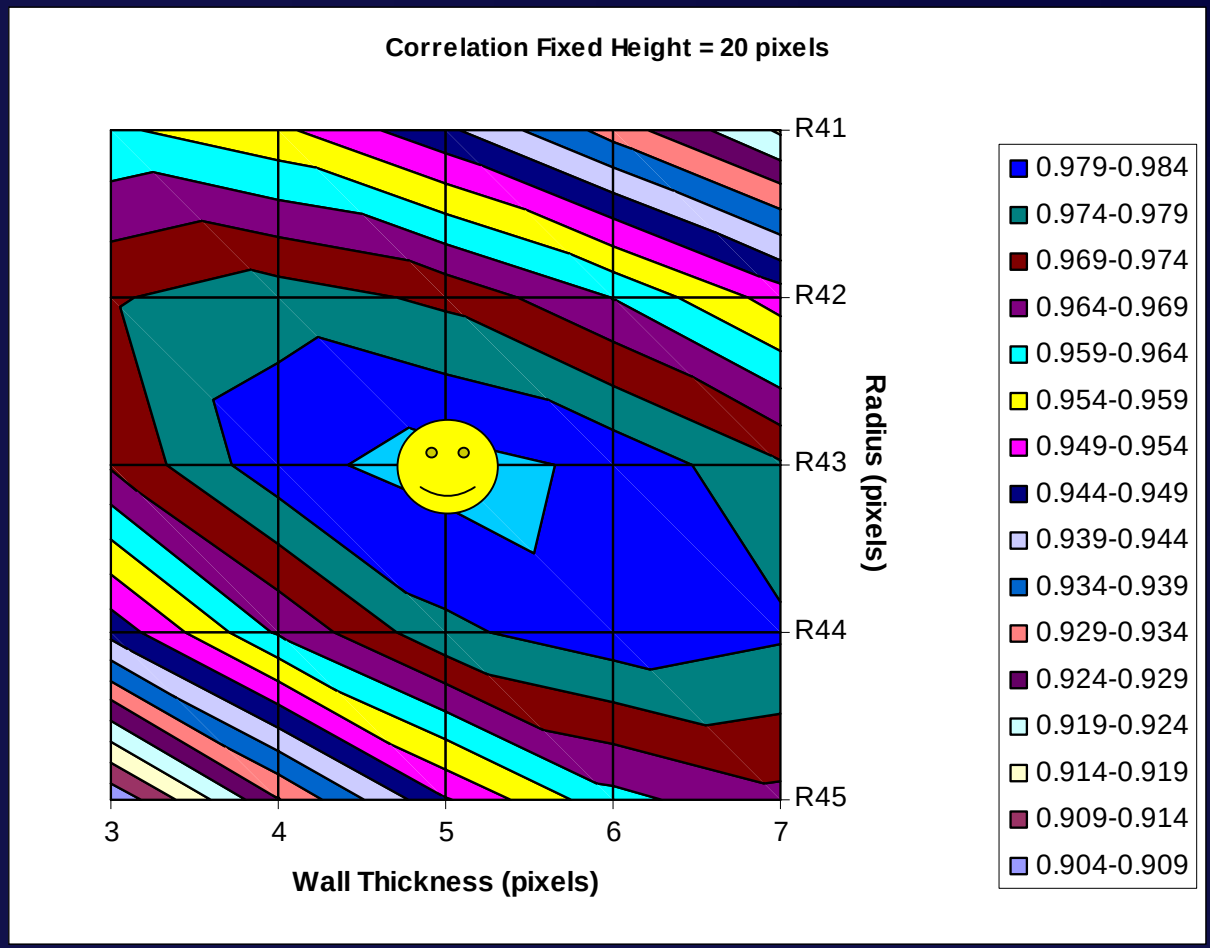
Correlation at 1:1 vs. correlation of image combined 4:1.

Correlation Results



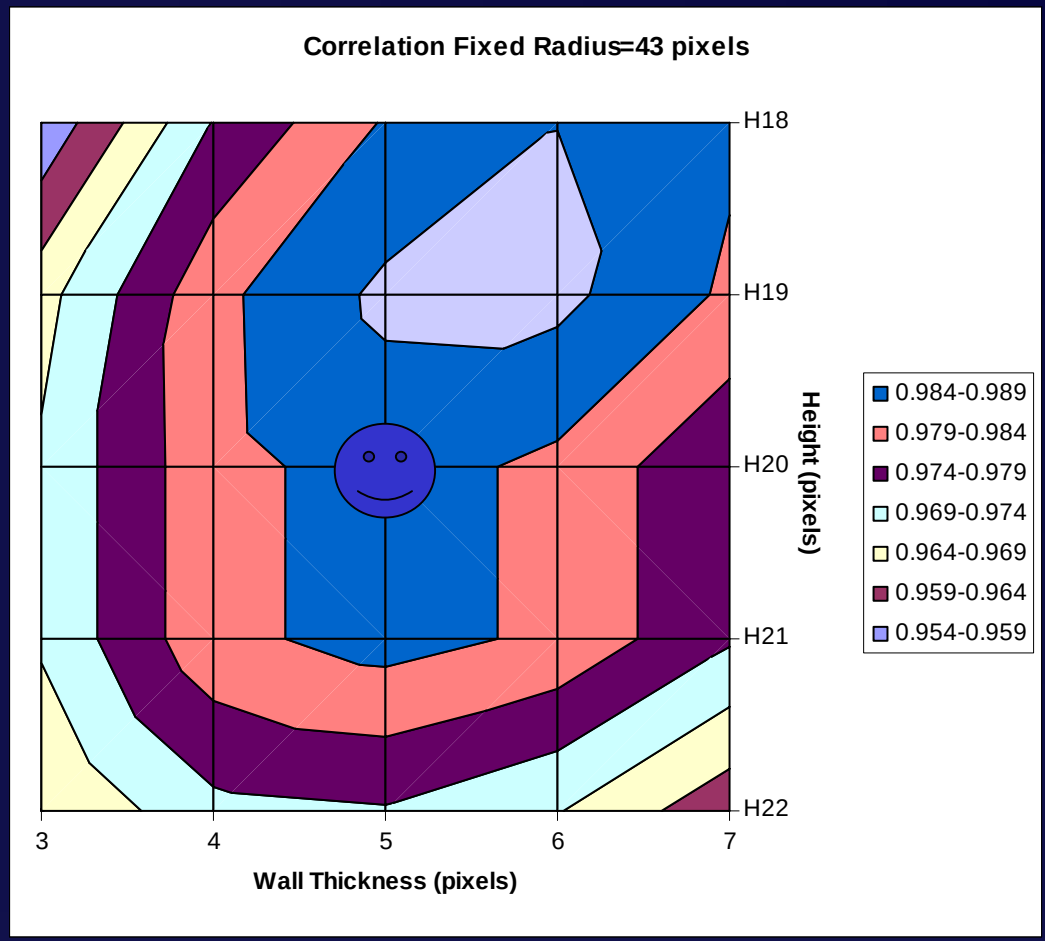
Correlation of disc (R43,H20,T5,A45) vs. same disc at angles from 0 to 90 degrees.

Combined Correlation



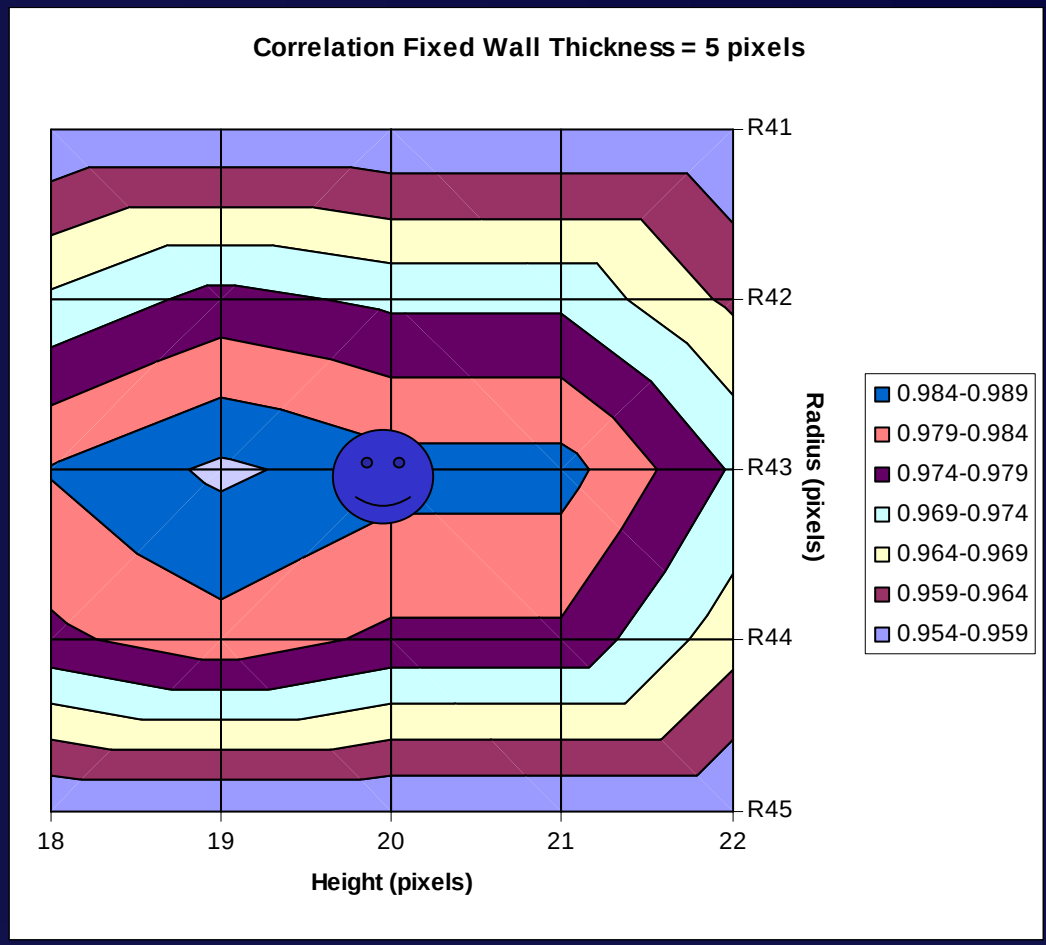
Correlation of disc (R43,H20,T5,A45) vs.
discs from R41-R45 and T3-T7 with H=20,A45 for all cases.

Combined Correlation



Correlation of disc (R43,H20,T5,A45) vs.
discs from H18-H22 and T3-T7 with R=43,A45 for all cases.

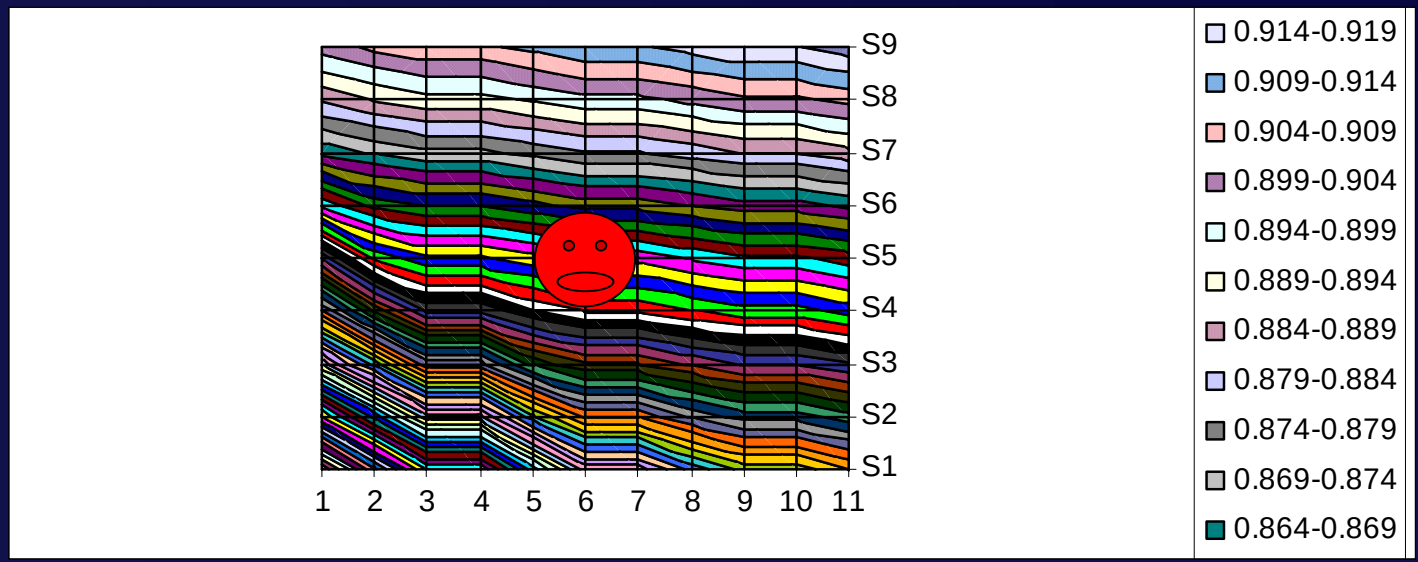
Combined Correlation



Correlation of disc (R43,H20,T5,A45) vs.
discs from R41-R45 and H18-H22 with T=5,A45 for all cases.

Correlation with FFT

Missed peak location (20x20 pixels) in analysis software☹
Need to re-run analysis☹☹☹ (on faster machine☺)



Correlation of disc (R43,H20,T5,A45) vs.
discs from R37-R49 and T1-T9 with H=20,A45 for all cases.
Correlation of disc (R43,H20,T5,A45) vs.
discs from H15-H25 and T1-T9 with R=43,A45 for all cases.
Correlation of disc (R43,H20,T5,A45) vs.
discs from R37-R49 and H15-H25 with T=5,A45 for all cases.

Project Setbacks

No access to the TEM…

- Documentation
  - SPS not as complete as hoped for in the acquisition mode
  - SRS incomplete for acquisition side
- Software
  - Had to create template images in software (and with Maya)
  - Spent lots of extra time on unnecessary file conversions
- Analysis
  - No design for acquisition assistance software
  - No prototype of acquisition assistance software
  - No images to test
  - Previous images had been saved as 8-bit instead of 12-bit

Future Work

- Analysis
  - Need further effort on “accuracy” of model matching
    - Slight shift in FFT-based processing
  - Need to find faster matching algorithms
- Software
  - Need to fine tune current analysis software
  - Need to design and develop acquisition assist
- Use
  - Need to get some real LDL images to understand the role of the sizes and shapes of LDL

Goal is to get funding to complete the software and do a complete analysis of shapes and sizes of VLDL, IDL, LDL, and maybe even HDL

Summary

- Cryogenic transmission electron microscopy has previously been used to measure lipoproteins
- This project…
  - Requirements and design for computer vision software to aid in acquisition and analysis of lipoprotein CTEM images
  - Analysis software has been developed to provide correlation of projected (2-D) images of 3-D objects
  - This software has been used to analyze artificially generated discoid objects to determine the analysis limits of the method
  - In general, correlation can determine sizes within a pixel or two

Matching by correlation, however, is very good at determining changes in small tilt angles (~5^{o}) so using small tilt angles seems to be a viable approach to determining geometric parameters of lipoproteins from CTEM images

Demonstration of Prototype Software

- Analysis Qualification Software
  - Generate grid of 91 discs (0^o to 90^o rotation)
  - Process against 91 individual discs (also 0^o to 90^o rotation)
  - Graphs of correlation effects

- Image Processing Analysis Software
  - Read in Maya image file
    - Process against a specified disc (radius, height, thickness, angle=45^o)
    - Detects best match for specified disc
  - Read in TEM image of blood
    - Process against a specified disc (radius, height, thickness, angle= 0^o)
    - Detects best match for specified disc