

Optimizing Image Acquisition

Getting the most out of your microscope and detectors and the math you need to do it

John Rubinstein

Molecular Structure and Function Program The Hospital for Sick Children Research Institute

Departments of Biochemistry and Medical Biophysics The University of Toronto





Outline

Image optimization

- Compare F20 to Titan Krios
- Mysterious optimization for DDDs

Math for DDDs

- Exposure weighting
- Introduce problem of frame alignment

<u>Refresher: waves and Fourier transforms</u>

- Padding and truncating in Fourier Space
- Fourier shift theorem
- Cross correlation functions

Applications of FFTs:

- Downsampling images
- (matrix multiplication review)
- Aligning whole frames
- Shifting frames
- Aligning individual particles

Getting good images

The NRAMM website...

2012 Worksh	Lectures × +				
scripps.edu/201	-workshop-lectures/			ê	
HSC VPN	📄 President's Choice 🔮 https://www.google 👩 Google Scholar 🗧 Home - PubMed 🦛 Wheel-Trans	s 🍈 LaTeX/Mathematics 👍 Google Drive 🔲 Rubinstein Laborat 🚺 Ge	ogle Caler	ndar	
×	Spotiton: A new approach to EM specimen preparation	l ilak jain	N ²		
Toolbar ail	DOLORS: Versatile Strategy for Internal Labeling and Domain Localization in Electron Microscopy	lan MacRae			
N nt's Choi	New innovations for capturing macromolecules	Debbie Kelly		1	
Scholar PubMe	Panel Discussion	David DeRosier (Chair)			
rans lathem Drive bin Labo	Opening windows into the cell: Focused ion beam micromachining of eukaryot cells for cryo-electron tomography.	otic Elizabeth Villa		N	
Calendar Menu Fags	Day 3: Tuesday November 13				
Book	Title	Speaker	Audio	Slides	;
kmark	Introduction and new approaches	Wah Chiu		₹	
irefox	Optimizing image acquisition	John Rubinstein		₹	
uct Biol microsc citations re	Direct Detectors Forum: Short contributions from people who have real life experience with these instruments.	David Agard (Discussion Leader)Yifan Cheng (K2),			
nce inet - G		Richard Henderson 🔁 / Sjors Scheres (Falcon), 🖻			

http://nramm.scripps.edu/2012-workshop-lectures/

Use your microscope appropriately...

	Tecnai F20	Titan Krios
Parallel	Use C2 aperture and lens setting that minimizes beam divergence	3rd Condensor Lens
Avoiding Lens Hystersis	Use over-focused diffraction for search mode	Constant power lenses
Stage Side Entry Cryoholder		Cryo-autoloader
Voltage	200 kV	300 kV

F20/Titan Krios cost analysis

Titan Krios/DDD: USD \$5M Tecnai F20/DDD: USD \$2M

Difference:	USD	\$3M
-------------	-----	------

"I think my time is worth ~£20/hr" - Richard Henderson (2001)

£20/hr in 2001 \simeq £29/hr in 2014 (£1 \simeq USD\$1.68)

≃ \$49/hr in 2014

62,000 hours of Richard's time "Official" work week = 35 h (34 years with Richard)

"Machines don't make discoveries, people do." - Lewis Kay Acknowledgments:

Tim Grant (JFRC) Alexis Rohou (JFRC) Niko Grigorieff (JFRC) Jianhua Zhao (Toronto) Samir Benlekbir (Toronto)

Thallous chloride crystal - 25 kx magnification setting



d=3.842 Å

FT of thallous chloride crystal image



Average of many thallous chloride FTs



Fitting of measured radius to an ellipse



FT of thallous chloride crystal image



Corrected FT(sinc interpolations)



Thallous chloride crystal



Corrected thallous chloride crystal



Average of many corrected thallous chloride FTs



Anisotropic magnification affects CTF estimation



- Anisotropic magnification appear different (worse) at low magnification
- Will look like objective lens astigmatism in power spectra

Easy way to check for anisotropic magnification (Jianhua Zhao)



Easy way to check for anisotropic magnification (Jianhua Zhao)



Is the problem widespread? (Yifan Cheng/Jianhua Zhao)



Math for DDDs

Signal to Noise ratio in averages and frames



Average of 30 frames: 30 e⁻/Å²

Individual frame 1 e⁻/Å²

Exposure weighting



Exposure weighting

The resolution dependence of optimal exposures in liquid nitrogen temperature electron cryomicroscopy of catalase crystals Baker *et al*.

Journal of Structural Biology 169 (2010) 431–437

An even more sophisticated approach would be to use the optimal exposures measured here to calculate weighted averages of frames in order to maximize the SNR at each spatial frequency.

Publication	Conditions	Conclusion	
Veesler <i>et al.</i> (2013) <i>JSB</i> 184, 193-202	200 kV, 20.6 e ⁻ /Å ² , ~4-6 Å, groups of frames	small effect	
Scheres (2014)ELife 3:e03665.		effect	
Wang <i>et al.</i> (2014) <i>Nat Comm</i> 5:5808	Baker <i>et al.</i> 2010 measured values + 30 %	effect	

Drift of movie frames



Sources of movement:

- Specimen stage drift
 - Long exposures necessary for Gatan K2 summit in counting mode (>5 sec)
 - Side entry cryoholders may have drift rates of ~1 Å/s
- Beam-induced movement
 - May cause shift of whole frame
 - May not be uniform within an image
 - Harder problem to solve

Waves and FFTs

Representing waves a vectors



The FT represents functions in terms of waves



Shifting waves causes a phase change



Phase change of Fourier components from shifting



Shifting in real space causes phase changes in Fourier space

Resolution encoded by different pixels in a FFT



Manipulating FTs: truncating in Fourier space



Truncating in Fourier space leads to downsampling in Real space

Manipulating FTs: padding in Fourier space



Padding in Fourier space leads to interpolation in Real space

Two dimension Fourier transforms



- The FT of real functions (e.g. images) are Hermitian: for every point (a+bi) there is a corresponding point (a-bi)
- For an N \times N pixel image, Fourier transform is N/2+1 \times N
- The positive Nyquist and negative Nyquist values are the same

Two dimension Fourier transforms



Phase change in 2D FFT upon shifting and image



$$F_{shifted} = F_{unshifted}(\cos\phi + i\sin\phi)$$
$$\phi = k_x(j) \cdot \Delta x \frac{2\pi}{N} + k_y(j) \cdot \Delta y \frac{2\pi}{N}$$

where Δx and Δy are the x and y shifts, respectively.

N is the extent in pixels in both the x and y direction of the $N\times N$ image.

 $k_x(j)$ and $k_y(j)$ are the distance of the Fourier component from the origin in the k_x and k_y directions, respectively.

Applying knowledge of FFTs to DDD images

Sometimes you may want to downsample your images



Ruskin, Yu, and Grigorieff (2013). JSB 184, 385-93.

Downsampling in Fourier space



Cross correlation functions



Aligning frames *Motioncorr*

Li ... Cheng (2013). Nat Methods 10, 584-90.

Matrix multiplication review

(A, B etc. represent matrices): AB \neq BA ABCD=A(B(CD))

The least squares method for aligning frames

Li ... Cheng (2013). Nat Methods 10, 584-90.



- Define Frame 1 as "unshifted" (0,0)
- Calculate vectors (*xshift*, *yshift*) that bring two frames into register
- Can use cross correlation to estimate 6 unique vectors for 4 frame movie:

Frame 1 vs Frame 2 Frame 1 vs Frame 3 Frame 1 vs Frame 4 Frame 2 vs Frame 3 Frame 2 vs Frame 4 Frame 3 vs Frame 4

Can calculate $(Z/2) \times (Z-1)$ cross-correlation functions for a movie with Z frames (e.g. 30 frame movie yields 435 CCFs)

The least squares method for aligning frames

 t_{NM} means true shift vector between frames N and M m_{NM} means measured shift vector (by cross correlation) between frames N and M

100		m ₁₂
1 1 0	t ₁₂	m ₁₃
1 1 1	t_{12}	m ₁₄
0 1 0	$\mathbf{t}_{23} = \mathbf{t}_{23}$	m ₂₃
0 1 1	134	m ₂₄
0 0 1		m ₃₄

Li...Cheng, Nature Methods

 $\begin{array}{l} m_{12} \!\!\simeq\!\! 1 \!\cdot\! t_{12} \!\!+\!\! 0 \!\cdot\! t_{23} \!\!+\!\! 0 \!\cdot\! t_{34} \\ m_{13} \!\!\simeq\!\! 1 \!\cdot\! t_{12} \!\!+\!\! 1 \!\cdot\! t_{23} \!\!+\!\! 0 \!\cdot\! t_{34} & \text{Once matrices are filled in standard} \\ m_{14} \!\!\simeq\!\! 1 \!\cdot\! t_{12} \!\!+\!\! 1 \!\cdot\! t_{23} \!\!+\!\! 1 \!\cdot\! t_{34} & \text{Iinear algebra can be used to find} \\ m_{23} \!\!\simeq\!\! 0 \!\cdot\! t_{12} \!\!+\!\! 1 \!\cdot\! t_{23} \!\!+\!\! 0 \!\cdot\! t_{34} & \text{values that best fit the data for} \\ m_{24} \!\!\simeq\!\! 0 \!\cdot\! t_{12} \!\!+\!\! 1 \!\cdot\! t_{23} \!\!+\!\! 1 \!\cdot\! t_{34} & t_{12}, t_{23}, t_{34} \\ m_{34} \!\simeq\!\! 0 \!\cdot\! t_{12} \!\!+\!\! 0 \!\cdot\! t_{23} \!\!+\!\! 1 \!\cdot\! t_{34} \end{array}$

Improvements to the least-squares approach (I)

 Subpixel accuracy for cross correlation peaks (padding in Fourier space leads to interpolation in Real space)



Improvements to the least-squares approach (II)

Minimum interval between frames
 (cross correlation functions for subsequent frames might have maxima too
 close to the origin to be reliable)

Improvements to the least-squares approach (III)

 Throw away equations with high residuals m_{NM} means measured shift vector (by cross correlation) between frames N and M
 c_{NM} means calculated shift vector between frames N and M

Shifting images in Fourier space

$$(a' + b'i) = (a + bi)(\cos \phi + i \sin \phi)$$

$$\phi = k_x(j) \cdot \Delta x \frac{2\pi}{N} + k_y(j) \cdot \Delta y \frac{2\pi}{N}$$

$$= (3)(-1.2)\frac{2\pi}{12} + (4)(2.3)\frac{2\pi}{12}$$

Aligning individual particles *alignparts_Imbfgs* Rubinstein and Brubaker (2014). *arXiv* 1409.6789

Global optimization for aligning individual particles

Fourier transform of individual particle movie Z frames J Fourier components (F_{jz})

Correlation of Fourier transforms

Fourier transform 1 Fourier transform 1

We only need to consider the *real* part of $F_{1j}F^*_{2j}$ because for every term:

 $(a_1 + b_1 i)(a_2 - b_2 i) = a_1 a_2 + b_1 b_2 + (a_2 b_1 - a_1 b_2)i$ There is a corresponding term:

$$(a_1 - b_1 i)(a_2 + b_2 i) = a_1 a_2 + b_1 b_2 - (a_2 b_1 - a_1 b_2)i$$

- Find an objective function that, when maximized, maximizes the sum of the correlations of each shifted frame with the sum of the shifted frames.
- Equivalently: find an objective function that, when minimized, maximizes the sum of the correlation of each shifted frame with the sum of the shifted frames.

Global optimization for aligning individual particles

Optimization of functions of many variables

Marcus Brubaker

(first of many important contributions)

Optimization of functions of many variables

 The function we are trying to optimize has 2 variables for every frame (x-shift and y-shift) **Ο(θ) dU(θ)/∂θ**_a Ο(θ) optimum (minimum) Conjugate gradients Broyden-Fletcher-Goldfarb-Shanno θ_a $\frac{\partial O(\Theta)}{\partial x_a} = -Re \sum_{i=1}^{J} \frac{2\pi i k_x(j)}{N} \left[F_{ja} S_{ja} \sum_{z=1}^{Z} F_{jz}^* S_{jz}^* - F_{ja}^* S_{ja}^* \sum_{z=1}^{Z} F_{jz} S_{jz} \right]$ $\frac{\partial O(\Theta)}{\partial y_a} = -Re\sum_{i=1}^J \frac{2\pi i k_y(j)}{N} \left[F_{ja} S_{ja} \sum_{z=1}^Z F_{jz}^* S_{jz}^* - F_{ja}^* S_{ja}^* \sum_{z=1}^Z F_{jz} S_{jz} \right]$

alignparts_Imbfgs.f90

Rubinstein and Brubaker (2014). arXiv 1409.6789

Improvement #1: disfavour unlikely trajectories

Second order smoothing

Rubinstein and Brubaker (2014). arXiv 1409.6789

Derivatives of penalty function

$$\frac{\partial P(\Theta)}{\partial x_{a}} = \begin{cases} 2\lambda \left(x_{a} - 2x_{a+1} + x_{a+2}\right), & a = 1, \\ 2\lambda \left(-2x_{a-1} + 5x_{a} - 4x_{a+1} + x_{a+2}\right), & a = 2, \\ 2\lambda \left(x_{a-2} - 4x_{a-1} + 6x_{a} - 4x_{a+1} + x_{a+2}\right), & a \in [3, Z-2], \\ 2\lambda \left(x_{a-2} - 4x_{a-1} + 5x_{a} - 2x_{a+1}\right), & a = Z-1, \\ 2\lambda \left(x_{a-2} - 2x_{a-1} + x_{a}\right), & a = Z. \end{cases}$$

$$\frac{\partial P(\Theta)}{\partial y_{a}} = \begin{cases} 2\lambda \left(y_{a} - 2y_{a+1} + y_{a+2}\right), & a = 1, \\ 2\lambda \left(-2y_{a-1} + 5y_{a} - 4y_{a+1} + y_{a+2}\right), & a = 2, \\ 2\lambda \left(y_{a-2} - 4y_{a-1} + 6y_{a} - 4y_{a+1} + y_{a+2}\right), & a \in [3, Z - 2], \\ 2\lambda \left(y_{a-2} - 4y_{a-1} + 5y_{a} - 2y_{a+1}\right), & a = Z - 1, \\ 2\lambda \left(y_{a-2} - 2y_{a-1} + y_{a}\right), & a = Z. \end{cases}$$

Rubinstein and Brubaker (2014). arXiv 1409.6789

Second order smoothing

Improvement 2: Enforce local correlation

Local averaging

Rubinstein and Brubaker (2014). arXiv 1409.6789

Comparison of drift correction methods

Approach	Correlation	Smoothing	Advantages/ Disadvantage
Least squares (<i>Motioncorr</i>)	Noisy images to noisy images	Over-determined problem (fitted trajectory, local correlation possible)	Over-determined/ low signal-to- noise in comparisons
Polishing (Relion)	Noisy images to map projections	Linear fit, rolling averages, enforce local correlation	Map projection v. high SNR/map projection may not match image
Non-global iterative	Noisy images to sums of noisy images	Fitted trajectory, enforce local correlation	Sum of images high SNR/No built in regularization
Global optimization (Alignparts_Imbfgs)	Noisy images to sums of noisy images	Penalize changes in trajectory, enforce local correlation	Non-linear trajectories/Map projections have higher SNR

- Thermoplasma acidophilum 20S proteasome (Kay lab)
- 1 grid frozen on a FEI Vitrobot in ethane/propane
- FEI F20 at 200 kV, Gatan 626 side entry cryoholder
- 30 µm C2 aperture
- Gatan K2 Summit in super-resolution mode
- movies captured at 5 e⁻/pix/sec
- 1.45 Å/pixel
- 30 frames, 15 seconds, 1 e⁻/Å²/frame
- 60 Movies (short afternoon session)
- downsampled by Fourier truncation (Alexis Rohou)
- local movement corrected with alignparts_Imbfgs.f90
- exposure weighting with alignparts_Imbfgs.f90
- magnification anisotropy correction in particles and CTF parameters
- 13,974 particles with D7 symmetry

Putting it all together (Michael Latham, Samir Benlekbir)

3.8 Å resolution

Apples to Oranges Comparisons

Prospects

Microscopes:

- 300 kV could be better than 200 kV:
 - Better DDD response
 - Less Ewald sphere curvature
 - Less charging effects
 - (Titan/JEOL 3200) more parallel beam
- 300 kV not currently needed for most projects

What we could use:

- Improved detectors (higher DQE at high resolution)
- Improved algorithms for conformational separation
- Improved spatial coherence
- Improved single particle motion correction
- Improved specimen preparation (prevent motion)

With the existing instruments and algorithms:

- Atomic resolution structures of homogenous specimens
- Sub-nanometer resolution structures of heterogeneous specimens
- Fewer blobs, fewer incorrect structures

Acknowledgements:

Current members: Samir Benlekbir Stephanie Bueler Michael Latham (Kay Lab) Mohammad Mazhab Jafari Jason Koo (Howell Lab) Martin Smith Jianhua Zhao Dan Schep Anna Zhou

Past members:

Lindsay Baker Shawn Keating Wilson Lau Chelsea Marr Nawaz Pirani Jana Tuhman

Collaborators Lewis Kay (Toronto) Marcus Brubaker (TTI)

Niko Grigorieff (JFRC) Alexis Rohou (JFRC) Tim Grant (JFRC) Yifan Cheng (UCSF)

alignparts_Imbfgs for β-galactosidase (Sjors Scheres)

Rubinstein and Brubaker (2014). arXiv 1409.6789