

Batch Processing of Tomograms with IMOD

Batch Processing Can Be Useful for a Wide Range of Tilt Series

- “Routine” plastic section tilt series can be fully reconstructed automatically with 80-95% success rate
- Some kinds of cryoET data can also be reconstructed automatically up to the final trimming
- Both the method of processing and the batch interface are designed to allow interactive steps to be intermixed with batch processing
 - Even difficult data that require intervention at several points can be handled more efficiently by processing them through the batch interface

Major Developments to Support Automated Processing

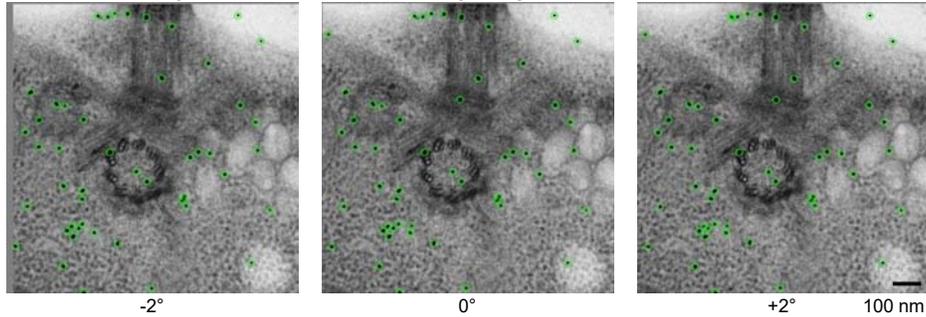
- Automatic seed model selection to get a well-distributed set of fiducials suitable for tracking

Automatic Seed Model Generation

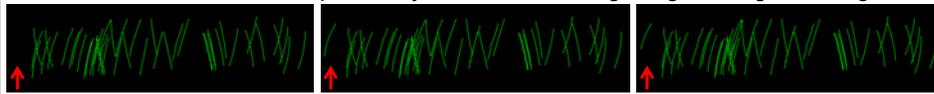
The goal is to find a well-distributed set of seed points suitable for tracking

Automatic Seed Model Generation

1. Find all gold beads on three starting images near 0°



2. Track the beads independently from each starting image through 11 images



Side view of tracked contours (Z is projection #)

3. Use tracking information to sort onto two surfaces, if appropriate



Side view of bead positions in 3D

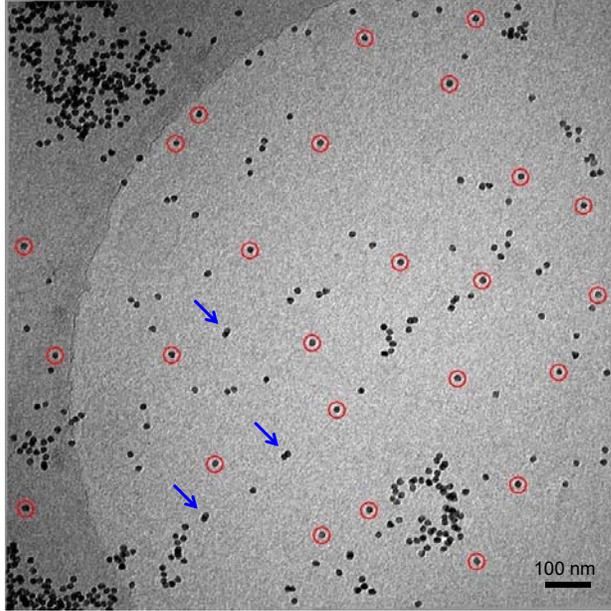
Automatic Seed Model Generation

The goal is to find a well-distributed set of seed points suitable for tracking

1. Find all gold beads on 3 starting images near 0°
2. Track the beads independently from each starting image through 11 images
3. Use tracking information to sort onto two surfaces, if appropriate
4. Score beads based on how well and consistently they tracked
5. Select a well-distributed subset of highest-scoring beads to achieve the desired density

Seed Model for Cryo Tilt Series

- Requested 25 beads
- It avoided clusters, distinct beads too close to others, and elongated (overlapping) beads



Tilt series of frozen-hydrated mammalian cell infected with bovine papilloma virus, from Mary Morphew

Major Developments to Support Automated Processing

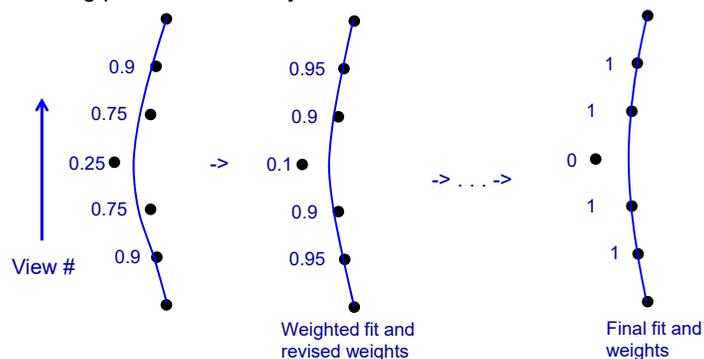
- Automatic seed model selection to get a well-distributed set of fiducials suitable for tracking
- Improvements in bead centering by “Sobel filtering” to enhance bead edges
 - For plastic section data, Sobel centering reduced error by a mean of 20% (0 – 0.25 pixel, n = 15)
 - For cryo data, it reduced error by a mean of 10% (0 – 0.1 pixel, n = 15)

Major Developments to Support Automated Processing

- Automatic seed model selection to get a well-distributed set of fiducials suitable for tracking
- Improvements in bead centering by “Sobel filtering” to enhance bead edges
 - For plastic section data, Sobel centering reduced error by a mean of 20% (0 – 0.25 pixel, n = 15)
 - For cryo data, it reduced error by a mean of 10% (0 – 0.1 pixel, n = 15)
- Robust fitting to reduce/eliminate effects of inevitable bad points

Robust Fitting – A Substitute for Fixing Bad Points

- Robust fitting follows these steps:
 - Fit as usual and get residual for each point
 - Give each point a weight based on how extreme the residual is
 - Fit again, minimizing *weighted* sum of errors
 - Get new residuals, new weights, repeat until convergence
- About 2% of points end up with zero weights, ~5% have weights < 0.5 weight
- If there is a sufficient excess of points, this should give as good an alignment as correcting positions manually

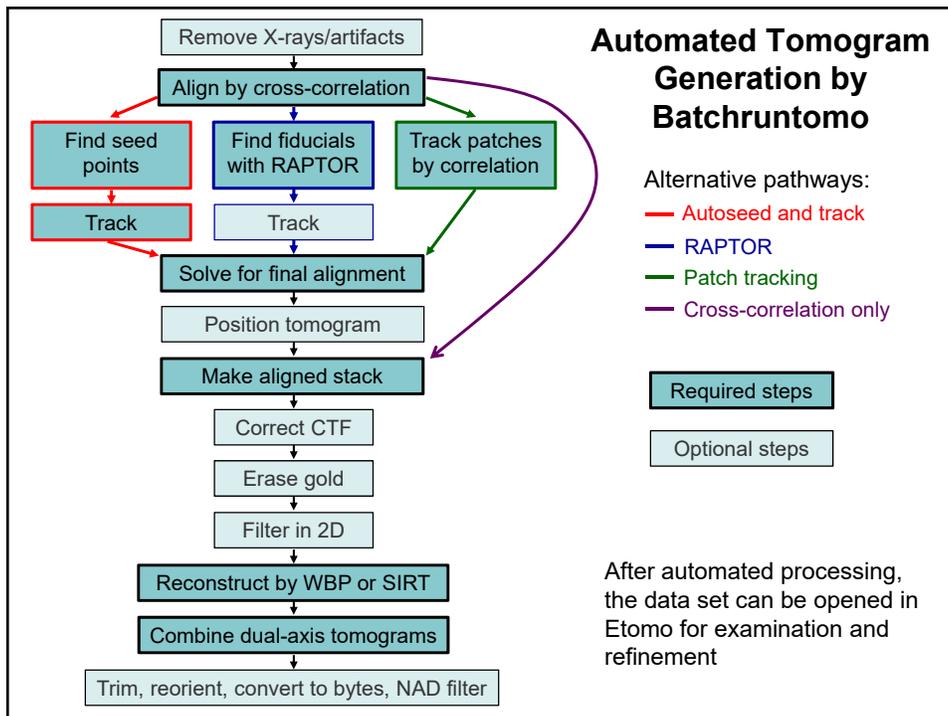
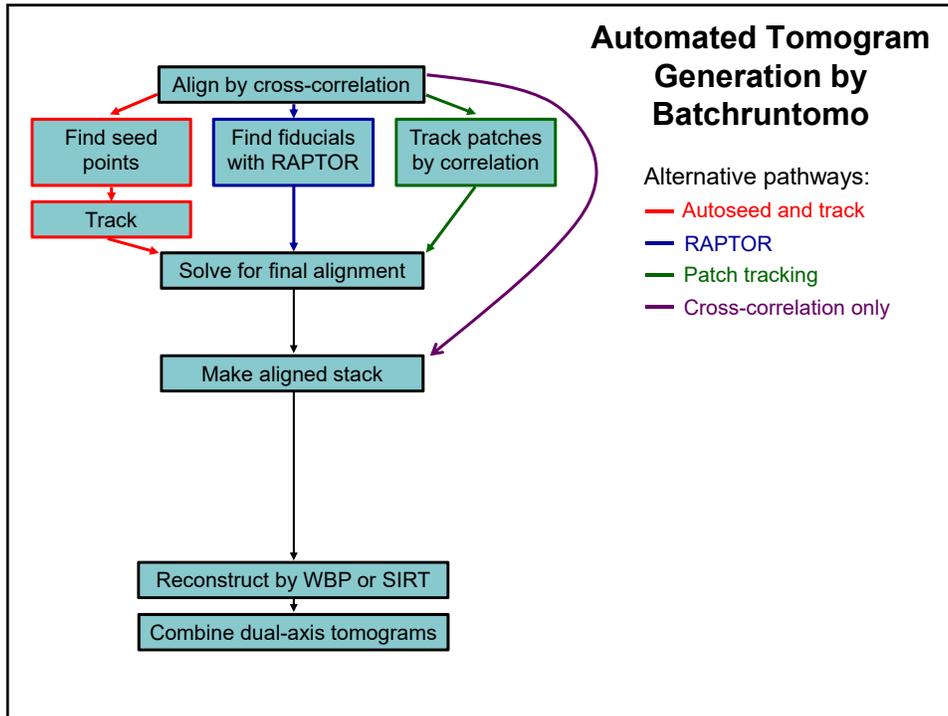


Major Developments to Support Automated Processing

- Automatic seed model selection to get a well-distributed set of fiducials suitable for tracking
- Improvements in bead centering by “Sobel filtering” to enhance bead edges
 - For plastic section data, Sobel centering reduced error by a mean of 20% (0 – 0.25 pixel, n = 15)
 - For cryo data, it reduced error by a mean of 10% (0 – 0.1 pixel, n = 15)
- Robust fitting to reduce/eliminate effects of inevitable bad points
- Detection of structure in specimen, for automating tomogram positioning, the alignment of dual-axis tomograms, and trimming of stained specimens

Major Developments to Support Automated Processing

- Automatic seed model selection to get a well-distributed set of fiducials suitable for tracking
- Improvements in bead centering by “Sobel filtering” to enhance bead edges
 - For plastic section data, Sobel centering reduced error by a mean of 20% (0 – 0.25 pixel, n = 15)
 - For cryo data, it reduced error by a mean of 10% (0 – 0.1 pixel, n = 15)
- Robust fitting to reduce/eliminate effects of inevitable bad points
- Detection of structure in specimen, for automating tomogram positioning, the alignment of dual-axis tomograms, and trimming of stained specimens
- A framework, script, and interface for running tracking, alignment, and reconstruction automatically



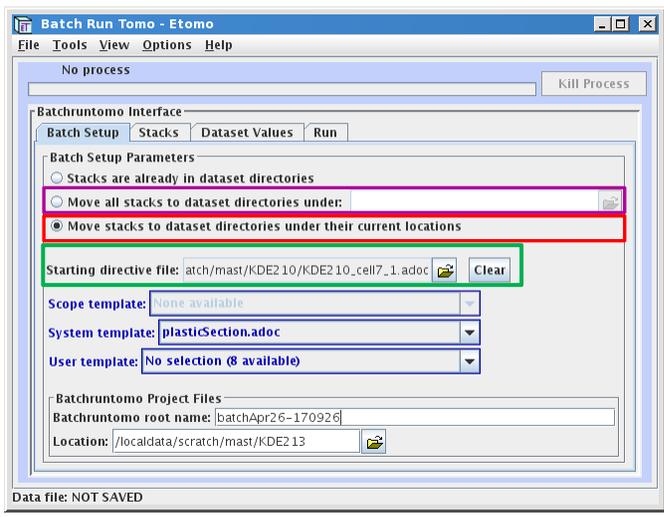
It's All Done With Directives

- Directives to control Batchruntomo are key-value pairs in a text file
- Can be options to programs, which can be handled generically
 - `comparam.track.beadtrack.RoundsOfTracking = 4`
- Can be directions interpreted by Batchruntomo or Etomo
 - `runtime.BeadTracking.any.numberOfRuns = 2`
- These are the same directives used in templates

Batch Interface: Initial Setup

- Stacks can be moved from where they are into separate directories **under one directory**, or **under their current location(s)**

- A **directive file** from a previous run or a processed data set can be used as starting point
- **Template files** can be chosen



Batch Interface: Setting Parameters (Top)

- It allows all the choices for **alignment method**
- The **stretching solution and local alignments** are **enabled** by the choices here, but Batchrun2010 will decide whether the number of fiducials is sufficient for each

Global Dataset Values

Image distortion file:

Mag gradient file:

Remove X-rays

Manual replacement model: Make in 3dmod

Coarse aligned stack binning - single frame: Montage:

Alignment Method

Autoseed and track Patch tracking
 Raptor and track Fiducialless

Bead size (nm): Target number of beads:

Patch tracking size: Break contours into pieces

Enable distortion (stretching) in alignment

Use local alignments

Parameter Settings for CryoET (Top)

- Using binning for the coarse aligned stack is a common way to improve bead tracking with a small pixel size and/or low SNR

Coarse aligned stack binning - single frame: Montage:

- **Stretching** would generally not be enabled but **local alignments** can be helpful for large areas with enough beads

Bead size (nm): Target number of beads:

Patch tracking size: Break contours into pieces

Enable distortion (stretching) in alignment

Use local alignments

Batch Interface: Setting Parameters (Middle)

- With **positioning for a plastic section**, the thickness is optional
- But thickness is required when **finding beads in 3D** for erasing gold
- The “**calculated thickness**” is usually best for plastic sections
 - It is based on distance between beads on surfaces if no positioning is done
 - The fallback is used if either analysis gives < 40% of this thickness

Do positioning for: Plastic section Cryo sample
 Tomogram thickness: Sample has gold beads Bead size (nm):
 Aligned stack binning:
 Correct CTF
 Autofit range and step
 Defocus:
 Fit every image
 Reconstruction
 Back-projection Thickness total (unbinned pixels):
 SIRT Thickness total (binned pixels):
 Both Calculated thickness (unbinned pixels):
 Leave iterations: With fallback:
 Scale to integers Plus (optional):

Parameter Settings for CryoET (Middle)

- Positioning must be done with the **Cryo sample** option and a **thickness** must be entered
 - If patch tracking is used and there are gold beads, this should be indicated

Do positioning for: Plastic section Cryo sample
 Tomogram thickness: Sample has gold beads Bead size (nm):

- CTF correction can be done by having Ctfplotter **autofit to ranges of angles** or **fit every image**
 - Voltage, spherical aberration, and noise configuration file must be set in a template or starting batch file

Correct CTF Erase gold
 Autofit range and step Use fiducial model
 Defocus: Find beads in 3D
 Fit every image Tomogram thickness (pixels):

Parameter Settings for CryoET (Middle)

- A [total reconstruction thickness](#) must be entered if positioning is not done

Reconstruction

Back-projection

SIRT

Both

Leave iterations:

Scale to integers

Thickness total (unbinned pixels):

Thickness total (binned pixels):

Calculated thickness (unbinned pixels):

With fallback:

Plus (optional):

Batch Interface: Setting Parameters (Bottom)

- With plastic sections, [postprocessing](#) can trim the tomogram if “[Find plastic section limits](#)” is checked, and convert to bytes if “[Fraction of Z slices](#)” is checked
- For cryo, [postprocessing](#) will just rotate the full tomogram so slices are in X/Y
- In the Datasets table, you can press [Open](#) to open a complete Dataset Values dialog specific to one data set
 - “Set” indicates there are specific values for a data set

Postprocessing

Postprocess with trimvol

Find plastic section limits and add:

Fraction of Z slices to analyze:

Datasets

#	Stack	Specific Values
1	KDE213_cell1_1a.mrc	Open
2	KDE213_cell1_2a.mrc	Open Set
3	KDE213_cell1_3a.mrc	Open
4	KDE213_cell1_4a.mrc	Open
5	KDE213_cell2_3a.mrc	Open
6	KDE213_cell2_4a.mrc	Open
7	KDE213_cell2_5a.mrc	Open
8	KDE213_cell2_6a.mrc	Open

Batch Interface: Running

- In the **Resources** area, a table for computers with CPUs opens if **Use multiple CPUs** is selected, and table for GPUs opens if **Parallel GPUs** is checked
- The default is to run every set to completion, but you can **stop, interact, and restart**

The screenshot shows the 'Batch Run Tomo - Etomo' window. The 'Resources' section contains two tables:

Computer	# CPUs	Load Average	GPU	# GPUs
bigfoot	3	0.0 0.09	bigfoot	1
eclipse	4	0.36 0.34	eclipse	1
druid	2	0.6 0.29		

Below the tables, 'CPU: 9' and 'GPU: 2' are displayed. The 'Run Actions' section has 'Use multiple CPUs' checked (highlighted in red), 'Local GPU' selected (highlighted in purple), and 'Parallel GPUs' checked (highlighted in pink). The 'Subset of steps to run' section has 'Stop after' checked (highlighted in green), 'CTF/gold detection' selected, and 'Finish CTF/gold' selected. The 'Datasets' table at the bottom shows 8 rows of data with columns for Stack, Status, Reached, Run, Set, Rec, and Log.

Batch Processing for CryoET: Running

- A common stopping point would be after CTF estimation and detection of gold
 - Both can be checked and then processing can be restarted to do CTF correction and gold erasing

This close-up shows the 'Run Actions' and 'Subset of steps to run' sections. 'Use multiple CPUs' is checked. 'Local GPU' is selected. 'Parallel GPUs' is unchecked. In the 'Subset of steps to run' section, 'Stop after' is checked, 'CTF/gold detection' is selected, and 'Finish CTF/gold' is selected. Below these are 'Run', 'Pause', 'Resume', and 'Reset' buttons. The 'Datasets' table at the bottom has two rows:

#	Stack	Status	Reached	Run	Open Set	Open Rec	Open Log
1	cryo1.mrc			<input checked="" type="checkbox"/>			
2	cryo2.mrc			<input checked="" type="checkbox"/>			

Batch Processing for CryoET: Running

- The status table shows the **current state** of the running job including the last step completed
- The Batchrun tomo log can be opened with the **Open Log** button as soon as it is available

Run Actions:

- Use multiple CPUs
- No GPU
- Local GPU
- Parallel GPUs

Subset of steps to run:

- Stop after
- Start from
- Tracking
- Fine alignment
- Fine alignment
- Positioning
- Aligned stack
- CTF/gold detection
- Finish CTF/gold
- Finished stack
- Reconstruction
- Enable starting from any step

Buttons: Run, Pause, Resume, Reset

Datasets:

#	Stack	Status	Reached	Run	Open Set	Open Rec	Open Log
1	cryo1.st	Running	Align	<input checked="" type="checkbox"/>			
2	cryo2.mrc			<input checked="" type="checkbox"/>			

Batch Processing for CryoET: Running

- All jobs are taken to the selected step for stopping
- Data sets can be opened in the reconstruction interface with **Open Set** when processing is stopped

Run Actions:

- Use multiple CPUs
- No GPU
- Local GPU
- Parallel GPUs

Subset of steps to run:

- Stop after
- Start from
- Tracking
- Fine alignment
- Fine alignment
- Positioning
- Aligned stack
- CTF/gold detection
- Finish CTF/gold
- Finished stack
- Reconstruction
- Enable starting from any step

Buttons: Run, Pause, Resume, Reset

Datasets:

#	Stack	Status	Reached	Run	Open Set	Open Rec	Open Log
1	cryo1.st	Stopped	CTF/gold	<input checked="" type="checkbox"/>			
2	cryo2.st	Stopped	CTF/gold	<input checked="" type="checkbox"/>			

Batch Processing for CryoET: Running

- After interactive work, you can start from the **next step** or repeat an earlier step
 - You can skip ahead, over steps done interactively, with **Enable starting from any step**

Run Actions

- Use multiple CPUs
- No GPU
- Local GPU
- Parallel GPUs

Subset of steps to run

- Stop after
 - Tracking
 - Fine alignment
 - Positioning
 - Aligned stack
 - Finish CTF/gold**
 - Reconstruction
- Start from
- Fine alignment
- Positioning
- Aligned stack
- Reconstruction

Enable starting from any step

Run Pause Resume Reset

#	Stack	Status	Reached	Run	Open Set	Open Rec	Open Log
1	cryo1.st	Stopped	CTF/gold	<input checked="" type="checkbox"/>			
2	cryo2.st	Stopped	CTF/gold	<input checked="" type="checkbox"/>			

Batch Processing for CryoET: Running

- When a data set is finished, the **Run** button is turned off
 - You have to turn Run back on for at least one set to enable starting from a chosen point
- The tomogram can be opened in 3dmod with **Open Rec** when one is available

Run Actions

- Use multiple CPUs
- No GPU
- Local GPU
- Parallel GPUs

Subset of steps to run

- Stop after
 - Tracking
 - Fine alignment
 - Positioning
 - Aligned stack
 - Finish CTF/gold
 - Reconstruction
- Start from
- Fine alignment
- Positioning
- Aligned stack
- Reconstruction

Enable starting from any step

Run Pause Resume Reset

#	Stack	Status	Reached	Run	Open Set	Open Rec	Open Log
1	cryo1.st	Done		<input type="checkbox"/>			
2	cryo2.st	Done		<input type="checkbox"/>			

Customizing Directives

- Almost every program option can be controlled in batch processing with a directive
- The master table of directives is available with:
imodhelp -d directives
 - But for programs listed there, any option can be specified with a comparand, not just the ones in the table
- To do CTF correction with the batch interface, you will need to supply microscope parameters as directives
 - A Scope template is a logical way to do this

Two Ways to Make a Scope Template for CTF Correction

1. Through the template editor:
 - In any data set open in Etomo, select File – Templates – Save Scope Template
 - Check “Show unchanged” in the Control Panel section
 - Turn on “Arguments to Copytomocoms” to open that section if necessary
 - Select “Microscope voltage” and “Spherical aberration” and fill in values
 - Select “Noise configuration file” and choose the file
 - Turn off other selected items if any – check “Show included” to see them better
 - Press the Save button and save it to a file with extension “.adoc”
2. With a text editor:
 - Make a file with extension “.adoc” with these entries modified for your case:
setupset.copyarg.voltage=300
setupset.copyarg.Cs=2.0
setupset.copyarg.ctfnoise=full_path_to_noise_file
- In either case, the file goes in /usr/local/ImodCalib/ScopeTemplate
 - Or more generally, \$IMOD_CALIB_DIR/ScopeTemplate for your setup