
X3dPreprocess

Program and notes by James Conway, 26-Mar-2001.

Program Version: 1.8.2n (19-Oct-2000)

Contact me:

Laboratoire de Microscopie Electronique Structurale
 Institut de Biologie Structurale
 41 rue Jules Horowitz
 38027 Grenoble Cedex 1
France
 Telephone: +33-(0)476-884568
 Fax: +33-(0)476-885494
 Email: James.Conway@ibs.fr

Contents:

1. Overview.....	1
2 Quick start.....	3
3 File formats.....	3
4 Image window.....	4
5 Hints.....	4
6 File menu.....	4
7 Options menu.....	5
8 Preprocess menu.....	5
9 Known problems.....	7
10 Future plans.....	7
11 Version history.....	7
12 References.....	10
13 Acknowledgements.....	10

1. Overview

An XWindows/Motif program in 'C' for selecting particles off scanned images and preprocessing them prior to determining orientations and calculating a three-dimensional reconstruction. You *must* be using an XWindows workstation to use this program. To run, type: `x3d`

The program has two phases of operation: selecting particles and preprocessing the selections. The first part is interactive in that an operator uses the mouse to click on particle centers to indicate which particles are to be preprocessed. The second part is automatic – the program applies a series of operations to copies of each selected area and writes the results to a packed image file, or a series of image files. The output file(s) will be used for subsequent reconstructions.

1.1 Input files

An input image is required to get started on the selection phase. This will usually be a digitized electron micrograph image containing many small particles. Once selections have been made, they may be saved in a special coordinate file (*.CRD) for future reference. This CRD file also includes the image filename and all the parameters used in preprocessing, and may be used as input to the program.

1.2 Output files

The particle origins and the preprocessing parameters can be saved in a coordinate file (*.CRD). This file can be read by the program so that selections or parameters can be modified or added to and preprocessing done on the new data.

The preprocessed images are saved into a single packed file, or as a series of files, all sharing a common filename prefix to which is appended a number, for example, IMG001.BP, IMG002.BP, IMG003.BP, etc. The starting number, and number of digits are both preprocessing parameters that can be set.

1.3 Selecting particles

Particles are selected by pointing the mouse near the particle center and clicking mouse button 1 (usually the left-most button) to make a selection. A circle is drawn at the specified radius to indicate the selection, and can be moved by dragging (pressing a mouse button within the circle, holding it down while moving the mouse, then releasing to leave the circle at the new position). The circle radius can be changed at any time – use the Options menu or the Radius button on the main display window – all selected circles will be changed.

Damaged particles or other ‘noise’ near good particles can be masked out by selecting them with mouse button 2 (usually the center button on a 3-button mouse). Alternatively, a previously ‘good’ selection can be changed to ‘bad’ by double-clicking on it with mouse button 2. A selection may also be deleted by double-clicking on the selection with mouse button 1, which brings up a dialog box offering the choices of deleting the selection, toggling the selection between ‘good’ and ‘bad’, or doing nothing.

The selections need to be made first so that particles close to each other can be masked out of neighbors background regions. Selections can be saved to a coordinate file at any time, and the coordinate file can be read back in to allow additional selections to be made.

An auto-centering option can be enabled to cause selections to be automatically centered on the target particles as they are selected. This can greatly speed up the selection process. Particles may be reduced in size (bin-ed) during calculations to improve performance, especially on slower workstations.

1.4 Preprocessing

Preprocessing is performed by selecting the Preprocess->Run menu item. A series of operations are performed on each selected particle and the results saved as separate image files. The operations are: removal of anomalously bright or dark pixels (caused by dust), calculation and correction for linear background gradient, masking out of neighboring particles from background, and normalization of pixel means and variances.

Parameters that affect preprocessing are entered in the dialog box (accessed with the Preprocess->Parameters menu item). These include the output filenames and numbering, box size for extracting particles and saving to file, background masking, and image normalization (detailed descriptions are in the help screens for the Preprocess menu).

2 Quick start

1. Run X3dPreprocess, and open an image file (BP; Perkin Elmar IMG/HDR; TIFF, Optronix, MRC or BIN) using the 'File' menu.
2. Select a particle. Then change the radius to a suitable value.
3. Continue selecting particles until done. At any time you may save the results in a coordinate file and quit the program. At a later restart, open the saved coordinate file to continue making selections.
4. When finished selecting particles, set the parameters using the 'Preprocess' menu. See the help for 'Preprocess menu' or click on the 'Help' button on the 'Parameters' dialog box.
5. Now is a good time to save your selections and parameters in a coordinate file as a record.
6. Select the Preprocess->Run menu item – this will extract, preprocess and save all the selected particles to image files.

3 File formats

Input image files can be read in the following formats:

BP.....Benes Trus's 8-bit 'PIC' format (OpenVMS systems only)
 IMG/HDR ...Perkin Elmar 16-bit format (from 'SCANS' program) as formatted after transfer to OpenVMS.
 PIF.....Portable Image Format from Purdue University
 MRCMRC modes 0 (8-bit), 1 (16-bit) and 2 (4-byte real).
 TIF.....Zeiss/Phodis output format (for example). Also the LeafScan/Photoshop TIFF files
 OPXOptronix scanner format, using 8-bit unsigned ints.
 BINJames' home-brewed simple binary format that works on OpenVMS and Unix systems. The first 32 bytes are a header:
 Bytes 0-7 (Ascii) X-dimension of image, eg 00001024
 Bytes 8-15 (Ascii) Y-dimension of image, eg 00000512
 Byte 16 (binary) Bits/pixel, 8 or 16, the top bit is 0 for little-endian, 1 for big-endian (2-byte pixels only)
 Bytes 17-31 Unused for now – all zeros

Note that the program expects files to have the extensions listed above so that it can recognize file types.

Output images (preprocessed) are saved in the BP, MRC mode-1, PIF or 8-bit BIN format (as selected). BP is only available on OpenVMS systems.

X3dpreprocess uses its own format for saving parameter data and particle coordinates – a *.CRD file. This is text and can be edited or printed.

4 Image window

Several convenience controls and status displays are located in the control area of this window. These include

- 1 buttons for zooming
 - 2 buttons for changing the radii
 - 3 sliders for changing the display contrast and brightness
 - 4 a count of selected particles and 'bad' selections
-

5 Hints

The standard Motif controls, such as scroll bars and sliders, allow several levels of fineness in their use. For example, a scroll bar can be:

- a) dragged by its 'thumb'
- b) 'nudged' by using the arrow controls at either end
- c) paged by clicking mouse button 1 in the gray area either side of the 'thumb'
- d) moved to an arbitrary position by clicking mouse button 2 in the same gray area

A slider can be moved in much the same way as the scroll bar, but in addition, if the slider is selected then the keyboard's cursor keys will move it incrementally for the finest of control.

Auto-centering will work better with good-contrast particles. The radius and extract box sizes will also affect the success, and the speed is better when the extract box can be kept under 256, or 128 or 64 pixels on each edge. Particles that intrude into the background can upset the centering if they are not masked out by being themselves selected. Generally you don't want to use high resolution information for centering – try 30-40 Ångstroms, and an edge of 5 Ångstroms.

6 File menu

The file menu allows opening of image or coordinate files, saving of coordinate data, closing of the main window, and quitting from the program.

Open Image	Open an image file in a window with no selections and default parameters set. This is the usual way to start preprocessing from a newly scanned micrograph.
Open Coords	Open a coordinate file, reading in all the preprocessing parameters, any selected coordinates, and the image file. This is the usual way to change preprocessing parameters, or to add new selections to the database.
Append Coords	Open a coordinate file, skip the preprocessing parameters, but add coordinates to those already selected.

Close	Close the main window. If selections or parameters have changed, you will be prompted to save the coordinate file with this updated information.
Save Coords	Save the selected coordinates and preprocessing parameters to a text-format *.CRD file. This file may be read back into the program at any later time. Coordinates are saved with id's that match those used for the output image file names.
Quit	Quit the program. If selections or parameters have changed, you will be prompted to save the coordinate file with this updated information.

7 Options menu

Display options are set here, as well as the circular selection radius.

Flip Rows	This option will flip the rows, in case the scan was backwards.
Flip Alternate Rows	Some scanners allow alternate rows to be scanned and saved in reverse order. This option will flip the alternate rows.
Invert Image Intensities	This option will invert the pixel values.
Forget Coordinates	Forget all selected particles – ie, wipe the slate clean.
Display E-boxes	In addition to the selection circle, a box the size of the initially extracted area is drawn. This size is set in the Parameters dialog box.
Display F-boxes	In addition to the selection circle, a box the size of the final area that will be saved to file is drawn. This size is set in the Parameters dialog box.
Display Ids	The ID of each particle is drawn in the center of the selection circle. The start number is set in the Parameters dialog box.
Update Ids	ID's are not automatically updated when a particle is (for example) deleted, or marked as bad. This option will redraw all the IDs on screen.

8 Preprocess menu

Preprocessing parameters are set here, and the operation is performed with the 'Run' item.

Parameters...	Preprocessing parameters are set in this dialog box. These are also saved in the coordinate file as a record and for any later re-processing.
Output Filenames	The preprocessed images are saved as a series of files, all sharing a common filename prefix to which is appended a number, for example, IMG001.BP, IMG002.BP, IMG003.BP,

etc. The prefix, starting numbers, and number width are set here. Also the output file format may be selected – currently BP, MRC, PIF and BIN are supported. Another option is to enable conversion from transmission values to optical densities (OD is on/off). Changing this will affect the display but not the input scan image, and the OD conversion will be performed on the particles (using floating point arithmetic) as they are extracted from the micrograph for preprocessing.

Particle Box Sizes Each particle is copied as a square array from the micrograph, centered (where possible) on the selection center. The initial size should be large enough so that a good estimate of any linear background gradient can be made. The final size is that saved in the output image file, and may be smaller than the initial size, but should not be smaller than the diameter (twice the selection radius).

Background Mask A circular mask is used for defining foreground (area containing particle) and background in the preprocessing operations. The 'Replace Background' option allows the background to be replaced by its mean value after other preprocessing is complete. The 'Radius' is the same value as that used for selecting particles.

Fade When the background is replaced by a constant value equal to its mean, a sharp edge is introduced. This circular edge may interfere with correlational alignment procedures. To avoid this problem, a fade width (in pixels) can be set. At the edge of the circular mask, the intensities are those of the image. Beyond this radius the pixels are blended with the mean according to a Cosine² function for a width set in the 'fade' field. The default is 0.0 pixels and the value can be fractional. Note that intensities within the mask radius are not affected by the fade operation, only those pixels from the mask radius out to a limit of (radius+fade) are a blend of image pixels and constant background.

Normalization Images are normalized to constant foreground mean (within circular mask) and variance (we use standard deviation, which is the square root of variance). Typical values might be mean of 127 and s.d. of 40. Increasing the s.d. will increase contrast but cause more pixels to be clipped at 0 (black) and 255 (white).

Set Centering Parameters Set low-pass filter parameters for the centering operations. Also set the bin-factor to reduce images by during calculations, to improve speed of centering: 1=no-bin, 2, 3...

Auto-Center Origins Particles are centered automatically after dragging. Use the left mouse button for this, but if a particle does not center properly then use the middle mouse button to drag it into place without auto-centering. Select this menu item again to disable it.

Auto-Pick Origins	Not working yet. This option should allow a fast start to selecting particles.
Re-Center All Origins	Particles are partially preprocessed, then a correlation alignment is calculated between each image and a copy rotated 180-degrees. The peak of the cross-correlation image indicates where the center of the particle is, assuming it is roundish or has a 2-fold in-plane symmetry axis. Selections are redrawn based on the new centers. This can take a considerable amount of time.
Run	This invokes the preprocessing phase. Selected areas are extracted (copied) from the image, preprocessing is performed (see Overview) and resulting image files are written out. Some status information is written to the DecTerm display.

9 Known problems

(Note – this is not a complete list!)

- * Screen updating is incomplete when scrolling while the image is partially obscured by another window. I haven't figured out how to get hold of the relevant 'Xgraphics-Expose' events yet.

10 Future plans

- * Additional selection shapes, especially a rotatable rectangle
- * Use any suitable colors in current color map – now, we ignore them
- * Respond to color-map-changed events
- * Try using RO color map before RW – now its the other way around
- * Draw graphics on temp image instead of directly
- * Streamline image updating – reduce unnecessary re-draws.
- * Zoom in as well as out
- * Progress bars for file input, and preprocessing
- * Don't redraw particle count titles, just numbers

Longer term aspirations:

- * Multiple populations, different sizes, shapes
- * Multiple windows
- * Replace the dreadful Motif default filename dialog box

11 Version history

v1.8.2n 20-Oct-2000

- * Changed display of particle and reject counts from labelWidget to textField, which avoids update flickering.
- * Attempting to enable auto-picking code (underway).

v1.8.2m 31-Aug-2000

- * Fixed TIFF_io.c for all cases of big/little-endian files with big/little-endian computers. The new Zeiss scanner (PhotoScan) hosted on a WindowsNT machine writes little-endian files, while the NIH Zeiss hosted on an SGI writes big-endian files.

v1.8.21 12-May-2000

- * Added menu item to flip all rows.

v1.8.2k 04-May-2000

- * Fixed input of TIFF files in 'strip' format (vs 'tile' format).
- * Changes to interpretation of MRC format.

v1.8.2j 29-Mar-99

- * Fixed the curve-fit routine in FitPoly2_2d, which occasionally suffered div-by-0 errors.
- * Also re-enabled the Auto-pick code, but its a bit primitive still.

v1.8.2i 30-Nov-98

- * Added a bin option for auto-centering to speed up this operation.
- * Added menu items to append coordinates from a file (ignoring parameters), and to forget the currently selected coordinates.
- * Auto-picking is coded, but not tested nor is it showing any signs of being useful (yet).
- * Added the image file name to the window title bar.

v1.8.2h 20-Oct-98

- * Added MRC and Optronix format for input.
- * Fixed bug in progress report of calculating statistics of input image.
- * Fixed single-click problem, which was selecting a particle but not timing out (waiting for double-click).

v1.8.2g 9-Apr-98

- * Fixed bug in OD conversion – it was being done all the time.
- * Also removed 'Dynamic background' check box in mail parameter dialog, cos it is obsolete (always off).
- * Normalization (const mean, const background) is too sensitive, so it has been changed (const mean, const variance).

v1.8.2f 24-Mar-98

- * Some tidy-ups.
- * Added acknowledgements to this file.
- * Added PIF-file input.
- * Added Optical Density conversion and Invert Image Intensities option.
- * Added LeafScan/Photoshop TIFF format for input (uses strips, where Zeiss/TIFF uses tiles).
- * Fixed memory bug in centering code (not free-ing copyarray).

v1.8.2e 16-Mar-98

- * Ported auto-centering code, including FFT, to C.
- * Added dialog box for lopass filter parameters.
- * Added the option of saving the results in a PIF-format file.

v1.8.2d 14-Mar-98

- * Fixed a few cosmetic display bugs, most noticeable when the good and bad particles were very different sizes.
- * Added an auto-center option – as soon as a particle is dropped it is run through the preprocessing steps, and cross correlated with a copy rotated by 180-degrees, and snaps to the center.
- * Fixed a bug that shows up only on the SGI (use of a static string).

v1.8.2c 10-Nov-97

- * Addition of TIFF file format for reading Zeiss scanner files directly.
- * Added centering option.
- * Removed dynamic background intensity calculations (which never worked properly anyway).

v1.8.2b 06-Mar-97

- * Another interim release.
- * Contrast has been moved from a menu item to a slider control, and an additional control for brightness has been added. The values of these two controls are saved in the coordinate file.
- * The radius menu item has been removed since there is a button for each of the two radii.
- * Likewise, the zoom menus have gone.

v1.8.2a 14-Nov-96

- * This is an interim release because changes for 1.8.3 have started but are taking longer than anticipated.
- * A new radius has been added for the bad selections as a quick-n-nasty way to mask out one of two populations that differ in size.

v1.8.2 19-Sep-96

- * Tidy up of globals in M_Menu - moved some routines into M_ImageDraw where they belong.

v1.8.1 17-Sep-96

- * SGI-compatible.
- * Made local routines and variables 'static', fixed up some 'extern's.

v1.8 14-Aug-96

- * Added mouse button 2 functions.
- * Bug fix to dust correction procedure.
- * Output files can have numeric field from 3 to 5 digits.
- * Coordinate file lists good particles in order with the same id number as the output files.
- * Bad particles are listed at the end.

v1.7.5 17-Apr-96

- * Bug fixed when re-sizing window.
- * Bug fix to output of binary image.

v1.7.4 12-Apr-96

- * Bug fixed in displaying final boxes.
- * Changed line width of bad circles to 1.

v1.7.3 11-Apr-96

- * Added menu item to flip alternate rows of image. Needed when scanning this way.

v1.7.2 09-Apr-96

- * Bug fix in zoomed screen display.

v1.7.1 22-Mar-96

- * Special one-off for Mario C.

v1.7.0 22-Mar-96

- * First distribution with source.

v1.6.1 06-Feb-96

- * Fixed new bug in 1.6.0 with normalization.

v1.6.0 02-Feb-96

- * Fixed problem causing some particles to end up with reverse contrast - caused by background being darker (or lighter) than foreground, when it should be the opposite.
- * Also changed particle dragging - particles are now picked up at the point where the mouse is clicked, not the circle center.
- * Particle ID's are now optionally displayed, as are the extents of the extraction box and the final saved box.

v1.5.4 21-Nov-95

- * Improved zoom-out - averages pixels instead of displaying top-left. Looked at cubic splines, but too slow.
- * Fixed bug in Help dialog (Unix).

v1.5.3 3-Nov-95

- * Enabled `backing_store` and `save_under`.

v1.5.2 31-Oct-95

- * Fixed bugs introduced into version 1.5.1.
- * 8-bit and 24-bit X-Displays fully supported on OpenVMS and Dec Unix.
- * 8-bit and 16-bit images displaying properly
- * selections display properly and are counted properly (good/bad)

v1.5.1 27-Oct-95

- * Fixed bug preventing output of BP files when working on a Perkin-Elmer input image.
- * Fixed screen update bug, made screen updates faster, and program now works on both 8-bit and 24/32-bit X-Displays.

v1.5.0 17-Oct-95

- * Changes for porting to Unix
- * conditional compilation for OpenVMS-specific file formats (BP, PE)
- * change all remaining Fortran code to C
- * add a new file format which will work on both Unix and OpenVMS systems (BIN).

v1.4.1 14-Sep-95

* Minor bug fixes

v1.3.2 30-Aug-95

* Added fade ring for circular masking.

12 References

An early frame-buffer version of the program (TrimNew) based on work by a summer student Yoli Yeung is described briefly in:

1. Conway JF, Trus BL, Booy FP, Newcomb WW, Brown JC & Steven AC (1993) The effects of radiation damage on the structure of frozen hydrated HSV1 capsids. *J Struct Biol* **111**, 222-233.

Other brief but more recent descriptions are in:

2. Conway JF, Duda RL, Cheng N, Hendrix RW & Steven AC (1995) Proteolytic and Conformational Control of Virus Capsid Maturation: The Bacteriophage HK97 System. *J Mol Biol* **253**, 86-99.
3. Conway JF, Trus BL, Booy FP, Newcomb WW, Brown JC & Steven AC (1996) Visualization of Three-dimensional Density Maps Reconstructed from Cryo-Electron Micrographs of Viral Capsids. *J Struct Biol* **116**, 200-208.
4. Conway JF & Steven AC (1999) Methods for reconstructing density maps of "single" particles from cryoelectron micrographs to subnanometer resolution. *J Struct Biol* **128**, 106-118.

http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10600565&dopt=Abstract

Some file formats used by X3d are described here:

5. Trus BL, Kocsis E, Conway JF, Steven AC (1996) Digital image processing of electron micrographs: the PIC system-III. *J Struct Biol* **116**, 61-67.
http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=8742724&dopt=Abstract
6. Crowther RA, Henderson R, Smith JM (1996) MRC image processing programs. *J Struct Biol* **116**, 9-16.
http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=8742717&dopt=Abstract
7. PIF (Purdue Image Format) is described on the following web-page:
http://bilbo.bio.purdue.edu/~baker/programs/doc/sys_info/pif.pdf

13 Acknowledgements

The genesis of X3dPreprocess was in a program called TRIMBP written largely by a Summer Student, Yoli Yeung, in 1993, supervised by Dr Benes Trus (CBEL, DCRT, NIH). This was an attempt to formalize and simplify the heavily manual methods in use at the time, and was written for a Gould 8000 Image Processor hosted by a DEC OpenVMS system. The first port was by JFC to an updated Gould 9000 with a 1024x1024 display acting as a moveable window onto a 2048x2048 frame buffer (TRIMNEW). This version added significantly to the interaction – selected particles were circled, selections could be deleted, the display area could be zoomed and scrolled. The second port was to XWindows/Motif, also by JFC, and included changing from Fortran to C as well as

removing all dependencies on OpenVMS so that the program would also compile on Dec Unix and SGI systems.

The following people have contributed code:

Rob Ashmore for PIF-format file I/O
Purdue University,
West Lafayette, IN
(rwa@baker1.bio.purdue.edu)

Brian Sheehan for MRC-format file i/o
The Scripps Research Institute
The La Jolla, CA
(sheehan@scripps.edu)

Elizabeth Hewat for more MRC-format file i/o, and Optronix i/o
Laboratoire de Microscopie Electronique Structurale
Institut de Biologie Structurale
Grenoble, France
(hewat@lmes.ibs.fr)

David Belnap for code to fit peaks (centering)
Lab of Structural Biology Research, NIAMS,
National Institutes of Health,
Bethesda, MD
(David_Belnap@nih.gov)

Useful ideas have been suggested by users in the LSBR and others.